Chapter 11
Nonparametric Econometrics

Daniel J. Henderson and Christopher F. Parmeter

Abstract Nonparametric econometrics emerged as a response to the limitations of traditional parametric modeling, offering tools capable of capturing economic relationships without imposing restrictive functional forms. Yet its history has been shaped as much by skepticism as by enthusiasm. Early critics questioned whether nonparametric methods could deliver meaningful identification, credible inference, and economically interpretable results, concerns amplified by data limitations, computational burdens, and the curse of dimensionality. Over time, advances in kernel smoothing, semiparametric modeling, mixed-data methods, shape restrictions, and instrumental-variable estimation, together with a deeper understanding of identification and inference in infinite-dimensional settings, transformed nonparametrics from a largely statistical import into an enduring part of econometric methodology. This chapter traces that development from its statistical pre-history through its econometric maturation and later interactions with structural modeling, causal inference, and machine learning. The chapter closes by assessing which contributions proved most durable in practice and which turned out to have narrower practical reach than early hopes suggested, with the aim of drawing broader lessons for econometric methodology.

11.1 The Beginnings

Nonparametric econometrics occupies a distinctive place in the history of empirical economics. From the outset, it promised to relax rigid functional-form assumptions and thereby allow empirical relationships to be learned more directly from the data. At the same time, its emergence challenged deeply held views about what

Daniel J. Henderson ✉
University of Alabama, Tuscaloosa, USA, e-mail: daniel.henderson@ua.edu

Christopher F. Parmeter
University of Miami, Coral Gables, USA, e-mail: cparmeter@bus.miami.edu

econometrics ought to be: a discipline closely tied to economic theory, organized around interpretable parameters, and grounded in tractable stochastic models. The tension between these two impulses — flexibility and structure — has shaped the field's development for more than half a century.

Because nonparametric ideas originated in statistics rather than economics, their adoption within econometrics required substantial conceptual reorientation. Econometricians were accustomed to environments in which theory and specification were tightly linked. Early nonparametric methods, by contrast, appeared to replace explicit functional forms with smoothness assumptions, kernels, and bandwidths. For many economists, those ingredients seemed opaque, arbitrary, or insufficiently connected to economic reasoning. The resulting skepticism was not merely technical. It reflected a broader concern about the role of empirical evidence in economics and about how far one could move toward descriptive flexibility without sacrificing theoretical discipline.

These concerns were reinforced by practical constraints. Early nonparametric estimators were highly sensitive to dimensionality, often requiring more data than empirical economists could realistically obtain. Inference was delicate, identification was often less transparent than under parametric structure, and computation could be burdensome in an era when even modest numerical tasks were nontrivial. In this sense, nonparametric econometrics developed under dual pressures: intellectual scrutiny and practical limitation. Yet those same pressures proved productive. They helped motivate semiparametric models, theory-guided shape restrictions, identification-aware regularization, mixed-data methods, and later the absorption of nonparametric ideas into machine-learning-assisted econometrics.

Today, nonparametric methods are best understood not as replacements for parametric econometrics, but as part of a broader methodological ecosystem. They serve as stand-alone estimators in some settings, as diagnostic tools in others, as components of semiparametric procedures, and as nuisance-function technologies in modern causal and structural work. Their history is therefore one of gradual absorption rather than abrupt methodological displacement. The aim of this chapter is to trace that history and to show how the field matured by confronting, rather than evading, the objections that initially limited its acceptance.

11.1.1 What 'Nonparametric' Meant — Then and Now

The meaning of 'nonparametric' has shifted over time. In the early stages of econometrics, the term often referred rather loosely to procedures that avoided a finite-dimensional parameterization, including rank-based methods, empirical distribution functions, and early smoothing devices. These methods were generally viewed as descriptive, diagnostic, or robust alternatives to conventional parametric analysis rather than as full-fledged econometric estimators. The prevailing view was that serious econometric work required explicit functional forms grounded in economic reasoning, a mindset that naturally placed nonparametric methods at the margin of the discipline.

As nonparametric density estimation and regression developed more fully, however, the meaning of the term became sharper. 'Nonparametric' came to refer not merely to the absence of a parametric form, but to a setting in which the object of interest is an unknown function or distribution rather than a finite-dimensional vector. This shift had important consequences. It reframed nonparametrics as a legitimate inferential framework with its own asymptotic theory, convergence rates, identification questions, and regularization problems. What had once looked like ad hoc smoothing increasingly became a mathematically principled approach to estimation.

The modern understanding of nonparametric econometrics goes one step further. It no longer suggests unconstrained flexibility. Contemporary practice makes clear that smoothness conditions, regularization, support conditions, and shape restrictions are central to rendering nonparametric estimation feasible and economically interpretable. In that sense, nonparametrics is now better understood as one end of a continuum that includes semiparametric, mixed-data, and shape-restricted methods. The issue is not whether structure is present, but what kind of structure is imposed and how transparent its role is.

This evolution has become even more visible with the growth of machine learning and high-dimensional statistics. Many modern methods — random forests, neural networks, boosting, Gaussian processes — are nonparametric in the broad sense that their complexity can grow with the data. Yet their goals, computational logic, and regularization strategies often differ from those of classical kernel and series methods. As a result, the boundary between nonparametric econometrics and predictive machine learning has become increasingly porous. What remains constant across these developments is the central question: how much structure does empirical economics actually need in order to learn something credible from data?

11.1.2 Why Nonparametrics Emerged in Econometrics

Nonparametric methods emerged in econometrics because researchers increasingly recognized that many empirical questions could not be addressed satisfactorily within the confines of standard parametric models. For much of the twentieth century, econometric practice relied heavily on linearity, separability, and fixed functional forms justified either by economic theory or by tractability. Those choices were often practical necessities, but they also limited the ability of researchers to capture nonlinearities, heterogeneous responses, and complex distributional features present in the data. As empirical work expanded into labor economics, consumer choice, industrial organization, program evaluation, and related fields, these limitations became harder to ignore.

The rise of nonparametrics also reflected changes in the data environment. The postwar expansion of household surveys, administrative records, and other richer sources of empirical information created settings in which more flexible methods became both more attractive and, eventually, more feasible. Nonparametric estimators promised to uncover heterogeneity, nonlinear structure, and distributional variation without requiring those features to be built in ex ante. Their appeal grew not simply because they were mathematically elegant, but because they aligned with an increasingly empirical research culture that valued flexible descriptions of observed relationships.

At the same time, developments in mathematical statistics supplied the conceptual foundations needed to make nonparametric methods credible to econometricians. Work on density estimation, kernel regression, empirical processes, consistency, and optimal convergence rates gradually transformed early smoothing devices into coherent estimation frameworks. As these ideas entered econometrics through graduate training, surveys, monographs, and methodological papers, nonparametric methods came to be seen less as foreign statistical curiosities and more as useful complements to economic modeling.

Perhaps most fundamentally, nonparametrics emerged because economists were searching for a middle ground between theoretical discipline and empirical realism. Strong parametric forms could clarify interpretation, but they could also drive results too forcefully. At the same time, few econometricians wanted to abandon structure altogether. Nonparametric methods offered an alternative: a way to relax some forms of structure while still retaining the possibility of disciplined empirical analysis. In this sense, the rise of nonparametrics reflected a broader methodological shift within economics — an acknowledgment that empirical credibility sometimes requires more flexibility than traditional parametric forms allow.

