

## WHAT IS SPURIOUS? AN INQUIRY ACROSS METHODS<sup>1</sup>

**Presented to:**  
**ECPR Budapest Conference**  
**Panel 20-5**  
**Looking at Methodological Issues Normatively**  
**9 September 2005**

### **ABSTRACT**

---

What findings qualify as negative, unexpected, mistaken or unpublishable under major methodological approaches, and what does this tell us about the state of methods and of the discipline? The paper will consider theoretical, empirical and normative propositions on ‘spuriousness’ applied to political science, the social sciences and other fields of research. ‘Spurious’ will be employed as a general term to refer not to what is false, falsified or falsifiable but rather to the disciplinary effects of a broader category of results that are ‘unpublishable’ whether false or not. The paper will consider the potential utility of schemes for returning the ‘by-products’ of the research process to the shared public sphere of social science.

David Lehrer  
Department of Political Science, University of Helsinki  
Visiting Researcher, WZB  
lehrer@wz-berlin.de  
www.jspurc.org

---

<sup>1</sup> The author wishes to thank the following persons for their critical comments on earlier versions of the paper and/or for input into the collaborative project that it reports: Natalia Ajenjo, Matthias Catón, Tom Cusack, Bernhard Kittel, Janine Leschke, Stefan Lhachimi, Will Lowe, Mikko Mattila, Jonathon Moses, Benoît Rihoux, Elmar Schlüter, Ana-Maria Vasiliu, Barbara Vis and Britta Weiffen. All errors and omissions remain the author’s own.

**PROLEGOMENA**

---

**MARKET FAILURES IN THE MARKETPLACE OF IDEAS?**

---

**Defining the Spurious**

**Why Is It a Problem?**

**Which Problem Is It?**

**FALSIFICATION, DEMARCATION, PUBLICATION**

---

**Method vs. Process**

**Processual Solutions**

**Are Data a Public Good?**

**TOWARD A MORE OR LESS SPURIOUS SOCIAL SCIENCE**

---

**Wholesale and Retail Ontologies**

**Clearing and Settling the Market**

**Conclusions, Prospects**

**REFERENCES**

---

Over 15 years ago, Bruce Charlton (1987) suggested in his article 'Think Negative' that many disciplines would benefit if negative results were given public airing. He argues that science needs reports of negative results for the simple reason that similar investigations, which are often costly and time-consuming, are frequently duplicated and produce the same negative result. Since then, other scientists have expressed similar sentiments (e.g., Knight 2003). Across the range of biological fields, table talk among ecologists and evolutionary biologists at all levels of academia at some point produces a sigh followed by 'If I had only known, I would have done things differently.' Recently, a number of the sciences have realised the gravity of the lack of published negative results, and have attempted to fill this publication void by producing journals that report negative results....

—**D.J. Kotze, C.A. Johnson, R.B. O'Hara, K. Vepsäläinen and M.S. Fowler, 2004, upon launching *The Journal of Negative Results in Ecology and Evolutionary Biology***<sup>2</sup>

The initiative to create a web journal of negative results is necessary, timely and presents some novel intellectual challenges....Negative results are not merely the inverse of positive results, and a journal of negative results is not simply the mirror image of conventional journals.

—**B.G. Charlton, 2004, upon realising his dream of a scientific journal of negative results**<sup>3</sup>

The importance of negative trials in the current research context is exceedingly high, as they provide highly useful information....Their importance can also be gauged by the fact that in 2002, finally, the *Journal of Negative Results in BioMedicine* materialized, something that Maxwell (1981) considered a 'joke' more than 20 years ago.

—**S. Malhotra, N. Shafiq and P. Pandhi, 2004, upon being asked by a student whether or not he should present a negative study to colleagues**<sup>4</sup>

Publication bias poses a potentially serious problem to those who endeavor to synthesize research findings in political science. Unfortunately, for any given literature it is difficult to know how many other studies were conducted but went unreported....If the discipline is to take seriously its large and growing research output, it must foster institutions such as registries that allow for meaningful synthesis of existing findings.

—**A.S. Gerber, D.P. Green and D. Nickerson, 2001, upon detecting publication bias in political science**<sup>5</sup>

Academia is a social enterprise that is usually most successful when individual researchers compete and collaborate in contributing toward common goals. In contrast, when we work in isolation on unrelated problems, ignoring work that has come before, we lose the benefits of evaluating each other's work, analyzing the same problem from different perspectives, improving measurement techniques and methods, and, most importantly, building on existing work rather than repeatedly reinventing the wheel.

—**G. King, 1995, upon describing the (related, but distinct) importance of replication of published studies to progress in the discipline of political science**<sup>6</sup>

Knowledge grows, inter alia, by acknowledging error. But political science hardly adheres to this rule.

—**G. Sartori, 1993, upon the demise of Sovietology (and of its object of study)**<sup>7</sup>

To err is not only human but also scientific.

—**G. Stigler, 1982, upon receiving the Nobel Prize in Economics**<sup>8</sup>

---

<sup>2</sup> Kotze, D.J., C.A. Johnson, R.B. O'Hara, K. Vepsäläinen and M.S. Fowler (2004) 'Editorial: The Journal of Negative Results in Ecology and Evolutionary Biology' *Journal of Negative Results in Ecology and Evolutionary Biology*, 1.

<sup>3</sup> Charlton, B.G. 'Why a Journal of Negative Results?' (2004) *Journal of Negative Results in Ecology and Evolutionary Biology*, 1, pp. 6-7.

<sup>4</sup> Malhotra, S., N. Shafiq and P. Pandhi (2004) 'Advantages of Reporting, Publishing, and Reading "Negative" Studies' *Medscape Pharmacists*, 5.

<sup>5</sup> Gerber, A.S., D.P. Green and D. Nickerson (2001) 'Testing for Publication Bias in Political Science' *Political Analysis*, 9, pp. 385-392.

<sup>6</sup> King, G. (1995) 'Replication, Replication' *PS: Political Science and Politics*, 28, pp. 443-499.

<sup>7</sup> Sartori, G. (1993) 'Totalitarianism, Model Mania and Learning from Error' *Journal of Theoretical Politics*, 5, pp. 5-22.

## PROLEGOMENA

---

How should scientific disciplines handle their mistakes?

The paper begins from the premise that the research process produces a plethora of results, findings, products and by-products, and argues that the totality of these results might be put to more effective use by the community of social science researchers than they are at present. The reasons for this neglect by political scientists may include:

- too much emphasis, across the discipline, on research *methods*, accompanied by too little emphasis on research *process*;
- too much emphasis on research individualism, and too little emphasis on building and maintaining a research community;
- and too much emphasis on innovation, coinciding with too little emphasis on extracting the full value of what has already been done.

The paper adduces examples from political science, from the other social sciences (particularly sociology and economics) and also from other research fields (including the natural and the clinical and applied sciences). The paper considers in passing whether the constellation of problems defined below as ‘spuriousness’ are homologous across diverse research communities and across the diverse approaches employed within the discipline of political science.<sup>9</sup>

The paper will consider three questions; one theoretic, one empirical, one normative:

- What are negative, unexpected, unwanted, or unpublishable results (and are all of these the same thing)?
- What do we do with them?
- What *ought* we to do with them?

---

<sup>8</sup> Stigler, G. (1982) ‘The Process and Progress of Economics’ *Nobel Lecture*, Stockholm.

<sup>9</sup> Lustick (1997); Gerring (2005).

Specifically, the first section will consider how best to define those results that are generated by the research process in political and other social sciences (particularly, but not only, sociology and economics) that are *not* disseminated or published. Are such results merely neglected, or negative, mistaken, unpublishable, spurious—and do these and related terms refer to the same or multiple types? Can and should results of this type be demarcated from morphologically similar results in the other social sciences and in other fields of research, including the ‘hard’ or experimental sciences?

The second section will detail various attempts to mitigate the ‘spurious’ as earlier defined that have been made in the social sciences and in other forms of research. It will also consider how the problem of ‘spuriousness’ is and is not addressed by philosophic theories of science such as falsificationism which continue to dominate the mainstream of social scientific and other scientific research.

The final section of the paper explores the question of what social scientists *ought* to do with ‘the spurious,’ and of what practice conventions, craft rules, criteria, tests, methods, processes, ontologic and epistemologic reorientations might be developed for analyzing and aggregating it, particularly within the context of the activity of individual researchers and research groups. This section also considers several concrete measures that could be taken at the disciplinary (research community) level to make alternative use of the range of products and by-products of the research process. Such shared disciplinary norms, institutions and practices remain notably undeveloped.

The paper relies throughout on the conceit of a marketplace of ideas as describing, and not only as metaphorically reflecting, essential aspects of the situation from which the problems it describes arise. The notion of the public sphere, the law or academia comprising sets of ideas in free competition may be traced back to the early decades of the last century, and continues to enjoy high standing today—perhaps particularly, but not only, in the context of the ‘Americanization’ of political science.<sup>10</sup>

---

<sup>10</sup> Moses, J., B. Rihoux and B. Kittel (2005) ‘Mapping Political Methodology: Reflections on a European Perspective’, *European Political Science*, 4, pp. 55-68.

## MARKET FAILURES IN THE MARKETPLACE OF IDEAS?

---

### Defining the Spurious

Despite widely accepted notions of the practices that ‘good’ social scientific research ‘should’ entail—i.e. the imperative to report one’s research findings—most ‘negative’ findings—those that do not fit with the researcher’s own theoretical presuppositions or with accepted tenets of methodological appropriateness—generally are not reported. Researchers do not submit these findings to journals, and journals do not seek to publish them. When they are published, it is often in the form of a footnote to an article focused on other, ‘positive’ results, or as ‘weak’ conclusions that are little cited. If such findings were consistently to be published, they might aid social scientists at large in refining theory or method, but at the same time they would likely undermine the individual careers of the few researchers who dared to publish them. Instead, ‘negative’ findings are buried in desk drawers, deleted from hard drives, or filed in circular files, and analyses are re-run or readjusted until more salutary results are produced.<sup>11</sup> Yet these rigorous but otherwise ‘unpublishable’ findings are, strictly speaking, necessary products of quantitative (and qualitative) social science research.

All methods contain an implicit normative dimension: application of a given method includes some results or statements as acceptable and excludes others, whether or not the criteria for doing so are highly codified or largely implicit within the practices of a particular research community. The implicit craft rules and conventions for applying methods (including qualitative as well as quantitative, theoretic, deductive and other ‘methods’) and for interpreting results are no less important to research practice than are the explicit, formal rules by which methods are defined. These normative dimensions of a given method (conventions such as the five percent test of significance, and craft rules governing which statistical test to apply to findings or how to assign relative analytic weight when multiple statistical tests are performed) may change over time.

