



**EJW**  
ECON JOURNAL WATCH  
Scholarly Comments on  
Academic Economics

# Econ Journal Watch

## Scholarly Comments on Academic Economics

### Volume 20, Issue 2, September 2023

#### COMMENTS

A Critique of “*The Birth of a Nation*: Media and Racial Hate”

*Robert Kaestner*

[214–233](#)

Temperature Shocks and Economic Growth: Comment on Dell, Jones, and Olken

*David Barker*

[234–253](#)

The Environmental Economy of the Soviet Famine in Ukraine in 1933: A Critique of Several Papers by Natalya Naumenko

*Mark B. Tauger*

[254–303](#)

Response to Professor Tauger’s Comments

*Natalya Naumenko*

[304–313](#)

The Limitations of Growth-Optimal Approaches to Decision Making Under Uncertainty

*Matthew C. Ford and John A. Kay*

[314–334](#)

Reply to “The Limitations of Growth-Optimal Approaches to Decision Making Under Uncertainty”

*Oliver Hulme, Arne Vanbooyweghen, Colm Connaughton, Ole Peters, Simon Steinkamp, Alexander Adamou, Dominik Baumann, Vincent Ginis, Bert Verbruggen, James Price, and Benjamin Skjold*

[335–348](#)

Reply to LeCraw, Montanera, and Mroz on Hospitals’ Malpractice Claims and Costs

*Maayan Yitshak-Sade, Allen Kachalia, Victor Novack, and Michelle M. Mello*

[349–356](#)

A Rejoinder on the Effects of a Communication-and-Resolution Program on  
Hospitals' Malpractice Claims and Costs  
*Florence R. LeCraw, Daniel Montanera, and Thomas A. Mroz* 357–362

Rejoinder to Barkowski and McLaughlin  
*Aaron M. Gamino* 363–367

## CHARACTER ISSUES

A History of Classical Liberalism in the Netherlands  
*Edwin van de Haar* 368–401

## WATCHPAD

*To Russia with Love: Boris Chicherin's 1857 "Contemporary Tasks of Russian Life"*

Foreword to "Contemporary Tasks of Russian Life"  
*G. M. Hamburg* 402–408

Contemporary Tasks of Russian Life  
*Boris N. Chicherin* 409–437

Pierre de Boisguilbert: Prime Excerpts and Some Correspondence  
*edited by Benoît Malbranque* 438–449

SSRN and medRxiv Censor Counter-Narrative Science  
*Jay Bhattacharya and Steve H. Hanke* 450–459

*Journal of Accounting Research's* Report on Its Own Research-Misconduct  
Investigation of an Article It Published  
*Daniel B. Klein* 460–464

What Are Your Most Underappreciated Works?: Second Tranche of Responses  
*Andrew Gelman, Robert Kaestner, Robert A. Lawson, George Selgin, Ilya Somin, and  
Alexander Tabarrok* 465–468



## A Critique of “*The Birth of a Nation*: Media and Racial Hate”

Robert Kaestner<sup>1</sup>

[LINK TO ABSTRACT](#)

*The Birth of a Nation* by D. W. Griffith was widely hailed as a major innovation in filmmaking and was placed on the National Society of Film Critics’ influential list of *100 Essential Films* (Carr 2002). Dave Kehr, a prominent movie critic, wrote the following in 2011 in the *New York Times* about the film:

A motion picture of unprecedented scale, ambition and formal assurance, “The Birth of a Nation” was the first to open at a legitimate theater on Broadway (the Liberty, absorbed not long ago into the franchised chaos of 42nd Street) with reserved seats, a 40-piece orchestra and an appropriately elevated ticket price: a shockingly high \$2, at a time when a typical admission charge was 15 cents. (Kehr 2011).

*The Birth of a Nation* was first shown in 1915. There were showings of the film across the country between 1915 and 1919 and then again during the 1920s and during the 1930s when sound was added to the film. The 1915 to 1919 showings of the film were a huge commercial success, with millions of viewers.

*The Birth of a Nation* has also been widely reviled for its racism, glorification of the Ku Klux Klan’s activities during the period of Reconstruction, and inaccurate history (Franklin 1979; Rogin 1985; Staiger 1992). Interestingly, at the same time *The Birth of a Nation* was being staged, there was a rebirth of the Ku Klux Klan in 1915, and this incarnation of the Klan is referred to as the second Ku Klux Klan to distinguish it from its Reconstruction Era counterpart (Mecklin 1924).

In a recent article in the *American Economic Review*, Desmond Ang (2023)

---

1. University of Chicago, Chicago, IL 60637.

combines these two facts to derive and test the hypothesis that screenings of *The Birth of a Nation* between 1915 and 1919 was a cause of the rise of the second Ku Klux Klan. Ang (2023, 1428) claims to provide causal evidence that the screening of the film *The Birth of a Nation* in a county between 1915 and 1919 is associated with:

1. an increase in the probability that a county had a lynching or race riot in months after the screening;
2. whether a county had a chapter ('klavern') of the second Klan before 1930; and
3. whether a county was home to a number of hate groups and was the site of hate crimes in the first two decades of the 21st century, approximately 80 years after the film was screened.

In this article I evaluate the validity of the three claims enumerated above. An important reason for doing so is that scholarly claims about things like film screenings causing hate crimes could provide support for laws or policies that restrict speech and repress voluntary association. Ang (2023, 1428) writes that research "examining political media has documented the persuasive effects of radio propaganda and slanted news on racist, nationalist, and xenophobic sentiments..." This paper [demonstrates] the socially harmful spillovers of entertainment media." If Ang were soundly to demonstrate harmful effects of cultural activity, such as film, that could be useful knowledge, although whether evidence from a film shown in 1915 would be applicable to today is questionable. But we should be concerned about such research claims being used to breach our sacred regard for the freedom of speech, the press, expression, and so on. Ronald Coase (1974) worried that the freedom of speech and expression could be assaulted using the kit of 'market failure' arguments used to assault freedom of trade and association. An allegiance to free speech does not deny that a book or movie can be bad for humankind. Rather, such allegiance is based on the fallibility and potential abuse of censors; the need to instill a toleration for disagreement, and the art of testing, challenging, and correcting bad speech with good speech. Coase teaches us to take responsibility for "the total effect" (Coase 1960, 2, 43–44) of a supposed fix to a problem. If you recognize that free speech sometimes goes wrong, recognize too that restricting speech sometimes goes wrong. The upshot of this discussion is that evidence purportedly demonstrating harms of free expression should be highly scrutinized, which is the purpose of this article.

To assess Ang's (2023) claims, I first review the historical scholarship documenting the causes of the rise and fall of second Klan to dispute Ang's claim that screenings of *The Birth of a Nation* between 1915 and 1919 caused the explosive

growth and equally dramatic decline of Klan in the 1920s (see Figure 1 in Ang 2023). Second, I evaluate the validity of Ang's empirical evidence on the effect of screenings of *The Birth of a Nation* between 1915 and 1919 on the rise of the Klan and the extent of hate crime in the first two decades of the 21st century. Finally, I critically review Ang's empirical analyses underlying his claim that screenings of *The Birth of a Nation* between 1915 and 1919 caused an increase in lynchings and race riots in the months after screenings.

Overall, I conclude, firstly, the extensive scholarly analysis that had already addressed the rapid rise (1919–1924) and rapid fall (1925–1927) of the second KKK already provides highly satisfactory explanations of the episode. Second, Ang's empirical analysis is ill-conceived and poorly executed, leading him to make unsupported and misleading claims about the role of *The Birth of a Nation* in the revival of the second Klan. The same faulty empirical framework invalidates Ang's conclusion that screenings of *The Birth of a Nation* between 1915 and 1919 had a persistent effect manifesting in a higher probability of having a Klan chapter in 1960 (the third incarnation of the Klan) and more hate crimes in the new millennium. I also conclude that Ang's evidence related to the effect of screenings of *The Birth of a Nation* between 1915 and 1919 on lynchings and race riots is based on low-quality data and analyses that lacked sufficient statistical power to be informative.

As described in a body of substantial and compelling historical analysis, the rise of the second Klan from 1919 to 1924 was the result of extensive recruiting activities that exploited existing racial hatred, anti-Catholic sentiment, and xenophobia. The historical record also makes clear that *The Birth of a Nation*'s role in the Ku Klux Klan was as a propaganda tool that had minimal effect on Klan membership between 1915 and 1919. Given that minimal effect between 1915 and 1919, the scientific plausibility motivating Ang's empirical analysis is weak, which underscores the need to thoroughly assess the validity of the evidence suggesting a causal link between screenings of *The Birth of a Nation* and the growth in second Klan membership.

## The historical scholarship

The historiography on the origin and growth of the second Ku Klux Klan is rich and extensive (see, among others, Mecklin 1924; Jackson 1992; MacLean 1994; Stokes 2007; Rice 2008; Weinberger 2011; Gordon 2017; Hernandez 2019). It is well documented, for example, that William Joseph Simmons, who was the founder of the second Klan in 1915, had long planned a revival of the Klan and that he was independently reviving the Klan prior to the screening of *The Birth of a Nation* in Atlanta in 1915. Here is John Mecklin (1924) describing Simmons:

Colonel William Joseph Simmons, the founder of the modern Klan, tells us that for twenty years he had given thought to the creation of an order standing for a comprehensive Americanism that would blot out the Mason and Dixon's Line. Fascinated as he was from boyhood by the romantic story of the old Klan of Reconstruction days, which is looked upon in the South as the savior of Southern civilization, he called the new order the Knights of the Ku Klux Klan. (Mecklin 1924, 20)

Simmons had applied for a state charter in October of 1915 and organized a meeting on Stone Mountain on Thanksgiving in 1915 to create the second Klan (Bernstein 2023). Simmons used the screening of *The Birth of a Nation* in Atlanta on December 6, 1915, to glorify his vision of the Klan. By the end of the run of the film in Atlanta, the Atlanta Klan had 92 members (ibid., 178). Mecklin (1924, 21) writes: "For five years the Klan seems to have passed an uneventful existence, spreading very slowly and making no great impression upon the country." Similarly, Melvyn Stokes (2007, 234–35) notes: "the really rapid growth of the Klan did not occur in the early years when *The Birth of a Nation* was at the peak of its influence and availability. By 1919, the Klan had only a few thousand members. Not until the summer of 1920, with the hiring of publicity agents Edward Young Clarke and Mrs. Elizabeth Tyler, did the real expansion of the Klan begin."

The outsized role of Clarke and Tyler in growing Klan membership is discussed in many works (Mecklin 1924; MacLean 1994; Rice 2008; Fryer and Levitt 2012; Gordon 2017). Indeed, Tom Rice (2008) documents how the Klan used Klan-owned newspapers (e.g., *Fiery Cross*, *The Watcher on the Tower*), Klan-produced films (*The Traitor Within* and *Toll of Justice*), and other films (*The Face at Your Window*) to recruit members. These media activities were undertaken during the early 1920s and coincide with the dramatic increase in Klan membership between 1919 and 1924. A notable media event that historians point to as a cause of the Klan's national expansion was the exposé of Klan activities published in the *New York World* in 1921 (Chalmers 1965; Jackson 1992; Lay 1992; MacLean 1994; Stokes 2007; Kneebone 2015). The *New York World* published twenty-one articles on consecutive days detailing Klan activities, and these articles were distributed by major newspapers across the country (e.g., the *Boston Globe*, *Plain Dealer*, *Houston Chronicle*, and *Seattle Times*). These articles resulted in Congressional hearings in which Simmons testified and brought attention to the Klan nationally, which facilitated Klan expansion.

Arguably, the most important cause of the growth in Klan membership was the Kleagles, or recruiters, that were enlisted by Clarke (Jackson 1992; MacLean 1994; Rice 2008). By 1921, there were 200 Kleagles around the country who were on a commission that allowed them to keep \$4 of the \$10 membership fee (Jackson 1992; Hernandez 2019). Clarke himself was paid \$2.50 for each recruit and repor-

tedly made \$30,000 some weeks (Jackson 1992). It was during this period and largely because of the Kleagles and the other organizing activities of the Klan that Klan membership and chapters expanded.

The rapid decline in Klan membership came after 1924, collapsing nearly to zero by 1928. The historical scholarship is, again, clear (Jackson 1992; Stokes 2007; Pegram 2011; Gordon 2017; Hernandez 2019). According to these historians, the decline in membership was due to the fratricidal struggles among the Klan leadership, revelations about the organization's financial irregularities, and high-profile criminal cases such as the prosecution of Klan members in Louisiana for the Mer Rouge murders, the *New York World's* exposure of the arrest of Clarke and Tyler for possession of alcohol during Prohibition, and the arrest of prominent Klan leader David Stephenson for rape and murder. Denunciation and mockery by opinion leaders such as H. L. Mencken (1924) no doubt also spurred the decline.

In short, screenings of *The Birth of a Nation* between 1915 and 1919 are not part of the traditional historiography of the rise and fall of the second Klan. If anything, showings of the film during its second run in the 1920s may have been used as a tool or mediator by Klan organizers (Simcovitch 1972; MacLean 1994; Stokes 2007; Rice 2008). John Hope Franklin writes: "Thus, Birth of a Nation was the midwife in the rebirth of the most vicious terrorist organization in the history of the United States" (1979, 431). Franklin's reference to a "midwife" is notable because it highlights the fact that *The Birth of a Nation* was a recruiting device of Klan organizing activities—i.e., a tool used by the Klan leadership. Even this characterization, however, does not apply to screenings between 1915 and 1919.

## **Did screenings of *The Birth of a Nation* between 1915 and 1919 increase Ku Klux Klan membership?**

Ang (2023) does not articulate the explanations of the second Klan's rapid rise and decline as found in the traditional historiography on the subject. He therefore neither integrates his emphasis on 1915–1919 screenings of *The Birth of a Nation* as a cause of the second KKK with traditional explanations, nor compares the persuasiveness of his explanation against the traditional explanations. This, perhaps, is not that surprising given the formulaic style that has come to dominate articles published in the American Economic Association (AEA) journals. The AEA style eschews critical literature reviews in favor of two or three paragraphs that mention how a study is related to other studies, even studies only tangentially related. For example, Ang (2023) simply references studies in "media economics"



such as a study of Fox news on voting (DellaVigna and Kaplan 2007) or how Serbian radio affected the presence of anti-Serbian graffiti in Croatian towns (DellaVigna et al. 2014). He does not discuss why these studies are relevant to the study of the rise and fall of the second Klan in the 1920s.

Instead of utilizing the extensive literature on the rise and fall of the second Klan, Ang (2023) argues that his study was necessary to address the lack of “causal evidence” on the relationship between screenings of *The Birth of a Nation* between 1915 and 1919 and the origin and growth of the second Klan. The claim that there is a lack of “causal evidence” supposes that there is a widely held view that local screenings were causal. This is not the case, although it is widely noted that the film’s depiction of the Klan influenced Klan’s behavior (e.g., the wearing of white robes) and that it was used by the Klan in its propaganda (Simcovitch 1972; MacLean 1994; Stokes 2007; Rice 2008). As such, it is a mediator between Klan organizers and leadership and Klan membership. Further, “causal evidence” here seems to mean evidence produced using econometric methods. Ang (2023) obtained estimates of the correlation between whether there was a screening of *The Birth of a Nation* in a county between 1915 and 1919 and the probability that a county had a chapter of the second Ku Klux Klan before 1930. Let me reiterate that by 1930 Klan chapters had all but disappeared entirely, so Ang’s dependent variable here is about the county in 1930 *having had* a chapter.

The disconnect between historical explanations of the rise and fall of Klan membership between 1919 and 1927 and Ang’s analysis, with its dependent variable anchored in 1930, is obvious. The motivation for Ang’s analysis is indicated in this passage: “the movie’s staggered and incomplete distribution provide an ideal setting for hypothesis testing” (Ang 2023, 1425). In short, Ang (2023) found what, to him, looked like a natural experiment that could be exploited using methods that would appeal to economist reviewers who value empirical analysis over wide-ranging historical scholarship, even when the scientific plausibility of the empirical analysis is questionable, as in this case. What’s more, despite saying that he had found “an ideal setting for hypothesis testing,” Ang’s analyses of the presence of a Klan chapter by 1930 and the extent of hate crimes in the new millennium do not, in fact, exploit the “staggered” screenings of *The Birth of a Nation*.

Ang is well aware that he faces a causality problem: It could well be that, rather than screenings being the cause of the rebirth of the Klan, there is another set of variables that causes both screenings and the probability of ever having a chapter of the second Klan by 1930. To obtain estimates of the association between the screening of *The Birth of a Nation* between 1915 and 1919 in a county and the probability that a county had a chapter of the second Klan before 1930, Ang (2023) used regression methods using county-level data. The key independent variable is on dates and location of screenings of *The Birth of a Nation*. Ang (2023, 1433)



describes the collecting of the data: “I collected data on newspaper advertisements for screenings of the movie from three online repositories of digitized historical newspapers... To supplement the newspaper data, research assistants also searched for “Birth of a Nation” in digitized copies of the *Moving Picture World* from 1915 to 1919.” Note that screenings that were unadvertised in a newspaper, or in a county without a newspaper, went unrecorded, which is a point I return to shortly.

Ang (2023) regressed an indicator for whether a county had had a Klan chapter by 1930 on an indicator for whether the county had a screening of *The Birth of a Nation* between 1915 and 1919. This analysis does not exploit the staggered screenings of the film because there were no Klan chapters prior to 1915. In its place, to address the causality problem, Ang (2023) uses an instrumental variable approach in which, instead of the actual screening of *The Birth of a Nation* in a county, he uses a *predicted probability* that there was a screening in a county. Ang predicted the probability that a county had a screening between 1915 to 1919 conditioned on whether a county had a theater in 1914, thus creating two sets of counties, namely, those with a theater as of 1914 and those without a theater. Supposing that a screening of *The Birth of a Nation* could only occur if a county had a theater,<sup>2</sup> the presence of a theater in a county by 1914 was strongly correlated with whether a county had a screening.

The validity of the instrumental-variables (IV) approach depends on whether there were other unmeasured factors associated with whether a county had a theater in 1914 that would also influence Klan membership and chapter formation. For example, given that counties with a theater in 1914 were relatively large and urban (Table 1 in Ang 2023), Klan organizing activities such as Kleagle recruiting efforts may have been focused on these counties. However, any number of unmeasured factors that differed by whether a county had a theater in 1914, such as socioeconomic status, could also affect Klan membership and chapter formation.

---

2. Ang (2023) seems to presuppose that a screening required a theater, but does not speak to whether some of the 1915–1919 newspaper advertisements of screenings were for screenings in locations in, for example, auditoriums, common rooms, etc., of schools, churches, lodges, civic centers, commercial buildings, and so on. A cursory search of the Newspaper.com website of historical newspapers turns up examples such as the *News & Observer* (Raleigh, N.C.), Nov. 12, 1915, p. 4, which reports a screening at the Academy of Music (not a conventional movie theater but instead an opera house); the Majestic Theater in La Crosse, Wis., was a vaudeville house when it screened *Birth of a Nation* (*La Crosse Tribune*, Nov. 12, 1915, p. 16); the Tremont Theater in Boston which showed the film 252 times in 1915 was a legitimate theater until 1947; the Victoria Theater in Dayton, Ohio, which had showings in 1917 was a legitimate theater that did not convert to full-time movie exhibition until 1930; the film’s run in Charlotte, N.C. was at the City Auditorium, a venue which “gave Charlotte a lofty space and 4,500 seats for all manner of civic gatherings from basketball to band concerts by John Philip Sousa and arias by tenor Enrico Caruso.” And as mentioned earlier, screenings that were not advertised would be missed. The resultant measurement error in screenings is likely to have an effect. Indeed, estimates in Ang’s Appendix Table A.3 show that alternative measures of screenings have a large effect on estimates, for example by a factor of 215 percent.

In fact, in Table 1 below, I show that counties with and without a theater in 1914 differed in several ways that are plausibly related to Klan membership (e.g., literacy rates).

We know that Klan members were more likely than the general population to be professionals, be Protestant, and of course be white (Fryer and Levitt 2012; Mecklin 1924). It is possible that some measure of diversity within the local population would provide salient scapegoats, inflame passions, and make Klan chapter formation more likely. My point is not to speculate on true factors, but simply to highlight the difficulty of trying to tease out, statistically, whether a variable like percentage Protestant or percentage white matters to Klan-chapter formation and other dependent variables.

Ang (2023) provides what he claimed was evidence against the possibility that theater counties were in some relevant way different than non-theater counties. First, he shows that a few demographic characteristics (e.g., black population in 1860) and political characteristics (Democratic vote share in 2012) prior to 1914 do not differ between counties with and without a theater in 1914, which was the predictor (i.e., instrument) for whether there was a screening of *The Birth of a Nation* between 1915 and 1919. This is a very limited assessment, as Ang himself acknowledges (2023, 1441). Ang examined only a few characteristics that are not particularly intuitive causes of Klan membership.

Second, Ang (2023) examined whether the presence of a theater in a county in Kansas, which banned the film, was associated with the probability of having had a Klan chapter by 1930 in that county. Standard errors from this analysis are so large as to make the entire analysis uninformative. For example, estimates suggest that the presence of a theater in a county in Kansas is associated with a 40 percent lower probability of having a Klan chapter, which is about the same magnitude, although opposite signed, as the estimate of the effect of having a theater on the probability of having a Klan chapter in counties outside Kansas. Despite the large magnitude of the estimate for Kansas, which suggests a faulty research design, it was statistically insignificant because of the lack of statistical power. In short, this analysis is severely under-powered, and unpersuasive.

The third piece of evidence Ang presents is an analysis correlating the year a theater opened in a county and the probability of having had a Klan chapter in that county by 1930. Ang (2023, Figure 5) reported that only if a theater opened before 1919 was there a higher probability of having had a Klan chapter by 1930. This simply shows that probability of having had a Klan chapter by 1930 is higher in counties that opened a theater earlier with, unintuitively, counties opening a theater prior to 1907 having the highest probability of having had a Klan chapter by 1930. The question as to why there was a non-trivial relationship between the year a theater opened in a county and the probability of having a Klan chapter by

1930 was unaddressed by Ang despite its bearing on the validity of the instrumental variables approach. There is no direct reason for the probability of having a Klan chapter by 1930 to be higher in counties with a theater by 1906 than counties that had a theater in 1914. The fact that there is such a relationship strongly suggests that *unmeasured influences* associated with both the year a county had a theater and the probability of having a Klan chapter by 1930 are present. The same problem with the main analysis examining the association between the presence of a theater by 1914 and a Klan chapter prior to 1930 plagues this analysis—counties that have a theater, or a theater by a given year, are likely to differ, for example, be exposed differentially to all the Klan’s early-1920s organizing activities.

Ang (2023) also conducted what he terms a placebo analysis in which he obtained the association between screenings of the 1918 film *Mickey* in a county between 1915 and 1919 and the probability that a county had had a Klan chapter by 1930. The expectation is that the placebo treatment—i.e., screening of *Mickey*—will not have an effect. Contrary to this expectation, Ang’s results (Online Appendix Table A.IX) from the exact same IV approach used for *The Birth of a Nation* indicate that screenings of *Mickey* between 1915 and 1919 in a county was associated with a large increase in the probability that a county in 1930 had had a Klan chapter. The magnitude of association was almost identical to that obtained for screenings of *The Birth of a Nation*. This suggests that a study focusing on theater screenings of *Mickey* and the rise of the second Klan would have led to the conclusion that *Mickey* caused the revival of the Klan—i.e., similar “causal evidence” underlying his claim about *The Birth of a Nation*. Yet, Ang (2023, 1452) comes to precisely the opposite conclusion of what the data show: “These results provide further confidence that differences in Klan support are not driven by unobserved factors correlated with theater locations or the arrival of popular movies more generally, but rather by *The Birth of a Nation*’s specific racist influence.” It is surprising that this damning evidence was glossed this way by Ang. Equally surprising is that this clear evidence of an invalid research design withstood editorial scrutiny.<sup>3</sup>

---

3. Ang (2023, 1451–1452) writes: “I perform placebo tests using screenings of *Mickey*, the most widely seen movie of the era. Online Appendix Table A.IX shows that while areas that screened *Mickey* were more likely to have a klavern in the future, this is due to the large overlap in distribution areas of the two films. Roughly 50 percent of counties that screened *Mickey* also screened *The Birth of a Nation*. Controlling for *The Birth of a Nation* screenings, the effect of *Mickey* showings is near zero and insignificant. Similarly, *Mickey* showings are not associated with increased Klan presence in Kansas (where *The Birth of a Nation* was banned) or in other counties where *The Birth of a Nation* was not shown. This is true for both OLS and IV models. These results provide further confidence that differences in Klan support are not driven by unobserved factors correlated with theater locations or the arrival of popular movies more generally, but rather by the *The Birth of a Nation*’s specific racist influence.” Ang (2023, 1451) claims that screening of *Mickey* was as strongly associated with whether a county had had a Klan chapter by 1930 as was screening of

There are many other placebos that can be used, and I tried some on the dependent-variable side of the equation, that is, whether the presence of a theater in 1914 is correlated with an outcome in 1930 that is clearly not caused by the presence of a theater, or the screening of a film *The Birth of a Nation* (or *Mickey*). I selected a sample of all males between the ages of 21 and 64 from the 1930 Census (5% sample). I then aggregated the data to the county level, and for each county I calculated the proportion of the sample that worked; was illiterate; owned a house; had a radio in the house; and among workers, the proportion that were self-employed. For each outcome, I estimated a reduced-form model almost identical to the one in Ang's Table 4. Specifically, I obtained estimates of the association between whether a county had a theater in 1914 and the outcomes just described using one of the same models as Ang (2023).<sup>4</sup> Table 1 presents the estimates.

**TABLE 1. Estimates of the association between having a theater in 1914 and socioeconomic outcomes**

	Worked	Illiterate	Owned a house	Had radio	Self-employed
Reduced form, any theater in 1914	-0.021** (0.002)	-0.009** (0.002)	-0.017** (0.004)	0.050** (0.004)	-0.096** (0.006)
IV=(Reduced form/0.277) (First stage=0.277 taken from Ang (2023, Table 3 Col. 2))	-0.076	-0.032	-0.061	0.180	-0.347
Mean of dep. variable in control counties	0.93	0.09	0.49	0.18	0.53
Number of observations = 3,075					
<i>Notes:</i> Estimates are from a model that includes all the variables in the model used by Ang (2023, Table 3 Col. 2). The implied IV estimate is equal to the reduced-form estimate shown in the first row of the table divided by the first-stage estimate of the effect of having a theater in 1914 on whether <i>The Birth of a Nation</i> was screened in that county between 1915 and 1919, which is taken directly from Ang's Table 3 Column 2. Standard errors are in parentheses. Standard errors were constructed using robust-cluster methods with clustering on the county. ** indicates p-value<0.05.					

Estimates in Table 1 provide consistent and strong evidence that the IV approach of Ang (2023) is invalid. For every outcome measured in 1930, estimates

*Birth* because approximately half the counties that screened *Mickey* also screened *Birth*. This argument does not refute the damning evidence that screenings of *Mickey* were strongly associated with the probability that a county had a Klan chapter by 1930. Instead, Ang's defense illustrates that the analysis is not informative or useful as a placebo because it is infeasible to separate the two relevant statistical associations (namely, the association of *Mickey* with the probability that a county had had a Klan chapter by 1930 and the association of *Birth* with the probability that a county ever had a Klan chapter by 1930). Further, as already discussed, evidence from Kansas is unpersuasive because the analysis is so lacking in statistical power that it is unable to reject huge associations and is therefore uninformative.

4. I used the data provided by Ang (2023), which allowed me to estimate a model that is identical to that in column 2 of Table 3 and includes: state fixed effects and the county level controls included by Ang (2023): a quadratic of total population, population density, Black population, Black population share, US-born population share, and the share of individuals who would have been of draft-eligible age during World War I. These data were made available by Ang (2023).

of the association between whether a county had a theater in 1914 and the outcome are statistically significant.<sup>5</sup> Implied IV estimates, which are estimates of the effect of a screening of *The Birth of a Nation* on these outcomes, suggest that some of the effects are quite large. For example, IV estimates indicate that the screening of a *The Birth of a Nation* in a county is associated with a 3.2 percentage-point, or approximately 33 percent, lower probability of being illiterate. The IV estimate for whether a person owned a house indicates that the screening of *The Birth of a Nation* in a county is associated with a 6.1 percentage-point, or approximately 12 percent, lower probability of owning a house. As there is no plausible reason why a screening of *The Birth of a Nation* would cause any of the outcomes in Table 1 (e.g., home ownership), the statistically significant estimates are compelling evidence of a faulty research design. Counties with a theater in 1914 are different from those without a theater in 1914 and these differences cause differences in a range of outcomes in 1930. The same factors that cause home ownership, work, and literacy to differ are also likely to influence Klan membership and chapter formation. Therefore, besides the compelling historical scholarship undermining Ang's claim that screenings of *The Birth of a Nation* caused the rise and fall of the second KKK, Ang's claim is further undermined because of a clearly flawed empirical analysis.

More doubt about the soundness of Ang's conclusions comes from assessing mechanisms that he highlights purportedly linking the screening of *The Birth of a Nation* in a county between 1915 and 1919 to the rapid growth and equally rapid demise of the Ku Klux Klan between 1919 and 1927. Any mechanism would have to explain both the rise and fall of the Klan membership. Ang (2023) mentions a few possible mechanisms. For example, Ang (2023) reported (Appendix Figure A.XXI) that there was an increase in boys named Benjamin—the protagonist of *The Birth of a Nation*—in the decade between 1915 and 1925 versus 1900 to 1904. Why this is evidence of “media imitation” (Ang 2023, 1452) that can explain the rapid rise and demise of the Ku Klux Klan remains elusive to this writer. Moreover, there is no statistical difference in the proportion of boys named Benjamin between 1910–1914, 1915–1919, 1920–1924 and 1925–1929 that would reflect some unstated and not obvious linkage between child naming and Ku Klux Klan membership.

Another mechanism mentioned by Ang (2023) is that *The Birth of a Nation* changed racial attitudes. Given the widespread and open racism at the time it seems unlikely that this mechanism was the link between the film and rise of the second Klan. Rather, it is far more believable that *The Birth of a Nation* was one of many representational forms that reinforced existing racial stereotypes, such that its

---

5. Estimates using county population as weights are qualitatively similar and always statistically significant, although magnitudes differed moderately. Ang (2023) did not weight by county population in his analysis.

general adherence to dominant cultural ideas rendered its impact minimal. The following quote summarizes the historical view on this issue:

What *Moving Picture World's* reviewer W. Stephen Bush called the film's 'undisguised appeal to race prejudices' was in itself unexceptional in a society which asserted racial difference as an absolute marker of identity and in which 'the figure of the depraved black criminal assumed a prominence in newspapers, fiction, plays, songs, and early films far out of proportion to his actual numbers.' Racial caricature pervaded middle-class white culture, from its advertising and its food packaging to its comics, drug store postcards and children's books, toys and games. 'Everywhere', notes Henry Louis Gates, Jr., '[Americans] saw a black image, that image would be negative.' Propelled by the racial sciences of Social Darwinism and eugenics, segregation was 'etched into the landscape of virtually every American town or city,' and embedded in President Wilson's 1913 resegregation of multiple federal government agencies and in state legislation segregating urban housing and facilities in factories. (Maltby 2023, 49)

Also, if, in fact, the film changed racial attitudes between 1915 and 1919, then why would this have resulted in a lengthy delay (as long as five years) before the increase Klan membership? If racial attitudes were significantly changed by screenings of *The Birth of a Nation*, then it cannot explain the rapid collapse of membership. Did racial attitudes suddenly change in 1925? Indeed, Ang (2023) argues that there was a persistent effect of screenings of *The Birth of a Nation* on racial attitudes. If so, then this would be consistent with a lasting effect on membership, although other forces may have hastened the decline after 1924. Moreover, other evidence of a persistent effect of screenings on racial attitudes in 1946 (Appendix Table A.X) and 1972 (Appendix Figure A.XXV) is not particularly robust (e.g., very large standard errors and many non-significant results) and based on the same faulty IV design described earlier.

Ang (2023) also cites as evidence of changes in racial attitudes a study by Ruth Peterson and L. L. Thurstone (1932) who conducted an experiment in 1931 in Crystal Lake, Illinois, in which 434 children in grades six to twelve were given an assessment of their attitudes toward "the Negro" a week before the children were shown *The Birth of a Nation* and then again, the day after the showing. Results showed a large, significant decrease (100 percent of a standard deviation) in having a favorable attitude toward "the Negro." It is unclear that the three-hour film's day-after effect on racial attitudes of adolescents in 1931 prove that the film significantly changed racial attitudes of adult males in 1915 to 1919 and caused significant behavioral changes years later. However, Douglas Cameron Moore (1971), in a Ph.D. dissertation not cited by Ang, conducted a similar study with



an arguably better research design in Crystal Lake, among 150 students in grades seven through twelve. The students were divided into treatment (100) and control (50) groups. Each group was given the same assessment of attitudes toward “the Negro” as used in Peterson and Thurstone (1933) and a newer assessment developed by Moore. The treatment group was shown *The Birth of a Nation* and the control group was shown an alternative film *Greed*. Results indicated that the screening of *The Birth of a Nation* was not statistically related to a change in attitudes toward “the Negro.” An interesting finding was that, among the treatment group, the screening increased favorable attitudes toward “the Negro,” although the increase was not statistically significant.

Despite the lack of a plausible mechanism, credible empirical evidence and a coherent explanation, Ang (2023, 1453) concludes the section on mechanisms with the following statement: “Though it is difficult to disentangle the extent to which these changes mediate the effect on historical Klan support—as opposed to being a consequence of it—they nonetheless suggest that *The Birth of a Nation* ultimately served to further radicalize racists.” The statement goes beyond the data and makes an unsubstantiated claim.

Ang’s analysis of the relationship between screenings of *The Birth of a Nation* between 1915 and 1919 and the rise of the second Klan is misleading in its conclusion that the former was a cause of the latter. Unfortunately, the patina of ‘credible’ econometric methods, which I have shown to be invalid, persuaded reviewers and editors to publish it in the *American Economic Review*.

## **Did *The Birth of a Nation* increase lynchings and race riots?**

A second purported contribution of Ang (2023) is the analysis of the relationship between screenings of *The Birth of a Nation* in a county between 1915 and 1919 and subsequent lynchings and race riots in that county. For this analysis, a difference-in-difference approach is used that compared the probability of a lynching or race riot before and after a screening of *The Birth of a Nation* in counties that did and did not have a screening. Ang used an event-study specification of the difference-in-differences approach in which the effect of a screening of *The Birth of a Nation* is allowed to differ by time (months) before and after the screening.

To begin, it is important to note that almost none of the statistical evidence presented by Ang of the effects of a screening of the film on lynchings and race riots meets standard thresholds of statistical significance. Consider results presented in Ang’s Figure 3 (2023, 1438). This figure shows difference-in-differences



(event-study) estimates of the effect of a screening of *The Birth of a Nation* on lynchings and race riots. Of the 24 estimates shown, only four are statistically significant. Moreover, two of the four significant estimates are in the pre-period and indicate that lynchings and race riots were higher prior to screening of the film in counties with a screening relative to counties without a screening—in short, there is evidence of a poor research design. The other two statistically significant estimates are for the period six or more months after the screening of the film. Notably, all the statistically significant estimates are found in what Douglas Miller (2023), in his recent review of event-study designs, refers to as “end caps,” which are subject to bias caused by differential pre-trends and dynamic treatment effects.

A second significant problem with the analysis and estimates presented by Ang in his Figure 3 is that there is inadequate statistical power. The confidence intervals for estimates that Ang focuses on as key evidence of an effect of screenings are so large that the analysis is unable to reject an effect size for lynchings that is approximately between 8 and 13 times the mean of the comparison group.<sup>6</sup> The lack of statistical power is even more severe for race riots. The lack of statistical power is so severe that it renders the analysis virtually useless. It should not have been undertaken given the lack of statistical power and it should not have been used as evidence of the effects of screenings of *The Birth of a Nation* on lynchings and race riots, especially because none of the estimates, as huge as they are, are statistically significant. Here too, Ang (2023, 1439) makes an incredible statement about the evidence: “Together, the findings demonstrate the violent racism that *The Birth of a Nation* incited in local communities.” This statement is, in fact, unsupported by the evidence, again raising questions about the effectiveness of the peer review process of the *American Economic Review*.

As for lynchings, estimates in Ang’s Figure 3 indicate that screenings of *The Birth of a Nation* resulted in a large, even legitimately characterized as gigantic, increase in lynchings in the three months after screening. However, by months four and five, the effect of screenings was virtually zero. Then the effect of screenings becomes positive (i.e., fewer lynchings) six months or more after screening. A similar pattern characterizes results for race riots, although in this case the effect goes to zero three months after a screening only to reappear after six or more

---

6. For lynchings, Ang (2023, 1438) highlights estimates in months 0, 1, 2 and 3 after a screening. While it is difficult to see precisely, standard errors of estimates for these periods are between 0.0006 and 0.001. Given these figures, the minimum detectable effect size is between 0.0012 and 0.002. How big are these minimum detectable size estimates relative to the mean of comparison group? Ang (ibid.) reported that the mean probability of lynching for the sample is 0.0003, but it must be significantly lower for the comparison group given the size of the estimated treatment effects (e.g., 0.001) and the size of the treatment group (approximately 20 percent of all counties). A reasonable estimate is that the mean probability of lynching in comparison counties in the pre-period is 0.00015. Therefore, the minimum detectable effect size is between 8 to 13 times the mean of the comparison group.

months. The sudden month-to-month change from a huge effect to virtually no effect is unintuitive and inconsistent with any hypothesis (theory) of the effect of the film and other claims in the article (e.g., persistent change in racial attitudes). However, such a pattern is not surprising when there is little statistical power and extremely poor quality of the data on lynchings and race riots, which is an issue I turn to next.

Data on lynchings used by Ang come from several sources. However, it is widely recognized by scholars that the data, despite best efforts collecting it, is error-ridden. The most salient summary of the problem is by Charles Seguin and David Rigby (2019), which is a source relied on by Ang: “Ultimately, we will never know the full and exact extent of lynching in the United States, as many lynchings are undoubtedly lost entirely to the historical record. As such, our data set remains a work in progress” (Seguin and Rigby 2019, 8). The poor quality of the data on lynchings goes unmentioned by Ang. One potentially important problem with the lynching data is that most of it is based on newspaper reports. However, as Nancy MacLean (1994, 18–19) notes, newspapers were not particularly inclined to report lynchings: “Indeed, newspaper editors in the South, like politicians, tended to quaver in the face of Klan’s power. Clearly, they did not view the order as an innocent analogue of other fraternal lodges. While local papers boosted these, most maintained an eerie silence regarding the Klan’s activities. Few offered outright support, yet neither would they investigate or expose it.” One might speculate on how a local screening of the film might affect the propensity *to report on* a local lynching.

As noted previously, the analysis of lynchings lacks statistical power, and measurement error will be exacerbated by this problem (Gelman and Carlin 2014; Loken and Gelman 2017). Consider the likely seriousness of the problem. As Ang notes, in the six months prior to a screening of *The Birth of a Nation*, there were only two lynchings in all 600 counties that had a screening. Given that Ang used weekly data in his analysis, the number of lynchings indicate that in only two of the 18,000 county-week observations was there a lynching. Ang doesn’t report how many lynchings in this period occurred in comparison counties, but given that there were only 50 or so lynchings a year across 3,000 counties during the period of analysis, and that there were two in the 600 treatment counties, one can guess that it may have been around 23—i.e., in only 23 out of 72,000 county-week observations was there a lynching among comparison counties.<sup>7</sup> As these small numbers suggest,

---

7. If there were 50 lynchings per year across all counties and two lynchings in 600 treated counties in six-month period, then it suggests that comparison counties had 46 lynchings in a year and 23 lynchings in six months. Of course, these are crude calculations, but they illustrate the small numbers of lynchings that underlie the analysis.

even a small number of lynchings missed, or lynchings included that were not actually lynchings, would have a significant impact on estimates. Clearly, lynchings are poorly measured and a rare (and horrible) event during this period. Given this, it was infeasible to conduct a credible statistical analysis. This, problem, however, did not deter Ang and did not seem to matter to reviewers.

The data on race riots used in the analysis of the effect of screenings of *The Birth of a Nation* between 1915 and 1919 are also plagued by measurement problems, and like lynchings, the measurement error problem is exacerbated by the lack of statistical power. One of the two data sources used by Ang is Paul Gilje (1996). Gilje's definition of a riot is "a riot is any group of twelve or more people attempting to assert their will through the use of force outside the normal bounds of law" (Gilje 1996, 4). Gilje justifies the use of the number twelve because it was used in the English Riot Act of 1715. Gilje is very clear that his count of riots is just one of many possible counts. In his Appendix titled "Counting Riots," Gilje provides his assessment: "How many riots were there in American history? For this study I examined over four thousand riots. That number, however, does not come close to the total of all riots... In some areas the gaps in my counting are glaring. As noted in Chapter 4 there were 5,112 persons lynched between 1882 and 1937. Since I only included lynchings where I read the details...my count only includes a few hundred of these" (Gilje 1996, 183).

These passages from Gilje make clear that the count of race riots used by Ang (2023) is surely error-ridden. As with lynchings this measurement error problem is exacerbated by the rarity of race riots, as shown in Figure 1. Except for 1919, the number of race riots is less than 10 per year. And the causes of the race riots of the Red Summer of 1919 have been well documented by historians and have nothing to do with screenings of *The Birth of a Nation* between 1915 and 1919 (Tuttle 1970; Gilje 1996; Voogd 2008; Whitaker 2009; McWhirter 2011). Returning Black soldiers who fought in WWI inspired an increase in unwillingness among Black people to accept the racist status quo. This combined with the mass migration of Black southerners to northern cities and the relatively scarce jobs and housing in those cities resulted in increased racial tension that exploded across the country in a series of race riots. Here again, Ang (2023) ignores the historical scholarship and uses low-quality data and ill-suited methods to come to a misleading conclusion.

## Conclusion

The rise and fall of the second Ku Klux Klan has been extensively studied by historians. The causes of the Klan revival and its quick demise, all happening in the 1920s, have been convincingly documented (Mecklin 1924; Jackson 1992;

MacLean 1994; Stokes 2007; Rice 2008; Weinberger 2011; Gordon 2017; Hernandez 2019). Similarly, the reception of *The Birth of a Nation* by audiences and its role in film history has been extensively studied (Staiger 1972; Stokes 2007; Stern 2014; Stokes and McEwan 2023). It is widely accepted that *The Birth of a Nation* was a wildly popular and racist film that was more propaganda than history (Franklin 1979). It is clear that *The Birth of a Nation* glorified the original Ku Klux Klan. But it remains unclear how large a causal role it can be ascribed for the rise of the second Klan. The plan to revive the Klan by William Simmons predated the 1915 screening of *The Birth of a Nation* in Atlanta, and screenings during the first run of the film between 1915 and 1919 had little effect on Klan membership. It was not until the 1920s that Klan membership exploded, and the growth of the Klan is largely due to the organizing activities of the Klan, most notably the incentivized Kleagles who were paid per recruit and numbered in the hundreds.

Ang (2023) manufactures a hypothesis that screenings of *The Birth of a Nation* between 1915 and 1919 caused the rise and fall of the second Klan in the 1920s. The motivation for this hypothesis is not a careful reading of the historical scholarship, an analysis of existing scholarship, or a contextually grounded review of what occurred, but the simple fact that screenings of the film were rolled out in a staggered manner across the country and therefore could be viewed as a natural experiment that could be exploited to obtain quasi-experimental estimates of the effect of these screenings on Klan membership, lynchings, race riots, and 21st-century hate sentiment.

In fact, given that there were no Klan chapters prior to 1915, the natural experiment of the film's rollout could not be exploited for the purpose of studying the cause of the Ku Klux Klan revival. Instead, Ang (2023) implements an instrumental-variables procedure that is implausible on its face—the presence of a theater in 1914 in a county was not a random event—and that is severely flawed, as I have showed. The results of this flawed analysis do not contribute to the wider historiography. Because the same approach was used to study the effect of screenings of *The Birth of a Nation* between 1915 and 1919 on 21st-century hate sentiment, those results are equally dubious.

Ang's (2023) empirical analysis of the effect of screenings of *The Birth of a Nation* between 1915 and 1919 on lynchings and race riots lack statistical power and are based on mismeasured data that render these results useless. Even if we ignore these problems and take the evidence as presented, it does not meet commonly used standards of statistical significance that would merit paying attention to it, nor does it even reveal a pattern that can be linked to any cogent explanation of the evidence. In short, it is unreliable and uninformative.

To sum up, the Ang (2023) article was arguably unnecessary and poorly conceived, and the evidence in it is unreliable.

## Data and code

Data and code used in this research are available from the journal's website ([link](#)).

## References

- Ang, Desmond.** 2023. *The Birth of a Nation: Media and Racial Hate*. *American Economic Review* 113(6): 1424–1460.
- Bernstein, Matthew H.** 2023. “At This Time in This City”: Black Atlanta and the Première of *The Birth of a Nation*. In *In the Shadow of The Birth of a Nation*, eds. Melvyn Stokes and Paul McEwan, 175–190. Cham, Switzerland: Palgrave Macmillan. [Link](#)
- Carr, Jay,** ed. 2002. *The A List: The National Society of Film Critics’ 100 Essential Films*. Cambridge, Mass.: Da Capo Press.
- Chalmers, David Mark.** 1965. *Hooded Americanism: The History of the Ku Klux Klan*. Durham, N.C.: Duke University Press.
- Coase, Ronald H.** 1960. The Problem of Social Cost. *Journal of Law and Economics* 3(1): 1–44.
- Coase, Ronald H.** 1974. The Market for Goods and the Market for Ideas. *American Economic Review Papers and Proceedings* 64(2): 384–391. Republished 2022 in *Econ Journal Watch* 19(1): 166–175. [Link](#)
- DellaVigna, Stefano, and Ethan Kaplan.** 2007. The Fox News Effect: Media Bias and Voting. *Quarterly Journal of Economics* 122(3): 1187–1234.
- DellaVigna, Stefano, Ruben Enikolopov, Vera Mironova, Maria Petrova, and Ekaterina Zhuravskaya.** 2014. Cross-Border Media and Nationalism: Evidence from Serbian Radio in Croatia. *American Economic Journal: Applied Economics* 6(3): 103–132.
- Franklin, John Hope.** 1979. “Birth of a Nation”: Propaganda as History. *Massachusetts Review* 20(3): 417–434. [Link](#)
- Fryer, Roland G. Jr., and Steven D. Levitt.** 2012. Hatred and Profits: Under the Hood of the Ku Klux Klan. *Quarterly Journal of Economics* 127(4): 1883–1925.
- Gelman, Andrew, and John Carlin.** 2014. Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors. *Perspectives on Psychological Science* 9(6): 641–651.
- Gilje, Paul A.** 1996. *Rioting in America*. Bloomington, Ind.: Indiana University Press.
- Gordon, Linda.** 2017. *The Second Coming of the KKK: The Ku Klux Klan of the 1920s and the American Political Tradition*. New York: Liveright Publishing Corporation.
- Hernandez, Miguel.** 2019. *The Ku Klux Klan and Freemasonry in 1920s America: Fighting Fraternities*. Abingdon, UK: Routledge.
- Jackson, Kenneth T.** 1992. *The Ku Klux Klan in the City, 1915–1930*. Chicago: Rowman & Littlefield.
- Kehr, Dave.** 2011. A Leap Forward That Tugs Backward. *New York Times*, November 25. [Link](#)
- Kneebone, John T.** 2015. Publicity and Prejudice: The New York World’s Exposé of 1921

- and the History of the Second Ku Klux Klan. VCU Scholars Compass (Virginia Commonwealth University, Richmond, Va.). [Link](#)
- Lay, Shawn**, ed. 1992. *The Invisible Empire in the West: Toward a New Historical Appraisal of the Ku Klux Klan of the 1920s*. Champaign, Ill.: University of Illinois Press.
- Loken, Eric, and Andrew Gelman**. 2017. Measurement Error and the Replication Crisis. *Science* 355(6325): 584–585.
- MacLean, Nancy**. 1994. *Behind the Mask of Chivalry: The Making of the Second Ku Klux Klan*. New York: Oxford University Press.
- Maltby, Richard**. 2023. Blackface, Disguise and Invisibility in the Reception of The Birth of a Nation. In *In the Shadow of The Birth of a Nation*, eds. Melvyn Stokes and Paul McEwan, 35–62. Cham, Switzerland: Palgrave Macmillan.
- McWhirter, Cameron**. 2011. *Red Summer: The Summer of 1919 and the Awakening of Black America*. New York: Henry Holt and Company.
- Mecklin, John Moffatt**. 1963 [1924]. *The Ku Klux Klan: A Study of the American Mind*. New York: Russell and Russell.
- Mencken, H. L.** 1924. Bravos in Bed-Sheets. *American Mercury* 2(May): 120–122.
- Miller, Douglas L.** 2023. An Introductory Guide to Event Study Models. *Journal of Economic Perspectives* 37(2): 203–230.
- Moore, Douglas Cameron**. 1971. *A Study in the Influence of the Film, The Birth of a Nation, on the Attitudes of Selected High School White Students Toward Negroes*. Ph.D. diss., University of Illinois at Urbana-Champaign. [Link](#)
- Pegram, Thomas R.** 2011. *One Hundred Percent American: The Rebirth and Decline of the Ku Klux Klan in the 1920s*. Chicago: Rowman & Littlefield.
- Peterson, Ruth C., and L. L. Thurstone**. 1933. *Motion Pictures and the Social Attitudes of Children: A Payne Fund Study*. New York: Macmillan.
- Rice, Tom**. 2008. “The True Story of the Ku Klux Klan”: Defining the Klan Through Film. *Journal of American Studies* 42(3): 471–488.
- Rogin, Michael**. 1985. “The Sword Became a Flashing Vision”: D. W. Griffith’s The Birth of a Nation. *Representations* 9: 150–195. [Link](#)
- Seguin, Charles, and David Rigby**. 2019. National Crimes: A New National Data Set of Lynchings in the United States, 1883 to 1941. *Socius* 5. [Link](#)
- Simcovitch, Maxim**. 1972. The Impact of Griffith’s *Birth of a Nation* on the Modern Ku Klux Klan. *Journal of Popular Film* 1(1): 45–54.
- Staiger, Janet**. 1992. The Birth of a Nation: Reconsidering Its Reception. In *Interpreting Films: Studies in the Historical Reception of American Cinema*, 139–153. Princeton, N.J.: Princeton University Press.
- Stern, Seymour**. 2014. *D. W. Griffith’s 100th Anniversary: The Birth of a Nation*, ed. Ira H. Gallen. Victoria, Canada: FriesenPress.
- Stokes, Melvyn**. 2007. *D. W. Griffith’s The Birth of a Nation: A History of the Most Controversial Motion Picture of All Time*. Oxford, UK: Oxford University Press.
- Stokes, Melvyn, and Paul McEwan**. 2023. *In the Shadow of The Birth of a Nation: Racism, Reception and Resistance*. Cham, Switzerland: Palgrave Macmillan.
- Tuttle, William M.** 1970. *Race Riot: Chicago in the Red Summer of 1919*. Champaign, Ill.: University of Illinois Press.



- Voogd, Jan.** 2008. *Race Riots & Resistance: The Red Summer of 1919*. New York: Peter Lang.
- Weinberger, Stephen.** 2011. "The Birth of a Nation" and the Making of the NAACP. *Journal of American Studies* 45(1): 77–93. [Link](#)
- Whitaker, Robert.** 2009. *On the Laps of Gods: The Red Summer of 1919 and the Struggle for Justice that Remade a Nation*. New York: Crown.

## About the Author



**Robert Kaestner** is a Research Professor at the Harris School of Public Policy of the University of Chicago. He is also a Research Associate of the National Bureau of Economic Research, an Affiliated Scholar of the Urban Institute and a Senior Fellow of the Schaeffer Center for Health Policy of USC. Prior to joining Harris, Kaestner was on the faculty of the University of Illinois, University of Illinois at Chicago, University of California, Riverside, the CUNY Graduate Center and Baruch College (CUNY). He received his Ph.D. in Economics from the City University of New York. He received his BA and MA from Binghamton University (SUNY). His research interests include health, demography, labor, and social policy evaluation. He has published over 150 articles in academic journals. Recent studies have been awarded Article of the Year by AcademyHealth in 2011 and the 2012 Frank R. Breul Memorial Prize for the best publication in *Social Services Review*. Dr. Kaestner has also been the Principal Investigator on several NIH grants focused on Medicare and Medicaid policy. Kaestner is on the Editorial Board of *Journal of Policy Analysis & Management*. His email address is [kaestner.robert@gmail.com](mailto:kaestner.robert@gmail.com).

[Go to archive of Comments section](#)  
[Go to September 2023 issue](#)





# Temperature Shocks and Economic Growth: Comment on Dell, Jones, and Olken

David Barker<sup>1</sup>

[LINK TO ABSTRACT](#)

In 2012, Melissa Dell, Benjamin Jones, and Benjamin Olken (DJO) published an article in *American Economic Journal: Macroeconomics* arguing that higher temperatures reduce the rate of economic growth. According to the Web of Science, the paper is in the top one percent for citations in academic economics and business publications. Google Scholar lists 1,895 citations. The same team of authors had published a briefer and related article in the *American Economic Review* (DJO 2009), which has 605 Google Scholar citations. The present commentary focuses on DJO (2012).

DJO (2012) was widely covered in the popular press and the authors are extremely well credentialed. Their result is important because the rate of economic growth is the most important determinant of future world wealth. The present paper does not take up the issue of whether temperature increases are coming. It only addresses the question of whether temperature increases reduce the rate of economic growth.

In a previous issue of this journal (Barker 2023) I commented on a Federal Reserve publication by Michael Kiley (2021), which is based on the methods and data of DJO, although it used more sophisticated econometrics. I found that Kiley's methods produced similar results from simulated data where no effect of temperature fluctuations on growth was present, and I showed that dropping countries with unusual events such as large oil discoveries and genocides reversed

---

1. Views expressed in this comment are my own and not those of the Iowa Board of Regents or any other organization I am affiliated with.

his results. In an earlier article I had debunked another study first published by the Federal Reserve that claimed to find effects of warm temperatures on economic growth using U.S. state-level data (Barker 2022; Colacito et al. 2019; 2018). The authors of the two criticized studies have been invited to reply but have not done so thus far; the invitation remains open.

The statistical methods of DJO are less complicated than in Kiley (2021) and their errors are not as subtle. Each of the following five statements is true as a single, standalone criticism or robustness check: (1) They use an untenable method of classifying countries by income; using more reasonable methods I find that their results disappear. (2) Their results are influenced by arbitrary methodological choices. (3) Their results are influenced by a small number of observations with unusual characteristics. (4) The inclusion of additional countries and more recent data weakens their results. (5) Alternative data does not support their hypothesis that high temperatures reduce economic growth.

## Description of DJO (2012)

DJO (2012) regress growth on temperatures and claim to find a causal relationship between them in which higher temperatures lower economic growth. In their abstract DJO say that “higher temperatures substantially reduce economic growth.” DJO uses annual data on average temperatures and rates of economic growth by country. The authors regress annual growth on annual average temperatures, with a sample of observations on 127 countries in years from 1961 to 2003 for a total of 4,924 observations. The panel is unbalanced, with some countries missing data for some years. There is no weighting of observations, so China, with a population of 1.2 billion, has the same influence on the results as St. Vincent and the Grenadines, with a population close to 100,000 and an area one-eighth the size of Rhode Island. Every country is assigned a single average temperature for each year.

DJO include a number of fixed-effect control variables in their model as independent variables. There are dummy variables for each country, and dummy variables for each year/region combination. For example, there is a dummy variable equal to 1 for all observations for which both of the following are true: the year is 1961 and the region is Middle East and North Africa; if one or both of those statements is not true, the dummy takes the value 0. There are also dummy variables for poor countries in each year; those dummy variables allow poor countries to have their own trend of per capita GDP growth that affects growth independent of temperature.

Average annual temperature is another independent variable. The main vari-

able of interest, however, is average temperature multiplied by a dummy variable equal to 1 for countries with per capita GDP below the sample median and 0 otherwise. In this way DJO picks up what they claim is an overall effect of temperature on economic growth, and a separate effect for poor countries. The entire claimed effect of temperature on growth for poor countries is the sum of the coefficients on average temperature and average temperature interacted with the income dummy variable.

The independent variables described above comprise all of the variables included in the regression. Annual per capita GDP growth is the dependent variable. The estimation method is ordinary least squares, with standard errors adjusted for clustering by both country and year.<sup>2</sup>

The primary results of DJO are reported in their second table, titled “Main Panel Results” (DJO 2012, 75). Table 1 shows the results from the second column of their table. Other columns in their table show the results including precipitation and an agricultural country dummy variable, but the results for temperature are very similar in each column. DJO (*ibid.*, 67–68) state that “Changes in precipitation have relatively mild effects on national growth,” so I focus on the results in column (2) of their Table 2. The t-statistics and p-values are not reported in DJO’s Table 2. Those shown in Table 1 below are from my replication of DJO. DJO report only whether a result is significant at the 1-, 5-, or 10-percent level.

**TABLE 1. Results from 2 column (2) in DJO (2012, 75), with replicated t-statistics and p-values**

Dependent variable: annual growth rate	
Temperature	0.261
Standard error	0.312
t-statistic	0.836
p-value	0.403
Temperature × poor	−1.655
Standard error	0.485
t-statistic	−3.410
p-value	0.001
Total effect on poor	−1.394
Standard error	0.408
t-statistic	−3.418
p-value	0.001
R <sup>2</sup>	0.223
Observations	4924

2. The text in Table 2 of DJO (2012, 75) says only that standard errors are clustered by country, but they are actually clustered by both country and year.

For countries that are above median GDP per capita, higher temperatures are associated with higher economic growth, but the effect is statistically insignificant. For poor countries, the total relationship is the sum of the two coefficients “Temperature” and “Temperature  $\times$  poor” in the table above; the sum is  $-1.394$  and is statistically significant. For poor countries, an increase in temperature of one degree Celsius is associated with economic growth that is 1.394 percentage points lower. The  $R^2$  for the regression is 0.22.

If we take DJO’s claim of a causal relationship between temperature and growth seriously, it is interesting to note that the coefficient estimates of the model, although not statistically significant, imply that higher temperatures will increase worldwide wealth. For countries in the upper half of per capita GDP, each degree of temperature increase would raise annual growth by 0.261 percentage points. For countries in the lower half of per capita GDP, a degree of warming would lower growth by 1.394 percentage points. I obtained the predicted temperature increase to the year 2100 for all countries assuming no change in CO2 emissions from the replication files of Marshall Burke et al. (2015). I multiplied this temperature change by 0.00261 for wealthy countries and  $-0.01394$  for poor countries and added the resulting growth adjustment to an assumed baseline growth rate of 0.02. I compounded the resulting growth rate over eighty years to obtain a growth factor up to the year 2100 for each country and multiplied that result by the country’s population. The sum is more than \$500 trillion higher with warming than without. This result is not noted or discussed in DJO.

Again taking seriously DJO’s claim of causality, it is also interesting to note that an implication of DJO’s result is that cooling temperatures would have a positive effect on economic growth in poor countries. If the positive effect of temperature on growth in non-poor countries is ignored because it is statistically insignificant, then spraying enough sulphate particulates into the air to lower temperatures by one degree Fahrenheit would increase annual growth in poor countries by 0.77 percentage points, increasing their annual GDP per capita by nearly 80 percent in the year 2100.<sup>3</sup>

## Replication

Data and replication code are available from Melissa Dell’s website (Dell 2013). I was able to exactly replicate the results in Table 2 of DJO. The Stata program the authors used named “cgmreg” is no longer easily available, but I was able to find and run an archived version. The program named “clus\_nway”

---

3. The cost of such a program has been estimated to be only \$2.5 billion per year (Smith and Wagner 2018).

has replaced `cgmreg`. It produces identical coefficient estimates, but the clustered standard errors are slightly different.<sup>4</sup>

The original data file produced by DJO named “climate\_panel.dta” contains missing values for a variable containing the initial GDP per capita for Myanmar, and so, even though all data are available to include the country in the sample, it is missing from the analysis. Adding Myanmar slightly strengthens DJO’s results.

In the original dataset used in DJO’s Stata program, called “climate\_panel,” the countries Pakistan and Bangladesh are separate observations, but both are coded as “PBD” for the variable called “parent.” As a result, standard errors are clustered by all countries except that these two are clustered together. In addition, the fixed-effect variables for these two countries are combined. Allowing these two countries to be separate clusters and to have different fixed effects slightly weakens DJO’s results.

DJO use different regional classifications of countries than are used by the World Bank. They create a single region combining North American countries, Australia, New Zealand, and European countries. In the World Bank database, North America is a separate region, Australia and New Zealand are part of South-east Asia, and Europe is combined with parts of Central Asia. DJO put Turkey into Europe, while the World Bank puts Turkey into the Middle East and North Africa region. Using World Bank regions slightly weakens DJO’s results.

## Classification of countries by income

A simple but crucial step in DJO’s analysis is the creation of a dummy variable indicating whether a country has per capita GDP above or below the world median. In the footnote to Table 2, they state “Poor is defined as a dummy for a country having below median PPP GDP per capita in its first year in the data.” In their sample, data for South Korea begins in 1961, when per capita GDP was \$1,168 in 2022 dollars. At that time, South Korea was a poor country, recovering from the Korean War. By 1977, however, South Korean per capita GDP was above the median of countries in DJO’s sample. South Korean per capita GDP was below the median for only 16 out of 43 years in the sample, yet it is classified as below the median for all 43 years. Jamaica, which was above the median in 1967, when data are first available, was below the median by the end of the sample period, yet it is classified as a non-poor country for all years.

I reclassified South Korea as poor for the years 1961–1976 and rich for the

---

4. Lines 37 and 38 of the file “maketable2.do” run without error when executed as a .do file, but need to be modified to run in the command line of Stata. The modification is simply to remove the comments in these two lines.

years 1977–2003. This single change nearly eliminates DJO’s result, as shown in the second column of results in Table 2. The t-statistic on the temperature-income interaction variable changes from  $-3.4$  to  $-1.3$ , and the p-value from 0.001 to 0.183. The p-value on the total association for poor countries is 0.086.

Why does South Korea have such a strong influence on DJO’s results? As South Korea emerged from poverty in the 1960s, its rate of growth of per capita income was high. As it became wealthy, its rate of growth declined. This pattern is consistent with convergence, or catch-up growth, predicted by commonly used economic models (Solow 1956). The average temperature in South Korea was higher in the later years of the sample. By classifying South Korea as poor for the years 1977–2003, DJO cause the regression to interpret the simultaneous higher temperatures and lower growth as evidence of high temperatures damaging growth in poor countries. Interestingly, the warmest months in South Korea, May–October, were cooler by 1.8 degrees from 1977–2003 than from 1961–1976. It was the cold months, November–April, that were 4.4 degrees warmer.<sup>5</sup> If high temperatures hindered economic growth, it seems likely that it would be a result of summer temperatures. In other words, classifying South Korea as rich from 1977 to 2003 eliminates DJO’s primary result, there is no justification for classifying it as poor during those years, and temperature likely had nothing to do with lower growth during those years. When all countries are allowed to have different income classifications in different years, DJO’s results disappear. This result is shown in the third column of Table 2 here. Neither the separate association of temperature and growth for poor countries nor the overall association for poor countries is statistically significant.

Allowing country classification to change requires GDP data in a common currency so that GDP per capita can be compared between countries. These data are missing in the World Bank database used by DJO for some years in some countries. DJO calculates GDP growth using inflation-adjusted local currency, but per capita income in dollar terms is not always available for the same years. This is why the number of observations is smaller for the third column of results in Table 2. DJO’s results hold up rerunning their regression using the same smaller sample, so the missing observations are not causing the change in the results. In most cases it is obvious whether a country is above or below the median when the data are missing. When I classified countries in these years in this way I was able to use the entire sample of 4924 observations and the results were similar. There are cases where a country changes from above or below the median, resulting in short periods of classification as poor or rich. In such periods of four years or fewer in a classification, I reclassified these years according to the classification of the

---

5. DJO do not use or discuss monthly temperature data. I explain in a later section how I obtained them.

surrounding years, and the results were also similar.

I also classified countries according to whether they are above or below the median for most years. The results are in the fourth column of numbers in Table 2 labeled “All mode.” I ignored missing observations to make this calculation, so that the full sample could be used. DJO’s results survive with slightly reduced statistical significance. This significance is eliminated, however, with a change to the set of fixed-effect variables used that is discussed in the next section of this paper.

**TABLE 2. Results allowing income classifications to change over time**

Dependent variable: annual growth rate				
	DJO	Korea mixed	All mixed	All mode
Temperature	0.261	−0.232	−0.284	0.128
Standard error	0.312	0.293	0.284	0.317
t-statistic	0.836	−0.792	−1.000	0.405
p-value	0.403	0.428	0.317	0.685
Temperature × poor	−1.655	−0.290	−0.005	−1.327
Standard error	0.485	0.218	0.035	0.519
t-statistic	−3.410	−1.330	−0.142	−2.556
p-value	0.001	0.183	0.887	0.011
Total effect on poor	−1.394	−0.522	−0.289	−1.198
Standard error	0.408	0.304	0.282	0.436
t-statistic	−3.418	−1.718	−1.024	−2.749
p-value	0.001	0.086	0.306	0.006
R <sup>2</sup>	0.223	0.220	0.227	0.221
Observations	4924	4924	4924	4924

Calculating median income using different starting points for different countries is clearly incorrect, but this error does not have a material effect on the results, as is shown in the fourth column of results in Table 2. Classifying countries as rich or poor for the entire 43-year period when many countries changed dramatically relative to other countries is also incorrect, and this error does have a material effect on DJO’s results. As was mentioned earlier, South Korea was poor in the early years of DJO’s sample, and rich in later years. Forcing South Korea to be one or the other for the entire sample is not reasonable. Even if income classifications of countries are constrained to be constant using a better method than DJO’s, their results depend on the inclusion of fixed effects with questionable justification, which will be discussed in a later section.

Instead of constraining the effect of temperature on growth to be discretely different between countries above and below median per capita GDP, it seems reasonable to check whether such an effect might vary continuously as income



changes. In section A22 of DJO's appendix, they show results from using quintile dummy variables instead of the binary dummy variable used in their main specification, and they show results from interacting temperature with initial GDP from their sample. They apparently do not try the obvious specification of directly interacting temperature with current per capita GDP. In this specification, the effect of temperature on growth would be greatest for poor countries, and less for rich countries, with the effect declining in a linear fashion as per capita GDP is higher. The results from this interaction are shown in Table 3. There is no statistically significant effect of temperature on growth.

**TABLE 3. Interaction of per capita GDP with temperature**

Dependent variable: annual growth rate	
Temperature	-0.144
Standard error	0.276
t-statistic	-0.521
p-value	0.602
Temperature $\times$ per capita GDP	0.018
Standard error	0.029
t-statistic	0.615
p-value	0.539
R <sup>2</sup>	0.222
Observations	4654

Another possibility is to allow the effect to vary continuously, but in a manner more like the discrete jump between per capita GDP classes in DJO's main specification using a logit function. Using a nonlinear estimation technique, I estimated the parameters  $\alpha$  and  $k$  shown in equation 1. The variable  $g$  represents growth in per capita GDP,  $T$  represents temperature,  $P$  represents the level of per capita GDP, and  $P_0$  represents median GDP per capita. If  $k$  is equal to zero, then there is no effect from the level of income, and the growth rate varies linearly with temperature. As  $k$  increases, the effect of  $T$  on  $g$  varies with  $P$ , at first linearly, and then with a small effect for countries below median GDP per capita, then a rapid jump in the effect, and a larger effect for countries above median GDP per capita.

$$g = \frac{\alpha T}{1 + e^{k(P - P_0)}} \quad (1)$$

I used OLS to calculate residuals from regressing per capita GDP growth on all of the fixed-effect variables, then I used the nonlinear estimation technique to regress the residuals on temperature and the level of per capita GDP. I also

checked to see if estimating DJO's equation on these residuals produced the same results, and doing so did not change their original results. The results of nonlinear estimation, shown in Table 4, show that neither parameter is statistically significant.

TABLE 4. Nonlinear estimate of equation 1

Parameter	Estimate
$\alpha$ estimate	-0.002
$\alpha$ standard error	0.007
$\alpha$ z statistic	-0.268
$\alpha$ p value	0.788
$k$ estimate	0.000
$k$ standard error	0.002
$k$ z statistic	0.054
$k$ p value	0.957
Observations	4654

## Fixed effects

One set of fixed-effect variables included in DJO's regressions is dummy variables indicating year multiplied by dummy variables indicating whether a country is above or below median income. Including these dummy variables in the analysis holds constant the pattern of average growth of poor countries over time. DJO (2012, 79 n.21) discuss the removal of these variables, saying that "dropping the poor  $\times$  year fixed effects produces similar estimated temperature effects in poor countries." These results are not reported in the paper, but in the online appendix Table A20, results are reported using urban-only data with and without the poor  $\times$  year fixed effects. The t-statistic on the effect of temperature on growth in poor countries drops from 2.41 to 2.01 when the poor  $\times$  year fixed effect is left out. This is not obvious in the table, since both results are only labeled as significant at the 5-percent level. In Table A30 the same specifications are reported for satellite temperature data, and the t-statistic without poor  $\times$  year is only 1.03.

Why do the poor  $\times$  year fixed-effect variables reduce the statistical significance of the result? I simulated a case in which high temperatures increase growth (as in DJO for rich countries), a few outlier poor countries offset this effect, and there is a warming trend in poor countries. Without the poor  $\times$  year fixed-effect variables, the regression shows no statistically significant relationship, but when they are included, the association is negative and statistically significant. In DJO's data, unweighted average temperatures in poor countries tend to increase from 1976 to 2003, and at the same time unweighted average growth rates in poor countries also increase. This correlation pushes against DJO's hypothesis. By inclu-

ding poor  $\times$  year fixed effects, DJO offset this effect, adding to the measured statistical significance. DJO discuss taking these variables out of the regression, and claim that it has no significant effect, but they never provide a justification for including them. Including these particular fixed-effect variables appears to be an arbitrary decision that happens to improve the statistical significance of their result.

TABLE 5. Results without poor  $\times$  year fixed effects

Dependent variable: annual growth rate				
	DJO	No poor $\times$ year FE	No poor $\times$ year FE, mixed	No poor $\times$ year FE, mode
Temperature	0.261	0.010	−0.280	−0.144
Standard error	0.312	0.329	0.287	0.326
t-statistic	0.836	0.031	−0.976	−0.441
p-value	0.403	0.976	0.329	0.660
Temperature $\times$ poor	−1.655	−0.884	−0.035	−0.465
Standard error	0.485	0.491	0.020	0.501
t-statistic	−3.410	−1.800	−1.705	−0.928
p-value	0.001	0.072	0.088	0.353
Total effect on poor	−1.394	−0.874	−0.315	−0.609
Standard error	0.408	0.398	0.285	0.421
t-statistic	−3.418	−2.199	−1.105	−1.445
p-value	0.001	0.028	0.269	0.148
R <sup>2</sup>	0.223	0.209	0.209	0.209
Observations	4924	4924	4924	4924

## Influential observations

In order to further assess the robustness of DJO's results I checked whether particular observations had unusual influence on the estimated coefficients using their income classification method and the same fixed effects variables that they use. The regression diagnostic statistic DFBETA, which stands for difference in beta values, can be used for this purpose. It gives the difference, as a fraction of a standard deviation of a regression coefficient, in the estimated coefficient if an observation is removed. David Belsey et al. (2004) suggest a cutoff of  $\pm 2/\sqrt{n}$  for this statistic, where  $n$  is number of observations. Observations above this cutoff should be examined to see if they may be affected by factors outside of the model being tested. DJO use 4,924 observations, so the cutoff value is  $-0.0285$ . I calculated DFBETA values for each observation using the specification in Table 1.<sup>6</sup>

6. I used simple OLS regression not adjusted for clustered standard errors to compute DFBETA values.

The most influential observation was for Rwanda in 1994 with a DFBETA value of  $-0.462$ , sixteen times the suggested cutoff value. Out of a population of 6.7 million, approximately 500,000 Rwandans were killed over a 15-week period in 1994 (McDoom 2020). GDP dropped by 63 percent. The average temperature in the country was 0.52 degrees Fahrenheit warmer than the previous year, 0.13 degrees cooler than the following year, and 1.66 degrees warmer than the average, and the second highest for years in DJO's sample. This combination of an extreme economic downturn and higher than normal temperatures substantially influenced the model to conclude that temperature affects economic growth. In Rwanda in 1994 the warmest month was September, but the political events which caused the drop in GDP took place in April.