11.1.3 Relationship to Parametric and Semiparametric Traditions

The relationship between nonparametric econometrics and parametric econometrics has always been both complementary and contested. Parametric models dominated early econometric analysis because they offered interpretable parameters, explicit links to economic theory, and familiar inferential tools. Those virtues remain important. What nonparametric methods challenged was not the value of structure itself, but the habit of embedding that structure too quickly in rigid functional forms.

This tension became especially productive through the development of semiparametric models. Semiparametric approaches preserved low-dimensional parameters of direct economic interest while leaving other components — such as regression functions, error distributions, or link functions — unspecified. In this way, they provided a bridge between the interpretability of parametric models and the flexibility of nonparametric ones. Historically, that bridge mattered a great deal. Many econometricians who were uneasy about fully unrestricted nonparametrics found semiparametric models to be a more acceptable route toward flexibility.

Seen from this perspective, the contrast among parametric, semiparametric, and nonparametric methods is best understood as a matter of how structure is allocated across the model rather than as a rigid taxonomy. Parametric methods place most of the structure up front; nonparametric methods relax more of it; semiparametric methods do so selectively. This more graduated view proved historically important because it helped the profession move away from a crude opposition between theory-driven structure and flexible estimation.

That reorientation also changed how nonparametrics itself was understood. Nonparametric methods increasingly came to be seen not as a rejection of structure, but as a framework for deciding more carefully which kinds of structure were worth maintaining. Smoothness, shape restrictions, support conditions, and regularization are not afterthoughts in this framework; they are part of the essential discipline that makes infinite-dimensional estimation possible. The same is true of semiparametric work, where nonparametric components are introduced precisely where parametric assumptions are least credible or least necessary.

In contemporary practice, the boundaries among these traditions are increasingly fluid. Nonparametric methods serve as diagnostics for parametric specification, building blocks for semiparametric estimators, tools for distributional and counterfactual analysis, and flexible nuisance estimators in causal and structural settings. Parametric and semiparametric models, in turn, often borrow their logic from nonparametric identification and robustness analysis. The result is a methodological ecosystem in which flexibility and structure coexist rather than exclude one another. This chapter takes that ecosystem as its starting point.

11.1.4 Scope and Organization of the Chapter

This chapter provides a historical and conceptual overview of nonparametric econometrics, tracing its development from statistical pre-history to its later integration into modern econometric practice. The emphasis is not encyclopedic. Our aim is not to catalog every method or contributor, but to identify the main intellectual shifts, methodological debates, and practical constraints that shaped the field's development. A central theme throughout is that many of the field's most important advances were responses to criticism: concerns about interpretability, identification, inference, dimensionality, and feasibility were not external obstacles, but part of what made the literature mature.

No chapter of this type can do full justice to every paper, author, or subliterature that contributed to the development of nonparametric econometrics. Inevitably, many important contributions receive only brief mention, and others are omitted from detailed discussion altogether. Such omissions should not be read as judgments about priority, quality, or importance, nor as implying any lack of awareness on our part. They reflect the ordinary editorial constraints of writing a coherent historical chapter rather than a comprehensive bibliography.

The chapter begins in Section 11.2 with the statistical foundations that predate the formal entry of nonparametrics into economics, including early work on density estimation, smoothing, empirical processes, and distribution-free inference. Section 11.3 then examines the first wave of econometric engagement, emphasizing both the promise of nonparametric methods and the skepticism they encountered. Section 11.4 turns to identification without functional forms, a development that forced the field to confront more explicitly what data and assumptions could jointly deliver. Section 11.5 addresses testing, confidence sets, and inference, showing how nonparametric econometrics became a more complete statistical framework rather than simply a collection of flexible estimators.

Section 11.6 considers semiparametrics as a bridge between rigid parametric specification and unrestricted flexibility, while Section 11.7 examines the move beyond the cross-section into time-series and panel environments. Section 11.8 treats nonparametric instrumental variables, where endogeneity, ill-posedness, and regularization meet directly. Section 11.9 examines shape restrictions and the way economic theory re-entered nonparametric econometrics through qualitative constraints rather than closed-form specification.

The next group of sections highlights domains in which nonparametric methods proved especially revealing. Section 11.10 considers inequality, welfare, and distributional analysis. Section 11.11 examines demand, auctions, and industrial organization, where economic theory provided especially rich qualitative discipline. Section 11.12 turns to discrete smoothing and mixed-data methods, an important step in making nonparametric econometrics usable on the kinds of support structures that empirical economists routinely confront. Section 11.13 then considers computational advances and practical adoption, tracing how improvements in algorithms, software, and computing power changed the empirical meaning of the field.

The chapter closes in Section 11.14 by asking which contributions proved most durable and which turned out to have narrower practical reach than early hopes suggested. That retrospective assessment is not an appendage to the history; it is part of the history itself. One of the goals of this chapter is precisely to understand how nonparametric econometrics changed by confronting the limits of its own early ambitions.

11.2 Pre-History: Statistical Foundations Outside Economics

The intellectual roots of nonparametric econometrics lie not in economics itself, but in earlier developments within mathematical statistics. Long before economists began debating whether flexible estimators could be reconciled with economic theory, statisticians were already studying how to estimate densities, regression functions, and distribution functions without committing to finite-dimensional parametric families. These developments did not arise in response to the problems that later came to dominate econometric discussion—identification, endogeneity, structural interpretation, or causal inference—but they created the mathematical and conceptual language in which such problems could eventually be reformulated; see, among others, Ullah (1988), Härdle and Linton (1994), and Pagan and Ullah (1999). In this sense, the pre-history of nonparametric econometrics is best understood not as an early phase of econometrics proper, but as a statistical literature that later proved adaptable to econometric purposes.

What makes this pre-history historically important is therefore not only the appearance of smoothing devices or distribution-free procedures, but the gradual emergence of a different inferential perspective: unknown objects of interest need not be finite-dimensional parameter vectors. Densities, regression curves, and distribution functions could themselves be treated as estimands. That shift was profound, but it did not immediately translate into econometric adoption. For economists trained in traditions that tied empirical specification closely to economic theory, these statistical advances initially appeared somewhat distant from the central tasks of econometric analysis; see the retrospective discussions in Härdle and Linton (1994) and Pagan and Ullah (1999). The history traced in this section is thus one of conceptual preparation rather than direct disciplinary uptake.

11.2.1 Early Density Estimation and Smoothing

Among the earliest statistical developments that later mattered for nonparametric econometrics was the problem of density estimation without distributional specification. By the middle of the twentieth century, statisticians had already begun to ask how one might recover an unknown density directly from the data rather than by selecting a member of a parametric family. The now-classic answers were given by Rosenblatt (1956) and Parzen (1962), who independently proposed what came to be known as kernel density estimators. Their contribution was not merely technical. It marked a shift in inferential emphasis away from parameterizing distributions and toward estimating an unknown function itself, a point emphasized in later retrospective treatments such as Härdle and Linton (1994) and Pagan and Ullah (1999).

In hindsight, this move appears central to the later development of nonparametric econometrics, but at the time it was primarily a development within statistics. Rosenblatt (1956) formalized the idea that the density at a point could be approximated by averaging observations in a neighborhood of that point, weighting them according to distance. Parzen (1962) broadened this framework and clarified the conditions under which such estimators are consistent, while also drawing attention to the role of the smoothing parameter—later universally referred to as the bandwidth—in governing the bias–variance tradeoff. What mattered historically was that the density itself had become the object of estimation.

