

Ciacci, Riccardo

Working Paper

Re-analysis of Adema et al. (2025a)

I4R Discussion Paper Series, No. 248

Provided in Cooperation with:

The Institute for Replication (I4R)

Suggested Citation: Ciacci, Riccardo (2025) : Re-analysis of Adema et al. (2025a), I4R Discussion Paper Series, No. 248, Institute for Replication (I4R), s.l.

This Version is available at:

<https://hdl.handle.net/10419/321375>

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.

No. 248
DISCUSSION PAPER SERIES

Re-analysis of Adema et al. (2025a)

Riccardo Ciacci

This paper responds to:

Adema, J., O. Folke, and J. Rickne. 2025. Re-Analysis of Ciacci, R. (2024). Banning the purchase of sex increases cases of rape: evidence from Sweden. *I4R Discussion Paper Series* No. 226. Institute for Replication.

This paper received a response:

Adema, J., O. Folke, and J. Rickne. 2025. Still no reform, just overfitted control functions and inflated precision. *I4R Discussion Paper Series* No. 249. Institute for Replication.

July 2025

I4R DISCUSSION PAPER SERIES

I4R DP No. 248

Re-analysis of Adema et al. (2025a)

Riccardo Ciacci¹

¹Universidad Pontificia Comillas, Madrid/Spain

JULY 2025

Any opinions in this paper are those of the author(s) and not those of the Institute for Replication (I4R). Research published in this series may include views on policy, but I4R takes no institutional policy positions.

I4R Discussion Papers are research papers of the Institute for Replication which are widely circulated to promote replications and meta-scientific work in the social sciences. Provided in cooperation with EconStor, a service of the [ZBW – Leibniz Information Centre for Economics](#), and [RWI – Leibniz Institute for Economic Research](#), I4R Discussion Papers are among others listed in RePEc (see IDEAS, EconPapers). Complete list of all I4R DPs - downloadable for free at the I4R website.

I4R Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

Editors

Abel Brodeur
University of Ottawa

Anna Dreber
Stockholm School of Economics

Jörg Ankel-Peters
RWI – Leibniz Institute for Economic Research

Re-analysis of [Adema et al. \(2025a\)](#)*

Riccardo Ciacchi[†]

July 7, 2025

Abstract

[Adema et al. \(2025a\)](#) critically reassesses the findings of [Ciacchi \(2024\)](#), arguing that the reported effects result from statistical errors and flawed empirical implementation. This response systematically addresses those claims by applying their proposed methodology to the corrected sample outlined in [Ciacchi \(2025\)](#). The analysis confirms that the large positive estimated coefficient linking the criminalization of sex purchase to reported rape cases is not attributable to technical or coding mistakes. Rather, the differences in robustness observed in [Ciacchi \(2025\)](#) arise from the estimation of a distinct target parameter and the implementation of alternative robustness checks, not from errors in the original study. In addition, this reply identifies methodological and procedural issues in [Adema et al. \(2025a\)](#), including comparisons across non-equivalent specifications and limited consideration of contextual assumptions. The response advocates for a more transparent, respectful, and collaborative approach to replication, particularly when addressing sensitive topics that demand scientific integrity and academic rigor.

*This article also offers a companion do file to replicate results. All errors are my own.

[†]Universidad Pontificia Comillas, Department of Economics and Business Management, rciacchi@icade.comillas.edu

1 Introduction

The recent re-analysis by [Adema et al. \(2025a\)](#) critiques my article [Ciacci \(2024\)](#) on both methodological and statistical grounds. While scientific debate is central to academic progress, the tone and approach taken in their article raise serious concerns. Their critique includes public accusations of misconduct based on partial access to materials and preliminary code, shared in response to informal social media interactions rather than through standard scholarly channels. Furthermore, their paper presents a number of analytical misrepresentations, incorrect statistical comparisons, and speculative claims regarding intent.¹

Substantively, their critique misrepresents the original study's empirical framework by (i) comparing regression discontinuity results across fundamentally different kernel and polynomial specifications without proper adjustment; (ii) neglecting changes in legal definitions of rape, which require sample adjustments or year fixed effects; and (iii) ignoring methodological corrections introduced in [Ciacci \(2025\)](#), including the appropriate treatment of seasonal effects and model specification under multicollinearity.

More concerning, [Adema et al. \(2025a\)](#) frames these disagreements not as scholarly debate but as a basis for calling for retraction, suggesting intentional misrepresentation without evidence. Such an approach undermines the norms of respectful academic discourse. It is important to emphasize that an independent editorial committee appointed by the Journal of Population Economics found no evidence of scientific misconduct. The scientific committee did not support the accusations levied in the re-analysis.

This reply also highlights multiple methodological and procedural shortcomings in [Adema et al. \(2025a\)](#)'s critique, including improper comparison across non-equivalent specifications, omission of key fixed effects and sample corrections, and disregard for contextual assumptions detailed in the original and follow-up work. More broadly, [Adema et al. \(2025a\)](#) engages in a replication effort that fails to meet basic standards of independence, transparency, and professionalism. The authors omit crucial contextual details, introduce ad personam criticisms, and rely on anecdotal evidence to question empirical findings that do not align with their ideological priors. This response calls for more rig-

¹A sample of this practice might be found in Section [Appendix A](#). A sample of how this practice was carried out on social media might be found in Section [Appendix B](#). Some of the worst tweets, such as writing that "I don't know how this paper made it past peer review" by Olle Folke, were deleted after several days of harassment and public shaming on social media. A year after they deleted those tweets they even claim they do not recall their disrespectful tone (Figure [A.6](#)). I would like to highlight I never refused to share the code, they asked me via email to send the code of the whole paper within 6 hours or they would post a terrible thread on myself on Twitter. I sent the code after 28 hours. I can show the relevant correspondence to prove this. The authors' prior public commentary on social media, raises reasonable concerns about the independence and objectivity of their judgment on this topic.

orous, respectful, and collaborative approaches to scientific replication, particularly on sensitive and policy-relevant topics.

This article responds point-by-point to the methodological concerns raised by [Adema et al. \(2025a\)](#), replicates their proposed specifications on the corrected sample, and demonstrates that, even under their own framework, the primary association between the reform and rape remains statistically significant in several configurations.

