https://doi.org/10.22370/pe.2026.20.5263

Evolving Perspectives on Breiman's Two Cultures: From Statistical Modeling to Contemporary Machine Learning

La evolución del debate de las dos culturas de Breiman: del modelado estadístico al aprendizaje automático contemporáneo

Cinthya Vergara Silva*
Universidad Adolfo Ibáñez

Resumen
Este artículo presenta una revisión crítica estructurada del debate entre el modelamiento estadístico y el modelamiento algorítmico, siguiendo el marco SALSA (Search, Appraisal, Synthesis y Analysis). La revisión se fundamenta en los trabajos de Breiman (2001a) sobre las "dos culturas" y de Athey e Imbens (2019) sobre el aprendizaje automático para economistas, complementados con literatura reciente identificada mediante Scopus, Web of Science y OpenAlex y evaluada mediante criterios explícitos de inclusión y exclusión. La síntesis analiza la evolución de la relación entre precisión predictiva e interpretabilidad, flexibilidad algorítmica e inferencia causal, así como sus implicancias para la investigación económica y la evaluación de políticas públicas. Asimismo, se examinan desarrollos recientes en ciencia basada en aprendizaje automático, inteligencia artificial explicable y confiable, y aprendizaje automático causal como extensiones contemporáneas del debate original. El artículo sostiene que el progreso futuro no dependerá del predominio de una única cultura metodológica, sino de la integración del razonamiento estadístico, los enfoques algorítmicos y el juicio científico mediante procesos transparentes de validación y decisiones metodológicas orientadas al problema.
Palabras clave: inferencia causal; aprendizaje automático; revisión crítica; econometría; interpretabilidad.

Resumo
Este artigo apresenta uma revisão crítica estruturada do debate entre a modelagem estatística e a modelagem algorítmica, seguindo o framework SALSA (Search, Appraisal, Synthesis e Analysis). A revisão é fundamentada nos trabalhos de Breiman (2001a) sobre as "duas culturas" e de Athey e Imbens (2019) sobre aprendizado de

*cinthya.vergaras@edu.uai.cl

máquina para economistas, complementados por literatura recente identificada por meio das bases Scopus, Web of Science e OpenAlex e avaliada segundo critérios explícitos de inclusão e exclusão. A síntese analisa a evolução da relação entre precisão preditiva e interpretabilidade, flexibilidade algorítmica e inferência causal, bem como suas implicações para a pesquisa econômica e a avaliação de políticas públicas. Além disso, são discutidos desenvolvimentos recentes em ciência baseada em aprendizado de máquina, inteligência artificial explicável e confiável, e aprendizado de máquina causal como extensões contemporâneas do debate original. O artigo argumenta que o progresso futuro não dependerá da predominância de uma única cultura metodológica, mas da integração do raciocínio estatístico, das abordagens algorítmicas e do julgamento científico por meio de processos transparentes de validação e escolhas metodológicas orientadas ao problema.
Palavras-chave: inferência causal; aprendizado de máquina; revisão crítica; econometria; interpretabilidade.

Abstract

This article presents a structured critical review of the debate between statistical modeling and algorithmic modeling, following the SALSA framework (Search, Appraisal, Synthesis, and Analysis). The review is anchored in Breiman's (2001a) "Two Cultures" and Athey and Imbens' (2019) perspective on machine learning for economists, with literature searches documented through Scopus, Web of Science, and OpenAlex and evaluated using explicit inclusion and exclusion criteria. The synthesis examines the evolving relationship between predictive accuracy and interpretability, algorithmic flexibility and causal inference, and their implications for economic research and policy evaluation. Recent developments in machine learning-based science, explainable and trustworthy artificial intelligence, and causal machine learning are discussed as extensions of the original debate. The review argues that future progress will depend not on the dominance of a single methodological culture, but on the integration of statistical reasoning, algorithmic approaches, and scientific judgment through transparent validation and problem-oriented methodological choices.
Keywords: causal inference; machine learning; structured critical review; econometrics; interpretability.

1. Introduction

Statistical analysis has traditionally relied on data models, in which knowledge is generated through explicit assumptions about the underlying data-generating process. Within this framework, researchers define the problem, select relevant variables, specify a model, and interpret the resulting estimates. Although this hypothesis-driven approach has produced major scientific advances, it also depends on analytical choices that may introduce researcher judgment and potential bias during study design and interpretation.

The expansion of computational capacity and data availability has promoted alternative approaches, including Knowledge Discovery from Data (KDD) (Fayyad et al.,1996), which emphasize the extraction of patterns from large datasets with fewer predefined structural assumptions. This shift contributed to the development of algorithmic modeling and machine learning (Sarker, 2021), challenging traditional statistical paradigms by placing greater emphasis on predictive performance, flexibility, and adaptation to complex data structures.

Rather than replacing statistical reasoning, these developments have expanded the range of available analytical tools. Modern approaches increasingly combine computational methods capable of capturing nonlinear relationships and high-dimensional patterns with theoretical frameworks that support interpretation, causal analysis, and decision-making. The resulting tension between prediction, explanation, and inference is at the center of the debate initiated by Breiman's (2001a) "Two Cultures" and its subsequent developments in economics and machine learning.

This article presents a structured critical review following the SALSA framework (Search, Appraisal, Synthesis, Analysis), anchored in Breiman (2001a) and Athey and Imbens (2019). The review is complemented by commentaries from the 20th-anniversary special issue of *Observational Studies* dedicated to Breiman's article, as well as additional works identified through a documented search protocol that extend the discussion toward deep learning, econometric identification, causal inference, and recent bibliometric evidence.