Social norms within the discipline prescribing what ought to be done with negative or unexpected results may also change over time. The relationship borne by new results to prior ones serves as the

---

<sup>11</sup> Anonymous (2005) ‘One in Three Scientists Confesses to Having Sinned’ *Nature* pp. 718-719.

basis of criteria for their dissemination across diverse evidentiary cultures, but in at times contradictory ways: either too much or too little novelty can doom otherwise valuable research.<sup>12</sup>

A set of results may be deemed publishable because it has been anticipated by theory and/or because it falls within the range of previously accepted ‘positive’ results. Results might also be validated procedurally, e.g. if they have been arrived at *via* scrupulously observed explicit and implicit rules defining the method employed, and if the data or objects analyzed are deemed to have been reliably observed or appropriately considered. Yet ‘negative’ or unpublished results may also have been arrived at *via* impeccable procedures that follow the implicit and explicit rules that define the particular methods employed. Validation of such findings largely by reference to prior results biases reporting across the discipline, privileging results arrived at earlier in time and diminishing researchers’ capacity to reconsider those earlier results in light of more recent ones.

Social norms and normative rules might both determine the evaluation of results in practice. Treatises on method might themselves include ‘pure’ normative claims about what constitute ‘good’ and ‘bad’ results (based on procedural rules, accepted thresholds, or algorithms for determining results’ internal consistency). Theories, and prior cumulated findings, underlying particular research projects may also introduce ‘applied’ normative claims that qualify and disqualify various sets of results.<sup>13</sup> Unpublished or unpublishable results may be methodologically sound, yet still might be discarded because they do not accord with the expectations of *theory*, or they may confirm hypotheses and yet fail for one reason or another to pass the explicit or implicit tests of the *methods* by which they were arrived at.

Differing evidentiary cultures across institutional settings might also determine which results are considered ‘negative’: some findings may not be submitted to journals by researchers from one group that would have been submitted had they been reached by a researcher from another group. While distinctions among craft cultures in the social sciences may be blurring, some cross-country studies indicate ongoing segregation of publishing communities along national lines<sup>14</sup>—which leaves open the possibility that results are inconsistently filtered. Competition for acceptance by journals through the peer-review process increases pressure on researchers to submit articles

---

<sup>12</sup> Collins, H.M. (1998) ‘The Meaning of Data: Open and Closed Evidential Cultures in the Search for Gravitational Waves’ *American Journal of Sociology* 104, pp. 293-338.

<sup>13</sup> They may also fail to do so, as Duhem and Quine have noted.

<sup>14</sup> Schmitter, P.C. (2001) ‘Seven (Disputable) Theses Concerning the Future of ‘Transatlanticized’ or ‘Globalized’ Political Science’, < <http://www.iue.it/SPS/People/Faculty/CurrentProfessors/bioSchmitter.shtml>>.

introducing both new data and new theory. This means that the less novel result might be neglected by researchers and not developed into an article, even when it might be formally sound and substantively interesting: not highly *marketable*, but rather confirmatory or otherwise *useful* to some disciplinary subgroup.

There is thus the problem that the unpublished or unpublishable result may not be ‘negative’ at all, but merely insufficiently interesting or compelling to merit attempting the hurdles of publication. This, however, is not the whole story. Not only what is defined as negative, but also whether or not a particular result is negative and whether this might be a ‘good’ or a ‘bad’ thing is context-dependent, and may change over time. This is true in experimental sciences such as the hard sciences, and it is no less true in political science and in other branches of social sciences that are by and large not experimental sciences. The paper attempts to resolve these antinomies, and perhaps mutual irrelevancies, of the ways in which the constellation of concepts surrounding ‘negative,’ mistaken, unexpected, and unpublishable results has been used in scientific research, and particularly in the social sciences. In doing so it attempts to proceed briefly through the major steps of min-max definition for concept formation in the social sciences identified by Gerring and Barresi: ‘(1) sampling usages; (2) typologizing attributes; and (3) constructing minimal and ideal-type definitions.’<sup>15</sup>

Sample usages of the term ‘negative results’ by researchers in clinical, experimental and natural sciences underline several of the above points.

As Bjorn Olsen of Harvard Medical School noted in the inaugural editorial of the *Journal of Negative Results in BioMedicine*, launched in the fall of 2002, ‘Sometimes results may be negative only in the context of current thinking and experience. These so-called negative results have not gotten the publicity they deserve.’<sup>16</sup>

In an early article on the problem of unpublished results in psychology, Mahoney similarly valorizes what he terms negative results as being of greater usefulness to the research community than positive ones:

---

<sup>15</sup> Gerring, J. and P.A. Barresi (2003) ‘Putting Ordinary Language to Work: A Min-Max Strategy of Concept Formation in the Social Sciences’ *Journal of Theoretical Politics* 15, pp. 201-232.

<sup>16</sup> Olsen, B. and C. Pfeffer (2002) ‘Editorial’ in *Journal of Negative Results in Biomedicine* 1.

Given that the researched question is relevant and the experimental methodology adequate, the obtained results—whatever they might be—should be of interest to the scientific community. Assuming that they are clearly and comprehensively described, the data should not be viewed prejudicially on the basis of whether they conform to current theoretical predictions. In fact, given that the logic of science should be more properly falsificational rather than confirmational, negative (or contratheoretical) results yield much more information than positive results.....<sup>17</sup>

This assertion is echoed by Allchin, writing about marine biology, who asserts that ‘Ultimately, then, negative results can be positive knowledge.’<sup>18</sup>

In the social sciences, McKeown similarly asserts that negative results may or may not be *useful*, but that whether or not they will partake of this status might depend on the normative and the methodological commitments within which the researcher is working and on the specific expectations he or she attaches to the particular study in question at the time that it is undertaken: ‘From [a Bayesian] perspective, King Keohane and Verba’s contention that a negative result is as useful as a positive one is only true if one originally thought that both results are equally likely.’<sup>19</sup>

This connotational, if not denotational, problem of the difficulty of discerning when negative results might or might not be desirable is compounded by, *inter alia*, problems of extending the term ‘negative’ result from the experimental sciences to largely non-experimental ones; by the diversity of approaches within political science (including non-positivist and non-quantitative approaches, to which the concepts discussed above may apply in different ways or to different degrees); by the question of when it is better to be ‘negative’ than it is to be ‘positive’ (falsifying rather than confirming); and by the status of results that may not be ‘negative’ but that are nonetheless unpublished (and that may or may not be worthy of consideration in the same context as results that are definitively ‘negative’ by some agreed-upon standard). Thus there are multiple possible issues at stake with respect to the broad category of results that exist in some private sphere (e.g. the researcher’s file drawer) but not in a more or less public one (e.g. the index of a scientific journal). Specifically, the following are possible candidate definitions for the species of result that this paper aims to investigate the role and status of in social scientific research: something both distinct from

---

<sup>17</sup> Mahoney, M.J. (1977) ‘Publication Prejudices: An Experimental Study of Confirmatory Bias in the Peer Review System’ *Cognitive Therapy and Research*, 1, pp. 161-175.

<sup>18</sup> Allchin, D. (1999) ‘Negative Results as Positive Knowledge, and Zeroing in on Significant Problems’ in Browman, H.I. (Ed) (1999) ‘Theme Section: Negative Results’ *Marine Ecology Progress Series*, 191, pp. 303-305.

<sup>19</sup> McKeown, T.J. (1999) ‘Case Studies and the Statistical Worldview’ *International Organization* 53, pp. 161-190.

and broader than the concept of ‘negative’ results as it appears in the ‘hard,’ clinical or experimental sciences:

- i) findings that do not meet conventional significance tests, that fail to reject the ‘null hypothesis’
- ii) findings that do not meet researcher expectations, Bayesian or otherwise
- iii) findings that are ‘positive’ but not sufficiently interesting or *marketable* to be judged worthy of submission for publication, particularly because the hurdles (temporal, financial, emotional) for many kinds of publication may be high (i.e. other findings, or other activities, may be prioritized instead)
- iv) findings that fail to prove or disprove, confirm or falsify; ambiguous or unclear findings
- v) findings that result from pushing the limits of methodological convention or standards—possibly verging on technically weak or questionable execution
- vi) findings conflicting with—or ‘merely’ confirming—mainstream disciplinary consensus
- vii) findings merely confirming (or conflicting with) a particular replicated study

Because the above list describes a set of results broader than the ‘negative,’ and because of the excessive dynamism in the connotation of ‘negative’ depending on context (as ‘negative’ results might at times be ‘positive’ (or desirable) and might at other times be unwanted), the paper employs the term ‘spurious’ for the whole set of phenomena it will discuss...although other terms for this might also do.

A minimal definition to which this term may be applied, based on the above samples and typology, may thus be dual:

- I) from the perspective of individual researchers, the above types of result may describe forms of results that are *unwanted* (even if they might subsequently prove themselves to be desirable), or *unmarketable*;

- II) from the perspective of the research community, the above types of result may function as *bottlenecks* or *gates* (to borrow terms from operations research) obstructing the goal of disseminating all results arrived at by all researchers and of thereby creating a comprehensive and transparent public sphere.

Because results that are initially *unwanted* may in fact turn out to be desirable and useful, and because a bottleneck in dissemination is likely to be a dynamic and moving one (e.g. it might appear as various single items in the list above or as combinations of them in different contexts and at different times), we may reduce the two definitions of ‘the spurious’ above to one ideal-type:

*The set of constraints, at any given time, on knowledge cumulation by a community via the research process.*

This tentative, working definition is stipulated to refer to the set of constraints ‘endogenous’ to individual research studies (e.g. the researcher’s decision whether or not and how to report results) rather than to those constraints ‘exogenous’ to such studies and possibly not easily altered by the researcher(s) themselves (e.g. decisions on funding and resource allocations for specific research programs).

The above definition may have the advantage of being both succinct and able to account for a diverse range of cases beyond (but in many of their effects close cousins to) a single type of ‘negative’ result, dogmatically defined (as, e.g. in definitions i or ii above). It may have the disadvantage of proceeding from a naïve view of the nature, possibility and value of knowledge cumulation in the social sciences. Undue ambiguity may also be attached to the term *constraints* employed in this definition, since, strictly speaking, it is researchers’, editors’, reviewers’ or readers’ *behavior* in selecting for and against certain results that acts as a constraint on knowledge cumulation; it is not the *results* themselves that do so. Nonetheless it may be a worthwhile exercise to attempt to define this concept and to typologize the constellation of problems it poses for research: It is the obstacle to knowledge aggregation and to the creation of a comprehensive and transparent public sphere among disciplinary research communities that arises from research practices. The following sub-sections on Why Is It a Problem? and Which Problem Is It? seek to

address the question of how aspects of the spurious, as defined above, have been recognized by social scientists as a problem for their work. The concluding section of this paper will argue that this set of constraints is not given and immutable; it can (and should) purposefully be addressed *via* standards and practices, rather than being (as the middle section of the paper will argue it has been) effectively ignored.