Other influential observations include Burundi in 1995, which was affected by events in Rwanda the previous year, and China in 1961, a year of famine caused primarily by agricultural mismanagement, and Kuwait in 1980, the first year that oil prices had ever fallen. That same year Kuwait reduced oil production by 25 percent<sup>7</sup> and GDP fell by 29 percent. In 1980 Kuwait had its fifth coolest year in DJO's sample of 35 years. Since Kuwait is classified as a rich country, low growth in a cool year there leads the model to increase the estimate of the negative association of temperature and growth in poor countries relative to rich countries, although it has little effect on the estimate of the total association between temperature and growth in poor countries.<sup>8</sup>

Dropping 13 influential observations out of 4,924 eliminates the statistical significance of the total association of temperature on poor countries at the 5-percent level, and dropping 16 observations eliminates it at the 10-percent level. That is, statistical significance disappears when we use that 0.997 portion of the set of observations. Dropping 23 observations eliminates both the total association of temperature and growth and the specific association for poor countries. These results are shown in Table 6. Sensitivity of DJO's results to the effects of a small number of observations does not necessarily invalidate their results. There are observations in the dataset that are influential in both directions, but the most influential observations that enhance DJO's results are more influential than those that diminish the results. The fact that there are such highly influential observations that have much more complicated stories than high temperature reducing growth adds uncertainty to DJO's results that are not captured in the calculated standard

---

7. *New York Times*, March 31, 1980, p. 63.

8. The effect of an observation on the total effect of temperature on poor countries is the sum of the coefficients on temperature and temperature multiplied by the dummy variable representing poor countries. To find the effect of each observation on the total effect of temperature it was necessary to recalculate the regression for each observation, since DFBETA is calculated for a single coefficient, not a linear combination of coefficients.

deviations of the coefficients that they estimate. At the very least, the existence of these highly influential observations indicates that checking DJO's results with additional data and alternative specifications is warranted.

**TABLE 6. DJO results and without influential observations**

Dependent variable: annual growth rate				
	DJO	Without 13 influential	Without 16 influential	Without 23 influential
Temperature	0.261	0.282	0.286	−0.021
Standard error	0.312	0.310	0.310	0.284
t-statistic	0.836	0.910	0.923	−0.075
p-value	0.403	0.363	0.356	0.940
Temperature × poor	−1.655	−0.969	−0.881	−0.593
Standard error	0.485	0.452	0.450	0.437
t-statistic	−3.410	−2.146	−1.958	−1.355
p-value	0.001	0.032	0.050	0.175
Total effect on poor	−1.394	−0.687	−0.595	−0.614
Standard error	0.408	0.369	0.367	0.373
t-statistic	−3.418	−1.864	−1.622	−1.644
p-value	0.001	0.062	0.105	0.100
R <sup>2</sup>	0.223	0.237	0.237	0.235
Observations	4924	4911	4908	4901

## Extended data

The source that DJO used for average annual temperatures for countries has, since the publication of DJO in 2012, released data through the year 2017. DJO only used data through 2003. Their source, *Terrestrial Air Temperature and Precipitation: 1900–2006 Gridded Monthly Time Series, Version 1.01*, provided monthly temperature data, interpolated from weather stations, for 85,794 sections of the globe, each one measuring  $0.5 \times 0.5$  degrees. Another data source, the Global Rural-Urban Mapping Project, provides population data for 741,312,000 sections of the world, each measuring 30 arc-seconds by 30 arc-seconds. DJO used GIS software to match these datasets and calculate population weighted temperatures for each country in each year. I used the raw data and matched it using my own Stata code. I also wrote a Java program to call a function called `coordinate_to_country` using the node.js environment. This allowed me to match each segment to a country. The temperature data have been updated since the publication of DJO, but the correlation coefficient of my calculated temperatures and those of DJO is 0.9978.

By calculating temperatures in this way, I discovered that DJO missed several small countries that do not incorporate the centroid of any 0.5-by-0.5 degree segment. Other small countries were included by DJO. For example, St. Vincent and the Grenadines is included in DJO's data, but Barbados is not, because the area of Barbados just misses the centroid of the surrounding global segments. I used the segment with a centroid nearest the capital city of these countries to provide temperature data.

I also used updated data for per capita economic growth from the World Bank's *World Development Indicators*. Data in this database is regularly changed by countries providing data, sometimes including the deletions of years of observations, sometimes for political purposes (Kadri 2016, 160). Data are also missing for other reasons. For example, GDP data are missing for Canada for the years 1970–1996 in the current version of the World Bank database, but they were available in the 2003 version of the database.

I was also able to expand the dataset by using dollar-denominated GDP from the World Bank database. DJO used the local currency unit-denominated GDP data. Data from some years in some countries are missing using the local currency unit data but are available using the dollar-denominated data. The growth rates are the same between the two series, because the IMF converts GDP to dollars in a base year and then applies local currency growth rates backwards to construct the time series of GDP for each country ([link](#)). For the purposes of categorizing countries by income, DJO used GDP data from the Penn World Table. These data are constructed to purchasing power parity instead of exchange rates.

Using the updated, recalculated data that I compiled, matching observations to those of DJO so that only years and countries used by DJO are included, I find results that are similar to DJO's results, indicating that the data adjustments I made are not driving the differences in results using extended data or alternative specifications.

All but the last column in the analysis reported in Table 7 use DJO's flawed method of using the same classification by income across years for countries. In the second column of results, using the extended data, the statistical significance of the total association of temperature and growth for poor countries is barely significant at the 5-percent level. Using World Bank regions and dropping the poor  $\times$  year fixed effects, the association is not significant at the 10-percent level. Using changing income classifications, the association disappears.

Going back to the original data source also allowed me to use monthly temperature data. DJO use only annual average temperature. If high temperatures cause a reduction in per capita GDP, it seems likely that the warmest temperatures of the year would have the greatest impact. Because some different countries have the highest temperatures in different months, particularly those in different hemi-

**TABLE 7. Comparing results from 2 in DJO (2012) with results using additional data**

Dependent variable: annual growth rate						
	DJO	Extended data	Without poor $\times$ year	WB regions	WB regions w/o poor $\times$ year	Changing income classes
Temperature	0.261	0.410	0.372	0.757	0.568	-0.215
Standard error	0.312	0.325	0.319	0.431	0.403	0.281
t-statistic	0.843	1.261	1.167	1.757	1.407	-0.766
p-value	0.403	0.207	0.243	0.079	0.159	0.444
Temperature $\times$ poor	-1.655	-1.199	-1.016	-1.571	-1.139	0.029
Standard error	0.485	0.497	0.419	0.579	0.493	0.052
t-statistic	-3.410	-2.410	-2.422	-2.714	-2.312	0.560
p-value	0.001	0.016	0.015	0.007	0.021	0.575
Total effect on poor	-1.394	-0.789	-0.644	-0.815	-0.571	-0.186
Standard error	0.408	0.401	0.334	0.410	0.374	0.276
t-statistic	-3.418	-1.967	-1.929	-1.989	-1.528	-0.674
p-value	0.001	0.049	0.054	0.047	0.126	0.500
R <sup>2</sup>	0.223	0.201	0.195	0.194	0.187	0.206
Observations	4924	9033	9033	9033	9033	8934

**TABLE 8. Monthly data**

Dependent variable: annual growth rate						
	Warmest	Coolest	Third-warmest	Third-coolest	Sixth-warmest	Sixth-coolest
Temperature	0.322	-0.127	0.412	-0.213	0.432	-0.195
Standard error	0.254	0.100	0.381	0.152	0.508	0.224
t-statistic	1.266	-1.271	1.083	-1.406	0.850	-0.871
p-value	0.205	0.204	0.279	0.160	0.396	0.384
Temperature $\times$ poor	-0.733	-0.426	-1.124	-0.697	-1.661	-1.170
Standard error	0.344	0.243	0.489	0.324	0.640	0.405
t-statistic	-2.131	-1.752	-2.300	-2.150	-2.595	-2.886
p-value	0.033	0.080	0.021	0.032	0.009	0.004
Total effect on poor	-0.411	-0.553	-0.712	-0.911	-1.229	-1.365
Standard error	0.245	0.264	0.327	0.351	0.423	0.464
t-statistic	-1.677	-2.092	-2.174	-2.594	-2.909	-2.942
p-value	0.094	0.036	0.030	0.009	0.004	0.003
R <sup>2</sup>	0.221	0.221	0.222	0.221	0.223	0.222
Observations	4924	4924	4924	4924	4924	4924

spheres, I ranked the months for each country and put the highest monthly temperatures in one variable, the second highest in another variable, etc. Using a variety of specifications, I was unable to find evidence that warmer months had any more



association with economic growth than cooler months. Table 8 shows the results for the warmest month, the coolest month, the warmest three months, the coolest three months, the warmest six months and the coolest three months.

## Alternative data

To see whether DJO's results can be replicated using different data, I used Robert Tamura's dataset containing estimates of output per worker by country going back to the 19th century (Tamura et al. 2019). Tamura's data is not annual; it is mostly decennial. I calculated the annualized average growth rate between observations and matched these growth rates with the average temperature over the same period. I did two different analyses, one using all of Tamura's data, and another that was restricted to 20th- and 21st-century data. For the 20th- and 21st-century analysis, I used the same temperature data as in the previous section. For the analysis that includes 19th-century data I used temperature data published by the National Oceanic and Atmospheric Administration (NOAA) called the "Global Historical Climatology Network daily" ([link](#)). This dataset contains monthly average temperatures from more than 100,000 weather stations from 180 countries over up to 175 years. I took averages of the temperatures by year and country.

To classify countries as rich or poor I ran a fixed-effect regression using year and country dummy variables and took the median coefficient on countries. Countries with coefficients above the median were classified as rich and countries below the median were classified as poor. I also restricted the sample to countries included in DJO's dataset and used their income classifications and obtained similar negative results. I also checked to see whether a continuous interaction of temperature and income would show an association between temperature and growth that varied by income, and it did not.

Table 9 shows the results. The first column shows the DJO results, the second column shows the results using 19th-, 20th- and 21st-century data, and the third column shows the results using 20th- and 21st-century data.

The alternative data show no statistically significant relationship between temperature and growth. This is true whether all fixed-effect variables from DJO are used, or if a subset of them are used, or if no fixed-effect variables are used at all.

TABLE 9. Alternative data

	DJO	19th–21st centuries	20th–21st centuries
Temperature	0.261	0.085	–0.286
Standard error	0.312	0.070	0.746
t-statistic	0.843	1.211	–0.384
p-value	0.403	0.226	0.701
Temperature x poor	–1.655	0.002	1.171
Standard error	0.485	0.032	3.045
t-statistic	–3.410	0.078	0.385
p-value	0.001	0.938	0.701
Total effect on poor	–1.394	0.087	0.885
Standard error	0.408	0.078	2.665
t-statistic	–3.418	1.121	0.332
p-value	0.001	0.262	0.740
R <sup>2</sup>	0.223	0.332	0.418
Observations	4924	933	1339

## Temperature and political economy

DJO claim that a mechanism through which temperature affects economic growth is that higher temperatures cause political instability. They use the same method of classifying countries by income as described in the previous sections of this paper. The dependent variables are a measure of political instability, with a dummy variable marking “years when the political system is in flux and no clear political regime has emerged,” and another that marks a year in which an “irregular” transition of power occurs, such as a coup. Table 10 shows the original DJO results and the results when different methods of classifying countries by income are used.

For the political instability dependent variable, simply reclassifying South Korea as poor when it was poor and non-poor in other years eliminates the statistical significance of the total association with temperature. The interaction of temperature and the dummy variable indicating whether a country is poor is statistically significant, but the total association, which is the sum of that coefficient and the coefficient of temperature alone is not; the sum just spoken of has a large standard error, and that is why the statistical significance is lost for the total association. The results also disappear when all countries are assigned a single income category based on whether the country is above or below the median in most years of the sample.

**TABLE 10. Political-economy effects: political instability**

Dependent variable: Any change in POLITY score				
	DJO	Korea mixed	All mixed income class	Mode income class
Temperature	−0.013	−0.006	−0.000	−0.016
Standard error	0.009	0.008	0.010	0.010
t-statistic	−1.447	−0.684	−0.042	−1.522
p-value	0.148	0.494	0.966	0.128
Temperature × poor	0.040	0.021	0.005	0.037
Standard error	0.016	0.007	0.003	0.018
t-statistic	2.456	3.142	1.572	2.022
p-value	0.014	0.002	0.116	0.043
Total effect on poor	0.027	0.016	0.005	0.022
Standard error	0.015	0.011	0.010	0.017
t-statistic	1.782	1.443	0.486	1.243
p-value	0.075	0.149	0.627	0.214
R <sup>2</sup>	0.156	0.156	0.157	0.154
Observations	5388	5388	5388	4734

Table 11 shows the results for irregular changes of government. In this case, if countries are given the same income classification over the entire sample period, the association between temperature and growth remains statistically significant. This is shown in the columns labeled “DJO” and “Mode income class.” The association is also statistically significant if South Korea’s classification is allowed to change during the sample period, as shown in the column labeled “Korea mixed.” If all countries are allowed to change, as shown in the column labeled “All mixed income class,” the differential effect of temperature on poor versus non-poor countries disappears, but the total association is still statistically significant. Eliminating only three influential observations, however, is enough to make the total association statistically insignificant at the 10-percent level. The observations are Ecuador in 1976, Paraguay in 1954, and Rwanda in 1994. A coup in Ecuador took place on January 11, 1976. The year 1975 was the second coolest in DJO’s sample period of 1951–2003, but 1976 was above the median temperature in the sample. By coding the coup as taking place in 1976, it appeared that the coup took place following warm temperatures, when in fact it followed cool temperatures. In Paraguay in 1954, monthly temperatures were all within 2.5 degrees of the mean temperature for that month from 1951–2003 except for November, which was 4.8 degrees warmer than average. The coup in Paraguay took place in May. As noted earlier, in Rwanda in 1994 the warmest month was September, but the irregular change of power took place in April.

**TABLE 11. Political economy effects: irregular changes of government**

	Dependent variable: Irregular leader transition				
	DJO	Korea mixed	All mixed income class	Mode income class	Mixed, drop three
Temperature	−0.005	0.009	0.012	−0.006	0.009
Standard error	0.004	0.006	0.006	0.005	0.006
t-statistic	−1.315	1.554	1.990	−1.089	1.364
p-value	0.188	0.120	0.047	0.276	0.173
Temperature × poor	0.050	0.012	0.002	0.064	0.001
Standard error	0.013	0.005	0.002	0.016	0.002
t-statistic	3.898	2.113	0.886	3.920	0.773
p-value	0.000	0.035	0.376	0.000	0.439
Total effect on poor	0.044	0.021	0.014	0.059	0.010
Standard error	0.013	0.007	0.006	0.016	0.007
t-statistic	3.482	2.792	2.166	3.548	1.522
p-value	0.000	0.005	0.030	0.000	0.128
R <sup>2</sup>	0.113	0.111	0.108	0.119	0.108
Observations	6677	6677	6677	5427	6674

## Conclusion

Each of the following five statements is true as a single, standalone criticism or robustness check of DJO (2012): (1) They use an untenable method of classifying countries by income; using more reasonable methods I find that their results disappear. (2) Their results are influenced by arbitrary methodological choices. (3) Their results are influenced by a small number of observations with unusual characteristics. (4) The inclusion of additional countries and more recent data weakens their results. (5) Alternative data does not support their hypothesis that high temperatures reduce economic growth. Thus, on the matter at hand, namely, whether higher temperatures reduce economic growth, DJO (2012) is not helpful and quite possibly has misled people. It is important that we not be misled on the matter at hand: If climate change does not reduce the rate of economic growth, then any likely effect of warming on the level of economic activity will be outweighed by long-term growth.

DJO's finding that higher temperatures would reduce growth is based on several arbitrary choices of method. One is the classification of countries as rich or poor for the entire 1961–2003 time period. Several countries changed from rich to poor or poor to rich during this period. Allowing them to change during the

sample period eliminates DJO's results. Another is the inclusion of fixed-effect variables that control for changes in temperature that are common to many poor countries. If high temperatures reduce growth, they should reduce growth even when they occur in many poor countries at the same time, but these fixed-effect variables control for this effect, causing the remaining influence of temperature to be exaggerated if the effects move in opposite directions, which they apparently do.

Adding additional data by time and country weakens DJO's results, and an alternative data set provides no support to DJO's hypothesis. A mechanism that DJO proposes for temperature to influence growth, political instability caused by heat, also fails robustness tests.

DJO began an important area of research. If higher temperatures significantly reduce the rate of economic growth, and if humans have the ability to lower temperatures, then doing so could greatly increase world wealth. DJO's hypothesis was worth testing, but their tests and those of subsequent authors were flawed, and do not support the hypothesis that higher temperatures reduce economic growth.

## Data and code

Data and code used in this research are available from the journal website ([link](#)).

## References

- Barker, David.** 2022. Temperature and U.S. Economic Growth: Comment on Colacito, Hoffmann, and Phan. *Econ Journal Watch* 19(2): 176–189. [Link](#)
- Barker, David.** 2023. Temperature and Economic Growth: Comment on Kiley. *Econ Journal Watch* 20(1): 69–84. [Link](#)
- Belsey, David A., Edwin Kuh, and Roy E. Welsch.** 2004. Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. Hoboken: John Wiley & Sons.
- Burke, Marshall, Solomon M. Hsiang, and Edward Miguel.** 2015. Global Non-Linear Effect of Temperature on Economic Production. *Nature* 527: 235–239. [Link](#)
- Colacito, Riccardo, Bridget Hoffman, and Toan Phan.** 2018. Temperature and Growth: A Panel Analysis of the United States. Working Paper 18-09. Federal Reserve Bank of Richmond (Richmond, Va.). [Link](#)
- Colacito, Riccardo, Bridget Hoffman, and Toan Phan.** 2019. Temperature and Growth: A Panel Analysis of the United States. *Journal of Money, Credit and Banking* 51(2–3): 313–368.
- Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken (DJO).** 2009. Temperature

- and Income: Reconciling New Cross-Sectional and Panel Estimates. *American Economic Review* 99 (2): 198–204. [Link](#)
- Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken** (DJO). 2012. Temperature Shocks and Economic Growth: Evidence from the Last Half Century. *American Economic Journal: Macroeconomics* 4(3): 66–95. [Link](#)
- Kadri, Ali**. 2016. *The Unmaking of Arab Socialism*. Anthem Press.
- Kiley, Michael**. 2021. Growth at Risk From Climate Change. Finance and Economics Discussion Series 2021-054. Washington: Board of Governors of the Federal Reserve System. [Link](#)
- McDoom, Omar Shahabudin**. 2020. Contested Counting: Toward a Rigorous Estimate of the Death Toll in the Rwandan Genocide. *Journal of Genocide Research* 22(1): 83–93. [Link](#)
- Smith, Wake, and Gernot Wagner**. Stratospheric Aerosol Injection Tactics and Costs in the First 15 years of Deployment. *Environmental Research Letters* 13(12): 1–11. [Link](#)
- Solow, Robert M.** 1956. A Contribution to the Theory of Economic Growth. *Quarterly Journal of Economics* 70 (1): 65–94.
- Tamura, Robert, Jerry Dwyer, John Devereux, and Scott Baier**. 2019. Economic Growth in the Long Run. *Journal of Development Economics* 137: 1–35.

## About the Author



**David Barker** taught economics and finance at the University of Chicago and the University of Iowa. His Ph.D. is from the University of Chicago and he worked as an Economist at the Federal Reserve Bank of New York. He currently runs a real estate and finance company in Iowa and is a member of the Iowa Board of Regents. His email address is [drb@barkerapartments.com](mailto:drb@barkerapartments.com).

[Go to archive of Comments section](#)  
[Go to September 2023 issue](#)



# The Environmental Economy of the Soviet Famine in Ukraine in 1933: A Critique of Several Papers by Natalya Naumenko

Mark B. Tauger<sup>1</sup>

[LINK TO ABSTRACT](#)

This paper analyzes and critiques the main arguments and evidence of three papers by Prof. Natalya Naumenko, an economist at George Mason University. One of the papers was published in the March 2021 issue of the *Journal of Economic History* (hereafter JEH), entitled “The Political Economy of Famine: The Ukrainian Famine of 1933” (Naumenko 2021a). The other two are on websites. One, an earlier version of the JEH article with the same title but from July 2018, is on the website of the Center for Micro-Economic Policy Research (CMEPR) at George Mason (Naumenko 2018). The other online paper is a National Bureau for Economic Research (NBER) working paper (meaning it may not be their final version), in which Naumenko collaborated with two other economists, Andrei Markevich and Nancy Qian, entitled “The Causes of Ukrainian Famine Mortality,” dated July 2021 and revised in 2023 (Markevich et al. 2023).

In the JEH article, which is the main subject of this paper, Naumenko rejected the argument that the famine was a genocide that intentionally targeted Ukrainians and, instead, argued that the evidence showing the Soviet regime’s “discrimination” against Ukrainians and Germans is limited. Her alternative explanation of the famine, which occupies most of her article, is extremely problematic. She attempts to discredit approaches that attribute the famine to environmental

---

1. West Virginia University, Morgantown, WV 26506.



disasters, which she reduces to “weather.” She attributes the famine mostly to Soviet policies: collectivization of agriculture and government policies favoring certain industries.<sup>2</sup> The CMEPR paper mostly takes the same approach as the JEH article, but is slightly more accepting of the argument that Ukrainians were targeted in the famine. It also cites substantial evidence that I published years before, but does not cite my work, as I discuss below. The NBER paper again repeats many of the earlier papers’ problematic arguments, but in contrast to the others argues that the famine specifically targeted Ukrainians because of “anti-Ukrainian bias” on the part of the Soviet leadership.

Naumenko’s arguments and analysis are based on major historical inaccuracies and falsehoods, omissions of essential evidence contained in her sources or easily available, and substantial misunderstandings of certain key topics. All of these characteristics reflect a biased approach to the history and issues that these papers address. All three papers reach their conclusions using statistical calculations based partly on selective use of Soviet evidence and partly on their own “estimates,” but they present only the results of their statistical calculations, and almost none of the actual evidence they used for those calculations, and the tiny amount of evidence that they do present is inaccurate. As a result, all three papers draw invalid conclusions from a grossly distorted history of this important crisis. In the following discussion I document these historical and evidence issues to show the problems with her arguments and to support the central importance of environmental factors in causing the famine conditions of 1931–1933.

I begin with comments about the introductory sections of Naumenko’s papers. After this introduction, Naumenko in her JEH article summarizes her arguments: that “available data” do not support “weather” as a major cause of famine deaths; that government policies, especially collectivization, explain most of the famine deaths; and that very little evidence indicates that Ukrainians were subject to discrimination. She elaborates these points in multiple sections of her article, so going through her article section by section would involve considerable repetition. To avoid that repetition, after discussing her introduction I will compile her points on each of these areas into one concise summary and then explain the problems with them. The CMEPR paper addresses most of these points but with certain problems in use of evidence, which I will discuss below. The NBER article also asserts similar points but then claims that Ukrainians were subject to bias in policies, which I will discuss in the final section.

---

2. This article also refers to tables in its appendices, but these appendices are not included in the article as published; as I found out later with help from the editors of this journal, those appendices are located on the internet (Naumenko 2021b).

## Naumenko's introduction

First of all, it is necessary to note that the titles of all three papers refer to “the Ukrainian famine of 1933,” but the first sentences refer to “the 1933 Soviet famine,” or in the NBER paper’s case, “the Great Soviet Famine (1932–33).” Several studies and document collections from Soviet archives, including some she cited, clearly show that during 1931–1933 many regions of the USSR endured famines (see, e.g., Danilov et al. 1999–2004; Antipova et al. 2009; Davies and Wheatcroft 2004). The Russian historian Nikolai Ivnitskii, who as a child lived through the famine in the Central Blackearth Oblast’ north of Ukraine, published a book documenting the famine in multiple regions of the USSR (Ivnitskii 2009). Yet Ivnitskii’s list was incomplete: Belarus many years ago posted a website on the “Holodomor in Belarus” ([link](#)), which stated that “Belarus, just like Ukraine, Povolzh’e and Kazakhstan, endured in 1932–1934 the tragedy of the Holodomor.” There are also studies of this famine in specific regions, such as the study of famine in Siberia by V. S. Poznanskii (2007). In addition to dealing with rural famine, the Soviet government in these years was providing food in the form of rations for approximately 40 million people in towns and cities, in part because of food shortages and high prices from crop failures (see Davies 1996, esp. 177–189, 530). The article more accurately should have referred to the topic as the Soviet famine in Ukraine.

Naumenko begins both the JEH article and the CMEPR paper with reference to an economic study in 1995 that found the USSR in 1928 was “one of the 30 richest countries in the world,” and asks: how could a severe famine have occurred there, especially in Ukraine that was “famous for its grain production” and the “grain-basket” of the USSR (Naumenko 2021a, 157; 2018, 1). These statements reflect a widespread bias and incomplete information that Naumenko does not address. In fact, the non-Ukrainian provinces of Russia and later the USSR always produced double or more than double the amounts of grain produced by Ukraine, simply because the non-Ukrainian agricultural regions were much larger. As an example of this, Table 1 shows the grain production of the Ukrainian and non-Ukrainian provinces of late Imperial Russia, based on estimates of the Central Statistical Administration, from a recent Russian study.

**TABLE 1. Grain production in Ukrainian provinces and in European Russia excluding the Ukrainian provinces, five-year averages, in million tons (calculated from original figures in puds, 16 kg units)**

Region	1896–1900	1901–1905	1906–1910	1911–1915
Ukrainian gubernias	9.68	14.0	13.95	16.1
Russia outside Ukrainian gubernias	25.1	27.68	27.8	32.2
<i>Source:</i> Ostrovskii 2013, 222.				

The Tsarist statistical agencies tended to inflate the harvest, but that inflation applied to all of the data, from both Ukrainian and non-Ukrainian provinces. While these data are thus approximate, and the five-year averages conceal considerable variations (as discussed below), they still are a valid indication of the relative shares of Ukrainian and non-Ukrainian grain harvests in the Russian Empire (Tauger 2001c, 24–29).

Descriptions of Ukraine as the ‘breadbasket’ of Russia have long been a conventional assumption that reflects Ukrainian nationalist views, but these data show that the non-Ukrainian regions of Russia were the main ‘breadbasket’ for Russia.<sup>3</sup>

Moreover, a detailed study of famines and famine aid in Tsarist Russia showed that from the serf emancipation era of the 1860s until 1909, during which the government provided famine relief almost every year to crop failure regions, the province that had the most frequent recourse to government famine relief was Tavrde gubernia, which included south-central Ukraine and Crimea. Other Ukrainian gubernias, including Ekaterinoslav in central Ukraine and Kherson in southwest Ukraine, also had frequent recourse to this relief (Ermolov 1909, pt. II 4–6). Ukraine gubernias did tend to have higher harvest yields in good years than most non-Ukrainian gubernias, but the Ukrainian regions were also subject regularly to the same environmental and agricultural problems and chronic crop failures and famines that struck the rest of Russia, and as I will show below, this situation had not changed by the early 1930s.

## Causes of the 1931–33 famine: “Weather”

Naumenko’s first main argument in her JEH article is that “weather” does not explain the famine adequately and is accountable only for a small share, 8.1 percent, of famine deaths. She supports this with discussions of three topics: grain harvest statistics; comparisons of 1931–1932 weather with weather in earlier periods; and the actual weather in 1931 and 1932. The NBER paper also dismisses “weather” as a factor in the famine, in an even more superficial manner (Markevich et al. 2023, 20).

---

3. Examples of this viewpoint include the following articles: “Futures Markets,” *Wall Street Journal*, 1 May 1986, p. 1, referring to Ukraine as “the Soviet breadbasket,” and “Bush, Yeltsin back Union Treaty...,” *Christian Science Monitor*, 5 August 1991, p. 1, stating “Ukraine is widely known as Europe and Eurasia’s bread basket.”

## Naumenko's problematic approaches to Soviet grain harvest statistics and earlier research

Naumenko's discussions of Soviet harvest data have the following problems. First, she inaccurately describes the sources of the official harvest data for 1931 and 1932. Naumenko (2021a, 175 n.36) wrote: "Officially, Soviet grain accounting changed in 1933: instead of barn yields, statisticians started using field yields that did not account for 15–20 percent harvesting losses. There is a debate on whether field or barn yields were used in 1931 and 1932 harvest reports; see the excellent discussion of the topic in Tauger (1991)." I appreciate her positive reference to my 1991 *Slavic Review* article, but a better citation would be my much more substantial and comparative Treadgold Paper from 2001 on the development of Soviet harvest statistics (Tauger 2001c; the following discussion is based on pp. 44–49 of this study, which in turn is based on both published and archival sources). Naumenko refers to the policy change in 1933, which established the central state commission for harvest yields in a central government decree in December 1932, with further decrees in early 1933 elaborating its structure. This agency organized pre-harvest forecasts of grain production based on samples from fields, which was called at the time the 'biological yield,' not 'field yields.'

In researching this commission, however, I found that the laws establishing it specified that all the previous agencies collecting pre-harvest data were to be abolished, clearly indicating that the central state commission was not the first agency to collect this data. I sought further evidence in Russian and Ukrainian archives, and found that during the collectivization period the government set up three previous agencies that also made pre-harvest forecasts, one each in 1930, 1931, and 1932, none of which were ever mentioned in previous scholarship. In particular, in May 1932 the regime established "interagency accounting-control commissions" at both district and oblast' levels, which performed the same pre-harvest sampling forecasts as the 1933 commission would conduct in subsequent years. As Soviet officials at the time admitted, these commissions inflated the sampling data because they distrusted lower figures. The commissariat of agriculture also conducted pre-harvest sample surveys of the harvests, and disputed the higher estimates of the interagency commissions. The full story of this dispute is complicated, but it was ultimately settled, as one might expect, by Stalin, who decided in a letter in September 1933 that the 1932 harvest should be defined as 698 million centners (69.8 million tons), about halfway between the two agencies' estimates, and which has been the 'official' figure for the 1932 harvest since then. This official figure was thus a dictated compromise between two pre-harvest forecasts, but in no sense was it actual harvest data.

This research shows, however, that Naumenko's assertion that 'field yields'

were officially used only from 1933 is incorrect, just as I showed that similar claims by other scholars about the ‘biological’ yield use only from 1933 were also incorrect. These pre-harvest exaggerated forecasts were used from at least 1930, and especially in 1932. In fact, most if not all earlier Soviet and pre-revolutionary Russian harvest data were based on pre-harvest forecasts (Tauger 2001c, 24–39).

Naumenko then acknowledges that there is more reliable harvest data, but in the JEH article she presents the harvest data very minimally, only in a few figures (Naumenko 2021a, 171). She presented the more reliable data, from kolkhoz annual reports, concretely only in the online 2018 version of the article. Yet I consider it important to point out that the two main points she makes about this data in both articles, and especially in the 2018 version, I made previously and in more explanatory detail in my first article in 1991 in *Slavic Review* and in certain later work, all of which were published before her 2018 paper and 2021 article (Tauger 1991b).

The first of the two main points Naumenko makes about the data is to note that the official harvest data cannot be correct because the “rural grain retention is too high in 1931 and 1932” (2018, 19; 2021a, 158). This is partly correct, but I documented this point with statistics in Table 2 of my 1991 article (Tauger 1991b, 74). However, I also showed that this statistical ‘retention’ was the result not only of inflated official harvest data, but also of substantial amounts of grain that the government returned to villages for food and seed, after collecting grain procurements. Naumenko did not cite my article and its data on either of these points. In my article I also cited numerous western and Soviet studies that asserted the official figures were too high (*ibid.*, 74–77). And as discussed above, I showed in 2001 why the official harvest data for 1931 and 1932 were too high, because they were based on inflated pre-harvest projections by special agencies set up to gather such data. Naumenko did not cite my work or evidence for any of these points in any of her papers.

The second main point Naumenko makes about the data is in her 2018 paper, where she specifically discussed an important source, the kolkhoz annual reports from two archival documents, and noted very briefly that they showed a much smaller harvest in 1932 in Ukraine. She noted that they came from 47.3 percent of kolkhozy in Ukraine. She also discussed their reliability and how representative they were, and concluded that “these data deserve serious consideration” (2018, 18). My 1991 article, which Naumenko does not cite in the 2018 paper, presented all of these data and more from those same archival sources, and made most of the same arguments about the data that she made, as well as other points (Tauger 1991b, 78–85). My article was the first western study to cite the annual report data, from the same archival sources that she cited. I presented tables on their coverage, which included the 47.3 percent figure that she cited, and their data on harvest yields and procurements for several regions including the six main oblasti (Soviet-

defined provinces) of Ukraine as well as totals for Ukraine as a whole, which was more data than what she cited from those sources. I also discussed their reliability and representativeness, citing the work of V. I. Zvavich, a Soviet scholar who wrote a dissertation and articles on these data, and who also endorsed their reliability and importance.

While Naumenko's points about these data are mostly correct, I made those points earlier, and I find it extremely problematic that Naumenko in her 2018 paper used the same data that I used and made most of the same arguments that I made about the data without citing my work on these points. In her JEH article she cites my article minimally and elliptically and does not identify the annual report data as a source, but her citation of my article proves that she knew about my work before 2021. I do not know whether she knew my article when she wrote her original paper in 2018, but that paper's presentation of the same data I used and the same points I made about that data clearly suggests that she did know my article and failed to cite it.

Naumenko also makes another error in her discussion of statistics in the 2018 article. In it she acknowledges that grain procurements from Ukraine fell from 7.1 million tons in 1931 to 4.2 million tons in 1932 (p. 19 and Table 2), but this missed two key points. First, that 4.2 million tons included 600,000 tons from sovkhozy in Ukraine, which were a smaller sector and not a major site of famine. The procurement plan for the main groups who suffered from famine, the kolkhoz peasants and non-collectivized peasants, were reduced to 3.77 million tons, or just over half of the 1931 target.<sup>4</sup> Second, all of these figures, the 4.2 million tons and 3.77 million tons, were the final plan after the government reduced procurements four times for Ukraine in 1932, but not the actual procurements gathered. The procurements the Soviet government managed to receive were less than that total. In mid-December the Ukrainian official press reported that kolkhozy had fulfilled the procurement plan 72.4 percent, sovkhozy 68.8 percent, and non-collectivized peasants only 39.5 percent (Slyn'ko 1961, 298, citing the newspaper *Za sotsialistichbnu perebudovu* from 18 December 1932). At that time, the regime had reduced the procurement plan for Ukraine three times; on 11 January the regime reduced the plan one more time, which enabled some regions to fulfill the plan (I discuss these reductions further below). But the total reduced plan was never fulfilled, and the regime on 29 January specified that most grain procurements still being collected at that point were to be allotted to the farms for seed (Pyrih 2007, 601–602, 625).

---

4. Calculated from the decrees reducing procurements recorded in documents in Pyrih 2007, 290–303, 355–360, 597–601. I discussed this in my review of Applebaum's book *Red Famine* on History News Network (Tauger 2018), but here I have corrected the data slightly.

Consequently, using the procurement plan to determine how much grain peasants and kolkhozy had after the harvest is misleading: the fact that they were starving even though they had not fulfilled the procurement plans surely indicated that the harvest was even lower than the annual report data indicated. Naumenko made a similar point in her CMEPR paper (2018, 20), but she claims that “weather data” do not support this harvest decline. This argument is problematic, as I will discuss below.

Finally, Naumenko made two more errors regarding my work in her discussions of statistics in her 2021 article. First, she wrote: “Since Davies and Wheatcroft (2004) and Tauger (2001) do not offer corrections for separate Soviet republics...” (Naumenko 2021a, 170–172). This is not correct: in my 1991 article I calculated weighted-average grain harvest yields for Ukraine and most of the other important grain regions of the USSR for 1932 (Tauger 1991b, 85 Table 10).<sup>5</sup> Then, discussing the larger 1933 harvest, Naumenko wrote “... historians (including Davies and Wheatcroft, although, remarkably, not Tauger) agree that the 1933 harvest was good” (Naumenko 2021a, 174). This is also incorrect. In my 1991 article three tables of data show that the 1933 harvest was much larger than the 1932 harvest: Table 6 has data from the kolkhoz annual reports for 1932 and 1933, showing much larger harvests in 1933 in Ukraine; Table 7 documents increased sown areas and increased yields in 1933 over 1932, again from annual reports; and Table 8 shows substantially larger harvests in 1933 than in 1932, from a 1935 Soviet publication that included annual report data (Tauger 1991, 80–82). I also document how the larger 1933 harvest substantially reduced grain prices on markets (Tauger 1991b, 87).<sup>6</sup> I do not know whether Naumenko’s false statements in these two cases were intentional, results of her failure to read those works carefully, or results of forgetfulness, but no matter what the reason they reflect at least incompetent work, and possibly an attempt to discredit other scholars whose work would challenge her conclusions.

## Comparisons of 1931–1932 weather with weather in earlier periods

Naumenko attempts to discredit the idea that the 1931 and 1932 harvests were low from environmental factors by comparing them with earlier periods when according to her weather statistics, weather conditions were similar.

Naumenko (2021a, 157, 175; 2018, 3) argues that the 1931 harvest was not

---

5. Davies and Wheatcroft also performed this calculation in a publication that came out 13 years after my initial presentation of it, also without citing my work; see Tauger 2006a.

6. See Tauger 2001b and 2001c for further documentation of these points.



terribly low because it was on the same level as 1924 and 1934 and those were not famine years, and in line with the average of 1924–1929, during which she states there were no famines. The truth is quite different: the USSR saw two substantial famines in this period, in 1924–1925 and in 1928–1929. The 1924 crop failure and the famine conditions that followed it affected many regions of the USSR, including Ukraine, and were not secret but were openly recognized in the Soviet press. Even the *New York Times* reported on them.<sup>7</sup> The Soviet government established a famine relief committee, led by top Soviet official Alexei Rykov, imported food and distributed it to famine victims, regularly reported on conditions both in certain publications and in internal sources now in Soviet archives, and in 1925 published a substantial collection of articles on the famine, its causes, relief efforts, and proposals to prevent future famines, *V bor'be s zasukhoi i golodom (In the Struggle with Drought and Famine)*. In 2007 I presented at a conference a substantial paper on this famine, based on both published sources and extensive archival research, and my study was published in France in a collection of my articles in 2017 (Rykov 1925; Tauger 2017b).

In 1928, a serious crop failure struck Ukraine, which was again not secret. The Ukrainian government again set up a relief committee, imported food, and provided relief to famine victims. I published an article on this famine in 2001, which was the first historical study of this famine in any language, as a later Ukrainian scholar acknowledged (Tauger 2001a; Hrynevych 2013, 7–8). Soviet scientists, including Nikolai Vavilov, published two studies documenting these crop failures. Other regions also had low harvests in 1927 and 1928 (Kuleshov 1929; Vorob'ev 1929). A Ukrainian leader, Stanislav Kosior, in a report to the Communist Party Central Committee from 5 February 1933, referred to the “catastrophe” of massive winterkill of crops in 1927–1928, indicating that Ukrainian officials remembered this crisis (Pyrih 2007, 642–643). The occurrence of these crop failures from diverse environmental disasters in multiple regions, and the government’s establishment of famine relief committees and distribution of food relief to crop failure regions from regions with better harvests and from imports, repeated decades of similar events and practices from the Tsarist regime.

Naumenko (2021a, 161–162; 2018, 7) briefly discusses this period with no reference to the crop failures, and instead attributes the urban shortages of the late 1920s to peasants’ decisions to shift away from growing grain to more expensive produce. This is an old conventional argument that was based on ignorance of actual agricultural conditions, especially the severe crop failures, and on incomplete evidence about peasants’ actions (Tauger 2001a, 164, 360; for an example of these

---

7. For example: “Reports famine sweeping Russia,” *New York Times* 7 September 1924, p. 3; “Famine again in Russia,” *New York Times* 9 September 1924, p. 28. For Soviet sources, see discussion below.

older views, see Lewin 1974, ch. 9). As the above cited sources show, the “grain crisis” of 1928–1929 was really at root a crisis of crop failures and famine (Tauger 2001a; 2006b). Naumenko concludes from her description of the 1924–1929 period as “non-famine years” that environmental factors were non-existent and unimportant for the 1931–1933 crisis. The reality is, however, that environmental factors were extremely important and caused serious crop failures and famines in 1924–1925 and 1927–1929, famines that the Soviet government openly recognized and organized relief, including importing food, to alleviate. Consequently, a historically accurate comparison of these years with 1931 would clearly have to indicate the central importance of environmental factors.

Naumenko makes a similar inaccurate historical comparison in a statistical calculation using ‘weather data’ from Russia in 1901–1915 to predict what the 1931–1932 harvests would have been without collectivization (2021a, 174–175; 2018, 16–17). She implies that the 1901–1915 data came from a period when Russia had good harvests and no famines, again in an attempt to prove that the 1932–33 famine did not result from ‘weather shock.’ She also makes a similar claim in the NBER paper using the “famine of 1892” (actually 1891–1892) and describing it as the “last large famine in the Russian Empire” (Markevich et al. 2023, 20).

Yet her implicit assumption in using this ‘data’ is again completely incorrect. I am in the process of writing a history of famines in Russia, and I have substantial evidence documenting crop failures in every one of those years. Several of those years had extremely severe crop failures that caused serious famines: in 1901–1902, 1905–1907, 1911–1912, and 1914–1915. These cases are documented in the western press, such as the *New York Times*, but especially in many Russian publications. The most important ones were a series of annual reports prepared by the Russian Ministry of Interior for the years 1906–1914, demanded by the new elected Russian Duma (parliament), established after the 1905 revolution. These reports were extremely detailed: the one for 1906–1907 had seven volumes, approximately 10,000 pages, mostly statistical data on the aid the government provided; the one for 1911–1912 was two volumes, approximately 1000 pages long (MVD 1907; 1908; 1913). All of them provided detailed information on weather and agricultural conditions, which were extremely complex and diverse and rapidly changing over the year each report covered, as well as estimated harvests of food and forage crops and estimated needs for relief in crop failure regions. To give an impression of how serious many of these crises were, the following table, prepared by a Russian scholar in a study of the 1911–1912 case, shows the amounts the government spent on famine relief in these years:

**TABLE 2. Russian government spending from state treasury for the most important food relief campaigns, in rubles.**

1881	3,229,324	1903	5,000,000
1891–92	146,500,000	1904	1,500,000
1897–98	35,000,000	1905–6	71,153,000
1900	699,698	1906–7	169,657,000
1901	20,000,000	1907–9	23,954,000
1902	18,620,000	1911–12	126,555,517
<i>Source:</i> Belokurov 2014, 17.			

The 1891–1892 famine is well-known in the literature as a particularly severe case, but this table, based in part on the substantial Ministry of Interior publications mentioned above, shows that the government spent much more on relief in 1906–1907, almost as much in 1911–1912, and very substantial sums in other years. These data as well as many more sources clearly show that the 1891–92 famine was not the last large famine in the Russian Empire.

Yet Naumenko never mentions that the years she used as a reference, 1901–1915, had crop failures every year and several serious famines; she implies that these years were perfectly normal, good harvest years, which was clearly not the case. She used some generalized weather data, but never mentions or cites these much more detailed and accurate sources, which described the weather and agricultural conditions as they changed week by week and region by region, and the concrete effects of them on farm production. Her claim that her calculations, using what in fact was very selective and unrepresentative data from 1901–1915, demonstrate that the harvests and famine in 1931–1933 were not the result of environmental conditions, is clearly incorrect and appears to reflect ignorance about both the real conditions in 1901–1915 and their comparability with conditions in 1931–1933. The Tsarist regime’s massive expenditures on relief efforts recorded in Table 2 were undertaken in response to low harvests and crop failures caused by weather extremes and other environmental factors. They clearly indicate that Naumenko’s assumption that the period 1901–1915 saw good regular harvests is wrong. Her assertion, derived from this assumption, that the same “favorable” weather conditions would have prevailed in 1931–1933 without collectivization is also wrong, because as the above table indicates, Russia (including Ukraine) had frequent and unpredictable environmental disasters and crop failures repeatedly in the past, and also because, as discussed below, the actual environmental conditions in 1931–1932 were also not favorable.

## The actual weather and other environmental conditions in 1931 and 1932

In discussing the ‘weather’ conditions in 1931 and 1932, Naumenko makes several assertions that ignored evidence in her own sources, in particular in my publications. First, there was a serious drought in the USSR in 1931, but she claimed in both papers that it did not affect Ukraine. In the 2018 paper she wrote that “raw weather data” did not confirm the drought; in the 2021 paper she wrote: “I demonstrate that there was no drought in Ukraine in 1931” (Naumenko 2018, 3; Naumenko 2021a, 158, 176).

She bases these arguments on published weather data, but these data are extremely problematic. The appendices to her 2021 article include a map show the locations of the weather stations that were the sources of her data (Naumenko 2021b, 14). This map indicated three main weather stations, and three smaller ones, all in three of the seven regions of Ukraine. Their coverage, in other words, was extremely sparse and incomplete.

Yet in my Carl Beck paper that Naumenko cited, I included a table from Ukrainian archives showing precipitation there in April–June 1932, crucial months for crop growth, and I reproduce it in Table 3 (Tauger 2001b, 12). These data came from 10 regions in Ukraine, in other words much wider and more comprehensive coverage than Naumenko’s data. These data came from the Ukrainian commissariat of agriculture, which had qualified observers and scientists all over Ukraine reporting on these conditions. These data are based on eyewitness, measured day-by-day reports, covering almost all of Ukraine, and hence are very reliable.

**TABLE 3. Summary of precipitation in Ukraine, April to the first half of June, 1931 and 1932**

Oblast'	Millimeters			Percent		
	Long-term average	1931	1932	Long-term average	1931	1932
Kiev	165	191	328	100	116	199
Vinnytsia	130	76	171	100	51	132
Sumy	150	113	178	100	75	119
Kharkov	118	99	233	100	84	198
Poltava	110	58	210	100	53	192
Zinov'ev	105	115	315	100	110	300
Odessa	80	77	191	100	96	240
Askenia	85	57	143	100	67	168
Novoluk	110	66	136	100	69	124
Iasinuvsk	140	56	210	100	40	150
<i>Source:</i> Tsentral'nyy derzhavnyy arkhiv vyshchikh orhaniv vlasty ta upravliniia Ukrainy (Central State Archive of Leading State Organs of Ukraine), f.27 o.13 d.213, ll. 37, 39; previously published in Tauger 2001b, 12.						

Table 3 shows that three regions—Kiev, Zinov’ev, and Odessa—had near-normal or above normal rainfall in 1931, but the remaining seven regions had lower rainfall than average, including three—Vinnitsya, Poltava, and Iasinuvsk—that had 40–53 percent of normal rainfall. Unless it was perfectly timed, such low rainfall would be likely to create drought conditions. These data thus document both drought or near-drought conditions in most of Ukraine in 1931 as well as extraordinarily heavy rainfall in 1932.

Naumenko ignored this table, even though she cited the publication on other points. These drought conditions in Ukraine clearly contributed to a low harvest in Ukraine in 1931. The official estimate of the harvest (of course prepared by one of those agencies that conducted pre-harvest projections) was 18.3 million tons, from a sown area of 21.1 million hectares, with a yield of 8.6 centners per hectare (Sels’skoe 1936, 248, 270). A Soviet Ukrainian scholar, however, cited archival sources showing that the harvest was actually about 30 percent lower, 13.8 million tons, from sowings of 19.5 million hectares, for a yield of 7 centners (Slyn’ko 1961, 285–287). Further evidence of the effects of this drought in Ukraine include Soviet decrees allocating 870,000 tons of food and seed aid to central and eastern regions of the USSR, the ones most severely struck by the drought, on 16 February 1932, but also 106,000 tons of grain as food aid to Ukraine on 15 May 1932, and additional aid over the following weeks (Pyrih 2007, 63–64, 156; Pyrih 1990, 162).

On the 1932 harvest, Naumenko presents inconsistent arguments in her JEH article and CMEPR paper, both of which ignore crucial data. In her 2018 paper, as discussed above, she cited, discussed, and basically accepted the kolkhoz annual report data showing a harvest in Ukraine much smaller than the official figure (Markevich et al. 2023, p.36, uses the same figure for “production” in 1932, in her footnote stating the data were “revised by the authors using archival sources,” which were clearly the sources I published in 1991). In this discussion, however, she relied exclusively on weather data and argued that the smaller harvest was not due to “weather”, again relying just on official rainfall and temperature sources, and her inaccurate comparison with 1901–1915 (Naumenko 2018, 15–20).

In her 2021 article, she very briefly acknowledged the low harvest estimates by me and by Davies and Wheatcroft (as noted above, overlooking our ‘corrected estimates’ for Ukraine), with a very brief reference to my much longer discussion of the other environmental factors of infestations of insects, plant diseases, rodents, and weeds. Then she discussed weather and its uncertainties, and applied the 1901–1915 and 1924–1929 comparisons, as discussed above, arguing that the weather predicted similar harvests to those earlier periods, again as noted without recognizing the crop failures and famines in those years. Then she concluded that “if there was a gap between the officially reported harvest and the true harvest (and there must have been, otherwise rural retention is too high), in Ukraine, this gap *is*

*not predicted by the weather*” (Naumenko 2021a, 175, her emphasis).

Then she wrote “Unfortunately, it is impossible to directly quantify the presumed damage from pests and grain diseases” (Naumenko 2021a, 175). She asserts that the weather data do not allow such quantification. Then she refers to “Appendix B.5,” which “studies how often published archival documents discussed famine, weather, grain diseases and pests” and “shows that weather or pests were not discussed more than usual in the years leading up to the 1933 famine” (ibid., 176). This is an invalid argument, as I will show below.

Her arguments regarding the 1932 harvest in the 2018 paper are based exclusively on weather, and completely ignore my paper that presented substantial scientific evidence on the other environmental factors. Her arguments in that paper are extremely problematic, because first of all, as the reports on the 1901–1915 period showed, the weather processes in this region were never simple, but involved many complex changes that cannot be reduced simply to selected temperature and precipitation readings. Second, any study of agriculture will show that in most cases one cannot reduce agricultural production exclusively to weather: there are also factors internal to the plants and in the surrounding environment that influence their growth and development. This was especially the case in 1932 all over the USSR. Her 2021 article unfortunately does not discuss the annual report data at all, which showed how serious the harvest decline was in 1932; instead, she basically argued that it was impossible to know how low the harvest was.

Naumenko’s assertion in the JEH article that “it is impossible to directly quantify the presumed damage from pests and grain diseases” (Naumenko 2021a, 175) ignores a table I cited from a Soviet scientific article, reproduced in Table 4 below, containing quantitative estimates of losses for six different crops from rust and smut that totaled 8.94 million tons, by a leading Soviet agronomist (Tauger 2001b, 17, citing Artemov 1933, 75). That article was a study of susceptibility of grain crops to fungal diseases, written by one of the leading specialists in that area, P. K. Artemov. I was informed about the article by a leading U.S. specialist, Professor Alan Roelfs of the University of Minnesota. The journal, *Trudy po prikladnoi botanike, genetike, i selektsii* (*Studies on Applied Botany, Genetics, and Selection*) was founded in 1908 and has continued publishing to the present day. At the time of this publication, 1933, the head of agricultural research in the USSR, and the head of VASKhNIL, the Soviet academy of agricultural sciences, was the internationally respected scientist Nikolai Vavilov; the fraudulent scientist Trofim Lysenko at this time worked at a research center in Ukraine, was still a follower of Vavilov and had not yet turned against genetics, and had no political power over agricultural research. Consequently, Artemov’s article is not political propaganda, not ‘lysenkoist,’ but a legitimate and reputable scientific article.

**TABLE 4. Harvest losses from smut and rust, 1932 (centners)**

Crop	Losses from Smut	Losses from Rust
Rye	457,000	---
Winter Wheat	1,313,000	40,000,000
Spring Wheat	5,771,000	11,000,000
Oats	3,056,000	20,000,000
Barley	1,422,000	---
Millet	1,406,000	---
Total	18,455,000	71,000,000

*Sources:* Artemov 1933, 75; Tauger 2001b, 17.

These data show that Soviet scientists recognized enormous losses just from these plant diseases and attempted to quantify them. In addition to ignoring this source, Naumenko wrote that published archival sources did not discuss these problems much if at all, but she did not seem to be aware that there are many unpublished archival sources by scientists and specialists who documented that these environmental factors had devastating effects on crops in 1932. I have a large collection of scans of these reports for 1931–1933, which I am translating and preparing for publication. There are many more of these reports for 1932 than for those other years, which indicates that specialists recognized how serious these problems were in 1932. Scientists from regional branches of the Soviet Institute of Plant Protection (VIZR) reported regularly on these conditions, with particular attention to Ukraine because of its importance for grain supply after the 1931 crop failures and because of the significance of these problems in that republic. Their report on Ukraine report from 10 July 1932 noted:

The general condition of grain crops in Ukraine, which at the present time are in the phase before the end of vegetative growth, should be characterized as satisfactory. Thanks to meteorological conditions of this year with presence of sufficient warmth and moisture in the important periods of development of plants unconditionally was reflected in the growth of grain crops. However, simultaneously with the rapid growth of cultivated plants, proceeded also rapid growth on them of parasitic microflora, which especially was observed in relation to rust, which by the end of June reached 100% development in a series of parts of Ukraine. The largest appearance of the parasite was observed in June, when the dynamics of the disease had the greatest intensity. The first 10 days of July noted further intensification of rust, which by this time implanted in the regions of the forest-steppe and Poles'ia. (TsGANTD f.356. o. 1-1 d.483, l.30)<sup>8</sup>

8. TsGANTD is the Central State Archive of Scientific and Technical Documentation, St. Petersburg.



Their report from 31 July 1932 noted that this year had especially severe infestations of rust, smut, and ergot (smut is documented in a quote further down):

Infestation of ergot this year evidently has exceptional character and should be attributed to favorable meteorological conditions (warmth and moisture), and also to poor agricultural measures. ... Rust has the character of almost uniform infestation in the whole territory of the Republic ... 1932 should be acknowledged unfavorable because of mass epiphytotic of rust, which encompassed almost the whole territory of Ukraine and stood out for 100% infestation of sowings. Rapid development of rust coincided with milky whiteness maturity of the crops, led to significant drying of leaf surfaces and interrupted the correct accumulation of the plant at the moment of its ripening. The observed rapid development of spring grain crops in connection with this was held back and the quality of the harvest was noticeably reduced. (TsGANTD f.356. o. 1-1 d.483, ll.43–44)

These reports all contained detailed tables showing the degrees of infestation of these diseases at the observation points, as well as detailed reports on several major insect infestations that damaged crops in Ukraine and elsewhere in the USSR. These agronomic reports show that Naumenko's emphasis on "weather," and her arguments that weather conditions were generally favorable, clearly failed to consider that "favorable" weather for crops can also be favorable for infestations that seriously reduce harvests.

The VIZR reports at the end of the year included detailed statistical reports on plant disease infestations in Ukraine, which concluded that "1932 should be noted for mass development of smut fungi on all grain crops, which led to significant lowering of the harvest," that "mass spread of ergot in Ukraine this year carried an exceptional character, in particular in some districts infestations of grain reached levels exceeding epidemic disease (above 0.5%)," and that "the year 1932 should be recognized as a year of mass development of rusts. Rust encompassed the territory of all of Ukraine and was distributed both on winter and spring grain crops." A recent study by a Ukrainian biologist, Nazar Nazarenko, also documents that Soviet agronomists were very concerned about the plant diseases and documented the unusually high degrees of infestations in scientific publications (TsGANTD f.356. o. 1-1, d.482, ll. 17, 19, 23; Nazarenko 2019, 189). There is much more evidence of these infestations in these documents, which I am preparing for publication.

There were also other environmental disasters that Naumenko did not men-

---

Russian Federation. In these citations, f = fond, the collection of documents for a specific agency; o = opis, inventory of documents in the fond; d = delo, the specific file; l = list, or page.

tion, most notably a vast weed infestation in 1932. The above-cited articles cited both published scientific and archival sources stating that while weeds had long been a serious problem in prerevolutionary Russian and Soviet farms, those of 1932 were much worse than previously observed. Even the Soviet press reported in July 1932 that weeds in Ukraine were “smothering” crops, providing breeding grounds for certain insects and plant diseases, and causing decreased harvests. Nazarenko cited several scientific studies showing that unusually high weediness on millions of hectares in most regions of the USSR greatly reduced yields and also made much harvested grain inedible because of lower quality and contamination with weed seeds. His article cited detailed studies showing high degrees of weed contamination of harvests of all the main grain crops in 1932 (Nazarenko 2019, 187–188; Tauger 2001b, 37–39).

The weed infestations resulted from favorable environmental conditions and from other historical, environmental, and economic factors. Until collectivization, most peasants, including in Ukrainian provinces, farmed their fields in strips, with borders between them that fostered weed growth (discussed below). Peasants usually did not sort their seed before sowing to eliminate weed seeds, peasants’ farming implements allowed only shallow plowing and did not uproot weeds, and peasants often never bothered to weed their grain crops. A Russian specialist noted in 1917 that these factors “partly explain the terrible infestation of our fields with weed plants” (Mozzhukhin 1917, 20–21). A Soviet scientist, A. I. Mal’tsev, documented the catastrophic effects of weeds in Soviet agriculture in a 1926 publication, from these and other causes (Mal’tsev 1926, 3–71). Russian grain exports were notorious in Europe for high proportions of impurities, including seeds from non-grain plants, which included weeds. Mal’tsev reported that weed contamination of Soviet grain frequently reached 20–30 percent, sometimes 50 percent, and that this contamination lowered the value and acceptability of Soviet grain exports (Rubinow 1908, 15ff.; Mal’tsev 1926, 8–9). This pattern also meant that after 1917, the USSR inherited Russia fields heavily infested with weed seeds that under favorable environmental conditions would grow rapidly. Mal’tsev refers to research documenting that Soviet soils were contaminated with hundreds of millions and even billions of weed seeds. Many scientific sources on weeds, both Western and Russian/Soviet, document this point (Mal’tsev 1926, 14; one example of a Western source is WSSA 2016).

In 1932, both Soviet and Western specialists observed and reported on weed infestations. They all argued that the infestations were exceptionally large and unprecedented. Soviet personnel recognized that many factors contributed to these infestations, including lack of crop rotations, unusually warm and wet weather (see Table 3 above), but also equipment and draft power shortages (Tauger 2001b, 38). At a national conference Soviet scientists held on the weed problem in December

1933, a specialist from the North Caucasus, which bordered the Ukrainian republic, stated that a major influence on weed growth was the moist weather in 1932 and 1933. A Ukrainian specialist noted that weediness was extreme in 1932 and even with better weeding, it was still significant in 1933 (Volkov et al. 1935, 26–27, 46). Another environmental aspect that they did not mention was that the enormous weediness was probably partly a bounce-back effect stimulated by above-average rainfall in 1932 after the 1931 droughts. The specialists also noted inadequate seed cleaning, and neglect of weeding and unwillingness to weed on the part of kolkhoz managers and peasants, which was in part a continuation of peasants' traditional practices. The government sent soldiers to Ukraine to help with weeding and harvesting in August 1932, but by then it was too late to have much of an effect (Tauger 2001b, 37–39; Nazarenko 2019, 186–188; Pyrih 2007, 290). These sources indicated that while more effective weeding work could have reduced these problems partially, the scale and severity of the infestations in 1932 would still have caused serious losses, and that fighting weeds involved much more than simply weeding work. The Soviets were also still in the early stages of developing chemical weed treatments (Volkov et al. 1935, 27–28, 179–196).

The weediness problem raises the issue of whether better incentives for peasants to conduct weeding could have affected the 1932 harvest and the famine. I would argue that such measures could have helped, but only to a limited extent. As noted above and will be noted further below, weeds were only one of a complex of environmental factors that reduced the 1932 harvest and overwhelmed both farmers and specialists. Also, as noted above, even before collectivization and under the Tsarist regime, peasants were known to weed poorly or not at all, even when they were managing their own crops for subsistence and sale. Collectivization did not eliminate incentives, because peasants earned labor days for their work and those who did more work would receive higher incomes in money and kind. The problem in 1932 was that the harvests were so small in many regions that even peasants who earned many labor days did not get much income. I documented in my dissertation that collective farms had many organizational problems, and also considerable autonomy, and sometimes re-created pre-collectivization working conditions for peasants (Tauger 1991a, 236–237, 249–250, 339–340, and elsewhere). Perhaps most important, the larger harvests in 1933 showed that kolkhoz peasants, including in Ukraine, also worked as hard as they could during the peak of the famine in spring and summer 1933 to produce a better harvest. Their “incentive” in this period was to produce enough to survive. In addition, they received considerable aid from several thousand industrial workers sent to the villages as “political departments” of the Machine-Tractor Stations, who worked hard and mostly successfully to improve labor organization and work in the kolkhozy (Tauger 1991a, 404–510). So better incentives and organization could

possibly have reduced somewhat the severity of the weed infestation in 1932, but probably not the other environmental factors.

Those other factors also included massive pest infestations that severely damaged crops in many areas. One of these was a widespread infestation of mice. Studies had shown that mouse infestations recurred every ten or eleven years; the last was in 1921–1922 (which contributed to the Soviet famine in those years), so the one in 1932 was part of that cycle. Delays in crop harvesting and moving cut grain off the fields and inadequate grain storage also enabled mice to consume large amounts of grain. Mice also consumed large amounts of forage that in many places left livestock starving. In some regions, officials and *kolkhozy* took measures to destroy the mice, but in other areas they minimized or overlooked these infestations (Tauger 2001b, 39–40; Nazarenko 2019, 188–189). There were also significant insect infestations, from locusts, beet weevils, meadow moths, and many other insect pests, that spread very rapidly because of the warm, humid weather in 1932 that followed the drought of 1931—again apparently a bounce-back effect. These insects infested millions of hectares in Ukraine, the Volga region, the Caucasus republics, Kazakhstan, and Central Asia, and reduced or destroyed harvests. Many farms and Soviet anti-pest agencies worked to eradicate them, but the scale of infestations overwhelmed them. They also had limited amounts of pesticides and those they had were often ineffective (Tauger 2001b, 18–20; Nazarenko 2019, 188). The documents I am preparing for publication contain hundreds of descriptions of all of these infestations.

Naumenko uses statistics from weather stations to argue that the weather was not a serious problem (Naumenko 2021a, 173–174, 180). Yet the issue was not simply weather, but all the environmental and agronomic factors discussed above, including weeds, insects, rodents, plant diseases, and more. Her calculations about predicted harvests are consequently invalid because they are based exclusively on temperatures and rainfall. Agriculture production, especially grain farming, is much more complex than that, which is why there are many university departments and institutes all over the world where researchers study every aspect of agricultural production, and companies that produce and sell many chemicals and other treatments to help plants resist a wide range of environmental threats. It is an invalid oversimplification, and in conditions of 1932 extremely inaccurate and misleading, to reduce Soviet harvests to temperature and rainfall. Consequently, her tables, calculations, and assertions, which ignore or minimize the decreased harvests and consequent famine conditions caused by these environmental disasters, are invalid because they omit the most important scientific evidence. Her arguments on these points are unjustified and incorrect, and seem to reflect some degree of ignorance about agriculture.

As noted above, Naumenko (2021a, 176) also asserts that published archival

documents do not discuss environmental factors more than usual, which was almost never. She reaches this conclusion based on a study of the frequency of the use of keywords, such as rust, plant disease, pests, and other terms related to famine, in two sets of published documents: a large collection of documents from the Soviet secret police produced by Russian scholars, and a broader collection of documents on “the tragedy of the Soviet village” from 1927–1939 produced by an international group of scholars (Berelovich and Danilov 1998–2012; Danilov et al. 1999–2004). In her appendices she presents numerous graphs showing the use of these words in these documents during the 1930s. These graphs show that the terms concerning infestations and pests, were used extremely rarely in these documents, especially compared to words referring to weather, like heat, drought, torrential rain, and especially human actions, like drunkenness, mismanagement, theft, and neglect (Naumenko 2021b, 22–28).

Yet the fact that these published archival documents did not discuss pests and infestations is not evidence that they did not occur. Virtually all of these documents were political documents, by officials and secret police personnel; virtually none of them were written by scientists. Their neglect or dismissal of such information reflects a deep-seated bias among Russian and later Soviet Marxists. During the first decades of the Soviet regime, Soviet leaders, like Marxists in the Tsarist regime, knew little about agronomic and environmental factors in farming, ignored or minimized them, and attributed famines to the actions of people, usually the Tsarist regime before 1917, and the peasants or their own lower-level personnel when they were in power. An example from the Tsarist era is the book on how Russian socialists should struggle with famine in 1906, by the founder of the Russian Marxist party, G. V. Plekhanov, which never even mentions the environmental factors that caused crop failures and famine that year, and blames Russian famines on the Tsarist regime (Plekhanov 1906). Stalin continued this pattern in his statement at a Central Committee plenum in January 1933 that the 1932 harvest must have been larger than that of 1931 because there was no drought in 1932, although there were certain “unfavorable climatic conditions” (which he did not identify) in 1932 (Tauger 2001b, 8; Stalin 1955, 13:220–221). Stalin did not mention any of the other factors that Soviet scientists recognized in 1932 as extremely damaging to the harvests, as noted above. Sometimes leaders did recognize natural disasters, as when Kosior referred in his February 1933 report to the Central Committee to the winterkill that caused the 1928 crop failure in Ukraine (mentioned above). But these leaders were not agronomists and were focused on blaming people rather than understanding the environment. And the sources cited above show that there are vast numbers of archival sources that document these problems that Naumenko did not use or even mention.

## Soviet policies

After rejecting ‘weather’ as an explanation of the famine, Naumenko, along with her collaborators in the NBER paper, attempt to attribute the famine and the deaths it caused to “Soviet policies”, mainly collectivization and to a lesser extent the lack of “Group A” industries in regions most severely struck by famine. In the following discussion I will focus on two major problems in her approach to Soviet policies, both based on false claims about or misrepresentation of those policies.

### **Private trade in food**

The topic of private trade is an important part of Soviet history, and Naumenko refers to it multiple times in all three papers, but all of those references are completely or almost completely wrong. Because this is a historical issue, I will discuss these references in historical order.

Both the CMEPR paper and the NBER paper contain a section entitled “Background,” and in both Naumenko asserts with no sources or evidence that the Bolshevik regime during the Civil War of 1918–1921 eliminated money and prohibited private trade. In the CMEPR paper, she refers to the Civil War and “experiments with ‘communism’ (abolishing money and the prohibition of private trade)” (Naumenko 2018, 6). The NBER paper states “War Communism [an important Soviet economic policy during the Civil War] banned money and trading of foodstuffs and *prodrazverstka* aimed to extract all ‘surplus’ grain from peasants. The peasants resisted by not working. Sown area in 1921 was 30% lower than the 1913 level” (Markevich et al. 2023, 8).

In fact, the Bolsheviks did not abolish money, they continued and intensified the increased currency emissions of the Tsarist regime during World War I. As a result, they printed more money than any previous Russian government, responding to the rapid inflation during the Civil War, and thereby created a hyperinflation that lasted beyond the Civil War. According to contemporary documentation, the Tsarist, Provisional, and then especially the Soviet government printed and distributed billions and trillions of rubles. The Russian economist S. S. Katzenellenbaum documented that during the World War, revolution, and Civil War the amount of paper rubles in circulation in Russia increased from 1.63 billion rubles in July 1914 to 18.9 billion rubles in October 1917 to 225 billion rubles in January 1920 to 1.168 trillion rubles in January 1921 to 17.5 trillion rubles in January 1922 to 178 quadrillion rubles in January 1924, when a new currency was finally introduced. The new regime printed so much because the government and the Red Army and ordinary people faced rapidly increasing inflation, and the regime had to print



ever more money to cover its expenses. In addition, many towns and regions in Russia began printing their own currencies, some of which the Soviet government authorized, and foreign currencies circulated in many border regions. The Bolsheviks even allowed use of the money issue by their opponents, the White Armies (Katzenellenbaum 1925, ch. 3, especially the table 56–58, that shows the amount of currency in circulation every month from July 1914 to May 1924; Marks 2014, 128–132; Efremov 2012, 16–17). Steven Marks also notes that Lenin’s objectives were not to eliminate currency but to establish a new currency, which the Soviet regime finally did in 1924 (Marks 2014, 132).

The Bolsheviks also did not stop private trade. After they came to power in November 1917, workers began sending representatives to villages to buy or barter food. The Bolsheviks in spring 1918 attempted to impose a “grain monopoly” to stop inflation, but it failed, and in August 1918 Lenin and the leaders returned to the Tsarist wartime food supply system of *prodrazverstka*, which imposed a kind of food tax to guarantee workers’ supplies, but still allowed private trade (Lih 1990, ch. 6–7). Lih (1990, 184) notes that officials recognized *razverstka* resembled a tax and admitted that it provided an incentive for peasants to produce more because larger production would leave a surplus for the peasants. The *New York Times* recorded the Tsarist regime’s introduction of the *prodrazverstka* (6 October 1915, p. 3, “Russia takes over food”). Lenin in a speech on 30 July 1919 admitted that in “a thorough study, it was found that this spring and summer the urban worker obtained about a half his food from the Commissariat of Food and had to buy the rest on the open market, as Sukharevka, and from the profiteers” ([link](#)). Sukharevka was a market in Moscow, but similar markets survived in most if not all other towns (see, e.g., the photo of the Nizhnii Novgorod market in Lih 1990, before p. 167). Even the western press reported on private food trade, “speculators,” and extremely high prices, which could not have happened if money had been banned.<sup>9</sup>

Finally, Naumenko’s claim that the peasants responded to the Civil War policies by “not working” is also not correct. First, as documented above and in the footnotes, private trade and speculation continued during this period despite the Bolsheviks’ attempts to control it, which clearly indicated that at least some peasants did attempt to produce for the market. A contemporary Russian economist who was an eyewitness to these events, L. N. Litoshenko, wrote a book

---

9. For example in the *New York Times*: February 18, 1918, p. 1, “Hunger and plague threaten Russia,” which notes that prices were “grotesque;” February 25, 1920, p. 2, “Russia under Reds much as of old,” notes that “profiteering is general,” and that trade continued despite the government policies against “speculation” and prices were “exorbitant;” August 21, 1920, p. 2, “Russia is stripped to supply Moscow,” notes that “Speculators operate despite stringent laws—prices beyond the workman’s income;” and June 26, 1922, p. 21, “Currency inflation on fantastic scale,” notes that “Russia’s new issues reported 58 trillions in April.”



about this period that was only published about 70 years later, in which initially he asserts that peasants cut back their crop sowings in response to the *prodrazverstka*, but then at the end of the book finally admits that they cut back their sowings mainly because they lacked sufficient seed and livestock to sow as much as they had before (Litoshenko 2001, 441; Tauger 2004). Also, during the revolution and Civil War, many peasants seized land from landlords and divided it among themselves, but peasants had lower productivity than landlords, so their land seizures resulted in lower farm production (on higher harvests on landlord estates, see Ermolov 1909, 144, citing a study by the Imperial Free Economic Society).