The contribution of this review is threefold. First, it provides an explicit and reproducible description of the literature search process across Scopus, Web of Science, and OpenAlex, including inclusion and exclusion criteria and verification of retrieved records. Second, it synthesizes the methodological dialogue between algorithmic modeling, econometrics, and causal inference, highlighting how machine learning can complement rather than substitute traditional analytical approaches. Third, it examines how the original debate has evolved toward contemporary challenges involving machine learning-based science, explainability, uncertainty, and responsible algorithmic decision-making. Section 2 describes the review protocol and search results; Section 3 discusses the two cultures debate and its economic interpretation; Section 4 examines contemporary challenges in prediction, inference, and scientific understanding; and Section 5 concludes by discussing implications for economic research, decision-making, and the future development of data-driven applications.

2. Methodological framework

2.1 Design

We conducted this review using the SALSA framework (Search, Appraisal, Synthesis, and Analysis) (Booth et al., 2021), a structured approach recommended for transparent literature reviews (García-Holgado et al., 2020) and aligned with the search stage of the PSALSAR methodology (Mengist et al., 2020). Given the conceptual and methodological objectives of the study, this is a structured critical review rather than a PRISMA-based systematic review or a meta-analysis.

The review is anchored on two seed articles representing complementary perspectives on statistical learning: Breiman's (2001a) Statistical Modeling: The Two Cultures and Athey and Imbens's (2019) Machine Learning Methods That Economists Should Know About. These works provide the conceptual foundation for examining how the debate has evolved in statistics, econometrics, and machine learning. The review therefore combines a structured search protocol with a qualitative comparison of methodological perspectives rather than quantitative evidence synthesis.

2.2 Search

Literature was identified through structured searches in Scopus (Elsevier Search API), Web of Science (Clarivate Starter API), and OpenAlex as complemented the indexed searches by providing reproducible citation networks and citation counts for the seed articles. Because cited-reference searches were unavailable through the API access level available for this study, identification relied on topic searches, document-type filters, citation metadata, and citation chaining from the seed articles. Searches were executed on 28–29 July 2026.

Given the conceptual objective of the review, a seed-based citation strategy was adopted rather than an exhaustive keyword search. The search was further anchored on the 20th-anniversary special issue of Observational Studies (Vol. 7, Issue 1, 2021), which provides a coherent set of methodological commentaries revisiting Breiman's original article. Detailed search strings, database outputs, record counts, and reproducibility materials are reported in Appendix A.

2.3 Study selection and synthesis

Studies were included when they (i) presented reviews, methodological syntheses, or commentaries on Breiman's two cultures or machine learning in economics; (ii) engaged substantively with the seed articles; (iii) were peer-reviewed publications or citable preprints with verifiable bibliographic metadata; and (iv) provided abstracts in English or Spanish. Purely empirical applications without methodological discussion and duplicate records were excluded.

Following the appraisal stage of SALSA, studies were evaluated according to their relevance to economics, methodological contribution, citation impact, and publication quality. The final synthesis contrasts the seed articles with complementary contributions addressing econometric identification, causal inference, deep learning, explainability, and recent bibliometric evidence.

Application of the appraisal criteria resulted in the selection of three complementary studies that extend the perspectives represented by the seed articles. Their contribution to the synthesis is summarized in Table 1.

Table 1: Complementary studies selected during the appraisal stage and their role in the review.

Study	Main contribution	Relation to Breiman's debate	Role in this review
Bhadra et al. (2021)	Revisits the statistical–algorithmic divide in the era of deep learning, emphasizing uncertainty quantification and the integration of statistical reasoning with flexible models.	Extends the "two cultures" debate to modern deep learning.	Supports Section 4.3 on machine learning–based science and the evolution of algorithmic modeling.
Bradić & Zhu (2021)	Discusses the implications of Breiman's arguments from an econometric perspective, emphasizing identification, structural assumptions, and interpretable models.	Bridges statistical learning and econometric methodology.	Supports Section 3.2 on Athey and Imbens's perspective.
Kennedy et al. (2021)	Examines causal estimation using modern machine learning, contrasting plug-in and targeted estimation.	Connects algorithmic prediction with valid causal inference.	Supports Section 4.2 on the prediction–inference trade-off.

The complete search protocol, appraisal process, and final corpus are documented in Appendix A.

2.4 Limitations

This review is designed as a structured critical synthesis rather than an exhaustive mapping of the literature. Because cited-reference searches were unavailable under the API access level used, study identification relied on topic searches, citation metadata, and citation chaining from the seed articles. Furthermore, the broad machine learning–economics query exceeded the Scopus API export limit and was used only to characterize the broader literature, not for study selection. Consequently, the findings should be interpreted as a qualitative synthesis of a defined corpus rather than as a comprehensive review of all publications in the field.

3. The two cultures: A Debate Beyond Its Time

3.1 Breiman's original perspective

Breiman (2001a) introduced the distinction between two emerging cultures within statistical practice: the *data modeling culture* and the *algorithmic modeling culture*. His central argument was not that one approach should replace the other, but that statistical research had become overly focused on models with explicit probabilistic assumptions and theoretical guarantees, while often underestimating the value of flexible predictive approaches.

According to Breiman, the data modeling culture is characterized by the specification of stochastic models intended to represent the underlying data-generating mechanism. These models provide interpretability and theoretical properties, but their validity depends strongly on whether the assumed structure adequately reflects the phenomenon under study. In many applications, identifying an appropriate model is itself a difficult task. The algorithmic modeling culture, in contrast, emphasizes the development of algorithms capable of extracting predictive patterns from data without requiring a predefined probabilistic structure. Although these methods may sacrifice some interpretability, they can provide greater flexibility and predictive accuracy, particularly in complex and high-dimensional settings.