### **Why Is It a Problem?**

Simply put, a tremendous amount of potentially useful information is currently lost to the social sciences through selective reporting of results.

While extensive study has been made of selection bias and of missing data on the input side of individual social scientific research studies and of strategies to counter these,<sup>20</sup> there has been only limited and inconclusive discussion of selection bias and missing data at the level of reported results. Yet the possibility exists that the level of information lost and the extent of conclusions distorted as a result of lacunae at the ‘macro’ level and on the ‘output’ side may be no less significant than those that are due to gaps in the ‘input’ side of individual studies. At least some of the information contained in or associated with ‘spurious’ results as defined above, results unlikely to be publicized, may potentially be useful—to other researchers, if not to the researcher who generated them. The more troubling possibility is thus that a great deal of information associated with these results is potentially useful, and is systematically disregarded.

Researchers in different institutions may repeatedly, and unknowingly, invest time and resources in pursuing research projects despite someone else having attempted something similar which has already led to results of demonstrably little utility.<sup>21</sup> Alternatively, the information associated with spurious results might be contained in some research institutions but not in others, or an informational stratification might emerge separating those who are sufficiently well-connected to learn about such failed initiatives *via* word of mouth and those who are not, and who therefore lack the means of avoiding such pitfalls.<sup>22</sup>

---

<sup>20</sup> See, for example Collier, D. and J. Mahoney (1996) ‘Insights and Pitfalls: Selection Bias in Qualitative Research’ *World Politics* 49, pp. 56-91.

<sup>21</sup> Prechelt, L. (1997) ‘Why We Need an Explicit Forum for Negative Results’ *Journal of Universal Computer Science*, 3.

<sup>22</sup> Charlton, B.G. ‘Why a Journal of Negative Results?’ (2004) *Journal of Negative Results in Ecology and Evolutionary Biology*, 1, pp. 6-7.

As a result of the reporting biases associated with spurious results, we inhabit a public space in which all of the researchers are successful, and all of the research is above average, but in which adequate meta-analysis is not possible; in which Bayesian priors and hermeneutic foreknowledge are unduly limited; and in which ‘stylized facts’ and theoretical certainties in various disciplinary subfields may well be based on the five percent of ‘positive’ results that have been reported, while the 95 percent of negative or conflicting results have not been.<sup>23</sup>

### **Which Problem Is It?**

The definition of the spurious advanced in the first section of the paper, built on a typology of multiple kinds of research results, is susceptible to the charge of reflecting a set of research problems and strands of literature in social science that is overly broad. Without responding directly to this charge, this section outlines several of the major debates on the implications and consequences of failing to deal adequately with the spurious. While there is no distinct body of literature on negative results in the social sciences, there are several strands of literature on the *effects* of the failure to deal adequately with such results in research.

There is also a burgeoning literature (and a developing infrastructure of journals and registries) on negative results in research fields outside the social sciences. The implications of this literature (in the natural, physical, clinical and applied sciences) for the social sciences will be discussed briefly in this section and in more concrete detail in the sections to follow.

### **Are social sciences sciences?**

Are the social sciences only poor cousins of the hard sciences? If not, then do conceptions of ‘negative’ results have any meaning for them? Specifically, need sciences be not only positive or inferential but also experimental to generate negative or spurious results worthy of the name? This paper argues that one need not make the argument that the social sciences are analogues of the hard sciences in order to assert that research communities in both worlds generate some forms of unwanted results; or to make the further argument that, even if what is negative or spurious may differ across fields, the problems that *the failure to process* such results jointly and appropriately lead to might tend to converge (i.e. such results in the natural and social sciences need not be homologous in their *essence* to be similar in their *effects*).

---

<sup>23</sup> Rosenthal, R. (1979) ‘The “File-Drawer Problem” and Tolerance for Null Results’ *Psychological Bulletin* 86, pp. 638-641.

### Spuriousness in diversity

Do negative results, spuriousness, and the problems associated with them even exist outside of positivistic, quantitative or scientific approaches? Is there a large class of approaches within the discipline of political science that does not have such problems...and that should therefore be valued more (or less, depending on one's preference) as a result of its divergence from the positivist model? One need not make the argument that quantitative and qualitative, or large- and small-N studies are analogues of each other<sup>24</sup> in order to consider the possibility that a wide range of approaches in political and other social sciences generate some set of results that display some of the characteristics in their epistemic effects that are defined here as spurious.

Examination of the problem of what is spurious across political science or the social sciences necessarily leads one into debates among 'quants' vs. 'quals,' 'American' vs. 'European' approaches, and 'inductivists' vs. 'deductivists,' and into various attempts to articulate 'third ways' (small-N comparisons,<sup>25</sup> mixed-method designs,<sup>26</sup> and analytic narratives<sup>27</sup>). The debate on the reconcilability of heterogeneous methods in the social sciences, which is now sporadic and fragmented in location across books, conferences and journals, has become increasingly salient and urgent. Perhaps it might be possible to reinvigorate dialogue among various approaches through the prism of 'spurious' results; their actual and potential effects; and what might be done to mitigate them. Perhaps, moreover, diverse traditions might find some common ground in their 'failure' that is missing in those results they consider to be their 'success.'

### Publication bias

As discussions of publication bias<sup>28</sup> and of the difficulties it poses for meta-analysis<sup>29</sup> have pointed out, it is possible that multiple, similar spurious results, independently arrived at, may collectively force reassessment of established orthodoxies in particular social science subfields. These orthodoxies might go unchallenged if multiple researchers arrive at conflicting results but do not communicate about them with each other and hesitate to submit them for publication because they

---

<sup>24</sup> King, G., R.O. Keohane and S. Verba (1994) *Designing Social Inquiry: Scientific Inference in Qualitative Research* (Princeton: Princeton University Press).

<sup>25</sup> Ragin, C.C. (2000) *Fuzzy-Set Social Science* (Chicago: University of Chicago Press).

<sup>26</sup> Tarrow, S. (1995) 'Bridging the Quantitative-Qualitative Divide in Political Science' *American Political Science Review* 89, pp. 471-474.

<sup>27</sup> *American Political Science Review* (2000) 'Symposium on Analytic Narratives and "Rational Choice History."'

<sup>28</sup> Gerber, A.S., D.P. Green and D. Nickerson (2001) 'Testing for Publication Bias in Political Science' *Political Analysis*, 9, pp. 385-392.

<sup>29</sup> Hedges, L.V. (1992) 'Modeling Publication Selection Effects in Meta-Analysis' *Statistical Science*, 7, pp. 246-255.

conflict with stylized facts or received wisdom in the field. While it is possible that informal networks of those working in particular subfields in different institutions might communicate some subset of this information in the absence of more formal and public dissemination, is this informal network combined with the opaque and uneven filtering process of the formal network sufficiently reliable across subfields and across national lines to ensure that an adequate portion of research ongoing and already conducted is contributing appropriately to knowledge?

### Observational boundaries

A spurious result might help to reinterpret or contextualize understanding of a phenomenon by providing an upper or lower bound of observation that researchers had previously been unaware of. Against the norms and normative rules for the testing, validation and dialogue of theory, methods, data and findings, the spurious, if widely disseminated, might serve as a critical benchmark: 'Positive' results can be interpreted, or reinterpreted, in terms of 'negative' ones. A 'negative' result might be rehabilitated to serve as the upper or lower limit of observations of a phenomenon.<sup>30</sup> One set of results (including 'theoretical' results or statements)<sup>31</sup> is typically validated in terms of its conformity with another; enhanced availability of a wider set of achieved results could therefore increase the scope for validating new findings through processes of *differentiation from* as well as of *conformity to*, thereby refining the joint exercise of social science researchers' judgment.

### Cumulation and Replication

Any concept defined in terms of its relationship to the cumulation of knowledge must necessarily enter debates on the possibility and value of such cumulation in the social sciences.<sup>32</sup> Related debates on replication in the social sciences are discussed in greater detail in the section Are Data a Public Good?

### Technical proficiency

The question of spurious results also touches on the profound (largely subtextual) debate in the political science literature on technical proficiency: the failure to understand or correctly to apply technically sophisticated methods. Certainly some spurious results (as defined above) are actually (or are feared to be by their originators) generated by unconventional application of sophisticated

---

<sup>30</sup> Collins, H.M. (2003) 'Lead Into Gold: The Science of Finding Nothing' *Studies in History and Philosophy of Science*, 34A, pp. 661-691.

<sup>31</sup> Kennefick, D. (2000) 'Star Crushing: Theoretical Practice and the Theoreticians' Regress' *Social Studies of Science*, 30, pp. 5-40.

<sup>32</sup> Kittel, B. (2005) 'The American Political Methodology Debate: Where is the Battlefield?' *Qualitative Methods. Newsletter of the APSA Section on Qualitative Methods*, pp. 12-19.

methods. It is not possible to know what portion of spurious results are of this type, or what proportion of results generated by other processes are feared to be of this type and are therefore not reported. It is the argument of this paper that even results actually generated by faulty application of particular methods may potentially contribute to the cumulation of social scientific knowledge and may therefore merit reporting.

### Choice of methods, choice of results

The opposite problem of too little technical proficiency is the problem of too much: mastery of excessively sophisticated and differentiated statistical and other methods may well make it possible for researchers routinely to match the methods they apply not to the ‘ontology’ of the object studied,<sup>33</sup> but rather to the desired results of particular studies. If this is the case, it then follows that ‘empirical’ research becomes more explicitly normative,<sup>34</sup> on the one hand, and that the scope for realizing ‘spurious’ results in the course of the research process may be reduced by ever-increasing choice, flexibility and agency in research design and interpretation.

## **FALSIFICATION, DEMARCATION, PUBLICATION**

---

### Method vs. Process

Many of the ways in which spuriousness is now dealt with in the social sciences remain (sporadically) at the level of method rather than (consistently) at the level of process. Methodologic approaches that have been developed to address the problem as it manifests itself *within individual studies* are in many cases interesting and useful, but as solutions to spuriousness *within the discipline at large* are necessarily self-limiting and piecemeal.

This paper seeks to call attention to the possibility that an overweening emphasis on research *method*, across the discipline of political science, in particular, has obscured critical view of the research *process*; and that, as a result, a great deal of time and ink have been devoted to highly sophisticated and specialized questions of research method, while pervasive, and basic, issues of

---

<sup>33</sup> Hall, P. (2003) ‘Aligning Ontology and Methodology in Comparative Research’ in J. Mahoney and D. Rueschemeyer, Eds. *Comparative Historical Analysis in the Social Sciences* (Cambridge: Cambridge University Press).