Naumenko makes similar false claims regarding trade in the late 1920s and collectivization period. In her JEH article, Naumenko claims at least twice that Soviet authorities “banned private trade in food” as part of the collectivization campaign (2021a, 158, 163). She also wrote that the government allowed *kolkhoz* peasants to trade at “*kolkhoz* markets” only after the famine (*ibid.*, 165). The CMEPR article also claims private trade was “mostly banned” from 1928 (Naumenko 2018, 8). The historical evidence, however, contradicts her claims and discredits some of her main arguments about the famine.

While the Soviet regime from 1928 did crack down on the ‘Nepman,’ the private producers and traders of the NEP period (1921–1928), it could not and did not stop peasants’ trade in agricultural products. During the first collectivization campaign from November 1929 to March 1930, some “zealous” local organizers (almost all of whom came from towns) went beyond their task and tried to introduce ‘socialism’ by closing peasant markets. When reports on these and other actions reached top Soviet authorities, Stalin called off the first collectivization campaign on 2 March 1930 and criticized those actions as “excesses” in his article “Dizzy with Success” published that day in most newspapers (Davies 1980a, 269ff.; *Pravda*, 2 March 1930). Two weeks later, on 15 March, the Communist Party Central Committee (TsK) published a long decree—“decree of TsK BKP(b) on the struggle with distortions of the party line in the *kolkhoz* movement”—in *Pravda* and other newspapers in the USSR that condemned these ‘excesses.’ It included the following statement: “Finally the TsK considers necessary to note completely impermissible distortions of the party line...also in the area of trade turnover between town and village. We have in view...the abolition in a series of places of markets and bazaars, leading to worsening supply of towns.” The TsK then ordered in this decree: “6. Prohibit closing of markets, restore bazaars and not restrict sale by peasants, include *kolkhozniki*, of their produce on the market” (Egorov and Bogoliubov 1984, 5:103–104).

In her JEH article, Naumenko on this point (Naumenko 2021a, 163) included a parenthetical reference to Davies, *The Soviet Collective Farm* (Davies 1980b), but with no page number, about the abolition of private trade. In fact, in that

book Davies wrote that the policy ideas of eliminating private food trade were “repudiated” in February 1930, and that the authorities authorized private trade, citing the 14 March decree discussed above and also citing the decrees that allowed kolkhoz peasants to retain private plots and livestock (1980b, 159–161). Naumenko did not mention this passage in Davies’ work at all, either overlooking it or choosing to ignore it and suppress the evidence that contradicted her false claim that trade was banned.

As a result of the 14 March decree, peasants’ trade in agricultural products on both rural and town markets continued. During 1930, Soviet economic publications and some official statements argued that because of growing inflation of food prices (again clear evidence of the continuation of private trade), the future would see “planned exchange.” On 10 May 1931, however, Sovnarkom (the Council of People’s Commissars, the top Soviet government institution) issued a circular that reversed this viewpoint and set the goal of promoting Soviet trade and ultimately abolishing rationing (Davies 1982, 22–23). The government also kept track of prices on these markets, and the indices of prices compared to 1928 as 100, reached 662 by 1931, i.e., grain prices more than sextupled by the beginning of the famine, mostly because of the low 1931 harvest and consequent food shortages in towns and industrial sites throughout the USSR, which I discuss below (Kerblay 1968, 122; on the low 1931 harvest and food shortages throughout the USSR, see Davies 1996, 176–196). The state planning commission (GOSPLAN) conducted a study of peasant food trade and estimated that peasants earned approximately 2 billion rubles from produce sales on private markets in 1929, 2.8 billion rubles in 1930, and 4.5 billion rubles in 1931 (Malafeev 1964, 130–131, 169; Davies 1982, 24). Another study showed that kolkhoz peasants earned more money from private trade than from work in kolkhozy in 1931–32 (Zvezdin 1968). In October 1931 the government issued another decree emphasizing the need to develop private trade by collective farms once they had fulfilled state procurements. By this time, kolkhoz private trading “had grown to such proportions that the authorities, rather than attempt to suppress it, decided to tax it” (Whitman 1956, 386). After that decree, local officials tried to encourage peasants to sell produce at “Soviet prices,” but with no success. By January 1932 at the 17th Party Conference, the party leader of the North Caucasus krai (another term for province), Boris Sheboldaev, stated:

Also evidence of the inadequacies of our work for improving turnover between town and village is that we still have bazaars that are not Soviet, but simply private bazaars—speculative bazaars. Almost in every stanitsa [village], almost in every town, I think, in all provinces there are bazaars, which trade again the same eggs, butter, milk and series of other relatively small goods, but these bazaars allow the existence of speculative elements who try to adapt to collectivization.... (VKP(b) 1932, 210; Davies 1982, 24)

Simultaneously with this pattern, in 1931 the regime imposed the highest grain procurements on farms up to that time, yet the drought, as discussed above, substantially reduced grain harvests in most of the main grain-producing regions. As a result, millions of people in rural areas (as well as in towns) were starving, and many kolkhozy, sovkhozy, and non-collectivized peasants did not have enough seed for the next year's sowings. In response, as discussed above, the regime issued a series of decrees to provide food and seed aid to villages over the following months. To obtain supplies for this purpose, the government had to reduce supplies allocated to rations for many of the 38 million people receiving rations that year in towns and industry, and cut in half planned exports of grains (Antipova et al. 2009; Pyrih 2007).

In response to the near-impossibility of stopping private food trade by peasants and kolkhozy, and the desperate need to supply the growing cities, Stalin and other Soviet leaders decided to compromise and took the changes of 1931 a major step further. In three decrees in May 1932, they reduced grain and other food procurements from kolkhozy and non-collectivized peasants, and explicitly authorized private trade by kolkhozy and peasants "at prices formed on the market." These decrees eliminated taxes on private trade by kolkhozy and collectivized peasants, and reduced taxes on non-collectivized peasants. These decrees also increased the range of food products that were removed from the rationing system and available only through peasant markets and Soviet cooperatives. As Davies pointed out, "In view of the huge scale on which the free market in food products was already operating semi-legally, this decree in terms of current practice merely regularized an existing activity" (Davies 1982, 25; Sharova 1957, 411–413, 416–419; Pyrih 2007, 149–152; on products removed from rationing, Davies 1996, 206; Tauger 1991b, 72). Naumenko's claim that the regime allowed private trade only after the famine is thus totally incorrect and suggests that she is ignorant about basic features of Soviet history.

The only constraint, which was imposed in the first of these laws, issued on 6 May, was that peasants and kolkhozy had to fulfill those reduced procurement goals and gather necessary seed for the coming crop year before beginning to trade grain. The decree specified that grain trade could not begin until 15 January 1933. Yet peasants, traders, and even some Soviet personnel ignored this regulation. The Politburo of the Ukrainian SSR issued a decree on 9 August reporting on speculation in grain trading and emphasized the need for local officials to control it. A decree of the Ukrainian Commissariat of Justice on 15 August reiterated this point and listed fines and other penalties for particular forms of grain trading. Finally on 16 September the central office of the Soviet security police (OGPU) issued a secret circular that stated that "according to available information, sale of grain and flour of the new harvest is being conducted at bazaars and markets almost

everywhere, despite the government decree prohibiting sale of grain from the new harvest until 15 January 1933” and that “OGPU up to now is not conducting sufficiently active struggle with this phenomenon, allowing trade in grain and flour.” The circular ordered a series of measures against grain trading. It is extremely important to note, however, that this circular explicitly directed these measures only against grain trading and not against all trade: the circular had 11 directives, and the 7th directive stated: “7. Categorically prohibit confiscation at bazaars of any other agriculture products, except grain and flour” (Pyrih 2007, 283–285, 287–289, 322–324). These decrees clearly aimed not to stop all private trade, but only trade in raw grain products, because grain was the primary food source that the regime needed to supply the population of the towns and industrial sites, and because they were still exporting a limited amount of grain to pay for imported supplies.

Finally, in case readers suspect that these Soviet sources are just ‘propaganda,’ we should note that Western observers in the USSR in these years clearly documented this private trade. The Third Secretary of the British Embassy to the USSR, J. M. K. Vyvyan, prepared a report based on his week-long tour of Ukraine and Crimea in July 1932 in which he wrote:

The institution of so-called collective farm trade, combined with the establishment in every town of large free markets encouraged by the State, is a dominant feature of the Soviet Government’s policy of decentralisation of distribution. ... [He noted that meat products were in short supply in those markets, but that] the collective farm markets are, however, thronged. In Simferopol, with a population of about 100,000, I do not think that less than 5,000 people were congregated in the bazaar in the morning. The produce sold consists in main of bread, meal, vegetables, milk, eggs, butter, potatoes and large quantities of fruit. (Carynnyk 1988, 90–91)

This description resembles the one Sheboldaev presented seven months earlier, and also supports the Gosplan study showing peasants earned billions of rubles from private market sales in the previous years, as well as all the other sources cited about peasant violations of restrictions on grain trading in 1932. In all three of these papers, Naumenko, and in the NBER paper her colleagues as well, repeat completely false claims that the Soviet regime banned money during the Civil War and private food trade both then and during collectivization. The above sources show that the regime was not nearly as repressive of peasants’ activities as they implied. Their false claims also suggests that all of these writers were ignorant about Soviet history, and worked from extremely biased assumptions about that history that they did not bother to verify.

## Procurements

The most important topic for the arguments regarding famine mortality in Ukraine in these three papers is grain procurements. Ukrainian nationalist literature views the grain procurements as the means by which the Soviet government imposed a ‘man-made’ famine. Naumenko’s papers present inconsistent viewpoints on this. In her JEH article, Naumenko asserts that the Soviet government “made an extreme effort to procure as much grain as planned” in 1932, which essentially repeats the conventional Ukrainian nationalist viewpoint (Naumenko 2021a, 164). She gives as a source Davies and Wheatcroft’s 550-page study *The Years of Hunger*, but she does not specify page numbers or provide any other explanation. This claim is part of her argument in the paper that the main cause of famine mortality was higher procurements from kolkhozy. This paper does not acknowledge, however, that the Soviet regime reduced procurements in 1932. The CMEPR paper acknowledges procurements as a factor in the famine but argues that kolkhozy were less productive in 1932 (using the sources I first presented, as discussed above) and that procurements were lower, but the article never presents the actual procurement data nor documents how and why the government reduced the procurements (Naumenko 2018, 3, 27–28). The NBER paper does acknowledge that the government reduced procurements in 1932, but emphatically and repeatedly argues that grain procurements were the direct cause of the famine, that the regime imposed harsher procurements on regions and kolkhozy populated by Ukrainians than on regions and farms with little or no Ukrainian population, and that these actions were the result of anti-Ukrainian bias of Soviet leaders and officials (Markevich et al. 2023, 1, 3, 5, 11, 36).

Yet again, all of these arguments in these papers are either incorrect or extremely misleading, because the changing Soviet policies toward private trade also affected grain procurements in this crisis year. It was mentioned above that the grain procurement plans were reduced in 1932, but in the following I will explain this in more detail because of what it shows about real Soviet policies compared to Naumenko’s claims.

Among the May laws regarding private trade discussed above, the one issued on 6 May 1932 reduced grain procurement quotas for every grain producing region of the USSR. In particular, the 6 May decree reduced the procurement quota for kolkhozy and non-collectivized peasants in Ukraine from 434 million puds (7.1 million tons) to 356 million puds (5.83 million tons), a reduction of 18 percent, much more than the reductions in any other region. The decree attempted to make up the difference with a marginal increase in procurements from sovkhozy, which were a small but growing subsector of agricultural production, and mostly by authorization and endorsement of private kolkhoz trade (Sharova 1957, 412ff.).

Naumenko (2021a, 164–165) does mention that the Soviet regime allowed peasants to sell private produce on “kolkhoz markets”—but her discussion makes it appear as if this decision came in the mid- or late 1930s, and she provides no date or documentation for this. In fact, as discussed above, kolkhozy were allowed to trade from the beginning of collectivization, and the law on kolkhoz markets came during the famine crisis, in May 1932, as documented here.

Yet this reduction set a precedent for more reductions. In August, once procurements began, Ukrainian leaders and leaders of other provinces appealed to the central government for more reductions in procurements. Stalin and the other leaders agreed to cut Ukraine’s grain procurement plan a second time, by 40 million puds (656,000 tons), over 12 percent of the plan that remained for Ukraine to fulfill. This proposal was approved (the decree specified that procurements were to be reduced by 39.5 million puds, holding back half a million puds of reduction in case further reductions were needed) and implemented over the next two weeks. In preparation of this measure, Stalin wrote to his subordinate in Ukraine Lazar Kaganovich and specified that this reduction was only for Ukraine, the other regions would have to wait:

As is evident from the materials, not only the Ukrainians but also the North Caucasus, Middle Volga, Western Siberia, Kazakhstan, and Bashkiria will speak with the Central Committee about reducing the grain procurement plan. I advise satisfying for the time being only the Ukrainians, reducing their plan by 30 million and only in extreme case by 35–40 million. As for the others, postpone discussion with them until the end of August. (Pyrih 2007, 290–298)

Stalin here clearly indicated that he considered reducing procurements for Ukraine a higher priority than for other regions. These are not the words of a leader who had a strong anti-Ukrainian bias. Naumenko never cites this source nor the document collection it came from.

The authorities reduced procurements for Ukraine a third time in late October 1932, after Stalin’s associate Viacheslav Molotov met with Ukrainian leaders, including Stanislav Kosior, the head of the Ukrainian republic Politburo, and the leaders of all the Ukrainian oblasti, and discussed in detail their views about how much the procurements should be reduced. Based on these discussions, he proposed to Stalin that Ukraine’s procurement target be cut by 70 million puds (1.1 million tons), almost double the previous reduction, including reductions for sovkhozy as well as for kolkhozy and individual peasants. This total reduction was rapidly approved and implemented (Pyrih 2007, 355–360). Here again, the regime reduced procurements substantially and only for Ukraine, in response to appeals by Ukrainian authorities, who were trying to alleviate the desperate situation of peasants and others in Ukraine. The regime reduced Ukraine’s grain procurement



plan a fourth time in January 1933, again in response to appeals from Ukraine and three other regions. In this decree, issued 12 January, the Soviet government reduced the grain procurement plan by 28 million puds (459,000 tons), in Ukraine, and much smaller amounts in three other regions (2 million puds in the North Caucasus, and half a million puds in the Urals and Kazakhstan) (Pyrih 2007, 597).

These documents on grain procurement reductions are drawn from a large document collection (over 1100 pages) on the “Holodomor” that was published in independent Ukraine in 2007 by Ukrainian scholars and students under the auspices of the Ukrainian Academy of Sciences and the Ukrainian Institute of History. They are genuine, previously secret, Soviet archival documents that reveal what actually happened in this crisis, and not some sort of Soviet or Russian propaganda (Pyrih 2007; some of these documents were also published in a major Russian-American document collection, Danilov et al. 1999–2004, 3:515ff.).

As noted above, these four substantial reductions reduced Ukraine’s procurement plan for collective farms and individual peasants from the 1931 level of 7.1 million tons to 3.77 million tons, just over half of the 1931 plan. The total final plan for 1932 was somewhat higher, 4.2 million tons, because it included procurements from sovkhozy, which as noted were also reduced in October 1932. No other Soviet region had its procurements reduced as frequently, by that percentage, or by anything close to that amount. That these decrees were implemented is confirmed by several documents. Ukrainian leader Stanislav Kosior, in his report to the Party Central Committee in 5 February 1933, confirmed the total reductions in the procurement plans after the May 6 decree, 138 million puds, and even admitted that those reduced grain procurements from Ukraine had not been fulfilled by the end of 1932, which confirmed the press reports shortly before, as noted above (Pyrih 2007, 642–643).

Naumenko’s assertion in the JEH article that the Soviet regime “was not willing to accept the low harvest estimates and made an extreme effort to procure as much grain as planned” (2021a, 164) clearly overlooks these reductions of grain procurements from Ukraine, which are documented in her sources and in her 2018 paper (they were documented in a source she cited: Tauger 1991b, 73 n.14). Her JEH article also does not present the actual amounts of grain procurements in Ukraine in any of these crisis years, except in the “appendices.” The procurement total for Ukraine in 1932 that she lists in her appendices is the planned total of 4.2 million tons, after the four reductions (Naumenko 2021b, 59), but as documented above, that planned total was never fulfilled and actual procurements were significantly lower. Her 2018 paper did recognize these points in general terms, but also did not discuss the process by which the procurement plan was reduced or the fact that the plan was not actually fulfilled. Most notably, the NBER article recognized that the regime reduced procurements, citing a table from Davies and



Wheatcroft's *Years of Hunger*, but the NBER paper brushed over these procurement reductions rapidly and superficially (Markevich et al. 2023, 9). The NBER paper attributed the reductions to reduced grain production, without any consideration of the implication that if the leaders reduced grain procurements for Ukraine because its production was lower, then they may not have been so "biased against" Ukraine.

At the end of the NBER paper is also an attempt to attribute the famine to issues raised in a study of the Chinese famine of 1958–61, referring to the government's inability and unwillingness to adjust procurements because central authorities did not trust local officials, and that bureaucrats did not have an incentive to report conditions truthfully (Markevich et al. 2023, 30). Yet the documents quoted and discussed above show that Stalin and Molotov did recognize food supply difficulties and did not arbitrarily procure "as much as possible" with no consideration for local needs. Molotov consulted with local and regional officials, trusted their information and views, and communicated this to Stalin, and Stalin and his associates reduced grain procurement targets for Ukraine three times after the May 6 general reduction, in response to their appeals. These sources clearly undermine the NBER paper's comparison with China and its claims about the total "anti-Ukrainian bias" of Soviet leaders and their actions.

These sources and data also contradict one of the NBER paper's main claims: that the regime imposed higher grain procurements on Ukraine-inhabited regions than on non-Ukrainian regions, with the result that Ukrainian regions retained less grain. The NBER paper presents none of the actual data to document these claims, but only statistical coefficients and similar calculation results. In Table 5, I present a similar calculation, but with the actual data, to show the per-capita grain procurements for Ukraine and other primary grain producing regions: the North Caucasus, the Middle and Lower Volga, the Central Blackearth oblast', and West Siberia. I use the rural population data from the 1926 census because this was the last census before the 1937 one, because all other estimates of population between these years are speculative and incomplete, and while the rural population grew by a few million between 1926 and 1932, several million of the rural population also left to work in the cities because of intensive recruitment of workers by the government and factory managers during the first five-year plan. Consequently, I would argue that the rural population in the 1926 census is not too far from the rural population in 1932.

**TABLE 5. Per-capita grain procurements from peasants and kolkhozy, 1932**

Region	1926 rural population, million people	Actual total peasant and kolkhoz grain procurements 1932, 1000 tons	Per-capita peasant and kolkhoznik grain procurements 1932, kg
Ukrainian SSR	23.66	3584	151
North Caucasus krai	5.92	1593	268
West Siberia krai	4.78	1054	209
Lower Volga krai	3.7	1185	320
Middle Volga krai	4.9	1159	236
Central Blackearth ob.	10.1	1797	180
<i>Sources:</i> Poliakov et al. 1991, 49–61 (table containing data for 1926 and 1937 censuses); Davies and Wheatcroft 2004, 478, Table 21.			

This table shows that the repeated procurement reductions for Ukraine, and the fact that Ukrainian peasants and kolkhozy did not fulfill the greatly reduced procurement quota, had the result that per-capita grain procurements in Ukraine were less, often significantly less, than the per-capita procurements from the five other main grain-producing regions in the USSR in 1932. These calculations, in which I use and present for readers the actual data, completely contradict the NBER paper's repeated claims that the regime imposed higher procurements on Ukraine than other regions, which that paper does not actually document with concrete data that readers can see and verify. The data and results in Table 5 further document the points made above that Soviet leaders did try to accommodate the food supply crisis in Ukraine by significantly reducing procurements below the levels of other regions, and that therefore the grain procurements in Ukraine did not reflect an "anti-Ukrainian bias." Finally, since Ukraine had such a severe famine despite much lower per-capita procurements, a central cause of the famine must have been a much lower harvest, which was mostly the result of the environmental disasters discussed above, and which Naumenko and her co-authors dismiss or ignore in all three papers.

To further understand the significance of these procurement reductions, it helps to view them in context. First, Ukrainian cities and industries were also part of the Soviet urban rationing system: most factories on the "special list" of sites to receive the highest rations were in Ukraine (Davies 1996, 178). This fact, which Naumenko never mentions, challenges her claims in JEH and the CMEPR paper that Ukraine did not have enough "Group A" industries and that lack of such industries led to famine, and the repeated references in the NBER paper to procurements being taken out of Ukraine (Markevich et al. 2023, 13, 24). In fact, Ukraine had many such enterprises. Consequently, much of Ukraine's procurements must have been used to feed Ukraine's town and industrial populations.

Second, as noted above, Soviet authorities reduced these grain procurements at a time when the towns and industrial sites were desperately short of food, and the regime had to reduce rations to make these accommodations for the rural population. Naumenko does not sufficiently take into consideration the essential point that the whole Soviet Union was enduring food crises in this period. Soviet procurements were not perpetrated arbitrarily to attack peasants, as Ukrainian nationalist interpretations usually claim and as the NBER paper tries to prove. Almost the entire urban population was living on inadequate rations provided by the Soviet government based on planned food procurements from the villages. Both Soviet and western sources documented the food crises in the towns: workers and their families were starving, fleeing from factory to factory in hopes of finding better food supplies, writing appeals to officials, and going on strikes and protests. Davies also presented evidence that urban mortality substantially increased in 1932 (Davies 1996, 176–192; Rossman 2005).

Naumenko's claim in the JEH article that the Soviet regime "made an extreme effort to procure as much grain as planned," and the NBER article's claims that the regime procured more from Ukrainian regions to kill Ukrainians, are false and misrepresent Soviet policies and actions. Soviet leaders clearly tried to balance the needs of towns and villages, both of which were enduring extremely serious crises of food supplies, by reducing procurements, especially in Ukraine, and by reducing supplies to certain urban groups. Both rural and urban populations had significantly higher mortality in this famine.

## Collectivization

In the JEH article and the CMEPR paper, Naumenko attempts to attribute most of the famine mortality in Ukraine in 1932–1933 to the Soviet policy of collectivization. Her arguments for this factor are based again on false data and historical misunderstandings.

First, she confused a crucial aspect of collectivization. She asserted that the regime allowed peasants in *kolkhozy* to have their own private plots only after the famine (Naumenko 2021a, 165), but this is completely wrong. From the beginning of the Soviet system in 1918 the regime allowed three types of *kolkhozy*: the *TOZ*, in which only a small part of the land was farmed collectively, and the rest farmed by peasants individually; the *artel'*, in which most of the land was farmed collectively, but the peasants retained small private plots and livestock, and the *kommuna*, in which all the land was collectively farmed. By fall 1917 about 100 collective farms of various types existed in Russia; by the end of 1918 there were at least 912, 61.8 percent of which were *kommuny* and 38.2 percent were *artely*. After that, however, the share of *kommuny* among the newly forming collective

farms fell, and the share of partly collectivized types, artely and TOZy, rapidly increased and greatly outnumbered the kommuny. Some kommuny became artely or TOZy and distributed part of their collective land to individual members of the collectives. By 1925, 59 percent of kolkhozy were artely, 28.9 percent were TOZy, and only 11.9 percent were kommuny (Danilov 1988, 291; Danilov et al. 1990–1991, 1:8, 2:81, 2:174–175). A report by the Commissariat of Agriculture on kolkhozy in 1927 showed the degree of “socialization” of kolkhoz assets by type. As presented in Table 6, in the great majority of kolkhozy by the late 1920s, half or more of the farms’ assets were not collectivized but were owned and used by members individually.

**TABLE 6. Percentage of socialization of kolkhoz assets by kolkhoz type, RSFSR 1927**

Kolkhoz asset	Kommuna	Artel <sup>a</sup>	TOZ
Land	100.0	55.5	32.6
Sown area	100.0	53.1	17.8
Tools	99.6	75.3	33.2
Draft livestock	97.3	13.5	2.9
Productive livestock	92.8	12.2	3.0

*Source:* Danilov 1988, 293, based on a 1929 publication of the Agriculture Commissariat.

During the first collectivization drive, there were attempts by some radical organizers to force new kolkhozy to be kommuny, but these were few cases, and the government considered them to be ‘excesses’ like the above-mentioned closing of markets. Stalin in his “Dizzy with Success” article that ended the first collectivization campaign, published on 2 March 1930 in virtually all newspapers, insisted that all kolkhozy be arteli as most appropriate for the present stage of peasant farming, because it involved both socialized farming and non-socialized household farms and livestock.<sup>10</sup> Simultaneously with his article, the regime issued the Kolkhoz Model Artel’ Statute on 1 March 1930, also published in virtually all the newspapers, that specified that kolkhozy were to be artely, and in articles II and III that kolkhoz members were to have private homes, private land plots, including gardens, orchards, and other sectors, and private livestock (Sharova 1957, 282–283). The kolkhoz system thus always had a private sector, and numerous studies both Soviet and non-Soviet have documented the great importance of the kolkhoz private plots in providing food supplies for the Soviet Union (Kerblay 1968; Whitman 1956; Wädekin 1973).

Then Naumenko elliptically and incompletely explained what a kolkhoz was. She wrote that most of the land, livestock, and implements belonged to the

10. A translation of Stalin’s article can be found at marxists.org ([link](#)).

kolkhoz, and that the peasants did not decide when or what to plant, but had to follow plans sent by the government (2021a, 162). This is misleading because it omits any comparison with the pre-existing traditional peasant farming system. Before collectivization, for centuries, peasants lived in corporate villages usually called a commune, or a ‘hromada’ in Ukrainian-speaking regions, in which village lands were divided into dozens and sometimes hundreds of strips, separated by borders to enable access without stepping on another peasant’s strip, and grouped into large fields by crop. This land use pattern was called interstripping; it dated back at least to 16th century Russian serfdom, but was maintained in the emancipated peasant villages after the emancipations of the serfs in the 1860s. Russian agronomists had long criticized it for inefficiency and low productivity, and also because those borders allow the spread of weeds and other infestations. In most of these villages, these strips would be redistributed every decade or so as family sizes changed. In Ukraine a minority of these villages had “household tenure” and did not redistribute strips, but this pattern was limited to western Ukraine, and was a matter of considerable conflict in that region; most peasants in the Ukrainian provinces lived in repartitional communes.<sup>11</sup>

In all of those villages, farm work was guided by the village as a whole: the village decided which fields would grow spring crops, winter crops, or lie fallow. Also, until the last two decades of the Tsarist regime, obligations such as taxes were imposed on the village as whole. These practices indicated that in many ways peasant villages already anticipated collective farms. This was one of the reasons why resistance to collectivization was relatively limited, and many peasants willingly formed or joined kolkhozy. In the JEH article, Naumenko wrote that kolkhoz members did not decide what to plant, because the government sent plans dictating what they were to grow (Naumenko 2021a, 161–162), but her statement makes it appear as if the regime forced the peasants to grow crops that were totally new and different. In fact, the kolkhozy grew almost exclusively the same crops they had grown for centuries, including rye, wheat, oats, and barley, as well as crops introduced in the 19th century such as sugar beets. The main change the kolkhoz brought was to replace interstripped fields with larger, consolidated fields, emulating especially farming in the United States, and it did improve efficiency: studies of kolkhozy in 1930 found that they farmed a larger area and produced

---

11. On this, see Holobuts’kyy (1970, 270), who noted that 51.5 percent of the farmland in the Ukrainian gubernias was under repartitional tenure, and that farms with household tenure were “exceptions.” Even in the western part of the Ukrainian Soviet republic, where most peasants held land in non-repartitional ‘household tenure,’ they farmed the land in interstripped patterns just as in regions where the repartitional commune was dominant as in eastern Ukraine and the Russian republic, and agronomists and officials saw this pattern as a problem that needed to be eliminated: see Tan 2000, especially 926–927. For an example of the conflicts about non-repartition, see Field 1989, ch. 3.

larger crops, employing only half the number of people, than the previous villages (Tauger 2005, esp. 75, 79). Kolkhozy did have many problems as they developed, but so did the previous communal villages, and the kolkhozy system facilitated improvement of agriculture more easily than the old villages. Kolkhozy also allowed large-scale mechanized farming with the Machine-Tractor Stations, similar to “custom cutters” in current U.S. farming (Slyn’ko 1961 and Miller 1970 document the important role of the MTS in Soviet farming even in this early period in Ukraine). Naumenko discusses the use of tractors only in the NBER paper, which acknowledges that the government increased the supply of tractors to Ukraine in the wake of the famine in 1933 to alleviate loss of labor and “boost production” in famine-stricken regions (Markevich et al. 2023, 6, 28–29)—another action that seems not to fit with an “anti-Ukrainian bias.”

Thus Naumenko’s claims that the kolkhoz system inherently led to famine are based on incorrect assumptions about the earlier character of peasant farming and the structure and operation of kolkhozy. Her sources on how the kolkhoz system allegedly led to famine are also problematic. For example, in the JEH article she cites an archival source, a Gosplan collection of statistical tables on the fulfillment of the first five-year plan in agriculture, according to which “in Ukraine in 1930, 27.9 percent of the harvest was extracted from collectives and 30.3 percent from individual peasants; in 1931, 42.8 percent was extracted from collectives and 32.4 percent from individual peasants; and in 1932, 45.1 percent was extracted from collectives and 40.6 percent from individual peasants” (Naumenko 2021a, 183–184; Markevich et al. 2023 also referred to this document, on p. 1 of its appendix). Yet she never explains what this source defined as “the harvest” and the actual amounts that were “extracted.” Fortunately, this Gosplan document collection is available on a website of Soviet historical documents (RGAE 1933). I reviewed the collection carefully, and nowhere does it specify the actual harvest for 1932. It has multiple tables with harvest data and percentages by sector for 1928–1931, and a few tables with columns for 1932, but with no data in those columns. The closest it comes is a table on page 14 that gives harvest yields for the whole Soviet Union for particular crops, but it gives a harvest yield for grain for 1932 that is the same as for 1930, 8.5 centners per hectare, which is most certainly false. The annual report data show clearly that the grain harvest and harvest yields in 1932 were much smaller than in 1930. This substantial inaccuracy clearly casts doubt on the reliability of this document collection.

Furthermore, as discussed above:

- the official harvest data for those years derived from pre-harvest sample forecasts and were matters of considerable dispute, and it was only at the end of 1932 that the government obtained substantial real harvest

- data from the kolkhoz annual reports;
- the procurement plans were reduced four times for Ukraine and in most regions even those lowered plans were not fulfilled, and procurements from February 1933 were shifted to be used as seed for kolkhozy;
- a significant part of the procurements went to feed other people in Ukraine, in cities and industries, so they were “extracted” from villages but not from Ukraine;
- substantial amounts of grain were returned to villages for food and seed (see sources above).

This Gosplan source does not address any of these facts, and is therefore useless for explaining famine deaths in 1932.

Then Naumenko in the JEH article describes her statistical calculations that she asserts connect famine mortality to collectivization. These calculations are extremely problematic because they treat key data as static when it was constantly and significantly changing. She asserts that collectivization in Ukraine reached 45 percent of peasant households in May 1930, then presents a graph that appears to show a decline later in 1930, but she never explains this or provides any data in the article about the changing level of collectivization (Naumenko 2021a, 162–163). She refers to “Appendix Table E1,” which indicates only that the percentage of rural households in kolkhozy was 16 percent on 1 October 1929, 38.2 percent on 1 May 1930 and 45.4 percent on 20 May 1930, 33.1 percent on 1 January 1931, and 69 percent on 1 June 1932 (Naumenko 2021b, 58). In her CMEPR paper, she has several tables showing statistical estimates of relationships between collectivization and mortality, but all of those numbers are statistical probabilities (Naumenko 2018, 65–71).

In the JEH article, in Table 2, “Size of collective farms and mortality: district-level estimates” (Naumenko 2021a, 185), she claims to base mortality in 1933 on collectivization data from 1930, three years earlier: the percentage of socialized land in 1930, and the per-capita sown area of kolkhozy and individual peasants in 1930. All of those factors, which she treats as static in 1930, actually changed significantly during 1930 and changed even more significantly by 1933. To take just one example, Table 7 below presents the official figures for the percentage of peasant households collectivized on certain dates in 1930, 1931, and at two points in 1932 for the USSR as a whole and for the Ukrainian republic:



**TABLE 7. Percentages of peasant households collectivized in Ukraine and USSR, 1930–1932**

1930	1 Jan	1 Feb	1 Mar	1 Apr	1 May	1 Jun	1 Jul	1Aug	1 Sep
USSR	18.1	31.7	57.2	38.6	28.0	24.8	22.5	21.9	21.5
Ukraine	15.9	30.5	60.8	46.5	41.5	36.3	31.5	29.6	28.8
1 Oct	1 Nov	1 Dec	1931	1 Jan	1 Feb	1 Mar	1 Apr	1 May	1 Jun
21.8	22.5	24.2		25.9	28.8	35.3	42.0	48.6	52.7
28.8	28.8	30.6		33.1	37.0	45.7	56.0	62.2	64.7
1 Jul	1932	1 Jan	1 Jun	1 Sep					
56.2		63.7	61.5	61.5					
65.6		69.2	69.0	69.7					
<i>Sources:</i> Davies 1980a, 442–443; Davies and Wheatcroft 2004, 488–489; Pyrih 2007, 307–308. The Russian-American document collection also has similar data on changing numbers of peasant households collectivized: Danilov et al. 1999–2004, 2:571–572, 674.									

As this table shows, the level of collectivization, in both the USSR as a whole and Ukraine in particular, reached a very high peak by 1 March. This level was partly only “on paper,” because many of those peasants had agreed, often under pressure, to sign a document that they had joined a kolkhoz, but had done little or nothing to start forming the kolkhoz before Stalin called off the campaign a few days later. Immediately after Stalin called off the first collectivization campaign, the number of households “collectivized” rapidly fell and continued to decline for the next 9 months. In addition, the number of households in kolkhozy significantly fell even during farm work, and many kolkhozy fell apart and divided their lands among individual households, which would also make averages meaningless (Davies 1980a, 279–281; Tauger 1991a, 133–159). Naumenko’s Table 2 showing “households per collective farm” in this context can only be an arbitrarily chosen number and not the reality, because as the data above show, there was no point from 1930 until late 1932 when there was a stable number of households per collective farm.

In Naumenko’s Table 2 are categories for sown area per capita in kolkhozy and “individual peasants” in 1930, but she does not indicate what date in 1930 is used for her “per capita” data (2021a, 185). Here again, in contrast to her assumptions, the area cultivated by kolkhozy and sovkhozy both massively increased in 1931–32, as shown in Table 8. This table has the same problem as her data, because it gives the total sown area in 1930 and 1931, years when the level of collectivization underwent significant changes. Yet the general pattern of a large increase in kolkhoz and sovkhoz sown area and a similarly large decrease in the area sown by non-collectivized peasants is fundamentally accurate. Naumenko in Table 2 tries to infer mortality in 1933 from the distribution of cropped land per capita between kolkhozy and non-collectivized peasants in 1930. Yet as Table 8 shows, the distribution of cropped land was utterly different in 1932–33 from in 1930. She also omits the sovkhozy, which had become an important sector by 1932.

TABLE 8. Soviet farmland by sector, 1930–1933 (million hectares)

Sector	1930 area	1931 area	1932 area	1933 area
Sovkhoz grain area	2.9	8.1	9.2	10.8
Sovkhoz total area	3.9	11.0	13.5	14.1
Kolkhoz grain area	29.7	61.0	69.1	75.0
Kolkhoz total area	38.1	79.0	91.5	93.6
Peasant grain area	69.1	35.3	29.5	15.7
Peasant total area	85.2	46.4	31.3	21.9
<i>Source:</i> Davies and Wheatcroft 2005, 453, drawn from Sel'skoe 1936, 259.				

Consequently, no matter how sophisticated her statistical calculations are, her attempt to connect conditions in 1930 with mortality in 1933 in both articles is just a numerical relationship that has nothing to do with the actual reality of the share of peasant households in kolkhozy and the cropped land by sector. This is the case because her presumed stable conditions in 1930 did not correspond to the reality of constant change that year, and because all of the conditions from 1930 that she uses in her calculations were completely different in 1932–1933, and her article never concretely addresses any of these differences. In her calculations in the JEH article and the CMEPR paper, she also omits from those calculations the fact, which as discussed above she partly documented only in the 2018 paper, that kolkhoz procurements in 1932 were much less than those in 1930. In the JEH article she claims that collectivization was the main cause of famine deaths because of grain procurements, but she does not recognize that the procurements from kolkhozy were substantially reduced that year, even though she documented that reduction in her 2018 article. The fact that a worse famine ensued in 1933 after much lower procurements in 1932 than the previous two years, in my view, can be explained only because of the much smaller 1932 harvest, caused first of all by the complex of environmental disasters in 1932, and also partly by the inadequacies of the government's, specialists' and peasants' responses to them.

She also seems to misunderstand or misrepresent the origins of some of the data she uses in her statistical calculations. For example, she wrote: “It is possible that Ukrainians just happened to live in lands better suited for grain production, and therefore more collectivized in 1930” (Naumenko 2021a, 191 n.51). Yet the Soviet regime planned that collectivization would be higher in the primary grain regions in 1930–1931, especially in Ukraine. Soviet leaders considered those regions better prepared for collectivization and they hoped such regions would be more productive as a result of collectivization, and many were in 1930 (Davies 1980, 167, 187–188; Tauger 2005, 79).

## “Anti-Ukrainian bias” and famine mortality

All three of these papers focus on the deaths of Ukrainians in the “Great Soviet Famine (1932–33),” as the NBER paper identified it on its first page. Yet on this issue the papers diverge significantly in their interpretations and ignore important implications of their data that contradict their arguments.

In her JEH paper, Naumenko questions claims, which she cites in the article, that the 1932–1933 “Ukrainian famine” was an intentional genocide against Ukrainians. She asserts that Ukrainians and Germans were “discriminated against” because “ethnic Ukrainians were more collectivized” and because both groups, according to her problematic statistics, had somewhat higher mortality in the famine, but she insisted that this was “*not* proof of genocide” (her emphasis). She concludes that “no strong evidence exists that government policies were enforced more harshly on ethnic Ukrainians or Germans” (Naumenko 2021a, 186–193). The CMEPR paper asserts that its data shows that the famine was “man-made” and “support those who argue that ethnic Ukrainians were targeted,” but admits that to prove genocide it would be necessary to show “that Stalin had the foresight that his policies would fail and lead to famine mortality years after they were introduced,” which she does not document (Naumenko 2018, 1, 39). In contrast to these papers, the NBER paper argues strongly that Ukrainian famine mortality was the direct result of “anti-Ukrainian bias,” which the authors claim to find in several aspects of the famine and for which they cite a recent book by the journalist Anne Applebaum, which is full of false claims from end to end (Applebaum 2017; for documentation of those false claims, Tauger 2018).

Yet the implications of what Naumenko and her collaborators wrote in these papers raise serious questions about these claims that they completely ignore. On mortality, the three papers present general data as follows: the JEH paper asserts that the “1933 Soviet famine ... killed six to eight million people” and “the 1933 Ukrainian famine killed as many as 2.6 million people out of a population of 32 million” (Naumenko 2021a, 1); the CMEPR paper asserts that “the 1933 Soviet famine killed six to eight million” and that “The famine of 1932–1933 in Ukraine killed as many as 2.6 million people out of a population of approximately 30 million” (Naumenko 2018, abs., 1); the NBER paper asserts that “During the Great Soviet Famine (1932–33), approximately seven million people perished” and that “approximately 2.1 to 3.15 million ethnic Ukrainians” died in the famine (Markevich et al. 2023, 1, 31).

Yet these different data imply two problematic points for the arguments of these papers. First, they indicate that total mortality is uncertain: despite all the archival sources that Naumenko and her collaborators use and all the sophisticated

statistical calculations they cite, they are still unsure how many people died in the famine in Ukraine and outside it. In the NBER paper, the authors acknowledge on several points, including Ukrainian mortality outside Ukraine, that they do not have actual data and they interpolate or estimate “data,” and they acknowledge “significant variation in famine mortality across and within provinces.” All of these points make their claims speculative and uncertain (Markevich et al. 2023, 13, 17, 18, 19).

Second, the numbers they present—2.1 million to 3.15 million Ukrainian deaths and 6 to 8 million total deaths—imply that non-Ukrainian deaths reached 3 million to 5 million, in other words from an equal number to more than double the number of Ukrainian deaths. If the large numbers of deaths among Ukrainians, in the view of the NBER authors, imply an “anti-Ukrainian bias,” then the even larger numbers of deaths of non-Ukrainians should imply a similar bias against non-Ukrainians, yet these three papers neither recognize nor discuss this clear implication from their data.

In the NBER paper, the authors admit that it is “beyond the scope of this paper to provide conclusive evidence on the motivations of anti-Ukrainian bias,” but they try to support and explain that bias with two “stylized facts”: that “Ukrainians offered stronger resistance to collectivization than other ethnic groups,” and that Ukrainian mortality was higher in places “more important for agricultural production,” which they claim implies targeting of Ukrainians in such regions (Markevich et al. 2023, 29).

On the first point, again the authors present the results of their statistical calculations but do not present the actual data regarding this resistance. That data is in a set of archival sources from the Soviet security police, the OGPU, compiled after the first collectivization drive. At the outset it is necessary to point out three aspects of this resistance that the authors ignore. First, in the great majority of cases, this “resistance” was non-violent, and basically consisted of peasants protesting against certain aspects of collectivization, such as transfer of livestock to the kolkhozy, and often those protests persuaded the people conducting collectivization to give in to peasants’ demands. Second, the total number of peasants and others involved in this resistance was on the order of 2.5 million, out of a total rural population in the USSR of at least 110 million, of whom about 70 million were 16 years old or older. In other words, the total number of protestors did not reach five percent of the adult rural population, which means that the vast majority of peasants did not protest against collectivization. Third, most of the protests took place in March 1930, in other words after the publication of Stalin’s article “Dizzy with Success,” which sharply criticized the personnel who carried out collectivization for using coercion and excesses, which provided peasants who read it with a basis for protesting and thereby stimulated the rise in protests (Viola

1996, 140ff.; Tauger 2005, 71–76). In other words, while the NBER authors try to show that protests resulted from Ukrainian nationality, they overlook the facts that protests were rare and were at least in part provoked by Stalin himself.

In Table 9, I present these data for Ukraine and other nearby regions and calculate the implied share of the total population involved in this “resistance”. Again I use the 1926 census data but I raise the figures slightly to account for population growth, and the figures for peasants involved in disturbances from the OGPU reports are also somewhat approximate as well, but these data still provide a basis for evaluating the relative degrees of resistance in Ukrainian and non-Ukrainian regions.

**TABLE 9. Share of population involved in resistance to collectivization in 1930 campaign**

Region	1930 rural population estimate, millions	Number of peasants in disturbances 1930	Percentage of rural population involved
Ukrainian SSR	24	956,587	3.99
North Caucasus krai	6	227,000	3.78
Central Blackearth ob.	10.5	315,035	3.0
Lower Volga	4	119,175	2.9
Middle Volga	5	140,383	2.8
<i>Sources:</i> Viola 1996, 140; Poliakov et al. 1991, 48–53 (section containing data from 1926 census).			

These data show that a slightly larger share of the population in the Ukrainian republic were involved in disturbances during the 1930 campaign than in the other regions, but the differences were clearly not very large, approximately 1 percent of the population, and the data indicate that in all of these regions, as in the USSR as a whole, the overwhelming majority of peasants, including in Ukraine, did not participate in this “resistance.”

Another source that weakens the NBER argument about “stronger resistance” among Ukrainians is the actual report produced by the OGPU on resistance in 1930, which was completed on 15 March 1931, and marked “completely secret:” “Dokladnaia zapiska Sekretno-politicheskogo otdel OGPU o formakh i dinamike klassovoi bor’by v derevene v 1930 g.” (Report of the secret-political department of OGPU on forms and dynamics of class struggle in the village in the year 1930) (Danilov et al. 1999–2004, 2:787–808). This report describes the various categories of peasants’ resistance during the collectivization campaign of late 1929 to March 1930, and in each category lists several regions where this resistance took place. The regions they list include Ukraine frequently but not always, just as frequently listed are the North Caucasus, the Volga regions, and the Central Blackearth region, and the lists of regions always end with “and others.” This report never states or even implies that Ukraine as a region stood out for a greater share of resistance. The report views this resistance as the result of “class struggle,” in other words

the efforts of certain groups within the villages—“kulaks,” members of opposition parties, and other “class enemies”—who could show up anywhere in the view of the OGPU.

All of this evidence makes it difficult to accept the NBER paper’s claims that Ukrainians offered “stronger resistance to collectivization” and motivated significant anti-Ukrainian bias (Markevich et al. 2023, 29). On the paper’s second “stylized fact” that more important agricultural regions suffered higher mortality, I would simply refer readers to the evidence that I presented above about the devastating infestations of plant diseases, insects, rodents, and weeds, which were documented in scientific publications and archival sources, and were particularly severe in Ukraine, especially because of the environmental conditions, including what Naumenko described as favorable weather, in 1932. Since the NBER authors, based on this “stylized fact,” attribute higher mortality in Ukraine exclusively to the facts that Ukraine was an important agricultural region and Ukrainians lived there, they are jumping to conclusions about Soviet leaders’ attitudes with little or no evidence for those attitudes—the NBER paper points out right at the beginning that “No direct documentary evidence that Stalin ‘ordered’ a famine has been uncovered” (ibid., 1)—and based on ignorance or suppression of legitimate scientific evidence of major environmental events that clearly were important causes of the famine.

Analysis of Soviet policies and environmental factors makes an even stronger point. Ukraine had a higher level of collectivization not because the Soviet government was trying to punish them. The Soviet leaders thought that collectivization would be an improvement: as I have documented from both published and archival evidence, they saw it as applying the U.S. model of modern mechanized farming in a socialist context. They did not see collectivization as “discrimination” against Ukrainians; they saw it as a reflection of—in the leaders’ view—Ukraine’s relatively more advanced farming skills that made Ukraine better prepared for collectivization (Davies 1980a, 166, 187–188; Tauger 2006a). And again, Naumenko’s minimizing of environmental factors and oversimplification of the environmental circumstances of 1932 leads her to view the higher mortality among Ukrainians as a sign of “discrimination,” when in fact it was the result of the devastating combined effects of the environmental disasters discussed above, which were particularly severe in Ukraine.

## Conclusion

The issues raised in this essay make two general points. The first one is the explicit point that Naumenko’s JEH article and two online papers are incorrect

about the most important aspects of the 1931–1933 famine and several related aspects of Soviet history that are important for those papers’ arguments. Her reductive analysis of “weather” overlooks the complexity and significance of much more diverse environmental factors behind the famine. Scientific evidence documented not only drought conditions in part of Ukraine in 1931 but also the serious effects of plant diseases, insects, weeds, and mice, that reduced the 1932 harvest greatly. All of these points were at least partly documented in her sources, but she never cited or acknowledged them, and instead inaccurately oversimplified, minimized, and ignored the crucial impact of these environmental factors on the famine, and the substantial evidence that I and Nazarenko published that documented their effects. Her attempt to blame Soviet policies for the famine involves overlooking or misrepresenting actual Soviet policies toward private peasant food trade, reduced grain procurements, especially in Ukraine, and the extreme variation and changes in collectivization in 1930–fall 1932. The claims in the NBER article that attribute famine mortality in the Ukrainian republic to “anti-Ukrainian bias” are based on statistical inferences from evidence that the authors never present in the article or that they “interpolate” and “estimate,” and the evidence regarding several of these issues that I present here or in earlier publications leads to completely different conclusions.

The second general point of this article, however, is the alternative conclusions regarding the 1931–1933 famine that these criticisms outline. Since the peasants freely traded food at village and town bazaars and markets from 1930 onward, since Soviet leaders compromised their communist principles and attempted to support peasants’ trade in May 1932 by cutting procurements and fully authorizing free trade by peasants, since peasants continued trade against Soviet restrictions and were not stopped, since the regime reduced Ukrainian peasants’ grain procurements three more times, after the May 1932 national reductions, to a level just over half of the previous year’s procurements and that were lower per-capita than procurements in the other main grain regions, and yet despite all these clearly pro-peasant policies and practices, the peasants and others in Ukraine, and peasants in much of the rest of the USSR, endured a much worse famine after those reduced 1932 procurements, the underlying explanation has to have been a small 1932 harvest. This conclusion seems especially unavoidable when one considers that the Soviets returned 5.7 million tons of grain to the Soviet countryside, including to Ukraine, in the first half of 1933, from their limited reserves and from grain procured in 1932, much more than in 1932 or 1934.<sup>12</sup>

---

12. Davies et al. 1995, 652–653, found that the regime provided to villages about 2 million tons of grain from 3 million tons held in grain reserves. The Central Statistical Administration prepared a table on the use of grain obtained in procurements and indicated that the government “returned to agriculture” some



Since these arguments lead to the conclusion that 1932 harvest must have been very small, especially in Ukraine, and since 1932 in most regions was not a year of drought like 1931, all of those specific scientific studies of agricultural conditions, including those cited above and others, provide the most immediate and persuasive explanations for a small 1932 harvest. These disasters, in addition, included not only the usual ones that had struck Russian farming for centuries, and as noted above also Soviet farming in 1921, 1924, and 1928: mainly severe weather like drought, extreme cold, winterkill, dust storms, and hailstorms. In 1932 Ukraine and other parts of the USSR were struck by disasters that the ostensibly favorable weather intensified: massive weed growth, widespread insect infestations, severe outbreaks of plant diseases, infestations of rodents. Now, to some extent, these factors were also intensified by collectivization, because that policy led to large, consolidated crop fields, which enabled many of these infestations to spread more easily and more widely, created more favorable conditions for rodents to hide and reproduce under large grain stacks, and made farm work more difficult for peasants because mechanization, which was the premise for collectivization, lagged behind. Yet these same problems also plagued peasant farming before collectivization, as documented for example in the Tsarist government's famine reports in 1906–1914, discussed above. They also plagued other places with large-scale farming, like the United States in the Dust Bowls of the 1930s. I documented in 2001 that several of these problems also plagued farms in Eastern and Central Europe in 1932, and caused food difficulties there as well (Tauger 2001b, 14).

Yet the environmental factors were fundamental, and the Soviet government and Soviet scientists responded to these crises with organizational and agronomic measures: not only increased production of tractors and other equipment, but also anti-pest measures, and especially development of plant varieties resistant to these threats. In noting these points, I am not trying to exonerate the Soviet regime completely for the famine. While the regime greatly reduced grain procurements in Ukraine, those procurements were clearly one of the important causes of the famine, and if the regime had not taken even that smaller amount grain from Ukrainian villages, the famine could have been greatly reduced or even eliminated. But if the regime had left that grain in Ukraine, then other parts of the USSR would have been even more deprived of food than they were, including Ukrainian cities and industrial sites, and the overall effect would still have been a major famine, even

---

3.5 million tons of grain from procurements; RGAE f.1562 op.74 d.19 (I have a photocopy of this archival table). These two figures are the basis for the overall total of 5.7 million tons of grain returned to the agricultural sector in Tauger 1991a, 74, 88, which were based on an earlier Soviet publication. Pyrih 2007 contains many documents reporting food and seed aid to Ukraine, as do other document collections like Danilov et al. 1999–2004, vol. 3. All of these data indicate larger transfers of grain to famine regions in 1933 for seed and food than Naumenko cites (e.g., Naumenko 2018, 9; Markevich et al. 2023, 10).

worse in “non-Ukrainian” regions. The biggest mistake of the regime in this period was to turn down the many offers of food imports from other countries, which they had previously accepted in the famines of 1921–23, 1924–25, and 1928–29. That decision made it impossible for Soviet relief efforts to alleviate the famine, both in and outside of Ukraine.

Still, many people in the USSR responded to the famine with efforts to prevent its recurrence. Perhaps the most important example of such an effort was the plant breeding work of the Soviet Ukrainian agronomist Pavel Panteleimonovich Luk’ianenko in the North Caucasus, the region where the 1932–1933 famine may have been most severe. In 1934 he published a study of the rust infestation of 1932 in his region and of the few wheat varieties that resisted it (Luk’ianenko 1934). He then set out to breed resistant varieties, using both Soviet and foreign wheat varieties, especially the semidwarf varieties developed by the Italian breeder Nazareno Strampelli (Tauger 2023). The culmination of this work came in the late 1950s when Luk’ianenko released his first semidwarf varieties, especially Bezostaia-1, which began a Soviet Green Revolution (Tauger 2017a; Luk’ianenko 1990; Rybalkin 2001). Luk’ianenko’s varieties were used widely in the USSR and several other countries, both for crop production and breeding, yet his work has been almost completely overlooked in western historiography. Yet the American leader of the Green Revolution, Norman Borlaug, recognized the quality of Luk’ianenko’s work.<sup>13</sup> His accomplishments were a long-term result of the 1932–1933 agricultural crisis and famine.

One basic definition of “political economy” involves the interaction of politics and economics. Yet to understand crises in agriculture, one must also take into consideration the environment in its broadest sense. I hope that future scholars who deal with these or other agricultural crises will be open-minded and aware enough to include the environment in their approaches.

## References

- Antipova, O. A., et al., eds. 2009. *Golod v SSSR 1930–1934*. Moscow: Federal Archival Agency.
- Applebaum, Anne. 2017. *Red Famine: Stalin’s War on Ukraine*. New York: Doubleday.
- Artemov, P. K. 1933. “K voprosu o porazhaemosti zernovykh kul’tur gribnymi bolezniami.” *Trudy po prikladnoi botanike, genetike i selektsii, Seriia A: Sotsialisticheskoi rastenievodstvo* 7, Leningrad.

---

13. Evidence of the wide use of Luk’ianenko’s first semidwarf variety, Bezostaia-1, can be found in Bonjean and Angus 2001, 202, 209, 270, 300, 304, 339–340, 356, 364, 387, 394, 721, 768, 840, 857, 859, 864, 887, 949, 963. For Borlaug’s recognition, see Tauger 2017a, 114.

- Belokurov, Evgenii Vladimirovich.** 2014. *Prodovol'stvennaia kampaniia 1911–1912 v Rossii: vlast' i obshchestvo v bor'be s golodom*. M.A. diss., Saint Petersburg State University.
- Berelovich, A., and V. Danilov,** eds. 1998–2012. *Sovetskaia derevnia glazami VChK-OGPU-NKVD 1918–1939: Dokumeny i materialy*, 4 vols. Moscow: ROSSPEN.
- Bonjean, Alain P., and William J. Angus,** eds. 2001. *The World Wheat Book: A History of Wheat Breeding*. Paris: Lavoisier.
- Carynnyk, Marco, et al.,** eds., 1988. *The Foreign Office and the Famine: British Documents on Ukraine and the Great Famine of 1932–1933*. Kingston, Canada: Limestone Press.
- Danilov, V. P.** 1988. *Rural Russia Under the New Regime*. London: Hutchinson.
- Danilov, V. P., et al.** 1990–1991. *Kooperativno-kolkhoznoe stroitel'stvo v SSSR 1917–1922: Dokumenty i materialy*, 2 vols. Moscow: Nauka.
- Danilov, V., R. Manning, L. Viola, et al.,** eds. 1999–2004. *Tragediia Sovetskoi derevni, Golod v SSSR 1929–1934*, 6 vols. Moscow: ROSSPEN.
- Davies, R. W.** 1980a. *The Socialist Offensive: The Collectivization of Soviet Agriculture, 1929–1930*. London: Macmillan.
- Davies, R. W.** 1980b. *The Soviet Collective Farm, 1929–1930*. Cambridge, Mass.: Harvard University Press.
- Davies, R. W.** 1982. Models of the Economic System in Soviet Practice, 1926–1936. In *L'Industrialisation de L'URSS dans les années trente*, ed. Charles Bettelheim. Paris: Éditions de l'École des hautes études en sciences sociales.
- Davies, R. W.** 1996. *Crisis and Progress in the Soviet Economy, 1931–1933*. New York: Palgrave.
- Davies, R. W., and S. G. Wheatcroft.** 2004. *The Years of Hunger: Soviet Agriculture 1931–1933*. New York: Palgrave.
- Davies, R. W., S. G. Wheatcroft, and Mark B. Tauger.** 1995. Soviet Grain Stocks and the Famine of 1932–1933. *Slavic Review* 54(3): 642–657.
- Efremov, Steven.** 2012. *The Role of Inflation in Soviet History: Prices, Living Standards, and Political Change*. M.A. Thesis, East Tennessee State University.
- Egorov, A. G., and K. M. Bogoliubov.** 1984. *Kommunisticheskaia Partia Sovetskogo Soiuza v Resolutsiakh i Resheniiakh S'ezdov, Konferentsii i plenumov TsK*. Moscow: Politizdat.
- Ermolov, A. I.** 1909. *Nasbi neurozha i prodovol'stvennyi vopros*, parts I and II. St. Petersburg: Tipografiia V. Kirshbauma.
- Field, Daniel.** 1989. *Rebels in the Name of the Tsar*. Boston: Unwin Hyman.
- Holobuts'kyi, V. O.** 1970. *Ekonoichna istoriia Ukraïnskoï RSR: dozbovtnennyi period*. Kyiv: Izdatel'stvo 'Vysha Shkola'.
- Hrynevych, Liudmyla.** 2013. *Holod 1928–1929 rr. u Radians'kyy Ukraïni*. Kyiv: Instytut istorii Ukrainy NAN Ukrainy.
- Ivnitskii, Nikolai Alekseevich.** 2009. *Golod 1932–1933 godov vs SSSR: Ukraina, Kazakhstan, Severnyi Kavkaz, Povolzh'e, Tsentral'no-Chernozemnaia oblast', Zapadnaia Sibir', Ural*. Moscow: Sobranie.
- Katzenellenbaum, S. S.** 1925. *Russian Currency and Banking, 1914–1924*. London: P. S. King & Son.
- Kerblay, Basile.** 1968. *Les Marchés Paysans en U.R.S.S.* Paris: Mouton.
- Kuleshov, N. N.,** ed. 1929. *Gibel ozimnykh khebor i meropriiatiia po ee preduprezhdeniiu: Trudy Vsesoiuznogo soveshchaniia po voprosu o gibel ozimnykh khebor v 1927–1928 v g. Khar'kov*

- [*Failure of Winter Grains and Measures to Prevent Them: Studies of the All-Union Conference on Question of Failure of Winter Crops in 1927–1928 in Khar'kov*]. Leningrad: All-Union Institute of Applied Botany and New Crops and Government Institute of Experimental Agronomy.
- Lewin, Moshe.** 1974. *Russian Peasants and Soviet Power. A Study of Collectivization*. New York: Norton.
- Lih, Lars.** 1990. *Bread and Authority in Russia, 1914–1921*. Berkeley, Cal.: University of California Press.
- Litoshenko, L. N.** 2001. *Sotsializatsiia zemli v Rossii*. Novosibirsk: Sibirskii Khronograf.
- Luk'ianenko, P. P.** 1934. O stepeni ugnetenii gibridov ozimoi psenitsy bujroi rzhavshinoi v 1932 g. v sviazi s rezul'tatami selektsii na immunitet, Krasnodar 1933 g. [On the Susceptibility of Winter-Wheat Hybrids to Puccinia Tritica Erikss during 1932 in Connection with Results Obtained in Breeding for Immunity]. Rostov-na-Donu: North Caucasus Selection Centre of the Seed-Growers Association “Edinenie,” Plant-Breeding Section, Winter Wheat Laboratory.
- Luk'ianenko, P. P.** 1990. *Izbrannye Trudy [Selected Works]*. Moscow: Agropromizdat.
- Malafeev, A. N.** 1964. *Istoriia tsenoobrazovaniia v SSSR (1917–1963 gg.)*. Moscow: Mysl'.
- Mal'tsev, A. I.** 1926. *Sornye rasteniia SSSR i mery bor'by s nimi*. Leningrad: All-Union Institute of Applied Botany and New Crops.
- Markevich, Andrei, Natalya Naumenko, and Nancy Qian.** 2023. The Causes of Ukrainian Famine Mortality, 1932–33. *NBER Working Paper* 29089. National Bureau of Economic Research (Cambridge, Mass.). [Link](#)
- Marks, Steven.** 2014. The Russian Experience of Money, 1914–1924. In *Russian Culture in War and Revolution, 1914–1922*, eds. Murray Frame, Boris Kolonitskii, Steven G. Marks, and Melissa K. Stockdale, Book 2. Bloomington, Ind.: Slavica.
- Miller, Robert.** 1970. *One Hundred Thousand Tractors*. Cambridge, Mass.: Harvard University Press.
- Ministerstvo vnutrennykh del (MVD).** 1906. *Obzor pravitel'stvennykh meropriiatii po bor'be s posledstviiami neurozhaia 1906 goda*. St. Petersburg: MVD.
- Ministerstvo vnutrennykh del (MVD).** 1908. *Otchet po prodovol'stvennoi kampanii 1906–1907 gg. Denezhno-material'naia chast'* [Report on Food Supply Campaign, Monetary-Material Part], t. 1–6. St. Petersburg: MVD.
- Ministerstvo vnutrennykh del (MVD).** 1913. *Otchet po prodovol'stvennoi kampanii 1911–1912 gg.*, books 1, 2. St. Petersburg: MVD.
- Mozzhukhin, I.** 1917. *Agraryi vopros v tsifrakh i faktakh deistvitel'nosti*. Moscow: Universal'naia Biblioteka.
- Naumenko, Natalya.** 2018. The Political Economy of Famine: The Ukrainian Famine of 1933. Center for Micro-Economic Policy Research, George Mason University (Arlington, Va.). [Link](#)
- Naumenko, Natalya.** 2021a. The Political Economy of Famine: The Ukrainian Famine of 1933. *Journal of Economic History* 81(1): 156–197.
- Naumenko, Natalya.** 2021b. The Political Economy of Famine: The Ukrainian Famine of 1933—Appendices. NatalyaNaumenko.com. [Link](#)
- Nazarenko, Nazar.** 2019. Sornaia rastitel'nost', bolezni i vrediteli kak factory goloda 1932–1933 godov. *Samarskii nauchnyi vestnik* 8(1): 186–193.

- Ostrovskii, A. V.** 2013. *Zernovoe proizvodstvo evropeiskoi Rossii v kontse XIX–nachale XX v.* St. Petersburg: Poltorak.
- Plekhanov, G. V.** 1906. *O zadachakh sotsialistov v bor'be s golodom v Rossii.* St. Petersburg: Schmidt.
- Poliakov, Iu. A., et al., eds.** 1991. *Vsesoiuznaia perepis' naseleniia 1937 g.: kratkie itogi.* Moscow: Academy of Sciences.
- Poznanskii, V. S.** 2007. *Sotsial'nye kataklizmy v Sibiri: Golod i epidemii v 20–30e gody XX v.* Novosibirsk: Izdatel'stvo SO RAN.
- Pyrih, Ruslan, et al., eds.** 1990. *Holod 1932–1933 rokiv na Ukraïni: ochyma istorykiv, movoin dokumentiv.* Kyïv: Vydavnytstvo politychnoi literatury Ukraïni.
- Pyrih, Ruslan, et al., eds.** 2007. *Holodomor 1932–1933 rokiv v Ukraïni. Dokumenty i materialy.* Kyïv: Kievo-Mohylians'ka akademiia.
- RGAE (Russian State Economic Archive).** 1933. *Plan financirovaniia kolkhozy systemy za 1-uiu piatiletku i ego vypolnenie.* [Link](#)
- Rossmann, Jeffrey.** 2005. *Worker Resistance Under Stalin.* Cambridge, Mass.: Harvard University Press.
- Rubinow, I. M.** 1908. *Russia's Wheat Trade.* Washington, D.C.: U.S. Department of Agriculture.
- Rybalkin, P. N.** 2001. *Pshebntsa i tritikale: Materialy nauchno-prakticheskoi konferentsii "Zelenaia revoliutsiia P.P. Luk'ianenko"* [*Wheat and Triticale: Materials of Scientific-Practical Conference "Green Revolution of P. P. Luk'ianenko"*]. May 28–30. Krasnodar: Sovetskaia Kuban'.
- Rykov, Alexei, ed.** 1925. *V bor'be s zasukhoi i golodom.* Moscow: Gosudarstvennoe izdatel'stvo.
- Sel'skoe.** 1936. *Sel'skoe khoziaistvo SSSR: Ezhegodnik 1935.* Moscow: Sel'khozgiz.
- Sharova, P. N., et al., eds.** 1957. *Kollektivizatsiia sel'skogo khoziaistva: Vazhneishe postanovleniia kommunisticheskoi partii i sovet'skogo pravitel'stva 1927–1935.* Moscow: Nauka.
- Slyn'ko, I. I.** 1961. *Sotsialistychna perebudova i technichna rekonstruktsiia sil'skoho hospodarstva Ukraïny (1927–1932 rr.).* Kyïv: Akademii Nauk Ukraïns'koi RSR.
- Stalin, Joseph.** *Works.* 1955. Moscow: Gosizdat.
- Tan, Graham.** 2000. Transformation versus Tradition: Agrarian Policy and Government-Peasant Relations in Right-Bank Ukraine 1920–1923. *Europe-Asia Studies* 52(5): 915–937.
- Tauger, Mark B.** 1991a. *Commune to Kolkoz: Soviet Collectivization and the Transformation of Communal Peasant Farming, 1930–1941.* Ph.D. diss., Department of History, University of California at Los Angeles.
- Tauger, Mark B.** 1991b. The 1932 Harvest and the Famine of 1933. *Slavic Review* 50(1): 70–89.
- Tauger, Mark B.** 2001a. Grain Crisis or Famine? The Ukrainian State Commission for Aid to Crop-Failure Victims and the Ukrainian Famine of 1928–1929. In *Provincial Landscapes: Local Dimensions of Soviet Power, 1917–1953*, ed. Donald Raleigh, 146–170. Pittsburgh: University of Pittsburgh Press.
- Tauger, Mark B.** 2001b. *Natural Disaster and Human Actions in the Soviet Famine of 1931–1933.* Carl Beck Papers in Russian & East European Studies, no. 1506. Pittsburgh: University of Pittsburgh.

- Tauger, Mark B.** 2001c. Statistical Falsification in the Soviet Union: A Comparative Study of Projections, Biases, and Trust. *Donald Treadgold Papers in Russian, East European, and Central Asian Studies* 34. University of Washington (Seattle).
- Tauger, Mark B.** 2004. Review of *Sotsializatsiia zemli v Rossii* by L. N. Litoshenko. *Russian Review* 63(2): 335–338.
- Tauger, Mark B.** 2005. Soviet Peasants and Collectivization, 1930–1939: Resistance and Adaptation. In *Rural Adaptation in Russia*, ed. Stephen Wegren, 65–94. London: Routledge.
- Tauger, Mark B.** 2006a. Arguing from Errors: On Certain Issues in Robert Davies’ and Stephen Wheatcroft’s Analysis of the 1932 Soviet Grain Harvest and the Great Soviet Famine of 1931–1933. *Europe-Asia Studies* 58(6): 973–984.
- Tauger, Mark B.** 2006b. Stalin, Soviet Agriculture, and Collectivization. In *Food and Conflict in Europe in the Age of the Two World Wars*, eds. Frank Trentmann and Fleming Just, 109–142. New York: Palgrave Macmillan.
- Tauger, Mark B.** 2017a. Pavel Panteleimonovich Luk’ianenko and the Origins of the Soviet Green Revolution. In *The Lysenko Controversy as a Global Phenomenon*, eds. William DeJong-Lambert and Nikolai Kremontsov, vol. 1, 97–128. London: Palgrave Macmillan.
- Tauger, Mark B.** 2017b. La famine soviétique “inconnue” de 1924–1925. In *Famine et transformation agricole en URSS*, 27–74. Paris: Editions Delga.
- Tauger, Mark B.** 2018. Review of *Red Famine* by Anne Applebaum. History News Network (George Washington University, Washington, D.C.). [Link](#)
- Tauger, Mark B.** 2023. Nazareno Strampelli and the First Green Revolution. *Journal of International Development*, April 12. [Link](#)
- Viola, Lynne.** 1996. *Peasant Rebels Under Stalin. Collectivization and the Culture of Peasant Resistance*. New York: Oxford University Press.
- VKP(b).** 1932. *XVII Konferentsiia Vsesoiuznoi Kommunisticheskoi Partii (b): Stenograficheskii otchet*. Moscow: Partiinoe izdatel’stvo.
- Volkov, A. N., A. O. Bertel’e, and M. F. Gladkii,** eds. 1935. *Bor’ba s sornoi rastitel’nost’iu. Materialy vtorogo vsesoiuznogo soveshchaniia po bor’be s sornakami pri VIZR v dekabre 1933 goda*. Moscow and Leningrad: Gosudarsvennoe izdatel’stvo kolkhoznoi i sovkhoznoi literatury.
- Vorob’ev, S. O.** 1929. *Katastroficheskaia gibel’ ozimnykh posevov na Ukraine 1927–1928* [Catastrophic Failure of Winter Sowings in Ukraine]. Khar’kov: Government of Ukraine.
- Wädekin, Karl-Eugen.** 1973. *The Private Sector in Soviet Agriculture*. Berkeley, Cal.: University of California Press.
- Whitman, John T.** 1956. The Kolkhoz Market. *Soviet Studies* 7(4): 384–408.
- WSSA (Weed Science Society of America).** 2016. WSSA Fact Sheet About Weed Seeds and Their Longevity. WSSA.net. [Link](#)
- Zvezdin, K. K.** 1968. Materialy obsledovanie denezhnykh dokhodov i raskhodov sel’skogo naseleniia v 1931–1932 gg. In *Istochnikovedenie istorii Sovetskogo obshchestva*, vol. 2, 319–338. Moscow: Nauka.



## About the Author



**Mark Tauger** grew up in Southern California. He earned a BA and MA in historical musicology at UCLA, then accepted a major fellowship there to do history, focusing on the history of agricultural development in the USSR. After he finished his Ph.D. on collective farms, he became a professor at West Virginia University, and a fellow at the Institute for Advanced Study at Princeton. He has written and published extensively on famines, agriculture, and agricultural history. His email address is [mtauger@wvu.edu](mailto:mtauger@wvu.edu).

[Natalya Naumenko's reply to this article](#)

[Go to archive of Comments section](#)

[Go to September 2023 issue](#)





# Response to Professor Tauger's Comments

Natalya Naumenko<sup>1</sup>

[LINK TO ABSTRACT](#)

First, I want to emphasize that this is a response *only* to the comments on the published 2021 *Journal of Economic History* article (Naumenko 2021). Comments on the earlier 2018 working paper version of this article will just confuse the readers. Similarly, I will not reply to the comments on Markevich, Naumenko, and Qian (2023); it is still a working paper. My co-authors, Andrei Markevich and Nancy Qian, and I thank Professor Tauger for the interest in our work, his comments on the working paper will help us improve the draft.

Before replying, I will quickly describe the structure of my 2021 JEH paper. It is organized into five sections: (1) Introduction, (2) Background, (3) Data, (4) Results, and (5) Conclusion. The Results section is further subdivided into (a) “Weather and grain accounting” (hereafter “Weather” for brevity), (b) “Government policies,” (c) “Mechanisms: Why Collectivization Increased Mortality,” and (d) “Ethnic Composition and Mortality.”<sup>2</sup> An Online Appendix provides additional results referred to in the paper. All data and code used to produce tables and figures in the paper and in the appendix are published in the accompanying replication files.<sup>3</sup> This organization is standard in the field of economics but might

---

1. George Mason University, Fairfax, VA 22030.

2. Per JEH style guidelines, sections and subsections are not numbered, but I add numbers here for the ease of reference.

3. As is customary, an Online Appendix is posted on the author's website: [natalyanaumenko.com](https://natalyanaumenko.com); it can also be downloaded directly [here](#). Replication files are uploaded to the journal's depository and are linked in the paper, I reproduce the reference here: Naumenko, Natalya. “Replication: The Political Economy of Famine: The Ukrainian Famine of 1933.” Ann Arbor, MI: Inter-university Consortium for Political and Social Research (distributor), 2020-12-8 ([link](#)). It is also not uncommon for interested readers to ask the author directly for appendices or replication files.

be unfamiliar to historians.