Breiman argued that technological advances and the increasing availability of data created the need to broaden the statistical toolbox beyond traditional modeling strategies. Rather than searching for a single optimal model, he emphasized the possibility that multiple models may provide useful representations of the same phenomenon. This idea is closely related to the Rashomon effect, where several models with comparable predictive performance can coexist. He also discussed the tension between simplicity and performance through *Occam's razor*, and highlighted the challenges introduced by high-dimensional datasets through Bellman's *curse of dimensionality.*

The discussion following Breiman's article further developed these ideas. Efron recognized the contribution of algorithmic approaches as a positive extension of statistical practice, arguing that their success could stimulate both methodological innovation and theoretical developments rather than represent a rejection of traditional statistics. Hoadley emphasized the practical importance of algorithmic methods while noting the limitations associated with both cultures, including model assumptions, overfitting, and the need to evaluate algorithms according to their usefulness for specific problems rather than solely by error minimization.

A central issue in Breiman's argument concerns the interpretability of complex models. He described these models as a form of "black box" information and raised the question of how meaningful scientific knowledge can be extracted from highly predictive but difficult-to-understand algorithms. Parzen's commentary reinforced this concern by warning against models that provide technically correct answers to incorrectly formulated questions, referring to this as the "error of the third kind". This observation remains relevant in current discussions surrounding deep learning and large language models, where predictive capability does not necessarily imply causal understanding or scientific explanation.

Overall, Breiman's contribution can be interpreted as a call for methodological pluralism rather than a strict separation between statistics and machine learning. His medical example illustrated that complex algorithmic models may reveal relevant patterns that traditional statistical models fail to capture, even when the latter offer greater interpretability. The broader message was that statistical practice should pursue both accurate prediction and meaningful understanding of the mechanisms generating data.

More than two decades later, the two-cultures debate remains relevant because modern approaches increasingly seek to combine these objectives. Developments such as causal machine learning, uncertainty quantification, and explainable AI represent attempts to integrate predictive performance with interpretability and reliable inference. The following sections extend this discussion by examining judgment, prediction–inference trade-offs, and ethical considerations in economic applications.

3.2 Athey and Imbens's perspective

Athey and Imbens (2019) revisited Breiman's two-cultures debate from the perspective of empirical economics, examining how machine learning methods can contribute to economic research. Rather than presenting machine learning as a replacement for traditional statistical or econometric approaches, they argue that its value depends on the objective of the analysis. In particular, they distinguish between prediction problems, where machine learning methods often achieve substantial advantages, and causal questions, where economic research requires additional assumptions related to identification, theory, and interpretation.

The authors review the growing adoption of machine learning techniques in economics, including supervised learning, unsupervised learning, high-dimensional prediction methods, and emerging applications such as reinforcement learning, online experimentation, recommendation systems, and text analysis. However, they emphasize that the direct transfer of machine learning methods from predictive applications to economic research is not sufficient. Economic analysis frequently involves challenges such as endogeneity, treatment-effect estimation, and the need to understand the mechanisms generating observed outcomes.

A central contribution of their perspective is the argument that machine learning and econometrics address different but complementary objectives. While machine learning traditionally focuses on predictive performance and flexible approximation of complex relationships, econometrics emphasizes the estimation and interpretation of causal parameters under explicit assumptions. Consequently, economists have developed adaptations that incorporate the strengths of machine learning while preserving the inferential goals of economic analysis. These include regularization methods, sample splitting, orthogonalization, and approaches such as double/debiased machine learning, which allow flexible algorithms to be used within causal estimation frameworks.

Athey and Imbens therefore advocate incorporating machine learning methods into the econometric toolkit rather than treating them as an alternative paradigm. They argue that economists should become familiar with these methods because they provide new possibilities for handling high-dimensional data and complex relationships, while economic theory and identification strategies remain essential for meaningful interpretation.

This perspective was further developed in the 20th-anniversary special issue of Observational Studies, where Imbens and Athey (2021) revisited Breiman's original distinction from an econometric viewpoint. They argued that the separation between the two cultures had become less pronounced in practice: machine learning methods are increasingly valuable in economics when combined with causal reasoning and domain knowledge. From this perspective, the two cultures have neither fully merged nor remained independent; instead, contemporary empirical research increasingly reflects a hybrid approach in which algorithmic flexibility is integrated with theory-based identification strategies. This interpretation supports the broader argument developed in Section 4: predictive accuracy and causal validity should be viewed as complementary objectives rather than competing methodological commitments.

4. Beyond the two cultures: prediction, inference, and scientific understanding

The discussion in Section 3 suggests that the original opposition between Breiman's two cultures has evolved into a broader methodological landscape in which prediction, inference, and explanation are increasingly interconnected. Figure 1 summarizes this transition: the tension between interpretability and predictive performance that characterized Breiman's debate has motivated contemporary approaches that combine statistical reasoning, algorithmic flexibility, and domain knowledge through causal machine learning, explainable artificial intelligence, uncertainty quantification, and policy-oriented forecasting.

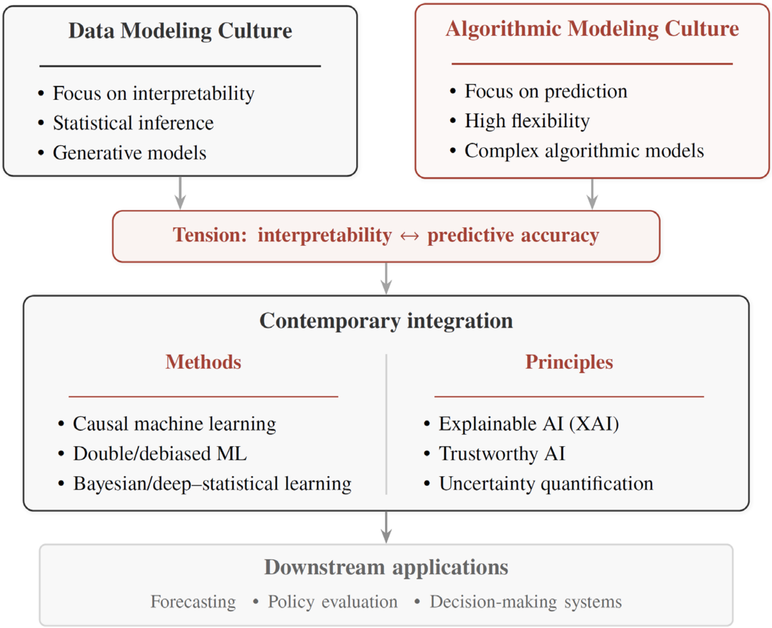

Figure 1: Conceptual evolution from Breiman's two cultures toward contemporary methodological integration. The figure highlights the transition from the interpretability–prediction tension to approaches combining algorithmic flexibility, causal reasoning, and principles of explainability and trustworthiness.