The findings of this article relate to a strand of the literature showing that criminalizing the purchase of sex can produce effects that are not only opposite to those of decriminalization but also substantially larger in magnitude. Specifically, [Gao and Petrova \(2022\)](#) estimate that the impact of criminalization is between four and six times greater, in absolute value, than the effect of decriminalization. This article engages with two main strands of the literature. First, recent empirical studies in economics provide causal evidence that decriminalizing prostitution, regulating sex work via licenses or increasing the availability of establishments related to sex work lead to a reduction in sex crimes, including rape ([Bisschop et al. 2017](#); [Cunningham and Shah 2018](#); [Ciacci and Sviatschi 2022](#)). Second, other studies highlight that criminalizing prostitution, whether targeting sellers or buyers, may lead to unintended consequences, such as increases in sexually transmitted infections and sexual violence ([Cunningham and Shah 2020](#); [Cameron et al. 2021](#); [Della Giusta et al. 2021](#); [Gao and Petrova 2022](#)).

2 Regression Discontinuity

[Adema et al. \(2025a\)](#) argue that the main regression discontinuity estimate in [Ciacci \(2024\)](#) is invalid due to misspecification and coding choices. While it is true that the presence of year and month fixed effects introduces collinearity with the treatment variable, [Adema et al. \(2025a\)](#) fail to replicate the correct sample and specification described in [Ciacci \(2025\)](#). Moreover, they fail to explore whether the large positive estimated effect at regional level in [Ciacci \(2024\)](#) is due to the collinearity issues or it is unrelated to it. Given the concerns raised in [Adema et al. \(2024\)](#), the scientific committee reviewing [Ciacci \(2025\)](#) established that year fixed effects should not be used in the regression discontinuity in time analysis and given this omission – and the legal changes in the definition of rape – the subsample spanning from 1997 to 2005 should be used for computations. [Adema et al. \(2025a\)](#) overlook these methodological issues and use the samples of [Ciacci \(2024\)](#) where year fixed effects were used.

Table 1 reports the re-estimated treatment effects using the specification proposed by

Adema et al. (2025a), applied to the sample of Ciacchi (2025).² Columns (1) and (2) show the treatment effects using first- and second-order polynomials, respectively. Both yield statistically significant and positive estimates, with rape increasing by 25–30%. Columns (3) through (6) assess the robustness of these estimates by removing control variables. Columns (3) and (4) remove police officers, while columns (5) and (6) remove region fixed effects. The estimates remain positive and statistically significant, regardless of whether police officers or region fixed effects are excluded. These findings invalidate the central claim of Adema et al. (2025a) that the results in Ciacchi (2024) are solely due to coding errors.

Figures 1 and 2 further explore robustness across bandwidth sizes using the checks introduced in Ciacchi (2025).³ These extensive robustness checks—which entail running around 70 regressions for each estimate and comparing pre and post treatment outcomes far from the optimal bandwidth while incorporating additional post treatment months given data availability—confirm that the results are not related to multicollinearity. The treatment effect remains stable up to an 8-9 month bandwidth, after which the pre/post treatment balance weakens. Given the optimal bandwidth of the considered regressions, this 8-9 month range roughly coincides with a balanced pre and post treatment window of 24 months. These results reinforce the substantive finding that the 1999 criminalization of sex purchase increased reported rape cases. Importantly, the estimates are not driven by coding artifacts and persist even under Adema et al. (2025a)’s methodology.

Critically, Adema et al. (2025a) compare estimates across specifications with different kernels (uniform vs. triangular) and polynomials (order 1 vs. 2), without stating so. This distinction is not acknowledged in their paper, such an omission might lead readers to an inaccurate interpretation of robustness.

To further investigate the robustness of the original results, Table 2 applies the same methodology used in Ciacchi (2025) to estimate a regional representative coefficient as in Ciacchi (2024). Specifically, it considers the following regression model:

$$\log(\text{rape}_{rmy}) = F(y; y < \text{Jan99}) + \beta_1 \mathbb{I}\{y \geq \text{Jan99}\} + F(y; y \geq \text{Jan99}) + \gamma \text{officers}_{ry} + \alpha_m + \varepsilon_{rmy} \quad (1)$$

where rape_{rmy} is the number of reported cases of rape in region r and month m of year y .⁴ $F(y; y < \text{Jan99})$ and $F(y; y \geq \text{Jan99})$ are the usual polynomial control functions

²The methodological framework employed in this analysis builds upon the regression discontinuity design developed by Hausman and Rapson (2018).

³Figures use the acronym RDiT to refer to the regression discontinuity in time framework.

⁴Namely, $\log(\text{rape}_{rmy})$ is the logarithm of one plus the number of reported cases of rape.

included in regression discontinuity frameworks. In this case, following [Ciacci \(2024\)](#), I show results for either a polynomial of order 1 or of order 2. $\mathbb{I}\{y \geq \text{Jan99}\}$ is the treatment variable, taking value 1 for observations after the entry into force of criminalization of the purchase of prostitution and 0 otherwise.

The control variable officers_{ry} is the number of police officers in region r in year y ; this variable only varies at region and year level since police officers are hired by regions on a yearly basis. α_m is a seasonal fixed effect that groups month according to the season to which they belong to. Namely, choice of the control variables is carried out relying on the assumed smoothness of time and the locality implicit in the regression discontinuity method.

Columns (1) and (2) of Table 2 explore the functional forms of the dependent variable using either logarithmic or inverse hyperbolic sine (hereafter, IHS) transformations and first-order polynomials, respectively. Columns (3) and (4) repeat the same analysis for second-order polynomials. In all cases, the estimated coefficients remain positive and statistically significant, confirming that the effect is not driven by multicollinearity between the treatment variable and fixed effects.

The lack of robustness in [Ciacci \(2025\)](#) stems from changes in robustness checks between [Ciacci \(2024\)](#) and [Ciacci \(2025\)](#), an issue unrelated to multicollinearity. Nevertheless, applying [Adema et al. \(2025a\)](#)'s methodology to the sample from [Ciacci \(2025\)](#) yields results that remain robust even under these stricter checks.

While [Adema et al. \(2025a\)](#) highlight the absence of any discernible jump at the cut-off date, their analysis relies on a different dataset than the one used in the original study. Here, I replicate their approach by aggregating data across year-months, but using both the sample and control variables specified in [Ciacci \(2025\)](#). Figure 3 presents these results.