The Introduction section states that the 1933 famine in Ukraine is an important topic for economists to study, summarizes the main findings, and places the paper in the context of broader literature. The Background section is an abbreviated and simplified description of the events that led to the famine; it doesn't present any new findings but summarizes the history for readers unfamiliar with the context. The Data section describes the main data sources and the structure of the data used in subsequent analysis. The Results are the meat of the paper; they can and should be discussed and debated. The Conclusion restates the main findings and suggests questions for future research. Professor Tauger mostly criticizes the sections (1) Introduction, (2) Background, and (4a) Weather, dismissing the regression analysis in the rest of the Results section: "all of those numbers are statistical probabilities" (Tauger 2023, 289). I will present Professor Tauger's arguments and my responses when discussing each of the sections.

## Comments on Introduction and Background

Criticism of the Introduction ("Descriptions of Ukraine as the 'breadbasket' of Russia have long been a conventional assumption that reflects Ukrainian nationalist views," Tauger 2023, 257) doesn't change the findings of the paper. I thank Professor Tauger for pointing out inaccuracies and oversimplifications in the Background section; it could certainly have been written better. In future work I will be more nuanced in describing the War Communism era, trading of food during the early years of the collectivization drive, and organization of collective farms. Nevertheless, the Background section doesn't present any new findings and doesn't affect the conclusions of the paper. At the same time, I would like to push back on some of Professor Tauger's remarks.

First, Professor Tauger argues that even during the early months of the collectivization drive Soviet officials understood the importance of food trade. Some officials may have understood it, but at the same time, trade was in legal limbo until the 'neo-NEP' reforms of May 1932. For example, during the All-Union Grain Conference in Moscow in June 1931, a 1930 slogan was reiterated again: "Not one kilogram of kolkhoz grain and the grain of the collective farmer to the private market for speculation" (Davies and Wheatcroft 2004, 82). Even when collective farm markets were officially legalized again in May 1932, peasants were allowed to sell their produce there only after the whole region where they lived fulfilled its compulsory grain procurement quota (*ibid.*, 289).

Second, Professor Tauger notes that after the publication of Stalin's "Dizzy with Success" article in early March 1930 most of the collective farms took the

form of *artely* (allowing private gardens and some livestock), not *kommuny* (aiming to eliminate all private production of food). It is true. However, if the remaining private gardens and livestock were sufficient for subsistence, there would have been no famine.

Next, Professor Tauger notes that “When reports of these and other actions [dekulakization, closing of peasant markets, organizing peasants into *kommuny* instead of *artely*] reached top Soviet authorities, Stalin called off the first collectivization campaign on 2 March 1930 and criticized those actions as ‘excesses’” (Tauger 2023, 276). It is important to remember that Stalin was directly behind these ‘excesses.’ For example, here’s what Professor Kotkin says: “Suddenly, however, in a speech on the last day (December 27, 1929) of a weeklong Congress of Agrarians-Marxists, Stalin preempted the commission, thundering in words Pravda carried two days later that ‘we have gone over from a policy of limiting the exploiting tendencies of the kulak to a policy of eliminating the kulaks as a class.’ ... Stalin also used his pencil to hand victory to the more rabid members of the Yakovlev commission: the partially socialized *artels* were no longer to be allowed as the main form of collectives indefinitely, but would be superseded by a leap to the ‘higher-form’ *kommunas*. Stalin also crossed out mention of farmers retaining minor implements, chickens, or a milk cow and wrote in that collectivization was to be completed in just one to two years (depending on region), using dekulakization. All this became a politburo resolution approved on January 5, 1930” (Kotkin 2017, 35–36). With the “Dizzy with Success” publication Stalin threw his subordinates under the bus, blaming them for his own excesses.

Finally, and most importantly, I find Professor Tauger’s views on collectivization puzzling. He notes that:

These practices [farming under peasant commune] indicated that in many ways peasant villages already anticipated collective farms. This was one of the reasons why resistance to collectivization was relatively limited, and many peasants willingly formed or joined *kolkhozy*. (Tauger 2023, 287)

and

Collectivization did not eliminate incentives, because peasants earned labor days for their work and those who did more work would receive higher incomes in money and kind. (Tauger 2023, 271)

The commune coordinated harvesting and many other agricultural activities, but peasant households were still independent production units working their strips of land on the communal fields. This was not the case in the collective farms. While many researchers argue that peasant communes may have hindered economic

development (e.g., Castañeda Dower and Markevich 2018), relative to collective farms, communes provided more economic freedom. It is a consensus among most respected historians that collectivization did disrupt incentives and I am puzzled by Professor Tauger's views. Here are a few quotes from historians who can hardly be accused of spreading 'Ukrainian propaganda:'

A study of collective farms in the Urals in late 1930 reported that in all cases labor discipline was lower on the collective farm than it had been under communal land tenure when peasants worked their own allotments of land. The study concluded that there was no stimulus for improved labor productivity. The lack of material incentives to foster labor discipline was and continued to be throughout Soviet history a central reason for the low productivity of collective farmers. (Viola 1996, 212)

and

But the arrangements for remunerating collective work did not offer adequate incentives to replace the stimulus provided to the individual peasants by the market, and by the need to produce food for their own consumption. The ingenious system of labour days successfully adjusted the level of payment of the peasants for their work before and during the harvest to the size of the income available for distribution after the harvest. But it proved very difficult to design work norms suitable for the immense variety of jobs and territory in the USSR, and still more difficult to penalize shoddy, and reward conscientious, work. In the autumn of 1930, the novel labour-day system was not yet much used in practice, and all these weaknesses seemed to be temporary administrative difficulties which would be overcome with experience; the authorities were confident that the system would work smoothly once it had fully replaced payment per eater. But in the ensuing decades these temporary administrative difficulties permanently haunted the kolkhoz. (Davies 1980, 168)

and

The state continued to compel both collective farms and individual households to surrender very large quantities of grain and other products, for a purely nominal payment, and offered virtually no economic inducement to the peasants to work on collective land. (Davies, Harrison, and Wheatcroft 1994, 16)

To sum up, I wish I wrote the Background section in a more nuanced and detailed way, but some of Professor Tauger's comments show deep misunderstanding of the nature of the comprehensive collectivization campaign, collective

farms, and of Stalin's role in the unfolding crisis. It is indisputable that collectivization disrupted agricultural production, and that Stalin was directly responsible for the early 'excesses.' And, most importantly, none of the above changes the findings and conclusions of my paper.

## Comments on Weather

The Weather section, first, presents rural grain retention (production minus procurement) calculated from official sources (Naumenko 2021, Figure 4). The section argues that the officially reported grain retention was too high in 1932, inconsistent with severe famine after the 1932 harvest. For illustration, it plots corrected retention figures based on corrected harvest estimates from Tauger (2001) and Robert Davies and Stephen Wheatcroft (2004) for the whole Soviet Union, and on my calculations for separate republics. Professor Tauger complains that I did not use his corrected figures for separate republics from Tauger (1991), only the figure for the whole Soviet Union; I wish I did, but I misread his 1991 publication thinking that he presented only estimates of the collective farms harvest, not total harvest. Not willing to impute numbers that he may or may not agree with, I instead cited him and explicitly followed his method in calculating corrected harvest by republic. Regardless, in complete agreement with Professor Tauger, the Weather section states that the official grain production figures were too high. Next, Professor Tauger argues that the procurement figure I use (4.2 million tons of grain procured from Ukraine in 1932) is planned procurement and not actual procurement. This is puzzling. Appendix Table E2 in Naumenko (2021) presents all numbers used to produce grain retention figures in the main text; notes to this table list the sources I used. In particular, 1932 procurement is from Table 33 in *Komitet po zagotovkam S.-Kh produktov pri SNK SSSR (1934, 19)*: total grain procurement from the 1932 harvest, including milling levy, by region. This is actual, not planned procurement, and Professor Tauger criticizes me for something that is not there in my text. That Table 33 also reports the total 1932 grain procurement for the whole Soviet Union (18.5 million tons), and this number is the same as the total 1932 procurement presented in Table 1 of Davies and Wheatcroft (2004, 449). My procurement numbers are also the same as presented in Table 14d of Davies and Wheatcroft (2004, 470—Grain collections by area supplying grain).<sup>4</sup>

---

4. Professor Tauger also claims that 5.7 million tons of procured grain were returned to the Soviet countryside in 1933 (he doesn't break this number into regions). If true, this would only strengthen the statement that the official 1932 harvest was too high: if some grain was returned, it would increase rural retention, making official numbers further incompatible with the subsequent famine.

To explain the presumably catastrophic harvest, Professor Tauger next shows that in Ukraine rainfall from April to the first half of June was below average in 1931 and above average in 1932 (Tauger 2023, Table 3). Figure 5 in Naumenko (2021) shows exactly that. However, rainfall for these months was even more below average in 1934, without catastrophic famine in 1935. Rainfall was also very low in May 1936, again without famine in 1937. Similarly, rainfall was above average in 1933 (higher than in 1932 in May, slightly lower in April and June), and Professor Tauger agrees that the 1933 harvest was good. Thus, picking just one factor that deviated from the average as the cause of harvest failure and famine is wrong. Many years had larger deviations in rainfall or temperature without millions dying. If I were to predict catastrophe on too-dry or too-wet spring months in Ukraine, I would not pick 1931 and 1932 as the worst years: other years had larger deviations in temperature and rainfall.

Moreover, Appendix Figure B8 in Naumenko (2021) uses available daily weather data to plot extreme weather days—share of days in April–July each year when temperature or rainfall was more than one standard deviation away from the day’s 1910–1950 average. It shows that 1931 and 1932 had no more extreme weather days than other years. This is why it is important to not just look at one selected factor (April–June rainfall), but at weather overall.

As Professor Tauger himself says, “the weather processes in this region were never simple, but involved many complex changes that cannot be reduced simply to selected temperature and precipitation readings” (Tauger 2023, 267). To capture this complexity, I then use 1901–1915 weather (fall, winter, spring, and summer temperature and precipitation, their pairwise interactions and square terms), soil quality, and area—all parameters that did not change rapidly under the influence of collectivization—into harvest and calculate grain production function. I then use this grain production function to predict harvest from 1916 to 1936 (Naumenko 2021, Figure 6). I show that the weather-predicted harvest is similar to the officially reported (too high) harvest in Ukraine. This does not mean that the officially reported harvest is true. Rather, the weather does not predict catastrophic harvest failure in 1931 and 1932. If the harvest was low, it was not because of the weather.

Professor Tauger argues that this exercise is not valid because the period from 1901 to 1915 saw many harvest failures and I ignore them. This is a misunderstanding on Professor Tauger’s side. I never claimed (nor thought) that 1901–1915 were years of only good weather and good harvests. On the contrary, having weather-driven harvest failures during 1901–1915 helps predicting weather-driven harvest failures in the early 1930s. Indeed, if the weather was only good from 1901 to 1915, a grain production function would not capture bad weather when it presumably occurred in the early 1930s. Thus, Professor Tauger’s argument that harvest failures occurred before only makes a grain prediction exercise stronger.

Next, Professor Tauger restates his earlier arguments that pests and grain diseases led to harvest failure and were the main cause of the famine. First, while historians generally agree that the official harvest figures were too high, few (except Professor Tauger) cite grain diseases and pests as the main cause; for example, Davies and Wheatcroft (2004) blame the presumed drought of 1931 and unfavorable weather in 1932. Second, to the extent that weather facilitated the spread of pests and grain diseases, the grain production function should capture it.<sup>5</sup> Third, Appendix Section B5 in Naumenko (2021) analyzes available published collections of documents prepared by international teams of respected historians (Berelowitch and Danilov 2000–2012; Danilov et al. 1999–2006). Neither collection indicates more often than usual mentions of pests and grain diseases in 1932 (Figures B11 and B12 in Naumenko 2021). Professor Tauger is correct that still-unpublished documents might show that there were more than usual number of pests or grain diseases. However, even if this were true, to explain why famine mortality is so strongly correlated with collectivization rate (as the estimates in my paper show) it will have to be the case that the geographic distribution of pests and grain diseases is correlated with collectivization—for example, if poor agricultural practices, disorganization, and low effort increased their spread. And in that case, the main cause of famine would be collectivization again.

Finally, here is a broad-picture argument why weather or grain diseases and pests cannot be the main cause of the 1933 famine. The last large famine under the tsarist regime, the 1891–1892 famine caused by drought and accompanied by a cholera epidemic, killed roughly half a million people (Wheatcroft 1992, 53–59). From 1892 till 1913 agricultural productivity was slowly developing.<sup>6</sup> By 1928, the Soviet economy roughly recovered to the pre-WWI level (Markevich and Harrison 2011). So let's say agricultural technology roughly returned to the 1913 level, or even to the 1891 level. How is it possible then that in 1933 drought, or mice, locusts, and grain rust, killed five to ten million people—an order of magnitude more than in 1892? Whatever unfavorable environmental conditions occurred, either they were of Biblical proportions (and should be easily spotted in the data, which they are not) or they were exacerbated by the government policies implemented at the time—by collectivization. Weather is almost never perfect, but the famine of that magnitude only occurred after the comprehensive collectivization

---

5. For example, Professor Tauger cites archival sources: "Infestation of ergot this year evidently has exceptional character and should be attributed to favorable meteorological conditions (warmth and moisture), and also to poor agricultural measures" (quoted in Tauger 2023, 269). Thus, weather might have been driving the spread of grain diseases, and therefore the impact of grain diseases should be captured by the weather-driven grain production function.

6. For example, Obukhov (1927) shows that, while volatile, on average harvest was increasing from 1883 to 1914 in the 50 provinces of European Russia, including future Soviet republic of Ukraine.



begun. And therefore the main cause of the famine must be government policies.

## Comments on Government Policies

It is regrettable that excepting the Weather section, Professor Tauger ignores most of the Results section. The gulf in training and methods between quantitative fields like political science and economics and qualitative fields like history is unfortunate as it prevents a more informed debate. Below I'll try to briefly describe the main results in my paper as they relate to Professor Tauger's explanation of the famine.

Table 1 in Naumenko (2021) presents a regression analysis of the relationship between collectivization rate and famine mortality. For brevity, I will only discuss Panel C Column 4, as this is the main estimate in the paper. It uses a sample of regions (*okruha*, January 1, 1925, administrative borders). Essentially, it compares how mortality changed from 1927–1928 to 1933 in regions that had similar pre-famine characteristics ("controls") but different collectivization rates.<sup>7</sup> It finds that, on average, if you compare two regions with similar pre-famine characteristics, one with zero collectivization rate and another with a 100 percent collectivization rate, the more collectivized region's 1933 mortality rate increases by 58 per thousand relative to its 1927–1928 mortality rate. Moreover, this increase in mortality is not only large in magnitude, but is highly statistically significant with p-value below 1 percent (in fact, that estimate's p-value is 0.2 percent).

Here we must talk about statistical probabilities and p-values. Imagine you had a coin and you wanted to test whether the coin was fair, that is, whether when you tossed it getting heads or tails was equally likely. Imagine you're tossing this coin and keep getting tails—perhaps, six times in a row. If the coin was in fact fair, the probability of getting six tails in a row (p-value) is approximately 1.6 percent—physically possible but not remotely likely. A reasonable person, after observing six tails in a row, would reject the assumption that the coin was fair. Similarly, the regressions presented in Table 1 show that if collectivization had zero impact on mortality, the probability of observing an increase in mortality of 58 per thousand is below 1 percent.<sup>8</sup> Therefore we reject the assumption that

---

7. The full list of controls is: number of Group A workers per capita in 1930, distance to a railroad, de-meaned June 1931 temperature, de-meaned June 1932 precipitation, average grain production per capita as of 1925, livestock per capita in 1925, value of agricultural equipment per capita in 1925, urbanization rate according to the 1926 census, Polissia region indicator, rural literacy rate according to the 1926 census, and population density as of 1926 census.

8. In fact, the p-value of the estimate in Table 1 Panel C Column 4 is 0.2 percent—the equivalent of tossing a fair coin and getting heads nine times in a row.

collectivization had no effect on mortality. Simply put, this is a too large increase in mortality observed too often in more-collectivized regions to be just a coincidence. Many assumptions go into calculating this p-value and they can and should be debated. But they can't be just dismissed and ignored.

Finally, baseline estimates use the 1930 collectivization rate. Professor Tauger complains that it is not clear when collectivization was measured, before or after the “Dizzy with Success” publication; the paper clearly states that this is May 1930 collectivization (Naumenko 2021, 13). In addition, Appendix Table C3 in Naumenko (2021) presents estimates that use 1932 collectivization rate and the results are even stronger—larger in magnitude and as statistically significant; the main text presents conservative estimates.

## Conclusion

I hope this response clarifies some of the questions and misunderstandings raised by my paper. I thank Professor Tauger for the productive and insightful discussion of my work.

## References

- Berelowitch, Alexis, and Viktor Danilov**, eds. 2000–2012. *Les campagnes Soviétiques vues par la Tchèque*—O.G.P.U.—N.K.V.D. 1918–1939: *Recueil de documents et matériels* [Sovetskaya derevnya glazami VChK—OGPU—NKVD. 1918–1939 gg. *Dokumenty i materialy*], 4 vols. Moscow: Rosspen.
- Castañeda Dower, Paul, and Andrei Markevich**. 2018. The Stolypin Reform and Agricultural Productivity in Late Imperial Russia. *European Review of Economic History* 23: 241–267.
- Danilov, Viktor, R. Manning, Lynne Viola, Robert W. Davies, Ha Yong-Chool, R. Jonson, V. Kozlov, A. Sakharov, V. Vinogradov, and Stephen G. Wheatcroft**, eds. 1999–2006. *The Tragedy of the Russian Village: Collectivization and Dekulakization—Documents and Materials 1927–1939* [Tragediya Sovetskoy derevni. Kollektivizatsiya i raskulachivaniye. *Dokumenty i materialy 1927–1939*], 5 vols. Moscow: Rosspen.
- Davies, Robert W.** 1980. *The Soviet Collective Farm 1929–1930 (The Industrialization of Soviet Russia*, vol. 2). Basingstoke, UK: Palgrave Macmillan.
- Davies, Robert W., Mark Harrison, and Stephen G. Wheatcroft**, eds. 1994. *The Economic Transformation of the Soviet Union, 1913–1945*. Cambridge, UK: Cambridge University Press.
- Davies, Robert W., and Stephen G. Wheatcroft**. 2004. *The Years of Hunger: Soviet Agriculture, 1931–1933 (The Industrialization of Soviet Russia*, vol. 5). Basingstoke, UK: Palgrave Macmillan.

- Komitets po zagotovkam S.-Kh produktov pri SNK SSSR.** 1934. *Yezhegodnik khlebooborota za 1931–32, 1932–33 i predvaritelnyye itogi zagotovok 1933 g.* Moscow.
- Kotkin, Stephen.** 2017. *Stalin, vol. II: Waiting for Hitler, 1929–1941.* New York: Penguin Press.
- Markevich, Andrei, and Mark Harrison.** 2011. Great War, Civil War, and Recovery: Russia's National Income, 1913 to 1928. *Journal of Economic History* 71: 672–703.
- Markevich, Andrei, Natalya Naumenko, and Nancy Qian.** 2023. The Causes of Ukrainian Famine Mortality, 1932–33. *NBER Working Paper* 29089. National Bureau of Economic Research (Cambridge, Mass.).
- Naumenko, Natalya.** 2021. The Political Economy of Famine: The Ukrainian Famine of 1933. *Journal of Economic History* 81: 156–197.
- Obukhov, Vladimir Mikhaylovich.** 1927. *Dvizheniye urozhayev zernovykh kul'tur v Yevropeyskoy Rossii v period 1883–1915 g.g.* Moscow: Rossiyskaya assotsiatsiya nauchno-issledovatel'skikh institutov obshchestvennykh nauk.
- Tauger, Mark B.** 1991. The 1932 Harvest and the Famine of 1933. *Slavic Review* 50: 70–89.
- Tauger, Mark B.** 2001. Natural Disaster and Human Actions in the Soviet Famine of 1931–1933. *Carl Beck Papers in Russian and East European Studies* 1506. University of Pittsburgh (Pittsburgh, Pa.). [Link](#)
- Tauger, Mark B.** 2023. The Environmental Economy of the Soviet Famine in Ukraine in 1933: A Critique of Several Papers by Natalya Naumenko. *Econ Journal Watch* 20(2): 254–303. [Link](#)
- Viola, Lynne.** 1996. *Peasant Rebels Under Stalin.* Oxford, UK: Oxford University Press.
- Wheatcroft, Stephen G.** 1992. The 1891–92 Famine in Russia: Towards a More Detailed Analysis of Its Scale and Demographic Significance. In *Economy and Society in Russia and the Soviet Union, 1860–1930: Essays for Olga Crisp*, eds. Linda Edmondson and Peter Waldron. Basingstoke, UK: Macmillan.

## About the Author



**Natalya Naumenko** is an assistant professor at the Department of Economics at George Mason University. She graduated from Northwestern University in 2018. Before joining Mason, she did a postdoc at the Political Theory Project at Brown University. Her email address is [nnaumenk@gmu.edu](mailto:nnaumenk@gmu.edu).

[Go to archive of Comments section](#)  
[Go to September 2023 issue](#)



# The Limitations of Growth-Optimal Approaches to Decision Making Under Uncertainty

Matthew C. Ford<sup>1</sup> and John A. Kay<sup>2</sup>

[LINK TO ABSTRACT](#)

A fundamental problem in the analysis of choice under uncertainty is that it is not obvious what constitutes optimal behaviour for an individual. Within economics assumptions have traditionally been made to justify various approaches: for example Expected Utility Theory (EUT) is justified by John von Neumann and Oskar Morgenstern's axioms about the preferences individuals have over lotteries. Weakening the axioms required by a model of decision making has clear attractions, and optimal growth theory represents a notable attempt to do so. John Kelly (1956, 925) showed that when returns compound, there is an asymptotically-optimal fraction of your wealth to offer in any given round of a repeated gamble, and that a gambler following this method "maximizes the expected value of the logarithm of his capital." This analysis has subsequently been extended to portfolio theory, in particular by Henry Latané. It has long been recognised, however, that this approach in general conflicts with EUT and requires different assumptions about what decision makers are aiming to achieve.

More recently, Ole Peters, who is a physicist, and collaborators, including the late Murray Gell-Mann, have drawn on ergodic theory to argue that optimal growth provides a general foundation for theories of decision making in economics (Peters

---

1. St. John's College, Oxford OX1 3JP, UK. We would like to thank Eric Briys for helpful feedback on an earlier draft. Some of the ideas in this paper benefitted from the discussion of Matthew's paper on 'What Work is Ergodicity Doing in Economics?' at the LSE's Hayek Programme in Economics and Liberal Political Economy seminar series.

2. St. John's College, Oxford OX1 3JP, UK.

and Gell-Mann 2016; Peters 2019). ‘Ergodicity economics’ (EE) focusses on time averages as an alternative to the use of the expectation operator in mainstream economic theories of decision making under uncertainty, in particular EUT. EE transforms accumulation dynamics into ergodic processes (for which the expected value and time average coincide).<sup>3</sup> EE then bases its predictions on the limit of the time average of the growth rate—as with other growth-optimal theories, the justification is that, in the limit, the process with the highest time-average growth rate will result in the most wealth (see Peters and Gell-Mann 2016, eqs. 2 and 5). EE is therefore distinct from other growth-optimal approaches for two reasons: it claims to provide an objective justification for decision making without the need to refer to individual psychology; and its proponents claim that it makes more metaphysical sense than alternative models of decision making.

This work on decision theory, and the associated criticisms of utility theories such as EUT, form the core of the ‘ergodicity economics’ research programme. The claims are forceful: readers are told that EE “resolve[s] the fundamental problem of decision theory, therefore game theory, and asset pricing” (Peters and Gell-Mann 2016, 8). We show in this critique that the model EE proposes is unsatisfactory in several respects and does not represent an improvement on utility theories. Furthermore, we show that their criticisms of utility theories are misguided and based on a misunderstanding of how they work. In particular, many of the distinctions drawn between utility theories and EE arise because, when analysing the same problem, Peters and Gell-Mann use different information sets, as well as different models. We argue that there is no justification for using different information sets and that EUT looks much more credible when it is given informational parity. As a result, we reject their claim to have resolved any fundamental problem, and we are sceptical of EE’s value for economists.

Most of the other work in the ‘ergodicity economics’ programme is based on the same fundamental approach and is therefore vulnerable to our critique. For example Peters (2011) claims to solve the St. Petersburg Paradox by altering the problem so that the individual faces multiple opportunities to participate in the gamble. His solution rests on the individual maximising the growth rate of their wealth and does not address what seems to us to be the core issue, which is that, in a world with truly unbounded potential payouts in terms of subjective satisfaction, it is unclear that the ‘paradoxical’ intuition we display—that we would not, in the real world, pay very much for St. Petersburg-type gambles—applies. It is therefore not a convincing demonstration of the limitations of EUT, or of the merits of

---

3. A “dynamic” is specified by the nature of the interaction of the random variable with the state variable, i.e., wealth. For example if the realisation of the random variable is added to wealth the dynamic is additive; if wealth is multiplied by the realisation it is multiplicative, etc.

EE. Another paper on insurance, co-written by Peters with Alexander Adamou (2017), similarly assumes that individuals wish to maximise their growth rate and contrasts this to an analysis based on risk aversion. They neglect to include the same information set in their EUT calculations (i.e., that individuals know that they will face similar situations repeatedly) and so mischaracterise that approach.

It is worth noting that ‘ergodicity economics’ has also been applied to questions about inequality, for example by Adamou and Peters (2016). This work is, conceptually, entirely distinct from the work on decision theory (although the mathematical background is similar), and it is not susceptible to the criticisms presented here. Danny Dorling (2016) has argued persuasively that the model of inequality they propose does not explain many of the data’s most interesting features, but this is an empirical question and it is possible that reformulations of the model may be more successful.

The key insight of optimal growth approaches is also their central flaw: they posit a theory of decision making under uncertainty *without* uncertainty. Paul Samuelson (1971) showed that for a certain class of utility functions the limit was an inappropriate guide for any finite process, but this limitation was also implicitly mentioned by Kelly (1956, 920): “It is surely true that if the game were to be stopped after  $N$  bets the answer to this question would depend on the relative values (to the gambler) of being broke or possessing a fortune.” However the situation is worse than this: optimal growth approaches *necessarily* violate EUT, in particular the axiom of completeness. In other words, they lead to inconsistent decision making.

## Terminology

The terminology used by different authors varies. Von Neumann and Morgenstern (1955) refer to a random drawing of one of a set of payoffs, weighted by a defined probability distribution, as a “lottery.”<sup>4</sup> When several lotteries are combined—for example if the payoff of a lottery is participation in another lottery—the overall structure is referred to as a “compound lottery.” In contrast, Peters (2019) refers to “gambles,” which are series of random drawings applied in sequence to a transformed variable, typically wealth.

Throughout this paper we will refer to individual drawings as lotteries and a series of lotteries, each applied to the wealth resulting from the previous lottery, as a gamble:  $G$ . Gambles therefore have a length  $l$ . A gamble of length 1 is just

---

4. Note that lotteries are over overall outcomes, so a coin toss to gain or lose \$100 is a different lottery if you have a starting wealth of \$1,000 or \$10,000.

a lottery. For concision we refer to approaches that result in optimal growth in the limit as ‘growth optimal.’ We note here, however, that for the finite cases we are considering this is a misnomer, as the growth rate is a random variable and a ‘growth optimal’ approach will not always lead to higher growth.

By ‘psychology,’ we mean the subjective valuation of objective experience: for example, different people will value different consumption bundles differently, leading them to make different choices.

## Optimal growth

For any gamble where returns compound total growth  $G$  after  $t$  draws is given by:

$$w_t = w_0 \cdot \prod_{i=1}^t g_i \Rightarrow \frac{w_t}{w_0} = G = \prod_{i=1}^t g_i \quad (1)$$

where  $g_i$  is the realisation of the  $i$ th lottery in the gamble. We can rearrange this to give the per-stage growth rate (a time average)  $\tau_t$ :

$$\tau_t = G^{\frac{1}{t}} = \left( \prod_{i=1}^t g_i \right)^{\frac{1}{t}} \quad (2)$$

If we then take the natural log of both sides, we have:

$$\ln(\tau_t) = \frac{1}{t} \left( \sum_{i=1}^t \ln(g_i) \right) \quad (3)$$

This is the sample average of the gamble transformed by the natural logarithm. If the sequence of growth rates  $g_i$  is stationary and ergodic and  $\mathbb{E}[\ln(g_i)]$  is finite then by the Ergodic Theorem, as  $t \rightarrow \infty$ ,  $\ln(\tau_t) \rightarrow \phi$  almost surely where  $\phi$  is some constant. From this it follows that  $\tau_t \rightarrow e^\phi$  almost surely. The growth rate will therefore be  $e^\phi$  with probability 1.

The time averages used in growth-optimal models correspond to a situation where there is no measurable uncertainty—final wealth will almost always be what the time average predicts. Decisions reduce to choosing the gamble which gives the highest wealth with (effective) certainty. But we do not live in the limit—to paraphrase only slightly the famous dictum of John Maynard Keynes, in the limit we are all dead.<sup>5</sup> Economic agents necessarily have finite lives, and may also be

---

5. Keynes’ original statement was “*In the long run we are all dead*” (Keynes 1924, 80).



interested in outcomes over shorter segments of that, as well as averages over the whole of these lives. What insight does growth optimality offer in these cases?

Peters (2019, eq. 2) describes a lottery (hereafter referred to as ‘Peters’ bet’) with an equal chance of increasing your wealth by 50 percent or decreasing it by 40 percent. If this is applied repeatedly over time to total wealth, it is very likely to leave one with almost nothing—the time-average growth rate is negative—even though the expected gain in each round is positive. If this seems counterintuitive, suppose you start with \$100 and play the bet for two rounds, and win the first round and lose the second. Winning the first round increases your wealth to \$150, but losing the subsequent round decreases it to \$90—less than you started with. However if we take the average of the four equally likely final outcomes—\$36, \$90, \$90, \$225—it is \$110.25. This ‘volatility tax’ is a well-known phenomenon in financial markets and is mechanically incorporated in an EUT analysis of choices as only final outcomes are considered (assuming, as with the simple growth-optimal models discussed here, there is no discounting and no consumption throughout the gamble). However one might expect that, as the gamble’s length goes to infinity, the predictions of EUT would approximate the growth-optimal predictions.

Samuelson (1971)’s contribution is to show that this is not the case for many utility functions. Noting that for any finite gamble the outcome is a random variable, rather than a deterministic number, he shows that an agent with a certain type of utility function—for example a risk-neutral agent described by  $u(w) = w$ —would want to take Peters’ bet regardless of its length (and regardless of whether or not they could decide to stop gambling after each stage). Although a risk-neutral agent taking the gamble will probably end up with almost nothing, for them the chance of a very large payoff will always outweigh that risk. In the limit case those high-value outcomes are suppressed, explaining why EUT might appear to coincide with growth optimality.

## Comparison to expected utility theory

We have seen that growth-optimal models can be justified by a riskless ‘limit’ world, and in such a world preferences over spreads of possible outcomes are unnecessary: the only preference ordering necessary is preferring more to less. In contrast, EUT can incorporate many preferences which give rise to risk aversion (in the sense that the more we have of something already the less value we tend to get from additional units—winning \$1M after going bankrupt makes more of a difference than winning \$1M when one has just become a millionaire). Yet risk-neutral EUT is not growth-optimal. So which preferences over risk do growth-optimal models assume?

Choosing a gamble to maximise the growth rate of wealth looks equivalent to choosing a specific utility function: if the dynamics are multiplicative we maximise the expected value of  $u(w) = \ln(w)$ ; if they are additive we maximise the expected value of  $u(w) = w$ . Peter Carr and Umberto Cherubini (2020) draw on this insight and vary a stochastic clock to show that a variety of utility functions can be justified in this way, and Peters and Adamou (2021) sets out a general way to find a correspondence between utility functions and dynamics (where this exists). However these ‘utility functions’ are artefacts of the transformation made to find the optimal growth rate of the gamble and depend on the dynamics of wealth accumulation: they are not representations of preferences over lotteries.

Given this, Samuelson (1971) can be extended to show that growth-optimal behaviour necessarily violates the axioms of EUT because it is incompatible with any utility function. Any non-linear utility function is incompatible with taking growth-optimal additive gambles: a strictly concave region means that, in that region, an individual would not take an additive gamble with a very slightly positive time average; and in a strictly convex region they would take a gamble with a very slightly negative time average. However a linear utility function means that an individual would accept Peters’ bet, which we have seen EE rejects.

A consequence of growth-optimal approaches contradicting EUT is that they necessarily violate at least one of von Neumann and Morgenstern’s axioms.<sup>6</sup> The axiom violated is completeness: they fail to represent (or, in more normative terms, establish) a preference ordering over many gambles. This occurs in three distinct ways.

Firstly, they are in general unable to rank gambles with different dynamics against each other. For example, how are we to decide between an additive gamble with a time average growth of 5 units per round and a multiplicative gamble with a time average growth of 1 percent of wealth per round?

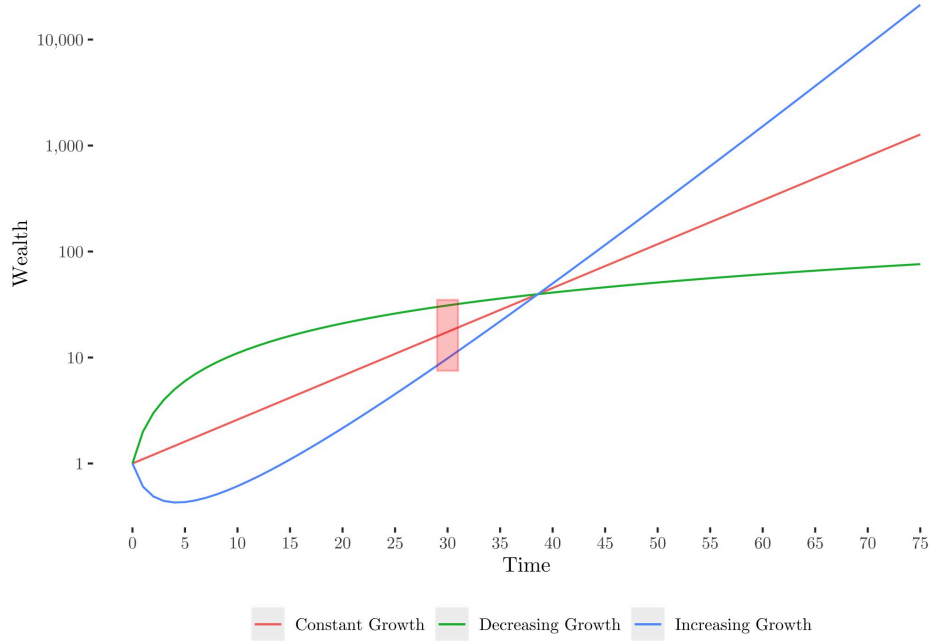
Secondly, all gambles of length 1 can be described with both additive and multiplicative dynamics, potentially giving different preference orderings compared to a ‘gamble’ which leaves wealth unchanged with certainty. For example Peters’ bet, when expressed as an additive gamble, is preferred to this static ‘gamble’; but when expressed as a multiplicative gamble it is not. But when both gambles are of length 1, they are exactly the same. Therefore some gambles must be at least some minimum length—which might vary depending on the gamble—to be comparable.

Finally, some time-varying gambles are not comparable. To show this, sup-

---

6. It is worth noting that Samuelson (1963, 110–111) shows that a system based on choosing the option most likely to leave you better off results in a violation of transitivity, but this is not necessarily the strategy used by EE.

pose every type of gamble has some minimum length  $t_{\min}(G)$  such that, when it is of length  $t \geq t_{\min}(G)$  it is comparable to all other gambles of sufficient length. As we saw above, for at least some gambles this must be greater than 1, and we will see later that we might want it to be very large.



**Figure 1.** Three deterministic gambles

Now consider three ‘deterministic’ gambles, i.e., ones where the change in per-period growth is certain, with multiplicative dynamics, shown in Figure 1.

The red gamble  $G_C$  grows at a constant rate of  $g$  per period. The blue gamble  $G_+$  has a growth rate  $\tau_+(t)$  which is increasing in  $t$  and tends towards  $\alpha g$  where  $\alpha > 1$ . It is defined recursively:

$$w_t = w_{t-1}(\alpha g)^{\left(\frac{t}{t+1}\right)} \quad (4)$$

which gives the general form

$$w_t = w_0(\alpha g)^{t \left(\frac{t!}{(t+1)!}\right)} \quad (5)$$

The green gamble  $G_-$  has a growth rate  $\tau_-(t)$  which is decreasing in  $t$  and tends towards 0. It is defined recursively:

$$w_t = w_{t-1} \left( \frac{t+1}{t} \right) \quad (6)$$

which gives the general form

$$w_t = w_0 \left( \frac{(t+1)!}{t!} \right) \quad (7)$$

Taking the time averages of these gambles, EE gives the ordering  $G_+ > G_C > G_-$ . However as Figure 2 shows for some values of  $t$ ,  $w(G_+(t)) < w(G_C(t)) < w(G_-(t))$ . A natural conclusion is that  $t_{\min}$  for time-varying gambles must be sufficiently high such that there is no constant deterministic gamble  $G'_C$  such that  $G'_C < G_+$  where  $w(G'_C(t_{\min})) > w(G_+(t_{\min}))$ . Otherwise it would be the case that growth-optimal models suggest that people prefer certainly receiving less money to certainly receiving more. However it will always be possible to create such a gamble by setting its growth rate marginally below  $\alpha g$ . Therefore there is no possible value of  $t_{\min}$  at or above which this gamble can be ranked, and so this type of gamble also cannot be compared by these models. As well as revealing a further violation of completeness, this demonstrates why using a limit result for finite problems can be—and in this case is—inappropriate: merely examining the time average growth rate results in a clear ordering with no warning that there is no finite length for which this ordering makes sense.<sup>7</sup>

## Psychology is fundamental to decision making

The contradiction between the predictions of EUT and growth-optimal models reveals that making psychological assumptions is fundamental to theories of decision making under uncertainty: when dealing with uncertain outcomes there cannot be a simple, generally applicable, and completely persuasive rule for what all decision makers should choose. As a result, claims that these approaches are valuable because they do not “appeal to an intangible psychology” (Peters 2019, 1218) become arguments *against* them.

Carr and Cherubini (2020, 3) run up against the same problem: they explicitly attempt “to reconcile Kelly and Samuelson” without appealing to utility. However their findings—interesting though they are—are again based on limit cases which

---

7. One way of avoiding this issue would be to make  $t_{\min}$  a function of both gambles being compared, rather than just an individual gamble. The problem with such an approach is that it would inevitably involve examining both entire distributions of potential outcomes at particular points in time and judging whether one was different enough to the other to justify distinguishing between the gambles. At this point the decision criteria is not really about the time average and is very close to EUT.

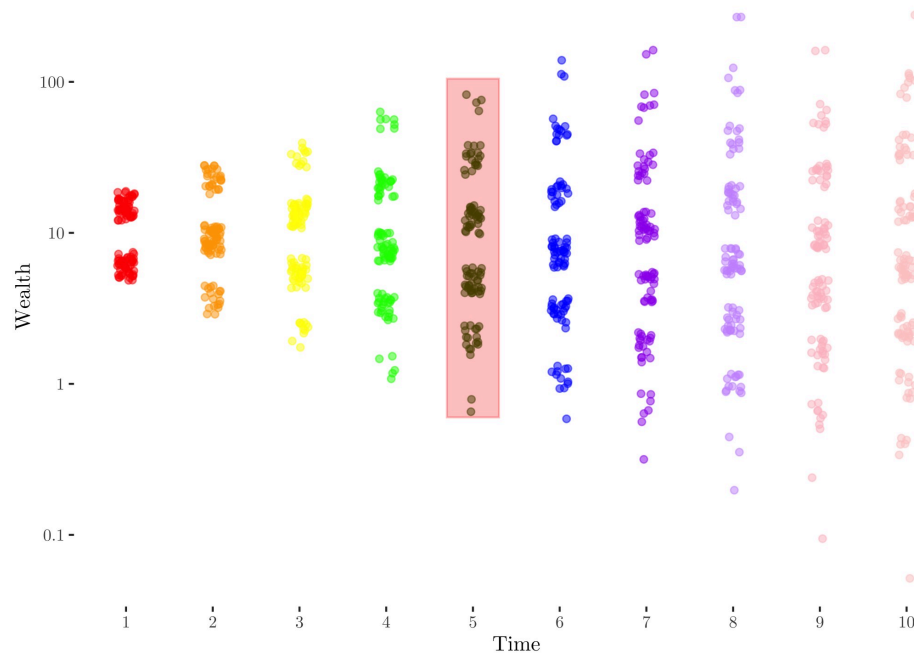
cannot be reached. In the real world, uncertainty of outcome remains, and different people will have different preferences across these. Samuelson and Kelly remain unreconciled.

We have seen that relying on the limit is insufficient as a justification of growth-optimal approaches. Sophisticated theorists have not attempted to claim that it is: for example Latané (1979, 310) notes that “My interest...is in the asymptotic qualities of  $G$  and the measurement of the probability of adverse dominance... It seems to me that this probability of adverse dominance is especially relevant to corporate and other investment decisions and portfolio management where individual subjective utilities of those involved are difficult if not impossible to determine.” As we noted above, more recently a novel justification has been offered by physicists with a background in ergodic theory.

Their claim is that decision theory must consider what actually happens to an individual, not what might happen to them. For this it needs to be an “ergodic observable”, because for such observables the ensemble and time averages coincide (Peters and Gell-Mann 2016, 1).<sup>8</sup> To create this observable we have to understand the dynamics of a gamble over time: “requiring the specification of a dynamic is requiring the admission that we live through time, not in a superverse of parallel worlds with which we can share resources” (ibid., 3). This is contrasted with EUT: “Expected utility theory computes what happens to a loosely specified model of my psychology averaged across a multiverse. But I do not live spread out across a multiverse, let alone harvest the average psychological consequences of the actions of my multiverse clones” (Peters 2019, 1218). EUT therefore stands accused of making two metaphysical errors: ignoring time; and assuming resources can be shared across multiple possible universes. We saw above, however, that it is wrong to say that EUT takes no account of time: time is implicitly incorporated into the potential outcomes explicitly considered, but only the time that will actually elapse during the decision problem. In contrast EE considers all the time that would elapse were the gamble to go on forever. Figure 2 illustrates this difference. (We discuss applying EUT to problems involving time in Appendix A.)

---

8. If we have a probability space  $(\Omega, \mathcal{F}, P)$  and a measure-preserving transformation  $T: \Omega \rightarrow \Omega$ , then we can think of each application of the transformation resulting in a new period in time. If the transformation is ergodic then Birkhoff's Ergodic Theorem states that the time average converges almost surely to the expected value (where it exists).  $T$  is ergodic for the measure  $P$  iff for any  $F \in \mathcal{F}$  such that  $T^{-1}(F) \subset F$  either  $P(F) = 0$  or  $P(F) = 1$ : intuitively, if we take a subset of the sample space resulting from the application of the transformation to some other subset, then the former subset will not be contained in the latter, except in trivial cases.



**Figure 2.** Different perspectives on the same bet, which increases wealth by 50 percent or decreases it by 40 percent, each with equal probability. An EUT analysis takes place at a specific point in time, illustrated here as  $t = 5$ ; an EE analysis includes the infinite unrealised future.

The suggestion that use of the expectation operator invokes multiple worlds is also incorrect. EUT is a *representation* of how people make decisions, not a realistic description. Decision makers do not form an average over outcomes because they think some physical equivalent of the averaging will occur; they can be represented as doing so purely because, if they are described by the axioms in von Neumann and Morgenstern (1955), their choices based on their preferences admit such a representation. In fact EE also uses the expectation operator following an appropriate transformation: this does not mean it is “assuming resources can be shared across multiple possible universes”; just that the expected value happens to be mathematically equivalent to the object it is interested in. Just like in EUT.

It is unclear why, in principle, we would ever want to use any form of summary statistic of the potential outcomes of a decision rather than the entire distribution of outcomes, except to make the problem more tractable: this applies equally to expected values and time averages. EUT is a special case where, given various axioms are met, the use of the expected value of a function happens to encode sufficient information to identify an agent’s preferred choice. It is not a general claim that expected values tell us all we need to know about decision

making. Furthermore, it is true to say that EUT considers multiple potential outcomes, only one of which will occur—but in that it is no different to EE, and we have seen that for real gambles (which necessarily occur in finite time) this fundamental truth cannot be avoided.

We suspect, however, that the limit is the implicit justification for EE, because—in marked contrast to Latané’s work—EE gives no guidance as to when a gamble is long enough for the theory to apply. The following example demonstrates the problems this causes. Consider two gambles on a realistic time scale: in the first, a fair coin will be tossed every month for the next ten years. Each time it lands heads you receive \$2; each time it lands tails you pay \$1. The expected gain is \$0.5 per month. The second is identical, except that if the coin lands heads you receive \$102 and if it lands tails you pay \$100. The expected gain per month is \$1. As this is a case of additive growth, EE predicts individuals will prefer this second gamble to the first. We predict that they would not, because it is much riskier—in general the sort of people who find such gambles attractive have not survived. Furthermore, the two variables which might change our prediction—if the individual is extremely wealthy, and if they are approximately risk-neutral, perhaps because they are a firm with diversified shareholders—are variables in EUT but not EE.

## The purpose of a decision theory

One point worth emphasising is the different purpose of these two decision theories. As we discuss in Ford and Kay (2023), EUT and its intellectual descendants are descriptive, not normative, theories: ‘of’ rather than ‘about’ decision making. In practice the distinction between descriptive and normative theories has been elided by many economists. To understand both the limits and benefits of EUT, it is helpful to distinguish the two.

The axioms of EUT and similar theories can be split into two groups: the first allow preferences to be described sensibly; the second restrict those preferences. For example, in EUT the axioms of completeness and transitivity allow for a description of an agent’s preferences, whereas the axioms of continuity and independence restrict them.

The descriptive axioms ensure that all of the objects in the agent’s choice set can be ordered against each other with the preference relation, and that there will, in any subset of this choice set, be at least one most-favoured option. What this means in practice is that, given any set of choices an agent faces, we always know what they want to do (or the collection of choices they are indifferent between but prefer to all the other choices). There is no need for any type of utility function to



tell us what they will choose. But it is important to note that this gives no insight into how or why the agent has this particular preference ordering. In that sense, it is psychologically naïve: it simply asserts that the agent already knows exactly what they want in any situation.

The restrictive axioms may or may not be plausible descriptions of an agent's preferences, but their purpose is to allow the preference ordering guaranteed by the descriptive axioms to be expressed with a utility function. For example, lexicographic preferences may be complete and transitive but they violate the axiom of continuity: this does not imply anything about the rationality of having those preferences, just that they cannot be represented by a continuous utility function.

EUT, therefore, is a descriptive theory which may or may not be a useful model of decision makers depending on the context of the situation they are placed in. It is not a normative theory unless we think the axioms themselves have normative force, and there is good reason to think that they do not. From this perspective, the theory's capaciousness is a benefit, not a drawback: it can accommodate many different sets of preferences, allowing it to describe many situations; it is no loss that it does not give precise predictions about what all agents will or should do since it makes no claims to be normative.

It is less clear whether EE is descriptive, normative, or both. Empirical work (discussed in Appendix B) suggests that there is hope that it is a descriptive theory, and one which produces contrasting results with EUT. There are also signs that it is a normative theory, on the grounds that optimising the growth rate is the sensible thing to do: for example Peters (2019, 1218) suggests that the author would make certain choices "because the payments correspond to a higher (additive) growth rate of my wealth" and "it's the growth rate I would optimize."

## Conclusion

Theories of decision making under uncertainty can be normative, descriptive, or both. Utility-type approaches are necessarily descriptive if the axioms are met, but have a much weaker claim to being normative: the axioms of EUT have come in for significant criticism, most persuasively by Matthew Rabin (2000). In contrast growth-optimal approaches appear to be normative, although as we have noted different theorists present different normative justifications for them.

Whether growth-optimal approaches are descriptive is a more difficult question. We are not aware of empirical studies examining how closely individuals' portfolios approximate the growth-optimal benchmark, but EE's metaphysical claims have led to a resurgence of interest in this general area, and two recent

experimental studies—one by David Meder et al. (2021) and another by Arne Vanhoyweghen et al. (2022)—attempt to distinguish empirically between EUT and EE. Unfortunately the too-frequent assumption that EUT ignores time and therefore can be represented as myopic—only looking at the next bet, even when the dynamics are known—leads to biased estimates in Meder et al. (2021). Vanhoyweghen et al. (2022) unwittingly present a situation where the predictions of EUT and EE are, at least in their formulation, identical! A more detailed analysis of both papers is provided in Appendix B.

We expect that, for real decisions which real people face, utility approaches and growth-optimal approaches are likely to give the same answer in many cases. This is especially true for short gambles where the dynamic is not given by the gamble but instead by what will happen to the person's wealth over their lifetime: for example if we assume that excess wealth gets put into a diversified portfolio then it is likely to grow multiplicatively. From this perspective, we could say that the dynamic is an aspect of someone's life, rather than a particular gamble, which would tie them down to a single utility. This remains the case for EUT models over time such as in Samuelson (1969) where consumption also depends on investment performance. This is because, as discussed above, EUT approaches automatically incorporate the growth dynamics of different investments: the only reason they differ is that EUT considers the spread of possible outcomes, and when most of the weight of this final probability distribution is over a very small range, we generally expect this to not be a very significant difference—although, as we have seen, it can be. In contrast, a third approach to investment—single-period models with particular objectives, such as Harry Markowitz (1952)'s pioneering variance-aversion model of portfolio choice—do not directly consider the effects of volatility on investments' performance over time (although variance aversion will implicitly result in similar behaviour). Here, therefore, we would expect the models' conclusions to diverge to a greater extent.

The great benefit of growth-optimal approaches to decision making is that they describe an underlying characteristic of gambles, which is true regardless of the psychology of any decision maker who happens to face them. However when individuals are faced with uncertainty, psychology will necessarily play a role: this may express itself in maximising the chances of getting rich; in extreme caution (e.g., minimax); or in other ways. There are many drawbacks to utility approaches, and we have discussed them at length elsewhere (Kay and King 2020). Their great benefit, however, is that they explicitly list the conditions under which they will necessarily be valid. We have shown that if these conditions are met then they necessarily contradict growth-optimal approaches, but this should not be seen as a criticism of the benefits growth-optimal perspectives bring, both scientifically and pragmatically. Instead this result should be seen as emphasising the fundamental

importance of psychological factors layered on top of a scientific understanding of the systems, and as emphasising how we cannot remove individuality from the equation when it comes to decisions made under uncertainty. To improve our understanding of decision making we should study these psychological factors, investigating how people actually make decisions.

## **Appendix A.**

### **Using expected utility theory over time**

In their reply to Peters (2019), Jason Doctor et al. (2020) note that EUT is a “static” theory and that economists use dynamic alternatives over time. However for cases like the ones discussed in this paper, where the only wealth individuals derive any direct benefit from is that which is produced at the end of a gamble, it is straightforward to analyse them with EUT. We noted above that EUT ‘mechanically’ incorporates the process by which potential final outcomes are reached. Indeed, EUT defines gambles like these as ‘compound lotteries’ and asserts that only the final outcomes of the compound gamble matter for the decision problem. If we consider a situation where individuals are offered a gamble of a particular length but can choose to participate or not at each stage the situation is a little more complicated, and we need to solve it recursively: so find the utility of every possible final outcome; then look at the preceding decision nodes and compare the utility of taking the lottery they represent to the utility of keeping that wealth for sure, and take the higher value as that node’s utility; etc.

The cases discussed in this paper are a very restricted subset of the ones we are likely to care about. In practice, investment decisions are not about wealth in a single future period, but about a stream of discounted consumption decisions influenced by both realised and expected investment gains. The natural way of setting up such problems is within a broader expected utility framework, such as Samuelson (1969). However for such situations there are well-known problems with EUT’s limitations in distinguishing between risk and time preferences.

## Appendix B.

### Empirical evidence for ergodicity economics

#### Meder et al. (2021)

Meder et al. (2021) is an attempt to test the predictions made by EE, EUT, and Prospect Theory. Participants had to choose between 312 pairs of 50-50 lotteries, and were told that a random selection of 10 of their choices would be realised and applied to an initial wealth of 1000DKK (about \$155) they had been endowed with. They would then receive this final amount. Each participant did this twice: on one day they faced additive dynamics, so each lottery was of the form ‘gain  $x$  or lose  $y$ ’; on the other they faced multiplicative dynamics, so each lottery was of the form ‘increase your endowed wealth by  $x\%$  or decrease it by  $y\%$ .’ To learn about the dynamics, they had a training session on each day where they saw the effect the gambles had on their wealth lottery-by-lottery; however they did not get this form of updating in the actual experiment. Their choices were then analysed and used to estimate an isoelastic utility function for them on each day:  $u(w) = (w^{1-\rho} - 1)/(1 - \rho)$ . They claim that EE predicts  $\rho = 0$  for all participants on the additive day and  $\rho = 1$  for all participants on the multiplicative day; EUT places no restrictions on  $\rho$  other than that each participant’s value of  $\rho$  should be the same on both days. The experiment found that EE’s predictions were met, and that EUT’s were not.

The study received criticism from several sources, most notably in the supplement to Doctor et al. (2020). They noted that EUT was used in a static rather than dynamic way “as if intermediate outcomes were actually received”; there was no analysis of the actual probability distribution faced over the whole experiment; and outcomes were ambiguous but no adjustment was made for ambiguity aversion, or for aversion stemming from the complexity of the probability calculus (Doctor et al. 2020, Supplement pp. 7–9). This critique is supported by our analysis above: EUT is capable of evaluating decisions over time, and ignoring the dynamic nature of the problem may lead to an incorrect prediction.

The primary response to this criticism has been that cognitive constraints prevent individuals in this experiment from ‘looking ahead’ even to the next gamble, and thus the myopic EUT used is appropriate. Meder et al. (2021) states “Whilst it is possible in principle to update one’s expected wealth as a function of the decisions already made, this is computationally implausible, especially under the

demanding cognitive constraints of the task. To compute expected wealth for a given trial, past choices must be recalled and integrated over all possible outcomes. This integration quickly becomes computationally implausible, especially for the multiplicative condition which must consider all the possible wealth trajectories up to the given point in time” (Meder et al. 2021, 11). In their reply to Doctor et al. (2020), published as a supplement to Meder et al. (2021) they state “If subjects are limited by experimental design, unable to compute outcomes of the future, the dynamic models, when faced with this task, effectively make the same predictions as the static models” (Meder et al. 2021, S4 p. 1).

The claim about computational limitations is therefore central, as the authors seem to have conceded that they are not using EUT entirely appropriately. There is no reason to disbelieve the neuroscience, but it does not follow that the defence holds. To see this, note that Meder et al. (2021) implicitly assume a binary: either participants have full knowledge of the structure of the experiment and unlimited processing power, or they are entirely myopic. There is surely a middle ground. For example, participants are certainly unable to calculate the moments of the probability distribution and update these as they move through the experiment, but they may well have intuitive estimates of them. They might consider the learning stage they experienced and remember it seemed quite generous, and therefore expect the mean gain per realised lottery would be somewhat above zero (Meder et al. 2021, 6). They might consider that their wealth seemed to jump around a lot—or a little—in the training session, and consider the implications of this when faced with a risky lottery. Similarly, skewness and kurtosis may have been considered with this informal but not uninformative method. Empirical work suggests that individuals are capable of making these inferences, although they are certainly less accurate than formal calculations.<sup>9</sup> Moreover, Meder et al. (2021) implicitly assume participants can tell the difference between the days, or they would have no reason to make different choices on them, which should raise the question of whether applying myopic EUT is sensible.

If participants can know something about the distribution of outcomes, then the analysis changes. Each choice they faced is best understood as a choice between two lotteries, each of which might be applied to an uncertain wealth level determined by the interaction of their experimentally-endowed wealth with 9 other lotteries. Both the additive and multiplicative days were generous, so participants

---

9. For example, Goldstein and Rothschild (2014) test how individuals can describe probability distributions, including the mean, and Nisbett et al. (1983) suggest that people can reason quite well statistically when primed to do so. There is also a large quantity of work showing how far from perfection individuals can be, such as Kahneman, Slovic, and Tversky (1982). Even if they were very inaccurate, we argue that the participants in Meder et al. (2021) would likely be have been able to make the distinctions discussed above.

had reason to believe that this wealth level would probably substantially exceed their 1,000DKK starting wealth. This alone implies that myopic EUT is misleading. On the additive day, experimental wealth is underestimated, and so a risk-averse participant with isoelastic utility would perceive the same lottery as less risky than the experimental analysis assumes. On the multiplicative day, experimental wealth is underestimated which also tends to overestimate the riskiness of the lottery faced—but there is a countervailing force, because the size of the experimental wealth the multiplicative lottery is applied to is much larger, and thus the riskiness of the lottery is underestimated. Note that if the participant did only possess their experimentally-endowed wealth and had isoelastic utility this would not matter because the relative amount of risk is independent of the wealth on which the lottery is applied. However, because participants had external wealth (which the experimenters did not gather information on) this result does not hold. Furthermore, ignoring external wealth (which is likely to be far in excess of experimental wealth, plausibly by at least two orders of magnitude) also biases the estimates, as lotteries appear more risky than they really are.

Due to a misunderstanding of EUT, Meder et al. (2021)'s estimation process is flawed. This is just as well for EE's supporters as it is very unclear that, had the estimates been unbiased, the results would have supported EE. This is because, as discussed above, decisions on the additive day are independent of each other. The growth dynamic outside of the experiment should therefore be taken into account—and this is likely to be multiplicative (since the individuals could bank their winnings, or invest them in the stock market, etc.).<sup>10</sup> But that would imply that EE would predict  $p = 1$ , rather than  $p = 0$ , on the additive day. For all of these reasons, the experiment does not provide evidence supporting EE over EUT.

### **Vanhoyweghen et al. (2022)**

Vanhoyweghen et al. (2022) is a response to some of the criticisms of Meder et al. (2021), but suffers from distinct problems. Eighty-one participants faced a series of eighty choices between two 50-50 lotteries, of which forty were in an additive setting and forty were in a multiplicative setting. These lotteries were applied to a starting 'wealth' of 1,000.<sup>11</sup> Unlike Meder et al. (2021), the participants faced all the lotteries in a single session, and were shown the lotteries explicitly, rather than associating them with fractals. Crucially, the choices were designed such that there was a 'safer' and 'riskier' lottery: both lotteries had the same expected

---

10. Peters and Gell-Mann (2016, 5) agree: "Wealth is often better modelled with multiplicative dynamics."

11. Unlike Meder et al. (2021) the wealth is purely within the experiment—it cannot be used or withdrawn by the participants.

value but the riskier one had a greater variance. (There were also ten “no-brainer” choices where one lottery is clearly better than the other, i.e., first-order stochastic dominance.) This experimental design was intended to discriminate between EUT and EE behaviour: “this set-up allows for easy differentiation between optimal and non-optimal behaviour according to time-average growth. The bets [lotteries] had the same expected value and time-average growth in the additive setting, which means none of the bet should dominate. However, bets with lower variance and equal expected value -the safer bets- always yielded a higher time-average growth than their more risky counterparts in the multiplicative setting” (Vanhoywhegen et al. 2022, 3).

They assert that their results “lend support to the theory that intuitive human decision-makers behave differently depending on whether their environment is ergodic or non-ergodic. In agreement with the findings of ergodicity economics, intuitive decision makers tend to optimize the time average of their wealth over the expected values” (Vanhoywhegen et al. 2022, 5).

There are three distinct problems with this experimental design. Firstly, we have seen above that approaching a dynamic problem as if each stage is a static decision is an incorrect application of EUT. Similarly to Meder et al. (2021), they argue that cognitive constraints remove the need for a dynamic analysis:

[W]e opted to use the risk argument (MC1) rather than extensions on utility theory which takes multiple periods into account. We believe this argument to be appropriate because respondents could not envisage the capital they would have at the end of the experiment and as such could not use this terminal capital as a heuristic within our experiment; they simply had too little information to increase their decision algorithm to multiple periods. Respondents were kept in the dark of the bets that were to come, including the amount of capital they possessed after each decision, bet outcomes, and the order used to determine terminal wealth, which removes the need for dynamic programming. (Vanhoywhegen et al. 2022, 5)

Curiously, they refer to Samuelson (1975; originally 1969) to support this claim, although that paper makes no such argument. In any case, it fails for precisely the same reasons we present in our analysis of Meder et al. (2021) above.

Secondly, even if we were to take the claim that the lotteries can be evaluated statically at face value, the experiment cannot discriminate between EUT and EE. When a pair of lotteries have the same expected value but differing variances, we can say that the ‘safer’ lottery with lower variance second-order stochastically dominates the ‘riskier’ lottery with higher variance. Under EUT, any risk-averse agent will avoid a second-order stochastically dominated lottery, whilst a risk-neutral agent will be indifferent between them. Therefore if participants are well-



described as risk-averse expected utility maximisers, we would expect no difference between their results and the predictions of EE.

Finally, the experiment's reward structure also muddies its conclusions. Participants did not benefit from their experimentally-endowed wealth in itself, but instead received a fixed prize if they were one of the six participants with the highest endowed wealth at the end of the experiment. If all participants followed the 'safer' strategy, then by symmetry their probability of receiving a prize would be 6/81. If they switched one lottery to its 'riskier' counterpart, they would improve or worsen their relative position, each with probability 0.5. However the payoff is asymmetric: improving their position plausibly increased their probability of winning a prize more than worsening it decreased their probability of winning a prize, due to the improbability of them winning a prize under the original strategy. Therefore participants were incentivised to maximise their chance of winning the prize, which is distinct from maximising, in expectation, their experimental wealth. We note that they try to deal with this problem by providing six prizes—"In order to motivate respondents and mitigate the excess risk-taking linked to winner-take-all games, six prizes could be won" (Vanhoywhegen et al. 2022, 8)—but as we show this is insufficient.

A misunderstanding of EUT has resulted in an experiment that cannot discriminate between EUT and EE, and therefore we cannot accept the claim that this experiment offers empirical support for EE.

## References

- Adamou, Alexander, and Ole Peters. 2016. Dynamics of Inequality. *Significance* 13(3): 32–35. [Link](#)
- Carr, Peter, and Umberto Cherubini. 2020. Generalized Compounding and Growth Optimal Portfolios: Reconciling Kelly and Samuelson. Working paper, April 6. [Link](#)
- Doctor, Jason N., Peter P. Wakker, and Tong V. Wang. 2020. Economists' Views on the Ergodicity Problem. *Nature Physics* 16: 1168. [Link](#)
- Dorling, Danny. 2016. Dynamics of Inequality: A Response. *Significance* 13(3): 36–37. [Link](#)
- Ford, Matthew C., and John A. Kay. 2023. Psychological Frameworks Augment Even Classical Decision Theories. *Behavioral and Brain Sciences* 46: e92. [Link](#)
- Goldstein, Daniel G., and David Rothschild. 2014. Lay Understanding of Probability Distributions. *Judgment and Decision Making* 9(1): 1–14. [Link](#)
- Kahneman, Daniel, Paul Slovic, and Amos Tversky, eds. 1982. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge, UK: Cambridge University Press.
- Kay, John A., and Mervyn A. King. 2020. *Radical Uncertainty*. London: Bridge Street Press.
- Kelly, John L. 1956. A New Interpretation of Information Rate. *Bell System Technical Journal* 35(4): 917–926. [Link](#)
- Keynes, John Maynard. 1924. *A Tract on Monetary Reform*. London: Macmillan.

- Latané, Henry A.** 1979. The Geometric Mean Criterion Continued. *Journal of Banking and Finance* 3: 309–311. [Link](#)
- Markowitz, Harry.** 1952. Portfolio Selection. *Journal of Finance* 7(1): 77–91. [Link](#)
- Meder, David, Finn Rabe, Tobias Morville, Kristoffer H. Madsen, Magnus T. Koudahl, Ray J. Dolan, Hartwig R. Siebner, and Oliver J. Hulme.** 2021. Ergodicity-Breaking Reveals Time Optimal Decision Making in Humans. *PLoS Computational Biology* 17(9): e1009217. [Link](#)
- Nisbett, Richard E., David H. Krantz, Christopher Jepson, and Ziva Kunda.** 1983. The Use of Statistical Heuristics in Everyday Inductive Reasoning. *Psychological Review* 90(4): 339–363. [Link](#)
- Peters, Ole.** 2011. The Time Resolution of the St Petersburg Paradox. *Philosophical Transactions of the Royal Society A* 369(1956): 4913–4931. [Link](#)
- Peters, Ole.** 2019. The Ergodicity Problem in Economics. *Nature Physics* 15: 1216–1221. [Link](#)
- Peters, Ole, and Alexander Adamou.** 2017. Insurance Makes Wealth Grow Faster. Working paper, July 13. [Link](#)
- Peters, Ole, and Alexander Adamou.** 2021. The Time Interpretation of Expected Utility Theory. Working paper, February 28. [Link](#)
- Peters, Ole, and Murray Gell-Mann.** 2016. Evaluating Gambles Using Dynamics. *Chaos* 26: 023103. [Link](#)
- Rabin, Matthew.** 2000. Risk Aversion and Expected-Utility Theory: A Calibration Theorem. *Econometrica* 68(5): 1281–1292. [Link](#)
- Samuelson, Paul A.** 1963. Risk and Uncertainty: A Fallacy of Large Numbers. *Scientia* 98: 108–113.
- Samuelson, Paul A.** 1969. Lifetime Portfolio Selection by Dynamic Stochastic Programming. *Review of Economics and Statistics* 51(3): 239–246. [Link](#)
- Samuelson, Paul A.** 1971. The “Fallacy” of Maximizing the Geometric Mean in Long Sequences of Investing or Gambling. *PNAS* 68(1): 2493–2496. [Link](#)
- Vanhoyweghen, Arne, Brecht Verbeken, Cathy Macharis, and Vincent Ginis.** 2022. The Influence of Ergodicity on Risk Affinity of Timed and Non-Timed Respondents. *Scientific Reports* 12: 3744. [Link](#)
- Von Neumann, John, and Oskar Morgenstern.** 1955. *Theory of Games and Economic Behavior*. Princeton, N.J.: Princeton University Press.

## About the Authors



**Matthew Ford** read History and Economics at St. John's College, Oxford, and then worked with John Kay and Mervyn King as a research assistant on their book *Radical Uncertainty*. Having got interested in decision theory, he returned to St. John's to write *Intellectual Capital*, an academic and financial history of the College, and to investigate how useful decision theories were in understanding the actual situation the College found itself in. His email is [matthewcford@icloud.com](mailto:matthewcford@icloud.com).



**John Kay** has been a Fellow of St. John's College, Oxford, since 1970. In 1979 he became research director and then director of the Institute for Fiscal Studies. In 1986 he founded London Economics, an economic consultancy. He was the first dean of Oxford's Saïd Business School and has held chairs at London Business School, the University of Oxford, and the London School of Economics. He is a Fellow of the British Academy and the Royal Society of Edinburgh. His email address is [johnkay@johnkay.com](mailto:johnkay@johnkay.com).

[Oliver Hulme](#), [Arne Vanhoyweghen](#), [Colm Connaughton](#), [Ole Peters](#),  
[Simon Steinkamp](#), [Alexander Adamou](#), [Dominik Baumann](#), [Vincent Ginis](#),  
[Bert Verbruggen](#), [James Price](#), and [Benjamin Skjold's](#) reply to this article

[Go to archive of Comments section](#)  
[Go to September 2023 issue](#)



# Reply to “The Limitations of Growth-Optimal Approaches to Decision Making Under Uncertainty”

Oliver Hulme<sup>1</sup>, Arne Vanhoyweghen<sup>2</sup>, Colm Connaughton<sup>3</sup>, Ole Peters<sup>4</sup>, Simon Steinkamp<sup>5</sup>, Alexander Adamou<sup>6</sup>, Dominik Baumann<sup>7</sup>, Vincent Ginis<sup>8</sup>, Bert Verbruggen<sup>9</sup>, James Price<sup>10</sup>, and Benjamin Skjold<sup>11</sup>

[LINK TO ABSTRACT](#)

We are writing in response to Ford and Kay (2023), where the authors criticise a decision theory which emerges in the field of ergodicity economics (EE), contrasting it with another from expected-utility theory (EUT). We will refer to the two models as EE and EUT, but we emphasize that both ergodicity economics

---

1. Danish Research Centre for Magnetic Resonance, Copenhagen University Hospital Amager & Hvidovre, Copenhagen, Denmark; Department of Psychology, University of Copenhagen, Copenhagen, Denmark.

2. Data Lab, and Mobilise, Vrije Universiteit Brussel, 1050 Brussels, Belgium.

3. London Mathematical Laboratory, London W6 8RH, UK; Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK.

4. London Mathematical Laboratory, London W6 8RH, UK; Santa Fe Institute, Santa Fe, NM 87501, US.

5. Danish Research Centre for Magnetic Resonance, Copenhagen University Hospital Amager & Hvidovre, Copenhagen, Denmark.

6. VesselsValue Ltd, London W14 0QL, UK.

7. Department of Electrical Engineering and Automation, Aalto University, Espoo, Finland; Department of Information Technology, Uppsala University, Uppsala, Sweden.

8. Data Lab, Vrije Universiteit Brussel, 1050 Brussels, Belgium; School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, US.

9. Data Lab, Vrije Universiteit Brussel, 1050 Brussels, Belgium.

10. Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK.

11. London Mathematical Laboratory, London W6 8RH, UK; Danish Research Centre for Magnetic Resonance, Copenhagen University Hospital Amager & Hvidovre, Copenhagen, Denmark.

and expected-utility theory are labels for fields much broader than what we discuss here. We believe that the criticism of the EE model is due to a misunderstanding on the part of the authors of Ford and Kay (2023), and we thank them for bringing this to our attention so that it may be clarified.

Instead of providing a point-by-point reply, we limit our response to two points, which we hope will unlock the key misunderstanding and clarify where other apparent disagreements come from.

First, we feel the authors have missed an important point about the relationship between EE and EUT. A mapping between the models exists, but the key condition which needs to be satisfied for this mapping to hold is that the utility function in EUT is chosen to be the ergodicity transformation of EE. It seems that the authors believe that the key condition is merely sufficiently long time scales, but this is not the case. We therefore clarify the relationship between the two fields by specifying exactly the mapping between EE and EUT.

Second, we believe we have identified a misunderstanding regarding convergent and non-convergent properties of random walks, which may have led to the first misunderstanding above as it led the authors to write that “final wealth will almost always be what the time average predicts” (Ford and Kay 2023, 317). In the stochastic processes typically studied in EE, the appropriately defined growth rate converges in the long run with probability one to its time average (and expected value). However, this does not imply, as the authors write incorrectly, that final wealth converges to its time average. Clarifying this misunderstanding will help resolve their concerns about other aspects of EE. We provide exact computations for the Peters coin toss discussed by Ford and Kay.