Rather than representing a complete convergence of the two cultures, this development can be understood as an expansion of the methodological space available to researchers. Modern analytical practice increasingly requires decisions about the purpose of modeling, the assumptions behind the data-generating process, and the type of knowledge expected from the analysis.

4.1 Methodological foundations: judgment, bias, and analytical frameworks

Data analysis is a process of continuous refinement in which conclusions depend not only on computational methods but also on how problems are formulated, data are collected, and assumptions are evaluated. The design of an analytical strategy therefore involves methodological judgments that influence what questions are asked, which variables are considered relevant, and how evidence is interpreted. Rigorous analytical practice does not eliminate judgment; rather, it makes such judgments explicit and evaluates their consequences (Majone & Quade, 1980; Tukey, 1962).

The role of researcher judgment has long been recognized as a fundamental aspect of scientific inquiry. Tukey (1962) distinguished several forms of judgment involved in data analysis: expertise derived from the domain generating the data, experience acquired through applying analytical techniques across different contexts, and understanding based on the theoretical and empirical properties of statistical methods. These forms of judgment remain relevant in contemporary machine learning applications, where choices regarding data representation, model selection, validation procedures, and interpretation directly influence results (Cumming, 2019; Kahneman, 2021).

The emergence of large-scale data analysis in the 1990s expanded this discussion by introducing new approaches for extracting patterns and knowledge from increasingly complex datasets. The Knowledge Discovery in Databases (KDD) framework (Fayyad et al., 1996) brought together data mining—whose methods have since evolved into much of what is now recognized as machine learning—alongside statistics, database systems, visualization, and high-performance computing to transform low-level data into higher-level knowledge. Rather than replacing traditional scientific reasoning, KDD broadened the analytical processes through which hypotheses, relationships, and predictions can be generated.

Despite differences among statistical modeling, machine learning, and causal inference, these approaches share a common objective: understanding relationships between outcomes (Y) and explanatory factors (X). However, they differ in the type of knowledge they seek. Causal models aim to identify mechanisms and estimate effects under explicit assumptions; machine learning methods primarily seek predictive accuracy and flexible representations of complex patterns; and analytical decision frameworks often combine mathematical, probabilistic, and empirical information to support actions under uncertainty (Angrist & Pischke, 2009; de Mast et al., 2023).

For applied economic research, this distinction leads to several methodological questions:

- **Research objective:** Is the primary goal prediction, causal inference, explanation, or decision support?
- **Interpretation requirements:** Will decisions require transparent mechanisms, regulatory justification, or post-hoc explanations?
- **Data characteristics:** Are available data sufficiently informative for flexible algorithms, or do sparsity and measurement limitations require stronger theoretical assumptions?
- **Ethical and institutional constraints:** Are fairness, accountability, privacy, or transparency requirements central to the application?
- **Validation criteria:** Should performance be evaluated only through predictive accuracy, or also through causal validity, robustness, uncertainty estimation, and generalization?

These questions illustrate that the central challenge is no longer choosing between statistical and algorithmic cultures but determining how their respective strengths can be combined according to the scientific objective of the analysis.

4.2 Balancing prediction and inference in economic applications

A central question emerging from the two-cultures debate is whether explicit deductive modeling is always required for every analytical objective. In many decision contexts, prediction may be the primary goal, whereas causal interpretation requires additional assumptions and methodological constraints that are not necessarily needed for forecasting. The challenge is therefore not to choose between prediction and inference, but to determine which objective is appropriate for the problem under study.

From an econometric perspective, subsequent discussions have emphasized that the value of machine learning depends on the research objective and the assumptions required for valid inference. Bradić and Zhu (2021) revisit Breiman's two cultures from this perspective, highlighting that predictive performance alone does not replace the importance of structural assumptions, identification strategies, and interpretable representations when the objective is scientific explanation or causal analysis.

This distinction is evident in algorithmic approaches such as random forests (Breiman, 2001b). Increasing model complexity through additional trees and deeper structures can improve predictive performance, but these models do not naturally provide the inferential framework associated with classical parametric approaches. Post-hoc interpretation methods, including feature importance measures and SHAP-based explanations, provide useful insights into model behavior, but they should not be interpreted as substitutes for formal causal or statistical inference (Hassija et al., 2024).

In economics, machine learning has expanded the possibilities for prediction and measurement by exploiting high-dimensional datasets and alternative data sources, including transaction records, digital traces, and text-based information. Applications in economic forecasting and policy analysis illustrate how algorithmic models can complement traditional econometric approaches (Athey et al., 2018; Chen et al., 2019; Woloszko, 2017). However, predictive performance alone does not resolve the central challenge of economic research: identifying the mechanisms underlying observed relationships.

Causal machine learning attempts to bridge this gap by combining flexible estimation with explicit identification strategies. Methods such as causal forests, double/debiased machine learning, and high-dimensional instrumental-variable approaches allow researchers to exploit complex patterns in data while preserving the inferential objectives of causal analysis. In the Breiman anniversary issue, Kennedy et al. (2021) connect this challenge directly to model complexity and causal interpretation, contrasting approaches such as plug-in and targeted estimation as strategies for adapting flexible machine learning estimators to causal questions while maintaining valid inference.