3 Difference in Discontinuity

[Adema et al. \(2025a\)](#) claim the observed discontinuity is a seasonal artifact, particularly the result of a December–January comparison. To address this, the difference-in-discontinuity approach in [Ciacci \(2025\)](#) includes a placebo test comparing December 1997 and January 1998—a period preceding the reform. Table 3 reports results from these placebo regressions. Across specifications, the estimated coefficients are not statistically different from zero. Estimates for the polynomial of degree two are even negative in point estimate. In other words, the placebo comparison does not exhibit the sharp increase observed in the post-reform period. If the found estimated coefficient were attributable to seasonal variation between December and January, an analogous pattern should appear

in the pre treatment period as well.

Furthermore, when computing the difference in discontinuity estimate suggested in [Ciacci \(2025\)](#), that is, the difference between the estimates in Table 2 and those in the placebo estimates shown in Table 3, the treatment effect is positive and even larger in size and statistically significant for the second-order polynomial. This also suggests that the effect is not a result of December vs. January seasonality.

Next, I estimate month and year fixed effects previous to the entry of the ban and I subtract the corresponding month from each cell of the dataset. The 1998 year fixed effect is subtracted from year 1998 data but also from post-treatment data to mimic the logic of the difference-in-discontinuity design.⁵ Then, I aggregate such data at national level for each year-month cell. Figure 4 displays the results of this analysis to evaluate data at the cut-off for the difference-in-discontinuity design.

Taken together, these findings provide further compelling evidence that the results in [Ciacci \(2024\)](#) are not driven by either collinearity issues or seasonal changes but rather by the policy intervention. Moreover, the point estimates are between 60 and 80%. Even larger than those estimated in [Ciacci \(2024\)](#).

Size of the effect

Figure 5 compares the main estimates from the regression discontinuity in time and difference-in-discontinuity in time frameworks. The results are statistically equivalent across specifications despite using polynomials of different orders and alternative kernels. Five of the six estimators indicate that the ban significantly increased reported rape cases, with effect sizes becoming larger when accounting for potential seasonal variations between December and January and ranging (in point estimate) from 25% to about 80%.

These results are consistent with previous findings in the literature, though somewhat smaller in magnitude. [Cunningham and Shah \(2018\)](#) document a 30% reduction in rape cases following prostitution decriminalization, while our analysis in Figure 5 suggests that criminalization leads to an increase of comparable or up to two and a half times greater magnitude. This effect size is moderately smaller than the more substantial impacts estimated by [Gao and Petrova \(2022\)](#), who find that criminalization produces effects four to six times larger, in absolute value, than decriminalization across different model specifications.

The regionally representative parameter estimated in [Ciacci \(2024\)](#) is particularly im-

⁵Subtracting fixed effects is as computing residuals for pre-treatment regression that uses as control variables only time (i.e. month and year) fixed effects.

portant for the Swedish context, as some counties have smaller populations where even absolute increases in rape cases represent substantial relative changes. This contrasts with the nationally representative parameter estimated in [Ciacci \(2025\)](#).

In sum, the observed increase in rape cases established in [Ciacci \(2024\)](#) cannot be attributed to either: (i) the multicollinearity issues discussed in [Ciacci \(2025\)](#), or (ii) seasonal fluctuations as suggested by [Adema et al. \(2025a\)](#). In fact, when accounting for seasonal fluctuations results are larger in size and suggest the ban increased rape by about 80%. Crucially, when estimating the regionally representative parameter across alternative methodologies, the results remain robust.

4 Fuzzy Difference-in-Differences

[Adema et al. \(2025a\)](#) argue that the Fuzzy DiD analysis in [Ciacci \(2024\)](#) uses an incorrect time variable. However, this critique overlooks the justification provided in the Supplementary Materials (Sections S.1.2.1 and S.5.3.2 of [Ciacci \(2024\)](#)), where the choice to aggregate data at the year level is motivated by the nature of officer hiring and budget cycles, which operate on an annual basis. Such choices are motivated by [de Chaisemartin and dHaultfoeuille \(2018\)](#); [de Chaisemartin et al. \(2019\)](#). Equation S.1 of [Ciacci \(2024\)](#) clearly specifies the empirical model, including fixed effects, and Figure S.6 of [Ciacci \(2024\)](#) supports the parallel trends assumption at the yearly level. This motivates the usage of the time variable at year level. Using year-month as a time variable would imply a different parallel trends assumption, not at year level but at year-month level. [Adema et al. \(2025a\)](#) claim such discussions are not present in [Ciacci \(2024\)](#). Yet, these discussions are present in the afore mentioned sections in the Appendix of [Ciacci \(2024\)](#).

It is worth noting that in the fuzzy difference-in-differences framework of [Ciacci \(2024\)](#), monthly heterogeneity is accounted for through month fixed effects. [Adema et al. \(2025a\)](#)'s interpretation fails to engage with the above mentioned justifications and instead argue that the annual aggregation choice was arbitrary. This approach falls short of the standards of respectful and constructive scientific discourse.

5 Event Study

The decision to cluster standard errors at the region-year-month level in the event study aligns with the treatment assignment mechanism, which operates at this same level. Specifically, the treatment variable is a binary indicator for whether at least one

fine event occurred in a given region-year-month. Given the non-standard nature of this event study design, clustering at the treatment assignment level required accounting for all three dimensions: region, year, and month. When I originally conducted this analysis, the clustering was implemented using the most up-to-date version at that time of Stata's user written `reghdfe` command, which supported three-way clustering.

In light of recent findings about biases in two-way fixed effects event studies, this is a specification I would no longer use if writing [Ciacci \(2024\)](#) today. However, while [Adema et al. \(2025a\)](#) criticize three-way clustering—arguing it artificially reduces standard errors due to the limited number of monthly clusters (only 12 levels)—they provide no empirical test to support this claim. Notably, when applying wild cluster bootstrap methods (using either Webb or Rademacher weights) to the event study from [Ciacci \(2024\)](#) with clustered variance at month level, the estimates remain statistically significant.

Given the current status of the literature, I would not use this methodology to evaluate the effect of the ban on cases of rape. However, in replicating this study, I would expect a thorough and intellectually honest discussion of how rapidly the literature on two-way fixed effects difference-in-differences estimators has evolved in recent years, along with the methodological improvements incorporated into the latest versions of `reghdfe`.

6 Instrumental Variables

The most problematic aspect of [Adema et al. \(2025a\)](#) is their reliance on anecdotal experience and subjective judgment in critiquing the instrumental variable strategy. For example, the assignment of main airports based on proximity to region centroids was transparently documented in [Ciacci \(2024\)](#) and intended as a consistent rule for instrumental variable construction.