<sup>34</sup> Ajenjo, N. (2005) ‘Panel Outline’ *European Consortium for Political Research Conference*, Panel 20-5, ‘Looking at Methodological Issues Normatively.’

research process have been largely ignored or fragmented into piecemeal debates that have not generated practicable solutions. The discussion of methodology in the discipline has expanded tremendously in recent years, and numerous specialized fora for refining and debating methods have emerged. Yet there has been little parallel discussion of research process; to this author's knowledge, there is no Standing Group, working paper series, symposium or online journal on research process.

Regardless of which side one stands on in relation to positivist/non-positivist, quantitative/qualitative, or large-N/small-N divides, an emphasis on research individualism, rationality, design and intention has led to implicit norms that the researcher's primary and perhaps sole task is to choose from a menu of methods and to construct a procedure that will deliver a set of publishable results. Short shrift has been given by individual researchers and by institutions to process and to what happens to findings and results once they are attained. In such an environment, it is not surprising that high-grade, premium quality, 'one-best' results are favored; and that results, models, and explanations that do not meet this quality standard are ignored or discarded...even when they might contain information that is useful. In this environment, each individual researcher chooses from an ever-widening menu of methods but there are few shared standards or controls and only limited and uneven shared involvement in the filtering and processing of results.

Sharp lines of demarcation between acceptable and unacceptable findings or results correspond to attempts to impose similar lines of demarcation, across substantive fields and sub-fields, between science and non-science. This set of demarcations has been a feature of epistemologies imported from the natural and physical sciences to the social sciences *via* falsificationism.

This section advances two arguments: First, that the salient line of demarcation separating social science from non-social science is best conceived of not as that which runs between theories and hypotheses that can and that cannot be *falsified*, but rather as the line that runs between research and conclusions that can and that cannot be *published* in academic journals, books and conference proceedings, regardless of whether or not they meet strict criteria of falsification. Second, it argues that demarcation lines have a tendency to disrupt the fields they demarcate (e.g. falsificationism necessarily and drastically reduced the set of potential research outcomes that are of interest), and that demarcation may more fruitfully be done *via* continua between two endpoints rather than *via* a single line: i.e. by shading from full membership in the set 'social science' to full membership in

the set ‘not social science,’ accompanied by researchers’ capacity to recognize all of the points and states in between.<sup>35</sup>

The Popperian stress on demarcation of science from non-science, developed primarily in the context of theorizing about the physical sciences and imported into the social sciences, is based on a logic of exclusion. This logic of exclusion has had the effect, in the social sciences, of narrowing the field of what is acceptable to report in a scientific context to the ‘one best’ finding.<sup>36</sup> In the resulting schema, research generates a relatively small and delimited signal against a vast field of noise: a premium saleable product and a great deal of waste. The paper argues in its ‘normative’ section that consideration should be given to replacing the notion of *lines of demarcation* which currently structures the research process and the interpretation of results with the alternative notion of *shades of usefulness* of results and findings. At present, the locus of filtering associated with the demarcation ontology of research rests largely with the individual researcher (although research institutions, publishers and peers also play roles, both direct and indirect), with some institutional but few disciplinary standards or checks to refer to. Under such a filtering regime, Caporaso asserts, ‘We rarely report results in incremental (value added) fashion, as additions to the existing capital stock. Instead, our results are presented as separate “findings.”’<sup>37</sup>

The methods-based attempts to make spuriousness more tractable, *within the context of individual studies and research projects*, has resulted in a number of overlays and addenda to the ‘one-best’ specification, theory, method, model and result. Interesting approaches, some now conventional, some little-used, have been developed to supplement the process of cumulation of knowledge *via* confrontation of ‘one best’ results with each other, with cumulation *via* incremental layering and aggregation of information of varying degrees of quality. These include currently standard procedures, such as various means of extracting information from the error term that accompanies ‘one-best’ results; interpreting confidence intervals; performing sensitivity analysis, and (more commonly in the clinical sciences and too rarely in the social sciences) undertaking multiple distinct tests of a study’s hypothesis. To some extent Bayesian approaches also enable extensions of the logic of inference to extract information from negative or unexpected results. Bayesian confirmation theory enables researchers to extract information from both positive and negative

---

<sup>35</sup> Ragin, C.C. (2000) *Fuzzy-Set Social Science* (Chicago: University of Chicago Press).

<sup>36</sup> Hjort, N.L. and Claeskens, G. ‘Frequentist Model Average Estimators’ *Journal of the American Statistical Association*, 98, pp. 879-899.

<sup>37</sup> Caporaso, J.A. (1995) ‘Research Design, Falsification, and the Qualitative-Quantitative Divide’ *American Political Science Review* 89, pp. 457-460.

*evidence, via* probabilities and expectations: Observations of a black raven and a non-black non-raven can both, to varying degrees, be interpreted as confirming a theory.

In addition to the above conventional methods for layering additional gradations of information around ‘one-best’ results, other forms of thick, fuzzy and probabilistic interpretations and layering have been developed.<sup>38</sup> One example is the recent suggestion of the ‘thick modeling’ approach in empirical economics, in which multiple specifications of a model are considered and used to generate output, which are then layered together and aggregated *via* a small-scale meta-analysis.<sup>39</sup> This approach, which has been applied to portfolio selection and which builds on earlier experience in economics of pooling forecasts, makes it possible for researchers to extract information from and report multiple specifications, findings and results simultaneously. Granger and Jeon argue that this approach is more appropriate to the study of some economic phenomena than it is to that of others. It should be noted, however, that while such procedures potentially expand the scope of results of research studies admitted to public discourse, they do not preclude the possibility of some set of results (even when these techniques are used) being systematically filtered, selected out and discarded. Other examples include algorithms for model averaging recently proposed by statisticians.<sup>40 41</sup> Such approaches succeed, to some extent, in dealing with spuriousness by pushing the problem back to the conceptual, theoretical or modeling level, and by vitiating the aura of precision that surrounds assumptions, concepts, models and findings under other approaches.

Other ways in which the spurious is handled, in the social sciences as in other sciences, include simply to ignore or discard unwanted or unmarketable results, to re-run analyses until appropriate ‘one-best’ results are achieved, or to ‘mis-present’ negative results as positive ones.<sup>42</sup> Finally, as noted above, negative results do in many cases see the light of day (although the proportion of attained negative results that do so is unknowable), appearing as ‘watered-down’ or ‘weak’ positive results; as footnotes to articles primarily disseminating other results; as shorter articles submitted to less prestigious journals; as arguments against the explanation that was being tested for or in favor of an alternative explanation; as folklore that is passed orally across disciplinary networks, etc.

---

<sup>38</sup> Nurmi, H., J. Kacprzyk and M. Fedrizzi (1996) ‘Probabilistic, Fatty and Rough Concepts in Social Choice’ *European Journal of Operational Research* 95, pp. 264-277.

<sup>39</sup> Granger, C.W.J. and Y. Jeon (2004) ‘Thick Modeling’ *Economic Modelling* 21, pp. 323-343.

<sup>40</sup> Hoeting, J.A., Madigan, D., Raftery, A.E. and Volinsky, C.T. (1999) ‘Bayesian Model Averaging: A Tutorial’ *Statistical Science*, 14, pp. 382-417.

<sup>41</sup> Claeskens, G. and Hjort, N.L. (2003) ‘The Focused Information Criterion’ *Journal of the American Statistical Association*, 98, pp. 900-916.

<sup>42</sup> Prechelt, L. (1997) ‘Why We Need an Explicit Forum for Negative Results’ *Journal of Universal Computer Science*, 3.

Finally, individual researchers have addressed the constellation of problems surrounding spuriousness by developing algorithms and techniques for testing the extent of its impact on the discipline. These techniques are typically employed in the context of meta-analysis of the literature in particular subfields. The utility of these techniques is debated, even by their own developers, as the quantity they seek to estimate—the number of results that remain unreported and that are selected out of the public sphere by those who generated them (and by editors, reviewers, peers, etc.) is simply unknown. Algorithms for estimating the size of the ‘file-drawer’ problem include, *inter alia*, the fail-safe file drawer analysis, which Scargle has argued is irrelevant and distortive of the size and character of the omitted studies.<sup>43</sup> Articles detailing attempts to refine algorithms to test for the file-drawer problem end, not infrequently, in pleas by the authors to develop shared, process-oriented solutions to the underlying problem at the disciplinary level.<sup>44</sup>

The above discussion does not assert that *no* filtering of results *ought* to take place during the research process, but rather seeks to open the question of what *criteria of filtering* are in fact being applied, particularly when filtering occurs at the ‘tail-end’ of the research process, in private and performed by individuals, under a research ontology based on demarcation and a set of norms and professional incentives to deliver, with few exceptions, premium-grade product and ‘one-best’ findings to the marketplace of ideas.

In contradistinction to the subsumption of the disciplinary problem of spuriousness under discussion of methods, and to the advancement of individual and research-methods-based approaches to dealing with collective and research-process problems, the following sub-section, Processual Solutions, reviews institutional, disciplinary, collaborative, and process-oriented attempts to address the problem of spuriousness in research. These initiatives are reviewed in a broadly comparative context.

### **Processual Solutions**

Other fields are now developing procedures to process the types of results defined herein as spurious, particularly the hard sciences. A number of these initiatives are quite recent, and the rate

---

<sup>43</sup> Scargle, J.D. (2000) ‘Publication Bias: The “File-Drawer” Problem in Scientific Inference’ *Journal of Scientific Exploration* 14, pp. 91-106.

<sup>44</sup> Gerber, A.S., D.P. Green and D. Nickerson (2001) ‘Testing for Publication Bias in Political Science’ *Political Analysis*, 9, pp. 385-392.

at which they proliferate appears to be increasing. If we accept the notion that epistemic trends in the social sciences follow (with some lag) those in the natural sciences, then it may only be a matter of time before similar initiatives, in some format or other, appear in political science and its sister disciplines.

### Natural and physical sciences

In the natural and physical sciences, the problem of rescuing and airing spurious results is now coming into its own. New journals are answering the calls for such fora in both the ‘hard’ and ‘life’ sciences that have been sounded in such prominent scientific journals as *Nature* and *New Scientist*.<sup>45</sup> Separate, specialized journals have arisen in specific fields (mirroring the substantive scope and distribution of conventional scientific journals of positive results), particularly in fields related to biology. In 1999, the *Marine Ecology Progress Series* published a special ‘theme section’ in its December issue on ‘Negative Results.’ The *Journal of Negative Results in Ecology and Evolutionary Biology* was launched online in 2004.