Pearl (2021), in the same special issue, provides a structural causal model perspective on Breiman's distinction. From this viewpoint, predictive accuracy alone cannot establish causal or counterfactual claims without an explicit representation of the causal structure generating the data. This interpretation echoes Breiman's concern about the "*error of the third kind*": answering the wrong question, even with a highly accurate model. For economic applications, the implication is that machine learning can enhance causal estimation, but it cannot replace the identification strategy required to define the causal question, specify what is being estimated and justify the assumptions under which causal conclusions are drawn.

4.3 Emerging frontiers: machine learning-based science, explainability, and ethics

Knowledge generation is increasingly incorporating machine learning not only as a predictive tool, but also to support scientific discovery, measurement, and hypothesis generation. The emerging paradigm of Artificial Intelligence for Science (AI4S) integrates artificial intelligence with scientific workflows, combining data-driven approaches with domain knowledge to accelerate discovery, automate parts of the research process, and address complex systems beyond the reach of traditional analytical methods (Wang et al., 2023). Rather than replacing traditional scientific approaches, machine learning-based science complements existing methodologies by expanding the available tools for exploring complex phenomena and extracting information from large-scale datasets (Kapoor et al., 2024; Ludwig & Mullainathan, 2024; Messeri & Crockett, 2024).

Bhadra et al. (2021) revisit Breiman's distinction in the context of deep learning, arguing that future analytical practices will require combining algorithmic flexibility with statistical reasoning, particularly when dealing with complex data-generating processes and uncertainty. Recent landscape reports further document the rapid integration of AI into scientific workflows and the expansion of these approaches across disciplines, emphasizing automated experimentation, knowledge discovery, and the integration of models, data, and scientific expertise as emerging components of contemporary research practices (Nature Research Intelligence, 2025).

This transition is also reflected in large-scale bibliometric evidence. Analyzing 4.9 million publications and 255 machine learning techniques indexed in *OpenAlex* between 1990 and 2026, Méndez Isla et al. (2026) identify increasing adoption of predictive approaches across scientific fields, including domains traditionally associated with interpretability and inference, such as the social sciences. These results provide large-scale evidence of changing methodological preferences that are consistent with the broader evolution of the debate initiated by Breiman, while also highlighting the need to understand how predictive tools interact with established standards of explanation and inference.

Emerging computational paradigms, including quantum machine learning, illustrate how continued increases in computational capability may introduce new methodological questions about the relationship between prediction, explanation, and scientific understanding. Although practical advantages of quantum-enhanced machine learning remain uncertain, these developments highlight a broader challenge: increasing computational power alone does not guarantee improved scientific insight, reproducibility, or interpretability (Castelvecchi, 2024). Thus, the central questions raised by Breiman—how to balance predictive performance, interpretability, and meaningful scientific understanding—remain relevant as machine learning becomes increasingly embedded in scientific discovery workflows and new computational paradigms emerge.

Generative artificial intelligence builds on earlier research on computational creativity and artificial generation, while recent advances have intensified the demand for explainability, transparency, and governance (van der Zant et al., 2013; Ooi et al., 2023; Kapoor et al., 2024). However, explainability alone does not guarantee trustworthy deployment. Recent surveys identify persistent challenges, including the standardized evaluation of explanations, alignment with legal and stakeholder requirements, and adaptation of explanations to specific decision contexts (Ali et al., 2023). In economic and policy settings, these challenges parallel the econometric emphasis on validated inference: post-hoc attributions may support accountability and interpretation, but they do not replace identification strategies, robustness analysis, or out-of-sample validation (Barocas et al., 2023; Hassija et al., 2024).

Thus, the central question raised by Breiman remains relevant: what does it mean to understand when models become increasingly predictive and complex? The evolution from statistical modeling to contemporary machine learning suggests that scientific progress will depend not on choosing between prediction and explanation, but on developing frameworks that clarify when, why, and under what assumptions predictive systems contribute to genuine understanding.

5. Conclusion

This review examined the evolution of the debate initiated by Breiman's "two cultures" through the perspectives of statistics, econometrics, and modern machine learning. The central tension between data modeling and algorithmic modeling has not disappeared; rather, it has evolved into a broader discussion about how prediction, inference, explanation, and decision-making can be combined in complex analytical environments.

The development of machine learning in economics illustrates this transformation. As emphasized by Athey and Imbens, machine learning methods provide valuable tools for high-dimensional prediction and flexible estimation, but their successful application in economics requires adaptation to the discipline's emphasis on causal structure, identification, and theoretical interpretation. Contemporary approaches such as causal machine learning and double machine learning represent attempts to integrate algorithmic flexibility with inferential validity rather than choosing between competing methodological cultures.

At the same time, emerging paradigms such as Artificial Intelligence for Science and generative artificial intelligence expand the role of machine learning from prediction toward scientific discovery, hypothesis generation, and knowledge production. These advances reinforce the importance of explainability, uncertainty quantification, reproducibility, and trustworthy deployment, particularly in domains where analytical results influence public policy, economic decisions, and social welfare.

The evidence reviewed here suggests that the future of data analysis is unlikely to be defined by the dominance of one methodological culture over another. Instead, progress will depend on selecting and combining methods according to the objectives of each problem: prediction when forecasting is required, causal identification when understanding mechanisms is essential, and transparent validation when decisions affect individuals and institutions.

Beyond the literature reviewed here, this synthesis is motivated by a recurring question in applied data analysis: identifying that a phenomenon occurs is only the first step; understanding the mechanisms behind it is what enables meaningful explanation and effective intervention. Statistics, mathematics, and computation should therefore not be viewed as competing traditions, but as complementary foundations that have progressively expanded our ability to analyze complex problems with greater speed, precision, and scope.

As machine learning and artificial intelligence continue to transform scientific discovery and decision-making, the central question is not whether algorithms will replace statistical reasoning, but how they can be integrated with theory, empirical evidence, and informed human judgment. Ultimately, the question raised by Breiman's two cultures was not which tradition should prevail, but how different forms of reasoning can be combined to transform data into reliable knowledge. The future of data analysis will depend on integrating prediction, understanding, and action: using algorithms to discover patterns, statistical reasoning to evaluate evidence, and scientific frameworks to guide better and responsible decisions.