## Formal setup

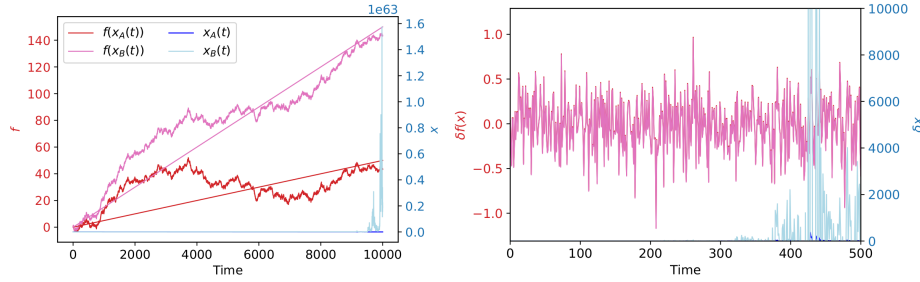
Both the EE model and the EUT model make use of a variable which represents the wealth of a decision-making agent. However, the way wealth is modelled is different in the two cases, and consequently, so is the way decision-making is modelled. A mapping between the models exists; that is, we can specify conditions under which they are equivalent. Generally, they are not equivalent, and it is important to state the exact conditions for the mapping to hold in order to specify the relationship between the models.

### EE model

The formal setup for the EE model is illustrated in Figure 1. It is a choice between two stochastic processes,  $x_A(t)$  and  $x_B(t)$  (Peters and Adamou 2018; Carr

and Cherubini 2020), representing wealth over time. The processes are chosen so that there exists a monotonically increasing transformation,  $f(x)$ , whose increments,  $\delta f(t) = f[x(t + \delta t)] - f[x(t)]$ , are ergodic. In particular, the time average of these increments is identical to the expected value of the increments,

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{\tau} \delta f(t + \tau \delta t) = \mathbb{E}[\delta f]. \quad (1)$$



**Figure 1.** Left: EE operates on stochastic processes,  $x_A(t)$  and  $x_B(t)$  (blue random lines, here geometric Brownian motions). It applies an ergodicity transformation  $f$  (here the logarithm) which produces  $f(x_A(t))$  and  $f(x_B(t))$  (red random lines, Brownian motions). These transformed processes are linear in time. Their slopes  $\delta f / \delta t$  converge to the time-average growth rates (slopes of the straight red lines) as  $\delta t$  becomes large. Right: increments  $\delta f$  over a single time unit (red lines, limited to the first 500 time units for clarity). These have the ergodic property that their expected value equals their time average. The increments in the original processes,  $\delta x$  (blue lines), do not have this property, are unstable and not suitable for many computations of interest.

Because of this mean-ergodicity property, the transformation  $f$  is called the ergodicity transformation. The rate of change of the ergodic increments is the appropriately defined growth rate for the process,

$$g = \frac{\delta f}{\delta t}. \quad (2)$$

*EE decision axiom.* According to the EE model, agents choose the process which maximizes the time-average or, equivalently (because of the ergodic property), the expected value of the growth rate  $g$ .

*Motivation.* Agents in this model can be thought of as representations of people who make decisions in a financial context, where  $x$  represents wealth. By maximizing the time average of equation (2), agents maximize the long-term

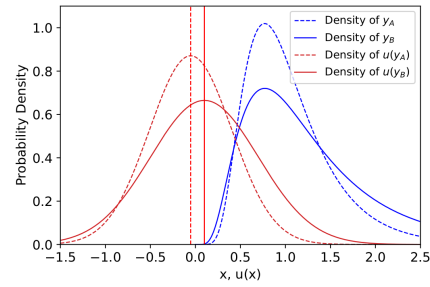
growth rate of their wealth. In the long run, agents who act in this way become wealthier than agents who act differently.

## EUT model

The formal setup for the EUT model is different; see Figure 2. Here, we deal with a choice between two random variables (not between two stochastic processes),  $y_A$  and  $y_B$ , representing wealth. A monotonically increasing transformation  $u(y)$ , called the utility function, is defined.

*EUT decision axiom.* According to the EUT model, agents choose the random variable which maximizes the expected value of the utility function,  $\mathbb{E}[u(y)]$ .

*Motivation.* Agents in this model can also be thought of as people who make decisions in a financial context, where  $y$  represents wealth at a future point in time. The utility function may be thought of as a quantification of how a given level of wealth relates to its subjective value. A concave utility function, for instance, represents a person who assigns less value to an extra dollar as wealth increases. By using the expected value of utility as their maximand, agents weight utility according to the probability of attaining it. Because the utility function can be freely chosen, this model can describe many different behaviours. In EE, both wealth and utility are maximized as time passes. There is no similar physical motivation for EUT; see the next section.



**Figure 2.** Under EUT, agents choose between two random variables,  $y_A$  and  $y_B$ , defined by their probability density functions in blue (here lognormals). The agent applies a transformation,  $u$ , which defines the new random variables  $u(y_A)$  and  $u(y_B)$  (red, normals) and chooses the option which maximizes the first moment of  $u$  (vertical red lines), here option  $B$  is chosen because its expected utility (solid vertical line) is greater than that of option  $A$  (dashed vertical line).

## Mapping EE and EUT

A stochastic process is a family of random variables parameterized by time. We can, therefore, move from the setup of EE to the setup of EUT by specifying the current time,  $t$ , at which we wish to evaluate wealth under the processes  $x_A$



and  $x_B$  as it will have evolved by the later time,  $t + \delta t$ , see Figure 3. To establish the mapping, we identify the random variables thus derived from the stochastic processes as the random variables required for an EUT treatment,

$$x_A(t + \delta t) = y_A \text{ and } x_B(t + \delta t) = y_B. \quad (3)$$

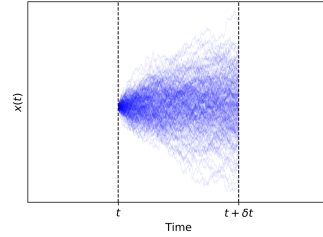
Under the EE model, we compute the time-average growth rates of wealth under  $A$  and  $B$  as the rate of change in the expected value of ergodicity-transformed wealth,

$$\bar{g} = \frac{1}{\delta t} \{ \mathbb{E}[f(x(t + \delta t))] - f(x(t)) \}. \quad (4)$$

The factor  $\frac{1}{\delta t}$  does not affect the ranking of the time-average growth rates for  $A$  and  $B$ , and we can drop it from equation (4). Further, because current wealth is identical for both processes  $A$  and  $B$ , subtracting  $f(x(t))$  in equation (4) does not affect the ranking either, and we can also drop it. Maximizing equation (4) is therefore equivalent to maximizing  $\mathbb{E}[f(x(t + \delta t))]$ .

Completing the mapping, we summarize that maximizing equation (4) is equivalent to maximizing expected utility in the special case where

- future-wealth random variables  $x_A(t + \delta t)$  and  $x_B(t + \delta t)$  considered in the EE model are the random variables  $y_A$  and  $y_B$  considered in the EUT model.
- the ergodicity transformation considered under EE is the utility function considered under EUT.



**Figure 3.** The random variable,  $y$ , required to map EE to EUT is the value of the stochastic process,  $x$ , of EE at some fixed time, here  $x(t + \delta t)$ , as in equation (3). We imagine the process to start at a known value at time  $t$  and use the density of its trajectories at  $t + \delta t$  to define the random variable  $y = x(t + \delta t)$  required for the EUT treatment.

## Relationship between EE and EUT

The EE model and the EUT model are equivalent under the restrictive conditions specified in the previous section. It seems interesting to us to highlight what differences emerge when only the second condition is violated, that is, when

the utility function is not the ergodicity transformation,  $f \neq u$ . In this case,

- Only agents who act according to EE and maximize the time-average growth rate of wealth,  $\bar{g}$ , also maximize utility,  $u$ , as time passes.
- Agents acting according to EUT and maximizing expected utility,  $E(u(y))$ , do not maximize utility,  $u$ , as time passes.

These points are rarely stated but they constitute an important limitation of EUT. EUT places great emphasis on defining utility as its quantification of subjective value. In particular, EUT holds that it is better to use utility than money when attempting such a quantification. One might expect the formalism of EUT to guarantee that utility itself—the object of desire by definition—would be maximized over time by agents acting according to its behavioral criterion. But this is not the case.

In contrast, EE focuses on maximising the time-average (or expected) growth rate of wealth. In doing so, it guarantees that EE agents, unlike EUT agents, maximize not only wealth but also utility as time passes. This follows from the assumption that utility is monotonically increasing in wealth (see section “Formal setup,” subsection “EUT model” above). As time passes, EE agents are guaranteed to do better than EUT agents, in terms of wealth and utility.

That EUT maximizes expected utility but does not maximize utility over time is a direct consequence of the non-ergodicity of the wealth process and the utility process it induces. Ergodicity implies that expected value and time average are identical, and therefore in an ergodic utility process it is guaranteed that optimizing the expected value of utility also optimizes the time average of utility. However, the processes usually considered in EE are not ergodic. These include standard models in finance and economics, for example Brownian or geometric Brownian motion. Here, the equality of expected value and time average does not hold, and consequently expected-utility theory does not optimize utility over time.

We feel that we should clarify another comment by Ford and Kay because it allows us to highlight the astonishing experimental results obtained by the Copenhagen group. EE, as Ford and Kay put it, violates the axiom of completeness and leads to “inconsistent decision making” (2023, 316). This is a strange way of saying that for a given pair of random variables  $y_A$  and  $y_B$ , EE can conclude that either  $A$  or  $B$  is preferable if the dynamic is left unspecified. Of course the same is true of EUT, if the utility function is left unspecified. It is unclear to us what practical problem arises from this, but this served the Copenhagen group in their attempts to put EE to the test.

Recall that the ergodicity transformation is given by the dynamic of the

stochastic process  $x(t)$ . It is a subtle but important detail that two different stochastic processes, both starting at  $x(t)$  can yield identical distributions at a later time  $x(t + \delta t)$ . In particular, the processes may have different ergodicity transformations. Therefore, an EE agent evaluating what looks to EUT as the same situation, namely the same random variables  $y_A = x_A(t + \delta t)$  and  $y_B = x_B(t + \delta t)$ , can arrive at different preferences depending on the process which generates the random variables.

From the perspective of EUT, because it does not take dynamic information into account, one could call this a violation of the completeness axiom of EUT. Put more prosaically, this is a case of a hidden variable, and once the dynamic is specified, the preferences of EE agents satisfy completeness. The case of the EE agent using different dynamics is equivalent to the case of an EUT agent using different utility functions. Of course, using different utility functions, the EUT agent can also arrive at different preferences, violating completeness in the same sense. But once a utility function is specified, also EUT preferences satisfy completeness.

However, the fact that EE preferences change according to dynamics enables experimental explorations of the theory. By manipulating the wealth process,  $x(t)$ , in simple gambling tasks in a laboratory setting, experimenters can control the ergodicity transformation. This makes it possible to test whether real human subjects behave according to idiosyncratic utility functions or according to circumstantial ergodicity transformations. To the great astonishment of most of us, the latter is often the case: fitting the EUT model to choices made under different dynamics reveals that people change their apparent utility functions to coincide with the relevant ergodicity transformation (Meder et al. 2021; Skjold et al. 2023).

By its construction, EUT does optimize *expected* utility, but because the wealth process is not ergodic, this object is different from what materializes for the decision-making agent. While the expected value is approximated by the average over large statistical samples, it is not generally a quantity of interest for an individual decision maker.

We illustrate this with the Peters coin toss. An agent is offered a repeated fair coin toss, where heads leads to a 50 percent rise in wealth and tails leads to a 40 percent drop. The coin toss is also discussed in Ford and Kay (2023), a video about it is available via [ergodicity.tv](https://ergodicity.tv) ([link](#)), and an interactive blog post is available on [ergodicityeconomics.com](https://ergodicityeconomics.com) ([link](#)). It is a special case of the multiplicative binomial process (Redner 1990).

First, we evaluate the gamble on offer using the EE model. Here, two stochastic processes are compared. The first stochastic process is trivial: if the agent rejects the gamble, wealth will be unchanged at its current level,  $x_A(t + \delta t) = x(t)$ .

The second stochastic process arises from the agent accepting the gamble,

and  $x_B(t)$  is a random walk in logarithmic space. The ergodicity transformation for this particular dynamic is the logarithm,

$$f(x) = \ln(x), \quad (5)$$

and both  $\delta \ln(x_A)$  and  $\delta \ln(x_B)$  are ergodic. This means the appropriately defined growth rate is

$$g = \frac{1}{\delta t} [\ln x(t + \delta t) - \ln x(t)] \quad (6)$$

and its time average is

$$\bar{g} = \frac{1}{\delta t} \mathbb{E}[\delta \ln x]. \quad (7)$$

To be explicit: it is the ergodic property of  $\delta \ln x$ , which allows us to maximize the time average of  $\delta \ln x$  (and thereby  $\bar{g}$ ) by maximizing the expected value  $\mathbb{E}[\delta \ln x]$ . This maximisation guarantees that we end up with greater wealth (and utility) in the long-time limit. Evaluating for both processes, we find  $\bar{g}(x_A) = 0$  per round and  $\bar{g}(x_B) \approx -0.05$  per round. The EE agent picks the process with the greater time-average growth rate, rejects the gamble and remains at  $x(t + \delta t) = x(t)$ .

Second, we evaluate the gamble on offer using the EUT model. To be able to do this, the agent needs to specify its utility function. To illustrate the problem with a simple example, let's say the agent has linear utility,  $u(y) = y$ , although the situation we're about to highlight also arises with many other utility functions.

For simplicity, we let the agent evaluate utility after one round, although nothing changes if the agent were to evaluate utility after an arbitrary number of rounds. Here, two random variables are compared. The first random variable is trivial, namely, wealth remains unchanged if the gamble is rejected, and  $y_A = x(0)$ . The expected utility associated with this random variable is, trivially,

$$\mathbb{E}[u(y_A)] = x(0). \quad (8)$$

The second random variable,  $y_B$ , takes the value  $1.5x(0)$  with probability  $1/2$  and  $0.6x(0)$  with probability  $1/2$ . The expected utility associated with  $y_B$  is

$$\mathbb{E}[u(y_B)] = \frac{1}{2}[u(1.5x(0)) + u(0.6x(0))] = 1.05x(0). \quad (9)$$

Since  $\mathbb{E}[u(y_B)]$  is greater than  $\mathbb{E}[u(y_A)]$ , the EUT agent with linear utility will always choose to participate in the gamble.

However, as we’ve seen in the EE analysis, the time-average growth rate of wealth is negative for this gamble: the probability that the agent loses money approaches 1 over time. Because the agent’s utility function is monotonically increasing, losing money means losing utility.

This illustrates that the EE agent maximizes utility over time, whereas the EUT agent only maximizes *expected* utility but not actual utility. In the simple example we’ve given, the EUT agent loses utility as time passes, whereas the EE agent does not. Ergodicity in multiplicative dynamics is broken in such a way that the expected value of many monotonically increasing utility functions does not indicate how utility actually behaves with probability 1 over time. In many cases, as in our example, an increasing expected utility  $\mathbb{E}[u(t)]$  is accompanied by systematically decreasing actual utility,  $u(t)$ .

## Wealth uncertainty diverges while growth-rate uncertainty vanishes

That the EE model produces such different outcomes from the EUT model is a profound consequence of uncertainty, which we believe was overlooked in Ford and Kay (2023, 317), where the authors write: “The time averages used in [the EE model] correspond to a situation where there is no measurable uncertainty—final wealth will almost always be what the time average predicts.”

This sentence seems to us to reflect a misunderstanding. The EE model uses the time average of the growth rate of wealth,  $\bar{g}$ , as its decision criterion. It does this because this quantity converges to a meaningful finite value, and such a simple scalar is needed to rank the stochastic processes  $x_A$  and  $x_B$ . However, this convergence does not imply that wealth itself,  $x(t)$ , converges to a value predicted by the time-average growth rate. For instance, in the Peters coin toss, if we average the growth rate over a finite time  $T$ , then its variance vanishes as  $1/T$ . The variance of wealth, on the other hand, diverges exponentially. Contrary to Ford and Kay’s statement, there is great uncertainty in final wealth, in the sense that it diverges in the limit  $T \rightarrow \infty$ , whether we measure it by variance or other relevant ways of measuring uncertainty (see [Appendix](#)). Specifically, in the Peters coin toss with

linear utility, the uncertainty leads to the ordering under EUT being different from the ordering under EE, whether after a finite or divergent number of rounds. This means that the uncertainty in terminal wealth is not only measurable but crucially important in the case under consideration.

We can only speculate here, but this misunderstanding may explain other comments by Ford and Kay (2023), which we find difficult to understand otherwise. For instance, the authors write:

- “[O]ne might expect that, as the gamble’s length goes to infinity, the predictions of EUT would approximate the growth-optimal predictions” (Ford and Kay 2023, 318).
- “[U]tility approaches [the EUT model] and growth-optimal approaches [the EE model] are likely to give the same answer in many cases” (ibid., 326).
- [The difference between the growth rate of the expected value and the time-average growth rate] “is mechanically incorporated in an EUT analysis of choices as only final outcomes are considered” (ibid., 318).

As we say in the “Mapping EE and EUT” section above, the two approaches *only* give the same answer if the utility function,  $u$ , is chosen to be the ergodicity transformation,  $f$ . However, the authors seem to be under the impression that the EE model essentially considers wealth far in the future and that this wealth is known with “no measurable uncertainty” (Ford and Kay 2023, 317). If this were the case, then their statement would be true: EE would compare known wealths  $x_A(t)$  and  $x_B(t)$  at some large  $t$ , and EUT would compare utilities  $u[x_A(t)]$  and  $u[x_B(t)]$  (we would be allowed to replace  $E[u[x_A(t)]]$  by  $u[x_A(t)]$  when there’s “no measurable uncertainty” in  $x_A(t)$ ). Because  $u(x)$  is assumed to be a monotonically increasing function, the preference orderings of A and B would be the same under both models. If this were true, one would presumably use neither EE nor EUT and just compare asymptotic wealth or, equivalently, utility. But none of this is actually the case, and uncertainty in  $x(t)$  grows beyond all bounds with  $t$ .

Wealth, in absolute terms, in the Peters coin toss goes to zero with probability one. In this sense, asymptotic wealth is known in this particular case. However, this is an asymptotic statement which must be interpreted carefully. The statement that terminal wealth is known with “no measurable uncertainty” is not correct, even in this special case as should be clear from the fact that expected wealth diverges while most probable wealth goes to zero. Uncertainty in terminal wealth diverges with time if we measure it by standard deviation; it also diverges if we use relative measures of uncertainty (see [Appendix](#)); most significantly, in the

present case (coin toss with linear utility), the uncertainty in terminal wealth leads to expected utility of terminal wealth being positively divergent, whereas utility of the terminal wealth which is approached with probability one is zero. This illustrates once more that EE is a very different model than EUT.

## Experiments and the role of psychology

Ford and Kay are critical of experimental work carried out to establish the realm of validity of the EE model (Meder et al. 2021; Vanhoyweghen et al. 2022). We agree with many of the criticisms and will provide a detailed response in a separate reply. We note that experimental design is always subject to constraints, such as ethical and financial considerations. Nor can a laboratory experiment ever be truly realistic, and in the process of designing, choices must be made. We have actively sought critiques of the designs of the existing experiments and have ourselves spent a great deal of time discussing weaknesses and alternatives. We have taken seriously all critiques we have received and incorporated them into the next phase of experiments (Skjold et al. 2023).

Controlled laboratory experiments are of interest to psychologists and neuroscientists. The experiments help clarify in how far EE provides a new behavioral baseline, deviations from which may yield insights into individual psychology. Ford and Kay's concern that EE "claims to provide an objective justification for decision making without the need to refer to individual psychology" (2023, 315) is therefore misplaced. It is not by accident that individual psychology (represented by idiosyncratic utility functions in EUT) is excluded from the EE model. Neuroscientifically, it is extremely exciting that behavior which used to be thought of as idiosyncratic, trait-like, and changeable only on evolutionary time scales can actually be altered by a simple intervention in the experimental environment on time scales of hours rather than millennia.

## Conclusion

We conclude with a quote from the memoir of Giorgio Parisi. Referencing the difficulty of transmitting ideas across traditional disciplinary boundaries, he wrote: "I believe that with a lot of good faith and a lot of patience, at least in the majority of cases, it is possible to arrive at shared conclusions. Or at the very least to clarify where our disagreements come from" (Parisi 2023, 18). We hope that this will be the case in the present context too.



## Appendix

We provide a technical analysis of the coin toss game in an appendix available at the journal's website ([link](#)).

## Data and code

All data shown here are for illustration purposes only and were randomly generated. Codes to reproduce all data and figures are available from the journal's website ([link](#)) and are archived at Zenodo ([link](#)).

## Funding statement

The ergodicity economics program at London Mathematical Laboratory is supported by Baillie Gifford, and its collaboration with the Danish Research Centre for Magnetic Resonance is supported by the Novo Nordisk Foundation (Exploratory Interdisciplinary Synergy Grant, ref NNF20OC0064869).

## References

- Carr, Peter, and Umberto Cherubini.** 2020. Generalized Compounding and Growth Optimal Portfolios: Reconciling Kelly and Samuelson. Working paper. [Link](#)
- Ford, Matthew C., and John A. Kay.** 2023. The Limitations of Growth-Optimal Approaches to Decision Making Under Uncertainty. *Econ Journal Watch* 20(2): 314–334. [Link](#)
- Meder, David, Finn Rabe, Tobias Morville, Kristoffer H. Madsen, Magnus T. Koudahl, Ray J. Dolan, Hartwig R. Siebner, and Oliver J. Hulme.** 2021. Ergodicity-Breaking Reveals Time Optimal Decision-Making in Humans. *PLoS Computational Biology* 17(9): e1009217. [Link](#)
- Parisi, Giorgio.** 2023. *In a Flight of Starlings: The Wonder of Complex Systems*. New York: Penguin Random House.
- Peters, Ole, and Alexander Adamou.** 2018. The Time Interpretation of Expected Utility Theory. Working paper. [Link](#)
- Redner, Sidney.** 1990. Random Multiplicative Processes: An Elementary Tutorial. *American Journal of Physics* 58(3): 267–273. [Link](#)
- Skjold, Benjamin, Simon R. Steinkamp, Colm Connaughton, Oliver J. Hulme, and Ole Peters.** Are Risk Preferences Optimal? Working paper. [Link](#)
- Vanhoyweghen, Arne, Brecht Verbeken, Cathy Macharis, and Vincent Ginis.** 2022. The Influence of Ergodicity on Risk Affinity of Timed and Non-Timed Respondents. *Scientific Reports* 12(1): 3744. [Link](#)

## About the Authors

**Oliver Hulme** is an Associate Professor at the Psychology Department of Copenhagen University and a group leader at the Danish Research Centre for Magnetic Resonance. His research investigates how the human reward system works, operating at the intersection of economics, psychology, neuroscience, and computational biology. His email address is [oliverh@drcmr.dk](mailto:oliverh@drcmr.dk).

**Arne Vanhoyweghen** is currently pursuing his Ph.D. in Business Engineering at the Vrije Universiteit Brussel and is a teaching assistant in mathematics, logistics, and operations management. His research focuses on ergodicity economics and mobility. His email address is [arne.vanhoyweghen@vub.be](mailto:arne.vanhoyweghen@vub.be).

**Colm Connaughton** is the Director of the London Mathematical Laboratory and emeritus professor of applied mathematics at the University of Warwick. His research interests are in statistical physics, mathematical modelling of socio-technical systems, and applied data science. His email address is [c.connaughton@lml.org.uk](mailto:c.connaughton@lml.org.uk).

**Ole Peters** is the Founder of the London Mathematical Laboratory and leads its ergodicity economics program. He is also an External Professor at the Santa Fe Institute. His email address is [o.peters@lml.org.uk](mailto:o.peters@lml.org.uk).

**Simon Steinkamp** is a postdoc at the Danish Research Centre for Magnetic Resonance. His research on the human reward system is at the intersection of cognitive computational neuroscience and experimental psychology. His email address is [simons@drcmr.dk](mailto:simons@drcmr.dk).

**Alexander Adamou** is Vice President of Data Science & Analytics at VesselsValue, a Veson Nautical company. Previously, he was a Fellow and Director of the London Mathematical Laboratory, where he conducted research in ergodicity economics. His email address is [a.adamou@lml.org.uk](mailto:a.adamou@lml.org.uk).

**Dominik Baumann** is an assistant professor in the Department of Electrical Engineering and Automation at Aalto University, Espoo, Finland, and a visiting researcher in the Department of Information Technology, Uppsala University, Sweden. His research interests revolve around learning and control for networked multi-agent systems. His email address is [dominik.baumann@aalto.fi](mailto:dominik.baumann@aalto.fi).

**Vincent Ginis** is an associate professor at the Vrije Universiteit Brussel and a

Visiting Professor at the Harvard School of Engineering and Applied Sciences. His research focus is complexity and machine learning. His email address is [vincent.ginis@vub.be](mailto:vincent.ginis@vub.be).

**Bert Verbruggen** is pursuing his Ph.D. in interdisciplinary sciences at the Vrije Universiteit Brussel, where he also serves as a teaching assistant, focusing on Mathematics. His research interests encompass Mathematics, Machine Learning, Learning Dynamics, and Physics. His email address is [Bert.Verbruggen@vub.be](mailto:Bert.Verbruggen@vub.be).

**James Price** is a Ph.D. student at the MathSys II CDT within the Mathematics Institute at the University of Warwick. His interests are time-based decision-making theories such as those studied in ergodicity economics. His email address is [james.price.2@warwick.ac.uk](mailto:james.price.2@warwick.ac.uk).

**Benjamin Skjold** is a Fellow at the London Mathematical Laboratory and an external researcher at the Danish Research Centre for Magnetic Resonance. His research focuses on ergodicity economics, with a particular emphasis on decision-making, cooperation, and insurance. His email address is [b.skjold@lml.org.uk](mailto:b.skjold@lml.org.uk).

[Go to archive of Comments section](#)  
[Go to September 2023 issue](#)



# Reply to LeCraw, Montanera, and Mroz on Hospitals’ Malpractice Claims and Costs

Maayan Yitshak-Sade<sup>1</sup>, Allen Kachalia<sup>2</sup>,  
Victor Novack<sup>3</sup>, and Michelle M. Mello<sup>4</sup>

[LINK TO ABSTRACT](#)

The previous issue of *Econ Journal Watch* features an article by Florence LeCraw, Daniel Montanera, and Thomas Mroz—henceforth, LMM (2023)—examining the 2018 article in *Health Affairs* by us and seven other coauthors on the outcomes of implementing a communication-and-resolution program (CRP) in six Massachusetts hospitals (Kachalia et al. 2018). For a brief explanation of the purpose and nature of CRP programs, we recommend the second paragraph of LMM’s article. As LMM go on to say in their third paragraph, the disagreement between them and us does not amount to different conclusions about the utility of CRPs for addressing patient injuries; both they and we conclude that the CRP approach is promising. Rather, the disagreement is over the analytic approach that should be used to generate evidence concerning CRPs’ effects (LMM 2023, 1–2).

Their paper caps over four years of engagement between our study team and LeCraw and Mroz to address their questions about our study. Their criticisms have been a moving target over time as we have addressed particular points; the latest round alleges “three major methodological errors” but discusses two. Both critiques are misplaced.

---

1. Icahn School of Medicine at Mount Sinai, New York, NY 10029.

2. Johns Hopkins Medicine, Baltimore, MD 21287.

3. Soroka University Medical Center, Beer-Sheva, Israel.

4. Stanford Law School, Stanford, CA 94305; Stanford University School of Medicine, Stanford, CA 94305.

First, LMM assert that when using an interrupted time series design, the hospitals in our sample that did not implement CRPs are inappropriate counterfactuals for those that did. Their concern is that the pre-implementation liability trends in the treatment and comparison hospitals were not identical. LMM (2023, 3) offer an alternative method of determining the outcomes of CRP implementation that directly compares one of the hospitals that implemented a CRP to other hospitals without CRPs. Importantly, we considered and rejected that design. The reason it was deemed inappropriate was that hospitals could not provide all of the data needed to adequately control for the range of ways in which they might differ. Because of the risk of residual confounding, we selected an interrupted time-series design, which uses each hospital as its own control. Non-implementing hospitals were included in the sample, undergoing the same analysis as a way of showing whether the changes observed over time at the implementing hospitals also held true at hospitals that did not implement CRPs.

LMM (2023, 4) report that the pre-implementation slopes of the implementing and comparison academic medical centers (AMCs) are not identical. Though we disagree with LMM's characterization of the nonparallelism as "severe" (2023, 9), it's reasonable to conclude that various differences among the sites should give pause about directly comparing sites. That is why our interrupted time-series approach was selected and is appropriate, while LMM's preferred approach is not appropriate.

As reported in our article's Online Appendix, because our analysis focused on the pre/post comparison within each hospital, we performed testing to verify that the requirements for conducting interrupted time-series analysis held in our data. We verified that the stationarity assumption was met for all the tested outcomes using the Dickey-Fuller test and the autocorrelation and partial autocorrelation functions. We concluded that the series were stationary and therefore correction for differencing was not needed.

Contrary to LMM's suggestion (2023, 9–10), we did not conduct our analysis within a causal modeling framework. Rather, our observational analysis yielded measures of association and is described as such in the article. LMM's claim that interrupted time series was deployed to provide "causal evidence" is incorrect, as is the assertion that we characterized our results as showing improved liability outcomes "due to CRP" (2023, 3, 9).

LMM (2023, 9) worry that we introduced "Difference in Nominal Significance" errors with our design. Specifically, they describe error arising from comparing the changes in outcomes over time between independent groups, where one result is significant and the other null, and concluding that the two groups' outcomes are different. Their critique exports arguments made in an article by Martin Bland and Douglas Altman (2015) to a different context where the argu-

ments have little traction. (The other article that LMM cited to buttress their criticism, by David B. Allison et al. (2016), doesn't discuss the issue.) Bland and Altman make a persuasive case for pooling data across units of analysis in the context of a randomized, controlled trial. But in the absence of randomization, if adequate control for confounding variables cannot be assured, there are strong arguments against an approach that relies heavily on comparing observational units rather than also using each unit as its own control. Further, of the three potential adverse consequences of this type of error described by Bland and Altman, only one—the possibility of declaring a difference significant when it is not—has any potential applicability to our study. Yet for four of the six outcome variables we concluded that there were no significant differences. Our study's core conclusion is not that there were significant improvements in liability outcomes associated with CRP implementation, but that hospitals can implement CRPs without experiencing worse liability outcomes.

LMM's second line of critique is that when they implemented a different modeling approach, they observed a “large spike in both of the liability outcomes” that occurred “as soon as CRP was implemented” (2023, 6). They allege that our article reports downward trends in liability outcomes but does not acknowledge large, immediate increases in the levels of liability outcomes in the implementing hospital. In our own models, too, some hospitals exhibited upwards shifts in the intercept for some outcome variables. This was reported in our article (Kachalia et al. 2018, 1840–1841), alongside our judgment that readers ought to focus on the changes in slope because what occurs immediately following a CRP's implementation date (i.e., the shift in level) is a less reliable and informative indicator of the program's effect than longer-term trends in the outcomes (i.e., the shift in slope).

There are two reasons why what occurs immediately following a CRP's implementation date is less informative than longer-term outcomes. First, a CRP's effects don't switch on overnight. A go-live date can be designated, but as with any complex quality-improvement program, full implementation takes a period of time. Based on our close monitoring of these programs, a ramp-up period of about three months took place at the implementing hospitals. To account for this, our main analysis imposed a one-quarter lag from the start of the post-implementation period. Second, CRPs take time to bring cases to resolution. Some cases may be resolved in a matter of days, but others—for example, those in which the insurer conducts a second review of the case for possible compensation—may require months. Consequently, one cannot observe the program's effect immediately.

An example of the problem posed by relying too heavily on intercept shifts and of using LMM's modeling approach is their finding of “a 200 percent increase in compensation costs (a tripling) immediately following CRP implementation”

at one of the AMCs (LMM 2023, 8). This result strains plausibility; we know of no mechanism through which it could occur. It arises from a model that LMM constructed from the parameter estimates we reported—not using actual data. As they report (2023, 2), the terms of our data use agreement with the involved liability insurers did not permit data sharing. Such terms are unfortunate because they can leave other scholars with few options for pressure-testing study results, leading to second-best strategies that generate information that is simply wrong.

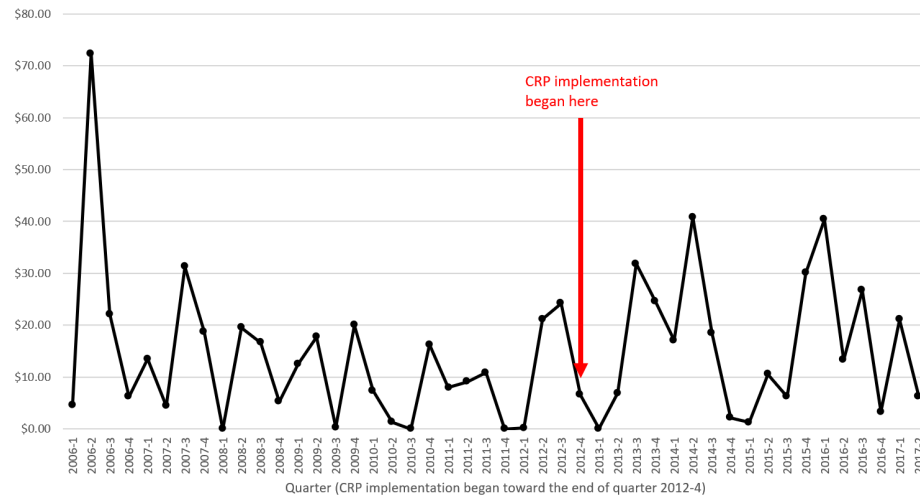
Simple descriptive plots showing the actual, unadjusted data on mean compensation cost rate in each quarter (Figure 1 here) demonstrate that LMM's model predictions do not match what actually occurred, which was a continuation of that hospital's historical experience of considerable quarter-to-quarter volatility in quarterly compensation cost rates. The same is observable in plots of the actual defense-cost rate (Figure 2) and rate of new claims (Figure 3).

Another problem with the approach taken by LMM relates to open claims—that is, claims that had been initiated but not yet resolved at the time data collection concluded. The fact that 16 percent of the claims in our sample remained open creates analytical difficulties, mostly affecting the post-implementation period. As described in our article, the problem is that if we were to impute costs for these unresolved claims using data from the past, it is likely those imputed costs would be too high, because they would be based on the hospital's old (pre-CRP) approach to resolving claims. If so, it could obscure the savings from CRPs. For example, in a sensitivity analysis (Kachalia et al. 2018, Appendix p. 18) we saw that for one academic medical center, excluding open claims increased the magnitude of the change in trend for new claims ( $-0.098$  without open claims vs.  $-0.066$  with open claims) and decreased the upward shift in level ( $0.547$  without open claims vs.  $0.766$  with open claims). Importantly, our sensitivity analysis (ibid., 18–20) showed that the exclusion of open claims affected the results differently across implementing and non-implementing hospitals. As we reported, excluding open claims did not affect the significance of changes in cost trends for implementing hospitals, but it did for non-implementing hospitals. Therefore, the open-claims problem may have affected the LeCraw team's findings, which arise from a direct comparison between the implementing and non-implementing hospitals.

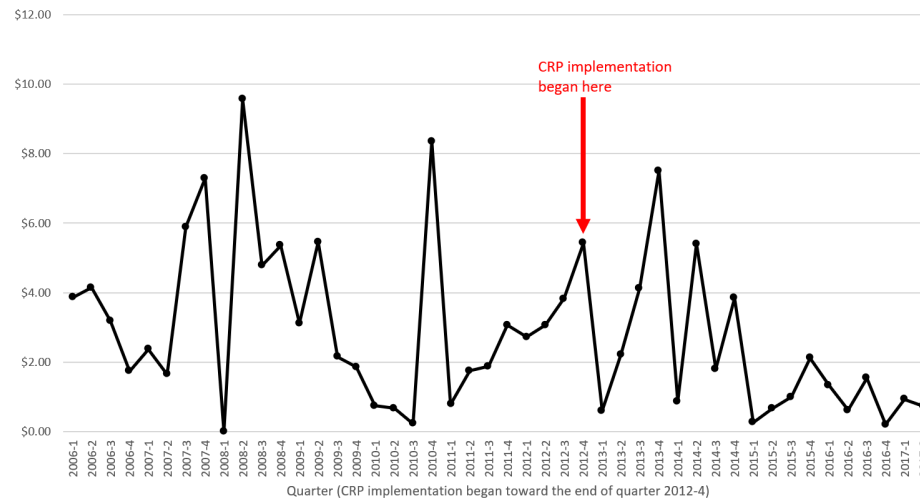
We agree that performing rigorous analyses of the effects of CRPs—using designs that fit the intervention, study questions, and data—is important because of the broad interest in these programs among U.S. healthcare organizations. Unfortunately, the work by LMM in the March 2023 issue of *Econ Journal Watch* is not such an analysis.

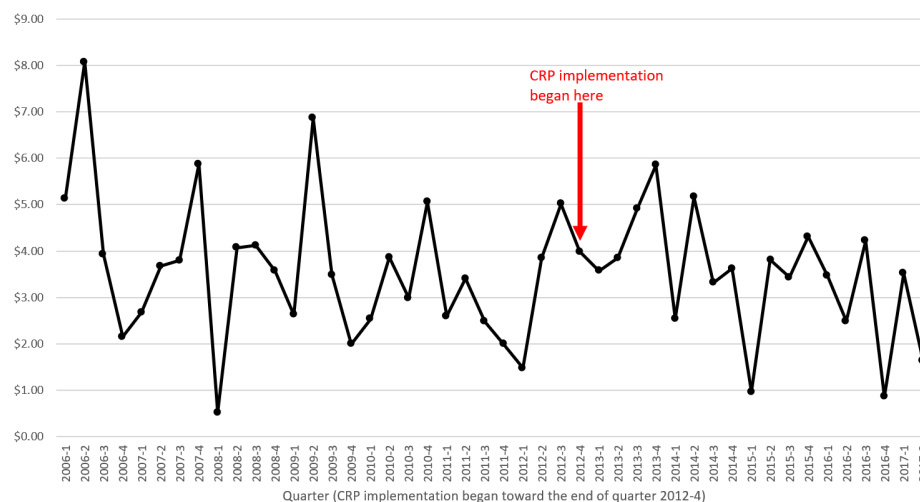


**Figure 1.** Compensation cost rate at AMC1 (unadjusted data)



**Figure 2.** Defense cost rate at AMC1 (unadjusted data)



**Figure 3.** New claims rate at AMC1 (unadjusted data)

## References

- Allison, David B., Andrew W. Brown, Brandon J. George, and Kathryn A. Kaiser. 2016. Reproducibility: A Tragedy of Errors. *Nature* 530: 27–29. [Link](#)
- Bland, J. Martin, and Douglas G. Altman. 2015. Best (but Oft Forgotten) Practices: Testing for Treatment Effects in Randomized Trials by Separate Analyses of Changes from Baseline in Each Group Is a Misleading Approach. *American Journal of Clinical Nutrition* 102(5): 991–994. [Link](#)
- Kachalia, Allen, Kenneth Sands, Melinda Van Niel, Suzanne Dodson, Stephanie Roche, Victor Novack, Maayan Yitshak-Sade, Patricia Folcarelli, Evan M. Benjamin, Alan C. Woodward, and Michelle M. Mello. 2018. Effects of a Communication-and-Resolution Program on Hospitals' Malpractice Claims and Costs. *Health Affairs* 37(11): 1836–1844. [Link](#)
- LeCraw, Florence R., Daniel Montanera, and Thomas A. Mroz (LMM). 2023. Reassessing the Effects of a Communication-and-Resolution Program on Hospitals' Malpractice Claims and Costs. *Econ Journal Watch* 20(1): 1–14. [Link](#)

## About the Authors



**Maayan Yitshak-Sade** is Assistant Professor in the Environmental Medicine and Public Health Department at the Icahn School of Medicine at Mount Sinai, New York. She is also a member of the Mount Sinai Institute for Exposomic Research and Transdisciplinary Center on Early Environmental Exposures. Dr. Yitshak-Sade received her bachelor's degree in nursing, her Master of Public Health, and her Ph.D. in epidemiology from Ben-Gurion University, Beer-Sheva, Israel.



**Allen Kachalia** is a general internist and Senior Vice President for Patient Safety and Quality at Johns Hopkins Medicine. He oversees patient safety, quality, and patient experience outcomes across the continuum of care for the Johns Hopkins Medicine enterprise. He is also the Director of the Armstrong Institute for Patient Safety and Quality, which is focused on improving healthcare for patients through scientific research, innovation, policy advocacy, and education. Allen is Professor of Medicine and the C. Michael and S. Anne Armstrong Professor of Patient Safety and Quality, where his own research focuses on how the law and legal policy affect health care, and in particular how liability system reform and the disclosure of medical error relate to the safety and quality of healthcare. His email address is [kachalia@jhu.edu](mailto:kachalia@jhu.edu).



**Victor Novack** is the founding director of Soroka-BGU Medical Research Institute and Head of the Research Authority at Soroka University Medical Center, the second largest hospital in Israel. Prior to his current position, Prof. Novack served as a Medical Director at Harvard Clinical Research Institute, Boston. He has designed and contributed to the design of more than 50 feasibility and pivotal trials for FDA approval. Prof. Novack is Vice Dean for Innovation and Professor of Medicine at Faculty of Health Sciences at Ben-Gurion University, Director of the Department of Medicine at Soroka University Medical Center, and Senior Scientist at Beth Israel Deaconess Medical Center, Harvard Medical School. He is an Associate and Statistical Editor of the *European Journal of Internal Medicine*. He has more than 300 scientific publications in peer-reviewed journals.

[Go to archive of Character Issues section](#)



**Michelle Mello** is Professor of Law at Stanford Law School and Professor of Health Policy at Stanford University School of Medicine. She conducts empirical research into issues at the intersection of law, ethics, and health policy. She has authored approximately 250 articles on medical liability, public health law, the public health response to COVID-19, pharmaceuticals and vaccines, biomedical research ethics and governance, health information privacy, and other topics. The recipient of a number of awards for her research, she was elected to the National Academy of Medicine at the age of 40. Mello holds an M.Phil. in Comparative Social Research from the University of Oxford, a J.D. from the Yale Law School, and a Ph.D. in Health Policy and Administration from the University of North Carolina at Chapel Hill. Her email address is [mmello@law.stanford.edu](mailto:mmello@law.stanford.edu).

[Go to archive of Watchpad section](#)



# A Rejoinder on the Effects of a Communication-and-Resolution Program on Hospitals' Malpractice Claims and Costs

Florence R. LeCraw<sup>1</sup>, Daniel Montanera<sup>2</sup>, and Thomas A. Mroz<sup>3</sup>

[LINK TO ABSTRACT](#)

There cannot be an effect without a cause.

– Alexander Hamilton, Federalist No. 31

The three pieces that precede the present rejoinder are:

1. Allen Kachalia, Kenneth Sands, Melinda Van Niel, Suzanne Dodson, Stephanie Roche, Victor Novack, Maayan Yitshak-Sade, Patricia Folcarelli, Evan M. Benjamin, Alan C. Woodward, and Michelle M. Mello, “Effects of a Communication-and-Resolution Program on Hospitals’ Malpractice Claims and Costs,” *Health Affairs*, 2018.
2. Florence R. LeCraw, Daniel Montanera, and Thomas A. Mroz, “Reassessing the Effects of a Communication-and-Resolution Program on Hospitals’ Malpractice Claims and Costs,” *Econ Journal Watch*, March 2023.
3. Maayan Yitshak-Sade, Allen Kachalia, Victor Novack, and Michelle

---

1. Georgia State University, Atlanta, GA 30303.

2. Grand Valley State University, Allendale, MI 49401.

3. Georgia State University, Atlanta, GA 30303. The views expressed here are the authors’ and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System.

M. Mello, “Reply to LeCraw, Montanera, and Mroz on Hospitals’ Malpractice Claims and Costs,” *Econ Journal Watch*, September 2023.

It appears that the replying authors agree (at least now) that the characteristics of the data and the circumstances described in the 2018 *Health Affairs* article were wholly inappropriate for advancing a causal interpretation for the original study’s results (Yitshak-Sade et al. 2023, 350). Our primary concern upon reading the original study was misleading statements about causal interpretation. While it is always better to head off such misinterpretation in the original article, the authors should be commended for making the correction now.

In this rejoinder we raise three matters for consideration. First, whether use of causal language, and the inclusion of comparisons to untreated non-CRP hospitals, are likely to induce the audience of *Health Affairs* to infer that the study results are causal. Second, even when examining associations instead of causality, whether large post-treatment spikes in outcomes of interest should be ignored in favor of changes in trend, simply because the spikes are difficult to explain. And, finally, how effectively data sharing agreements insulate original analyses from the scrutiny of reanalyses based on the published information.

## Causal misunderstanding

The original authors say they never meant for the non-CRP hospitals to be used as control groups. As they state in their response to our article, the non-CRP hospitals were simply displayed “as a way of showing whether the changes observed over time at the implementing hospitals also held true at hospitals that did not implement CRPs” (Yitshak-Sade et al. 2023, 350). This seems, to us, remarkably like the role of a control group in causal inference, and so confusion among the audience about causality seems not only understandable but probable. For future reference, in studies where the authors do not wish the audience to compare CRP and non-CRP hospitals, they should avoid framing their research question by “examining before-and-after trends in claims volume, cost, and time to resolution and *comparing them to* trends among nonimplementing peer institutions” (Kachalia et al. 2018, 1836, italics added).

Their reply states “we did not conduct our analysis within a causal modeling framework” (Yitshak-Sade et al. 2023, 350). We agree and appreciate this clarification for the audience. To prevent future misunderstandings, we recommend further rewrites to the original article:

Original article text (all italics added)	Recommended rewrite
“ <i>Effects of A Communication-And-Resolution Program on Hospitals’ Malpractice Claims and Costs</i> ” (Kachalia et al. 2018, article title)	“Hospitals’ Malpractice Claims and Costs Before and After a Communication-and-Resolution Program”
“We evaluated the liability <i>effects</i> of CRP implementation at four Massachusetts hospitals...” (ibid., 1836)	“We study liability outcomes at four Massachusetts hospitals before and after CRP implementation.”
“implementing a communication-and-resolution program <i>does not expand</i> liability risk and <i>may, in fact, improve</i> some liability outcomes” (ibid., 1843)	“We find no increases in liability risk following CRP implementation. There may, in fact, be decreases in some outcomes following implementation.”

We note and appreciate the replying authors’ attempt to clarify in their response to our comment: “Our study’s core conclusion is not that there were significant improvements in liability outcomes associated with CRP implementation, but that hospitals can implement CRPs without experiencing worse liability outcomes” (Yitshak-Sade et al. 2023, 351). If the word “can” is used simply to allow for possibility, then the latter offered conclusion is one we can get behind. But if “can” suggests a promising potentiality, then the latter remains causal; and it does little more than add a double negative, implying something like, “We’re not saying they’ll get a good outcome with CRP; we’re saying they won’t get a bad outcome.” More precise wording could resolve confusion about the meaning of this statement.

## Spikes at implementation

On the post-implementation spikes, we agree that there should be latitude around leads and lags and, notwithstanding clear justification and the scrutiny of the referees, they are at the discretion of the study authors. Discounting the estimated spikes because they are difficult to explain, or seem implausibly large to have resulted entirely from the treatment, is another matter. Could such spikes be indicative of a model misspecification in the original analysis? What if the treatment caused most of the spike? Or some of it? At what degree of the ‘treatment responsibility’ does the spike become problematic to the conclusion? These are a few reasons why many standard modern methods for inferring causality do not ignore level effects. It cannot be ruled out that the spikes have something to do



with CRP implementation, and so they must be accounted for in the confidence intervals. Similarly, as the original authors mention, the existence of open claims are problematic and have a bearing on estimates and their interpretations. We have encountered these issues in our own original research, and understand the challenges, but the solution cannot be the jettisoning of reliable causal inference methods from our studies, particularly when causal language is retained.

## Data darkness ought to raise the bar

As the replying authors mention, agreements disallowing the distribution of data to groups holding original research up to scrutiny are not ideal. Rather than a shield, as the replying authors seem to regard it, we contend that a data sharing restriction places an even greater burden on original studies to acknowledge and accept criticism from reanalysis studies. While the authors (now) do not want the audience to interpret the results as causal, we thought actual causal investigation was warranted given the causal language permeating the original study and its invited comparisons to untreated subjects. In the absence of the original data, we evaluated as best we could, using *their* estimated models' coefficients and some data means,<sup>4</sup> what their estimation procedure revealed when viewed through a causal lens. The resulting interpretation is quite different. The only way our reanalysis could have been improved was if, instead of us estimating the standard errors, the original authors had supplied them to us (which they declined to do). It is important to note that the replying authors do not dispute our calculations using their coefficient estimates. They merely state that "This result strains plausibility" (Yitshak-Sade et al. 2023, 352). We agree, but those calculations are the implication of *their* estimated coefficients. So, while our reanalysis may be "second-best" to an identical one with access to the original data (*ibid.*), it may still offer more reliable

---

4. The approximately 200 percent increase in costs we reported in our original article *follows directly* from the coefficient estimates reported by Kachalia et al (2018) in their Online Appendix. Consider, for example, *their* estimate of the post-implementation "change in intercept" for Compensation Cost at AMC 1 reported in their Appendix Exhibit A.9 (and also in their Appendix Exhibit A.10), which equals 1.240. Since the quasi-Poisson, General Additive Model for their estimations uses a log link function, from just this one coefficient "turning on" in the post-CARE implementation period, the predicted mean Compensation Cost shifts gets multiplied by a factor of 3.46 [ $3.456 = \exp(1.240)$ ]. The impact of this post-CARE intercept shift for Compensation Costs at AMC 1 can be interpreted as an increase in Compensation Cost of 246 percent [ $245.6 = 100 \times (3.46 - 1)$ ]. *Their* reported estimates clearly imply a huge increase in predicted mean Compensation Cost post-CARE implementation, even after adjusting for the "change in slope" parameter they report as  $-0.035$ . The comparable change in Compensation Cost at the comparison AMCs from this "one coefficient" is just 16 percent [ $15.7 = 100 \times (\exp(0.146) - 1)$ ]. See LeCraw, Montanera, and Mroz's (2023) Online Appendix for complete details on how we carried out our calculations.

insights than an initial investigation that plays loose with causal language.

In conclusion, our uncontroversial view is that a study making causal statements must hold up to scrutiny based on standard and accepted causal inference methods. If authors are not prepared to suffer this scrutiny, or do not wish to convey a causal interpretation, then the causal language should be removed and an alternative interpretation of the study's contribution must be reinforced. A study cannot have it both ways. We think it is clear to the audience what happened here with the original study, reanalysis, and rejoinder. If causal inference was not intended, why use the causal language? If only before-and-after analysis was intended, why make a comparison to non-CRP hospitals? Why include the non-CRP hospitals at all? We may never receive answers to these questions. The most important outcome of all this is that, with the assistance of the replying authors, any confusion around causal interpretations of their study seems to have been resolved.

Furthermore, flaws in the editorial process (the details of which are apparently not in dispute) at *Health Affairs*, a major health services research journal, have been brought to light and can be used for organizational learning moving forward. While it may be a small step for academia, it is at least taken in the right direction.

## References

- Kachalia, Allen, Kenneth Sands, Melinda Van Niel, Suzanne Dodson, Stephanie Roche, Victor Novack, Maayan Yitshak-Sade, Patricia Folcarelli, Evan M. Benjamin, Alan C. Woodward, and Michelle M. Mello.** 2018. Effects of a Communication-and-Resolution Program on Hospitals' Malpractice Claims and Costs. *Health Affairs* 37(11): 1836–1844. [Link](#)
- LeCraw, Florence R., Daniel Montanera, and Thomas A. Mroz.** 2023. Reassessing the Effects of a Communication-and-Resolution Program on Hospitals' Malpractice Claims and Costs. *Econ Journal Watch* 20(1): 1–14. [Link](#)
- Yitshak-Sade, Maayan, Allen Kachalia, Victor Novack, and Michelle M. Mello.** 2023. Reply to LeCraw, Montanera, and Mroz on Hospitals' Malpractice Claims and Costs. *Econ Journal Watch* 20(2): 349–356. [Link](#)

## About the Authors



Healthcare System. Her email address is [frlwatts@gmail.com](mailto:frlwatts@gmail.com).

**Florence LeCraw** was a practicing anesthesiologist for over 35 years. In the later part of her career, she specialized in health economics, which she now pursues full time. She collaborates with economists and other specialists to investigate questions about patient safety, medical liability, and healthcare professional burnout. She is an Adjunct Professor at the Andrew Young School of Policy Studies at Georgia State University in Atlanta. She is on the hospital staff at Northside



His email address is [montaned@gvsu.edu](mailto:montaned@gvsu.edu).

**Daniel Montanera** is an Assistant Professor in the Department of Economics in the Seidman College of Business at Grand Valley State University. His research investigates defensive medicine and the impact of medical malpractice liability costs on health care spending and access to care, as well as auction theory in health insurance exchanges. His work has been published in such journals as *Information Systems Research*, the *Journal of Law, Economics, and Organization*, *Manufacturing &*



His research assesses the reliability of estimators of behavioral economic relationships for explanations of health, labor, and demographic outcomes. His email address is [tmroz@gsu.edu](mailto:tmroz@gsu.edu).

**Thomas A. Mroz** holds the Bernard B. and Eugenia A. Ramsey Chair of Private Enterprise and economics professor at Georgia State University's Andrew Young School of Policy Studies. He is a Research Fellow at the Federal Reserve Bank of Atlanta. He develops econometric models for estimating the impacts of household behaviors on dynamic human development, health, and fertility processes that incorporate incomplete information and factors unobservable to the researcher as key determinants of these processes.

[Go to archive of Comments section](#)  
[Go to September 2023 issue](#)



# Rejoinder to Barkowski and McLaughlin

Aaron M. Gamino<sup>1</sup>

[LINK TO ABSTRACT](#)

In the previous issue of this journal, I published “Health Insurance Mandates and the Marriage of Young Adults: A Comment on Barkowski and McLaughlin” (Gamino 2023). That piece by me is a critique of Scott Barkowski and Joanne Song McLaughlin’s “In Sickness and in Health Interaction Effects of State and Federal Health Insurance Coverage Mandates on Marriage of Young Adults,” published in the *Journal of Human Resources* (2022). The previous issue of this journal also featured a reply by Barkowski and McLaughlin (2023). I appreciate that Barkowski and McLaughlin have engaged with my commentary.

In their reply, B&M point out that I provide a minimal explanation for why I think their models are misspecified. Also, they fail to address my primary criticism about their specifications, and offer several justifications for their specifications, justifications that I consider to be inadequate. In this rejoinder, I revisit the misspecification problem and address B&M’s justifications for their model specification.

## The model’s problem restated

Barkowski and McLaughlin (2022) use IPUMS-USA 2000–2015 American Community Survey data and estimate the following “DD-style” model (equation 5 in their paper):

$$Y_{iast} = \beta_1 ELIG_{ast} \times ACA_t + \beta_2 ELIG_{ast} + X'_{iast} \gamma + a_a + \delta_{st} + u_{iast}. \quad (1)$$

---

1. Middle Tennessee State University, Murfreesboro, TN 37132.

The issue with this model centers on the  $ELIG_{ast}$  term, which is an indicator taking a value of one for individuals who reside in a state with a dependent mandate in place at time  $t$  and are age-eligible for the dependent mandate. To clarify the issue, it may be helpful, in assessing state mandates, to consider the use of equation 2 here for the years before the ACA—the ACA term falls out, resulting in the following model:

$$Y_{iast} = \beta_2 ELIG_{ast} + X'_{iast} \gamma + a_i + \delta_{st} + u_{iast}. \quad (2)$$

It is now apparent that this requires a DDD setup to identify the difference in marriage due to state eligibility,  $\beta_2$ . Matching this up with the basic DDD framework, as presented by Jeff Wooldridge (2007), the necessary terms for age-by-state and age-by-year are omitted.

B&M interpret  $\beta_2$  as “the marriage gap before the implementation of the ACA between those who were eligible for state-mandated coverage and those who were not” (B&M 2023, 36). This interpretation is only valid when  $\beta_2$  is identified in a valid framework, which requires additional terms. In the base specification, where only state-by-year interactions are included,  $\beta_2$  is going to be improperly picking up differences that exist at the age-by-state and age-by-year levels. Rather than isolate the difference due to state mandate eligibility, this term estimates that along with unintended biases.

Returning to Equation 1, the addition of  $ELIG_{ast} \times ACA_t$  does not address the problems that  $\beta_2$  suffers from. The interaction of the ACA term and the problematic term  $ELIG_{ast}$  extends this problem to the intended interpretation of  $\beta_1$  as the “difference between the post-ACA and pre-ACA marriage gaps” (B&M 2023, 37). The interaction term instead measures the difference between the marriage gap due to  $\beta_2$ , a difference that does not capture the marriage gap resulting from state mandate eligibility. When the complete set of interaction terms is included, the main results on the likelihood of being married drastically differ and lose significance (Gamino 2023, 19–21).

## Addressing B&M’s arguments for their model choice

### Loss of identifying variation

B&M (2023) make two assertions relating to identifying variation. First, they assert that introducing age-by-year fixed effects absorbs the remaining identifying

variation (B&M 2023, 38–40). The standard errors obtained in the models with and without the proper set of interaction terms are consistent in size, which contradicts the claim that the additional terms absorb too much variation. The standard errors should be larger if the additional terms absorb too much variation.

Second, B&M assert that “the fixed effects are absorbing the identifying variation in one model [marital state], but not in the other [marital entry]” (B&M 2023, 40). At the age-state-year level, the identifying variation is the same for both models—namely, the variation in the mandates. That sameness can be seen by applying Theorem 1 of Andrew Goodman-Bacon (2021) to the DDD-to-DD transformation in Andreas Olden and Jarle Møen (2022). B&M state that “the variation explained in our marriage entry outcome by our DD-style models is only about 15 percent as much as the variation explained in marital state” (B&M 2023, 39). I am not sure what they are referencing by “the variation explained in our marriage entry outcome.” However, if this refers to the R-squared—a measure of how much variation in the outcome variable the model explains—that is not the same as identifying variation.

### **Bias and the “practical implementation problem”**

B&M argue that bias is introduced by the addition of age-by-year interactions and the inclusion of these interactions lead to a “practical implementation problem” (B&M 2023, 35–40). Their argument relies solely on the direction of predicted effects provided in their Table 1, for combinations of age (young or old) and state mandate status (mandate or non-mandate) (B&M 2023, 37). Despite the importance of accurately predicting the signs to justify their model choice, they do not provide a theoretical model. Instead, B&M make predictions based on a single possible mechanism: whether a spouse or parent is a source of better insurance. The marriage decision is complicated, which makes it difficult to have confidence in an argument based on one of many possible channels. In the following paragraphs, I provide cases for effects in directions other than those predicted by B&M.

First, consider the case of an older (ineligible) uninsured young adult residing in a state with a mandate. Suppose this ineligible young adult is in a relationship with a younger (eligible) adult. The decision to marry depends upon both parties. The younger individual’s disincentive to marry is lessened following the ACA because the marriage prohibition is removed. As a result, this type of couple could see an increase in the likelihood of being married. This example is at odds with B&M’s predicted effect for an older individual in a mandate state.

Next, consider the case of an older (ineligible) uninsured adult residing in a state without a mandate. After the ACA, this ineligible individual is insured and

less susceptible to negative wealth shocks from an adverse health event. It is not difficult to imagine scenarios where having increased wealth increases her likelihood of being married (e.g., her preferred ceremony remains within her budget, and she does not delay marriage; or she is likelier to marry now that she has higher levels of wealth<sup>2</sup>). Again, this example is at odds with B&M's predicted effect.

These two possible cases give rise to predicted effects opposite to those relied on by B&M, undercutting the argument justifying their model choice. Furthermore, B&M's expected increase in marriage should increase the likelihood of having coverage as a policyholder or through a spouse. In Table 4 (Gamino 2023, 23), I find no empirical support for the single channel considered by B&M. In their models, I see no changes in the likelihood of being on a spouse's ESI plan or being an ESI policyholder.

## Concluding remarks

In this rejoinder I have briefly restated the criticism in Gamino (2023) that remained unaddressed in B&M's response (2023). I demonstrate straightforwardly how their eligibility term requires a DDD framework for proper identification. I show the deficiencies in their justification for their model choice. Specifically, I point out that the additional fixed effects do not absorb the identifying variation. The weak argument that bias is introduced through additional fixed effects is based on a woefully inadequate consideration of the marriage decision.

## References

- Barkowski, Scott, and Joanne Song McLaughlin (B&M).** 2022. In Sickness and in Health Interaction Effects of State and Federal Health Insurance Coverage Mandates on Marriage of Young Adults. *Journal of Human Resources* 57(2): 637–688.
- Barkowski, Scott, and Joanne Song McLaughlin (B&M).** 2023. Response to Gamino. *Econ Journal Watch* 20(1): 34–44. [Link](#)
- Gamino, Aaron M.** 2023. Health Insurance Mandates and the Marriage of Young Adults: A Comment on Barkowski and McLaughlin. *Econ Journal Watch* 20(1): 15–33. [Link](#)
- Goodman-Bacon, Andrew.** 2021. Difference-in-Differences with Variation in Treatment Timing. *Journal of Econometrics* 225(2): 254–277.
- Olden, Andreas, and Jarle Møen.** 2022. The Triple Difference Estimator. *Econometrics*

---

2. Our hypothetical individual could also be likelier to marry because she has higher health stock following an adverse event post-ACA. Such an example further illustrates that there are too many possible arguments to capture in a simple 2-by-2 table.



*Journal* 25(3): 531–553. [Link](#)

**Wooldridge, Jeff.** 2007. What's New in Econometrics?: Lecture 10 Difference-in-Differences Estimation. Presented at the NBER Summer Institute (Cambridge, Mass.), July 31. [Link](#)

## About the Author



**Aaron M. Gamino** is an assistant professor in the Economics and Finance Department of the Jennings A. Jones College of Business at Middle Tennessee State University. He received his Ph.D. and M.A. in Economics from Vanderbilt University. He received his B.A. in economics and philosophy from the University of Illinois at Urbana-Champaign. His interests include demography, health, history, and labor. His email address is [aaron.gamino@mtsu.edu](mailto:aaron.gamino@mtsu.edu).

[Go to archive of Comments section](#)  
[Go to September 2023 issue](#)



# A History of Classical Liberalism in the Netherlands

Edwin van de Haar<sup>1</sup>

[LINK TO ABSTRACT](#)

The Netherlands has a reputation as a liberty-loving, free trading country. In some respects, classical liberal ideas about the economy and personal freedom seem deeply engrained in Dutch culture, from the Frisian Freedom in the Middle Ages, to the embrace of commerce before the Dutch Republic, and, down to recent times, to gay marriage and liberal attitudes about personal choice in sex and drugs. In other respects, however, the Dutch story is one of proto-liberal leadership and great promise up through the 17th century, and then a loss of classical-liberal footing and indeed a rather sorry showing through the 20th century and up to today. This article will sketch the Dutch story over many centuries.

Liberalism as a political outlook only achieved self-conscious coherence in Europe from the 18th century onwards. For classical liberalism, however, freedom of religion, conscience, commerce, and personal lifestyle are but expressions of the basic idea of freedom from coercive restrictions imposed by government. Adam Smith spoke of “allowing every man to pursue his own interest his own way, upon the liberal plan of equality, liberty, and justice” (Smith 1976/1776, 664). Such a presumption of liberty is characteristic of classical liberalism, as understood in the present paper.

With this study, I do not mean to suggest that things told of here led by necessity to the development of classical liberalism. Following Quentin Skinner (2002) and J. G. A. Pocock (1989), I acknowledge that ideas should be seen in the context of the time and circumstances they were put forward. At the same time, some ideas are of a perennial nature, even when they are not completely stable

---

1. I would like to thank Patrick van Schie, Erwin Dekker, and the anonymous referees for comments and suggestions on earlier versions of this article.

in meaning. The 17th-century idea of freedom of conscience is not the same as our modern freedom of religion. Yet the core idea has remained recognizable and fairly stable in meaning, namely the idea that an individual should be allowed to decide for his or herself about religious belief (Van de Haar 2009, 14–15; 2015, 17–18). Zooming in at the specific Dutch historical circumstances, I follow Marius Wessels’ suggestion of the existence of a ‘Dutch tradition of freedom.’ He emphasizes that before the 18th century, when liberalism became a more or less coherent doctrine, a number of developments in economics, politics, and political and moral philosophy can safely be labelled ‘proto-liberal’ (Wessels 1998). Still, thinkers who advanced ideas that appear proto-liberal may have had aims and sensibilities that differed from those of Smith and other expositors of “the liberal plan.”

Compared to other variants of liberalism, classical liberalism stands out for its realistic view of human nature, in which both reason and emotion have a place in the explanation of human behavior, but the first cannot always subdue the second. Classical liberalism entails negative individual freedom, or the entitlement of individuals to a large private domain, in particular vis-à-vis the state. Associations should be voluntary, while classical natural rights, in particular those to life, property, and liberty, are critical for the preservation of individual liberty. These natural rights entail freedom of speech, the press, religion, association, et cetera. The state only has a small number of tasks (judiciary, defense, some public goods), while societal order must depend to a large degree on spontaneous ordering processes, such as the free market. Governments are bound by the rule of law, based on constitutional limits to their power.

Another liberalism spoken about in this article is social liberalism, a 19th-century variant, which shares some affinity with classical liberalism, for example in abstaining from the endorsement of outright socialism, communism, and fascism. But it is at variance with classical liberalism in abandoning the presumption against the governmentalization of social affairs, or at least abandoning the breadth with which classical liberalism upholds that presumption. Social liberals are far less opposed to government intervention in private lives and the economy, and rather supportive of extended social welfare arrangements. Sometimes I use ‘liberalism’ to cover both variants (for more detail see Van de Haar 2015), but my focus here is on classical liberalism.

The historical development of the country now known as Kingdom of the Netherlands is central to this article. It has of course seen many geographical and border changes throughout the ages. It has been part of the German Holy Roman Empire, the lands of the House of Burgundy, and the Habsburg Empire. For long periods it included the Southern Netherlands, which now covers Belgium, Luxembourg and some parts of northern France, and between 1815 and 1830 only the current Belgium area. The northern part of the province of Brabant became

Dutch after the Peace of Westphalia in 1648, while parts of the province of Limburg were also part of the German League until its demise in 1867 (Andeweg et al. 2020, 1–8; Kennedy 2017, 3–6). It goes beyond the purposes of this article to account for all these different situations.

Proceeding chronologically, I start around the year 1000 and end in our time. The story attempts to distinguish the main economic, social, philosophical and political developments in the Netherlands related to classical liberalism and to a lesser extent social liberalism.

## Middle Ages to 1550

The foundations for Dutch political culture and economy were laid in the late Middle Ages, as Maarten Prak and Jan Luiten Van Zanden (2023) argue. An important reason for the Dutch tradition of freedom lies in the geographical fact that the country is located in the northwest of Europe, where rulers from Austria, Germany, and France could less easily exercise control. It was also swampy, sandy, and generally less attractive land to hold or occupy. Being peripheral to the centers of great power led to the growth of relatively strong local towns and regions, which would evolve into the Dutch provinces. In the center and the west of the country, an important feature was the draining of wasteland, done at first to harvest peat for heating and later to use the drained land for farming. Local rulers, including the Bishop of Utrecht, accorded rights to developers. The laws included provisions for future taxes to be paid to the rulers. Local communities emerged after the reclamation works, which set up local drainage or dike boards (*beemraden*), which still exist. In return for the duties paid, elected representatives had a say in affairs. During 1000–1350 the Netherlands saw a relatively strong civil society consisting of guilds, draining boards, and autonomous villages and cities, where regular meetings and elections were established.

Feudalism also played an important role in the development of the Netherlands. The trust-based reciprocity between lord and vassal was more flexible than one might suppose, and it fostered cooperation to grow the pie. People could adapt to changing circumstances. Feudal times saw an explosion of agricultural produce (dairy and cattle farming) and peat, which led to rapid population growth. The Church also played a role, as it ensured some European unity in norms and values and was an economic actor as well. Utrecht was the seat of the Archdiocese, which attracted luxury trade. Another factor was church construction. From the fourteenth century, the Netherlands caught up with and surpassed other European countries in the building of churches. The cities in the eastern part of the country were far more important than those in the west. Trade with German cities on

the rivers Rhine and IJssel was a source of richness in cities like Tiel, Deventer, Zutphen, Zwolle, Harderwijk, and Hattem (some of these became part of the Hanseatic League), while the North and East Seas allowed for international trade abroad (Prak and Van Zanden 2023, 14–57).

In Friesland, which included the lands north of Amsterdam, which was only a minor village at the time, there was no feudal tradition. The land was owned by ‘free farmers’ who largely governed and defended themselves, traded along the German and Baltic coasts, and had their own silver and gold coins. The arrangement and the period are known as the ‘Frisian Freedom.’ The Counts of Holland also stimulated the development of cities, by granting special rights of self-rule and autonomy, such as in Dordrecht. One effect was that the influence of the more local nobility decreased and that central political authority was weak at best. The guilds had some political influence (more so in the Southern Netherlands), although they could be a protectionist force in the economic sense. In the fourteenth century, as in other countries, when rulers needed money to wage wars, they called upon representatives of the regions (‘States’), which led to the development of the States-General, another rudimentary form of civil influence on state decision-making (Wessels 1998, 9–20).

In the late Middle Ages (1350–1566) the western part of the country—Holland and Zeeland—became more important, politically and economically, although Flanders and Brabant remained most important. This period saw a rapid urbanization, made possible by the improvement of the waterways, which enabled the growth of cities such as Amsterdam, Haarlem, Leiden, Delft, Gouda, and Rotterdam. Dairy production needed fewer people, and the products traded for wheat from France and the Baltics. The surplus of workers moved to the cities. Around 1500, 45 percent of Dutch people lived in cities, a high figure compared to other European countries. Those remaining on the land were not all small farmers, as about half of them worked as salaried workers, for example in textiles, fisheries, and digging peat. In 1514 the share of agriculture in the income of Holland decreased to less than 20 percent, produced by a quarter of the work force, while fisheries and peat cutting accounted for 15 percent of the work force. Prak and Van Zanden say that in this period “capitalism was born,” although we are of course looking at developments of a number of factors, in different paces. The capitalist development was based on trusting market relationships to continue to provide food and other primary goods, not least through trade in textiles, beer, peat, and fish. Dutch ships already dominated trade routes on the East Sea. Importantly, property rights were well protected, while capital markets functioned well, with interest rates decreasing from 12 percent in 1350 to 5–6 percent in 1450 (also see McCloskey and Nash 1984). Villages were free to develop economically, as they were largely independent from adjacent cities. Politically, there was balance

between the several cities, and all local rulers had to share power. In short, a modern market economy developed, with commercialization of production, labor, capital, and agricultural land (Prak and Van Zanden 2013, 25–76).

In this context there was a growing demand for human capital. The relatively high literacy of the Dutch population was another crucial factor to its economic and political development. Important was the influence of the movement Brethren of the Common Life (also known as Modern Devotion), a religious community founded by Geert Grote in Deventer in 1374. Although the movement operated within the bounds of the Roman Catholic Church, Grote was dissatisfied with moral decline in the church. To him religion was a personal matter, which meant that individuals should be able to read the Bible and other books in the vernacular. Hence the Brethren first copied and later printed books and texts, and founded Brethren houses, schools, monasteries, and communities. It was ‘education for all,’ all over the country, and in parts of Westphalia as well. This led to increased economic development, not only in Deventer but throughout the Netherlands, which had lasting economic effects (Akçomak et al. 2016).

Jumping forward a century we must highlight the famed thinker Desiderius Erasmus of Rotterdam (1467–1536), who spread a number of ideas now associated with liberalism. Erasmus was associated with Christian Humanism and went to the famous Latin School in Deventer. His best-known book, *The Praise of Folly* (1511), was a critique of many things—of national pride, of people who think too much of themselves, of scholastics, and of the wrongdoings of the church, the monasteries, and popes in particular (Russell 2001, 543–548). Erasmus believed in the individual’s capacity to improve herself. He wrote many pamphlets on the importance of education and in his politics called for arbitration instead of war. Humans had a talent for piety and had the moral duty and actual capacity to do good, even when their power was limited and dependent on divine grace (Rummel and MacPhail 2021). Erasmus broaches the Reformation. In the balance of authority between church hierarchy and scripture, Erasmus lightened the first with his criticisms and added heft to the second with his new Latin and Greek editions of the New Testament.