[Adema et al. \(2025a\)](#) criticize how Google Maps identify airports and suggest alternatives based on their personal experience traveling to Arlanda Airport. Such anecdotal counterfactuals do not meet academic standards for methodological critique. Their adjustments fail to account for the broader instrumental variable logic and introduce measurement error of their own. They even mention changes in airport that took place in the very last year of the sample period considered in [Ciacci \(2024\)](#).

7 Conclusion

This paper establishes that the central claim of [Adema et al. \(2025a\)](#)—that the main results in [Ciacci \(2024\)](#) are driven by coding errors and seasonal fluctuations—does not hold when their own methodological corrections are implemented on the appropriate sample established in [Ciacci \(2025\)](#). The large and positive estimated effect of the 1999 sex purchase ban on reported rapes remains statistically significant under several reasonable specifications. These results align with a growing body of literature showing that criminalizing prostitution can have effects not only opposite to those of decriminalization but also significantly larger in magnitude ([Cunningham and Shah 2020](#); [Cameron et al. 2021](#); [Della Giusta et al. 2021](#); [Gao and Petrova 2022](#)).

It is important to emphasize that in the spirit of scientific inquiry, [Ciacci \(2025\)](#) re-examines a related but distinct research question from [Ciacci \(2024\)](#), estimating a different target parameter (i.e. nationally representative – weighted by population); their results should therefore be interpreted not as invalidating the earlier findings but rather as refining them through continued investigation – precisely as the current study aims to do. Further research employing more suitable methodologies, improved data quality, or both will be essential to fully elucidate this important policy question.

This reply also underscores the importance of respectful scientific discourse. Public accusations of misconduct, made prior to formal peer review and based on partial evidence, have a chilling effect on academic inquiry. While robustness checks and further research are essential, they must be conducted and communicated in a manner consistent with scholarly norms.

More broadly, I hope this debate encourages the academic community to reflect on the importance of maintaining respectful and constructive discourse, particularly when disagreements arise. It is vital that all scholars—and especially early-career researchers—can engage in scientific dialogue without facing undue pressure or dismissive commentary, particularly from more senior colleagues, who play a key role in modeling professional standards.

I also hope that this reply helps to shed light on an important and sensitive topic. The appropriate response to methodological disagreements is not censorship of former articles but rather the promotion of better research practices and broader scholarly engagement. It remains unclear why [Adema et al. \(2025a\)](#), despite their institutional affiliation with Swedish universities and access to more precise administrative data, have not pursued analyses using geographically more disaggregated and/or higher frequency data. As a researcher not affiliated to a Swedish institution, I do not have access to those data,

but I would welcome the opportunity to analyze them if made available. Alternatively, I encourage other scholars with access to such information to contribute to this research area. It is essential that these topics not be treated as taboos but instead studied rigorously through a plurality of methods and perspectives.

Finally, it is important that replication efforts adhere to principles of independence, respect, and transparency. While I welcome scrutiny of my work, I believe that some aspects of the commentary surrounding this replication have not fully aligned with those standards, and this has at times detracted from what could have been a more collaborative and constructive scientific exchange.

Such practices undermine both the integrity of replication work and the norms of scholarly debate. Replication efforts should adhere to the principles of scientific dialogue: engaging constructively, contacting the original author to clarify uncertainties, and relying on formal academic channels rather than informal or social media platforms. In this case, since their article [Adema et al. \(2024\)](#), the authors of [Adema et al. \(2025a\)](#) never contacted me with any questions or concerns, despite claiming to engage in scientific replication. This approach weakens the credibility of their critique and detracts from the collaborative spirit of academic inquiry. The limited engagement around the replication process, together with public statements made in advance of receiving the replication materials, may give the impression that the discussion was shaped by pre-existing views. I believe this context is relevant for readers to fully assess the tone and framing of the replication effort.

Reply to their response [Adema et al. \(2025b\)](#)

After receiving my reply, the authors of [Adema et al. \(2025a\)](#) submitted their response [Adema et al. \(2025b\)](#), to which I now provide this counterresponse:

I hope this discussion contributes constructively to advancing our understanding of the impact of sex purchase bans on sexual violence. In the case of Sweden, there is evidence suggesting that the implementation of the ban varied significantly across regions, both in terms of enforcement strategies and allocated resources ([Ekberg 2004](#)). This heterogeneity highlights the importance of considering covariates in regression discontinuity in time analyses—not merely to improve precision, but also to mitigate potential bias ([Hausman and Rapson 2018](#)). When the treatment effect may not be constant, methods such as the two-step procedure and local linear regressions, particularly within narrow bandwidths, can offer more reliable estimates than global polynomials ([Hausman and Rapson 2018](#)). These approaches focus on the immediate vicinity of the discontinuity and are less influenced by long-run trends away from the threshold. If [Adema et al. \(2025b\)](#) suspect that the treatment effect evolves over time, dynamic controls (e.g., lagged outcomes) or adopting an augmented local linear framework may be more effective ([Hausman and Rapson 2018](#)).⁶ This analysis further highlights, as in [Ciacci \(2025\)](#), that findings in [Ciacci \(2024\)](#) were not due to multicollinearity. That is, findings in [Ciacci \(2024\)](#) are not incorrect. However, as usual, more disaggregated data might be needed to further explore robustness of results and mechanisms. While aspects of this debate have at times been disheartening, it has nonetheless opened the door to valuable research finding no evidence supporting a precise zero or a decreasing effect; an important contribution in itself.

⁶[Hausman and Rapson \(2018\)](#) note that regression discontinuity in time designs are conceptually similar to event studies, interrupted time series and simple pre/post comparisons, as all exploit known treatment timing. However, they emphasize that time-varying treatment effects and reliance on global polynomials can introduce bias (especially when using data far from the threshold) suggesting that approaches like donut may be less informative in such settings.