Acknowledgments

The author would like to express her gratitude to Javiera Barrera for feedback and support in preparing this review. The author also acknowledges the use of AI-assisted tools for language editing and manuscript organization during the preparation of the manuscript. The analysis, synthesis, interpretation of the literature, and conclusions presented in this article remain the sole responsibility of the author.

Conflict of interest

The author declares that she has no conflicts of interest.

Appendix A. Search protocol and reproducibility materials

This appendix documents the search protocol and appraisal procedures underlying the SALSA (Search, Appraisal, Synthesis, and Analysis) framework described in Section 2. It reports the bibliographic sources, search strategy, representative queries, record counts, screening criteria, and the studies retained for qualitative synthesis. All searches were executed on **28–29 July 2026**. Machine-readable exports, complete query syntax, and supplementary reproducibility materials are available from the corresponding author upon request.

A.1 Search strategy

Bibliographic sources

Literature was identified through three complementary bibliographic databases:

- **Scopus** (Elsevier Search API)
- **Web of Science** (Clarivate Starter API)
- **OpenAlex** (REST API)

OpenAlex was additionally used to obtain reproducible citation counts and citation networks for the two seed articles. ArXiv and Google Scholar were consulted exclusively to verify bibliographic metadata and DOI information for selected preprints included in the final synthesis. It was not used as an independent bibliographic search source.

Full-text retrieval

Studies selected after screening were retrieved and evaluated using full-text versions obtained through institutional library subscriptions, DOI-resolved publisher websites, and open-access or preprint versions where available. These access routes supported qualitative appraisal and synthesis but were not considered additional search databases beyond the documented API-based searches.

Seed articles

The review was anchored on two foundational contributions representing complementary perspectives on statistical learning:

- **Breiman (2001a)** *Statistical Modeling: The Two Cultures* DOI: 10.1214/ss/1009213726 OpenAlex Work ID: **W2084341220**
- **Athey and Imbens (2019)** *Machine Learning Methods That Economists Should Know About* DOI: 10.1146/annurev-economics-080217-053950 OpenAlex Work ID: **W2922705140**

Search strategy

Given the conceptual objective of this review, literature identification followed a **seed-based citation strategy** rather than an exhaustive keyword search. The objective was to identify methodological contributions that explicitly engage with the conceptual debate initiated by the seed articles.

The search combined:

- topic-based searches,
- document-type filters,
- citation metadata,
- forward and backward citation chaining,
- OpenAlex citation networks anchored on the seed works.

To provide a coherent methodological corpus, the search was further anchored on the **20th-anniversary symposium of Breiman's article** published in *Observational Studies* (Volume 7, Issue 1, 2021). This symposium was identified through Scopus topic searches and independently verified using OpenAlex citation metadata.

Cited-reference searches

Direct-cited-reference searches could not be fully implemented because of API limitations. Specifically,

- the **Clarivate Starter API** does not support cited-reference (CR=) searches;
- Scopus REF() and REFDOI() queries returned no usable results under the available API access level.

Consequently, study identification relied on topic searches, document-type filtering, citation metadata, and OpenAlex cites: queries anchored on the seed articles.

A.2 Representative search queries

The complete query syntax is available from the corresponding author upon request. Representative queries are summarized below.

Appendix Table A.1. Representative search queries

Label	Database	Representative query
breiman_review_topic	Scopus	TITLE-ABS-KEY("two cultures" AND Breiman) + review document types
athey_review_topic	Scopus	Athey, Imbens, "machine learning", econom* + review document types
ml_econ_review	Scopus	"machine learning" AND econom* + Review + PUBYEAR > 2018
dt_review_ml_econ	Web of Science	TS=("machine learning" AND econom*) AND DT=(Review) AND PY=(2019-2026)
dt_review_two_cultures	Web of Science	TS=("two cultures") AND DT=(Review)
Breiman citing set	OpenAlex	filter=cites:W2084341220
Athey–Imbens citing set	OpenAlex	filter=cites:W2922705140

A.3 Search results

Appendix Table A.2. Search results

Query	Database	Records	Reproducibility evidence
Breiman (2001a) record	OpenAlex	4,308 cited-by	Seed record (W2084341220)
Athey & Imbens (2019) record	OpenAlex	1,064 cited-by	Seed record (W2922705140)
Works citing Breiman	OpenAlex	4,269	Full export (4,269 records)
Works citing Breiman with "two cultures" in title	OpenAlex	22	Full export
Review articles citing Breiman	OpenAlex	35	Full export
Works citing Athey & Imbens	OpenAlex	1,074	Full export
Review articles citing Athey & Imbens	OpenAlex	5	Export metadata
breiman_review_topic	Scopus	25	Full export
athey_review_topic	Scopus	2	Full export
ml_econ_review	Scopus	54,666 (5,000 exported)	API export limit
dt_review_ml_econ	Web of Science	2,255	Full export
dt_review_two_cultures	Web of Science	15	Full export

Source: Author compilation from API exports (28–29 July 2026).

Note. An earlier exploratory Web of Science search retrieved nine review records related to "two cultures". After the search strategy was standardized and documented, rerunning the query (TS=("two cultures") AND DT=(Review)) returned fifteen records. Because many of these records refer to unrelated uses of the expression "two cultures", the symposium corpus was defined using the Scopus breiman_review_topic search and independently cross-checked against OpenAlex citation metadata.

A.4 Screening and appraisal

Inclusion criteria

Studies were included when they:

- presented a review, commentary, or methodological synthesis related to Breiman's *Two Cultures* and/or machine learning in economics;
- engaged substantively with the conceptual contributions of the seed articles;
- were peer-reviewed publications or citable preprints with verifiable bibliographic metadata and DOI;
- provided an abstract in English or Spanish.