Other initiatives that have arisen in the natural sciences to control the problems associated with spurious results (selection bias, publicationism, missing information at the disciplinary level) include attempts to control bias in reporting *via* registries, and statistical adjustments to meta-analyses of the literature that seek to account for unreported studies and results.

### Clinical and applied sciences

In the clinical and applied sciences, such as medicine, pharmacology, clinical psychology, cognitive and computer science and software engineering,<sup>46</sup> journals and special sections of journals publishing negative results have begun to proliferate. In the clinical sciences (i.e. health and medical research), registries for publishing abstracts prior to undertaking studies, and for submitting results following their completion, are more highly developed than in any other discipline and serve the purpose that specialized journals for reporting such results do in other disciplines. The Cochrane Collaboration was founded in 1993 to systematically synthesize clinical data from around the world. In addition to registries, extensive activity and supporting institutions (including disciplinary norms) for reporting ‘negative’ results can be found in the clinical sciences. As

---

<sup>45</sup> Kotze, D.J., C.A. Johnson, R.B. O’Hara, K. Vepsäläinen and M.S. Fowler (2004) ‘Editorial: The Journal of Negative Results in Ecology and Evolutionary Biology’ *Journal of Negative Results in Ecology and Evolutionary Biology*, 1, pp. 1-5.

<sup>46</sup> Prechelt, L. (1997) ‘Why We Need an Explicit Forum for Negative Results’ *Journal of Universal Computer Science*, 3.

mentioned above, the *Journal of Negative Results in BioMedicine* was launched by BioMed Central in the fall of 2002. The *Journal of Negative Results in Speech and Audio Sciences* was launched at Carnegie-Mellon University in the winter of 2003.

In psychology (in large part an experimental and treatment-oriented discipline drawing on the full institutional apparatus of laboratories and maintaining close links to biology, chemistry, and other natural and physical sciences), the *Journal of Articles in Support of the Null Hypothesis* appeared online in the summer of 2002.

In the applied sciences and engineering, the ‘Forum for Negative Results (FNR)’ was inaugurated as a permanent special section of the *Journal for Universal Computer Science (J.UCS)*, an online journal published by Springer-Verlag, in 1997.

#### Social, policy and management sciences

As discussed below in the sub-section Are Data a Public Good? below, a repository for *data* (but not for negative results) of published positive articles in empirical economics was created by the *Journal of Money, Credit and Banking* in the 1980s but was later abandoned. Attempts have been mounted to aggregate political science data (again for conventional published studies of ‘positive’ results) by ICPSR, APSA and other professional associations and journals, with mixed results.

Along with the collection and dissemination of *data* (rather than of findings and results), attention in the social sciences to issues related to spuriousness has focused largely on the publication of attempts to replicate previously published, positive studies. *Political Analysis*, the journal of the Society for Political Methodology of the American Political Science Association, regularly includes a section on ‘Replications and Extensions’ of previously published work. *Empirical Economics*, the journal of the Institute for Advanced Study in Vienna, like the *Journal of Applied Econometrics*, includes a section containing brief ‘replication studies of important results in the literature’—with both positive and negative results. The *Quarterly Journal of Business and Economics* states on its website that

There is a neglected body of research involving tests of the validity of theories, methodologies, and data used in previously published studies. There is also a need for more articles that synthesize the literature on important topics, particularly those topics that

bridge several disciplines. The editorial philosophy of this journal will be to give priority to replicatory articles.<sup>47</sup>

Replications of prior studies serve an important purpose in advancing social scientific theory and methods that bears some similarities to the purpose that might be served by publishing original negative results, but one that remains distinct. While the publication of replications has gained some acceptance in the social sciences in recent years following attention to the issue by methodologists,<sup>48</sup> it is not widespread, and it does not include the wider range of spurious results defined above.

Registries (e.g. the Campbell Collaboration, launched in the winter of 2000) have been developed in some subfields, particularly in applied branches of social science (criminology, education, social welfare), to enable the conduct and dissemination of systematic reviews of published literature. Initiatives such as the Campbell Collaboration do not, however, ensure that studies introducing negative or spurious results will be included in such reviews (i.e. it does not obviate the file-drawer problem in the subfields it addresses).

Algorithms continue to be proposed to test and adjust for publication bias in meta-analysis (which remains less practiced in the social than in the clinical sciences). Some scholars have suggested that concrete, cooperative measures be undertaken to publish various forms of spurious result:

It would be helpful if researchers automatically published a synopsis of their findings regardless of the outcome. More realistically, professional associations within political science could create a central registry of abstracts for proposed studies, akin to what exists in the medical sciences....<sup>49</sup>

Yet to the author's knowledge, this has yet to be done, and such suggestions have not been accompanied by detailed proposals. In the first half of the present decade, approaches to the problem of spuriousness in the social sciences have ranged from fragmented to conjectural to non-existent.

---

<sup>47</sup> *Quarterly Journal of Business and Economics*, <<http://www.qjbe.unl.edu/editorial.html>>, accessed 18 September 2005.

<sup>48</sup> King, G. (1995) 'Replication, Replication' *PS: Political Science and Politics*, 28, pp. 443-499.

<sup>49</sup> Gerber, A.S., D.P. Green and D. Nickerson (2001) 'Testing for Publication Bias in Political Science' *Political Analysis*, 9, pp. 385-392.

In the summer of 2005, *The Journal of Spurious Correlations* was founded, in cooperation with the Standing Group Political Methodology of the European Consortium for Political Research, to provide a forum for original, neglected results in the social sciences.

### **Are Data a Public Good?**

In the late 1980s, the *Journal of Money, Credit and Banking* undertook a project to collect the data analyzed in the studies it published. Responses to requests for data were mixed. A decade later, those datasets that were successfully collected had not led to extensive replication of the original studies by other researchers. Similarly, an initiative funded by the National Science Foundation in the US to encourage other social science journals and dedicated institutional archives to act as repositories for data was not received with great enthusiasm by journal editors or by researchers.<sup>50</sup>

The problem that the *JMCB* initiative addressed is that of cumulation of knowledge within the subset of *published* social science studies. Making the data associated with published studies available, it was reasoned, would enable replication and extension of previously accepted ‘positive’ results. This issue of enabling replication of positive results *via* data availability has yet to be resolved in the discipline. The additional, and distinct problem of ‘data selection bias’—i.e. data sets associated with unpublished *findings* may themselves remain unpublished—is one form of spuriousness that may appear as a secondary effect of the biased reporting of results (i.e. results and data often go together, and sometimes suffer the same fate).

Data is an intermediate good that many researchers process on their way to results, findings, stylized facts, explanations, refinements to theory etc. Data sets from which spurious results ultimately emerge might in turn become ‘spurious’ themselves, if neither the data nor a summary of the characteristics that led to the results in question are publicized. Data sets associated with spurious results might not be published or disseminated, or their dissemination might be considerably delayed, for reasons similar to those discussed above: the most ‘saleable’ articles may be those that introduce new data, new theory and new findings concurrently (i.e. datasets not previously disseminated might be withheld from the market until sufficiently positive results can be coaxed from them).

---

<sup>50</sup> Anderson, R.G. and Dewald, W.G. (1994) ‘Replication and Scientific Standards in Applied Economics a Decade After the *Journal of Money, Credit and Banking* Project (St. Louis: Federal Reserve Bank of St. Louis).

If new data sets are not published because analyses performed by the researchers who compiled them fail to yield publishable results—or, as above, because researchers may view with trepidation the possibility of replication—then all or part of those data, along with the analysis of the data, might never see the light of day. In such cases one researcher’s rubbish may indeed be another’s dinner, if previously unpublished data sets were to be employed by other researchers in novel ways. This problem of ‘data selection bias’ describes the state of affairs in which only data sets accompanied by ‘successful’ analyses are likely to be published and disseminated, even if data that has been compiled and used ‘unsuccessfully’ by one researcher might be analyzed ‘successfully’ by another.

## TOWARD A MORE OR LESS SPURIOUS SOCIAL SCIENCE

---

### **Wholesale and Retail Ontologies**

As noted under the sub-section Methods vs. Process above, a great deal of methodological practice and debate in the social sciences implicitly addresses the demarcation problem. The present sub-section advances recommendations for mitigating spuriousness (the effects of which are greatly exacerbated by conventional demarcation practices) at the individual, researcher, level, primarily by reorienting our ontology of research: specifically by advocating the emancipation of research from its exclusive focus on premium-grade information, one-best outcome sets and high-end methods at the expense of transparency, consistency and comprehensiveness in reporting.

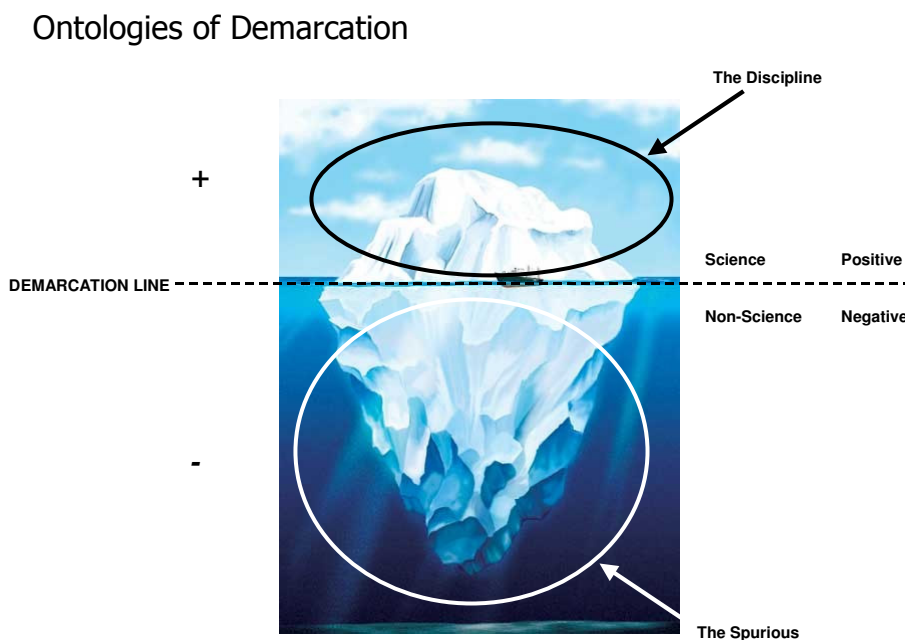
How, then, should scientific disciplines handle their mistakes?

At present the time-honored method of flaying, or of maintaining the standing threat to flay, the individual researcher who publicly disseminates a result that might be susceptible to being labeled a mistake (under any number of possible criteria), remains the default procedure.