The foundations for liberal commercial society laid in this period thus consisted of a mix of ideas, practices, and customs. Political power was dispersed, which led to the development of several social institutions, with input of different groups such as the church, guilds, farmers, and laymen. Economically, trade was of major importance already, fostered by dairy farming and shipbuilding abilities. Dutch human capital also developed early, fostered by the work of Geert Grote and Erasmus, who also provided a philosophical base with a degree of liberty at its center.

## Dutch Golden Age (1550–1700)

On these proto-liberal foundations the Dutch republic became the most powerful and influential country on the globe, and remained so for about a century. Trade, empire, innovation, and, compared to other countries, large degrees of personal, societal, and religious liberty were the most important underpinnings.

In 1568 began the Dutch Revolt against the Spanish rulers, a struggle only fully resolved in 1648, at the Peace of Westphalia. The Revolt was directed against the harsh Spanish Catholicism of Philip II. His attempt to lay down strict Catholic rules met with fierce resistance, by Protestants but also by moderate Catholics. Freedom of conscience and religion was a major point in a 1579 treaty among the Dutch provinces, the Union of Utrecht, although the freedom of religious service was still a point of contention. Besides religion, the other main point of contention was a number of new taxes imposed by the Spanish, which were often used for warfare against other powers such as the Ottoman Empire. On 26 July 1581, the States General of the Northern Netherlands moved further, and declared themselves to be free from the rule of Philip II, due to his unjust rule, in the *Acte van Verlatinghe* (Act of Abjuration), nearly two centuries before the American Declaration of Independence of 1776. The Dutch sovereign state was born, and it was based on the strong desire for liberty in multiple senses of that term (Wessels 1998, 34–43).

Despite continued warfare with the Spanish, the new Republic quickly became a great success. The Dutch became an economic powerhouse for almost a century, dominating trade in Europe and in many other parts of the globe as well. Indeed the proceeds made the war possible. Amsterdam became the leading port of the Baltic Sea trade, trading specialized goods from all over for Eastern European grain and wood. In 1583, 84 percent of goods shipped out of Danzig, and 73 percent shipped into that most important Baltic trading city, were transported by Dutch ships. Atlantic coastal trade was also strong. In 1590, the Spanish themselves had to end their embargo on Dutch trade, because they were too dependent on trade in wood and grain. Dominance in the herring trade was also relevant. Hence, the Dutch Golden Age was built on free commerce: importing goods, refining or refinishing them for sale and exporting them again. Important for this success were a number of factors: immigration of labor, especially from the Southern Netherlands (Antwerp), which continued under Habsburg rule, but also Jews from Portugal and later from Germany and Scandinavia; urbanization, with Amsterdam becoming by 1670 the third-largest city in Europe after Paris and London; financial investment; a free, individualized business climate; and technological innovation, such as the wind-powered sawmill, which was crucial for the production of cheap



and easy-to-build *fluyt* merchant ships, produced by the hundreds annually (Kennedy 2017, 142–146).

The economic boom was enabled by financial innovation. Besides the first stock exchange, two Italian-inspired governmental interventions saw light in Amsterdam: a public exchange bank, called *Wisselbank* (with the municipality as guarantor), and a public credit bank (*Bank van Lening*), both forerunners of today's central banks. They were meant to combat instability arising from speculation and the manipulation of exchange rates, and also to control and stabilize the monetary and financial systems. The Dutch East India Company (VOC), from 1602 onwards, can be seen as one of the first modern companies in the world: a limited liability company, with shares traded on the stock exchange, speculation on these shares, and a division between ownership and management, which led to all kinds of conflicts (see Gelderblom and Jonker 2004). The VOC needed and attracted huge amounts of money, used for building trading posts, harbors, fortresses, infrastructure, and so on. It remained in operation for over 200 years, being a stable, well-funded global commercial enterprise. It was dominant in large parts of the trade between Asia and Europe and between Asian ports. The main traded commodities changed throughout these years: from spices to Indian textiles, followed by coffee and sugar from Java and tea from China. The major factor in the slow but certain downfall of the company, which took almost the whole of the 18th century, was that profits were no longer invested but paid out in dividends to the shareholders (Prak and Van Zanden 2023, 90–143).

However, it must be underlined that neither the VOC nor the West Indies Company were classical-liberal highlights. They used slavery in Asia, South Africa, the Caribbean, and South America, and abused and murdered many people there. So, while enjoying freedom at home, they operated on the basis of unfreedom abroad.

Compared to other places in Europe, England included, the Dutch Republic knew a relatively large freedom of opinion and expression. However, “virtually nowhere, not even in England or Holland after 1688, was full tolerance the rule and hardly anyone subscribed to the idea that the individual should be free to think and believe as he or she thought fit” (Israel 2001, 17), and the Catholic faith was now officially forbidden (although tolerated by the authorities after specific payments were made), and members of faiths other than the Dutch Reformed Church could not hold official public positions. There was less censorship than in other places, and Dutch publishers supplied northern Europe, especially France, with forbidden books that had to be smuggled in. The Dutch Republic was also the center of (French-language) learned journals, which were important carriers of cultural and intellectual change, from the late 17th century onwards. Of the nearly 30 learned journals with international standing in Europe in 1746, two were based

in France, several in Germany and Italy, one in England, and no less than 18 in the Netherlands. The Republic was also a refuge for thinkers, such as René Descartes, Pierre Bayle, and John Locke. In terms of intellectual tendency, the Netherlands became the chief source for the spread of Cartesian ideas and mechanistic thinking around northern Europe. The high prestige of Dutch universities attracted many foreign students (Israel 2001, 23–58, 104, 116–118, 149, 295–327), not least from Scotland.

## **Dutch (proto-)classical liberal thinkers**

This boom in intellectual activities is also reflected in Dutch contributions to the development of (proto-)liberal thought. The most famous Dutch liberal thinkers, or at least those who introduced, defended, or fostered ideas that would become part of classical liberal thought, lived in and around the Golden Age. Although many more Dutchmen participated in the lively public debate in Europe in the 16th and 17th centuries, the five best known to us are Grotius, Spinoza, Bernard Mandeville, and to a lesser extent Pieter and Johan de la Court. Spinoza and Grotius were most influential, and of importance for the early modern foundations of classical liberalism (Collins 2011). All five wrote pleas for greater individual freedom, including free trade, economic liberty, and personal liberties, not least of conscience. It made them controversial, if not infamous, among contemporaries. Hence Doug Den Uyl's remark (1987), that "historically Spinoza and Mandeville have at least one thing in common: their writings caused such a furor of controversy that one would have thought the whole fabric of Western civilization was jeopardized by their work."

Generally, Dutch thought of this period was a hymn to freedom, also in the works of the lesser thinkers not discussed here. In the 17th century, the Dutch already laid out all the essential political ideas of the Enlightenment of a century later (Kossmann 2000, 128–129).

This section briefly introduces Grotius, Spinoza, Mandeville, and Pieter and Johan de la Court, with an emphasis on the liberal aspects in their writings.

### **Hugo de Groot (Grotius) (1583–1645)**

Scholar, advocate, politician, refugee, and diplomat Hugo de Groot, better known internationally by his Latin name Grotius, wrote about a number ideas central to liberal thought, not least natural rights, natural law, and free trade. His writings on international law also had a direct effect on the thought of David Hume and Adam Smith (Van de Haar 2008; 2009, 41–74; 2013a; 2013b).

Born in Delft, Grotius' extraordinary talents were discovered at young age, entering Leiden University at the age of 11. King Henry IV of France even called him "the miracle of Holland." Nowadays, Grotius is perhaps best known for his work *De Jure Bellis ac Pacis* (1625), the declared purpose of which is to treat justice between nations, but to do so Grotius first discussed justice between individuals, so the massive work covers much more than international law. He also published many treatises and books throughout his life, not least in theology. His earlier book *Mare Liberum* (1609) discusses the limits of the rights of sovereigns to restrict travel and shipping on open waters. His first years he spent as advocate-



Hugo Grotius (1583–1645)

fiscal, before he quickly became an important politician in the Dutch Republic. He was pensionary, the most influential official, in Rotterdam, and at the national level he was a staunch ally of Johan van Oldenbarnevelt, the most influential politician of the country. Both were ousted in 1618, when the Counter-Remonstrant party succeeded in getting to power. Van Oldenbarnevelt was beheaded, but Grotius was sentenced to life imprisonment and confiscation of property. In 1621 he fled from his prison, Loevestein Castle, famously hidden in a book chest, and he became an exile in Paris. This was not the happiest period of his life, although he succeeded in entering the learned circles in Paris. His hopes for a return to the Netherlands were dashed time and again, and in 1634 he became the Swedish Ambassador to Paris. He was dismissed in 1645, and died on his way back from Stockholm, in the German city of Rostock on 28 August 1645 (Lesaffer and Nijman 2021, 17–87).

It is said that "no history of the rise of individual rights can be told without Grotius" (Somos 2021, 113). David Schmidtz and Jason Brennan (2010) emphasize that he saw rights as bound to a person, not to a property or a relation, such as medieval serfdom had been. Liberty was an inalienable property belonging to individual men, and people do not have the right to give up that liberty by placing themselves in bondage, hence they should not be allowed to give up freedom for slavery. Grotius built the philosophical foundation of liberalism because, as Schmidtz and Brennan put it, he argued that "the legal idea of a right was also an infrastructure of moral thinking about how a person ought to be treated." By birth, an individual had rights, to life, limb, and liberty. This is the idea of natural law, to be respected by everyone, including legislators, either Christian or non-Christian. Grotius did not argue for secular natural law himself, but he laid the groundwork for that idea by extending natural law from the realm of theology to that of philosophers and lawyers. He claimed his theory would be valid "even if

we were to grant what we cannot grant without the greatest wickedness, namely that there is no God, or that human affairs are of no concern to him” (Schmidt and Brennan 2010, 106–109). In his political theory, Grotius attempted to prove, by going back to the Batavian times, that sovereignty resides in the hands of the people, not in the king, prince, or stadtholder (Weststeijn 2013, 30).

Grotius is also famous for fostering free trade, which was a “major issue in the political and economic debate—and warfare—between England and the Dutch Republic throughout the 17th century, from Grotius’ *Mare Liberum* onwards” (Weststeijn 2012, 227–228). *Mare Liberum* (*The Free Sea*) made a splash, but it was originally a chapter of *De Jurea Praedae* (*The Law of Prize and Booty*), a book that was only discovered in the 19th century. Influenced by Spanish legal scholar Francisco de Vitoria, Grotius’ goal was to make a positive case for the Dutch being able to trade in southeast Asia, opposing the Portuguese claim that they held property of the East and thus were within their rights to exclude the Dutch from entering that area. This was a violation of the fundamental right to preserve oneself, Grotius argued (Fitzmaurice 2021; for more detail see Armitage 2004).

## Pieter (1618–1685) and Johan (1622–1660) de la Court

Pieter de la Court and his brother Johan were wealthy cloth manufacturers from Leiden. They moved in republican circles and published a number of treatises, often in the vernacular instead of Latin. They openly criticized the monarchy, fanatically calling for a republic without the House of Orange, equating monarchy with tyranny. Besides prominence in the Netherlands, they also gained international fame, influencing such diverse thinkers as Samuel von Pufendorf, Gottfried Wilhelm Leibniz, Jean-Baptiste Colbert, Marie De Gournay, Anne Robert Jacques Turgot, Algernon Sidney, and their fellow Dutchmen Spinoza and Mandeville. After Johan’s death in 1660, Pieter used and built upon his work, making it hard to distinguish between the ideas of one or the other (Weststeijn 2012, 65–68, 349–357).



Pieter de la Court  
(1618–1685)

The brothers De la Court argued that sovereignty in a state always originates in the people, going back to the agreement people reach when leaving the state of nature (Weststeijn 2013, 65). In 1661, Pieter de la Court published *The Interests of Holland* (it also contained two chapters by Johan de Witt, the most powerful contemporary politician, who took a keen interest in de la Court), and the book was

an immediate bestseller. But it was also very controversial, leading to disciplinary measures by church and state. One year later Pieter revised and expanded the book, changing its title to *Political Maxims of the State of Holland* (*Aannysing der heilsame politike gronden en maximen van de republike van Holland*). The book provides a mix of religious, political, and economic arguments. For example, while writing against sovereign princes and monarchs, he argued in favor of maximum liberty for the population—although, as was custom those days, that foremost meant the educated male part of the population. The true interest of any state was the joint welfare of the governor and governed. Good leaders will aim to expand the public welfare and will recognize that their population cannot be commanded around like horses. “Where there is liberty, there will be riches and people,” de la Court said in his defense of “a free commonwealth government.” The people thrive because freedom results in the growth of commerce, manufacturing, fishing, arts, and the population. Free trade and free fisheries are related, and they were the cornerstones of Dutch prosperity, as was the prohibition or limitation of guilds, monopolies (such as the West and East Indies Companies), and protectionism in general. Referring to his brother’s *Political Discourses* (1662), Pieter argued that the economic success of Holland also depends on the number of inhabitants. To attract foreigners it was crucial to have freedom and toleration of religious service, but also freedom of occupation: “strangers without freedom of earning their bread and seeking a livelihood cannot live amongst us” (de la Court 2003).

Commerce was the means for the preservation and increase of the polity, and for it to thrive there was a need for liberty. De la Court was rather radical in this. Liberty of trade, occupation, and enterprise, and also immigration, lead to commercial greatness for a country. Trade monopolies should be abandoned, such as the Dutch East India Company (VOC), whose privileges the de la Courts tried to overturn. Humans should enjoy the greatest amount of natural liberty within the boundaries of the law, including the freedom of religion, study, trade, manufactures, arts, and citizenship. The greatest degree of freedom, including low taxes, makes a city or place attractive to immigrants with knowledge and goods, and makes the city or place competitive as a result. The concept of liberty in the thought of de la Court includes individual freedom as non-interference and independence from arbitrary domination (Weststeijn 2012, 224–237).

### **Baruch de Spinoza (1632–1677)**

Spinoza’s family were Sephardic Jews. His ancestors fled from the Spanish Inquisition and moved to Amsterdam. Aged 24, he was excommunicated for no longer observing Jewish standards, rejecting the Jewish-Christian dogmas, and, worst of all, spreading his thoughts. He was banned from the Amsterdam Jewish



Quarter, and then lived in Rijnsburg, near Leiden (making lenses for microscopes, besides his scientific activities), Voorburg, and the Hague, where he died at the age of 45. Spinoza kept a wide scientific network all over Europe, including Henry Oldenburg, the secretary of the English Royal Society, and Gottfried Leibniz. Spinoza's principal works are the *Ethics*, *The Principles of Cartesian Philosophy*, and his *Theologico-Political Treatise* (Feldman 1992, 1–5; Scruton 2000, 9–28).

Spinoza was the origin of much debate and ultimately change of opinion in the period 1650–1750, and all over Europe, although Hobbes was more influential in Britain. Spinoza was widely seen as an atheist. He denied the existence of miracles—a major issue at the time—and was a radical free thinker. Spinoza believed that motion is inherent in matter and that Nature is self-moving; he rejected divine providence and the idea that a God governs man's destiny. Spinoza also argued in favor of natural liberty, which included individual freedom, liberty of thought, and radical toleration. He was the first major European thinker to embrace democratic republicanism, including political freedom for all citizens (Israel 2001, 159–294; 2007, viii–ix). He is seen as the



Baruch de Spinoza  
(1632–1677)

first modern thinker (as opposed to medieval), even more so than Descartes, because Spinoza “cut the ties with religious tradition as a source of information, instead relying upon natural means to arrive at the philosophical truth,” although he was no atheist. In his view of human nature he saw the passions and reason as two more or less equal sources of human conduct, without one necessarily being superior over the other (Feldman 1992).

Spinoza's ideas on economics are not well-known, yet some point in a classical liberal direction, albeit with important exceptions. On the one hand, Spinoza saw commerce and movable wealth as benign, because they foster interests that are either interdependent or require the same means for their furtherance. On the other hand, Spinoza thought property of real estate should be in the hands of the state, to avoid unresolvable disputes and unextinguishable envy (Spinoza 2000, 67–68, 80). Spinoza's politics and economics are interdependent. Order is needed for economic prosperity, as it will foster higher productivity through cooperation, specialization, and the division of labor, certainly when compared to the anarchy of the state of nature. Money is helpful in that it is mobile and can give access to any concrete good. A money-based system is dynamic and cooperative, he held, while a land-based one is static and antagonistic. The state needs to provide security and freedom of trade and contract, while the market will pacify the natural rivalry

among people. A harmony of interests will develop in commercial society, if the citizens can achieve their income from commerce. Yet if the state overregulates property rights, this will destabilize the whole order, as Spinoza said: “he who seeks to determine everything by law will aggravate vices rather than correct them. We must necessarily permit what we cannot prevent” (Wagener 1994). Hence economic well-being depends on liberty and order.

Democracy and liberty enabled people to live together in relative peace and harmony, despite religious and other differences. The Bible should not be read literally, nor should one opinion be imposed on all. The Scriptures and the writings of the Jewish prophets were helpful to teach people faith and benevolence, but were not sources of truth, which could only be found in reason (Schmidt and Brennan 2010, 109–111). In his politics Spinoza should be considered as an evolutionary theorist, who (perhaps paradoxically) respected the lessons from practical politicians more than the abstract ideas of philosophers, as would Mandeville after him. Human nature was ruled by the passions, although Spinoza did not exclude the possibility of a life of reason. His thought (and Mandeville’s) remains an effective rebuttal to rationalistic enthusiasm in politics and social theory (Den Uyl 1987).

### **Bernard Mandeville (1670–1733)**

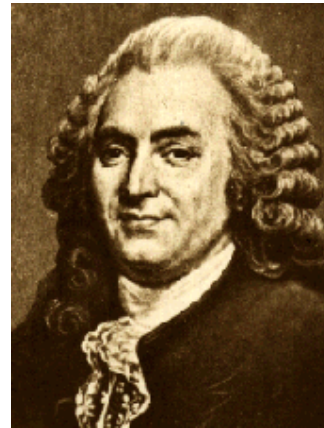
Born in Rotterdam and educated in philosophy and medicine in Leiden, Bernard Mandeville moved to London in his early twenties and would stay there for the rest of his life. A practicing medical doctor in what today we would call psychiatry, ‘The Dutch Doctor’ also published in philosophy, ethics, economics, and social, political, and religious commentary. Yet he did not build a coherent philosophical system. He became well-known, if not notorious, among contemporaries in England and the broader Anglo-Saxon world, far more than he was (and is) in his home country. Mandeville fought against hypocrisy of the clergy and the population at large, especially on topics of morality and sexuality. He even published a defense of brothels and prostitution, criticizing attempts to prohibit them and the double moral standards by lawmakers and people in the London Society for the Improvement of Morals (Mandeville 2006; Willemsen 2022, 13–66; Jansen 2006). He also famously argued that passions often regarded as negative, such as pride, selfishness, and lust, have an upside to them (Willemsen 2022, 67–96).

Mandeville is best known for his *Fable of the Bees, or Private Vices and Publick Benefits* (Mandeville 1988), published in two parts, in 1714 and 1729 respectively. It suggested spontaneous ordering mechanisms, as later developed by Adam Ferguson, David Hume, Adam Smith, and Friedrich Hayek, who praised Mandeville



exactly on this point, noting that “Mandeville did not show exactly how an order formed itself without design, but he made it abundantly clear that it did” (Hayek 1978, 249–266). Mandeville’s notion of private vices and public benefits did not mean, as Mandeville’s contemporaries argued, that everything that is vicious is also beneficial. Crimes should be punished. The only vice to be encouraged is useful vice. So, not all evil was a public benefit. The exact boundary was of course hard to establish. The growth of society and morals was gradual and evolutionary (Kaye 1988, lx, lxvi). The state originated in the desire for protection against wild animals, against protection from dangerous humans, and the gradual development of language as a basis for laws and regulations for a particular group of people (Willemsen 2022, 99–101).

Mandeville’s views on economics were well-known among contemporaries, not least because he defended luxury, which was rather uncommon. Luxury did not necessarily corrupt people nor was it the result of the waste of resources. Mandeville argued that national frugality was the result of certain economic conditions, which had nothing to do with morality. Great states and luxury were related, through production, trade, and commerce. The safeguarding and fostering of trade were prime interests of the state, and the inevitable result would be an increase in luxury. That many people objected to luxury but were fully engaged in fostering trade was a contradiction, Mandeville suggested. Luxury was another case of private vices and public benefits.



Bernard Mandeville  
(1670–1733)

The other main aspect of his economic thought was his strong defense of free trade, both domestically and internationally. Different from other English and Dutch writers on the issue who focused on state welfare, he emphasized that the selfish good of the individual would also be beneficial to the state (Kaye 1988, xciv–ciii). In his last work, *A Letter to Dion* (1732), Mandeville points at weaknesses of the market system but still is a powerful advocate of it: “to this emulation and striving to outdo one another it is owing, that...there is still a plus ultra left for the ingenious; it is this, or at least the consequence of it, that sets the poor to work, adds spurs to industry, and encourages the skillful artificer to search for further improvements” (quoted in Prendergast 2016, 121). Mandeville’s ideas fit with the classical-liberal preference for limited government and limited public institutional interference in the economy (Prendergast 2016; Van de Haar 2015). In *A Search into the Nature of Society* (1723) he described the working of the division of labor, a term he coined in his sixth dialogue. Mandeville always emphasized the economic role of

government (“a capable politician”), especially the protection of property rights. Yet he emphasized that individuals had a right to make their own (economic) choices, aiming to satisfy their self-interests and earn profits. While he was no economic theoretician, and also made a number of mercantilist remarks (for example on the balance of trade), there can be no doubt that in the main he contributed to the developing Dutch and English laissez-faire tradition (Willemssen 2022, 121–136).

It has been noted that Mandeville exercised influence on the ethical and economic thought of Hume (see, e.g., Mossner 1980, 49, 74; Harris 2015) and Smith (see, e.g., Schliesser 2017; Hanley 2016; Ross 2010). This is not to deny they also took pains to distance themselves from him, due to his notoriety but also because they attempted to offer a fuller moral theory. Still, they shared Mandeville’s emphasis on the importance of the selfish elements in human nature.

After 1730 no Dutchmen made any additional major philosophical contribution. This is not to deny that minor contributions were made, or that modern academics added to their fields of specialization. However, in general the Dutch transformed from thought leaders to followers of ideas and events initiated elsewhere.

## Liberalism droops (1700–1840)

After 100 years of richness and power, the Dutch Republic started to lose military, political, imperial, and economic power. Yet this should be seen in perspective: the Dutch remained fabulously rich, with one of the highest incomes per capita, and continually improving the quality of life of its citizens (McCloskey 2019, 228). It took other Western European countries until 1870 to catch up with Holland, measured in average income per head (McCloskey 2011, 194). Still, from 1672 onwards, the cities decreased in number of inhabitants, and the housing and art markets turned downward. Due to the expensive wars, the public debt had risen to unsustainable levels, with the Dutch state effectively going bankrupt in 1715. During the ensuing decades, political decision-making was slow, tax rates high, and social arrangements, such as publicly funded poor relief, were cut (Prak and Van Zanden 2013, 152–165). A stationary Dutch economy, as Adam Smith put it, was therefore predominant in the 18th century, even if it was on a relatively high level of wealth.

Intellectual input came from abroad now. Yet the great classical liberal thinkers were not very popular. The writings of the Scottish Enlightenment were generally welcomed, as the Dutch liked their political and moral moderation, as well as the focus on moral-philosophical questions. Yet the thinkers most famous to us

were not the most popular then. None of the writings of Francis Hutcheson, only a few writings of Hume (*The History of England* and the *Political Discourses*), and just a part of the *Wealth of Nations* were translated into Dutch, although the learned part of Dutch society would be able to read these authors in French. Most often they hardly read or spoke English (see Wilhelm 2018). Despite the influence of Grotius and Spinoza, Hume's philosophical writings were too skeptical to the Dutch taste. One might have expected that Smith's political economy fit the Dutch like a glove, but it was actually published during a protectionist time, when the Dutch were more concerned with preserving their economic conditions. If *The Wealth of Nations* was used in academic circles, it was mostly for the statistics in it. The evidence of other uses of it, such as by bankers, businessmen, or political writers, is limited. Smith received at best a sympathetic reception in the Low Countries, becoming a foreign member of two learned societies, but his fame would have to wait for later centuries. As for his moral theory, *The Theory of Moral Sentiments* was largely overlooked despite several positive reviews after its publication. The dissemination of the ideas in *The Theory of Moral Sentiments* depended on a few enthusiasts, but the moral theory of Hutcheson was far more prominent (Hengstmengel 2021).

The most famous 18th-century Scottish thinker in the Netherlands was James Beattie (1735–1803), a professor of moral philosophy and logic at Marischal College in Aberdeen. All his works would be translated into Dutch. In the United Kingdom, Beattie was also well-known, especially for his *Essay on the Nature and Immutability of Truth* (1770). He belonged to the first generation of the Scottish Common Sense school, together with Thomas Reid, George Campbell, and James Oswald. Common Sense philosophy, especially in Beattie's version, defended morality and religion against the perceived skepticism of Hume and George Berkeley which, in the Dutch view, undermined the foundation of morality. Beattie argued that there exist intuitive principles, or axioms, that are beyond reasonable doubt, whose truth can be perceived by man's faculty of common sense. It is not human reason that forms the ultimate criterion of truth, but instantaneous and instinctive feeling, which is in line with human nature and the Creator (Hengstmengel 2020). Beattie was more explicit in his moralism than Thomas Reid, and much closer to Adam Ferguson than most of his contemporaries or predecessors had been (Wood 1990).

Eighteenth-century Netherlands hardly produced classical-liberal highlights. The one exception is the Patriot Movement, because of its emphasis on political freedom. It arose against the background of the demise of the Dutch Republic, exemplified by the loss of the Fourth Anglo-Dutch War (1780–1784). The Patriot Movement was primarily directed against the stadtholder and the rule of the House of Orange. In 1781, Baron Joan Derk van der Capellen tot den Pol wrote an anonymous pamphlet, *To the People of the Netherlands*, calling on the Dutch to defend

their rights against the stadtholder, challenging corrupt public institutions and reclaiming old local rights against the central government, including local militias, named Free Corps. The pamphlet was inspired by the American War for Independence, but as Jonathan Israel points out, in fact had its roots in the thought of 17th-century thinkers such as Grotius and De la Court, while explicitly calling on a “national feeling,” which was a relatively new phenomenon at the time. The Patriots wanted people to retake their freedom, and they fostered bottom-up democratic practices, foremost being popular participation in civic and provincial government. Militias should protect this popular freedom and take control of the state. The anti-monarchical Patriot Movement was multi-religious, including Catholics, Lutherans, Remonstrants, and other denominations (Israel 1995, 1098–1112). The Patriots became a political movement, with its main centers in Utrecht and the province of Gelderland. The appeal to the middle classes of this democratic movement was considerable, and the Patriots secured power in a number of key towns and various provinces, and through them in the States-General in The Hague. They ultimately stripped stadtholder Willem V of much of his political power. He was only restored into power in 1787, after the Prussian King had intervened, with British support. Despite their previous rhetoric the Patriots more or less vanished without a fight (Kennedy 2017, 258–260; Schama 1998).

From 1795 to 1814 the Netherlands were under French influence, first as the semi-independent Batavian Republic and then as part of Napoleon’s empire. The Batavian Republic was not very stable, but it was relatively democratic. Its founding law contained a number of classical liberal elements, making all people equal before the law regardless of their political views or their religion. It turned the federal Republic into a unitary state (Aerts 2013). After the defeat of the French, the Netherlands became a united kingdom, with King William I as its rather autocratic ruler. The southern part, now Belgium, seceded in 1830, a secession formalized in 1839.

In short, the period between 1700 and 1840 saw a stationary state, without many (classical) liberal highlights. The economy did fairly well, also due to the income from colonies, but overall much worse than before. Dutch influence on the world stage diminished and the country ended up as part of Napoleon’s empire. After 1813 Dutch independence was restored, but the new King was an authoritarian anti-liberal.

## **A so-called liberal age (1840–1918)**

The year 1848 was revolutionary in terms of constitutional developments. All over Europe, people demanded democratic reforms, and the Netherlands was no

exception. This was the result of the increased strength of the liberal movement, which had slowly commenced from the 1820s onwards. Under influence of liberal movements in France, Germany, Switzerland, and England, the Dutch liberal voice also became stronger. A central figure was the outspoken lawyer Dirk Donker Curtius, who rejected the remnants of the pre-1795 regime still visibly present, such as the influence of the old aristocracy, clientelism, and the autocratic ways of the King. He called for greater freedom of the press, greater religious freedom, direct elections for national parliament, a fully independent judiciary power, transparent public finances, and private commercial railways. Donker Curtius supported the Belgian secession, calling for public recognition of the new state by the Northern Netherlands, also demanding a new constitution, now that the structure of Kingdom of the Netherlands had fundamentally changed (Stuurman 1992, 95–134). Compared to the Patriot Movement, Donker Curtius and other liberals presented a more coherent program of economic and political liberalization, with individual freedom at its center (Van der List and Van Schie 1993, 1–4).

At the end of the 1830s, Johan Rudolf Thorbecke (1798–1872) started his ascent as the most influential liberal. Thorbecke would remain the central liberal figure in Dutch politics until the early 1870s. Yet his rise to political power took the whole 1840s. The new constitution of 1840 did not bring real change, much to liberal public dismay. It was not until early 1848 that King Willem III “turned into a revolutionary, overnight,” sacked his conservative cabinet, and gave Thorbecke orders to draft a new constitution. At the end of 1848, the new constitution was agreed to by parliament, with a leading role for Donker Curtius and his “radical liberal” friends, as Stuurman says. After a good deal of political turmoil, and much to the dismay of the King, who would continue to have bad relations with Thorbecke for the next twenty-three years, Thorbecke became prime minister in 1849 (Stuurman 1992, 135–170; Drentje 1998, 104–106).



Johan Rudolf Thorbecke  
(1798–1872)

Thorbecke was a Zwolle-born professor (at Ghent and Leiden) who drifted from academia into politics in the 1830s and 1840s. He was much influenced by the German Romantic view that saw individual and state as an organic unity. He thought natural rights and natural law were nonsense, and Rousseau’s social-contract theory too. As a professor in constitutional law he was advisor to parliament in the process leading to the slightly changed constitution of 1840, which was needed after the secession of Belgium. In the years following he became

more of a public figure, and he demanded a fundamental change of constitution. His moment came when the revolutionary spirit embraced Europe in 1848, and the King also concluded there was a need for a change. The popular idea that Thorbecke wrote a proposal by himself in ten days is exaggerated; he was part of a larger committee, and while he was the main author and also got his most important ideas included in the text, the draft did not include many of his insights and demands. During parliamentary approval (where he did not play an official role) even more amendments were made, much to his publicly displayed dismay. Thorbecke was completely convinced of himself and his ideas, and he never gave in. This made him unpopular to say the least, although friends and foes admired his intellect, willpower, and ability to resolve arguments in his favor. The 1848 constitution is still the basis for the current Dutch constitution, not least in the division into three public layers—national, provincial, and municipal, which is known as “Thorbecke’s House”—though the European level has been added. The most important change from a classical liberal perspective was that the King, and the government, was made subordinate to parliament, hence the executive power was controlled by directly elected politicians, as part of the fuller implementation of the division of powers. Elections for the Lower House were now direct, which decreased the power of the regional and local elites. Also, the official divide between religion and the state was made clearer, and other freedoms better protected, such as the freedom of association (Aerts 2020, 177, 261–430).

In his first government (1849–1853), Thorbecke worked out the constitutional provisions in lower legislation, aiming to restructure politics and public administration. He wanted to get rid of the old oligarchic ways of aristocratic local and regional government, while he continually fought turf wars with the King. Most of these battles, often over appointments of officials, were won by Thorbecke. Opposition to his plans, exacerbated by his rather rude and merciless political behavior (including toward former friends and allies), led to the fall of his first cabinet. Yet the liberal achievements were lasting, as the constitution would last despite some severe tests, most often about the ministerial responsibility for the King’s behavior. Although Thorbecke did not return to government until 1862–1866, he remained the most important and influential politician. In his second government he focused again on restructuring, but now literally: the construction of additional waterways and a nationwide railway network were top priorities—some commercial, some with public money. He also focused on general guidelines for health care, education, and culture, but actual implementation of these was outside the national state’s realm and had to be decided at the lowest level possible, often municipal. In the case of culture Thorbecke was stricter: the government could not have an opinion about the contents of the arts, and no budget either. Thorbecke’s third cabinet commenced in 1871, but he died while in office the next



year. At that time he was already past his political prime, and a younger generation impatiently waited to take over the helm and develop towards social liberalism, although many would also hold classical liberal views on particular points. Thorbecke's Romantic liberalism of organic societal order was no longer popular. It was characterized by a largely abstinent state, although this depended on specific circumstances, as Thorbecke did not always strictly adhere to his own liberal tendencies (Aerts 2020, 431–755).

While it was an uphill battle all along, liberals ensured that Dutch society modernized, and renewed its public institutions. Most notably they installed a parliamentary monarchy, with ministers that were accountable to parliament, open public debate, freedom of the press, direct elections for the Lower House, uniform legislation, and reduced influence of the King (Stuurman 1992, 361–367). These reforms were done with a focus on law and the constitution. The label *liberalism* was used in Dutch politics from the 1820s onwards, and it became more popular and common throughout the century (ibid., 110–112). The Dutch liberals were not often guided by books and other intellectual contributions, national or foreign, except perhaps John Stuart Mill's *On Liberty*. They were jurists, who wanted a strong rule of law embedded in the constitution, to ensure that politics would remain in a domain largely separated from the private sphere (Te Velde 2008b).

Liberals were the most powerful political factor for over 70 percent of the time between 1848 and 1901, which resulted in a number of classical liberal policies in the field of economics and taxation, while an increasing number of children received an education, which was also continually improved (Van Schie 2005, 22–33). They got rid of most existing protectionist measures, lowered tariffs, eradicated export restrictions, offered less protection to the Dutch commercial fleet, and initiated the (commercial and public) construction of waterways, railroads, and other infrastructure. They also took on colonial policy, which thus far had focused on enhancing overall profits without much care for the people in particularly the Dutch East Indies (Kossmann 2012, 220–228). These profits amounted up to a fifth of the national budget in 1866. The liberals fostered increased transparency in the colonial budget. This “Liberal Offensive,” as Jan Luiten van Zanden and Arthur van Riel (2004, 168–187) call it, resulted in a reform of public finance and a change in the institutional structure of the Dutch economy, which had been unusually centralized and interventionist under the powerful King Willem I. National debt dropped and government expenditure as a percentage of national income also decreased, not least because of lower interest payments. In 1869, the mercantilist Patent Act was abolished, as was the prohibitive printing tax, while the Anglo-French Cobden-Chevalier Treaty of 1860 fostered further trade liberalization in the Netherlands.

In the last quarter of the 19th century, Dutch classical liberalism was, at



least to a large extent, replaced by several forms of social liberalism, demanding increasing governmentalization of social affairs (De Beaufort and Van Schie 2014). According to Jos de Beus (1996, 77–80), this was also due to the influence of German economists in the Verein für Sozialpolitik who wanted to find a third way between classical liberalism and Marxism. These were also the economists who were opposed by Carl Menger, Ludwig von Mises, and other Austrians. Yet the development towards social liberalism was not unique to the Netherlands; it occurred quite generally throughout Europe (Freeden 1978; Weinstein 2007). The Dutch social liberals of this period, such as Johan Kappeyne van de Capello and Samuel van Houten, introduced the first important social legislation, notably limiting child labor and the maximum working hours of women. Social liberals after them, such as Tak van Poortvliet, Nicolaas Pierson, and Goeman Borgesius, would introduce more interventionist measures, while the question of general (male) suffrage also became prominent in the 1890s. After the turn of the century the two most prominent social liberals were Treub and Cort van der Linden (Stuurman 1992, 294–318; also see de Beaufort and Van Schie 2014). The latter would be the last ‘liberal’ prime minister (1913–1918) before the current Dutch prime minister Mark Rutte took office in 2010. Recently, Rutte announced he will leave Dutch politics after a new government is formed after the November 2023 general elections. Negotiations for a new government are increasingly time-consuming in the rather fragmented Dutch political system, therefore his actual exit may be sometime late in 2024.

The last quarter of the 19th century also saw the beginning of the age of political parties. To counter religious parties, liberals also organized themselves, but it would take until the early 1920s before more or less stable liberal parties would emerge (Van Schie 2005). The first party was the Liberale Unie (1884), and most of its members embraced social liberal ideas. The more classical liberal members would leave the party in 1894, reuniting as Bond van Vrij-Liberalen in 1906. In 1921, they and most members of the Liberale Unie would merge into the Vrijheidsbond, in 1928 renamed Liberale Staatspartij De Vrijheidsbond, but mainly known as Liberale Staatspartij. At the other side of the liberal spectrum, the Radicale Bond, constituted in 1894, united the progressive liberals who leaned most towards socialism. Five years later, in 1899, many proponents of direct implementation of general male suffrage left the Liberale Unie and united with the Radicale Bond into the social-liberal Vrijzinnig Democratische Bond (VDB). They fought for immediate implementation of general male and female suffrage, and were willing to make other points subordinate to that goal (Van Putten 1995, 62–64). Until the Second World War, the VDB and Liberale Staatspartij were the two principal ostensibly liberal parties, albeit with decreasing influence and declining seats in parliament (Lipschits 1982, 33–39).

The differences among these ‘liberals’ were largely of degree, not principle. The two main dividing issues were about the degree of state intervention, with all liberals of this period favoring relatively extensive intervention, and they differed about the proper moment of implementation of general suffrage. The remaining classical liberals could be found in the *Bond van Vrij Liberalen* and a miniscule *Liberaal Partij* (Van der List and Van Schie 1993, 12–18). Male universal suffrage would be implemented in 1918, and female suffrage from 1919 onwards, with the first female participation in national elections in 1922. Lizzy van Dorp was one of the first parliamentarians, and in contrast to many of her contemporaries, she did have clear classical liberal credentials, which also showed in her contacts with Ludwig von Mises and Friedrich Hayek (De Beaufort and Van Schie 2019; Dekker and Cornax 2022b).

Van Schie (2005, 377–435) notes that the demise of classical liberalism was not unique for the Netherlands. In most European countries classical liberalism lost influence in the period 1900–1940. Indeed, liberals of all descriptions were in decline. All Dutch liberals combined took around 37 percent of the vote in the first election of the 20th century, but only 9.9 percent in the last election before WWII. The same goes for Germany and England, with only the Belgian liberals remaining a notable force at around 16 percent of the vote. For the Netherlands, the main factors were the introduction of universal suffrage (although Dutch liberalism was by no means an elite phenomenon). Especially Dutch women voted for the liberals to a far lesser extent, thus not rewarding the liberal efforts in the previous decades. The strength and appeal of collectivist political ideologies was felt, although compared to the Socialists, the Christian parties were stronger direct competitors at the ballot box. World War I—The Great War—ended the era of liberal optimism, even in neutral Netherlands. Socio-economically, classical liberal policies were criticized, and the free market and free trade came under fire, especially after the Great Depression. Planning and other direct governmental interventions were the rage.

The first decades of the liberal age (from 1840 onwards) saw a number of classical liberal measures implemented, foremost in the constitution, in the field of personal and economic freedom. Yet the most important liberal, Thorbecke, did not prioritize individual liberty and, partly as a result of his organic world view, did not steadily resist the further governmentalizing of social affairs. From the 1870s onwards social liberalism took over. Hence, the so-called liberal age (1840–1918) saw some classical liberal measures, but cannot be counted as a classical liberal age. The Austrian influence on the economists was an exception (see below), but the influence of economists on public policy was much smaller than it would be after 1945. Around that time, the classical liberal influence among economists was over as well.

## 1945 to the present

Although some revolutionaries from the left had hoped that everything would change after World War Two, the Dutch quickly returned to established patterns (De Liagre Böhl 2013, 298–303). This also entailed the politics of pillarization and pacification. This meant that society was divided between pillars of socialists, different Christians, and to a lesser extent liberals. These groups would live almost exclusively among their own people, and had among others their own trade unions, employers federations, sports and leisure clubs, newspapers, broadcasting stations, churches, and political parties. To avoid violent fragmentation of society, the leaders of the pillars collaborated, also in coalition governments. This pacification brought stability in politics and in society. From the late 1960s onwards, this system slowly disintegrated, with many mergers between all parts of the pillars (Lijphart 1992), although remnants of it can still be seen, for example in the organization of public broadcasting.

Political decision-making was mostly a matter of consensus, prepared in collaboration with trade unions and employers federations. The employers federations have not stood up for classical-liberal principles. The unions, employers federations, and independently appointed ‘Crown Members’ formed the tripartite Socio-Economic Council. This was supported by a purported depoliticization of the main economic decisions. Expected effects of policy proposals were stated beforehand by the Central Planning Bureau (CPB), as it literally translates, the independent fiscal institute for economic policy analysis, which also creates macroeconomic analysis and forecasts that are the basis of the national budget. Uniquely, before national elections, the CPB also forecasts the economic effects of the election manifestos of most main parties, thereby setting the parameters of macroeconomic policy in Dutch politics (Van de Haar 2016; also see Dekker 2021). CPB economists and econometricians largely employ a mix of Keynesianism with neoclassical methodology. They believe in the power of macroeconomics, econometric modelling, and the need for broad welfare outcomes of policy processes, which includes accounting for all kinds of externalities and market failures. Yet they also maintain some regard for market dynamics to allocate scarce resources, while maintaining the old Dutch preference for free international trade.

In terms of social and economic policy, the marginalization of classical liberalism continued. In the first three postwar decades of pillarized society, classical liberals were not influential, although they were sometimes junior partners in governments. The principal liberal party is the VVD (People’s Party for Freedom and Democracy), founded in 1948, which originated from the prewar *Liberaal Staatspartij*, although its immediate predecessor was the *Partij van de Vrijheid*

(Freedom Party, 1946–1948; currently the right-wing party of Geert Wilders has almost the same name). Most members of the prewar VDB fused into the new Labor Party (*Partij van de Arbeid*). The VVD and the liberals became more successful in the 1970s when the pillarization system slowly broke down, due to the social change initiated in the 1960s (Kennedy 1995).

In 1966, a new party, D66 or Democrats '66, was founded, which initially called for change to the party system and to institute a number of direct elective measures, such as the introduction of referenda and the direct election of the prime minister. In other fields it wanted to be pragmatic. Its greatest appeal was to the non-religious urban intelligentsia, and in its politics it leaned towards the left, not unlike the VDB in the first decades of the 20th century. Hence, both parties diverged on socio-economic issues, and the VVD was more the right-wing party of law and order, but they agreed on most questions of individual liberty, such as the right to abortion, and later also gay marriage, euthanasia, etc. Both parties have also supported the loose Dutch drugs policies, especially the toleration for the use (not the trade) of so-called soft drugs, such as marijuana and MDMA. Between 1994 and 2002, D66 and VVD formed the so-called purple coalition with Labor, which was the first coalition in over 70 years without Christian-democratic parties. They were then able to draft and implement legislation on these issues. Initially, the two liberal parties were not mass parties, but both would increase their electoral base over time, albeit that D66 saw great changes in electoral results over time (Daalder and Koole 1988). Instead of 'blowing up the political system,' as its catchphrase used to be, D66 became a regular part of the system. In the 1990s, the party adopted the label 'social liberal,' making clear it should not be seen as classical liberal (Van der Land 2003; also see Brummer and Boomsma 2019). D66 has been a partner in many coalitions, including the present one.

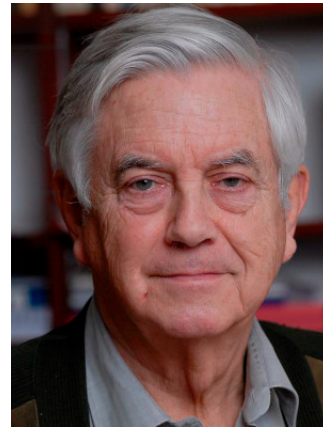
This leaves the question of whether the VVD should be seen as the best representative of classical liberalism in Dutch politics. The answer cannot be in the affirmative. It is a mixed picture at best. The founding principles of the party contain classical liberal ideas and principles, but these are not often put into practice. This started right away. From 1948 to 1963, under the leadership of P. J. Oud, the party strongly opposed socialism, but also defended Dutch colonialism, strongly opposing Indonesian independence and later the handover of New Guinea to the Indonesians. This was a position contrary to the anti-imperialism of many classical liberals including Hume and Smith (Van de Haar 2023). Until at least the 1970s the VVD appealed to higher income classes, farmers, and owners of small and medium-sized enterprises, which gave it an elitist and right-wing profile in Dutch politics. Contrary to classical-liberal ideas, the VVD embraced state intervention in the economy, such as macroeconomic steering of the economy, monetary policies, industrial policies, a drastic increase of governmental interfer-

ence in social welfare and health care, education, spatial planning, and public housing. The postwar intervention state was supported by the VVD, and it hardly attempted to justify its support, not even along the lines of Ordoliberalism. In general, the VVD hardly ever discusses its theoretical foundations. Its internal culture is anti-intellectual. The party values loyalty to the leadership (especially when in government, which is often the case) and acts like a social club. Sporadic initiatives to change this status quo have come and gone since the early 1960s, without much lasting effect (De Beus 1996, 88–93).

In the 1970s the VVD assumed an anti-left posture when the young leader Hans Wiegel took the helm, supported by chairwoman Haya van Someren and senator Harm van Riel. Wiegel appealed to a broader electorate, sharply polarizing against the socialist left. After a term in government with the newly formed CDA (Christian Democratic Appeal), between 1977–1981, Wiegel left national politics. The VVD would remain the junior partner in CDA-led governments during most of the 1980s, and was also internally in turmoil for most of the decade, initially under the leadership of social-liberal-leaning Ed Nijpels (Koole 1995, 292–309). Wiegel and Nijpels were pragmatic leaders, not interested in liberal theory, let alone classical liberalism. If they did support a classical liberal idea, it was likely out of coincidence or political opportunity (see Sijpersma 2020; Nijpels 2022). Their persistent calls for financial austerity, low taxation, and a critical position towards some (certainly not all) of the ever-increasing governmental interference in Dutch society fit this picture. In the end the VVD maintained the interventionist state. Most years when the VVD was in government, the budget was actually in deficit, although this was of course also the result of coalition government, always needed in Dutch politics. Foreign policy is another example of VVD inconsistency. Without having the room to analyze all these topics (see Van de Haar 2009; 2015; 2023), the VVD always mixes classical liberal, conservative, and social liberal viewpoints. Examples are the (initial) defense of imperialism, coupled with a strong concern for defense and Dutch NATO membership. As Cold War hawks, VVD's concern for human rights abuses was mainly reserved for those in communist countries, while the VVD always supported mandatory military service. It became critical of development aid only in the 1970s, and since the 1990s it has been internally divided over the need for further European integration. Largely, the issue was whether the European Union should develop into a federation and get more tasks on an increasing number of policy domains or, alternatively, that it should shift back tasks and powers to the national member states, in particular those not related to the internal market (Van der List 1995).

The leadership of Frits Bolkestein (1989–1998) was the classical liberal exception in the history of the VVD, until this day. Bolkestein, a former Shell manager, joined the VVD in parliament in 1977, was a minister of defense in the

mid-1980s, and became the VVD leader in 1989. He differed from any VVD leader before and after him in a number of ways. First, he did have a clear classical liberal compass, and widely published on classical liberal topics, which are collected in over 20 books (see, e.g., Bolkestein 1990; 2008; 2011; 2019). Second, he was an agenda-setting politician, going against the rather cozy Dutch consensus in foreign and European affairs, and particularly on the topic of immigration and the integration of minorities. Thirdly, he was one of the architects of the so-called purple coalitions of the 1990s, which were tripartite governments of Labour, D66, and VVD. These were the first governments without Christian-democratic input in more than 70 years. Bolkestein also had wide electoral success (Koole 1995, 306–309). After his national career, he became a European Commissioner, newspaper columnist, and part-time academic.



Frits Bolkestein (1933–)  
(Source: [Wikimedia Commons](#)/Pieter Boersma)

Despite his success, Bolkestein did not leave a lasting classical liberal legacy (Te Velde 2008a). His chosen successor, Hans Dijkstal, was his opposite in many ways, including a lack of interest in classical liberal issues. More or less the same goes for Jozias van Aartsen and Gerrit Zalm, the leaders after Dijkstal. In 2006, Mark Rutte became party leader. On many accounts, he is a remarkably talented politician, who has been able to remain the undisputed leader of his party while leading four coalition cabinets, with different parties, in unusual and demanding circumstances. He is mainly pragmatic, and his record in office is dismal from a classical liberal perspective. Partly, but not solely, under influence of his coalition partners, he let governmental interference in society grow, as well the share of the state in the economy, with increased taxes. Rutte fully stands in the tradition of Dutch social liberals.

## Classical liberalism and Dutch economists, since 1880

Outside politics, classical liberalism had a more favorable reception. Indeed, for more than 50 years, between 1880 and 1930, the Austrian school was quite dominant among Dutch economists. Nicolaas Pierson (1839–1909) contributed to the socialist calculation debate, and J. G. Koopman (1900–1958) worked on



equilibrium theory and neutral money. G. A. Verrijn Stuart (1865–1947) was the most important Dutch Austrian during this era, not least due to his editorship of the main periodical *De Economist*. Like the aforementioned Lizzy van Dorp, he corresponded with Mises and Hayek. The Austrians also referred to some of the Dutch economists in their writings. For instance, Mises regarded Verrijn Stuart's *Die Grundlagen der Volkswirtschaft* (1923) as one of the best introductions to economics (Mises 1996, 195), and in 1935 an article previously published by Pierson was included in Hayek's edited book *Collectivist Economic Planning* (1963). With the exception of Pieter Hennipman (1911–1994), the Austrian influence largely waned in the 1930s, due to the dominance of collectivist thinking in the Great Depression of the 1930s (Dekker and Cornax 2022a; 2022b).

The Austrian school would never be influential again. Other classical liberal traditions, those of Chicago and Virginia, did find inroads in academic economic sciences, especially since the 1980s, but Keynesianism would remain the most influential economic theory, and Keynes also the most admired thinker among Dutch economists. Although many leading economists are publicly known as Labor Party members, the majority of economists votes D66. Most of them believe a market society is better at generating wealth and growth than a socialist society, and that tariffs and quotas decrease economic welfare. Still, the question remains how they define a market society, because they also think taxes and government expenses can be effective in stabilizing the economy, half of them think the spread of income should be more even in developed economies, and that capitalism has a built-in tendency towards crisis. A majority also rejects Milton Friedman's idea that inflation is mainly a monetary phenomenon. The political-ideological preferences of economists play an important role in their assessment and advice on public policies (Van Dalen et al. 2016).

The lack of classical-liberal presence in the Netherlands is also seen by the scantiness of the Dutch participation in the Mont Pelerin Society. There has never been a Dutch officer (Butler 2022), and there have never been many Dutch members (often below five, and at present there are three). There was no Dutch participant at the first meeting (Caldwell 2022, 35), or at incorporation, although four Dutchmen attended the second meeting in Seelisberg in 1949 (Hartwell 1995, 51, 88). Surprisingly, the third General Meeting was held in Bloemendaal in 1950, and there was a regional meeting in Amsterdam in 1977. According to the report on the Bloemendaal meeting, Dutch MPS members at the time were A. de Graaf, J. Jitta, H. Keus (who wrote an article in the report of the meeting), J. Meyer, and G. M. Verrijn Stuart. In 1977, the meeting was organized by members from Belgium (Van Noltén) and Luxemburg (Hamilius). Dutch members at the time were De Graaf, Hennipman, Renooij, Spat, Gerrit Meijer, and Arnold Heertje, at the time the best-known economist of the country but a card-carrying member of



the Labour Party (see the Mont Pèlerin Society files, Hoover Institution).

In comparison to other countries the Netherlands is also unique for the absence of any substantially privately funded political think tank. Most existing think tanks are related to political parties and also depend on them for their (public) funding. These are mostly very small outfits (around five to ten employees) and not very influential. The Telders Foundation is the think tank associated with the VVD, and it is the only Dutch think tank that publishes books and articles on classical liberalism or with classical-liberal viewpoints. Outside the political-party orbit there are only a few additional organizations or websites. Of these, there is hardly any organization with a classical-liberal profile. There is a minuscule Libertarian Party (without a think tank), that never won seats in national or regional elections, and the tiny Mises Institute is largely focused on education in Austrian economics. In short, the development of new classical-liberal ideas mostly relies on foreign sources, and those ideas are not actively brought into Dutch public debate.

## Conclusion

The Netherlands is in some respects a classical-liberal country (trade, personal freedoms), yet deeply collectivist. In the terms of Michel Albert it has been a Rhineland country, without really having had a season as an Anglo-Saxon market economy. The most important classical-liberal contributions were made before Adam Smith's time. The social and economic developments between 1000 and 1650, in particular the Dutch roots of commercial society, are the greatest heritage, together with the contributions of the big Dutch thinkers. Some of these contributions persisted over time, most notably the preference for open commerce and important aspects of personal liberty, not least of conscience. Hayek was right to note that the Dutch Republic played an important role as example of a country with great individual liberty (Hayek 2011, 232 n.1).

Johan Thorbecke, by far the most important Dutch politician of the 19th century, cannot be seen as a classical liberal, although his basic attitude of governmental constraint had important aspects in common with the classical-liberal program. Generally, classical liberals have been a rare species in the Netherlands. I have focused on the well-known figures, yet the list is limited to some 19th-century liberals, most of the prewar economists, and Frits Bolkestein. Sure enough, classical-liberal policies were sometimes implemented, yet hardly ever grounded in a classical-liberal program or strong conviction, again policies towards the expansion of personal liberties excepted. In the Netherlands, social liberalism dominates.

## References

- Aerts, Remieg.** 2013. Een Staat in Verbouwing: Van Republiek naar Constitutioneel Koninkrijk, 1780–1848. In *Land van Kleine Gebaren: Een Politieke Geschiedenis van Nederland*, ed. Remieg Aerts, Piet De Rooy, Herman De Liagre Böhl, and Henk Te Velde, 13–108. Amsterdam: Boom.
- Aerts, Remieg.** 2020. *Thorbecke Wil het: Biografie van een Staatsman*. Amsterdam: Prometheus.
- Akçomak, I. Semih, Dinand Webbink, and Bas Ter Weel.** 2016. Why Did the Netherlands Develop So Early? The Legacy of the Brethren of the Common Life. *Economic Journal* 126(593): 821–860.
- Andeweg, Rudy B., Galen A. Irwin, and Tom Louwers.** 2020. *Governance and Politics of the Netherlands*. London: Red Globe Press.
- Armitage, David.** 2004. Introduction to *The Free Sea*, by Hugo Grotius, xi–xxi. Indianapolis: Liberty Fund.
- Bolkestein, Frits.** 1990. *De Engel en het Beest: Opstellen over Politiek*. Amsterdam: Prometheus.
- Bolkestein, Frits.** 2008. *Overmoed en Onverstand: Beschouwingen over Politiek*. Amsterdam: Bert Bakker.
- Bolkestein, Frits.** 2011. *De Intellectuele Verleiding: Gevaarlijke ideeën in de Politiek*. Amsterdam: Bert Bakker.
- Bolkestein, Frits.** 2019. *Bij het Scheiden van de Markt: Beschouwingen over Politiek en Samenleving*. Amsterdam: Prometheus.
- Brummer, Coen, and Daniel Boomsma,** eds. 2019. *De Canon van het Sociaal-Liberalisme*. Amsterdam: Boom.
- Butler, Eamonn.** 2022. *Scaling the Heights: Thought Leadership, Liberal Values and the History of the Mont Pèlerin Society*. London: Institute of Economic Affairs.
- Caldwell, Bruce,** ed. 2022. *Mont Pèlerin 1947: Transcripts of the Founding Meeting of the Mont Pèlerin Society*. Stanford, Cal.: Hoover Institution Press.
- Collins, Jeffrey.** 2011. The Early Modern Foundations of Classical Liberalism. In *The Oxford Handbook of The History of Political Philosophy*, ed. George Klosko, 258–281. Oxford, UK: Oxford University Press.
- Daalder, Hans, and Ruud Koole.** 1988. Liberal Parties in the Netherlands. In *Liberal Parties in Western Europe*, ed. Emil J. Kirchner, 151–177. Cambridge, UK: Cambridge University Press.
- De Beaufort, Fleur, and Patrick Van Schie.** 2014. *Sociaal-liberalisme*. Amsterdam: Boom.
- De Beaufort, Fleur, and Patrick Van Schie.** 2019. *De Liberale Strijd voor Vrouwenkiesrecht*. Amsterdam: Boom.
- De Beus, Jos.** 1996. Oorsprong en Wederkeer van de Liberalen. In *De Ideologische Drieboek: Nederlandse Politiek in Historisch Perspectief*, ed. Jos de Beus, Jacques van Doorn, and Piet de Rooy, 52–97. Amsterdam: Boom.
- De la Court, Pieter.** 2003. Political Maxims of the State of Holland (1662). In *Commerce, Culture, and Liberty: Readings on Capitalism Before Adam Smith*, ed. Henry C. Clark, 9–36. Indianapolis: Liberty Fund.

- De Liagre Böhl, Herman.** 2013. Consensus en Polarisatie: De Kwaliteit van de Democratie 1945–2000. In *Land van Kleine Gebaren: Een Politieke Geschiedenis van Nederland 1780–2012*, ed. Remieg Aerts, Herman De Liagre Böhl, Piet De Rooy, and Henk Te Velde, 283–346. Amsterdam: Boom.
- Dekker, Erwin.** 2021. *Jan Tinbergen: Een Econoom op zoek naar Vrede*. Amsterdam: Boom.
- Dekker, Erwin, and Willem Cornax.** 2022a. A Brief History of the Austrian School in the Netherlands. Working paper.
- Dekker, Erwin, and Willem Cornax.** 2022b. Elizabeth van Dorp and the Women’s Question at the Intersection of Bourgeois Ideals and Liberal Economics. *Oeconomia* 12(3): 581–604.
- Den Uyl, Douglas J.** 1987. Passion, State, and Progress: Spinoza and Mandeville on the Nature of Human Association. *Journal of the History of Philosophy* 25(3): 369–395.
- Drentje, Jan.** 1998. *Het Vrijste Volk ter Wereld: Thorbecke, Nederland en Europa*. Zwolle, Netherlands: Waanders Uitgevers.
- Feldman, Seymour.** 1992. Introduction to *The Ethics, Treatise on the Emendation of the Intellect and Selected Letters*, by Baruch de Spinoza, 1–20. Indianapolis: Hackett.
- Fitzmaurice, Andrew.** 2021. Property, Trade and Empire. In *Cambridge Companion to Hugo Grotius*, ed. Randall Lesaffer and Janne E. Nijman, 275–290. Cambridge, UK: Cambridge University Press.
- Freedon, Michael.** 1978. *The New Liberalism: An Ideology of Social Reform*. Oxford, UK: Clarendon Press.
- Gelderblom, Oscar, and Joost Jonker.** 2004. Completing a Financial Revolution: The Finance of the Dutch East India Trade and the Rise of the Amsterdam Capital Market, 1595–1612. *Journal of Economic History* 64(3): 641–672.
- Hanley, Ryan P.,** ed. 2016. *Adam Smith: His Life, Thought, and Legacy*. Princeton, N.J.: Princeton University Press.
- Harris, James A.** 2015. *Hume: An Intellectual Biography*. Cambridge, UK: Cambridge University Press.
- Hartwell, Ronald M.** 1995. *A History of the Mont Pelerin Society*. Indianapolis: Liberty Fund.
- Hayek, Friedrich A.,** ed. 1963. *Collectivist Economic Planning: Critical Studies on the Possibilities of Socialism* by N. G. Pierson, Ludwig von Mises, Georg Halm and Enrico Barone. London: Routledge & Kegan Paul.
- Hayek, Friedrich A.** 1978. *New Studies in Philosophy, Politics, Economics and the History of Ideas*. Chicago: University of Chicago Press.
- Hayek, Friedrich A.** 2011. *The Constitution of Liberty: The Definitive Edition (The Collected Works of F. A. Hayek, vol. XVII)*, ed. Ronald Hamowy. Chicago: University of Chicago Press.
- Hengstmengel, Joost.** 2020. ‘I Am Greatly Obligated to the Dutch’: James Beattie’s Dutch Connection. *Journal of Scottish Philosophy* 18(1): 67–90.
- Hengstmengel, Joost.** 2021. Adam Smith Goes Dutch: The Reception of Smith in The Netherlands, 1759–1800. *Adam Smith Review* 12: 329–376.
- Israel, Jonathan I.** 1995. *The Dutch Republic: Its Rise, Greatness and Fall, 1477–1806*. Oxford, UK: Clarendon Press.
- Israel, Jonathan I.** 2001. *Radical Enlightenment. Philosophy and the Making of Modernity*,

- 1650–1750. Oxford, UK: Oxford University Press.
- Israel, Jonathan I.** 2007. Introduction to *Theological-Political Treatise*, by Baruch de Spinoza, viii–xxxv. Cambridge, UK: Cambridge University Press.
- Jansen, Arne C.** 2006. Het Leven van Bernard Mandeville. In *De Wereld Gaat aan Deyd Ten Onder*, by Bernard Mandeville, ed. Arne C. Jansen. Rotterdam: Lemniscaat.
- Kaye, Frederick B.** 1988. A Commentary: Critical, Historical and Explanatory. In *The Fable of the Bees or Private Vices, Publick Benefits*, ed. Bernard Mandeville. Indianapolis: Liberty Fund.
- Kennedy, James C.** 1995. *Nieuw Babylon in Aanbouw: Nederland in de Jaren Zestig*. Amsterdam: Boom.
- Kennedy, James C.** 2017. *A Concise History of the Netherlands*. Cambridge, UK: Cambridge University Press.
- Koole, Ruud.** 1995. *Politieke Partijen in Nederland: Ontstaan en Ontwikkeling van Partijen en Partijstelsel*. Utrecht: Het Spectrum.
- Kossmann, Ernst H.** 2000. *Political Thought in the Dutch Republic: Three Studies*. Amsterdam: Koninklijke Nederlandse Academie van Wetenschappen.
- Kossmann, Ernst H.** 2012. *De Lage Landen: Twee Eeuwen Nederland en België—Deel I 1780–1914*. Amsterdam: Olympus.
- Lesaffer, Randall, and Janne E. Nijman**, eds. 2021. *The Cambridge Companion to Hugo Grotius*. Cambridge, UK: Cambridge University Press.
- Lijphart, Arend.** 1992. *Verzuiling, Pacificatie en Kentering in de Nederlandse Politiek*. Haarlem, Netherlands: Becht.
- Lipschits, Isaac.** 1982. *Politieke Stromingen In Nederland: Inleiding tot de Geschiedenis van de Nederlandse Politieke Partijen*. Deventer, Netherlands: Kluwer.
- Mandeville, Bernard.** 1988. *The Fable of the Bees, or Private Vices, Public Benefits*. Indianapolis: Liberty Fund.
- Mandeville, Bernard.** 2006. Een Bescheiden Verdediging van Openbare Bordelen, of Een Essay Over de Hoererij Zoals Zij Nu in Deze Koninkrijken Wordt Bedreven. In *De Wereld Gaat aan Deyd Ten Onder: Verzameld Werk Deel 1*. Rotterdam: Lemniscaat.
- McCloskey, Deirdre N., and John Nash.** 1984. Corn at Interest: The Extent and Cost of Grain Storage in Medieval England. *American Economic Review* 74(1): 174–187.
- McCloskey, Deirdre N.** 2011. *Bourgeois Dignity: Why Economics Can't Explain the Modern World*. Chicago: University of Chicago Press.
- McCloskey, Deirdre N.** 2019. *Why Liberalism Works: How True Liberal Values Produce a Freer, More Equal, Prosperous World for All*. New Haven, Conn.: Yale University Press.
- Mises, Ludwig von.** 1996. *Liberalism: The Classical Tradition*. Irvington-on-Hudson, N.Y.: Foundation for Economic Education.
- Mossner, Ernest C.** 1980. *The Life of David Hume*. Oxford, UK: Clarendon Press.
- Nijpels, Ed.** 2022. *Nooit wachten op Later: Opgetekend door Thomas van den Bergh*. Amsterdam: Alfabet.
- Pocock, John Greville Agard.** 1989. *Politics, Language, and Time: Essays on Political Thought and History*. Chicago: University of Chicago Press.
- Prak, Maarten, and Jan Luiten Van Zanden.** 2013. *Nederland en het Poldermodel: Sociaal-economische geschiedenis van Nederland, 1000–2000*. Amsterdam: Bert Bakker.

- Prak, Maarten, and Jan Luiten Van Zanden.** 2023. *Pioneers of Capitalism: The Netherlands 1000–1800*. Princeton, N.J.: Princeton University Press.
- Prendergast, Renee.** 2016. Bernard Mandeville and the Doctrine of Laissez-Faire. *Erasmus Journal for Philosophy and Economics* 9: 101–123.
- Ross, Ian Simpson.** 2010. *The Life of Adam Smith*, 2nd ed. Oxford, UK: Oxford University Press.
- Rummel, Erika, and Eric MacPhail.** 2021. Desiderius Erasmus. In *The Stanford Encyclopedia of Philosophy*, Winter 2021 ed., ed. Edward N. Zalta. Stanford University (Stanford, Cal.). [Link](#)
- Russell, Bertrand.** 2001. *Geschiedenis van de Westerse Filosofie: Vanuit de Politieke en Sociale Omstandigheden van de Griekse Oudheid tot in de Twintigste Eeuw*, trans. Rob Limburg. Utrecht: Servire.
- Schama, Simon.** 1998. *Patriots and Liberators: Revolution in the Netherlands, 1780–1813*. New York: Harper Perennial.
- Schliesser, Eric.** 2017. *Adam Smith: Systematic Philosopher and Public Thinker*. Oxford, UK: Oxford University Press.
- Schmidtz, David, and Jason Brennan.** 2010. *A Brief History of Liberty*. Malden, Mass.: Wiley-Blackwell.
- Scruton, Roger.** 2000. *Kant*. Rotterdam: Lemniscaat.
- Sijpersma, Pieter.** 2020. *Hans Wiegel: De Biografie*. Amsterdam: Atlas Contact.
- Skinner, Quentin.** 2002. *Visions of Politics, vol. I: Regarding Method*. Cambridge, UK: Cambridge University Press.
- Smith, Adam.** 1976 [1776]. *An Inquiry Into the Nature and Causes of the Wealth of Nations*, ed. R. H. Campbell and A. S. Skinner. Oxford, UK: Oxford University Press.
- Somos, Mark.** 2021. Virtue. In *The Cambridge Companion to Hugo Grotius*, ed. Randall Lesaffer and Janne E. Nijman, 91–117. Cambridge, UK: Cambridge University Press.
- Spinoza, Baruch de.** 2000. *Political Treatise*, trans. Samuel Shirley. Indianapolis: Hackett.
- Stuurman, Siep.** 1992. *Wacht op onze Daden: Het Liberalisme en de Vernieuwing van de Nederlandse Staat*. Amsterdam: Bert Bakker.
- Te Velde, Henk.** 2008a. De Partij van Oud en Wiegel: Leiderschap in de VVD en het Primaat van het Electoraat. In *Zestig Jaar VVD*, ed. Patrick van Schie and Gerrit Voerman, 27–53. Amsterdam: Boom.
- Te Velde, Henk.** 2008b. The Organization of Liberty: Dutch Liberalism as a Case of the History of European Constitutional Liberalism. *European Journal of Political Theory* 7(1): 65–79.
- Van Dalen, Harry, Arjo Klammer, and Kees Koedijk.** 2016. De Waardevrije Econoom bestaat niet. *Me Judice* (Tilburg, Netherlands), March 8. [Link](#)
- Van de Haar, Edwin R.** 2008. David Hume and International Political Theory: A Reappraisal. *Review of International Studies* 34(2): 225–242.
- Van de Haar, Edwin R.** 2009. *Classical Liberalism and International Relations Theory: Hume, Smith, Mises, and Hayek*. New York: Palgrave Macmillan.
- Van de Haar, Edwin R.** 2013a. Adam Smith on Empire and International Relations. In *The Oxford Handbook of Adam Smith*, ed. Christopher J. Berry, Maria Pia Paganelli, and Craig Smith, 417–439. Oxford, UK: Oxford University Press.

- Van de Haar, Edwin R.** 2013b. David Hume and Adam Smith on International Ethics and Humanitarian Intervention. In *Just and Unjust Military Intervention: European Thinkers from Vitoria to Mill*, ed. Stephano Recchia and Jennifer M. Welsh, 154–175. New York: Cambridge University Press.
- Van de Haar, Edwin R.** 2015. *Degrees of Freedom: Liberal Political Philosophy and Ideology*. New York: Routledge.
- Van de Haar, Edwin R.** 2016. CPB Netherlands Bureau for Economic Policy Analysis: Dutch (Economic) Policy-Making. In *International and Interdisciplinary Insights into Evidence and Policy*, ed. Linda Hantrais, Ashley Lenihan, and Susan Macgregor, 182–190. London: Routledge.
- Van de Haar, Edwin R.** 2023. *Human Nature and World Affairs: An Introduction to Classical Liberalism and International Relations Theory*. London: Institute of Economic Affairs.
- Van der Land, Menno.** 2003. *Tussen Ideaal en Illusie: De geschiedenis van D66, 1966–2003*. The Hague: Sdu.
- Van der List, Gerry.** 1995. *De Macht van het Idee: De VVD en het Nederlandse Buitenlands Beleid, 1948–1994*. Leiden, Netherlands: DSWO Press.
- Van der List, Gerry, and Patrick Van Schie.** 1993. Inleiding. In *Van Thorbecke tot Telders: Hoofdpersonen uit de geschiedenis van het Nederlandse Liberalisme vóór 1940*, ed. Gerardus Van Der List and Patrick Van Schie. Assen, Netherlands: Van Gorcum.
- Van Putten, Jan.** 1995. *Politieke Stromingen*. Utrecht: Aula/Het Spectrum.
- Van Schie, Patrick.** 2005. *Vrijheidsstreven in Verdrukking: Liberale Partijpolitiek in Nederland 1901–1940*. Amsterdam: Boom.
- Van Zanden, Jan Luiten, and Arthur Van Riel.** 2004. *The Strictures of Inheritance: The Dutch Economy in the Nineteenth Century*. Princeton, N.J.: Princeton University Press.
- Wagener, Hans-Jürgen.** 1994. Cupiditate et Potentia: The Political Economy of Spinoza. *European Journal of the History of Economic Thought* 1(3): 475–493.
- Weinstein, David.** 2007. *Utilitarianism and the New Liberalism*. Cambridge, UK: Cambridge University Press.
- Wessels, Marius.** 1998. *De Nederlandse Traditie van Vrijheid: Een vruchtbare voedingsbodem voor de hervormingen van 1848*. Assen, Netherlands: Van Gorcum.
- Weststeijn, Arthur.** 2012. *Commercial Republicanism in the Dutch Golden Age: The Political Thought of Johan and Pieter de la Court*. Leiden, Netherlands: Brill.
- Weststeijn, Arthur.** 2013. *De Radicale Republiek: Johan en Pieter De La Court—Dwarse Denkers uit de Gouden Eeuw*. Amsterdam: Prometheus Bert Bakker.
- Wilhelm, Frans.** 2018. Foreign Language Teaching and Learning in the Netherlands, 1500–2000: An Overview. *Language Learning Journal* 46(1): 17–27.
- Willemsen, Renatus.** 2022. *Bernard Mandeville: Een ondeugende denker?* Gorredijk, Netherlands: Noordboek.
- Wood, Paul.** 1990. Science and the Pursuit of Virtue in the Aberdeen Enlightenment. In *Studies in the Philosophy of the Scottish Enlightenment*, ed. M. A. Stewart, 127–149. Oxford, UK: Clarendon Press.