Closely related ideas soon appeared in nonparametric regression. Although regression would later become the main entry point for nonparametrics in econometrics, the earliest contributions again came from statistics rather than economics. The estimators proposed independently by Nadaraya (1964) and Watson (1964) extended kernel ideas from density estimation to conditional means, yielding what is now known as the Nadaraya–Watson estimator. Instead of fitting a global linear relationship, these procedures estimated the regression function through local averaging, allowing the conditional mean to vary flexibly over the support of the regressors. For later econometric discussions of why this mattered, see Ullah (1988), Härdle and Linton (1994), and Pagan and Ullah (1999).

From the standpoint of later econometric developments, the importance of these contributions lies less in immediate application than in conceptual reorientation. They showed that regression functions, like densities, could be treated as unknown objects to be estimated directly. Yet they were not developed with the structural or causal concerns of econometrics in mind. Their original setting was one of statistical recovery and smoothing, not one of identification, endogeneity, or behavioral interpretation. This gap between statistical possibility and econometric relevance helps explain why the importation of such methods into econometrics was neither immediate nor seamless.

11.2.2 Kernel Methods, Local Averaging, and Statistical Intuition

The methods introduced by Rosenblatt (1956), Parzen (1962), Nadaraya (1964), and Watson (1964) can also be situated within an older and more intuitive family of smoothing procedures based on local averaging. Long before kernel estimation acquired formal asymptotic theory, statisticians and time-series analysts used moving averages to reduce noise and reveal latent patterns in data. Such procedures were especially familiar in descriptive studies of trends and business cycles, where the immediate goal was often not formal inference but the extraction of smoother underlying movements from irregular observations. While this older smoothing tradition was not yet nonparametric econometrics in any modern sense, it provided an intuitive precursor to later formal developments; see the retrospective discussions in Härdle and Linton (1994) and Pagan and Ullah (1999).

Kernel methods may be viewed as a mathematically refined extension of this local-averaging logic. Rather than assigning equal weight to all observations within a fixed window, kernel estimators allow weights to vary smoothly with distance from the evaluation point. This made the procedure both more flexible and more amenable to formal analysis. In regression settings, the Nadaraya–Watson estimator effectively applied the moving-average idea in covariate space rather than in time, with nearby observations receiving greater weight than distant ones. As later surveys emphasized, this local weighting principle became foundational for large parts of the nonparametric literature; see Ullah (1988), Härdle and Linton (1994), and Fan and Gijbels (1996).

Historically, however, the significance of kernel methods lay not only in what they made possible, but also in the difficulties they made visible. Even in these early forms, smoothing methods raised questions that would later become central in econometrics: how should the degree of smoothing be chosen, how sensitive are conclusions to that choice, and how quickly do such methods deteriorate as the number of explanatory variables increases? These problems did not yet amount to an econometric research program, but they foreshadowed many of the objections that economists would later raise when nonparametric techniques began to move into econometric practice.

11.2.3 Empirical Distribution Functions and Glivenko–Cantelli

If kernel estimators illustrated how unknown functions might be smoothed from data, the empirical cummulative distribution function (ECDF) provided an even earlier and more fundamental example of nonparametric reasoning. Given a random sample, the ECDF estimates the underlying cumulative distribution function by placing mass $1/n$ on each observed data point. Its conceptual importance lies in its simplicity: rather than characterizing a distribution through a finite collection of parameters, it treats the distribution function itself as the estimand.

The theoretical legitimacy of this perspective was established by the Glivenko–Cantelli theorem, proved independently by Glivenko (1933) and Cantelli (1933). The theorem showed that the empirical distribution function converges uniformly to the true distribution function. This was a decisive result in the history of nonparametric statistics because it established that an entire function could be consistently estimated without parametric specification. The later inequality of Dvoretzky, Kiefer and Wolfowitz (1956) sharpened this insight by providing explicit probability bounds on the maximal deviation between the empirical and true distribution functions.

In retrospect, these results mattered to econometrics because they helped normalize the idea that infinite-dimensional objects could be estimated and studied rigorously. They did not arise from econometric questions, and they did not solve specifically econometric problems. But they formed part of the broader theoretical environment from which nonparametric econometrics eventually drew; see Pagan and Ullah (1999). Later work in empirical process theory would build on precisely this foundation, and that literature would become increasingly important once econometricians began to ask about the asymptotic behavior of estimators defined through functions rather than finite-dimensional parameter vectors.

11.2.4 Optimal Rates and the Curse of Dimensionality

By the 1970s and early 1980s, the nonparametric statistics literature had moved beyond existence and consistency questions to confront a more demanding issue: how well can unknown functions actually be estimated from finite samples? This was a critical step in the pre-history of nonparametric econometrics because it turned attention from conceptual possibility to statistical feasibility. The answer, as it turned out, was sobering.

The central contributions here were those of Stone (1980, 1982), who derived optimal convergence rates for nonparametric regression estimators under smoothness assumptions. Their results made precise a fact that had already been sensed informally: estimating an unknown function is statistically much harder than estimating a finite-dimensional parameter vector. Whereas parametric estimators often converge at the familiar $n^{-1/2}$ rate, nonparametric rates depend on both smoothness and dimension, and they deteriorate rapidly as the dimension of the covariate space increases. It was this body of work that gave formal expression to what became known as the *curse* *of dimensionality*; see also the later econometric discussions in Härdle and Linton (1994) and Pagan and Ullah (1999).

The historical importance of Stone (1980, 1982) is difficult to overstate. The curse of dimensionality was not simply an inconvenience or a technical footnote. It supplied one of the most powerful reasons for econometric skepticism toward nonparametric methods. Many empirical economic settings involve multiple regressors, interaction effects, and heterogeneous responses, exactly the types of environments in which unconstrained nonparametric estimation becomes statistically demanding. For economists interested in applied work rather than purely asymptotic possibility, the curse of dimensionality helped define the limits of what nonparametric methods could plausibly deliver.

Seen in historical sequence, this was a turning point in emphasis. The early statistical literature had shown that flexible function estimation was possible; the optimal-rate literature showed just how costly that flexibility could be. Later econometric developments—including semiparametric models, additive structures, and dimension-reduction strategies—can be understood in part as responses to this realization.

11.2.5 Rank-Based and Distribution-Free Inference

A different branch of early nonparametric statistics developed not around function estimation but around hypothesis testing and inference procedures that avoided strong distributional assumptions. The central idea in this literature was that useful inference could often be based on the ordering of observations rather than on their numerical magnitudes. By relying on ranks, such methods achieved validity across broad classes of distributions and thereby offered an alternative to Gaussian-based parametric procedures; for early and synthetic treatments see Hoeffding (1948), Hájek and Šidák (1967), Hettmansperger (1984), and Randles and Wolfe (1979).

Historically, this literature predates much of the smoothing tradition that later became most closely associated with nonparametric econometrics. Early landmarks include Spearman (1904) on rank correlation, Wilcoxon (1945) on rank-based testing, and Mann and Whitney (1947) on two-sample distribution-free inference. The theoretical status of such procedures was strengthened by Hoeffding (1948) and later synthesized in the monographs of Hájek and Šidák (1967), Hettmansperger (1984), and Randles and Wolfe (1979), which helped establish rank-based methods as a rigorous inferential framework rather than as a loose collection of ad hoc robust tests.