Tables and Figures

Table 1: Replicating Adema et al. (2025)

	(1)	(2)	(3)	(4)	(5)	(6)
RD_Estimate	0.253**	0.298**	0.232**	0.283**	0.231 *	0.288 *
Observations	2268	2268	2268	2268	2268	2268
BW above cutoff	9.889	15.30	9.687	15.22	10.28	15.72
BW below cutoff	9.889	15.30	9.687	15.22	10.28	15.72
Effective N above cutoff	210	336	210	336	231	336
Effective N below cutoff	189	315	189	315	210	315
P-value	0.0104	0.0136	0.0194	0.0195	0.0696	0.0637

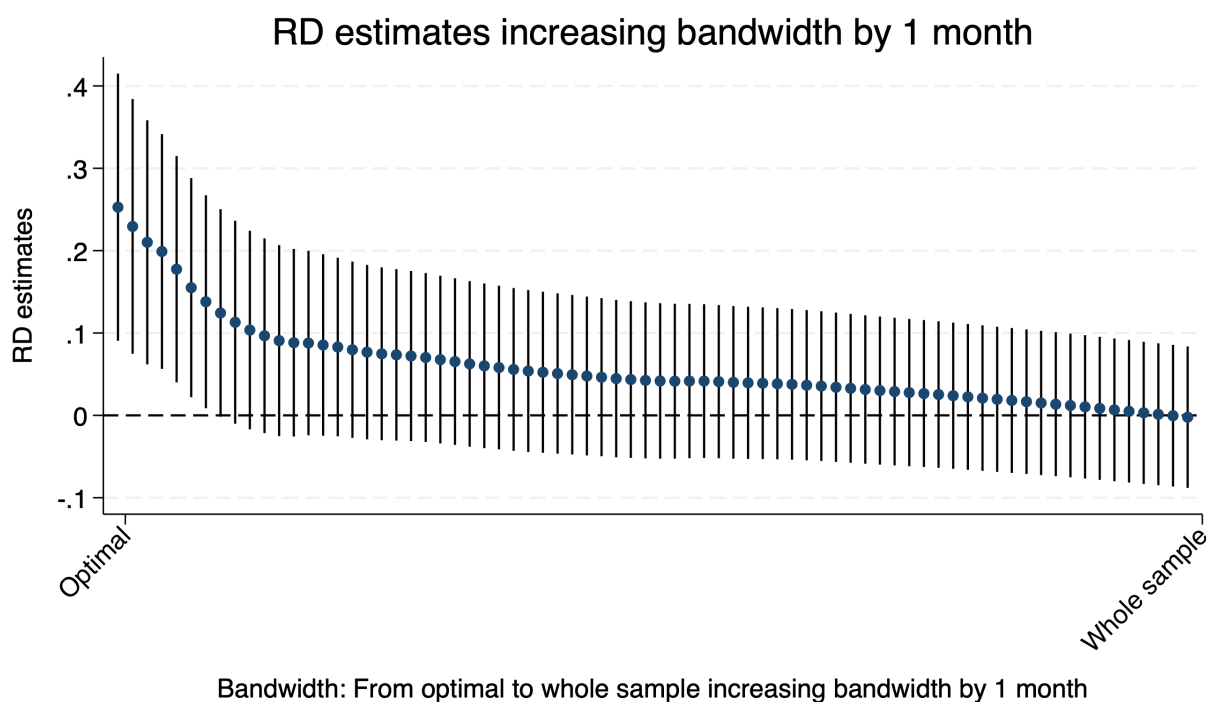
Table 2: Main results of Ciacci (2024) using methodology in Ciacci (2025)

	(1)	(2)	(3)	(4)
	Logs	IHS	Logs	IHS
RD_Estimate	0.257*	0.334*	0.558**	0.711**
Observations	2268	2268	2268	2268
BW above cutoff	6.329	6.336	7.169	7.215
BW below cutoff	6.329	6.336	7.169	7.215
Effective N above cutoff	147	147	168	168
Effective N below cutoff	126	126	147	147
P-value	0.0853	0.0800	0.0174	0.0177

Table 3: Placebo and diff in disc

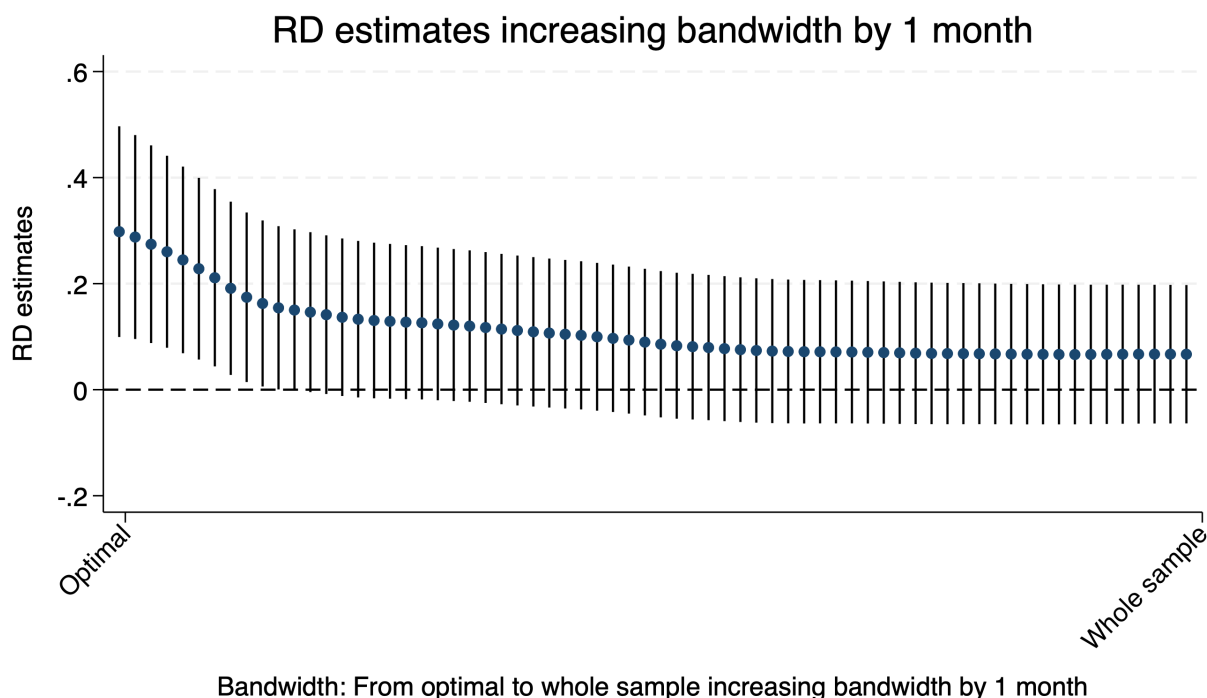
	(1)	(2)	(3)	(4)
	Logs	IHS	Logs	IHS
Placebo RD Estimate	0.128	0.162	-0.069	-0.086
Observations	2268	2268	2268	2268
BW above cutoff	12	12	12	12
BW below cutoff	12	12	12	12
Effective N above cutoff	273	273	273	273
Effective N below cutoff	252	252	252	252
P-value	0.234	0.234	0.672	0.680
Diff in disc	0.129	0.172	0.627**	0.797**
P-value of Diff in disc	0.483	0.464	0.0282	0.0290