## About the Author



Dr. **Edwin van de Haar** is an independent scholar who specializes in the liberal tradition in international political theory. He has been a (visiting) lecturer at Brown University, Leiden University, and Ateneo de Manila University. Van de Haar is the author of *Classical Liberalism and International Relations Theory. Hume, Smith, Mises and Hayek* (2009), *Beloved Yet Unknown: The Political Philosophy of Liberalism* (2011, in Dutch), *Degrees of Freedom: Liberal Political Philosophy and Ideology* (2015), and *Human Nature & World Affairs: An Introduction to Classical Liberalism and International Relations Theory* (2023). He has contributed to several books, such as *The Oxford Handbook of Adam Smith* (2013) and *The Liberal International Theory Tradition in Europe* (2020), while his articles on liberal ideas and liberal thinkers have appeared among others in *Review of International Studies*, *International Relations*, and *International Politics*. His website is [www.edwinvandehaar.com](http://www.edwinvandehaar.com) and his email address is [edwinvdhaar@gmail.com](mailto:edwinvdhaar@gmail.com).

[Go to archive of Character Issues section](#)  
[Go to September 2023 issue](#)





# To Russia with Love: Boris Chicherin's 1857 “Contemporary Tasks of Russian Life”

[LINK TO ABSTRACT](#)

## Foreword

G. M. Hamburg<sup>1</sup>

The reception in Russia of Scottish political economy dates to the 1760s, when Empress Catherine II financed the studies at Glasgow University of Ivan Tret'iakov and Semen Desnitskii. As Mikhail Alekseev has shown, after they returned from Scotland, Tret'iakov and Desnitskii inspired two generations of intellectuals to read Adam Smith's *The Wealth of Nations*. By the 1810s, Smith's views had become central to political philosophers such as Aleksandr Kunitsyn and Nikolai Turgenev, and important to Russia's great poet Aleksandr Pushkin (Alekseev 2018). Fifty years later, Smith and his followers had gained a still wider constituency among Russian thinkers but had also become more controversial. This was so because Smithian economics became entangled in the increasingly bitter debate over the future of Russian serfdom.

In imperial Russia, Boris Chicherin (1828–1904) was the pivotal figure in the development of classical liberalism. From 1861 to 1867, he was a professor of law at Moscow University, and thus a lineal descendant of Tret'iakov and Desnitskii. From late 1881 to 1883, he served as Moscow's elected mayor, a position that lent him national prominence. He was also a well-regarded member of the elective land

---

1. Claremont McKenna College, Claremont, CA 91711.

councils (the *zemstvos*) in Tambov province for most of the period from 1867 to 1900. Among Russians in his generation, he probably had the deepest experience in public institutions, including local self-government. As an intellectual, Chicherin first made his reputation as a historian. In 1856, he wrote a watershed history of the centralized state and of the law before the reign of Petr I (Chicherin 1856a). Later, between 1869 and 1902, he published a profound analysis of European political thought, in which he included chapters on Adam Smith and on his socialist critics (see Chicherin 1869; 1872; 1874; 1877; 1902). Chicherin's polemic against socialism, *Property and the State* (1882; 1883), was likely the most forceful statement of classical liberal principles to appear in old-regime Russia. Finally, Chicherin's magnum opus, his *Philosophy of Right* (1900), was an attempt to marry Smith's notions of civic freedom with those of Georg Hegel and Immanuel Kant.

Today readers of English can find a convenient selection of Chicherin's writings, including sections from *Property and the State*, in a volume titled *Liberty, Equality, and the Market* (Chicherin 1998).

Chicherin's career is striking not only for his practical experience and far-ranging erudition, but also because many of his writings on problems of political economy were clandestine, composed 'for the drawer.' That is, they were written to circulate in manuscript, without attribution to the author, or were published abroad anonymously to circumvent Russian censorship. According to the bibliographer Ignatii Gul'binskii (1914, 119–143), seventeen of Chicherin's 112 writings circulated privately as manuscripts and/or were published abroad pseudonymously. Chicherin's memoirs, composed in the late 1880s/early 1890s, were intended for publication only fifty years after his death, largely because they contained accounts of his clandestine writings and because they reproduced many of these secret texts (Chicherin 1929a; 1929b; 1932; 1934). In light of this record, it is clear that, throughout his life as a classic liberal, Chicherin resorted to clandestine writing when the political situation demanded he do so.

Let us take the essay "Contemporary Tasks of Russian Life," published in 1857. The essay made the case that the Russian state had been the "driving force in our development and progress," that there was "no nation in Europe where the government is more powerful than ours." Chicherin saw this fact not as a warrant of the Russian government's invincibility but instead as proof of its vulnerability: over time, the country's rulers had isolated themselves from the people and therefore had come to lack reliable information on popular life. The byproducts of this situation were despotic rule, "the omnipresent official lie," the "pandemic corruption of officials," bureaucratic incompetence, the diminution of civic life and atrophy of public discussion, as well as the "primitive condition" of the economy.

Chicherin's proposed solutions were: promulgation of freedom of con-

science; the liberation of the peasants from servile status; freedom of speech, of the press, and of the academy; publication of governmental policies; and public access to legal proceedings. These proposals closely interrelated to core elements of a liberal program as advanced elsewhere in Chicherin's writings. For millions of nonconforming subjects, freedom of conscience was just as necessary as was the end of serfdom. Freedom of expression in its various forms was the linchpin of civic life, because open discussion of Russia's problems was a vital source of information to the government and the best safeguard against official arbitrariness. These civic freedoms in turn made necessary gains in governmental transparency. Public access to court proceedings would help ensure against the government's abuse of power and therefore called for gains in administrative transparency. The logic of enhanced civil rights pointed toward two further, unarticulated political rights: the institution of representative government and the promulgation of a constitution. Thus, according to Chicherin, the classical liberal objective of economic freedom was inextricably linked to the granting of civil and political rights—that is, to sweeping reforms.

From Chicherin's memoirs, we know that, just after Nicholas I's burial in 1855, Chicherin was in Petersburg. There he met his former teacher, the abolitionist Konstantin Kavelin. Chicherin had written a short essay on Russian diplomacy and the Crimean crisis. Kavelin showed him a manuscript by Nikolai Mel'gunov critical of nearly every aspect of Nicholas I's reign of 30 years.<sup>2</sup> In conversation with Kavelin and others, including the future minister of war Dmitrii Miliutin and the prominent Slavophile Aleksei Khomiakov, Chicherin resolved "to say everything that tortured and disturbed thinking people in Russia, to express their dissatisfaction about the past and their plans for the future." Chicherin observed: "No one at that point even thought about changing our form of government. Everyone understood that, under serfdom and given the centuries-long humiliation of educated society, that [a change in the form of government] was impossible. The one thing we all passionately desired was...intellectual and civic liberty" (1929a, 161–162).

Chicherin therefore wrote "Contemporary Tasks of Russian Life." In the handwritten 1855 manuscript, he asserted that a transformation in the system of government "would be the ultimate result of the demanded reforms." Indeed, he wrote: "In the manuscript version, unencumbered by censorship considerations, I wrote with complete candor." However, Kavelin told Chicherin that, for now,

---

2. Chicherin's manuscript, "Vostochnyi vopros s russkoi tochki zreniia," circulated from hand to hand. It was published as an appendix to *Zapiski kniazia S. P. Trubetskogo* (1907, 125–153). Chicherin wrote another essay, "Sviashchennyi soiuz i avstriiskaia politika," which Kavelin also decided to circulate in Petersburg and later published (see Chicherin 1856b). For the critique of Nicholas' reign, see Mel'gunov 1856, 67–164.

"it would be better to keep silent about the distant goal" (1929a, 163). Chicherin agreed to remove the reference to the future abolition of autocracy.

After Chicherin had edited "Contemporary Tasks of Russian Life" to suit Kavelin's scruples, Kavelin circulated it in Petersburg and eventually sent it to London, where the socialist Aleksandr Herzen had established the Free Russian Press. Herzen published the article in Russian along with many other manuscripts in the anthology *Voices from Russia* (1856–1860).

Chicherin's involvement in the production and publication of clandestine literature fit a pattern of evading censorship almost as old as Gutenberg's printing press. According to the historian Jonathan Israel, there occurred "a decisive broadening of such activity from around 1680, after which it fulfilled a crucial function in the advance of forbidden ideas for over half a century" (2001, 684). A leading historian of the Enlightenment book trade, Robert Darnton, has claimed that, "to French readers in the eighteenth century, illegal literature was virtually the same as modern literature" (1995, xix).<sup>3</sup>

In eighteenth-century Russia, clandestine literature tended to exist in manuscript form tightly controlled by the author, as the example of Denis Fonvizin's (1959) "Discourse on Fundamental State Laws" suggested: Fonvizin showed his draft only to his patron Nikita Panin.<sup>4</sup> Clandestine authors often engaged in esoteric writing, such as Aleksandr Radishchev in his *Journey from Petersburg to Moscow* (1790).<sup>5</sup> In post-Napoleonic Russia, clandestine writing was common, especially among the Decembrist movement of the 1820s.

One prominent Decembrist living in Western Europe at the time of the 1825 uprising but sentenced to death for his part in its preparation was Nikolai Turgenev, who published his memoirs of the movement in Paris, under the subtitle *Memoires d'un proscrit* (1847). A second political exile, Aleksandr Herzen, left Russia in 1847 to live in Western Europe. Herzen established two journals on European soil—*Kolokol* (*The Bell*) and *Poliarnaia zvezda* (*Polar Star*)—and published in addition *Voices from Russia*, in which Chicherin and Kavelin placed their manuscripts.

The young Chicherin was well aware of the French literary underground of the eighteenth century. Growing up in Tambov, Chicherin discovered that his neighbor, Nikolai Krivtsov, had been a Decembrist. In university, Chicherin learned details of the Decembrists' uprising, read Turgenev's *La Russie et les russes*, and became familiar with Herzen's radical past. By the time Chicherin wrote

---

3. Both Israel and Darnton were cited by Melzer (2014, 247–248). My thanks to Chris Nadon for this reference to Melzer.

4. On the history of this document see Safonov 1974.

5. Radishchev's denunciation of Catherinian Russia was printed on Radishchev's private press (ownership and operation of private presses were then legal in Russia) and cast as a travelogue to lull suspicions of the censor.

“Contemporary Tasks of Russian Life,” circumventing the Russian censorship became for him a natural process, part of being a Russian intellectual.

“Contemporary Tasks of Russian Life” was both clandestine and esoteric, as defined by Arthur Melzer (2014, 247)—clandestine because it circulated in private channels and abroad without attribution of authorship, and esoteric because the demand for political rights was between the lines. According to Melzer, virtually all late seventeenth- and early eighteenth-century political thinkers “found it necessary—in the very pursuit of popular enlightenment and freedom of speech—to employ a tactical esotericism. ... They were not strong enough—or reckless enough—to declare their true beliefs openly and all at once” (ibid., 254–255).

It is perhaps odd for twenty-first century admirers of Adam Smith to imagine societies in which discussion of free labor could not be unhindered, because economic liberty was bound firmly to the need for civil and political rights. Despite Russia’s early exposure to the Scottish Enlightenment, however, Smith’s economic and jurisprudential ideas were fully accessible there only in the political underground. Let us hope that by historical imagination modern readers will come to understand such constraints.

## References

- Alekseev, Michael P.** 2018. Adam Smith and His Russian Admirers of the Eighteenth Century. *Econ Journal Watch* 15(3): 351–364. [Link](#)
- Chicherin, Boris N.** 1856a. *Oblastnye uchrezhdeniia Rossii v XVII veke*. Moscow: Tip. A. Semena.
- Chicherin, Boris N.** 1856b. Sviashchennyi soiuz i avstriiskaia politika. In *Golosa iz Rossii*, vypuskh I, knizhka I, 145–164. London: Trübner and Co.
- Chicherin, Boris N.** 1869. *Istoriia politicheskikh uchenii*, vol 1. Moscow: Tipografiia Gracheva i komp.
- Chicherin, Boris N.** 1872. *Istoriia politicheskikh uchenii*, vol 2. Moscow: Tipografiia Gracheva i komp.
- Chicherin, Boris N.** 1874. *Istoriia politicheskikh uchenii*, vol 3. Moscow: Tipografiia Gracheva i komp.
- Chicherin, Boris N.** 1877. *Istoriia politicheskikh uchenii*, vol 4. Moscow: Tipografiia Gracheva i komp.
- Chicherin, Boris N.** 1882. *Sobstvennost' i gosudarstvo*, vol. 1. Moscow: Tipografiia Martynova.
- Chicherin, Boris N.** 1883. *Sobstvennost' i gosudarstvo*, vol 2. Moscow: Tipografiia Martynova.
- Chicherin, Boris N.** 1900. *Filosofiia prava*. Moscow: Tipografiia I. I. Kushnerev.
- Chicherin, Boris N.** 1902. *Istoriia politicheskikh uchenii*, vol 5. Moscow: Tipografiia I. I. Kushnerev i komp.
- Chicherin, Boris N.** 1907. Vostochnyi vopros s russkoi tochki zreniia. In *Zapiski kniazia S. P. Trubetskogo*, by S. P. Trubetskago, 125–153. Moscow: Tipografiia Sirius.

- Chicherin, Boris N.** 1929a. *Vospominaniia Boris Nikolaevicha Chicherina: Moskva sorokorykh godov*. Moscow: Izdatel'stvo M.i S. Sabashnikovykh.
- Chicherin, Boris N.** 1929b. *Vospominaniia Boris Nikolaevicha Chicherina: Moskovskii universitet*. Moscow: Izdatel'stvo M.i S. Sabashnikovykh.
- Chicherin, Boris N.** 1932. *Vospominaniia Boris Nikolaevicha Chicherina. Puteshestvie za granitsu*. Moscow: Izdatel'stvo M.i S. Sabashnikovykh.
- Chicherin, Boris N.** 1934. *Vospominaniia Boris Nikolaevicha Chicherina. Zemstvo i Moskovskaia дума*. Moscow: Kooperativnoe izdatel'stvo "Sever".
- Chicherin, Boris N.** 1998. *Liberty, Equality, and the Market: Essays by B. N. Chicherin*, ed. G. M. Hamburg. New Haven and London: Yale University Press.
- Darnton, Robert.** 1995. *The Forbidden Best-Sellers of Pre-Revolutionary France*. New York: W. W. Norton.
- Fonvizin, Denis.** 1959. Rassuzhdenie o nepremennykh gosdarstvennykh zakonakh. In *Sobranie sochinenii v dvukh tomakh*, tom 2, ed. G. P. Makogonenko, 254–266. Leningrad: Gosdarstvennoe izdatel'stvo khudozhestvennoi literatury.
- Gul'binskii, Ignatii Vladislavovich.** 1914. Boris Nikolaevich Chicherin: Biobibliograficheskii ocherk. In *Bibliograficheskie izvestiia*, year II, ed. V. S. Podorskii, 119–143. Moscow: Izdanie Russkim Bibliooficheskim Obshchestvom pri Moskovskom Universitete.
- Israel, Jonathan.** 2001. *Radical Enlightenment: Philosophy and the Making of Modernity, 1650–1750*. Oxford, UK: Oxford University Press.
- Mel'gunov, Nikolai Aleksandrovich.** 1856. Mysli vslukh ob istekshem tridtsatiletii Rossii. In *Golos iz Rossii*, vypusk I, knizhka I, ed. A. I. Gertsen, 67–164. London: Trübner and Co.
- Melzer, Arthur.** 2014. *Philosophy Between the Lines: The Lost History of Esoteric Writing*. Oxford, UK: Oxford University Press.
- Radishchev, Aleksandr.** 1790. *Puteshestvie iz Peterburga v Moskvu*. St. Petersburg.
- Safonov, Mikhail.** 1974. Konstitutionnyi proekt N. I. Panina–D. I. Fonvizina. In *Vspomogatel'nye istoricheskie distsipliny*, ed. Leningradskii institut istorii, 261–280. Leningrad: RAN.
- Turgenev, Nikolai.** 1847. *La Russie et les russes: Memoires d'un proscrit*, 3 vols. Paris: Imprimerie de Guiraudt et Jouaust.

## About the Author



**G. M. Hamburg** is Otho M. Behr Professor of Russian History at Claremont McKenna College. His books include: *Boris Chicherin and Early Russian Liberalism, 1828–1866* (Stanford University Press, 1992); as editor and translator, *Liberty, Equality, and the Market: Essays by B. N. Chicherin* (Yale University Press, 1998); *Russia's Path Toward Enlightenment, 1500–1801* (Yale University Press, 2016); and as editor, *Russia in War and Revolution: The Memoirs of Fyodor Sergeyevich Olferieff* (Hoover Institution Press, 2021). Hamburg is now writing a book on the poet Alexander Pushkin as prophet. His email address is [gary.hamburg@claremontmckenna.edu](mailto:gary.hamburg@claremontmckenna.edu).



# Contemporary Tasks of Russian Life<sup>1</sup>

Boris N. Chicherin  
*translated by G. M. Hamburg*

In the lives of individuals and of peoples there are moments when, shaken by unexpected events, they seem to awaken from a long slumber, look around themselves and arrive at a clear understanding of their situation. Such a moment has arrived for Russia. For a long time we assured ourselves that our fatherland is great and powerful, and we carelessly followed the course which the government set for us. A few thoughtful individuals realized with sorrow where we were headed. They saw how, under the influence of a faulty system of administration, the state itself was becoming corrupt; how all the foundations of civic life were gradually being undermined and destroyed; how all the forces [of our national vitality] were collapsing; and how the people were condemned to stagnate and grow dull in mute bondage. These individuals saw that we were imperceptibly drawing near an abyss, and suddenly the abyss opened before our very eyes. Unexpectedly, like a flash of lightning, a crisis burst upon us [in the Crimea], and in that moment, animated by the danger to our fatherland, we surveyed the scene and we began to see that a faulty system of government had undermined the might of Russia, that we had neither the personnel nor the technical means to defeat our foreign enemies,

---

1. Reprinted with the permission of Yale University Press from *Liberty, Equality, and the Market: Essays by B. N. Chicherin*, edited by G. M. Hamburg, pp. 110–140. Copyright © 1998 by Yale University. *Note by G. M. Hamburg about this text:* “Contemporary Tasks of Russian Life” was written in summer 1855 for K. D. Kavelin’s liberal “manuscript literature” campaign. Because the essay was intended for limited private circulation, Chicherin expressed his political views “with absolute candor.” The original draft of the essay even included a brief discussion of the possibility of constitutional rule in Russia. Chicherin understood the unlikelihood of a constitutional regime at present but foresaw that liberal reforms might facilitate the establishment of such a regime in the future. When Kavelin reviewed the manuscript in early 1856, he suggested that references to the liberals’ “distant [constitutional] goal” be removed from Chicherin’s essay because such references might prove politically counterproductive. Chicherin “agreed with this [assessment] and rewrote the article.” See MSG, 163–172. Kavelin sent the second, rewritten variant of Chicherin’s essay to London, where Alexander Herzen and Nikolai Ogarev published it in the *Voices from Russia* anthology. See “Sovremennye zadachi russkoi zhizni,” *Golosa iz Rossii* (London, 1857), 4: 51–129. The English text is a translation of the published version of Chicherin’s essay. Editors of the Soviet facsimile edition of *Golosa iz Rossii* described “Contemporary Tasks of Russian Life” as “the most precise formulation of the...theoretical foundations of the [Russian] liberal program.” See “Kommentarii i ukazateli,” *Golosa iz Rossii* (Moscow, 1975), 10: 177.

and that all our efforts would remain futile, given the corruption that had spread throughout the social organism.

The hour of reckoning has come. The voices of preachers, poets, and journalists call us to repentance. We all have become aware of our internal blemishes and shortcomings, although perhaps not everyone understands clearly whence they originated and how they contributed to our general misfortune. The time has come to explain these matters to ourselves, and, having devoted ourselves to introspection and having gathered our strength, to cleanse ourselves of corruption and set out anew with a rejuvenated spirit and a rational self-consciousness.

Our plight is difficult and dismal to contemplate: from abroad we face a terrible threat, the *levée en masse* which we are in no condition to match, while at home there is universal disorder and disintegration. Everywhere there is lawlessness, everywhere oppression, everywhere complaints and dissatisfaction. Nowhere in the public realm is there a gratifying development that one might survey with satisfaction and from it derive reassurance. There was but a single city on the periphery of the realm [Sevastopol] where all the heroism of the Russian people was concentrated, and which, by its courageous defense, redeemed the disasters of a difficult era. But even this city fell at last in its unequal struggle. Despondency spread everywhere and the future came to seem, if that was possible, darker than ever before.

How did we come to this humiliating pass? How could young and powerful Russia fall into such desperate straits?

Let us turn to the past and try to find in it the key to comprehending the present. Let us examine the natural conditions under which the Russian people have developed, what are their traits, what is their history, and then we will perhaps better understand what kind of turning point the people have reached and what may extricate them from their current, sorry plight.

The territory of Russia extends over the immeasurable distance from the Black Sea to the White and Baltic Seas in a monotonous steppe interrupted only by large rivers and, in the North, by broad lakes. On the steppe there are no mountains, only insignificant, hardly noticeable eminences that are lost in the vast flatland. The sparse population is sprinkled over the plains; villages and cities lie enormous distances apart, the roads are often barely passable, and there is no link between the various towns. Yet on the other hand, there are no natural barriers to divide the population and facilitate their coalescence in separate, closed alliances. Everything is dispersed across a vast space, nowhere finding a natural concentration. Rich soil easily yields food: a man does not have to work hard and contend against nature. There is nothing demanding exceptional activity and the straining of mental and physical energies, there are no striking vistas or natural phenomena arousing the imagination, no variety for developing different aspects of the human

spirit. Man seems lost in the vast steppe; he lives under the influence of monotonous nature and a limitless terrain that awakens in him an urge to wander like a nomad and a boldness that has no direction.

In such a milieu the national character took shape. The Russian nation is not composed of discrete elements, each living its own life and bearing its own peculiar stamp; it consists of a more or less homogeneous mass, dispersed across a vast area. Only Ukraine, separated from Russia as a result of historical circumstances, seems to be a discrete entity with its own characteristics. Furthermore, this homogeneous mass, dispersed over limitless territory, is nowhere concentrated. In it there are no central points that draw together scattered interests and facilitate the people's development and progress. Natural conditions led to the formation of a vast realm, but there are no tight social connections between the whole and its parts, no vital and closely woven relationships promoting human activity and constituting the life of society. The populace, scattered over the plains, lives under the influence more of natural determinants than of social ones. In Europe there is therefore no people in whom the civic spirit is so little developed as among the Russians; everyone lives aloof from everyone else, and no one troubles about common needs and interests. A Russian does not gladly leave that private sphere into which birth and environment have thrust him. He loves the nomadic life but not concerted activity. Being naturally talented, he may outdo and outperform others, but these moments are rare, evanescent flashes. In general, a Russian does not move untiringly forward, but does everything haphazardly, by chance, accidentally and lazily. He does not know how to create from his own resources an intellectually multifaceted world, and he does not tear himself away from the influence of his surroundings. In him the force of custom and tradition is astonishing; it is amazing how he uncomplainingly and humbly submits himself to the sovereign authority he recognized centuries ago.

These traits clearly suggest that the Russian character is more passive than active. But this passivity makes a Russian quite capable of accepting foreign influences, and when he is once forced to step out of his usual rut, he may easily, and with the same extreme devotion, submit to the new, just as earlier he clung stubbornly to the old. It is no wonder that under such circumstances he did not develop a multifaceted culture, he did not develop fully his talents in the fields of science, art, industry, and elegance of manners. Another factor in the development of the Russian character was Russia's distance from Rome, which passed on to Western nations the age-old achievements of the ancient world. The Russian people took from Rome the Christian religion, but accepted it, so to speak, in a pagan fashion. The Russian man did not assimilate Christianity's abstract spiritual content, its separation of the spiritual from the mundane, nor did high philosophical questions occupy him; Christianity did not prompt him to engage

in indefatigable activity for the sake of his neighbor. No, he accepted mainly the external, ritualistic side of Christianity, which became for him just as strong a tradition, just as sacrosanct as other centuries-old customs. Trivial rituals became for him more important than substantive truths; from smokers of tobacco he turned with revulsion as from apostates; the change of a letter in a book became the basis of the most important schisms. And to this day among the people pagan rituals survive, so mixed up with the Christian as to be indistinguishable. It is sufficient to point to marriage and other rites, still practiced today as the fresh, eternal flowers of popular life.

Of course, given such natural circumstances, given such a national character and the estrangement of Russia from the ancient world, Russian history could not be as rich in great phenomena as the history of the West. But the Russian people are capable of developing; they belong to the family of European peoples, and, despite all their peculiarities and the utter modesty of their historical achievement, the Russian people have developed parallel with the Europeans. Patriarchal life, based on blood relations, prevailed originally in both the East and the West. But no tribe submitted so profoundly as the Slavic tribe to this natural form of organization. Its residual effects are still felt among us. On the other hand, no tribe was less able to move beyond the patriarchal life by means of internal differentiation and to create from its own resources new forms of social organization. Whereas in the West loyalty to the feudal prince [*družinnoe nachalo*], along with the related concepts of personal freedom [*lichnost'*] and voluntary associations, developed organically, for us it was an alien notion. Patriarchal society was shaken by an influx of foreign influences, and when its internal insolvency finally showed, the Slavic tribes proved incapable of creating new forms of civic organization. To do this they needed to rely on an external authority, so they were forced to summon the Varangians, their old enemies. They left it to the Varangians to rule over them and to establish order.

This remarkable event, which has merited the attention of all historians, clearly demonstrates the passive qualities of the Slavic tribe and the inability of patriarchal leaders to unite themselves in an alliance based on their own forces and activity.

The Varangian princely retinue finally destroyed society based on the clan, and little by little replaced blood ties with voluntary alliances based on the will of individuals. But as we have noted, the principle of individuality [*lichnost'*] was unable to generate a rich culture or [to serve as a foundation for] significant social agglomerations. Torn from their natural social moorings, individuals plunged themselves into the nomadic life, and it was then that began the chaotic popular nomadism that characterized the so-called appanage period. Everyone—princes and boyars, servants, merchants, and peasants—surrendered to nomadic impulses; they all moved about the entire expanse of broad Russia, nowhere stopping to

settle, nowhere shaping solid and durable forms of social organization. Even the free cities of Pskov and Novgorod, where social life was more developed, were examples of such disorder as could rarely be found elsewhere [in Europe].

Princes were the first to settle down, and, by so doing, they became the true founders and builders of the Russian land. Chaotic nomadism could not continue; order was needed, and so there appeared sovereign princes [*gosudari*] who arrogated to themselves political power, gradually reduced the scope of nomadism, and finally compelled the nomads to submit to the state's authority. There was no protracted, spirited resistance [from below], but it was hard to force the wandering masses to settle down and assume permanent obligations. Everyone fled from his place: the boyars fled to Lithuania, urban tradespeople [*posadskie*] and peasants fled from the lands to which they had been bound. It was no easy matter to establish a [modern] state in the Russian land. Ivan IV had to arm himself with all the fury of a dread king, Boris Godunov had to employ all the guile of a clever politician in order to put an end to nomadism. But the sovereign princes had barely managed to introduce a certain order when all the downtrodden elements again rebelled and broke out of it. The Time of Troubles arrived. Pretenders, boyars, Cossacks, slaves, peasants, Polish and Russian hordes—all wandered across Russia, roaming everywhere without a goal, without a plan, without rational design of any kind. But this was already the last festival of the old way of life; the people had grown sufficiently mature to prefer living in a modern state to chaotic wandering. Moreover, the Time of Troubles threatened the very foundations of national identity—the Russian nationality and the Orthodox faith. Invasion by foreigners was a bitter experience for Russians. The people rose up, this time of their own accord, without compulsion from above and, after a desperate effort, expelled the Poles and selected their own tsar. Afterward, the people again sank back into their lazy and submissive state, trusting their future fate to the government they had established.

The first Romanovs reintroduced order, but it soon became apparent that to rule in the old way was impossible, that a strong state cannot survive without thought and enlightenment. Before Peter the Great, Russia succeeded in elaborating a rudimentary form of social organization; now was the time to add rationality and proportion to this system—to enliven it by thought. So Peter the Great turned, as a pupil, to the Western peoples, the guardians of thought and enlightenment. The Russian tsar donned a sailor's jacket in order to serve the cause of [secular Western] education, which he wished to transplant into his fatherland. This great and noble enterprise demarcated a new direction in Russian history; since that time, secular education has spread and become ever more important in Russia. The Russian people are now part of the European family and are one of the peoples governing the course of historical events. Russia's external power has grown remarkably; domestically, the realm has developed and become

stronger—an army and navy have appeared as if by miracle, the resources of the treasury have multiplied, education has spread, and there has appeared a literature in which talented individuals display their gifts, glorifying the Russian name. Whereas earlier the government was the source of all social institutions, of all measures for order and improvement, so [under Peter the Great] it now again led the new movement and fostered education. The government compelled Russians to study grammar and forcibly implanted on virgin soil science, art, industry, new manners, and customs. But it was also evident that the Russian man easily submitted to foreign influence; once having been forcibly divorced from ancient prejudices, having set off on a new course, he completely enslaved himself to the new way; he disavowed his old identity, borrowed a new language and customs, Frenchified himself, and, with a frenzy, now persecuted everything native, everything that earlier had been thought sacred. On one hand, the lower classes, which the reform did not affect, remained a stubborn negation of the foreign; on the other hand, the higher Estate transformed itself into a senseless negation of everything Russian, and, of course, in its still childish stage of development, it at first adopted only the superficial trappings of Western life and only later, little by little, did it investigate their rational content.

This quick glance at Russian history is enough to demonstrate that the government has always been a driving force in our development and progress. The Russian people, given their passivity, were not so constituted as to develop on their own initiative, without the state's interference, a multifaceted culture. The government led the people by the hand, and the people blindly followed its guide. For this reason there is no nation in Europe where the government is more powerful than ours. At no time in Russian history have the people posed a real threat to our government. The government should have fought harder against ignorance, stupidity, centuries-old and deeply rooted customs than it did against [the prospect of] serious, energetic opposition [from below]. The government's actual enemies were the seditious boyars, the wandering Cossacks, the ignorant defenders of the past. But the boyars had no roots in the people and could oppose the tsars only by secret intrigues and regionally based cabals; it was easy to deal with the Cossacks by instituting a standing army, while the devotees of the old ways [the Old Believers] could act only by silent and dull opposition and not by open force. There were only a few free cities; with one of them [Novgorod] there was a short struggle, but another [Pskov] surrendered without murmur. It was not so much independent centers and alliances that interfered with governmental activity as it was the absence of social connections, the universal disorder, and the enormous distances that made difficult the extension of the central government's power. But as these obstacles were overcome, the government's importance increased. With the growth of the population, with the establishment of stable social relationships,



with the development of the state apparatus, with the growth of the government's resources, obstacles collapsed and the government grew stronger and stronger. Today it is stronger than ever before: an enormous army submitting to the state's desires alone, an enormous bureaucracy spreading across the Russian land and everywhere carrying out the center's decisions and orders, a universal uncomplaining submissiveness taught to us by centuries of Russian history—all these factors make the government's power limitless and unconditional. Any criticism is immediately silenced, any complaint crushed, any whim [of the sovereign] may be carried out in the most distant parts of the empire.

The government's great power is also the source of its vulnerability. There is no surer axiom in politics than that the state must never act with too much persistence. Statecraft is nothing but a series of uninterrupted compromises. Its secret consists in knowing how to reconcile various demands and interests of society, in satisfying each as far as others will permit and thus establishing universal unity, in integrating these demands and interests into an overall plan based on the entire society's welfare. Only by such activity can the government attain real strength, for through it the government will generate support everywhere. Once having received its legitimate due, every social interest becomes a defender of the established legal order and of the government. But if the government, proceeding selfishly, stubbornly supports one social interest and ignores all others, it will inevitably go to extremes, alienate its popular support, and arouse dissatisfaction. Precisely by seeking to make its social foundation unassailably strong and durable, it will undermine this very foundation. Superficially it may look ever more formidable, but in fact it will grow weaker, for in the process of acquiring the external trappings of greater authority, it will lose its base of support.

This is an inescapable axiom which must be borne in mind in governing any polity. But if, as a general proposition, each social interest, each substantive demand deserves an appropriate hearing, it is all the more crucial never to ignore the totality of needs and interests constituting popular life. The government and the people—these are the two basic elements of which any polity is composed. Each has its place and each must have the freedom of activity requisite to it. The people compose the body politic, but the government is the head and master. The former live and act, give birth to various designs, demands, interests; the latter makes coherent this varied activity, fosters harmony in society, prevents one set of private goals from interfering one with another, facilitates their development, and ultimately makes the people aware of what is good for the whole. But political unity does not have to develop at the expense of diversity. Governmental activity must not preclude the autonomy of the people, for popular autonomy is a basic precondition of public life. Of course, the government cannot permit a particular interest group to take up arms against another and to compel it to behave in a



certain way, for this would be to legalize anarchy; yet to harbor and express views and demands different from the government's views and demands should be a people's legal right, for without it the people will lose all autonomy and forfeit all political significance. Every social interest around which individuals coalesce must live and develop in its own way, according to the laws of its own nature and not according to standards imposed from without. To develop properly it needs freely to express its grievances and its attitude toward the existing governmental system. Only then will it be able to assume its appropriate place in popular life. For the government to establish norms of behavior and opinion and to bend everything to these norms, to render the people voiceless and silent before the government, is to kill in the people any life and to destroy one of the fundamentals of society. The suppression of popular activity by the government is the suppression of diversity for the sake of unity, the substitution of a dead machine for a living organism, the substitution of external forms for the inner development of social forces. For, once having suppressed the people's autonomy, the government itself will inevitably become a mere external form and a dead machine. Only from the people can it draw spirit and strength and life. The people live as autonomous subjects, and it is for their sake that the government itself is established, having no other purpose than to promote the common welfare. The people are that very society for whose sake all governmental institutions exist. Why deprive the people of [the opportunity to] participate in their own affairs?

To do so would obviously contradict good political sense. Both elements [the government and the people] should function side by side; they should help one another and should collaborate in working for the common public purpose. For this reason the dominance of either the government or the people always has deleterious consequences. As soon as one begins to outweigh the other, society experiences discord; a reaction sets in and there commences a historical process tending to restore the requisite equilibrium. The entire political life of states consists in the interaction of these two elements, in the eternal search for balance between them. History presents us with an edifying spectacle that all those governing the fates of peoples must study, for a short-sighted, thoughtless concern for the present alone will lead rulers to ruinous delusions. Within such a narrow perspective, transient concerns may appear to be vitally important, while great historical forces that have developed continually over centuries may seem to be matters of relative insignificance. Only study of the past can provide us the key to understanding the present or can afford us the possibility of divining the future. Let us see what that greater teacher reveals to those attentive to her voice.

If we turn to the origin of the history of contemporary societies, to that period when, having abandoned their original patriarchal structure, they began their historical lives, we see the almost total dominance of the popular element. In

the Middle Ages the state as such did not exist, governmental authority was almost nonexistent, social ties were to the highest degree unstable and capricious. But this extreme circumstance engendered a powerful countervailing force. Lacking government, the people were but a chaotic mass subject to the reign of unbridled arbitrariness. Instead of relationships based on mutual benefit and justice, there prevailed the right of the strong; the majority was trampled under the feet of mighty personalities. Endless civil strife was ignited and anarchy ruled, universal and unbearable. [Strife and anarchy] generated a historical reaction, a movement to establish government that would introduce order and peace. The desire to avert endless troubles created a reaction against fractiousness and anarchy. The modern state arose, and history's main agenda became the gradual development and strengthening of the heretofore missing element in society—namely, the government. Anarchy diminished, order was established, the state apparatus was constructed, the government's resources were augmented, the authority of political agencies increased. The popular element became more and more secondary, and the government became ever more powerful.

Yet this dispensation also went to an extreme. Accustomed to following its own plan, the government began to consider itself a surrogate for the rest of society and to lose sight of society's other constituent element. The people became completely estranged from participation in public affairs. Every popular murmur was suppressed, and the government's voice alone filled the general void. At that point discord became evident in society; again it became apparent that there was some kind of important shortcoming, that the fullness of life had disappeared, that something essential was lacking in society. Everything became unstuck. No matter what measures were taken, they proved useless, for between legislation and its implementation there was a vast gap. The government became essentially powerless, and among the people dissatisfaction was reborn.

This is the situation in which we find ourselves today. Since the fifteenth century our state has evolved and the government's authority has grown; but only in our time has it reached an unbearable extreme. In Muscovite Russia popular life was not so restricted as it is now. At that time there was such disorder in both society and in government that the government was shackled hand and foot. With every movement it felt its own impotence. It could neither learn of social problems nor undertake any successful action [to remedy them] without the people's co-operation, and it was constantly forced to summon land councils [*zemskie dумы*] for advice on the management of public affairs. Individuals found it easy to avoid not only oppressive authority but also their legal obligations. They simply resettled themselves in other places, and, given the existing chaos, it was impossible to pursue or to find them. Entire communes not infrequently made use of this freedom. If cities or villages found too onerous the authority of an appointed governor,

they simply sent money to Moscow and obtained for themselves favorable charters that freed them from the burdensome administrator. It even happened that junior officials, without authorization [from the central government], changed superiors, and this was thought to be nothing unusual. Government was so distant and so weak that it could not resolve justified complaints and had to allow all kinds of license and disorders.

Peter the Great's reforms put an end to all this. The state had built itself up and now ceased to tolerate the unauthorized actions of junior officials as well as exceptions in the general rules of administration. But the government had not yet acquired the kind of dominance [over national life] that would have made it oppressive to the people. The government's attention was directed toward general public concerns; it was building the army and navy, stimulating industry, implanting education. Citizens carried heavy responsibilities for the sake of Russia's greatness, but the government did not intrude into everyday life. Citizens did not ceaselessly confront some administrative body set up to supervise their every action. Authority gradually concentrated itself in the center, but in the provinces people still lived freely and easily, worrying little about the authorities. The government was thought something distant and glorious, something raising Russia's might and lifting high the fatherland's banner. On the throne Russians saw the genius of Peter, the liberal wisdom of Catherine II, and they embraced the throne as the source of the fatherland's glory and prosperity.

In the nineteenth century everything changed. The age-old process of state building finally completed itself: the administrative apparatus, gradually extending itself from its trunk like a tree, pushed its branches into every region of the country, while centralization crowned the whole edifice and made administration the obedient tool of a single will. The new bureaucratic machine enabled government power to flow unhindered like a stream from center to peripheries and to return from the peripheries to the center. Now humble citizens saw the government close up, face to face. The administrative apparatus grasped us from all sides, and the more it expanded itself and branched out, the more restricted was popular activity. The government was all-encompassing, dominating everywhere, penetrating everywhere, and the people grew ever weaker and finally shrank before it. Today it is impossible to take a single step without bumping into some official, without seeing in front of oneself a manager, a director, a supervisor. On every street one meets a policeman or gendarme; on village roads, where formerly one never saw official uniforms, now rings the bell of the district police officer. And this does not occur without harm. For the humble citizen is powerless and voiceless before the most petty official, who to us is not a defender and protector but a person completely foreign and, in addition, most vile. We find ourselves narrowly and tightly bound in administrative swaddling clothes covering us from all sides, and

with every new day we sense ever more acutely that something important has gone out of life.

Meanwhile, the government's frightful dominance over national life grew ever greater as a result of its adoption of a faulty theory of administration. The inner logic of Russian history, the uncomplaining obedience of the people, bred in the government habits of despotism and arrogance; meanwhile, the conservative system borrowed from the Austrian court lent legitimacy to these habits and transformed them into theoretical conviction. Conservatism inculcated in government the notion that, besides itself, no other social element should exist in the realm, that every independent manifestation of life is lawless, something leading to troubles and revolutions, that a government desiring to preserve itself must suppress in the people every movement and all autonomous activity.

It was not difficult to create such a system in Russia. There were no obstacles; it was only necessary for the government to keep moving, oblivious to everything else, upon the same course. [The new ideology] demanded only that the government avoid distractions of the sort that may naturally issue from excessive solicitude for the popular welfare. Previously the government had been concerned about stimulating intellectual activity among the people, about spreading education; but now it was explained to the government that education engenders liberalism and atheism, that an educated people will inevitably want to participate in public affairs, and that revolution occurs when the government prohibits such harmful doctrines. The lesson was reinforced, of course, by obvious examples, and the government, giving in to such convincing proofs, renounced its former deeds. Public welfare, thought, intellectual development are compatible only with a certain degree of governmental power. These enterprises derive their vitality from the people and therefore require [that the people possess] a certain autonomy vis-à-vis the state; as soon as this autonomy is eliminated, they become estranged from and opposed to the government. It is only necessary [for the government] to cross a certain line, to move further away from the people, and these interests take on a completely different light; whereas formerly they were [the government's] allies, now they are inimical to it. Our government bravely crossed this line, and the consequence was the general decay of the state apparatus. From this came those social problems from which we now suffer. Let us try to enumerate them.

Once impregnated by the conservative spirit, the government's first concern was to find obedient servants [to staff the bureaucracy]. This was not difficult; it was only necessary to turn to the habitual slaves and cringing lackeys who, even without it being demanded, are always ready to satisfy the authorities. The government sought obedient tools and found them: it staffed itself with officials personally devoted to the tsar. Statesmen were replaced by courtiers. But the state did not benefit from this. Do courtiers care about the popular welfare and keep

the government safe from delusions? They expect nothing from the people, but for them the tsar is the source of all favors, and personal advantage prompts them to indulge the tsars and to uphold the tsars' delusions. Only by such means do the courtiers become powerful in the state; only in an atmosphere of universal silence does their voice have a certain significance. They substitute themselves for public opinion. The government inquires of them when it wishes to seek support, and their judgments pass for the voice of Russia. Through them the government receives information about the nation's condition; they decide what to tell the government and what to conceal. In other words, given the suppression of popular autonomy, they are the only mediators between tsar and people.

It cannot be said, however, that these mediators are very conscientious. They were summoned to serve the crown because of their loyalty and obedience, and they do not go beyond that. Their only goal is to please the tsar personally, to keep away from him every unpleasant sight, and by servile bows to win favors for themselves. Standing at the summit of society, they pay attention only to the sun shining above them, and they care little about what goes on below. If you plunge yourself into the middle of an ocean wave, you see below monsters moving about and eating the small denizens of the deep; yet on the surface of the water everything is placid and calm! And in fact everything in Russia is improbably placid and calm. The most scandalous abuses are carefully concealed; crying grievances are kept from the tsar's hearing; everything is arranged, smoothed over and presented in such a favorable light that it looks beautiful. The truth is inexorably persecuted because it exposes both tsars and courtiers. A triple barrier of high courtiers keeps the truth from the throne, and the government remains completely ignorant about what Russia is actually thinking and doing. Sealed off from the people by servile courtiers, the throne stands isolated, and the people have ceased to regard the throne as a sacred thing to be trusted and loved.

The government itself senses its estrangement from the people, or, rather, it considers this estrangement to be a necessary rule of politics. The conservative system proclaims that the people are dangerous, that they harbor revolutionary ideas, and that the government must find support not in the people's affection but in an enormous army powerful enough to defeat any movement. Naturally, in accordance with these principles, the government has increased the size of its army to the outer limits and has devoted to the army its greatest attentions. The army has become the model for the entire state apparatus; everything must run on a military basis. Military discipline has become the fundamental principle of government; it has been introduced everywhere—in the civil service, in private relationships and in the public sphere. Ignorant generals have been appointed trustees of universities; military discipline and instruction have been introduced in educational institutions.

Yet this enormous army was not really necessary. No one seriously contem-

plated a foreign war, and, domestically, everyone obeyed without reference to the army. Hence the army became a toy. Instead of worrying about improving the army's fighting capacity, its material basis, the soldiers' morale, or officers' education, the government devoted its attention to superficial matters. Inspections, parades, maneuvers served the tsar as amusements; his concern was with decorous uniforms and elegant maneuvers. When finally, as a result of rashness [in Russian foreign policy], war burst upon us unexpectedly [in the Crimea], when there was a real challenge for the army, then its inadequacy became apparent at every step. The adversary encroached upon Russia's borders, and yet the Russian government could not oppose it with equal forces. The government had at its disposal neither sufficient material means for effectively waging war nor able generals. Constant levying of recruits had exhausted the people, while the badly equipped recruits themselves have died by the thousands without advantage to the state. The elite guards have been held far from the war theater, while hastily recruited levies have been sent to the Crimea—levies that have proved worthless in the actual fighting.

Similar shortcomings are felt in civil administration. The government's goal has been to transform the bureaucracy into a machine that will always operate reliably as an obedient tool of central authority. The rank system has functioned in such a way as to reinforce obedience; the failure to follow established bureaucratic routine, the systematic transmission of orders from the top to the bottom of the apparatus, has led to the replacement of otherwise capable officials whose independence and convictions did not harmonize with the government's will. The government actually has achieved its goal. A bureaucratic machine now exists; its methods of operation, its protocols governing written communications, have been refined to express the most precise nuances; bureaucratic discipline is observed as never before. Yet the bureaucratic machine lacks a soul, and from the perspective of public affairs it is practically worthless. Respect for rank leads officials to cultivate the favor of their superiors rather than to think about the public welfare or systematically improving matters under their jurisdiction. State service now means serving one's superiors. Instead of applying their talents and zeal for the nation's benefit, officials give priority to personal relations, for, beginning with the highest officials in the government and descending to the most petty ones, each may profit more from sycophancy than from actual performance of particular service obligations. Meanwhile, the bureaucratic routine has led to an appalling multiplication of paperwork, which has now replaced real business. To take care of a matter means to make a formal written reply to a request, to write something on a piece of paper; it means to copy down and pass along data. The upper bureaucracy must make do with such documents and take for granted that the reported data are accurate, for it has no other means of gathering information. So on the one hand, the circle of high officials obscures and replaces public opinion, while on the other



hand, bureaucratic paperwork obscures and replaces an understanding of the real situation.

From this issues one of Russia's greatest problems—the omnipresent official lie. One can say without exaggeration that every official statement is nothing other than a lie. All reports and communiqués of high state officials are essentially lies; all reports and communiqués of governors and other provincial authorities are lies; all statistical data are lies; all assurances of loyalty to and affection for the throne are lies; all public ceremonies honoring higher officials, such as, for example, government-authorized presentations of awards, are lies; finally, even most patriotic declarations are nothing but lies. After all, we have watched private individuals give money to the government for patriotic causes, but nine-tenths of these gifts have been given under duress. We have read certain statements by members of the nobility in which nobles have assured the government of their readiness to sacrifice everything for the tsar and the fatherland, but when it was time to select levies for army duty, every noble strove to decline this honor. Under a faulty system of government the most sacred of human sentiments are perverted into disgraceful flattery and servile fear. And how could it be otherwise? When everyone must bow silently before the government, who will dare speak the truth? Only free public opinion can serve as a check on the public utterances of officials whose personal advantage always forces them to flatter the government. When public opinion is suppressed by all available means, when all who would dare to disagree with the servile choir of official praises and assurances are forced to be silent, then the lie necessarily insinuates itself into every aspect of public life. The government deprives itself of the possibility to learn of the people's real condition and to exercise restraint over the actions of its own agents, while the people lose all trust in the government.

Another no less serious problem in the existing order is the pandemic corruption of officials, both civilian and military. Abuses are the rule rather than the exception in our administrative system. As we noted earlier, it is enough [from the official perspective] to implement a law on paper; as far as the government is concerned, how the law is implemented in fact remains a mystery. Taking advantage of this state of affairs, the authorities permit themselves to do anything that strikes their fancy. They know that the cry of the oppressed will never reach the throne, and that at the very foot of the throne they will find themselves powerful protectors. Thus the power entrusted to them becomes an instrument of oppression; the law, which should be the people's defense, serves only as a means for enriching officials. Every government measure, every encounter with citizens becomes a pretext for officials to extract more money [to line their own pockets]. And subordinates remain defenseless and without recourse before their superiors because they are in no way protected against administrative arbitrariness, and



complaints are themselves considered a violation of discipline. It is the duty of the highest state officials both to supervise those junior officials who actually execute the law and to prosecute those who abuse the law, yet wherever possible, these higher officials try to cover up abuses. On one hand, a desire to please the sovereign compels them to present everything in the rosiest light; on the other hand, flattery prompts them to look leniently upon their subordinates and to conceal lawlessness. Finally, many of these same high officials are themselves no strangers to bribes. For convenience they take money from their subordinates, selling positions and extracting from the takers an annual fee [*obrok*]; in turn, their subordinates take money from their juniors and so on until finally citizens pay out of their own pockets the salaries not allocated by law to high as well as low government officials.

A priori one might assume that government is established for the benefit of the governed, but in reality it turns out to exist solely for the benefit of the rulers who constitute a bureaucratically organized corporation based on the principle of mutual assistance in the robbery and oppression of their subjects.

It is, of course, difficult for a genuinely noble and capable person to submit to a system that transforms him either into a machine for sending letters or into the wheel of a hydraulic press designed to squeeze all life and resources out of the people. Moreover, noble and capable persons are not given a chance [to affect policy], except in very unusual cases when they enjoy the personal protection of an important high official. Consequently, there is a third serious problem—the widespread incompetence of those in government. The present crisis has demonstrated how few people there are in our government with expertise and talent. The few who have these qualities have been found either in the lower ranks or far from the central sphere that ruins everyone who enters its charmed circle. It is as if incompetence has become a prerequisite for receiving a responsible office. For the most part, the incompetent person has neither an independent character that would be unpleasant to the authorities, nor convictions in disharmony with prevailing views. He agrees with everything and serves as the most uncomplaining and obedient tool for implementing imperial orders on paper. One could entrust to him an entire ministry and everything would go on as before. It is only necessary that under his command there be someone initiated in the secrets of bureaucratic routine and knowing how cleverly to arrange matters so that the interests of high officials will not clash. People of this last type constitute the business section of administration, just as the policy and executive sections consist of those whose courtier's tact replaces statecraft.

After this it is surprising that our legislation has failed to advance and that the most crucial political and social problems remain unresolved? We have a Senate, a remnant from an earlier institutional structure, which does not fit at all with the most recently created agencies. Its multitudinous and utterly superfluous areas of

jurisdiction should have been abolished long ago, while its genuinely important role as Russia's final court of appeals should have been emphasized. Yet the Senate has become a refuge for those high officials who are judged as incompetents even in our administration, and thus Russia's supreme court has become a place of exile for old men and idiots. We have a trial system that protects no one and that delays cases for decades by requiring submission of endless documentation. We have a financial system that rests all its weight on the lower classes and is a model of injustice and inefficiency in taxing the nation's resources. Finally, we have the serf system, an ulcerous wound, an institution swallowing an entire third of the nation's populace. Both equity and the interests of the state demand the abolition of serfdom. The illegality and harm generated by serfdom are so apparent as to be obvious even to a conservative government.

The government senses the incongruity of maintaining serfdom and would like to ameliorate this problem but does not know how to go about it. In general, recent legislation has been confused and contradictory. Every minute new laws are published, so that a decade's addendum to the *Code of Laws* almost exceeds the original code in length; yet among all these laws there is not one that can bring real benefit to the state, not one that can cure any of the terrible wounds of contemporary life. Individual decrees and half-measures often contradict one another; they are published today and are repealed or ignored tomorrow. Timid, venal, and incompetent hands have transformed our body of laws into chaos. There is no one in the government who can point to true principles, to the essential needs of the people; and if someone can do so, he keeps silent, so that by servile and mute obedience he may purchase for himself the tsar's favor.

Such are the political consequences of the hypertrophied development of governmental authority. Obviously, such a system could not help but have a deleterious impact on the people. The main goal of the conservative doctrine was, as I indicated above, to suppress popular initiatives of any kind and to thwart popular independence. In a sense the goal was commendable: by destroying opposition the government hoped to avoid internal dissension and to preserve its unity with the people. Unfortunately, the means were inappropriate to attainment of the ends. If the people could have been completely stricken from public life or transformed into a mute machine, then the conservative system would have been irreproachably correct. But fortunately, that could not be done. One cannot shake the people off the face of the earth; yet, if one forbids the people to speak, if one strives to create harmony by force, the people quietly become alienated from the government and refuse to cooperate with it.

That is what happened in Russia. What could have been easier for government than to uphold an alliance with the people, an alliance fortified by centuries? And yet the foolishness of the conservative system managed to weaken even this

centuries-old loyalty. Having become conservative, the government began mercilessly to persecute every independent initiative; meanwhile, the idea of the people fostered by the government itself grew ever stronger and bolder and demanded for itself greater latitude. The very growth of governmental power generated a popular demand for greater freedom. When the government was but a distant authority, it was possible not to get entangled in its grasp, it was possible to overlook rare cases of arbitrariness. But when the government's presence was felt in every aspect of daily life, it inevitably provoked a political response; legal protection from administrative arbitrariness became a necessity. Indeed, how can one not criticize the actions of an authority that one sees everywhere around oneself! How can one tolerate restrictions and abuses that one encounters at every step? For us freedom of speech which exposes the shortcomings of our body of laws and reveals the abuses of our rulers has now become essential, a necessity, yet the exercise of this freedom has been more and more circumscribed. The consequence is discord between the government's aims and the people's demands, discord that has led to the destruction of the ancient alliance [between crown and people].

Popular alienation from the government has had an impact in every sphere of public life. All thinking and enlightened persons have concluded that between themselves and the government there is nothing in common. They hunger for freedom of thought, for the enlightened and free activity of the mind, but this freedom is suppressed by all available means. They involuntarily recoil from a government which has trampled on all their fundamental beliefs and convictions. The lower classes view the government as an oppressive agency gripping them from all sides and resting upon them with all its weight. There is no hut, no corner in the Russian nation where the government's unclean hand has not reached; and wherever it has been, it has never come away empty; wherever it has intruded itself, there invariably is heard the moan of the oppressed. For the lower classes the government may be characterized in a single word: the treasury—a word ominous and frightening to the people. It is, in the eyes of the people, a monster swallowing the money and sucking the blood of poor citizens. The treasury is the people's natural enemy. Ignoring just demands and complaints, the treasury's sole ambition is to squeeze out of the humble as much money as possible, and, for their part, the humble do not consider it a crime to steal from and deceive the treasury when this can be done without fear of punishment. It never occurs to the people that the treasury exists for their benefit. The interests of treasury and people are so different that they resemble two opposing parties that, despite formal assurances of mutual love and respect, secretly try to play on each other as many dirty tricks as possible.

Yet how touching are official expressions of loyalty and submission! How the people crowd all official ceremonies! How movingly the bribe-taking and servile writers describe the Russian people's love for the established authorities

and the felicity which they enjoy under the blessed rule of autocratic monarchs! Officially everyone is improbably happy, yet if one could but lift this false curtain, how much different would be the spectacle unfolding before our eyes! It is rare for a private citizen to meet anyone contented with the actual state of affairs in Russia. Everywhere you hear endless complaints about oppression and bribery. Such is the theme of all conversations about public affairs; such are the innumerable stories which circulate everywhere, in the capitals and in the provinces, both in public meetings and around the domestic hearth. Such stories have educated the younger generation, who from their early childhood have learned to fear the government and to distrust the legal order. Under this influence a new civic life develops, a civic life full of bitterness and mistrust. It has replaced the former patriarchal ways, so carefree and tranquil, the former life on the plough so characteristic of the Russian people. How frightfully Russian society has changed in recent times! Everything that our fathers loved and to which they bowed has become foreign to us. Love for the throne and trust in the government have been beaten violently out of us. We have no memories like those that formerly delighted our grandfathers. Even the fatherland has lost its charm for us, for our public life is full of things that vex the soul and insult all noble human aspirations. No one dares say the word aloud, but secretly everyone grumbles and grows indignant.

Everyone understands, however, that grumbling and indignation lead nowhere. Active opposition, much less revolutionary opposition, is not at all in our character. For that we are too soft, compliant, and submissive. Our entire history has facilitated the growth of governmental authority, and before that power we are reduced to nothing. We know that the government is everything and we are nothing, and we see no way out of this predicament. Everyone is dissatisfied, but everyone understands that this dissatisfaction cannot be assuaged and so everyone resignedly views the situation as hopeless. Consequently, indifference to public affairs is universal. Everyone sees that he can do nothing for the fatherland's benefit, and therefore he concentrates on his private life and cares only about his personal affairs. If an individual runs into a problem with the law, he tries, where possible, to get around the law; if he has to make a monetary sacrifice for the nation, he tries to give as little as possible, and even this he gives reluctantly and under duress. There is no public spirit, no public interest, and this imparts to all public life the stamp of triviality and boredom. Not being drawn into nobler activity, most people surrender themselves to trivial calculations and vulgar passions, which are poor substitutes for intellectual passion and political engagement; those rare individuals with a more elevated vision are bored and succumb to despondency amid a monotonous, colorless, and petty life that offers no spiritual nourishment.

In general, intellectual life suffers much under the existing governmental system. Enlightenment cannot spread without a greater or lesser degree of liberty.

Science and art are disciplines that do not bend to the government's prescriptions; they have their own methods and sources of inspiration over which no one has power. Science cannot change at the government's insistence the logical coherence of its concepts or the incontrovertible meaning of certain facts. Art seeks inspiration in life; it depicts the surrounding good and evil as they appear to the artist's eyes, not as the government understands them. The government is no judge of art at all. It cannot order an artist to think this way or that, to study a phenomenon in such a way but not in another way, to depict this rather than that side of life. Science and art do not subordinate themselves to such demands, for they are by their very nature free. You can destroy them, but you cannot direct them in an arbitrary fashion.

The government itself seems finally to have arrived at this realization. Having no tolerance for freedom itself, the government cannot tolerate those things that are even indirectly connected with freedom. But cowed by the opinion of the educated world, the government has not dared completely to eradicate science and literature; rather, it has simply pressured them whenever possible, so that they have come to a very sorry pass. The censorship regulations are so strict that one cannot write anything having a humane content. Every thought is hunted like contraband, and even facts are expunged from publications that might throw a not-altogether-favorable light on the existing political order or on those political, religious, and moral principles that are now accepted as the official norms. In educational institutions military instruction and military discipline have been introduced; ignorant generals have taken charge of our national system of education; universities have lost their former prerogatives and even the number of students has been reduced. One wonders what could be better for the state than the largest possible number of educated people? One wonders what to inculcate in the younger generation if not a desire to study? And yet this most exalted human aspiration is the object of prohibitions; the government regards education as dangerous to itself, and it limits the number of students.

Consequently, is it really surprising that intellectual interests in society have diminished, that zeal for science and art has diminished? Instead of the several hundred students who formerly entered our major universities each year, the universities now field hardly enough students to fill even the limited quotas. And the cause of this is not the war, for [restrictions on university enrollment] predated the war. The cause is governmental measures that impede intellectual activity. Educationally, Russia is going backward and the government will gain nothing by this, for [its hostility to education] has produced more dissatisfied youths than could have been produced by all the journal articles in the world. Everyone who has the least respect for education must involuntarily lose respect for a government that suppresses education. Only the most patent vulgarity can enable one to describe

the government as enlightened or as a protector of science and art. Such a claim is a lie, refuted by facts at hand.

Education is also connected with Russia's material welfare. Industry cannot get along without science, but science in Russia is in disfavor. It is futile for the government to funnel personnel into those branches of science that have practical benefits. Science does not submit itself to such a disingenuous approach. Scientific advancement accrues from concentrated attention to basic science, not from [superficial attention to] its practical applications. It is necessary that there be a lively public participation in science, but this participation must spring from forces other than the government's dictates. One must first grant freedom to science, and then science will yield practical benefits.

In Russia, however, even the practical application of science suffers the same disregard as theory, and this disregard naturally is reflected in the primitive condition of industry. The ignorance of our managers and manufacturers is vast; there is only a handful of exceptions to the general rule. In most cases, when knowledge or a special skill is required, a foreigner is quickly sent for, and the administration of the firm in question is entrusted to him. This habit is very expensive; moreover, a foreigner, especially one not knowing the Russian language, usually finds it hard to deal with our workers. Those factory owners who do not have the means to hire a foreigner usually follow an age-old routine, make the most elementary mistakes, and needlessly waste time, money, and labor. Obviously, our industry is bound to suffer from this lack of enlightenment.

Other factors also contribute [to the primitive condition of our industry], factors that also ultimately arise from the same source. For the successful development of industry the first prerequisite is civil liberty, which provides the opportunity to use one's energies and talents in the most profitable fashion. Moreover, industrial growth presumes an active, energetic populace, for otherwise there will be no entrepreneurial spirit; it presumes the existence of legal safeguards against governmental interference and, in general, a system of government that will not hamper commercial exchanges. Finally, it presumes the development of society's material resources such as credit, communications, and so on. And we lack all these prerequisites for industrial growth. Serfdom diverts from work a significant portion of the people's energies and constitutes an obstacle to any improvement in agriculture; the financial system rests all its weight on the lower classes; the suppression of popular initiative and independence undermines popular industriousness; the faulty ideology adopted by the government generates ruinous restrictions for industry and trade, and it destroys private credit without which commercial exchanges are impossible; finally, the absence of good roads hinders market activity and depresses productivity everywhere. The government does not permit private entrepreneurs to build a communications network because



it wants to do everything itself, and it does everything slowly and badly. It has at its disposal insufficient capital [for building a proper road network], and of those funds which are earmarked for the communications network, two-thirds goes into the pockets of officials and contractors.

Thus both the intellectual and material interests of the state suffer from the government's short-sightedness. Indeed, this short-sightedness has become so pronounced as to affect all aspects of life. The passive qualities of the Russian people and the nature of the historical development to which these qualities have contributed have led us to a situation in which the government is all-powerful. But instead of cautiously taking advantage of this omnipotence and, in consciousness of its strength, permitting the people to exercise their autonomy, instead of stimulating in Russian society the initiative that is wanting in it and which, given the national character, cannot be dangerous to the central government, the government has concentrated all its efforts on increasing the state's power and has ignored the other elements of public life. Little by little, the government has completely separated itself from the people. In the army it has created for itself a base of support separate from the citizenry; in the bureaucracy it has an obedient instrument; the tsar is surrounded by officials who are personally loyal to him but who have nothing in common with the people. From all sides the government has fenced itself off and shielded itself from the light. It lives and moves in a separate sphere, a sphere into which nothing penetrates from the lowlands where there is not audible even a whimper of civic life. Inside the government's charmed circle everything is unique, utterly distinctive, bearing no resemblance to things on the outside. There is official loyalty, official flattery, official falsehood, official satisfaction, official welfare, official information, official articles, official discipline, official routine. All of this is created and maintained with great effort for the tsar's amusement, but, unfortunately, all is illusion. It is a soap bubble which has taken shape around the omnipotent lord; it blinds him by its rainbow colors. Meanwhile, the people live their lives, mute and submissive, but not at all happy and contented. But who cares? Courtiers treat the people with arrogant contempt. The people are duty-bound to be satisfied with their lot. Dissatisfaction is something pernicious and seditious, something that cannot be tolerated, and, if anyone dares to say that Russia is unhappy, then he is classified as a Jacobin who ought to be thrown into prison or exiled to a distant province of the empire. If not for this purpose, then why does the [political police, the] Third Section of His Majesty's Chancellery exist?

Under such circumstances one cannot be surprised that Russia finds it difficult to wage war against foreign invaders. What good is all the great valor of the Russian people when their energies are sapped by the general corruption of the state apparatus, by the virtually universal corruption they see on all sides?

Russia has reached a critical moment in its historical life. The course that it



has pursued for several centuries turns out to be incorrect; Russia has pursued that course to its logical limits and now suffers from the most ruinous consequences. Such is the law of history. The people (and along with them the government) have long been walking down one road, obsessed by a single ideal embodying their deepest desires. But as they reached the end of the road, as they grasped all the implications of their ideal, they suddenly realized that the road was perhaps not the correct one, that the idea did not make full use of all the nation's resources. Extremism inevitably entails absurd consequences. The very process of achieving an extremist program reveals the bankruptcy of that program and suggests the need to alter one's course, to pursue another, more rational objective. Then comes the turning point. Everywhere discord and frustration make themselves felt; today it is apparent that the very constituent elements of the body politic that have heretofore been held in contempt by the government are precisely those elements needed to bring about improvements in the common welfare. Happy is the nation whose government is sufficiently prudent to recognize its shortcomings, to sense its mistakes, and to lead the people onto a new course. How much misfortune and civil strife this will save the people!

For Russia such a turning point has now arrived. Heretofore in Russia the body politic has existed in name only; now the spirit of public life must be suffused into this body. Heretofore the only actor on stage has been the government; now the people must be allowed to play their role as well. Centuries of obedience have atoned for the people's anarchistic impulses. The people have submitted to the state order; they have been educated for political life. Now they must be treated not as a child wrapped in swaddling clothes but as an adult who thinks and acts independently. We will remain grateful to our mentor and prepared to love him as before, if only he will understand our needs and satisfy our just desires. We do not demand rights peculiar to our social station, nor the limitation of tsarist authority—a limitation about which no one in Russia even thinks. We need freedom! We want the opportunity freely to express and develop our thoughts, so the tsar will know what Russia is thinking and doing and can govern us with a clear understanding of social conditions and with a rational love for his people.

Liberalism! This is the slogan of every educated and sensible person in Russia. This is the banner which can unite individuals from all spheres, all social Estates, all convictions. This is the word that can shape a powerful public opinion, if only we will shake off our self-destructive laziness and indifference to the common cause. This is the word that can heal our profoundest wounds, that can heal our inner corruption, that can give us the opportunity to stand next to other peoples and, with renewed strength, to set off on a great course a harbinger of which lies in the great achievements of the Russian nation. In liberalism is the whole future of Russia. May both the government and the people rally around this banner

with trust in each other and with the firm intention of reaching the avowed goal.

But what should be understood by the term *liberalism*? *Freedom* is a vague word. It may be either unlimited or limited, and, if unlimited freedom cannot be permitted, then in what should limited freedom consist? In short, what measures should a liberal government adopt and what should the liberal party in society desire?

Let us attempt to enumerate the chief principles connoted by the term *liberalism* and those measures which, in our view, are essential for Russia's prosperity.

1. *Freedom of conscience*. This is the first and most sacred right of a citizen, for if the government begins to concern itself with matters of conscience, then what will remain safe from it? A person's religious convictions are a sanctuary into which no one has the right to intrude. They constitute the inner world of his soul and do not fall under the jurisdiction of civil law, for law is public prescription applying only to citizens' public activities. Law defines citizens' rights and duties toward one another and toward the government, but a citizen's attitude toward God remains a matter of conscience. The means a person considers best for the salvation of his soul is no concern of the state. In this sphere one can act only by persuasion, not by force; persuasion is a matter for the church, to which must be left the right of counseling souls for salvation's sake. The law has no authority whatsoever here. It cannot compel a person to believe in one thing but not in another; coercion fosters only hypocrisy, not sincere worship. Legal coercion can have but one result: peaceful citizens, who heretofore have strictly discharged their public duties, will be insulted by interference in these most intimate matters and will become enemies of the law. Such a result would doubtless harm the state just as it would harm religious belief.

The legal code itself recognizes this principle, and it proclaims freedom of conscience as the fundamental right of all citizens in the Russian empire. Unfortunately, this freedom is limited in practice. Several millions of schismatics, [the Old Believers], constituting perhaps the most industrious and developed part of the Russian peasantry, not only do not have freedom of religion but are subjected to every kind of persecution and oppression. While pagans are allowed to bow to their idols without interference, those sects closest to Russian Orthodoxy are persecuted in every way. Such an unjust contradiction in our institutions must be eliminated. Moreover, it is also necessary to abolish legal sanctions meted out both for failure to carry out religious duties and for conversion from Orthodoxy to other confessions. This is again a matter of conscience in which the civil law has no right to meddle. Finally, justice demands that oppressive restrictions be lifted from the Jews, for freedom of conscience is a right that must not be taken from any citizen of the Russian empire. No one should suffer for his religious convictions.

2. *Emancipation from servile status*, from one of the greatest evils now afflicting

Russia. Serfdom reflects not on a lesser but on a grander scale the ideology that dominates the whole nation. The seigneur wields almost the same kind of omnipotent and irresponsible mastery over his peasants that the sovereign wields over his subjects. The seigneur has more than the power over life and death: neither the person nor the property of the serf is in any way immune from the lord's arbitrariness.

The position of the two authorities [with respect to their subordinates] is similar, their abuses similar, and the very arguments that they employ to justify themselves the same. The government speaks in favor of the autocrat's prerogatives, landlords speak in favor of the right of property. The government says that to permit freedom and opposition in the state would violate the patriarchal union based on mutual love for the sake of artificial relations based on political calculations; landlords say that destruction of their authority is tantamount to replacing patriarchal care (!) [of the peasantry] with a commercial system of free wages. The government asserts that the people to whom freedom is being granted are disposed to anarchy; the landlords assert that the liberated peasant goes to seed through drink and becomes a ne'er-do-well. The government says that every concession leads to revolution; the landlords say that peasant emancipation will be a signal for the general slaughter of the lords. The government wants to act by palliative measures, by eliminating abuses, but it wants to preserve the system that generated the abuses. Likewise, the landlords do not want to hear about the eradication of the problem itself; instead, they support sanctions for violating legal statutes that actually protect no one. The government eloquently describes the bliss being savored by subjects under its power; the landlords just as eloquently prove that their peasants are better and more fortunate than all free peoples in the world. The government points to Western peoples as an example of the unhappy effects of alternate forms of political administration; the landlords depict the lot of state peasants as a misfortune that their serfs have avoided. But neither the one nor the other authority wishes to take into account what their very own subjects consider the greatest happiness. If their subjects' circumstances should even begin to resemble the people's vision of happiness, it would serve the authorities as a lesson and a warning against ruinous schemes. Both the government and the landlords claim that the liberal projects amount to nothing more than imitations of Western models that cannot be adapted to Russia. Russia, in their opinion, is a country completely unlike any other, a country having such peculiarities as to make the existing order the only conceivable one. But nobody has taken the trouble to enumerate these strange peculiarities. Apparently, they are generally ones most convenient for the authorities.

Consequently, the arguments are uncannily similar; indeed, these are the arguments of all oppressors. But there is an important difference—namely, the

government is the representative of society, a representative possessing necessary and legitimate authority which it now uses in a one-sided fashion. The landlord is but a private person invested with power, power that has no just and legal foundations. Authority over a person can belong only to the state, and private authority over other citizens is a violation of the state's prerogatives. It is repugnant to ethical sense, to justice and to the public good. It is a reward to one social Estate at the expense of another, because all citizens should equally enjoy the blessings of civic life. Therefore, there is nothing stranger than to hear a landlord speaking about his prerogatives, about government oppression affecting him, when he himself is a perpetrator and defender of the greatest oppression. The landlord has no justification whatsoever to complain about the existing order, for the abuses of seigniorial authority are much more serious than the abuses of the conservative government, and the harm done by the government is nothing in comparison with the harm done by serfdom, which ties the hands of all social Estates, perverts all social relations, hinders all institutional improvements, removes from the law's jurisdiction an entire third of the population, spoils the national character, destroys in the people a sense of human dignity, and fosters in the state apparatus immorality, illegality, and corruption. If the government is primarily responsible for the nation's sorry plight, then doubtless the burden of guilt also falls on the nobility, the primary supporters of the throne, for the nobility could have helped improve the sorry situation but did not do so. Given the prevailing moral laxity, if there is one social Estate that might still render true service to the fatherland, that Estate is the nobility. By emancipating their peasants, nobles can render a far more useful service to Russia than they render by the futile complaints and pointless lamentations now heard on all sides.