This strand of the nonparametric tradition is worth recalling in a history of econometrics even though it was not the main route through which nonparametric regression entered economics. Its importance is broader and more intellectual. Rank-based procedures helped legitimize the idea that valid inference need not depend on full parametric specification, thereby contributing to the wider statistical environment in which nonparametric thinking became respectable. At the same time, their relative distance from structural economic modeling helps explain why they did not initially become central to econometric practice.

11.2.6 Computational Constraints and Early Feasibility

The history of nonparametric methods cannot be told solely as a sequence of theoretical advances. It is also a history of feasibility. Many early nonparametric procedures were computationally burdensome relative to the hardware and software environments in which mid-century empirical work was conducted. Kernel smoothing, repeated local averaging, iterative bandwidth choice, and resampling-based inference could be mathematically elegant while remaining difficult to implement routinely in applied work. Later overviews of the field repeatedly note that advances in computation were an important part of the transition from theoretical possibility to empirical usability; see, for example, Härdle (1990), Härdle and Linton (1994), Li and Racine (2007), Pagan and Ullah (1999), and J. S. Racine, Su and Ullah (2014).

What is notable historically is that this concern was not voiced only by skeptics outside the field. Even leading contributors to nonparametric methodology acknowledged the practical burden of implementation. Discussing bootstrap-based procedures, Härdle (1990) noted that the relevant algorithm "requires quite a bit of computer intensive resampling," adding that "this computational burden" could only be mitigated through further numerical devices. By the late 1990s, the same concern could be stated in even broader terms. Surveying the use of nonparametric regression in economics, Yatchew (1998) remarked that "nonparametric regression techniques are computationally intensive and they require large (in some cases astronomically large) data sets." These observations captured a central feature of the field's early development: nonparametric methods were often judged not only by their asymptotic properties, but also by whether they were practically feasible with the data and computing resources available at the time.

This practical constraint mattered greatly for the eventual econometric uptake of nonparametric methods. In principle, a statistical method may be available long before it is empirically usable on the kinds of datasets and computing environments typical of the profession. In practice, many of the procedures developed in the statistical literature remained closer to theoretical demonstrations or small-scale illustrations than to standard applied tools. This gap between theoretical availability and routine feasibility helps explain why nonparametric ideas diffused into econometrics more slowly than their formal elegance might suggest.

The situation changed gradually as computing power expanded in the 1970s and 1980s and as numerical methods and software environments improved. These developments did not by themselves guarantee econometric adoption, but they altered the practical conditions under which economists evaluated flexible estimators. Once implementation became less prohibitive, the substantive objections to nonparametric methods became clearer: the debate could shift from whether such methods were computationally possible to whether they were econometrically useful, interpretable, and identifiable. In that sense, computational change did not end the controversy surrounding nonparametrics; it made that controversy more genuinely econometric.

11.3 The First Wave in Econometrics (1960s–1970s)

The first encounters between nonparametric methods and econometrics took place in a discipline still strongly shaped by parametric modeling, structural interpretation, and relatively limited computational capacity. By the 1960s and 1970s, economists had at their disposal a growing statistical literature on smoothing, density estimation, and local averaging, yet the econometric meaning of these methods remained unsettled. The central question was not simply whether unknown functions could be estimated flexibly, but whether such flexibility could be reconciled with the profession's prevailing standards of interpretability, identification, and theoretical discipline. Later surveys make clear that the early econometric reception of nonparametric methods was cautious rather than revolutionary; see, among others, Ullah (1988), Härdle and Linton (1994), and Pagan and Ullah (1999).

What entered econometrics in this first wave was therefore not a fully formed research program, but a set of statistical tools whose relevance to economic modeling had yet to be established. Some researchers saw in them a promising way to relax restrictive functional-form assumptions and to diagnose misspecification. Others regarded them as insufficiently tied to theory, difficult to interpret, and impractical in the kinds of data environments common in applied economics. The history of this period is thus best understood as one of tentative entry, methodological skepticism, and gradual clarification of what nonparametric methods could reasonably contribute to econometric practice.

11.3.1 Nonparametric Regression as a Challenge to Linear Models

The first important point of contact between nonparametric methods and econometrics came through regression analysis. For much of the discipline's earlier development, empirical work had relied heavily on linear regression models, often supplemented by transformations of regressors chosen for convenience, tractability, or rough theoretical plausibility. Such models were attractive precisely because they provided a compact and interpretable representation of economic relationships. Yet they also imposed strong functional-form restrictions, and by the 1970s economists had become increasingly aware that these restrictions could themselves shape substantive conclusions; see the retrospective discussions in Ullah (1988), Härdle and Linton (1994), and Pagan and Ullah (1999).

The statistical foundations for flexible regression estimation had already been laid by Nadaraya (1964) and Watson (1964), but their econometric significance became clearer only later, as researchers began to treat the regression function itself as an object of interest rather than merely as a vehicle for estimating a finite-dimensional coefficient vector. A key contribution in this transition was Stone (1977), which provided rigorous consistency results for nonparametric regression estimators and helped establish that regression functions could, under suitable conditions, be estimated directly from the data without prespecifying a global form. Among the earlier contributions that brought these ideas more explicitly into econometrics was Bierens (1983), who established uniform consistency of kernel regression estimators under conditions geared toward econometric applications. These contributions mattered not because they displaced linear regression but because they made visible an alternative conception of what regression analysis could be.

For econometricians, this possibility posed both an opportunity and a challenge. On the one hand, nonparametric regression offered a way to detect nonlinearities, interactions, and local features that might be obscured by conventional linear specifications. Used cautiously, it could serve as an exploratory device for asking whether familiar parametric models were imposing too much structure on the data. On the other hand, a flexible estimated curve did not carry the same immediate interpretive content as a finite vector of coefficients, nor did it align as easily with theory-driven modeling. What statisticians viewed as a strength—minimal functional-form commitment—many econometricians saw as a possible weakness.

This helps explain why the earliest econometric uses of nonparametric regression were often framed not as substitutes for parametric models but as complements to them. Flexible estimators were appealing precisely because they could reveal what a parametric model might be missing, yet they were rarely presented as complete replacements for theory-guided specification. That diagnostic and supplementary role would prove historically important, since it provided one of the first ways in which nonparametric methods could enter econometrics without immediately confronting the profession's strongest objections; see also Ullah (1988).

11.3.2 Specification Testing and Model Diagnostics

One of the earliest and most natural roles for nonparametric methods in econometrics was in the diagnosis of parametric misspecification. Rather than replacing conventional regression models outright, early researchers used flexible estimators as benchmarks against which parametric specifications could be judged. The intuition was straightforward: if a parametric model provided an adequate description of the conditional relationship between variables, then its fitted values should not differ systematically from those produced by a sufficiently flexible nonparametric alternative. This diagnostic logic would later become a central theme in nonparametric econometrics; see Ullah (1988), Härdle and Linton (1994), and Pagan and Ullah (1999).

This perspective also fit naturally within pre-existing econometric concerns about functional-form misspecification. Classical procedures such as the RESET test of Ramsey (1969) had already emphasized that parametric regression models could fail not only because of omitted variables or stochastic assumptions, but because the chosen functional form was itself inadequate. Nonparametric methods extended this logic by providing a flexible comparison object under relatively weak structural commitments. As Pagan and Ullah (1999) note, for example, joint normality of the dependent and independent variables is sufficient to imply linearity of the conditional expectation, so observed departures from linearity can be interpreted as departures from that underlying distributional structure.