Figure 1: RDiT varying bandwidth, with covariates and unweighted



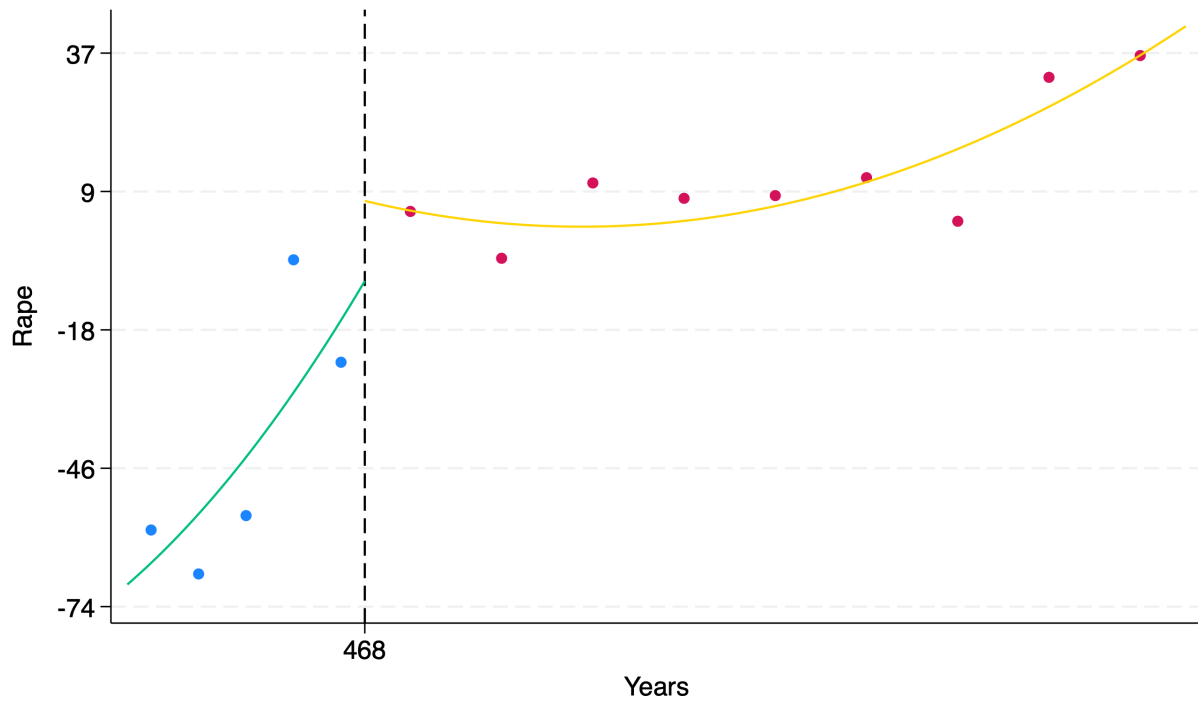
Notes: This figure shows the estimated coefficients and 90% confidence intervals of using the methodology suggested in [Adema et al. \(2025a\)](#) with a polynomial of order 1 to the sample considered in [Ciacci \(2025\)](#). The first estimate corresponds to the optimal bandwidth choice, each column later on adds an extra month and displays the corresponding estimated coefficient. Since there are only 24 months before the introduction of the law, after reaching that amount of months only months on the right of the threshold might be added and are added.

Figure 2: RDiT varying bandwidth, with covariates and unweighted



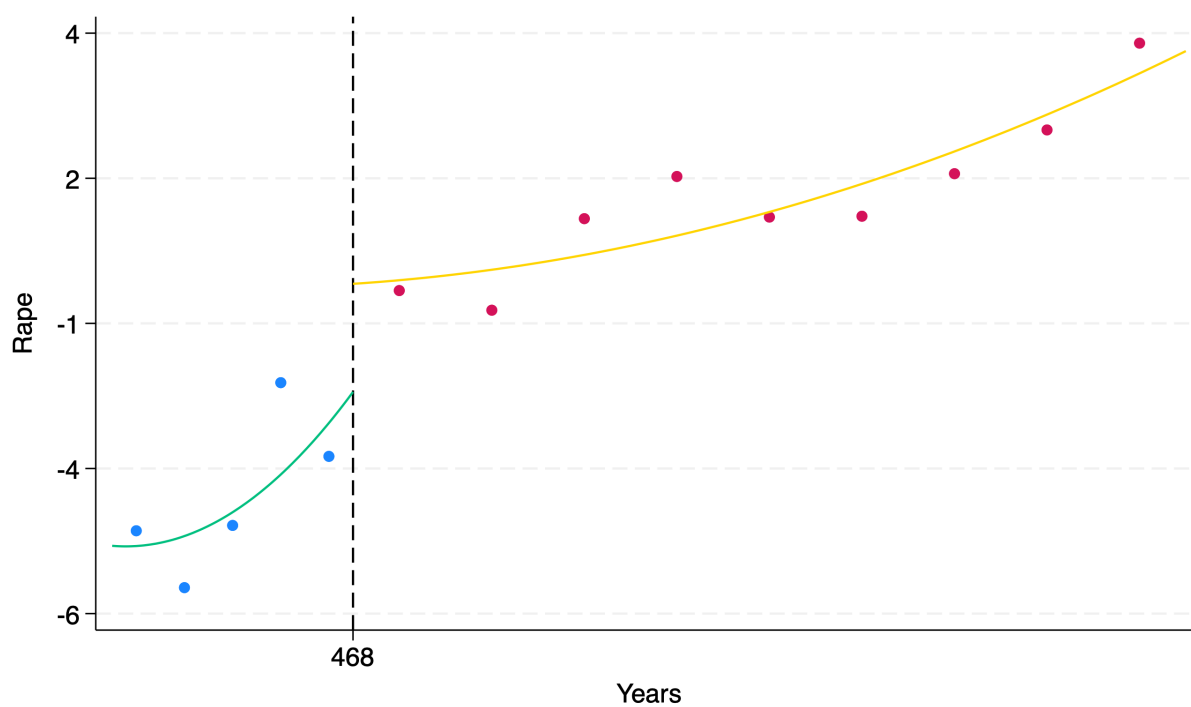
Notes: This figure shows the estimated coefficients and 90% confidence intervals of using the methodology suggested in [Adema et al. \(2025a\)](#) with a polynomial of order 2 to the sample considered in [Ciacci \(2025\)](#). The first estimate corresponds to the optimal bandwidth choice, each column later on adds an extra month and displays the corresponding estimated coefficient. Since there are only 24 months before the introduction of the law, after reaching that amount of months only months on the right of the threshold might be added and are added.