Is there a prescription for handling ‘spurious’ results that all researchers ought to follow, or is there a set of prescriptions varying by disciplinary subfield or research method employed? Or is this an area not properly within the realm of disciplinary debate and standard-setting? Is the current set of practices arguably what we should be doing?

An ontology of research based on a demarcation line and a logic of exclusion leads us to value only those results and that information that is manifested ‘above’ the line. The demarcation line in the ‘iceberg of knowledge’ bequeathed to the social sciences by falsificationism (Figure 1, below) separates ‘science’ from ‘non-science,’ and also ‘positive’ or publishable and desirable from ‘negative’ or unpublishable and unmarketable results (defined above as spurious). Results must exceed the most stringent tests in order to remain above the demarcation line and within the realm of legitimate scientific discourse; any hole poked in a finding or statement may be sufficient to submerge the entire finding and all those other findings and ancillary statements associated with it beneath the demarcation line.<sup>51</sup> If (as argued above) there may in fact be useful information (albeit information that is not of premium grade or of the highest quality) contained in results that fall ‘below’ the line—results that will never come into view in a public discourse that treats the line as a lower limit—then all of that information will be lost.

**Figure 1. Falsification, Demarcation, Publication**



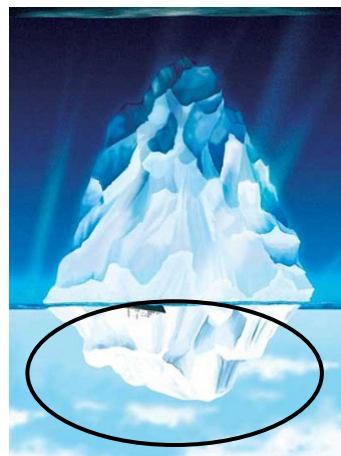
The ‘iceberg of knowledge’ in Figure 1 may also be viewed in another way, one that is considerably more troubling, i.e. as an ‘upside-down’ iceberg (Figure 2). When lines of demarcation lead to

<sup>51</sup> Soberg, M. ‘The Duhem-Quine Thesis and Experimental Economics: A Reinterpretation’ *Discussion Paper* 329 (Norway: Statistics Norway Research Department).

some results being published and others being submerged, it is possible that an entire subfield of inquiry might have based its conclusions on a set of published results that ‘ought’ to have fallen on the opposite side of the line. In airing only results that tend to confirm those that have gone before (one of the problems, in some research communities and subfields, that leads to spuriousness), we increase the risk that an entire field can have it wrong: i.e. that the five percent of positive studies confirming a theory have been published, while the 95 percent of studies testing and failing to confirm the same explanation are languishing in file drawers.<sup>52</sup> If spuriousness is defined, as in the first section of this paper, as the constraint at any given time on knowledge cumulation in a research community, then published results, inappropriately selected *for* due to history, habit, ideology or some other reason, may be no less spurious than those results that have remained unpublished because they conflict with those that have. In other words, it may be possible to be ‘negative,’ or more broadly, spurious, on both sides of the demarcation line at once under a research ontology of ‘hard’ demarcation and in an environment of the opacity of a significant set of results.

## Figure 2. Why Is It a Problem?

What Is Spurious?



If all participants in the research community aim to disseminate only premium-grade, ‘above the line’ results, and if some potentially useful information is contained in what might be called the

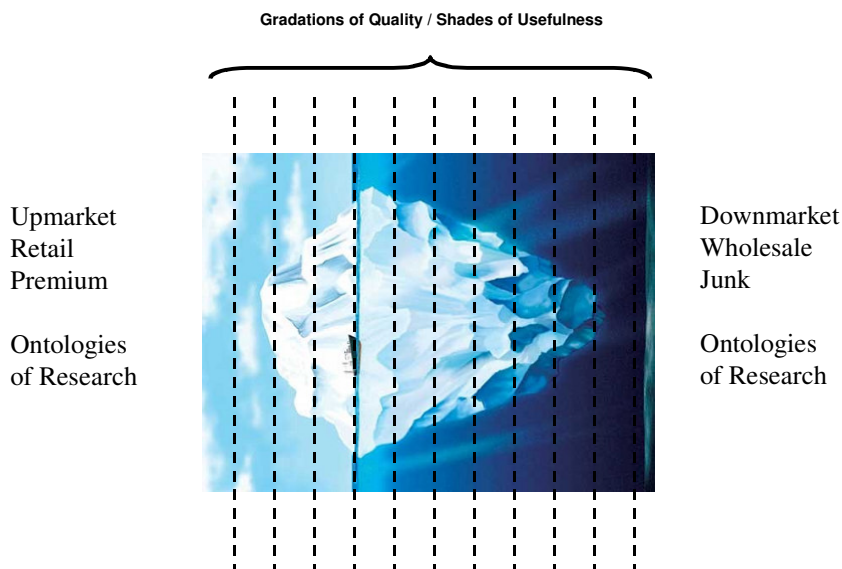
---

<sup>52</sup> Rosenthal, R. (1979) ‘The “File-Drawer Problem” and Tolerance for Null Results’ *Psychological Bulletin* 86, pp. 638-641.

mid-range of quality (just below the surface of the demarcation line), then there is obviously some value to be gained by moving the line of demarcation down. If, as may well be the case, there is useful information to be found even in the nether depths of informational quality toward the bottom of the knowledge iceberg, then moving the line of demarcation down to bring that information into public view will vitiate the demarcation project entirely.

**Figure 3. Wholesale and Retail Ontologies of Research**

### Undemarcating the Iceberg of Knowledge



The paper proposes instead that research reorient itself, by reporting on results that fall below the most stringent line demarcating informational quality. In turning the iceberg on its side (Figure 3), perhaps we can speak of a range of gradations of quality of the information to be extracted, ranging from ‘upmarket’ results at the very high end of the market, to wholesale or even junk results, findings, information and knowledge at the other.

In its broad form this proposition is nearly as old as the institution of the demarcation line itself. As Hands notes, Carl Hempel called in 1965 for elision of the line of demarcation—conceived of as the line separating ‘theory’ from ‘observation’ in the epistemology of the natural and physical sciences:

Furthermore, there remains no satisfactory general way of dividing all conceivable systems of theoretical terms into two classes: those that are scientifically significant and those that are not; those that have experiential import and those that lack it. Rather, experiential, or

operational, significance appears as capable of gradations....experiential significance presents itself as capable of degrees, and any attempt to set up a dichotomy allowing only experientially meaningful and experientially meaningless concept systems appears as too crude to be adequate for a logical analysis of scientific concepts and theories.<sup>53</sup>

Is it necessary for research communities to produce and consume only premium-grade knowledge all the time, or is there value to be gained by going downmarket now and then?

### **Clearing and Settling the Market**

The present sub-section advances recommendations, based on the comparative review of the state of the art in other research disciplines under Processual Solutions (above), for addressing spuriousness collaboratively within the discipline.

#### **Central Registry**

**An online registry**, in addition to publishing synopses of ‘unreported’ studies and ‘spurious’ results, might provide a venue for introducing datasets that were not previously available because the researchers who developed them have not yet published articles analyzing those data, as well as the data sets of published studies for further use and replication by others.

The challenge that runs across the spectrum of initiatives from an online registry to a print journal is the trade-off between comprehensiveness and honesty in reporting, on the one hand, and information overload, infeasibility and irrelevance on the other. It is the same trade-off that individual researchers, editors, reviewers and readers now face in drawing the present ‘demarcation line’ of publication and the public sphere of scholarly discourse. Reliably filtering, exerting quality control over and disseminating useful and interesting information from the entire mass of ‘junk’ currently to be found in the desks, hard drives and wastebaskets of social scientists would be no mean task. We may indeed be thankful that convention requires individual researchers to take responsibility for this filtering themselves.

At its best, journals and even online registries might create a ‘second-tier’ of informational quality and availability, while still leaving the large mass of findings in the file-drawer precisely where they

---

<sup>53</sup> Hempel, C.G. (1965) *Aspects of Scientific Explanation* (New York: The Free Press), cited in Hands, D.W. (2001) *Reflection Without Rules: Economic Methodology and Contemporary Science Theory* (Cambridge: Cambridge University Press), p. 83.

now sit. Even were this the case, however, such an exercise might still move the discipline closer toward insight into the scope, magnitude, shape and effects of spuriousness than it is at present.

#### Specialized journals and journal sections

A **print journal**, or a series of journals segmented by disciplinary subfield, or regular sections introducing spurious results appended to existing journals, might embrace all manner of results on the frontier of social scientific research that might not otherwise have found a place within existing disciplinary journals in their current forms. Such results ought not to be limited to a particular set of subfields, methods or approaches, and ought to be peer-reviewed and vetted for research quality and scientific and heuristic value. The value of such a project might lie in its publication of *original* (rather than replicated) spurious results. As such it might also enable editors, contributors and readers avoid the anxiety associated with witch-hunting and revision of previously published studies. Fora in which to expose others' 'mistakes,' real or imagined, already exist in every discipline. Yet there remain few such venues in which to expose our own.

#### Routine individual online notification

Another approach, which might work only if widely adopted, might be simply to **decentralize the registry**. This describes a world in which researchers routinely post all of their results on their websites (or in a similar form), perhaps accompanied by a label rating those results' 'informational quality,' and perhaps in formats conducive to automated meta-analysis or other forms of aggregation or review.

There may indeed be no 'one-best' solution to the problem of spuriousness. There may, however, be paths and practices that would be better than those currently traveled and employed. The discussion of such alternative paths and practices might best be conducted across social scientific disciplines, and indeed across research communities in the other sciences as well.

#### Conclusions, Prospects

The paper advanced a definition of 'the spurious' as the set of constraints on joint knowledge cumulation arising from the research process at a particular moment in time.

The paper also considered the present state of technology in the discipline of political science for processing the discipline's own waste: the discarded by-products of its collective research process.

The paper considered analogous practices of greater and lesser degrees of advancement in other disciplines and research communities. While economics has developed well-articulated theories of production and consumption, it has little to say about waste. What is true of economics is no less true of sociology, political science, and history. Some have argued that civilizations can be defined by the ways in which they process their own waste.<sup>54</sup> Scientific disciplines, and research communities, may be no different.