Formal specification tests using nonparametric ideas began to appear more clearly in the econometric literature in the 1980s. Bierens (1982) developed a consistent specification test for regression models that did not depend on a specific parametric alternative, and Bierens (1990) extended this line of work through tests based on conditional moment restrictions capable of detecting broad classes of functional-form failure. These contributions mattered historically because they showed how nonparametric reasoning could be integrated into econometric practice without asking researchers to abandon parametric models altogether. They also helped establish specification testing as one of the earliest econometrically acceptable uses of nonparametric methods; see also Ullah (1988).

The broader significance of this development was pragmatic. Nonparametric methods first entered econometrics not primarily as stand-alone estimators of economic structure, but as tools for checking whether parametric assumptions were credible. In this form they were easier for the profession to absorb. They helped econometricians relax the rigid opposition between parametric and nonparametric methods by demonstrating that flexibility could be useful even within a fundamentally parametric workflow.

11.3.3 Early Skepticism Among Econometricians

The early reception of nonparametric methods in econometrics was marked by substantial skepticism. That skepticism was not merely inertia or conservatism, though it certainly reflected established disciplinary habits. More fundamentally, it arose because nonparametric methods seemed to challenge several of the profession's deepest methodological commitments. Econometrics had long been organized around the idea that empirical specification should be guided by economic theory, and that interpretable parameters should be linked to explicit economic mechanisms. In that setting, procedures imported from mathematical statistics and justified primarily by smoothness conditions could appear alien to the central aims of econometric analysis; see the retrospective accounts in Ullah (1988), Härdle and Linton (1994), and Pagan and Ullah (1999).

This tension is easiest to understand against the background of the Cowles Commission tradition and the broader emphasis on theory-based measurement. Koopmans (1947) famously criticized "measurement without theory," and later reflections by Christ (1994) make clear how strongly Cowles-era econometrics presupposed that the functional form of structural equations was known from theory or imposed as part of the modeling exercise. As Christ (1994) puts it, questions about the functional form of the equations themselves were largely not part of the Cowles agenda. Against that backdrop, nonparametric methods could look like a retreat from structure rather than a refinement of empirical practice.

A central objection was that nonparametric procedures did not eliminate assumptions so much as relocate them. Proponents stressed the relaxation of rigid functional-form restrictions, but critics noted that smoothness conditions, bandwidth choice, and kernel selection were themselves substantive ingredients in the analysis. Unlike the coefficients of a linear or structural model, however, these ingredients were not naturally interpretable in economic terms. To some economists, this made nonparametric methods appear less transparent rather than more so. Similar concerns about the fragility of empirical conclusions under alternative modeling choices were articulated more broadly in econometrics by Leamer (1983).

A second source of skepticism was feasibility. As discussed in the previous section, the curse of dimensionality and the computational demands of smoothing methods meant that unconstrained nonparametric estimation often looked impractical in the kinds of multivariate settings economists actually studied. Econometric models typically involved multiple regressors, possible endogeneity, unobserved heterogeneity, and modest sample sizes. Under those conditions, flexible estimation could appear theoretically elegant but empirically fragile. The later formalization of these issues by Stone (1980, 1982) sharpened concerns that had already been present in intuitive form, while Yatchew (1998) would later summarize the issue bluntly by noting that nonparametric regression techniques were computationally intensive and could require, in some settings, “astronomically large” data sets.

There was also unease about inference and communication. Early nonparametric estimators were associated with slower convergence rates, nonstandard asymptotic theory, and a heavy dependence on tuning parameters. By contrast, parametric models delivered familiar test statistics, readily reportable coefficient estimates, and a vocabulary that aligned naturally with prevailing empirical standards. From this perspective, nonparametric methods risked producing results that were statistically sophisticated but difficult to summarize economically. These concerns later helped motivate semiparametric approaches that preserved finite-dimensional parameters of direct interest while allowing selected model components to remain flexible; see, for example, Powell (1994).

A further line of criticism involved identification. Even when nonparametric procedures produced statistically precise estimates, it was not always clear that the underlying economic object was point identified without functional-form restrictions. Econometricians had long been attentive to identification as a foundational issue, beginning with classical work such as Hurwicz (1950). From that vantage point, greater statistical flexibility could obscure rather than solve the identification problem. Later developments in partial identification and set-valued inference, especially those associated with Manski (1995), can be read in part as a more explicit response to concerns that were already latent in this early reception period.

Finally, skepticism reflected broader disciplinary boundaries. Econometrics in the 1960s and 1970s still placed heavy weight on structural systems, simultaneous equations, and explicit economic mechanisms. Nonparametric methods, by contrast, often appeared to offer statistically refined descriptions without comparable behavioral content. As a result, many of the earliest contributions were received more readily in statistics than in economics. In retrospect, however, this skeptical reception was productive. It forced the nonparametric literature, once it became more fully econometric, to confront problems of identification, inference, and economic interpretation directly rather than treating them as secondary concerns.

11.3.4 Bandwidth Selection as a Conceptual Obstacle

One of the most persistent obstacles to the early acceptance of nonparametric methods was the problem of bandwidth selection. Kernel estimators require the researcher to choose a smoothing parameter that determines the size of the neighborhood over which local averaging is performed. A small bandwidth produces a curve that follows the data closely but may be highly variable, while a large bandwidth yields a smoother estimate at the cost of increased bias. This bias–variance tradeoff is fundamental to nonparametric estimation and quickly became one of the defining practical and conceptual issues in the field; see the discussions in Härdle (1990), Härdle and Linton (1994), Pagan and Ullah (1999), and Fan and Gijbels (1996).

For many econometricians encountering these methods in the 1960s and 1970s, bandwidth choice seemed to introduce precisely the sort of arbitrariness that parametric modeling was supposed to avoid. Once a parametric form had been chosen, the estimator itself was usually well defined. Nonparametric methods, by contrast, appeared to require a second layer of judgment, and one whose empirical consequences could be substantial. Critics therefore argued that bandwidth selection merely replaced explicit functional-form assumptions with implicit tuning decisions that were difficult to defend on economic grounds. This concern was not trivial. Historically, it was one of the clearest examples of how nonparametric flexibility could be perceived as coming at the cost of transparency.

The response to this objection was itself historically important, because it generated a large literature devoted to principled selection methods. Stone (1974) proposed early cross-validation procedures for choosing the smoothing parameter by minimizing estimated prediction error. Later work by Rudemo (1982) and Bowman (1984) extended related ideas to kernel density estimation, and Silverman (1986) provided an influential synthesis that helped standardize implementation and made practical bandwidth choice appear less ad hoc than it had initially seemed. These developments did not eliminate disagreement, but they changed the terms of the discussion by showing that the choice of smoothing parameter could be disciplined statistically rather than left entirely to subjective judgment.

A parallel literature developed around direct plug-in methods, in which pilot estimates of unknown functionals were substituted into asymptotic MISE-optimal formulas. Advocates of these methods argued that, particularly in density estimation, they avoided the tendency of cross-validation to undersmooth and exhibited lower sampling variability. An influential contribution in this tradition was Sheather and Jones (1991). The resulting debate between cross-validation and plug-in methods continued well beyond the first wave of econometric uptake. In economics, however, cross-validation became especially prominent because of its natural extension to regression problems and, later, to settings involving mixed discrete and continuous regressors. In this sense, the bandwidth problem did not simply impede the spread of nonparametric methods; it also helped generate one of the literatures that made their subsequent adoption more credible.