Exclusion criteria

The following records were excluded:

- purely empirical applications without methodological discussion;
- duplicate records across databases (one record retained per DOI or title–year combination);
- publications lacking sufficient bibliographic metadata for verification.

Appraisal procedure

The appraisal stage followed the SALSA framework and considered four complementary dimensions:

- relevance to economics and econometrics;
- methodological contribution;
- citation impact (OpenAlex citation counts, July 2026);
- publication quality.

Three studies were selected for detailed qualitative synthesis because they provide complementary methodological perspectives directly extending the Breiman–Athey debate:

- **Bhadra et al. (2021)** — statistical learning and deep learning;
- **Bradić and Zhu (2021)** — econometric interpretation of Breiman;
- **Kennedy et al. (2021)** — causal inference using modern machine learning.

Additional symposium commentaries (Imbens & Athey, 2021; Pearl, 2021) and one recent bibliometric study (Méndez Isla et al., 2026) were incorporated where they directly supported the arguments developed in Sections 3 and 4.

The broader Scopus, Web of Science, and OpenAlex searches were used to characterize the evolution of the review literature but were not screened exhaustively record by record.

A.5 Studies included in qualitative synthesis

Appendix Table A.3. Final synthesis corpus

Study	Venue	Role in the synthesis	DOI
Breiman (2001a)	*Statistical Science*	Seed article	10.1214/ss/1009213726
Athey & Imbens (2019)	*Annual Review of Economics*	Seed article	10.1146/annurev-economics-080217-053950
Imbens & Athey (2021)	*Observational Studies*	Econometric perspective	10.1353/obs.2021.0028
Bradić & Zhu (2021)	*Observational Studies*	Econometric commentary	10.1353/obs.2021.0019
Kennedy, Bonvini & Mishler (2021)	*Observational Studies*	Causal inference	10.1353/obs.2021.0001
Pearl (2021)	*Observational Studies*	Causal reasoning commentary	10.1353/obs.2021.0008
Bhadra et al. (2021)	arXiv preprint	Deep learning extension	10.48550/arXiv.2110.11561
Méndez Isla, Larivière & Kozlowski (2026)	arXiv preprint	Bibliometric extension	10.48550/arXiv.2606.20995

Source: Author compilation from Scopus, Web of Science, and OpenAlex API exports. Bibliographic metadata and full texts were verified through DOI-resolved publisher records, institutional library access, and arXiv where applicable.

A.6 Scope of the review

This appendix documents the procedures undertaken to enhance the transparency and reproducibility of the review. Consistent with the objectives of a **structured critical review**, rather than a PRISMA-style systematic review or meta-analysis, the search strategy was designed to identify, appraise, and synthesize the principal methodological contributions to the debate initiated by Breiman (2001a) and extended by Athey and Imbens (2019), rather than to provide an exhaustive inventory of every publication related to machine learning in economics.

The resulting synthesis should therefore be interpreted as a transparent and reproducible qualitative assessment of a bounded corpus of influential methodological literature, rather than as a comprehensive mapping of the entire field.

References

ALI, Sajid; ABUHMED, Tamer; EL-SAPPAGH, Shaker; MUHAMMAD, Khan; ALONSO-MORAL, José M.; CONFALONIERI, Roberto; GUIDOTTI, Riccardo; DEL SER, Javier; DÍAZ-RODRÍGUEZ, Natalia y HERRERA, Francisco (2023). "Explainable artificial intelligence (XAI): What we know and what is left to attain trustworthy artificial intelligence". Information Fusion, 99, 101805. https://doi.org/10.1016/j.inffus.2023.101805

ANGRIST, Joshua D. y PISCHKE, Jörn-Steffen (2009). Mostly Harmless Econometrics: An Empiricist's Companion. Princeton: Princeton University Press.

ATHEY, Susan; AGRAWAL, Ajay; GANS, Joshua y GOLDFARB, Avi (2018). "The impact of machine learning on economics". En The Economics of Artificial Intelligence: An Agenda, pp. 507–547. Chicago: University of Chicago Press.

ATHEY, Susan y IMBENS, Guido W. (2019). "Machine learning methods that economists should know about". Annual Review of Economics, 11(1), 685–725. https://doi.org/10.1146/annurev-economics-080217-053433

BAROCAS, Solon; HARDT, Moritz y NARAYANAN, Arvind (2023). Fairness and Machine Learning: Limitations and Opportunities. Cambridge, MA: MIT Press.

BHADRA, Anindya; DATTA, Jyotishka; POLSON, Nicholas; SOKOLOV, Vadim y XU, Jianeng (2021). "Merging two cultures: Deep and statistical learning". arXiv preprint arXiv:2110.11561. https://doi.org/10.48550/arXiv.2110.11561

BOOTH, Andrew; ST. JAMES, Maxine; CLOWES, Mark y SUTTON, Anthea (2021). Systematic Approaches to a Successful Literature Review. Londres: SAGE Publications Ltd.

BRADIĆ, Jelena y ZHU, Yinchu (2021). "Comments on Leo Breiman's paper 'Statistical modeling: The two cultures' (Statistical Science, 2001, 16(3), 199–231)". Observational Studies, 7(1), 21–31. https://doi.org/10.1353/obs.2021.0019

BREIMAN, Leo (2001a). "Statistical modeling: The two cultures (with comments and a rejoinder by the author)". Statistical Science, 16(3), 199–231. https://doi.org/10.1214/ss/1009213726

BREIMAN, Leo (2001b). "Random forests". Machine Learning, 45, 5–32. https://doi.org/10.1023/A:1010933404324

CALIN-JAGEMAN, Robert J. y CUMMING, Geoff (2019). "The new statistics for better science: Ask how much, how uncertain, and what else is known". The American Statistician, 73(sup1), 271–280. https://doi.org/10.1080/00031305.2019.1616221

CASTELVECCHI, Davide (2024). "The AI-quantum computing mash-up: Will it revolutionize science?". Nature. https://doi.org/10.1038/s41586-024-08888-7

CHEN, Jenny C.; DUNN, Abe; HOOD, Kyle; DRIESSEN, Alexander y BATCH, Andrea (2019). "Off to the races: A comparison of machine learning and alternative data for predicting economic indicators". En Big Data for 21st Century Economic Statistics, pp. 123–145. Chicago: University of Chicago Press.