How one might accomplish the emancipation is not a matter to be examined here; that would digress too far from the issue at hand. It is enough to insist on the need for this step to be taken.

3. *Freedom of speech.* Our entire discussion suggests that freedom of speech is essential in Russia. Therefore, one should bring it to the forefront as the cornerstone of liberal politics. Let each Russian regard himself as a citizen of his fatherland who is called to assist the common cause. The government not only should not suppress political life among the people but is obliged by all means to arrange matters so that each individual may understand as clearly as possible the laws of Russia, its domestic condition, and its system of government. But this cannot be achieved by any other means than by granting to all the right freely to express their opinions and convictions. A person who for each oppositional word, for each bravely expressed thought, can be seized and subjected to arbitrary punishment is not a citizen but a slave. He cannot use his strengths and talents according to his convictions but can only silently obey another's will. But we have

seen where silent obedience has led us. It is time to put it aside, and in its place to raise freedom of speech, which can be the best and most reliable aid to a government that has in mind the popular good. Free speech alone can uncover the truth, search out capable people, rouse the public to action, and finally impel the government itself to adopt necessary reforms. Free speech is the expression of the people's aspirations. Allow it at first to be shaky, immature, uncoordinated; it cannot be otherwise after centuries of silence. But freedom will educate and strengthen public opinion, and ultimately the government will find in the public its best ally.

4. *Freedom of the press*, a necessary consequence of the freedom of speech. It is not enough to have the possibility to express one's ideas in conversations; conversations are private exchanges. If popular opinion is to have a public impact, then it should be publicly accessible and widely known, and that is only possible through the press. Until censorship which subjects every expression of ideas to governmental approval is abolished, only the government's opinion will be known and the people will remain silent. The abolition of censorship is the foundation of every liberal system desiring to rely on free speech and to give the people a certain autonomy. Only by abolishing censorship can the government prove that it does not intend to escape from trouble by ringing phrases alone, but that it wants to adopt genuine reforms. It has nothing to fear from the political tricks of the opposition. Opposition not only cannot weaken but will actually strengthen the government, provided the government knows how to behave prudently. The existence of an opposition will mobilize the government's defenders, who are today held in low esteem, for no one dares to contradict them; in the future these defenders will become more powerful and more important. Opposition will provide a legal channel for expressing the now-suppressed aspirations of the people; dissatisfaction once expressed loses half its force. Instead of a secret and widespread irritation extending to every aspect of life, whether justifiably or not, opposition will develop among the people a genuine understanding of things; it will educate their political instincts; and to the government it will reveal truth, it will indicate the government's shortcomings, it will proclaim the various desires of the people. Opposition consists precisely in expression of these heterogeneous aspirations; the nation's political life consists of debate and struggle between advocates of such [conflicting] perspectives. It is only necessary that this struggle be peaceful and legal, and that can be assured only by granting to it free channels for development. Let even absurd opinions be expressed; they will expose themselves as absurdities. The less rational the opposition, the less support it will find in public opinion and the stronger will be the government. For the government's moral strength does not consist in a populace that is outwardly silent but inwardly resentful; rather, the government finds strength as the majority joins it with trust

and affection, under conditions where the possibility of opposition exists. Only by this means can government be certain that truth is on its side; today the government has only the power of physical coercion on its side. The government should not fear prudently expressed criticism; only a call to violate the law, only an attempt to overthrow public institutions cannot be permitted. But in such a case the government may rely on law for punishing the guilty party. It can always bring a writer to court or prohibit a journal that is engaged in destructive propaganda. The law must punish actual crimes but should not prohibit every activity that presents an opportunity for abuse. There will always be abuses, but from this it does not follow that everything must be prohibited.

5. *Academic freedom.* Scholarship must develop independently, and the government should not intrude its own views into scholarship. The government must limit itself to making certain that academic departments do not become centers of political and religious propaganda rather than vehicles for promoting scholarly instruction. Meanwhile, all measures should be repealed that now hinder public education, and public education ought to be based on liberal principles and not on military discipline.

6. *Publication of all governmental activities* whose disclosure would not be injurious to the state, and especially publication of the budget, state income, and expenditures. The people should know what is happening in the central government: public business is the people's own business, and a government that genuinely cares about the people's welfare cannot fear publishing a record of its actions.

7. *Public legal proceedings.* In order to guarantee fair verdicts and speedy trials, to eliminate the innumerable abuses now hidden under the cover of darkness, this measure is vital. Plaintiffs and defendants will find in it protection against oppression; judges will find in it an encouragement to render just verdicts; the government and the people will find in it the recognition of the judiciary's rightful place as one of the most important branches of government. In addition, the daily spectacle of trial and punishment will nurture in the citizenry a respect for rights and legality that is the ultimate basis of rational public life but which is unfortunately moribund among us today. How to implement this reform is a matter about which we do not now have space to speak.

These are the chief measures which should be the objects of an enlightened government's solicitous protection and which are the desire of the liberal party in Russia. In liberalism, as we said earlier, is Russia's future; it alone can awake Russia to new life and provide the opportunity to develop the nation's slumbering potential. Therefore, everyone who genuinely loves Russia, every enlightened citizen must rally around this banner. Having cast off our ruinous indifference toward the common cause, we must, by every means permitted under the law, strive to support the great and salutary principle of freedom. Let us proclaim loudly



our desire for liberty, for our conviction is not the result of an intellectually empty, hyperemotional impulse toward license, but of genuine love for Russia and of a desire to extricate it from the sorry condition into which it has fallen. Perhaps certain people will be forced to suffer for their candor, but to suffer in a just cause is no hardship. It is time for us to dispense with our habitual servility and humiliating fear of authority and to recognize that a noble firmness of conviction is alone worthy of a great people. Civic courage is a virtue that has almost disappeared among us, but it is essential for everyone who wishes to accomplish something useful. We must act on our own, not expecting anything from the government and not blaming it for all our misfortunes. We ourselves are to blame for much in the current state of affairs in Russia. Only our own inactivity, our silent indifference to the public welfare, our unpardonable timidity could have permitted the government such a degree of blindness. Seeing no obstacles around it, having no admonitions to guide it, the government imagined that it was pursuing the proper course and it assumed that the people were satisfied with its rule. We must finally raise our voices and make evident our desire for improvements. Only when we begin boldly to tell the truth without fear of being exiled or punished, only when we ourselves act without waiting to be prompted by the government, only then will we have the right to say that we are a people that has within us the makings of a great future. And who among us does not want that for Russia?

In conclusion, let me say one more word. This exposure of our inner corruption, this public confession should serve as an answer to the call for repentance that is now heard on all sides. If it is the duty of each citizen who truly feels the shortcomings of his fatherland to speak about them for the benefit of his fellow citizens and to suggest ways to correct these shortcomings, then this duty is now doubly sacred. The call of our spiritual elders has not been made in vain. More than once from among the people the voice of true confession has been heard. We repent of our sins and we repent from the depths of our souls, with heartfelt grief, with the desire that all be well in our fatherland. But now it is the government's turn to confess. Why do our preachers, who speak so eloquently about the sins of the people, not call upon the government to repent? Do not the same transgressions that weigh down so heavily upon us also press upon the government's conscience as well? Until now we have depended on the government for everything; for centuries it has been our ruler and guide. It has now led its lost sheep to such a pass that it has called down upon itself God's punishment. Let it now make its public confession; let it show us an example of atonement. And this the government may do so by abolishing the ruinous system of administration now in place, by renouncing egoistic attempts to strengthen its authority, and by granting to the people that political life without which no enlightened state can survive.



## About the Author



**Boris Chicherin** (1828–1904) was a professor of law at Moscow University. From late 1881 to 1883, he served as Moscow's elected mayor, a position that lent him national prominence. He was also a well-regarded member of the elective land councils (the *zemstvos*) in Tambov province for most of the period from 1867 to 1900. Among Russians in his generation, he probably had the deepest experience in public institutions, including local self-government. As an intellectual, Chicherin first made his reputation as a historian. Between 1869 and 1902, he published a profound analysis of European political thought, in which he included chapters on Adam Smith and on his socialist critics. Chicherin's polemic against socialism, *Property and the State*, was likely the most forceful statement of classical liberal principles to appear in old-regime Russia. His magnum opus, *Philosophy of Right* (1900), was an attempt to marry Smith's notions of civic freedom with those of Georg Hegel and Immanuel Kant. Today readers of English can find a convenient selection of Chicherin's writings in *Liberty, Equality, and the Market: Essays by B. N. Chicherin* (1998).



# Pierre de Boisguilbert: Prime Excerpts and Some Correspondence

Pierre de Boisguilbert  
edited by Benoît Malbranque

## [LINK TO ABSTRACT](#)

It is impossible to understand the history of economic thought if one does not pay attention to the fact that economics as such is a challenge to the conceit of those in power. An economist can never be a favorite of autocrats and demagogues. With them he is always the mischief-maker, and the more they are inwardly convinced that his objections are well founded, the more they hate him.

—Ludwig von Mises (1966, 67)

*Laissez-faire* (leave things be) is an ideal which presents itself in various forms in French literature prior to the end of the 17th century, notably in the writings of Rabelais and Montaigne, but whose first great advocate in economics is Pierre de Boisguilbert (1646–1714). Boisguilbert should be seen as *the* first in a long line of liberal economists in France active through 1776, including Richard Cantillon, marquis d’Argenson, Vincent de Gournay, marquis de Mirabeau, G.-F. Le Trosne, abbots Nicolas Baudeau, André Morellet, and Condillac, L.-P. Abeille, P.-S. Dupont de Nemours, marquis de Chastellux, and A. R. J. Turgot (see Malbranque 2023).

Unlike the humorous Rabelais or the esoteric Montaigne, Boisguilbert was frank, challenging, even indignant. The method of taxation in use over the decades up to 1695 was ruinous, so he wrote to the minister Michel Chamillart that “your predecessors were well-intentioned, but they acted as if they were paid to precipitate the downfall of the King and his people” (Boisguilbert 2023, II:51). Similarly,

restricting the grain trade was “like stabbing every year a great amount of people” (ibid., II:61). “Please be kind enough to open your eyes to the real situation of the kingdom,” he wrote to the Controller-General (ibid., II:92).

One might liken Boisguilbert to such challengers as Thomas Paine, William Lloyd Garrison, Lysander Spooner, Ludwig von Mises, and (to a lesser degree) Frédéric Bastiat. But for the most part, those later authors did not face censorship and persecution to the degree that Boisguilbert did. Boisguilbert was censored and persecuted merely for calling good and bad, true and false, as he saw them.

Until recently, Boisguilbert’s main economic works had not been translated into English (although they had been into Italian, German, and Chinese). The present article provides some excerpts of English translations done by me and offered more fully by Institut Coppet ([link](#)). The excerpts provided here are from two works. Also provided here are three letters from 1707 between Boisguilbert and his persecutors.

## The gist of Boisguilbert

There are two chief planks of Boisguilbert’s economic writings: (1) The call for freer trade across jurisdictions, externally and *internally*—it is important for the modern reader to understand that jurisdictions internal to France (towns, regions known then as *généralités*, or collectively as *pays d’états* and *pays d’élection*, depending on the nature of their local governance) imposed their own restrictions, and that the internal restrictions were more damaging than the external restrictions. (2) Tax reform; the reality of the tax system in 1695 is difficult to comprehend, being full of complexities, but suffice it to say that it was collected by the agents of the Ferme générale, forming a state within the state, and that all in all it was arbitrary in implementation and enforcement, inefficient, burdensome, corrupt, and often abusive. It is worth noting that one hindrance to economic progress that Boisguilbert did *not* address was the guilds (the first to take those on in a significant way was Gournay; see Gournay 2023).

Boisguilbert authored nearly 100 letters sent to ministers in the space of twenty years, as well as several short books, published anonymously and illegally, with the aim of winning public support for free trade and fair taxation. He reiterated his pleas with boldness, and it is hardly surprising that ministers did not listen to him. In 1695, when he first released some of his writings to the public, he dared to rename his first book, *The State of France (Le Détail de la France)*, to the rather punchier *France Being Ruined Under the Reign of Louis XIV, by Whom and How: Along with the Solution for a Quick Recovery*. It was avidly read, to the point that it can be named one of the economic best-sellers of the century (Carpenter 1975, 11).

The state of France, according to Boisguilbert, was in spoilage. Grapevines and fruit trees were being abandoned, fields were left to ravens and owls, and all lands had lost half or even three quarters of their value (Boisguilbert 2023, II:25, II:100, I:25). Such was “the corpse” of France. One then had to find the culprits and punish them (*ibid.*, I:156). If the people of France were living a life of near-poverty, it was because consumption was made impossible by two faulty schemes: first, unfair and unpredictable taxation, which discouraged land owners from investing and farmers from working; second, the restriction on trade, of grains, in particular (*ibid.*, I:30–31).

People responsible for the collection of taxes, Boisguilbert says, behave as if France were a conquered foreign country. The conquerors do not find it extraordinary that a man is entirely ruined, his house and all his belongings seized, to satisfy their demands (Boisguilbert 2023, I:79). Because of unstable, unpredictable, and unfair taxation, people across the country were discouraged from being rich and getting richer. Anyone who did not appear to be extremely poor, or paid his taxes on time without fuss, could be sure to be overtaxed the next year (*ibid.*, I:36). Therefore, every man hid his possessions and pretended to be on the verge of poverty. “Only a daily life of bread and water,” Boisguilbert notes, “can keep a man safe. If his neighbor sees him buying a piece of meat or some new clothes, he will be made to pay; if by accident he receives money, he must hide it, for if this was ever known, he would be a lost man” (*ibid.*, I:180). Since the best way to appear poor was to be and remain poor, no agricultural improvement was being made. “Everything that a man could earn in addition would not be for him; thus if he understands where his interest stands, he will do his best to stay still, as much as possible” (*ibid.*, II:250; see similarly Vauban 2007, 768).

France was being ruined also, Boisguilbert wrote, because trade was restricted in many ways. Well-intentioned measures were in place, for instance, to prohibit the free flow of grains across the country. Regions where good crops were expected or obtained could not exchange with others where harvests had been bad. A series of controls and permits were making the whole business of trade very wearisome. Once a producer has experienced what it takes to sell his commodity, Boisguilbert alleged, seeing how the government employees abuse the business, the producer will make sure never to conduct trade again (Boisguilbert 2023, I:46). Consequently, wine producers would rather lose some of their wine than to bother transporting it with carts and horses and be subjected to the abuse (*ibid.*, I:48; see also Vauban 2007, 768).

Boisguilbert offered solutions. Taxation was to be established in such a way, he asserted, that people will be once again allowed to work and to spend their income (Boisguilbert 2023, II:66; see also Vauban 2007, 762). The current trend had to be reversed, in favor of fair and predictable taxation. Taxation is coercive in

essence, so even the best schemes are fraught with difficulty. But Boisguilbert said that taxes should also be made to flow directly from taxpayers to the state treasury, without an army of intermediaries and stakeholders being involved (Boisguilbert 2023, I:185). As to trade, it had simply to be made free. The freedom of roads is necessary for consumption and consequently for incomes,” Boisguilbert insisted. “They cannot rise from a situation where one encounters at every step people who are paid to prevent two countries [*pays*—but could be understood as different parts of France] from trading with one another” (ibid., I:225).

The economic policy that Boisguilbert espoused can be summarized in one motto: *laissez faire la nature*, that is, let nature operate freely. Because of this conclusion—as well as some brilliant passing remarks, for example on what after Adam Smith would be known as the “invisible hand” idea (Boisguilbert 2023, II:236)—Boisguilbert emerges as one of the earliest of the liberal economists. He calls for the disengagement of government from virtually all economic affairs, which are meant to function “without the intervention of any superior authority, which must be barred from any sort of production whatsoever, because nature, far from obeying the will of men, is constantly rebelling, and always finds a way to punish those who held her in contempt, by means of famine and desolation” (ibid., I:156).

*Notes on the texts:* I have tried to remain as faithful and as close to the original as possible. All footnotes and insertions in square brackets [like these] are my own.

## ***Le Détail de la France 1695:*** **Excerpts from Boisguilbert’s** **25-point summary of the book**

Boisguilbert enumerates 25 points to summarize his book *Le Détail de la France* (*The State of France*), published in 1695. I omit a majority of the 25 points, beginning with the first eleven. In those eleven points, Boisguilbert says that “French soil is excellent” but that the extent and quality of its cultivation is stunted and indeed far less than in the past. He suggests that the kingdom has lost “half of its wealth in thirty or forty years, and amidst no spread of plague, no earthquake, no war at home or abroad, and no such incidents that bring monarchies down in ruins.” In point 11 he says: “It all comes down to finding the cause of this abandonment.” I now reproduce the four points that come next:

12. There can only be two reasons why a man would be prevented from cultivating his land: either it is because cultivation requires funds that he is not able to muster, neither by himself nor by borrowing; or it is because after

cultivating he would not be able to sell his productions, as he did in the past, and therefore that he would risk losing all his funds: and therefore it becomes in his interest to leave his land bare.

13. This is precisely what happens regarding the arbitrary *taille*. It is not extraordinary to witness a large enterprise paying virtually no *taille*, while a poor man, who can only count on the resource of his two arms to provide daily sustenance for himself and his family, is under intense pressure; and the reason why he is not pressured more is that if higher taxation was levied on him the payment would never be successfully collected; and thus, if he were to plow the soil, currently left bare, the harvest would not be for him, and he would lose in addition the costs of cultivation, which are very high.

14. As for the second obstacle—that which stops him from cultivating, for after the harvest selling the productions would be impossible—duties of *aides* and customs on import and export [between jurisdictions within France, but also from and to other neighboring countries], being four times higher than what the commodity can bear, which is the reason why these duties become useless for the King himself, since nothing is levied, such duties have resulted in a situation where consumption was reduced by a factor of four in the course of thirty or forty years; and one is not surprised to see a whole county drinking only water, when in the neighboring county grapevines and fruit trees are pulled up; and far from resulting in an increase of the sums collected by the King, this has prevented them from doubling since 1660, as they did every thirty years, from 1447 until the said year 1660.

15. The remedy to all of this is easy, as long as one will only concern oneself, in the issue of taxation, with the interests of both the King and his people. One must examine whether there cannot be a system in which sums are immediately transferred from the hand of the people, to that of the King, which would have a rule and maintain a balance between all conditions, so that the poor pay as poor and the rich as rich, and this without the involvement of judges or authorities, whose intervention implies fees and a loss of time, amounting to the primary tax burden itself.

Boisguilbert then gives the next several points principally to the issue of taxation, suggesting reforms that he says would make taxation simpler, fairer, and less burdensome. His call for fairer taxation based on ability to pay is not to be understood as a call for progressive taxation, but rather an objection to regressive taxation, which resulted in part from the widespread exemption from taxation enjoyed by nobles, clergy, and office holders. I proceed by giving the last three of the 25 points and Boisguilbert's two paragraphs that follow his 25-point listing:

23. One should not expect tax collectors to ever suggest another system, for their intention being to receive large remittances, they put all their hope in a system resulting in a difficult and hence ruinous recovery of taxes: such a recovery enriches them to the same extent as it impoverishes the poor, since the expenses made to forcefully collect taxes are shared between them, ushers and *recors*<sup>1</sup>, who give them large discounts on what is submitted to taxes.

24. All these truths, being denied by tax collectors and their protectors, who are in much greater number than it is believed, will be certified by all important persons, either in administration or in trade, who live in the country; yet those whose interest it is to cause the ruin of everything being the only ones listened to, no attention is given to those who would like to save everything, and who could not even ask to be heard without risking their own security.

25. This summary in articles is provided here so that the bad faith of those who would wish to deny their consequence will appear more clearly: for it being impossible for them to dispute any of these articles without showing their lack of reasoning power or good faith, they must assert, despite their lack of what was previously mentioned, that the King can indeed enrich himself and his people, in fifteen days, whenever he will decide that no longer will we accept that some make fortunes by causing his ruin, and that of his people, but that he will now collect all that is necessary for the present war, without being cause for despair to his people: which is what happens when a man sees his property seized and sold as a result of taxes ten times higher than what he can bear, which forces him and his family to rely on charity, and yet does not give anything to the King, as happens everyday.

All of this, without making dangerous moves, but only by enforcing the regulations of *tailles*, which provide that this tax will be calculated according to the capabilities of each person, and by joining a part of the custom duties known as *aides*, as explained previously, and as was the case thirty years ago: and this is four times less far-reaching than the *capitation* tax was.

In this way, it is maintained that the people will possess two hundred million more worth of property, by this liberation of their possessions previously seized. And since the King needs sixty million per year in addition to his normal revenue, there are a thousand ways to obtain them from people who would just have seen their wealth increased by four times this amount, not to mention the future, in which it will double again, in less than the two or three years that are needed to collect the funds.

---

1. Public officials in charge of seizures and executions.



**“Are we to wait for the end of the war?”**  
**A supplement to *The State of France***  
**Published in *The Real Situation of France (1707)*<sup>2</sup>**

Boisguilbert's second book, *Factum de la France* (1707), contains various memorials. In one, Boisguilbert employs the rhetorical device known as anaphora, starting each of a successive series with the phrase “are we to wait for the end of the war.” He is objecting to the excuse of waiting to the end of the war to make the needful reforms he has proposed. I quote selectively, using ellipses where material has been omitted.

Are we to wait for the end of the war to cultivate lands in all parts of the kingdom, where they are idle for the most part, due to the low price of wheat making it impossible to bear the cost of cultivation, and similarly where seeds are not sown in the other fields, which causes the country a loss of more than 500,000 muids<sup>3</sup> every year, and a loss of 500 million in the revenue of people, because it stops the circulation of this first commodity [wheat], which sets in motion every other industry, all living and dying together...? ...

Are we to wait for the end of the war to stop uprooting grapevines, as every day is the case, when three quarters of the people are drinking water only, because of heavy taxation of alcohols, exceeding four or five times the price of the commodity...?

Are we to wait for the end of the war to enforce a new rule for the *tailles*, such that they will be fairly distributed across the country, and will not allow large fortunes to be barely subjected to taxation, when a poor man who has nothing but his two arms to earn a living for himself and his family, sees not only the sale of his deteriorated furniture or instruments, necessary for his livelihood, as it happens regarding the *ustensile* tax, which is based on the *taille*, but also witnesses his doors and bed base being taken away from his house in order to pay a tax four times greater than what he can bear? ...

Are we to wait for the end of the war to save the lives of two or three hundred thousand creatures who die every year from poverty-related causes, and this especially at a young age, for less than half of all children will ever be old enough to earn a living, because mothers lack breast milk, due either to lack of food or to excessive workload; and when those who reach an older age

---

2. The original French text is available from Institut Coppet ([link](#)).

3. One muid represented a different amount in Rouen, where Boisguilbert lived, and in Paris. In any case it was a little under 2m<sup>3</sup> or 2,000 liters (70.6 cubic feet or 67,500 ounces).

have only bread and water, but no bed, no clothes, no remedy in case of illness, and no sufficient strength to perform their work, which is nonetheless their only source of income, and therefore die before even having walked half of the road?

Are we to wait for the end of the war to end similarly the war that is made on properties, which can happen instantly, if the King would only declare that taxes will now be raised according to the capacity of each taxpayer, as is currently the case in England, in Holland, and in every country in the world, and as was even the case in France for one thousand one hundred years...?

Are we to wait for the end of the war to enable the King to fund officers at the right time, so that they will be able to recruit their soldiers in a timely manner?

Are we to wait for the end of the war to give sufficient funding to the King so that through a large scheme soldiers will be voluntarily recruited, and that we will not see anymore men being dragged into the army, with their hands behind their backs, like we would do for convicts being sent to galleys or even to the gibbet; all of which, according to Mr. Sully, in his Memoirs, only makes other soldiers feel discouraged, and the armed forces, as well as the nation, despised, as all soldiers desert the military at the first opportunity, or die of sorrow?

Are we to wait for the end of the war to stop putting the King and the State in debt, and at such a pace that when the war ends the interest on the borrowing will cost the people more than the war itself, and therefore that they [the people] will have to fight a perpetual war [that is, to find in themselves the resources to pay back]? ...

Are we to wait for the end of the war to stop selling properties, and offices most of all, every single day, with the right to enjoy and freely dispose of them, and a special privilege for those who will have loanable funds to this end, and then, sometime later, to resell this new title to another person, without compensation of any sort to the original purchaser, nor to the lender: which, by destroying confidence, the heart and soul of commerce, severs all ties between the prince and his subjects, and causes money alone, thanks to its being able to avoid such storms, to be estimated the one and only good, and therefore is hidden in the most obscure places of refuge that one can imagine, and causes the full termination of all sorts of consumption, to which this money is nothing but the very humble servant? ...

## Three letters on the exile and persecution of Boisguilbert

The last set of materials I share with you are letters from 1707, concerning Boisguilbert's exile and persecution.

The new book (*Factum de la France*) was banned (by an *arrêt* from March 14, 1707), and Boisguilbert was to be sent to Brive-la-Gaillarde, in Auvergne, a region located in the very center of France. At the time, he was living in Rouen (Normandy), about 130 kilometers northwest of Paris. Auvergne is 400 kilometers south of Paris, and was generally destitute at the time. What Boisguilbert did instead was to fly outside the country, perhaps to England or the Netherlands, we do not know. He came back to France and his family and himself made pleas. It is believed that he came back because the government refused that his son would take over his office of lieutenant-général<sup>4</sup> and receive the treatment of his father. He was promised a shorter stay in Auvergne, of only six months. He made his way to Auvergne and stayed two to three months. By the end of 1707 Boisguilbert was back in Rouen. During the course of 1707, seven editions of his *Factum de la France* were published. Boisguilbert died seven years later, in 1714.

*From Boisguilbert to the Controller-General (Michel Chamillart), 17 March 1707.*

This 17 March [1707].

Your excellency<sup>5</sup>,

I would very humbly ask for your forgiveness if 112,000 livres of taxes paid by myself during your ministry, and the same sum being presently required of me, have made me lose my head enough to disobey your orders, in the hope that the public would be kind enough to join its pleas to mine, to obtain from you some policies about which for a long time you have agreed to receive my opinion. I have been given an order to go to Brive-la-Gaillarde. I am convinced, your excellency, that my sentence would be smaller than my crime, if my situation was that of every other man; but as to me, leaving Rouen means asking a wife and many children to beg for food, now that nothing can be obtained from the lands, and the daily emoluments of my office being my only source of income. I started removing, everywhere that I could, copies

---

4. Boisguilbert held a charge of lieutenant-général, which is somewhat similar to a préfet in today's France, or a district attorney in the United States.

5. French has *Monseigneur* (Monsignor).

of my works, and sending to the fire my very numerous manuscripts; and if, your excellency, following the example of God, you would be merciful to me, you would be shown in the future that my repentance is genuine, and in such instances the voluntary penitence of the author is more striking than any sentence he may receive. I implore you not to let my wife and my children be punished for a crime which is only my own, and I urge your character full of goodness to grant me pardon, so that my future silence may bring you the proof of my acknowledgment.

It is with a very high respect that I am, your excellency, your very humble and very obedient servant.<sup>6</sup>

Boisguillebert.<sup>7</sup>

*From Boisguilbert to the Controller-General (Michel Chamillart), 11 April 1707.*

This 11 April [1707]

Your excellency,

I have the honor to repeat to you the pledge that I made previously while staying in a foreign land, which is: to cease speaking and writing, in any manner whatsoever, about the affairs of the state, except to you only, when and if you would only grant me a permission to do so, hoping that following the example of God, who forgets all about the past when giving pardon to sinners, you will be kind enough to allow me to come and salute you when I will be in Paris. I have burnt all my manuscripts, which were very numerous, except for a copy of the Memoirs by Mr. de Sully, in eight volumes, with my notes and some papers attached to the pages—there are one hundred of them only—thanks to which without even leafing through or opening these books, one can understand in half an hour the sort of policy put in practice by a horseman aged 35, with no prior study, to restore in three months a kingdom which was in a more pitiful state, following wars both at home and outside, than it is today, and all of this while having the whole government and the court as his sworn enemies, to the point that they wanted him to be murdered, as he was made aware by the King himself. The first principle of his policy was the free export of grain, without any taxes, permissions or passports; and in fact the king Henri IV explains in a letter of his own writing, that everything is lost when it is decided otherwise. This sole article costs today four times more than the war itself, due to the cultivation of half our lands having stopped. I have heard that you are well aware of this, but that the King argues the opposite. Perhaps if His Majesty

---

6. This is nothing but the common closing formula in letters of the time.

7. The name is given with various spellings in the archives: Boisguilbert, Boisguillebert, or even Boisguilbert (like the town, in Normandy). *Boisguilbert* is the most commonly used by historians.

would see the opinion of his ancestor, he will prefer following it than that of the first president of Paris and Mr. d'Argenson<sup>8</sup>, especially given the advice to the contrary by you as well as the public.

It is with a very high respect that I am, your excellency, your very humble and very obedient servant.

Boisguillebert.

*Answer by the Controller-General (Michel Chamillart), to the previous letter.*

Since you are still addressing me after having given to the public all your eccentricities, the only good advice that I can give you is to throw to the fire your remarks on the memoir by M. de Sully, and to convince yourself once in your lifetime that one can only make use of examples from the past when the situation is nearly in the same proportion, and when a kingdom is rich enough to bear the charges that the Kings wish to establish. If you understand well what I am saying, which is not very difficult to grasp, you will now focus on administering justice and you will stop working on the affairs of the State.

## References

- Boisguilbert, Pierre de.** 2023. *Écrits Économiques* [Economic Writings], 2 vols. Paris: Institut Coppet.
- Carpenter, Kenneth E.** 1975. *The Economic Bestsellers Before 1850*. Boston: Baker Library.
- Gournay, Vincent de.** 2023. Gournay Lives! The Freedom of Labor in a Three-Part Exchange of the 1750s. *Econ Journal Watch* 20(1): 159—203. [Link](#)
- Malbranque, Benoît.** 2023. French Liberal Economics, 1695—1776. *Just Sentiments* (Adam Smith Works, Liberty Fund, Carmel, Ind.), June 28. [Link](#)
- Mises, Ludwig von.** 1966. *Human Action: A Treatise on Economics*, 3rd ed. Chicago: Henry Regnery.
- Vauban, Sébastien le Prestre, maréchal de.** 2007. *Projet de dîme royale*. In *Oisivetés de Monsieur de Vauban*, intégrale ed., ed. Michèle Virol. Paris: Champ Vallon.

---

8. Father of the marquis known for his laissez-faire stance.

## About the Author



**Pierre de Boisguilbert** (1646–1714) was the first economist to articulate a theory of laissez-faire, or non-intervention, in France. He promoted his challenging ideas in letters sent to ministers and senior officials, as well as in unauthorized books, censored and widely read. Already translated in Italian, German and even Chinese, writings of his can now be read in samples in English in *Econ Journal Watch*.

## About the Editor



**Benoît Malbranque**, a research fellow at Institut Coppet, is studying French classical liberalism. He is the author of *Les théoriciens français de la liberté humaine* (Institut Coppet, 2020). His email address is [benoit.malbranque@institutcoppet.org](mailto:benoit.malbranque@institutcoppet.org).

[Go to archive of Watchpad section](#)

[Go to September 2023 issue](#)



# SSRN and medRxiv Censor Counter-Narrative Science

Jay Bhattacharya<sup>1</sup> and Steve H. Hanke<sup>2</sup>

[LINK TO ABSTRACT](#)

Chapter XI of Friedrich Hayek's *The Road to Serfdom* is titled "The End of Truth." Hayek develops the idea that to function and maintain power, totalitarian regimes must use propaganda to establish an official doctrine:

The need for such official doctrines as an instrument of directing and rallying the efforts of the people has been clearly foreseen by the various theoreticians of the totalitarian system. Plato's "noble lies" and Sorel's "myths" serve the same purpose as the racial doctrine of the Nazis or the theory of the cooperative state of Mussolini. They are all necessarily based on particular views about facts which are then elaborated into scientific theories in order to justify a preconceived opinion. (Hayek 1965, 157)

To be effective, propaganda must exhibit monopoly power, with no dissent allowed and "all information that might cause doubt or hesitation...withheld" (Hayek 1965, 160). Hayek grew up in Vienna and lived there until 1931. He had first-hand experience with totalitarianism and studied its operations. He dedicated much of his professional life to understanding the causes of bad government and combatting its brutalities.

So, we may ask: has covid brought an end to truth? We answer by accounting our covid research experiences, as well as those of others. Before doing so, we reflect on how matters of public health, such as covid, fit into Hayek's "The End of Truth."

There is perhaps nothing that opens the door to censorship wider than the fear of disease and the prospect of an early death. Indeed, there is nothing that

---

1. Stanford University, Stanford, CA 94305.

2. Johns Hopkins University, Baltimore, MD 21218.



matches a looming pandemic to generate fear. And there is nothing like fear to grease the skids of censorship.

In the present article, we use the word *censorship* in the sense of the first definition of the transitive verb *censor* given at Wiktionary ([link](#)): “To review for, and if necessary to remove or suppress, content from books, films, correspondence, and other media which is regarded as objectionable (for example, obscene, likely to incite violence, or sensitive).” This definition does not confine censorship to activity undertaken by governmental players. We believe that a line that is connoted by *censorship*, a line that distinguishes censorship from plain and honest content curation, has been crossed by SSRN and medRxiv. That line has to do with violating terms and conditions, even if those are only tacitly understood and are derived from established convention, and with the violations deriving from motives like those that drive government censorship. Indeed, we would not rule out the possibility governmental pressures are playing a role in the censorship perpetrated by SSRN and medRxiv.

So, how does covid fit into this picture? A signal event in the timeline of Western covid lockdowns occurred on March 16, 2020, with the publication of the Imperial College London covid report (Ferguson et al. 2020). Its frightening predictions sent shock waves around the world. The next day, the government threw the United Kingdom into lockdown.

The impact of the report was amplified by the United Kingdom’s soft-power machine, the BBC. Its reach has no equal: broadcasting in 42 languages, reaching 468 million people worldwide each week, and efficiently disseminating its message (Barber 2022). With the BBC in full cry and the public genuinely alarmed, there was little room or tolerance for dissent. In the United Kingdom, the government put its recently established Counter Disinformation Unit on full covid alert, to stamp out dissent (Investigations Team 2023).

A copycat cascade then took hold, with the United States and other countries embracing the UK government’s messaging and policies. The result was a policy based on a defective model (see Herby et al. 2023a, 28–29) that originated at Imperial College London under the leadership of Professor Neil Ferguson, who is the director of Imperial College’s School of Public Health.

UK policymakers should have been aware that Professor Ferguson’s Imperial College team had a history of defective modeling and a track record littered with what are little more than fantasy numbers. To put the blunders of the Imperial College London’s epidemiological fear machine into context, consider the numbers generated by the modelers in 2005, when Professor Ferguson suggested that “up to around 200 million” could die from bird flu globally. He justified this claim by comparing the lethality of bird flu to that of the 1918 Spanish flu outbreak, which killed 40 million (Sturcke 2005). By 2021, bird flu had only killed 456 people

worldwide (WHO 2021). And, there were other huge misses by the Imperial College London's modelers: foot and mouth in 2001, mad cow in 2002, and swine flu in 2009 (Hanke and Dowd 2022).

## SSRN censors Herby, Jonung, and Hanke

The Social Science Research Network (SSRN), operated by Netherlands-based publishing house Elsevier, is an enormous platform in academia. SSRN is extremely important to professors for disseminating their research results and for advancing within the academic community. Indeed, as of September 25, 2023, the SSRN provided 1,271,271 research papers from 1,381,280 researchers in 70 disciplines, per its homepage ([link](#)). SSRN is the leading site for academic working papers. It platforms research irrespective of whether the research has been submitted, accepted, or published in journals. Such platforming of research is vital for giving space to all voices, to make scholarship contestable and challengeable, and without delay or stonewalling. The gatekeeping role of the journals is also a vital facet of the common enterprise of scholarship—but it is separate from the realm of the working paper, and it should continue to be kept separate. If SSRN wishes to play the role of scholarly gatekeeper, it is obliged to assume that role responsibly, which includes clarifying its gatekeeping procedures and living up to the procedures it purports to follow, avoiding double-standards.

Here is what SSRN says (as of September 25, 2023) about its policies regarding “Contributed Content” ([link](#)):

Contributed content should be relevant to the subject scope of SSRN. Content may not be illegal, obscene, defamatory, threatening, infringing of intellectual property rights, invasive of privacy or otherwise injurious or objectionable. Elsevier does not pre-screen or regularly review any contributed content, but Elsevier has the right (though not the obligation) to monitor submissions to determine compliance with these Terms and any operating rules to satisfy any law, regulation, or authorized government request.

Needless to say, none of the SSRN-censored research products mentioned in the present article runs afoul of the guidelines against material that is “illegal, obscene, defamatory, threatening, infringing of intellectual property rights,” or “invasive of privacy.” As for “otherwise injurious or objectionable,” if that is the grounds for SSRN's censorship, it is something they should declare and defend. Thus far, SSRN has failed to do so.

Jonas Herby, Lars Jonung, and Steve H. Hanke (hereafter HJH) undertook a major meta-analysis of the effects of lockdowns. Before conducting their study,

HJH wrote up a protocol for the methods they would use, so as to head off charges about cherry-picking. On July 15, 2021, SSRN published the protocol (HJH 2021). The protocol laid out in detail what HJH proposed to do, how they were going to do it, and what, in fact, they did do. That protocol is still available at SSRN.

HJH completed their research as outlined in the protocol. They published the first edition as a working paper, in the series *Johns Hopkins Studies in Applied Economics*, in January 2022 (HJH 2022a). On February 2, 2022, Dr. Marty Makary, a Johns Hopkins professor of medicine, appeared on the Fox News Channel television program *Tucker Carlson Tonight* and praised the HJH meta-analysis of covid lockdowns (Manno 2022). Two days later, a reporter asked about the HJH meta-analysis in the White House (White House 2022).

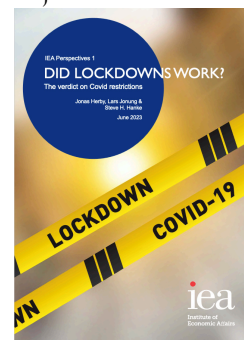
The HJH study reached the conclusion that lockdowns only had a tiny impact on covid mortality and came with huge economic and social costs. The HJH meta-analysis estimated that lockdowns only saved between 6,000 and 23,000 lives in Europe and between 4,000 and 16,000 lives in the United States (see HJH 2023a). For context, lockdowns prevented relatively few deaths compared to a typical flu season. In Europe, 72,000 flu deaths occur (WHO 2023) and in the United States, 38,000 flu deaths occur during a typical flu season (CDC 2022). As a result, HJH concluded that covid lockdowns were a major public policy blunder.

On March 21, 2022, the SSRN posted a critique of HJH (2022a), a critique by Nicolas Banholzer, Adrian Lison, and Werner Vach (2022).

Then, HJH published a second edition of their working paper in the *Johns Hopkins Studies in Applied Economics* series in May 2022 (HJH 2022b). Following its publication, HJH repeatedly requested (on May 25, 2022, June 15, 2022, July 23, 2022, and July 5, 2023) that the second edition of their working paper be published by the SSRN. But SSRN refused, citing “the need to be cautious about posting medical content” (SSRN Author Comment Notification, email message to author, June 15, 2022; August 5, 2022). We infer from SSRN’s actions that the HJH study was deemed “injurious or objectionable.” But neither a critique of the HJH study nor the protocol for the study had been deemed “injurious or objectionable.”

HJH polished and expanded the second edition of their working paper. Following a heavy peer-review, it was published in June 2023 as a book *Did Lockdowns Work? The Verdict on Covid Restrictions* by the Institute of Economic Affairs in London (HJH 2023a). That book received considerable press coverage in the United Kingdom, but only light coverage in the United States.

In August 2023, Herby, Jonung, and Hanke published another working paper in the *Johns Hopkins Studies in Applied Economics* series (HJH 2023b). It was a reply to Banholzer,



Lison, and Vach's criticism that had been posted at the SSRN. HJH attempted to post this reply at SSRN. But on August 16, 2023, SSRN refused to post it, using the same language that was used in the prior HJH working paper rejections: "Given the need to be cautious about posting medical content, SSRN is selective on the papers we post. Unfortunately, your paper has not been accepted for posting on SSRN" (SSRN Author Comment Notification, email message to author, August 16, 2023).

How is the censorship perpetrated by SSRN related to the larger theme of recklessly meandering down a road that leads to the end of truth? The experience of HJH illustrates what, to us, anyway, seems to be a syndrome followed by officialdom and allied factions. (Recent revelations make it clear that government actors pervasively exercise clandestine influence over the media.) First come the "fact-checkers" who produce unfounded, irrelevant verbiage that lacks critical sense or analytical insight (for example, Evon 2022). Next come hit pieces that echo the claims of the so-called fact-checkers. The perpetrators hope that a cone of silence, aided by censorship at preprint servers, will descend on the counter-narrative scientific findings. They lead people down a road that leads to the end of truth.

## **SSRN and medRxiv censor Vinay Prasad and coauthors**

Dr. Vinay Prasad, a physician-epidemiologist and professor at the University of California at San Francisco medical school, documents (Prasad 2023) how SSRN systematically suppressed scientific papers from his laboratory, papers that contained findings that were at odds with the government's policies and pronouncements on the covid vaccines, on mask mandates, and even on standard of evidence appropriate for meta-analyses (see Figure 1). Papers on cancer research and oncology by Prasad and colleagues have never been censored at SSRN. But, with two exceptions, all of his covid papers have been censored.

SSRN even censored an article about SSRN's censorship! They found it objectionable, no doubt. That paper was written by Prasad and his colleague, epidemiologist Dr. Alyson Haslam (Haslam and Prasad 2023).

In every case, the reason provided to Prasad by the SSRN for its decision to censor his papers is the same: "Given the need to be cautious about posting medical content, SSRN is selective on the papers we post."

**Figure 1.** SSRN decisions on scientific covid research from Dr. Vinay Prasad's laboratory

DECISION	TITLE, DATE, METRICS*	REASONING
<b>SSRN</b>		
Removed	Mask mandates and COVID-19: A Re-analysis of the Boston school mask study July 19 2023	"Given the need to be cautious about posting medical content, SSRN is selective on the papers we post."
Removed	A comprehensive analysis of articles submitted to preprint servers from one laboratory (VKPrasad Lab at UCSF): Download statistics, rates of rejection, and reasons for rejection: Are preprint servers acting fairly or playing politics? July 18 2023	"Given the need to be cautious about posting medical content, SSRN is selective on the papers we post."
Removed	Falsification Endpoints and the Pitfalls of Using Observational Studies in Reviews of Effectiveness of COVID-19 Vaccines: Critique of a Systematic Review and Meta-Analysis of BNT161b2 vaccination of 5-11-Year-Olds April 19 2023	"Given the need to be cautious about posting medical content, SSRN is selective on the papers we post."
Removed	Interpretation of Wide Confidence Intervals in Meta-Analytic Estimates: Is the 'Absence of Evidence' 'Evidence of Absence'? April 13 2023	"Given the need to be cautious about posting medical content, SSRN is selective on the papers we post."
Accepted	Statistical and Numerical Errors Made by the US Centers for Disease Control and Prevention During the COVID-19 Pandemic March 14 2023 40,213 Views 7,959 Downloads	
Removed	COVID-19 vaccines: history of the pandemic's great scientific success & flawed policy implementation November 14 2022 3,072 Views 1,028 Downloads**	"Given the need to be cautious about posting medical content, SSRN is selective on the papers we post."
Accepted	COVID-19 Vaccine Boosters for Young Adults: A Risk-Benefit Assessment and Five Ethical Arguments against Mandates at Universities August 31 2022 417,488 Views 90,252 Downloads	
Submitted by journal	A Systematic Analysis of Post-Protocol Therapy in First Line Checkpoint Inhibitor Trials October 29 2021 283 Views 19 Downloads	
Accepted	Why Is Research in Early-Stage Cancer Research so Low? A Re-Assessment of Budish, Roin and Williams June 6 2017 2,454 Views 239 Downloads	

Source: Prasad 2023 ([link](#)).

SSRN is not the only preprint server that has suppressed scientific papers. During the pandemic, medRxiv, the largest preprint server in medicine and health sciences, also systematically suppressed scientific findings that it deemed at odds with government covid narratives. Again, Haslam and Prasad (2023) provide documentation for Prasad's laboratory (see Figure 2). While medRxiv accepted a larger proportion of papers that Prasad's group submitted to the server, it rejected two counter-narrative papers. The first was an evaluation of errors in statistics committed by the U.S. CDC during the pandemic (Krohnert et al. 2023). The medRxiv proprietors refused to post the paper claiming that the paper "is not a systematic evaluation with reproducible methodology." By that standard, medRxiv should have refused to publish a large proportion of papers currently published on its website. (Meanwhile, over at SSRN, that piece was *not* censored.) The second paper medRxiv rejected was the Haslam and Prasad (2023) paper that documents censorship by medRxiv and the SSRN. (As already noted, that one was, evidently, deemed objectionable over at SSRN, as they too refused to post it.)

**Figure 2.** medRxiv decisions on scientific covid research from Dr. Vinay Prasad's laboratory

medRxiv			
Removed	A comprehensive analysis of articles submitted to preprint servers from one laboratory (VKPrasad Lab at UCSF): Download statistics, rates of rejection, and reasons for rejection: Are preprint servers acting fairly or playing politics?		
	July 14 2023		
Accepted	Interpretation of wide confidence intervals in meta-analytic estimates: Is the 'Absence of Evidence' 'Evidence of Absence'?		
	July 11 2023	277 Views	52 Downloads
Accepted	Changes in Masking Policies in US Healthcare Facilities in the First Quarter of 2023: Do COVID-19 Cases, Hospitalizations, or Local Political Preferences Predict Loosening Restrictions?		
	July 11 2023	502 Views	435 Downloads
Accepted	Characteristics and quality of studies pertaining to masks published in the Morbidity and Mortality Weekly Report		
	July 11 2023	2,148 Views	1,797 Downloads
Accepted	Analysis of tweets discussing the risk of Mpox among children and young people in school (May-Oct 2022): Public health experts on Twitter consistently exaggerated risks and infrequently reported accurate information		
	May 16 2023	1,364 Views	920 Downloads
Submitted by journal	Cross-sectional analysis of Open payments for physicians at designated hemophilia centers in the US (2018-2020)		
	March 7 2023	659 Views	92 Downloads
Accepted	Current landscape of disparity-focused research: a bibliometric analysis of 260 research articles		
	March 7 2023	510 Views	133 Downloads
Removed	Statistical and numerical errors made by the US Centers for Disease Control During the COVID-19 Pandemic		
	March 6 2023		
Accepted	An empirical analysis of lay media coverage on influenza prevention pre- and post-COVID 19: Mask recommendations were previously rare, now ubiquitous		
	February 14 2023	1,306 Views	202 Downloads
Submitted by journal	Estimation of time cost of anti-cancer drugs approved based on comparisons to best supportive care: a cross sectional analysis medRxiv		
	June 22 2022	1,090 Views	361 Downloads

Source: Prasad 2023 ([link](#)).

## medRxiv censors Bhattacharya and coauthors

Jay Bhattacharya, too, has had medRxiv refuse to post one of his working papers. In late 2020, a team that included Bhattacharya, Christopher Oh, and John Ioannidis and led by Stanford University infectious disease professor Eran Bendavid conducted an empirical analysis of the effectiveness of early 2020 shelter-in-place orders and business closures on the spread of the pandemic. Using sub-national data, the analysis compared places that did not have mandatory orders, like Sweden and South Korea, against places that did. The paper failed to show any statistically significant effect of mandatory orders on covid spread. The authors of this paper uploaded it to medRxiv as a preprint, while simultaneously submitting it for peer review. The medRxiv refused to post the piece, telling the authors that the topic was too sensitive to permit the publication of a preprint, even though the site teemed with modeling analyses purporting to demonstrate the efficacy of lockdowns in limiting the spread of covid. In early January 2021, the peer-reviewed



journal *European Journal of Clinical Investigation* published the paper (Bendavid et al. 2021). To date, the article has garnered 245 citations and an Altmetrics score that places it in the top 56 papers among the 24.5 million papers tracked by Altmetrics.

Before the pandemic, medRxiv provided little to no content-based screening of the preprints it published on its site. However, in May 2020, the site announced that it would no longer permit the preprint publication of “manuscripts making predictions about treatments for covid-19 solely on the basis of computational work” (Kwon 2020). But, computational simulations were at the heart of government lockdown campaigns. Preprint servers had had no problem posting such work, even as the results from many such studies were, in our judgment, patently preposterous. But medRxiv deemed it too dangerous to permit computation biologists to engage with clinical trialists and other doctors and scientists regarding potential treatments for covid.

What SSRN and medRxiv need to do to rectify the current state of affairs is not complicated. What they need to do is repent and desist from censorship.

However, as C. S. Lewis explained in *Mere Christianity* (ch. 4), those most in need of repenting are often the least capable of repenting.

To help us return to the right path, Stephen Walker has created a website called Vicegerents.org ([link](#)), where scholars can tell of their experiences of being censored by SSRN, medRxiv, or other preprint servers. The website takes its name from a passage in Adam Smith’s *The Theory of Moral Sentiments*, a passage displayed atop the frontpage of the Vicegerents.org site. In that book, Smith (1790, 337) also wrote: “Frankness and openness conciliate confidence.” If you have been censored by SSRN, medRxiv, or other preprint servers, you might wish to visit Vicegerents.org.

**Note from the editors of EJW:** *We invite officers of SSRN and medRxiv to provide a reply to this article, for publication in a future issue of the journal.*

## References

- Banholzer, Nicolas, Adrian Lison, and Werner Vach.** 2022. Comment on ‘A Literature Review and Meta-Analysis of the Effects of Lockdowns on COVID-19 Mortality.’ Working paper (SSRN), February 11. [Link](#)
- Barber, Lionel.** 2022. The BBC at 100—the Past and Future of a British Institution. *Financial Times*, February 8. [Link](#)
- Bendavid, Eran, Christopher Oh, Jay Bhattacharya, and John P. A. Ioannidis.** 2021. Assessing Mandatory Stay-at-Home and Business Closure Effects on the Spread of COVID-19. *European Journal of Clinical Investigation* 51(4): e13484. [Link](#)
- CDC.** 2022. Disease Burden of Flu. October 4. Centers for Disease Control and Prevention



- (Atlanta). [Link](#)
- Evon, Dan.** 2022. Here's What We Know About 'Johns Hopkins Study' on Lockdowns. *Snopes*, February 6. [Link](#)
- Ferguson, Neil M., et al.** 2020. Impact of Non-Pharmaceutical Interventions (NPIs) to Reduce COVID-19 Mortality and Healthcare Demand. March 16. Imperial College London. [Link](#)
- Haslam, Alyson, and Vinay Prasad.** 2023. A Comprehensive Analysis of Articles Submitted to Preprint Servers from One Laboratory (VKPrasad Lab at UCSF): Download Statistics, Rates of Rejection, and Reasons for Rejection: Are Preprint Servers Acting Fairly or Playing Politics? Working paper (Zenodo), July 24. [Link](#)
- Hanke, Steve H., and Kevin Dowd.** 2022. Imperial College's Fear Machine. *National Review*, March 30. [Link](#)
- Hayek, Friedrich A.** 1965 [1944]. *The Road to Serfdom*. Chicago: University of Chicago Press.
- Herby, Jonas, Lars Jonung, and Steve H. Hanke (HJH).** 2021. Protocol for 'What Does the First XX Studies Tell Us about the Effects of Lockdowns on Mortality? A Systematic Review and Meta-Analysis of COVID-19 Lockdowns.' Working paper (SSRN). [Link](#)
- Herby, Jonas, Lars Jonung, and Steve H. Hanke.** 2022a. A Literature Review and Meta-Analysis of the Effects of Lockdowns on COVID-19 Mortality. *Johns Hopkins Studies in Applied Economics* 200. Institute for Applied Economics, Global Health, and the Study of Business Enterprise, Johns Hopkins University (Baltimore, Md.). [Link](#)
- Herby, Jonas, Lars Jonung, and Steve H. Hanke.** 2022b. A Literature Review and Meta-Analysis of the Effects of Lockdowns on COVID-19 Mortality – II. *Johns Hopkins Studies in Applied Economics* 210. Institute for Applied Economics, Global Health, and the Study of Business Enterprise, Johns Hopkins University (Baltimore, Md.). [Link](#)
- Herby, Jonas, Lars Jonung, and Steve H. Hanke.** 2023a. *Did Lockdowns Work? The Verdict on Covid Restrictions*. London: Institute of Economic Affairs. [Link](#)
- Herby, Jonas, Lars Jonung, and Steve H. Hanke.** 2023b. A Reply to "Comment on 'A Literature Review and Meta-Analysis of the Effects of Lockdowns on COVID-19 Mortality.'" *Johns Hopkins Studies in Applied Economics* 237. Institute for Applied Economics, Global Health, and the Study of Business Enterprise, Johns Hopkins University (Baltimore, Md.). [Link](#)
- Investigations Team.** 2023. Secret Government Unit Collected Telegram Posts about Covid Policy Critic. *Telegraph*, June 7. [Link](#)
- Krohnert, Kelley, Alyson Haslam, Tracy Beth Høeg, and Vinay Prasad.** 2023. Statistical and Numerical Errors Made by the US Centers for Disease Control and Prevention During the COVID-19 Pandemic. Working paper (SSRN), March 7. [Link](#)
- Lewis, C. S.** 2002 [1952]. *Mere Christianity*. In *The Complete C. S. Lewis Signature Classics*, 1–177. San Francisco: HarperOne.
- Manno, Adam.** 2022. Johns Hopkins Professor Blasts His College and Media for Down-playing Study on Covid Lockdowns. *Daily Mail*, February 3. [Link](#)
- Prasad, Vinay.** 2023. Preprint Servers Have Repeatedly Censored Our Work on COVID-19 Policy. *Sensible Medicine* (sensibled.substack.com), July 27. [Link](#)

- Smith, Adam.** 1976 [1790]. *The Theory of Moral Sentiments*, ed. D. D. Raphael and A. L. Macfie. Oxford, UK: Oxford University Press.
- Sturcke, James.** 2005. Bird Flu Pandemic ‘Could Kill 150M.’ *Guardian*, September 30. [Link](#)
- White House.** 2022. Press Briefing by Press Secretary Jen Psaki. February 4. White House (Washington, D.C.). [Link](#)
- WHO.** 2021. Cumulative Number of Confirmed Human Cases for Avian Influenza A(H5N1) Reported to WHO, 2003–2021. October 1. World Health Organization (Geneva). [Link](#)
- WHO.** 2023. Estimating Disease Burden of Influenza. World Health Organization (Geneva). [Link](#)

## About the Authors



**Jay Bhattacharya** is a professor at Stanford University Medical School currently teaching in the Department of Health Policy. He holds an M.D. and Ph.D. in economics, earned at Stanford University. He is a research associate at the National Bureau of Economic Research and a founding fellow of the Academy of Science & Freedom at Hillsdale College. Dr. Bhattacharya co-wrote the Great Barrington Declaration, a focused-protection alternative to lockdowns. Dr. Bhattacharya’s research focuses on the health and well-being of vulnerable populations, emphasizing the role of government programs, biomedical innovation, and economics. He has published over 165 papers in top peer-reviewed journals in medicine, epidemiology, economics, statistics, health policy, and public health. His email address is [jay@stanford.edu](mailto:jay@stanford.edu).



**Steve H. Hanke** is a Professor of Applied Economics and Founder and Co-Director of the Institute for Applied Economics, Global Health, and the Study of Business Enterprise at the Johns Hopkins University in Baltimore. Hanke holds seven honorary doctorate degrees and is an honorary professor at four foreign institutions. Hanke served on President Reagan’s Council of Economic Advisers and has served as adviser to five foreign heads of state and five foreign cabinet ministers. He was President of Toronto Trust Argentina in Buenos Aires, the world’s best-performing emerging market mutual fund in 1995. His email address is [hanke@jhu.edu](mailto:hanke@jhu.edu).

[Go to archive of Watchpad section](#)  
[Go to September 2023 issue](#)



# *Journal of Accounting Research's* **Report on Its Own Research- Misconduct Investigation of an Article It Published**

Daniel B. Klein<sup>1</sup>

[LINK TO ABSTRACT](#)

I comment here on a document by *Journal of Accounting Research* (JAR). The document presents itself as a report on its own research-misconduct investigation of a problem in an article it published in 2020. Here I refer to the document as the JAR Report. JAR's investigation was undertaken in response to a call by Stephen Walker, a call which he made publicly in the September 2022 issue of *Econ Journal Watch* (EJW) (see [here](#)).

On June 5, 2023, JAR emailed the JAR Report to Walker. Walker forwarded it to me. I refrain from quoting the JAR Report, which is 1,077 words, because doing so is not necessary to clarify its nature and the judgment it conveys, because it is not published, and because the first one to publish it ought to be JAR itself.

The JAR Report is a response to a request submitted by Walker to JAR to investigate a problem that Walker identified in the 2020 JAR article.

The present comment is not to allege misconduct on the part the author group of the 2020 JAR article. The issue here is JAR's conduct. Walker called for a research-misconduct investigation and JAR took up the call and sent Walker a report on the investigation. The question is: Did JAR conduct and report on a research-misconduct investigation in a way that is at all satisfactory? I will argue that it did not. I go further and say the JAR Report is a disgrace.

The issue is of cardinal importance to science, because if a problem in research raises the question of research misconduct and that question is not meaning-

---

1. George Mason University, Fairfax, VA 22030.

fully and faithfully addressed by the guardians of science, the scientific project becomes a sham.

In the present comment I speak of “the author group,” because the focus here is on the JAR Report, not that author group.

Walker made his call for a research-misconduct investigation, by email, to JAR on August 11, 2022. Eight weeks later, JAR emailed Walker, clarifying how to submit such a request. Walker promptly replied to JAR’s satisfaction on October 5, 2023. Eight months later, on June 5, 2023, JAR sent the JAR Report to Walker.

Walker has published three articles on this matter, all three in EJW. The author group has published one article in EJW. Here is a list of the six articles that preceded Walker’s request for a research-misconduct investigation:

1. The 2020 JAR article ([link](#)).
2. Walker’s March 2021 critique in EJW of the 2020 JAR article ([link](#)).
3. The author group’s March 2021 reply in EJW to Walker’s March 2021 critique ([link](#)).
4. Walker’s September 2021 rejoinder in EJW to the author group ([link](#)). (Before the rejoinder was published, the author group was invited to reply to it, either for concurrent publication or anytime afterward, but they declined.)
5. The author group’s September 2022 “Erratum” in JAR ([link](#)).
6. Walker’s September 2022 critique in EJW of the “Erratum” ([link](#)). (Immediately after Walker’s critique of the “Erratum” was published, the author group was invited to reply to it for subsequent publication, but they have not done so; the invitation remains open.)

In the author group’s “Erratum” in JAR, the opening words of the first paragraph (p. 1635) are as follows: “Walker (2021a and 2021b) identified an error **in the program codes** of Bao et al. (2020) posted on Github that led to an overstatement of model performance metrics. This erratum corrects this error...” (boldface added).

The paragraph goes on to say: “Bao et al.’s (2020) **intent was to recode** all spanning serial fraud observations in the training period as nonfraud. However, the **codes** posted on Github **contained an error** that resulted in approximately 10% of these spanning serial fraud cases in the training period not being recoded as zero, which in turn helped model performance” (boldface added).

Thus, in the “Erratum” the author group confess that there was a problem—of some kind—in the 2020 JAR article.

Later in the “Erratum” (p. 1636) they speak of the problem as a “coding error.” Also, the JAR Report itself employs the expression “coding error” to refer

to the problem.

Whether the problem is aptly denominated a “coding error” is one thing that the JAR investigation should have investigated and reported on. But the JAR Report offers nothing on the problem that is at the heart of the entire matter. Again, the JAR Report is 1,077 words. Of that, 492 words are given to the title and a recounting of the materials in question, a description of what JAR has done in their putative investigation, including a listing of eleven documents distributed to Reviewers. Next come the presentations of two quotations from Reviewers of the materials, taking 197 words; both Reviewer quotations express an overall judgment that research misconduct has not been definitively established but not giving a single word to investigation of the “coding error.” Next come 57 words stating that the JAR senior editors agree and that the case is closed. Lastly, there are 331 words given to the matter that, in the “Erratum,” once the author group corrected the core problem, the preferred test sample period changes from 2003–2008 (as in the 2020 article) to 2003–2005. This matter is downstream of correcting the core problem and does nothing toward investigation of the core problem itself. The segments of 492, 197, 57, and 331 words sum to 1,077. There are zero words given to investigating the core problem.

It is now agreed by all that there was a core problem in the 2020 JAR paper. Walker has described at length the nature of that problem. Neither the author group nor the JAR Report, however, so much as identify what the core problem was, why it occurred, nor explain why it should be called a “coding error.” The author group simply denominate the core problem a “coding error” and then proceed to correct it.

At the outset of his 2022 critique of the “Erratum,” Walker quotes the “Erratum” where it says that he (Walker) “identified an error in the **program codes**.” Walker then says:

The part in boldface is false. I discovered no errors in their program codes. Rather, the ‘error’ was in their dataset—specifically, how they identified fraud cases in their sample. A coding error may be accidental, but the method used to define their fraud cases was deliberate. To this date, the authors have offered no explanation as to why they did what they did, nor offer any explanation as to what prompted their misidentifying of fraud cases. (Walker, p. 190 [here](#))

Later in the piece, Walker writes:

In fact, ‘error’ is not the correct word to describe the issue. The code worked fine. Rather, it was the manipulation of the dataset itself where 17 unique fraud cases in the dataset received two identifiers, a manipulation of the underlying data which made no sense. (Walker, p. 193 [here](#))

Walker is pointing out some of the questions that a real research-misconduct investigation would ask:

1. What is the nature of the problem? That should be clearly laid out in any Report on a research-misconduct investigation.
2. When the actions giving rise to the problem were taken, what was intended, as opposed to what was actuated? As an analogy, if someone mistypes a word, what was the word he intended to type? If someone misquotes a source, what is the quotation he intended to reproduce? If the problem arises from an error in judgment, what was the state of mind when the person erred? What was he thinking? In other words, why did the problem arise?
3. Walker in 2022 notes: "If an erratum were in order, why didn't they say that in their March 2021 reply to me? In that reply, they act as though nothing in their 2020 JAR article is amiss. Furthermore, when invited by *Econ Journal Watch* to supply a second reply, why didn't they take that opportunity to say that an erratum was in order?" JAR should have put these very questions to the author group and reported on their replies.

When a problem that arouses suspicion of misconduct occurs, investigators must ask and report on the questions posed above. If the parties in question say, "Sorry, we made an innocent mistake," or a "coding error," and the investigators do not probe the credibility of those excuses, they simply are not investigating. If all research-misconduct investigations are handled in such a way that misconduct is never to be found, misconduct is not handled justly. There must be a line somewhere, and the responsibility of investigators is to test matters against that line. In doing so, they both establish and evaluate evidence in the case at hand, and they help to clarify and establish the line that we all need to preserve integrity in our common project of scholarship and science.

My present remarks echo some of the insights and concerns expressed by Ian Gow in his 2022 SSRN paper "Should Bao et al. (2020) Be Retracted?" ([link](#)) and in his heavily downloaded 2023 SSRN paper "The Elephant in Room: P-Hacking in Accounting Research" ([link](#)). The latter includes the following block quotation. The quotation refers to things that I have not referenced but makes several of the points made here:

That the research community is relatively understanding about coding errors perhaps explains efforts by researchers to attribute issues in their papers to "coding errors." For example, the code repository for Boissel and Matray (2022; retraction [here](#)) included the line *replave*  $B = B/1.8$  if  $t > -3$  &  $t < 0$ . This code modifies two coefficients in a way [that] enhances a plot used to support

a claim of “parallel trends.” While the responding author attributes this to a “coding error” it is difficult to imagine what the correct version of this line of code would be. The “coding error” in Bao et al. (2020) differs from that in Boissel and Matray (2022) in a number of respects. First, no code containing the claimed error was provided. Second, producing code with this issue accidentally seems even less plausible than with the line above from Boissel and Matray (2022; see Gow, 2022 for an attempt to replicate the “coding error”). Third, the authors’ efforts to attribute the issue to a “coding error” is belied by earlier efforts to suggest it was an appropriate research design choice (Bao et al., 2021). (Gow 2023 [here](#), p. 12 n.31)

Again, I encourage JAR to make its Report public. Also, EJW invites JAR to reply to my comment here.

In conducting and reporting the research-misconduct investigation, JAR is, in a significant way, acting as judge in its own case, for if misconduct occurred, it reflects on the editors and reviewers at JAR, not only of the 2020 article but also the 2022 “Erratum.”

In England in 1775, Parliamentarians and other magistrates had to act as judge in their own case. These words from Edmund Burke ([link](#)) seem apropos:

We are indeed, in all disputes with the Colonies, by the necessity of things, the judge. It is true, Sir. But I confess, that the character of judge in my own cause is a thing that frightens me. Instead of filling me with pride, I am exceedingly humbled by it. I cannot proceed with a stern, assured, judicial confidence, until I find myself in something more like a judicial character.

## About the Author



**Daniel Klein** is the chief editor of *Econ Journal Watch* and professor of economics and JIN Chair at the Mercatus Center at George Mason University, where he and Erik Matson lead a program in Adam Smith. His email address is [dklein@gmu.edu](mailto:dklein@gmu.edu).

[Go to archive of Watchpad section](#)  
[Go to September 2023 issue](#)





# What Are Your Most Underappreciated Works?: Second Tranche of Responses

Econ Journal Watch

[LINK TO ABSTRACT](#)

prologue by **Daniel B. Klein**

At a scholar's Scholar Profile page at [Google Scholar](#)—for example [this one for Angus Deaton](#)—one finds a list of the scholar's publications, the citation count for each, and the scholar's h-index. The h-index is the largest number  $h$  such that  $h$  publications have at least  $h$  citations.

We posted an [open invitation](#) to scholars working primarily in the social sciences and/or humanities with at least 4,000 Google Scholar citations, asking them to identify one or two publications with publication date 2012 or prior, for which the citation count is lower than their present h-index, that they consider underappreciated. We encouraged them to remark briefly on why they selected the publication, and to provide a link to it.

We believed that this project would be useful, first, because a scholar herself is likely to be a good judge of what work of hers is underappreciated and therefore this project will alert people to works worthy of greater attention, and, second, because the selection she makes here will inform understandings of that scholar herself.

In the September 2022 issue of *Econ Journal Watch* we published a first tranche of responses received from Doug Allen, Niclas Berggren, Christian Bjørnskov, Peter Boettke, Nick Bostrom, Bryan Caplan, Joshua Gans, Terri Griffith, Zoe Hilton, Dan Klein, Douglas Noonan, Michael Ostrovsky, Sam Peltzman, Eric Rasmusen, Paul Rubin, Steve Sheffrin, Stefan Voigt, and Richard Wagner ([link](#)).

Here now are additional responses from Andrew Gelman, Robert Kaestner, Robert A. Lawson, George Selgin, Ilya Somin, and Alexander Tabarrok.

## response from Andrew Gelman

Gelman, Andrew. 2004. Treatment Effects in Before-After Data. In *Applied Bayesian Modeling and Causal Inference from an Incomplete Data Perspective*, ed. A. Gelman and X. L. Meng, 195–202. London: Wiley.

It is standard practice to fit regressions using an indicator variable for treatment or control; the coefficient represents the causal effect, which can be elaborated using interactions. My article from 2004 argues that this default class of models is fundamentally flawed in considering treatment and control conditions symmetrically. To the extent that a treatment “does something” and the control “leaves you alone,” we should expect before-after correlation to be higher in the control group than in the treatment group. But this is not implied by the usual models.

My article presents three empirical examples from political science and policy analysis demonstrating the point. The article also proposes some statistical models. Unfortunately, these models are complicated and can be noisy to fit with small datasets. It would help to have robust tools for fitting them, along with evidence from theory or simulation of improved statistical properties. I still hope to do such work in the future, in which case perhaps people will see in this work the merit that I am hoping it has.

## response from Robert Kaestner

Kaestner, Robert, and Jeffrey H. Silber. 2010. Evidence on the Efficacy of Inpatient Spending on Medicare Patients. *Milbank Quarterly* 88(4): 560–594. [Link](#)

The abstract, from 2010, follows: It is widely believed that a significant amount, perhaps as much as 20 to 30 percent, of health care spending in the United States is wasted... This article uses Medicare claims data to study the association between inpatient spending and the thirty-day mortality of Medicare patients admitted to hospitals between 2001 and 2005 for surgery (general, orthopedic, vascular) and medical conditions (acute myocardial infarction [AMI], congestive heart failure [CHF], stroke, and gastrointestinal bleeding). Estimates from the

analysis indicated that except for AMI patients, a 10 percent increase in inpatient spending was associated with a decrease of between 3.1 and 11.3 percent in thirty-day mortality, depending on the type of patient. Although some spending may be inefficient, the results suggest that the amount of waste is less than conventionally believed, at least for inpatient care.

## response from Robert A. Lawson

Lawson, Robert A., and Jayme S. Lemke. 2012. Travel Visas. *Public Choice* 153: 17–36. [Link](#)

Travel visas seem to really have big effects but very little attention goes to this policy.

## response from George Selgin

Selgin, George, and John L. Turner. 2011. Strong Steam, Weak Patents, or the Myth of Watt’s Innovation-Blocking Monopoly, Exploded. *Journal of Law and Economics* 54(4): 841–861. [Link](#)

The patent that James Watt and Matthew Boulton secured, and then had extended, for Watt’s external condenser, is often treated as Exhibit A in arguments to the effect that industrial patents hinder innovation. Historians and others often say that, in this case, the patent delayed the advent of high-pressure steam technology. The hitch in this argument, John Turner and I point out, is that high-pressure engines don’t require condensers, separate or otherwise—a point surprising numbers of economic historians appear to have overlooked. The truth, we explain, is that high-pressure steam technology was considered too dangerous to toy with—Watt actually tried to have it outlawed. It was only for the sake of *evading* Boulton and Watt’s patent that Richard Trevithick risked experimenting with high-pressure steam, after others had given it up as hopeless. In short, far from hindering the development of ultimately superior and, eventually, safe high-pressure steam technology, Boulton and Watt’s patent actually served, inadvertently, to inspire it.

## response from Ilya Somin

Somin, Ilya. 2011. What if *Kelo v. City of New London* Had Gone the Other Way? *Indiana Law Review* 45: 21–39. [Link](#)

The article challenges the conventional wisdom on the impact of one of the Supreme Court’s most controversial modern rulings, which held that government could take property for private “economic development.” Some have suggested that, because defeat led to a major political backlash against eminent domain abuse, the decision actually aided the cause of property rights protection. I argue that property rights advocates would have been better off had they won the case. In addition, I develop a more general framework for using counterfactual analysis to assess the impact of court decisions. The latter is relevant far beyond the specific context of the *Kelo* case.

## response from Alexander Tabarrok

Tabarrok, Alexander. 2013. Private Education in India: A Novel Test of Cream Skimming. *Contemporary Economic Policy* 31(1): 1–12. [Link](#)

Students in private schools often outperform their counterparts in public schools. But what drives this difference, superior teaching or selective admission of high-performing students? If selective admission (or ‘cream skimming’) is the explanation for the higher scores on standardized tests, then the average test score across all students—both public and private—should remain constant as the share of private schooling increases. Just shuffling students around wouldn’t change the average. But if better teaching is the explanation, the average score should increase as more students attend private schools. I test this hypothesis in India, a country where the percentage of students in private schools varies dramatically between districts—from as low as 5% to as high as 70–80%. The public-private average score increase as the share of students in private schools increases, suggesting that better teaching methods are the driving factor.

It’s simple, but I think telling. I hope the paper finds its audience. Private schooling in India *per se* is important.

[Go to archive of Watchpad section](#)  
[Go to September 2023 issue](#)