The paper has attempted to indicate some of the ways in which the problems of spuriousness may impact research and its use across diverse research communities within and beyond political science. It is likely, though it perhaps remains to be demonstrated both empirically and theoretically, that one need not be positive to be spurious—i.e. that the concept of negative, unexpected, unpublishable or unwanted results applies not only to ‘quantitative’ or ‘positive’ approaches in political science but also to other approaches, traditions, and styles of research, just as it need not apply only to the experimental sciences. Rather, the spurious may be a partially hidden feature of *the application of* all approaches within the discipline of political science, be they positivist, postpositivist, formal, informal, quantitative, discursive, analytic, hermeneutic, critical, applied, inductive, deductive, ideographic, nomothetic, fuzzy, emancipatory, historical or otherwise. While early attention to the problem of *negative* results in specific originated in the natural, ‘hard,’ experimental sciences, attention to the broader problem of what is spurious or unpublishable in the social sciences need not be taken as a step on the path toward refashioning the social sciences in the image of the ‘harder’ sciences, nor as mere imitation. The discussion of spurious results is nonetheless implicitly a discussion of the boundaries between and among ontologies and epistemologies, hard sciences and soft, as well as an opportunity for dialogue among the various methodologies and methods, approaches and paradigms employed by those who investigate social and political life.

It is hoped that, at a minimum, the paper has called attention to the need for a wider discussion of what constitutes an analytic or research ‘mistake’ and of how these ‘mistakes’ ought to be handled within the discipline. It has considered several possibilities for enhancing researchers’ opportunities to learn from our own ‘mistakes’ by creating shared venues in which to subject individual mistakes (which in a wider context might prove not to be ‘mistakes’ at all) to the critical scrutiny and scholarly discussion of our peers.

---

<sup>54</sup> Fagan, H. (2002) ‘Grounding Waste: Towards a Sociology of Waste Networks’ *NIRSA Working Paper Series*, 18, National University of Ireland.

At present, there is no single place where the best of the discarded data and analysis, the most interesting and relevant of the flotsam and detritus of social scientific inquiry, is gathered and subjected to critical review.

Who knows? In aggregate, it might well have something useful to tell us.

## REFERENCES

---

- Ajenjo, N. (2005) 'Panel Outline' *European Consortium for Political Research Conference*, Panel 20-5, 'Looking at Methodological Issues Normatively.'
- Allchin, D. (1999) 'Negative Results as Positive Knowledge, and Zeroing in on Significant Problems' in Browman, H.I. (Ed) (1999) 'Theme Section: Negative Results' *Marine Ecology Progress Series*, 191, pp. 303-305.
- Allchin, D. (2000) 'The Epistemology of Error' Paper presented at Philosophy of Science Association Meetings, Vancouver, November 2000.
- Allchin, D. (2001) 'Error Types' *Perspectives on Science*, 9, pp. 38-59.
- Almond, G.A. (ed.) (1990) 'A Discipline Divided: Schools and Sects in Political Science (USA: Sage).
- American Political Science Review* (2000) 'Symposium on Analytic Narratives and "Rational Choice History."
- Anderson, R.G. and Dewald, W.G. (1994) 'Replication and Scientific Standards in Applied Economics a Decade After the *Journal of Money, Credit and Banking Project* (St. Louis: Federal Reserve Bank of St. Louis).
- Anonymous (2005) 'One in Three Scientists Confesses to Having Sinned' *Nature* pp. 718-719.
- Anderson, D.A., W.A. Link, D.H. Johnson and K.P. Burnham (2001) 'Suggestions for Presenting the Results of Data Analyses' *Journal of Wildlife Management*, 65, pp. 373-378.
- Bhaskar, R. (1998) *The Possibility of Naturalism: A Philosophical Critique of the Contemporary Human Sciences (3rd edition)* (London: Routledge).
- Bonilla, J.P.Z. (2002) 'Scientific Inference and the Pursuit of Fame: A Contractarian Approach' *Philosophy of Science*, 69, pp. 300-323.
- Browman, H.I. (Ed) (1999) 'Theme Section: Negative Results' *Marine Ecology Progress Series*, 191, pp. 301-309.
- Campbell Collaboration*, <<http://www.campbellcollaboration.org>>.
- Caporaso, J.A. (1995) 'Research Design, Falsification, and the Qualitative-Quantitative Divide' *American Political Science Review* 89, pp. 457-460.
- Charlton, B. (1987) 'Think Negative' *New Scientist*, 116, p. 72.
- Charlton, B.G. 'Why a Journal of Negative Results?' (2004) *Journal of Negative Results in Ecology and Evolutionary Biology*, 1, pp. 6-7.
- Chatfield, C. (1995) 'Model Uncertainty, Data Mining and Statistical Inference (With Discussion)' *Journal of the Royal Statistical Society Series A*, 158, pp. 419-466.
- Claeskens, G. and Hjort, N.L. (2003) 'The Focused Information Criterion' *Journal of the American Statistical Association*, 98, pp. 900-916.
- Cleophas, R.C. and T.J. Cleophas (1999) 'Is Selective Reporting of Clinical Research Unethical as Well as Unscientific?' *International Journal of Clinical Pharmacological Therapy*, 37, pp. 1-7.
- Clogg, C.C. and G. Armingier (1993) 'On Strategy for Methodological Analysis' *Sociological Methodology* 23, pp. 57-74.
- Cochrane Collaboration*, <<http://www.cochrane.org>>
- Collier, D. and J. Mahoney (1996) 'Insights and Pitfalls: Selection Bias in Qualitative Research' *World Politics* 49, pp. 56-91.

- Collins, H.M. (2003) 'Lead Into Gold: The Science of Finding Nothing' *Studies in History and Philosophy of Science*, 34A, pp. 661-691.
- Collins, H.M. (1998) 'The Meaning of Data: Open and Closed Evidential Cultures in the Search for Gravitational Waves' *American Journal of Sociology* 104, pp. 293-338.
- Coppedge, M. (1999) 'Thickening Thin Concepts and Theories: Combining Large N and Small in Comparative Politics' *Comparative Politics* pp. 465-476.
- Delamont, S. and P. Atkinson (2001) 'Doctoring Uncertainty: Mastering Craft Knowledge' *Social Studies of Science*, 31, pp. 87-107.
- DeLillo, Don. (1997) *Underworld*. (New York: Simon & Schuster).
- Dewald, W.G., J.G. Thursby and R.G. Anderson (1986) 'Replication in Empirical Economics: The *Journal of Money, Credit and Banking Project*' *American Economic Review*
- Duncan, O.D. and M. Steinbeck (1989) 'Panels and Cohorts: Design and Model in the Study of Voting Turnout' in *Sociological Methodology 1989*, 18 (Washington: American Sociological Association).
- Empirical Economics* (Heidelberg: Springer-Verlag).
- Fagan, H. (2002) 'Grounding Waste: Towards a Sociology of Waste Networks' *NIRSA Working Paper Series*, 18, National University of Ireland.
- Feyerabend, P., I. Lakatos and M. Motterlini (1999) *For and Against Method: Including Lakatos's Lectures on Scientific Method and the Lakatos-Feyerabend Correspondence* (Chicago: University of Chicago Press).
- Flyvbjerg, B. (2000) *Making Social Science Matter: Why Social Inquiry Fails and How it Can Succeed Again* (Cambridge: Cambridge University Press).
- Freedman, D.A. (1991) 'Statistical Models and Shoe Leather' *Sociological Methodology* 21, pp. 291-313.
- Friedman, M. (1953) "The Methodology of Positive Economics." in *Essays in Positive Economics*, M. Friedman (ed.) (Chicago: University of Chicago Press).
- Gerber, A.S., D.P. Green and D. Nickerson (2001) 'Testing for Publication Bias in Political Science' *Political Analysis*, 9, pp. 385-392.
- Gerring, J. (2005) 'Causation: A Unified Framework for the Social Sciences' *Journal of Theoretical Politics* 17, pp. 163-198.
- Gerring, J. and P.A. Barresi (2003) 'Putting Ordinary Language to Work: A Min-Max Strategy of Concept Formation in the Social Sciences' *Journal of Theoretical Politics* 15, pp. 201-232.
- Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P. and Trow, M. (1994) *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies* (London: Sage).
- Goldsmith, J. and A. Vermeule (2002) 'Empirical Methodology and Legal Scholarship' *The University of Chicago Law Review* 69, pp. 153-167.
- Gould, S.J. (1993) 'Cordelia's Dilemma' *Natural History* 2, pp. 10-18.
- Granger, C.W. (2003) 'Time Series Analysis, Cointegration, and Applications' *Nobel Lecture*, Stockholm.
- Granger, C.W.J. 'Modeling, Evaluation and Methodology in the New Century' *Economic Inquiry*
- Granger, C.W.J. and P. Newbold (1974) 'Spurious Regressions in Econometrics' *Journal of Econometrics*, 2, pp. 111-120.

- Granger, C.W.J. and Y. Jeon (2004) 'Thick Modeling' *Economic Modelling* 21, pp. 323-343.
- Habermas, J. (1976). 'A Positivistically Bisected Rationalism' in T.W. Adorno, H. Albert, R. Dahrendorf, J. Habermas, H. Pilot, and K.R. Popper, Eds. *The Positivist Dispute in German Sociology* (Heinemann, London).
- Habermas, J. (1973) 'A Postscript to Knowledge and Human Interests' *Philosophy of the Social Sciences*, 3, pp. 157-189.
- Haig, B.D. (2003) 'What Is a Spurious Correlation?' *Understanding Statistics*, 2, pp. 125-132.
- Hall, D. (1999) 'Evidence Against One Hypothesis Turns Out to Be Evidence for Alternative Hypotheses' in Browman, H.I. (Ed) (1999) 'Theme Section: Negative Results' *Marine Ecology Progress Series*, 191, pp. 301-309.
- Hall, P. (2003) 'Aligning Ontology and Methodology in Comparative Research' in J. Mahoney and D. Rueschemeyer, Eds. *Comparative Historical Analysis in the Social Sciences* (Cambridge: Cambridge University Press).
- Hands, D.W. (2001) *Reflection Without Rules: Economic Methodology and Contemporary Science Theory* (Cambridge: Cambridge University Press).
- Hedges, L.V. (1992) 'Modeling Publication Selection Effects in Meta-Analysis' *Statistical Science*, 7, pp. 246-255.
- Hempel, C.G. (1965) *Aspects of Scientific Explanation* (New York: The Free Press).
- Hendry, D. (1980) 'Econometrics—Alchemy or Science?' *Economica*, 47, pp. 387-406.
- Hjort, N.L. and Claeskens, G. 'Frequentist Model Average Estimators' *Journal of the American Statistical Association*, 98, pp. 879-899.
- Hoeting, J.A., Madigan, D., Raftery, A.E. and Volinsky, C.T. (1999) 'Bayesian Model Averaging: A Tutorial' *Statistical Science*, 14, pp. 382-417.
- Ioannidis, J.P.A. (2005) 'Contradicted and Initially Stronger Effects in Highly Cited Clinical Research' *Journal of the American Medical Association*, 294, pp. 218-228.
- Iyengar, S. and J.B. Greenhouse (1988) 'Selection Models and the File-Drawer Problem (With Discussion)' *Statistical Science*, 3, pp. 109-135.
- Journal of Applied Econometrics* (UK: John Wiley & Sons, Ltd.).
- Journal of Articles in Support of the Null Hypothesis*, <<http://www.jasnh.com>>.
- Journal of Negative Results in BioMedicine*, <<http://www.jnrbrm.com>>.
- Journal of Negative Results in Ecology and Evolutionary Biology*, <<http://www.jnr-eeb.org>>.
- Journal of Negative Results in Speech and Audio Sciences*, <<http://journal.speech.cs.cmu.edu>>.
- Journal of Spurious Correlations: Qualitative and Quantitative Results in the Social Sciences*, <<http://www.jspurc.org>>.
- Kane, E.J. (1984) 'Why Journal Editors Should Encourage the Replication of Applied Econometric Research' *Quarterly Journal of Business and Economics*, 23.
- Kanazawa, S. (2001) 'Science vs. History: A Reply to MacDonald' *Social Forces* 80, pp. 349-352.
- Kennefick, D. (2000) 'Star Crushing: Theoretical Practice and the Theoreticians' Regress' *Social Studies of Science*, 30, pp. 5-40.
- King, G. (1995) 'Replication, Replication' *PS: Political Science and Politics*, 28, pp. 443-499.