11.4 Identification Without Functional Forms

Nonparametric identification marked a major conceptual turning point in econometrics. Earlier debates over smoothing, flexibility, and computational feasibility had already raised doubts about how useful nonparametric methods could be in practice. But the identification literature shifted the argument to a deeper level. The issue was no longer merely whether one could estimate a regression function without imposing a parametric form; it was whether the underlying economic object of interest was even recoverable once functional-form assumptions were withdrawn. In this sense, nonparametric identification forced econometricians to confront a distinction that had often remained implicit in earlier work: flexibility in estimation does not by itself guarantee clarity in interpretation. By the mid-1990s, more general semiparametric treatments were already emphasizing the centrality of identifying restrictions under weak stochastic assumptions; see Powell (1994). Later surveys make clear that the nonparametric identification literature helped transform flexible econometric methods from a collection of estimation tools into a more mature field organized around explicit questions of what data and assumptions could jointly deliver; see, among others, Matzkin (2007), Härdle and Linton (1994), and Pagan and Ullah (1999).

Historically, this development mattered because it reoriented the field away from a simple contrast between parametric and nonparametric methods. Once identification became central, the relevant question was not whether one was willing to relax functional-form assumptions, but what took their place. Support conditions, completeness, monotonicity, separability, exclusion, and other structural restrictions became central not because they restored parametric structure, but because they determined whether the objects of economic interest could be learned from the observable distribution at all. The history of nonparametric identification is therefore also a history of econometrics becoming more explicit about the informational content of assumptions.

11.4.1 What Identification Means in a Nonparametric World

Identification has long occupied a central place in econometric theory. In the classical parametric setting, it refers to whether a finite-dimensional parameter vector can be uniquely recovered from the probability distribution of the observable data. Early formal treatments, such as Hurwicz (1950), were developed in the context of simultaneous equations and structural models, where the central concern was whether economic restrictions were sufficient to distinguish one parameter vector from another. In that setting, identification was largely understood through algebraic restrictions, normalization conditions, and exclusion restrictions tied to structural interpretation.

The emergence of nonparametric econometrics required this concept to be reformulated. Once the object of interest becomes an unknown function rather than a finite-dimensional vector, identification can no longer be understood simply as the uniqueness of a point in parameter space. Instead, the question becomes whether that function can be uniquely recovered from the joint distribution of the observables. In this environment, identification depends not only on economic restrictions, but also on properties of the observable distribution itself, including support conditions, independence assumptions, and the behavior of conditional expectation operators. Early formal discussions of this shift appear in Roehrig (1988), while Matzkin (2007) later emphasized how nonparametric identification moves attention away from algebraic solvability and toward the informational content of the data-generating process.

This was more than a technical change in vocabulary. It had direct implications for empirical research. In parametric models, identification can often be obtained through exclusion restrictions together with functional-form assumptions that sharply limit the class of admissible models. In nonparametric settings, by contrast, such assumptions are weakened or removed, and the burden shifts toward richer variation in the data and more subtle structural restrictions. As a result, many economic models that appear point identified in parametric form become only weakly identified, set identified, or unidentified once the same functional-form commitments are abandoned. Early work by Matzkin (1993) made this point concrete in the context of polychotomous choice models, showing that identification in flexible structural settings could still be pursued, but only by paying careful attention to the economic and probabilistic structure that remained. Historically, this was one of the moments at which nonparametric econometrics ceased to be merely a flexible estimation strategy and became a challenge to the way econometricians understood what could be learned from data.

11.4.2 Completeness, Support Conditions, and Ill-Posedness

A central lesson of the nonparametric identification literature is that recovering unknown functions from observable data often requires strong conditions on the distribution of the underlying variables. In many settings, identification depends on whether the regressors or instruments vary sufficiently over their support and on whether conditional distributions are rich enough to pin down the unknown function uniquely. These requirements differ in character from the exclusion restrictions familiar from parametric econometrics. Rather than simply ruling out certain direct effects, they concern the informational adequacy of the observable environment itself.

For later syntheses of this point, see Matzkin (2007), Newey and Powell (2003), and Carrasco, Florens and Renault (2007).

Among the most important ideas to emerge in this literature was completeness. In nonparametric models based on conditional moment restrictions, completeness ensures that the relevant conditional expectation operator is injective, so that distinct structural functions generate distinct observable implications. In this way, completeness serves as a functional analogue of uniqueness conditions in more classical identification arguments. The role of completeness was brought into particularly sharp focus in the nonparametric instrumental variables literature, especially in Newey and Powell (2003), and later developed further by Carrasco et al. (2007) and Darolles, Fan, Florens and Renault (2011). What made this line of work historically significant was that it revealed how identification in flexible models depended on features of the conditional distribution that had no close analogue in simpler parametric settings.

A closely related development was the recognition that many identification arguments in nonparametric econometrics could be expressed as inverse problems. Florens (2003) was especially important in making this perspective explicit in econometrics, showing in the instrumental variables setting how the recovery of a structural function from observable moments could be framed as the inversion of an operator. This formulation helped clarify why identification and estimation could not be neatly separated: the same operator that governs uniqueness of the solution also governs the stability of its recovery from sample information.

Even when identification could be established in principle, therefore, a second problem immediately emerged: estimation could still be unstable in practice. Many nonparametric identification problems take the form of inverse problems in which the mapping from the unknown structural function to the observables does not admit a continuous inverse. In such cases, small disturbances in the sample analogue of the operator can produce large movements in the recovered function. The resulting ill-posedness meant that uniqueness alone was not enough. A model could be point identified in theory and yet remain extremely difficult to estimate reliably in finite samples. This insight helped move the literature beyond a narrow notion of identification and toward a broader understanding of the interplay between identification, regularization, and inference.

Historically, this was a decisive clarification. Earlier discussions of nonparametric methods had sometimes blurred the line between flexible estimation and recoverability of the underlying economic object. The completeness and ill-posedness literature made clear that these were distinct issues. A unique structural function might exist, yet the observable problem could still be statistically fragile. In that sense, nonparametric identification sharpened rather than resolved the profession's earlier concerns about ambiguity. It also opened the door to adjacent developments in regularization, shape restrictions, and partial identification.

11.4.3 Nonparametric versus Parametric Identification

The contrast between parametric and nonparametric identification provides one of the clearest windows into the role played by functional-form assumptions in econometric analysis. Parametric models often achieve point identification through comparatively simple restrictions on finite-dimensional parameters. Exclusion restrictions, normalization conditions, and maintained distributional assumptions can together produce unique parameter values even when the available variation in the data is limited. But such identification may rest heavily on the maintained structure of the model itself.

Removing functional-form assumptions changes this balance. Once the mapping between observables and the underlying economic structure is permitted to vary over a much broader class of functions, uniqueness can no longer be taken for granted. Models that seemed well identified in parametric form may become only weakly identified or fail to be point identified altogether in a nonparametric framework. This was one of the most important historical lessons of the identification literature: some econometric certainty had been purchased, often implicitly, through functional-form assumptions whose identifying role had not always been made fully explicit.

At the same time, the nonparametric literature also showed that parametric and nonparametric identification were not simply opposites. Point identification without functional-form restrictions could sometimes be restored through other kinds of structure, including monotonicity, separability, independence, or triangular relationships among observables. As Matzkin (2007) documents, many models became identifiable not because they reintroduced conventional parametric forms, but because they imposed weaker, more transparent structural restrictions tailored to the economics of the problem. Earlier work such as Matzkin (1991) illustrated this clearly by showing how monotonicity and concavity restrictions on utility could support semiparametric recovery in polychotomous choice settings, while Chesher (2003) showed how triangular systems could deliver nonparametric identification in models with endogeneity.