Figure 3: RDiT plot



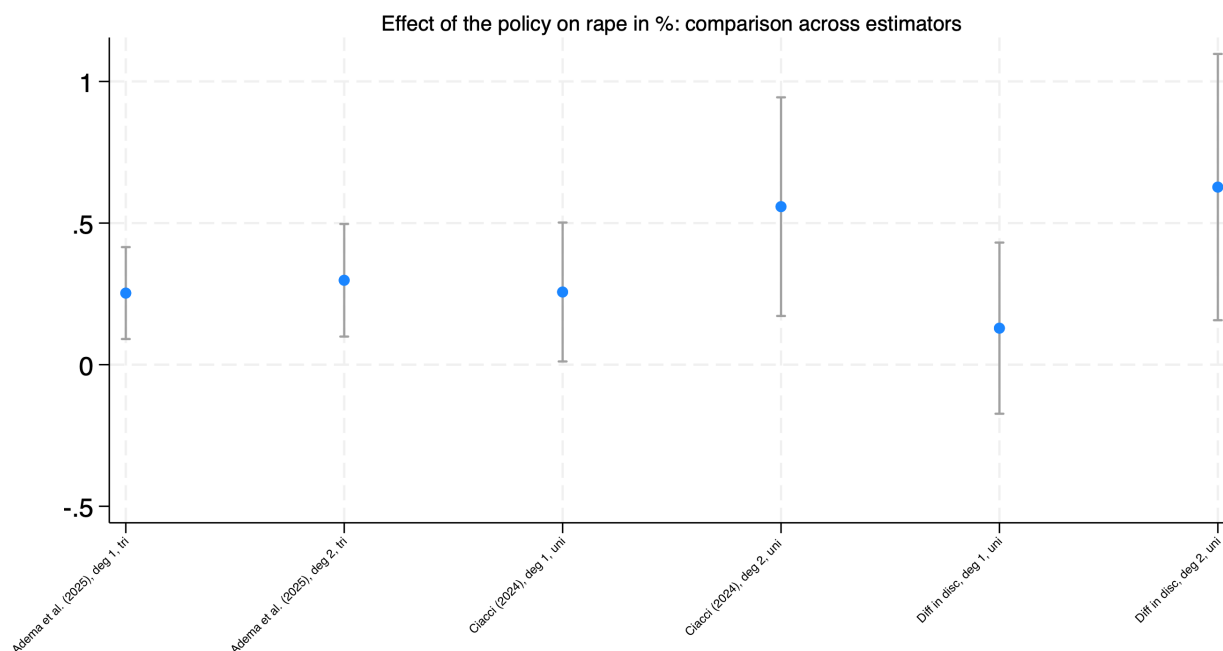
Notes: This figure shows the regression discontinuity plot related to equation (1) a polynomial of order 2 at the aggregated level as in [Adema et al. \(2025a\)](#).

Figure 4: Difference in discontinuity plot



Notes: This figure shows the regression discontinuity plot related to estimating a difference in discontinuity design with equation (1) and a polynomial of order 2 at the aggregated level as in [Adema et al. \(2025a\)](#). Pre-treatment year fixed effects (1998) are subtracted from data of 1998 onwards.

Figure 5: RDiT plot



Notes: This figure compares the main estimates for the regression discontinuity in time and the difference in discontinuity in time frameworks and their corresponding 90% confidence intervals.

Results are statistically equal across estimators. Note that they use polynomials of different orders and different kernels. Five out of six estimators find that the ban led to an increase in rape. The increase in cases of rape due to the ban is not related to the multicollinearity issue discussed in [Ciacci \(2025\)](#) nor to seasonal changes as claimed by [Adema et al. \(2025a\)](#). [Ciacci \(2025\)](#) estimates a different target parameter (nationally representative) than that of [Ciacci \(2024\)](#) (regionally representative). Yet, when the regionally representative parameter is estimated across different methodologies results are robust.

References

- Adema, J., O. Folke, and J. Rickne (2024, March). Rates of reported rapes are higher in January than in December. Technical report.
- Adema, J., O. Folke, and J. Rickne (2025a). RE-ANALYSIS OF Ciacci, R. (2024). Banning the purchase of sex increases cases of rape: evidence from Sweden. *Journal of Population Economics*, 37(2), 1-30. Institute for replication discussion paper series, 226.
- Adema, J., O. Folke, and J. Rickne (2025b). Still no reform effect, just overfitted control functions and inflated precision. Institute for replication discussion paper series, forthcoming.
- Bisschop, P., S. Kastoryano, and B. van der Klaauw (2017). Street prostitution zones and crime. *American Economic Journal: Economic Policy* 9(4), 28–63.
- Cameron, L., J. Seager, and M. Shah (2021). Crimes against morality: unintended consequences of criminalizing sex work. *The Quarterly Journal of Economics* 136(1), 427–469.
- Ciacci, R. (2024). Banning the purchase of sex increases cases of rape: evidence from Sweden. *Journal of Population Economics*: 37.
- Ciacci, R. (2025). Additional evidence on the effects of banning the purchase of sex on cases of rape in Sweden. *Journal of Population Economics*: 38.
- Ciacci, R. and M. M. Sviatschi (2022). The effect of adult entertainment establishments on sex crime: evidence from new york city. *The Economic Journal* 132(641), 147–198.
- Cunningham, S. and M. Shah (2018). Decriminalizing indoor prostitution: Implications for sexual violence and public health. *The Review of Economic Studies* 85(3), 1683–1715.
- Cunningham, S. and M. Shah (2020). Economics of sex work and policy considerations. *Handbook of Labor, Human Resources and Population Economics*, 1–19.
- de Chaisemartin, C. and X. d'Haultfoeuille (2018). Fuzzy differences-in-differences. *The Review of Economic Studies* 85(2), 999–1028.
- de Chaisemartin, C., X. Dhaultfoeuille, and Y. Guyonvarch (2019). Fuzzy differences-in-differences with stata. *The Stata Journal* 19(2), 435–458.
- Della Giusta, M., M. L. Di Tommaso, S. Jewell, and F. Bettio (2021). Quashing demand or changing clients? evidence of criminalization of sex work in the united kingdom. *Southern Economic Journal* 88(2), 527–544.