- King, G., R.O. Keohane and S. Verba (1994) *Designing Social Inquiry: Scientific Inference in Qualitative Research* (Princeton: Princeton University Press).
- Kittel, B. (2005) 'The American Political Methodology Debate: Where is the Battlefield?' *Qualitative Methods Newsletter of the APSA Section on Qualitative Methods*, pp. 12-19.
- Knight, J. (2003) 'Negative Results: Null and Void' *Nature*, 422, pp. 554-555.
- Kotze, D.J., C.A. Johnson, R.B. O'Hara, K. Vepsäläinen and M.S. Fowler (2004) 'Editorial: The Journal of Negative Results in Ecology and Evolutionary Biology' *Journal of Negative Results in Ecology and Evolutionary Biology*, 1, pp. 1-5.
- Lakatos, I. (1970) 'Falsification and the Methodology of Scientific Research Programmes' in I. Lakatos and A. Musgrave (Eds) (1970) *Criticism and the Growth of Knowledge* (Cambridge: Cambridge University Press), pp. 91-196.
- Lakatos, I. (1974) 'Popper on Demarcation and Induction' in P.A. Schlipp, ed. *The Philosophy of Karl Popper: Book I* (La Salle, Illinois: Open Court), pp. 241-273.
- Lindsey, J.K. (1999) 'Some Statistical Heresies' *Journal of the Royal Statistical Society: Series D (The Statistician)*, 48, pp. 1-40.
- Longford, N.T. (2005) 'Editorial: Model Selection and Efficiency—Is 'Which Model...?' the Right Question?' *Journal of the Royal Statistical Society Series A*, 168, pp. 469-472.
- Lustick, I. (1997) 'The Discipline of Political Science: Studying the Culture of Rational Choice as a Case in Point' *PS: Political Science and Politics* pp. 175-179.
- Mahoney, J. and G. Goertz (2004) 'The Possibility Principle: Choosing Negative Cases in Comparative Research' *American Political Science Review*, 98, pp. 653-669.
- Mahoney, M.J. (1977) 'Publication Prejudices: An Experimental Study of Confirmatory Bias in the Peer Review System' *Cognitive Therapy and Research*, 1, pp. 161-175.
- Malhotra, S., N. Shafiq and P. Pandhi (2004) 'Advantages of Reporting, Publishing, and Reading "Negative" Studies' *Medscape Pharmacists*, 5.
- Mason, W.M. (1991) 'Freedman Is Right, As Far As He Goes, and It's Worse. Statisticians Could Help' *Sociological Methodology* 21, pp. 337-351.
- Massie, B.M. (2003) 'Editorial: Negative, Neutral, and Discordant Results: What Should an Editor Do?' *Journal of Cardiac Failure*, 2, pp. 77-79.
- Maxwell, C. (1981) 'Clinical Trials, Reviews, and the Journal of Negative Results' *British Journal of Clinical Pharmacology*, 1, pp. 15-18.
- Maxwell, G. (1974) 'Corroboration Without Demarcation' in P.A. Schlipp, ed. *The Philosophy of Karl Popper: Book I* (La Salle, Illinois: Open Court), pp. 292-321.
- Mayo, D. (1996) *Error and the Growth of Experimental Knowledge* (Chicago: University of Chicago Press).
- McGregor, J.P. (1993) 'Procrustus and the Regression Model: On the Misuse of the Regression Model' *PS: Political Science and Politics*, 26, pp. 801-804.
- McKeown, T.J. (1999) 'Case Studies and the Statistical Worldview' *International Organization* 53, pp. 161-190.
- Meehl, P.E. (1997) 'The Problem is Epistemology, Not Statistics: Replace Significance Tests by Confidence Intervals and Quantify Accuracy of Risky Numerical Predictions' in L.L. Harlow, S.A. Mulaik and J.H. Steiger (Eds) (1997) *What If There Were No Significance Tests?* (New Jersey: Lawrence Erlbaum Associates).

- Mestre, M.P. (2002) 'Editorial: The Relevance of Research Papers With "Negative Results"' *Revista Clínica Española*, 202, pp. 575-576.
- Miller, D. (2000) 'Sokal & Bricmont: Back to the Frying Pan' *Pli*, 9, pp. 156-173.
- Mittelstaedt, R. and T.S. Zorn 'Econometric Replication: Lessons from the Empirical Sciences' *Quarterly Journal of Business and Economics*, 23.
- Moe, T.M. (1979) 'On the Scientific Status of Rational Models', *American Journal of Political Science*, 23, pp. 215-243.
- Moses, J., B. Rihoux and B. Kittel (2005) 'Mapping Political Methodology: Reflections on a European Perspective', *European Political Science*, 4, pp. 55-68.
- Nurmi, H., J. Kacprzyk and M. Fedrizzi (1996) 'Probabilistic, Fatty and Rough Concepts in Social Choice' *European Journal of Operational Research* 95, pp. 264-277.
- Olsen, B. and C. Pfeffer (2002) 'Editorial' in *Journal of Negative Results in Biomedicine* 1.
- Petticrew, M. and H. Roberts (2005) *Systematic Reviews in the Social Sciences: A Practical Guide* (UK: Blackwell).
- Phillips, P.C.B (1986) 'Understanding Spurious Regressions in Econometrics' *Journal of Econometrics*, 33, pp. 311-340.
- Popper, K.R. (1963) *Conjectures and Refutations: The Growth of Scientific Knowledge* (UK: Routledge).
- Popper, K.R. (1971) 'Philosophical Comments on Tarski's Theory of Truth' in Popper, K.R. (1972) *Objective Knowledge: An Evolutionary Approach* (UK: Oxford).
- Prairie, Y.T. and D.F. Bird (1989) 'Some Misconceptions About the Spurious Correlation Problem in Ecological Literature' *Oecologica*, 81, pp. 285-288.
- Prechelt, L. (1997) 'Why We Need an Explicit Forum for Negative Results' *Journal of Universal Computer Science*, 3.
- Putnam, H. (1974) 'The "Corroboration" of Theories' in P.A. Schlipp, ed. *The Philosophy of Karl Popper: Book I* (La Salle, Illinois: Open Court), pp. 221-240.
- Quarterly Journal of Business and Economics*, <<http://www.qjbe.unl.edu>>
- Quine, W.V. (1974) 'On Popper's Negative Methodology' in P.A. Schlipp, ed. *The Philosophy of Karl Popper: Book I* (La Salle, Illinois: Open Court), pp. 218-220.
- Ragin, C.C. (2000) *Fuzzy-Set Social Science* (Chicago: University of Chicago Press).
- Ragin, C.C. and J. Sonnett (2004) 'Between Complexity and Parsimony: Limited Diversity, Counterfactual Cases, and Comparative Analysis' in S. Kopp and M. Minckenberg (Eds) *Vergleichen in der Politikwissenschaft* (Wiesbaden: VS Verlag für Sozialwissenschaften).
- Rapoport, A. (1958) 'Various Meanings of Theory' *American Political Science Review* 52, pp. 972-988.
- Rosenthal, R. (1979) 'The "File-Drawer Problem" and Tolerance for Null Results' *Psychological Bulletin* 86, pp. 638-641.
- Salmon, W. (1984) *Scientific Explanation and the Causal Structure of the World* (Princeton: Princeton University Press).
- Sartori, G. (1993) 'Totalitarianism, Model Mania and Learning from Error' *Journal of Theoretical Politics*, 5, pp. 5-22.
- Sartori, G. (2004) 'Where Is Political Science Going?' *PS: Political Science and Politics*, 4, pp. 785-787.

- Scarborough, E. and E. Tanenbaum (1998) *Research Strategies in the Social Sciences: A Guide to New Approaches*. (Oxford: Oxford University Press).
- Scargle, J.D. (2000) 'Publication Bias: The "File-Drawer" Problem in Scientific Inference' *Journal of Scientific Exploration* 14, pp. 91-106.
- Schmitter, P.C. (2001) 'Seven (Disputable) Theses Concerning the Future of 'Transatlanticized' or 'Globalized' Political Science', <<http://www.iue.it/SPS/People/Faculty/CurrentProfessors/bioSchmitter.shtml>>.
- Sharpe, D. (1997) 'Of Apples and Oranges, File Drawers and Garbage: Why Validity Issues in Meta-Analysis Will Not Go Away' *Clinical Psychology Review* 17, pp. 881-901.
- Shepperd, M. (2003) 'Empirically-Based Software Engineering' *Upgrade: The European Journal for the Informatics Professional*, 4, pp. 37-41.
- Sigelman, L. (1999) 'Publication Bias Reconsidered' *Political Analysis*, 8, pp. 201-210.
- Silvertown, J. and K.J. McConway (1997) 'Does "Publication Bias" Lead to Biased Science?' *Oikos* 19, pp. 167-168.
- Soberg, M. 'The Duhem-Quine Thesis and Experimental Economics: A Reinterpretation' *Discussion Paper* 329 (Norway: Statistics Norway Research Department).
- Stigler, G. (1982) 'The Process and Progress of Economics' *Nobel Lecture*, Stockholm.
- Tarrow, S. (1995) 'Bridging the Quantitative-Qualitative Divide in Political Science' *American Political Science Review* 89, pp. 471-474.
- Ward, A. (2004) 'How One Mistake Leads To Another: On the Importance of Verification/Replication' *Political Analysis*, 12, pp. 199-200.