Historically, this tension proved productive. It did not lead to the abandonment of parametric thinking, nor did it establish nonparametric identification as a universal substitute. Instead, it forced econometricians to ask more clearly which conclusions depended on functional-form assumptions and which survived under weaker forms of structure. In doing so, it helped reshape the methodological discussion around transparency of assumptions rather than around a crude opposition between rigid parametrics and unrestricted flexibility.

11.4.4 Partial Identification and Set Identification

One of the most important consequences of the nonparametric identification literature was the recognition that point identification is not the only intellectually coherent outcome of an econometric model. Once strong functional-form assumptions are relaxed, the data may be consistent with a range of admissible values rather than with a unique point. In such cases, the appropriate object of analysis is not a parameter estimate in the usual sense, but an identified set. This insight, later developed systematically in the partial identification literature, marked an important widening of econometric perspective.

In historical terms, partial identification can be read as a natural outgrowth of the nonparametric critique of overconfident point identification. If parametric assumptions had often supplied identifying power more quietly than economists acknowledged, then removing those assumptions would not always leave behind a fully determined object. The work of Manski (1990) and, more broadly, Manski (2003) made this implication explicit by treating the limits of what can be learned from data as a first-order question rather than as a failure of technique. In this framework, the econometrician's task becomes one of characterizing the set of values consistent with the data and maintained assumptions, rather than forcing a spurious precision onto an underidentified problem.

This development was especially consequential for nonparametric econometrics because it clarified that ambiguity need not be regarded as methodological defeat. On the contrary, set identification could be the honest expression of what the available information warranted. Formal work by Horowitz and Manski (2000) and Tamer (2003), among others, extended this perspective to economically important classes of models in which the data generating process did not deliver unique outcomes or unique parameter values under weak assumptions. In that sense, partial identification did not simply emerge alongside nonparametric econometrics; it represented one of the most intellectually important ways in which the field forced econometricians to confront the informational limits of their models.

11.4.5 Treatment Effects as an Identification Laboratory

One of the most important applied arenas in which nonparametric identification issues became visible was the treatment effects literature. Program evaluation and causal inference created strong demand for methods that could recover policy-relevant parameters without relying heavily on parametric outcome equations. In that sense, treatment effect analysis offered a natural setting for nonparametric and semiparametric ideas. Causal parameters such as average treatment effects could be defined as functionals of potential outcome distributions rather than as coefficients in a parametric regression model, and this opened the door to flexible estimators based on matching, reweighting, and related methods.

At the same time, this literature made especially clear that relaxing functional-form assumptions does not by itself solve the identification problem. In the potential outcomes framework associated with Rubin (1974), and building on the earlier setup (first published in 1923) of Neyman (1990), identification depends on assumptions about the assignment mechanism rather than on a parametric specification of the outcome equation. The importance of this distinction was sharpened by Rosenbaum and Rubin (1983), who showed that under conditional independence, treatment assignment is independent of potential outcomes conditional on the propensity score. This result was historically important not only because it simplified implementation, but because it illustrated a recurring theme in nonparametric econometrics: flexibility is often made feasible by some form of dimension reduction or auxiliary structure.

Subsequent work on matching, weighting, and semiparametric efficiency in treatment effect models reinforced this lesson. Contributions such as Heckman, Ichimura and Todd (1997) and Hirano, Imbens and Ridder (2003) showed that one could weaken functional-form assumptions substantially while still obtaining economically meaningful causal estimates. But they also made plain that the credibility of those estimates rested on assumptions such as unconfoundedness and common support, not on nonparametric estimation alone. In this respect, the treatment effects literature served as a kind of identification laboratory for nonparametric econometrics more broadly: it demonstrated with unusual clarity that flexibility can reduce specification risk, but cannot substitute for credible design or defensible identifying assumptions.

11.4.6 Informational Content of Economic Structure

The difficulties encountered in nonparametric identification ultimately redirected attention to the informational role of economic theory itself. If flexible statistical methods made clear that data alone often do not pin down the objects of interest, the natural follow-up question was what additional identifying power economic structure might supply. This was a crucial historical development because it helped move the field beyond the simplistic idea that nonparametrics meant freedom from assumptions. The real issue was never whether assumptions could be avoided, but which assumptions were economically meaningful, empirically defensible, and transparent in their identifying role.

As Heckman (2001) argued, economic theory contributes more than functional-form restrictions narrowly understood. It can also provide information about behavioral mechanisms, institutional environments, equilibrium restrictions, or selection processes that help interpret and identify empirical relationships. Earlier work by Matzkin (1994) had already made this point directly in the nonparametric context by emphasizing that restrictions drawn from economic theory could be incorporated into flexible methods without collapsing them back into conventional parametric models. In this view, theory is not opposed to nonparametric reasoning; it is one source of identifying information among others.

Once framed this way, the relationship between nonparametric and parametric methods looks less like a contest between flexibility and structure and more like a search for the right combination of statistical openness and economically grounded restriction. This interaction between statistical flexibility and economic structure has become one of the enduring legacies of the nonparametric identification literature. Rather than encouraging a wholesale rejection of structured econometric modeling, it helped reveal which aspects of structure were doing real identifying work and which had been maintained mainly for convenience. The most productive developments in later nonparametric econometrics would often follow precisely this path: retaining enough economic discipline to identify interpretable objects, while using flexible methods to avoid unnecessarily restrictive functional forms.

11.5 Testing, Confidence Sets, and Inference

One of the central challenges in the history of nonparametric econometrics was not merely how to estimate unknown functions, but how to conduct valid inference about them. Flexible estimation could reveal nonlinearities, heterogeneity, and other features that parametric models suppressed, but unless those estimates could be accompanied by trustworthy tests and confidence statements, nonparametric methods would remain methodologically incomplete. In this respect, the inferential literature was indispensable to the maturation of the field. It demonstrated that nonparametric econometrics was not just a collection of smoothers and exploratory devices, but a discipline capable of rigorous statistical evaluation under weak structural assumptions. For broader retrospective discussions of this transition from flexible estimation to full econometric methodology, see Pagan and Ullah (1999).

Historically, the difficulty arose because the objects of interest in nonparametric econometrics are typically infinite-dimensional. The distributional approximations familiar from parametric econometrics do not carry over automatically when the target is a function rather than a finite-dimensional vector, when convergence is slower than root-n, or when the relevant statistics depend on the supremum of an empirical process rather than on a fixed set of coordinates. These complications forced the literature to develop new tools: empirical process methods, nonstandard asymptotics, bootstrap procedures adapted to smoothing problems, and minimax criteria for evaluating performance over classes of functions rather than at isolated points. Together these developments supplied the inferential foundation the field needed in order to become a coherent statistical framework.

11.5.1 Hypothesis Testing in Infinite-Dimensional Spaces

One of the earliest inferential lessons of nonparametric statistics, later inherited by econometrics, was that hypothesis testing in infinite-dimensional settings often leads to limiting behavior quite unlike that of familiar parametric test statistics. In parametric models, likelihood ratio, Wald, and score statistics typically converge to normal or chi-squared limits under standard regularity conditions. In nonparametric settings, by contrast, the target of inference is often an entire function, and test statistics are frequently based on supremum norms, integrated squared deviations, or empirical process functionals whose limiting laws are nonstandard.