- Ekberg, G. (2004). The swedish law that prohibits the purchase of sexual services: Best practices for prevention of prostitution and trafficking in human beings. *Violence against women* 10(10), 1187–1218.
- Gao, H. and V. Petrova (2022). Do prostitution laws affect rape rates? evidence from europe. *The Journal of Law and Economics* 65(4), 753–789.
- Hausman, C. and D. S. Rapson (2018). Regression discontinuity in time: Considerations for empirical applications. *Annual Review of Resource Economics* 10, 533–552.

Appendix A A sample of unprofessional and disrespectful passages in [Adema et al. \(2025a\)](#)

Abstract (Page 1)

“The paper’s **faulty** policy evaluation and top Altmetric score in the *Journal of Population Economics* underscore the importance of retraction.”

The use of “faulty” is unnecessarily pejorative, and mentioning the Altmetric score is irrelevant to methodological critique.

Footnote (Page 1)

“...new results by the author (with the **potentially misleading title** ‘Additional evidence on the effects of banning the purchase of sex on cases of rape in Sweden’; Ciacci 2025).”

The phrase “potentially misleading” implies deceptive intent.

Conclusion (Page 14)

“The conclusion and policy advice in Ciacci (2024) are **erroneous and misleading** for policymakers and researchers. We **strongly recommend** that the journal retract this paper.”

This characterization construes the critique as an ethical condemnation rather than a substantive methodological debate. Notably, an independent scientific review committee, upon thorough examination of the evidentiary record, determined that retraction of the article was unwarranted.

Methodologically problematic passages

Main analysis (Page 4)

“A fundamental error renders both these estimates invalid. Equation (1) is impossible to estimate due to perfect multicollinearity...”

Here the key point should be to explore whether multicollinearity affects estimates. While, a priori knowledge of [Adema et al. \(2025a\)](#) determinnes the estimates are invalid.

Difference in discontinuity(Page 7)

“Ciacci’s claim that his Diff in RDiT analysis removes seasonality is **misleading**.”

“Misleading” suggests intentional deception rather than potential methodological misunderstanding.

Event study (Page 10)

“The chosen method differs from all other analyses in the paper...”

This implies inconsistency without considering legitimate reasons for methodological variation. Namely, [Adema et al. \(2025a\)](#) omit that the event study analysis relied on treatment variation not comparable to the other analyses in the paper. This variation arose from the exact timing of fine events across regions.

Unsubstantiated claims

IV Analysis (Page 13)

“Since we do not have access to complete flight data for each airport, we do not fully correct this analysis.”

This admission undermines the strength of their critique while still making definitive claims elsewhere.

Appendix B A sample of unprofessional and disrespectful passages on social media

Figure A.1: A sample of their online posts

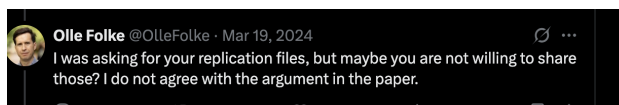


Figure A.2: A sample of their online posts

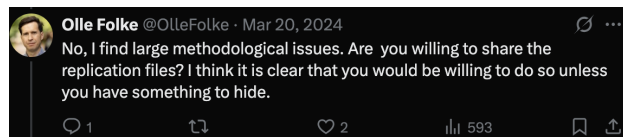


Figure A.3: A sample of their online posts

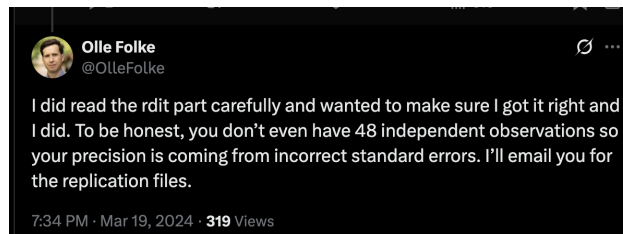


Figure A.4: A sample of their online posts

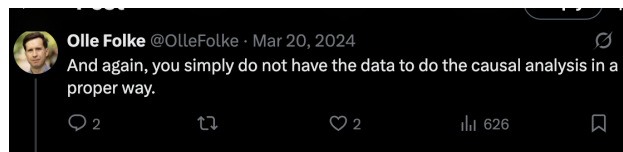


Figure A.5: A sample of their online posts

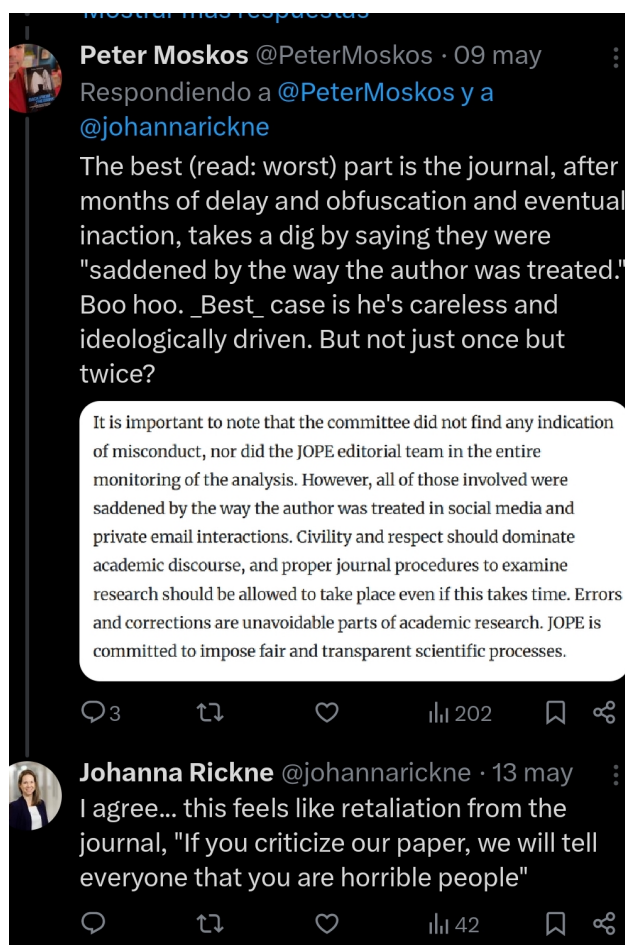


Figure A.6: A sample of their online posts