DE MAST, Jeroen; STEINER, Stefan H.; NUIJTEN, Wim P. M. y KAPITAN, Diederik (2023). "Analytical problem solving based on causal, correlational, and deductive models". The

American Statistician, 77(1), 51–61. https://doi.org/10.1080/00031305.2023.2100701

FAYYAD, Usama; PIATETSKY-SHAPIRO, Gregory y SMYTH, Padhraic (1996). "From data mining to knowledge discovery in databases". AI Magazine, 17(3), 37–54. https://doi.org/10.1609/aimag.v17i3.1464

GARCÍA-HOLGADO, Alicia; MARCOS-PABLOS, Samuel y GARCÍA-PEÑALVO, Francisco (2020). "Guidelines for performing systematic research projects reviews". International Journal of Interactive Multimedia and Artificial Intelligence, 7(1), 19–26. https://doi.org/10.9781/ijimai.2020.02.002

HASSIJA, Vikas; CHAMOLA, Vinay; MAHAPATRA, Adhar; SINGAL, Abhinandan; GOEL, Divyansh; HUANG, Kaizhu; SCARDAPANE, Simone; SPINELLI, Indro; MAHMUD, Mufti y HUSSAIN, Amir (2024). "Interpreting black-box models: A review on explainable artificial intelligence". Cognitive Computation, 16(1), 45–74. https://doi.org/10.1007/s12559-024-09803-w

IMBENS, Guido W. y ATHEY, Susan (2021). "Breiman's two cultures: A perspective from econometrics". Observational Studies, 7(1), 127–133. https://doi.org/10.1353/obs.2021.0028

KAHNEMAN, Daniel; SIBONY, Olivier y SUNSTEIN, Cass R. (2021). Noise: A Flaw in Human Judgment. Londres: Hachette UK.

KAPOOR, Sayash; CANTRELL, Emma M.; PENG, Kaicheng; PHAM, Tom H.; BAIL, Christopher A.; GUNDERSEN, Odd Erik; HOFMAN, Jake M.; HULLMAN, Jessica; LONES, Michael A.; MALIK, Momin M. et al. (2024). "REFORMS: Consensus-based recommendations for machine-learning-based science". Science Advances, 10(18), eadk3452. https://doi.org/10.1126/sciadv.adk3452

KENNEDY, Edward H.; BONVINI, Matteo y MISHLER, Alan (2021). "Comment on 'Statistical modeling: The two cultures'

by Leo Breiman". Observational Studies, 7(1), 145–156. https://doi.org/10.1353/obs.2021.0001
LUDWIG, Jens y MULLAINATHAN, Sendhil (2024). "Machine learning as a tool for hypothesis generation". The Quarterly Journal of Economics, 139(2), 751–827. https://doi.org/10.1093/qje/qjz016
MAJONE, Giandomenico y QUADE, Edward S. (1980). Pitfalls of Analysis (vol. 8). Nueva York: John Wiley & Sons.
MÉNDEZ ISLA, Malena; LARIVIÈRE, Vincent y KOZLOWSKI, Diego (2026). "From inference to prediction: How machine learning is reconfiguring science (1990–2025)". arXiv preprint arXiv:2606.20995. https://doi.org/10.48550/arXiv.2606.20995
MENGIST, Wondimagegn; SOROMESSA, Teshome y LEGESE, Getachew (2020). "Method for conducting systematic literature review and meta-analysis for environmental science research". MethodsX, 7, 100777. https://doi.org/10.1016/j.mex.2019.100777
MESSERI, Lisa y CROCKETT, Molly J. (2024). "Artificial intelligence and illusions of understanding in scientific research". Nature, 627(8002), 49–58. https://doi.org/10.1038/s41586-024-08900-w
NATURE RESEARCH INTELLIGENCE. (2025). AI for Science 2025 [Report]. Nature Portfolio. https://www.nature.com/collections/bfefgbacag
OOI, Keng-Boon; TAN, Garry Wei-Han; AL-EMRAN, Mostafa; AL-SHARAFI, Mohammed A.; CAPATINA, Alexandru; CHAKRABORTY, Abhishek; DWIVEDI, Yogesh K.; HUANG, Tzu-Ling; KAR, Arpan K.; LEE, Voon-Hsien et al. (2023). "The potential of generative artificial intelligence across disciplines: Perspectives and future directions". Journal of Computer Information Systems, 1–32. https://doi.org/10.1080/08874417.2023.2174927
PEARL, Judea (2021). "Causally colored reflections on Leo Breiman's 'Statistical modeling: The two cultures' (2001)". Observational

Studies, 7(1). https://doi.org/10.1353/obs.2021.0008

SARKER, Iqbal H. (2021). "Machine learning: Algorithms, real-world applications, and research directions". SN Computer Science, 2(3), 160. https://doi.org/10.1007/s42979-021-00439-x

TUKEY, John W. (1962). "The future of data analysis". The Annals of Mathematical Statistics, 33(1), 1–67. https://doi.org/10.1214/aoms/1177704711

VAN DER ZANT, Tijn; KOUW, Matthijs y SCHOMAKER, Lambert (2013). Generative Artificial Intelligence. Berlín: Springer.

WANG, H., Fu, T., DU, Y. et al. (2023). Scientific discovery in the age of artificial intelligence. Nature 620, 47–60 https://doi.org/10.1038/s41586-023-06221-2

WOLOSZKO, Nicolas (2017). "Making better economic forecasts with machine learning". OECD Economics Department Working Papers. https://doi.org/10.1787/0b9a7614-en