A foundational contribution here was Bickel and Rosenblatt (1973), who studied goodness-of-fit tests for kernel density estimators and showed that supremum deviations could converge to extreme-value limits rather than to Gaussian ones. Historically, this was important because it revealed at an early stage that nonparametric inference would require its own asymptotic machinery rather than a straightforward extension of parametric testing principles.

These ideas later entered econometrics more directly through the literature on specification testing. Bierens (1990) developed consistent specification tests based on integrated conditional moment restrictions, providing procedures with power against broad classes of alternatives rather than only against fixed parametric departures. Härdle and Mammen (1993) proposed tests comparing a parametric fitted curve with a nonparametric alternative through weighted integrated squared differences, using bootstrap methods to obtain critical values. Historically, these contributions mattered because they helped shift nonparametric econometrics from a merely descriptive role toward one in which it could also discipline parametric modeling through formal testing.

11.5.2 Uniform versus Pointwise Inference

A fundamental distinction in nonparametric inference concerns whether statistical statements apply at a single point or uniformly over an entire domain. Pointwise confidence intervals are often relatively easy to derive: one fixes a covariate value and studies the distribution of the estimator there. But such intervals do not answer the question that many empirical researchers actually care about, namely whether an estimated function is uniformly close to the truth over a range of economically relevant values. Once the target becomes the function as a whole, pointwise inference is no longer sufficient.

The historical importance of this distinction lies in the fact that it made clear how much stronger function-level inference really is. Uniform confidence bands require control of the supremum of a stochastic process, not merely its marginal behavior at a point. This brought empirical process theory into the center of nonparametric inference; see, for example, van der Vaart and Wellner (1996). Hall (1992) developed bootstrap-based uniform confidence bands for kernel density estimators, providing one of the clearest demonstrations that the inferential demands of nonparametric econometrics were inseparable from its functional targets.

More broadly, the pointwise-versus-uniform distinction helped shape the way the field thought about uncertainty. It showed that inferential statements about functions are inherently global objects and that the width of valid confidence bands reflects not just variance at a point, but the complexity of controlling an entire random function over its support. In this way, uniform inference became one of the places where the ambition of nonparametric econometrics to estimate function-valued objects met the full force of its statistical consequences.

11.5.3 Bootstrap Methods and Their Limitations

Bootstrap methods became especially important in nonparametric econometrics because many nonparametric estimators have asymptotic distributions that are difficult to derive analytically and awkward to approximate in finite samples. Resampling offered a practical alternative, and for many applied researchers it became one of the main ways nonparametric inference became operational. Yet the bootstrap also forced the field to confront another historical lesson: procedures that work well in smooth parametric settings cannot simply be assumed valid once smoothing, regularization, and nonstandard rates of convergence enter the problem.

The theoretical foundations of bootstrap methods in econometrics were synthesized in sources such as Hall (1994), Horowitz (2001), and Horowitz (2003), all of which treated the nonparametric and semiparametric case as requiring special attention. A key insight from this literature is that bootstrap validity may depend sensitively on the smoothing structure of the estimator, on the bandwidth sequence, and on the way the resampling scheme reproduces the relevant bias and variance terms. In nonparametric settings, therefore, the bootstrap is not merely a computational convenience. It becomes part of the inferential design.

This also meant that the bootstrap had visible limitations. In some cases, standard bootstrap procedures fail because the estimator is not a sufficiently smooth function of the empirical distribution. In such settings, alternatives such as subsampling or the m-out-of-n bootstrap provide valid inference. A distinct failure mode arises when smoothing parameters induce bias distortions that naive resampling does not reproduce correctly; this problem is typically addressed through undersmoothing, explicit bias correction, or specialized bootstrap procedures. Sperlich (2014), for example, emphasized how the choice of regularization and bandwidth parameters in specification testing can materially alter the behavior of bootstrap procedures, with direct consequences for empirical size and power. More recent work such as Hall and Horowitz (2013) further showed how bootstrap methods could be adapted to the construction of nonparametric confidence bands in the presence of bias and smoothing complications. Historically, this literature mattered because it made clear that nonparametric inference could not simply borrow computational tools from parametric econometrics without rethinking their validity from first principles.

11.5.4 Honest Confidence Sets

One of the deepest insights in nonparametric inference concerned the construction of confidence sets with reliable coverage properties. In finite-dimensional settings, it is often possible to pursue procedures that adapt automatically to the unknown complexity of the problem while still maintaining nominal coverage. In nonparametric settings, however, adaptation turns out to be much more limited. Confidence procedures that exploit unknown smoothness to become narrower in favorable cases cannot, in general, maintain correct coverage uniformly over large classes of functions.

This impossibility was made especially clear by Low (1997), who showed that adaptive confidence intervals cannot simultaneously achieve nominal coverage and full adaptation over broad nonparametric function classes. Historically, this was a decisive result because it made explicit one of the hidden costs of flexibility: reliable inference requires the econometrician to be transparent about the function class over which coverage is sought. In other words, honest confidence sets are possible, but only relative to an explicitly maintained smoothness or complexity class.

The subsequent literature turned from impossibility to construction. Recent work such as Armstrong and Kolesár (2020) developed simple and honest confidence intervals for nonparametric regression, showing how valid coverage can be achieved once the smoothness class is made explicit and the bias-variance tradeoff is handled directly. This line of work is historically important because it illustrates the field's broader response to Low's negative result: rather than abandoning inference, researchers learned to make the underlying smoothness commitments transparent and to build confidence procedures that are honest with respect to those commitments.

The importance of this literature extends beyond any single technical result. It forced the field to abandon the hope that one could always have both adaptation and uniform validity at no cost. More broadly, it reinforced a recurring theme in the history of nonparametric econometrics: whenever one weakens parametric assumptions, the resulting inferential guarantees must be tied more explicitly to the class of admissible objects under study.

11.5.5 Minimax Perspectives

A further step in the inferential maturation of nonparametric econometrics came through minimax reasoning. Rather than evaluating an estimator under a single data-generating process, the minimax framework assesses its worst-case performance over a class of admissible functions. This was historically important because it provided a principled way to think about the fundamental limits of estimation and inference in nonparametric problems. If parametric econometrics often took root-n consistency as a natural benchmark, minimax theory showed that no such benchmark could be universal once the object of interest was an unknown function.

The foundational results of Stone (1980, 1982) established optimal convergence rates for nonparametric regression and density estimation under smoothness restrictions, thereby giving precise form to the curse of dimensionality and to the bias–variance tradeoff that governs flexible estimation. Later work, including Donoho (1994), extended minimax ideas to broader inferential settings and clarified how the attainable performance of an estimator depends on the geometry of the underlying function class. Historically, these contributions were crucial because they gave the literature a set of sharp performance benchmarks against which new methods could be judged.

In econometrics, minimax thinking did more than justify specific estimators. It changed how researchers evaluated the ambition of nonparametric procedures themselves. Instead of asking whether an estimator performed well at a single convenient model, the relevant question became how well it could possibly perform across the full range of functions compatible with the maintained assumptions. In this way, minimax theory helped complete the inferential foundation of nonparametric econometrics by linking estimation, testing, and confidence construction to explicit notions of statistical difficulty.

11.6 Semiparametrics as a Bridge

Semiparametric econometrics emerged from the recognition that flexibility was most useful when deployed selectively rather than indiscriminately. Fully nonparametric models relaxed functional-form restrictions, but they often did so at substantial cost in terms of convergence rates, interpretability, and vulnerability to the curse of dimensionality. Parametric models, by contrast, retained the advantages of root-n consistency, asymptotic normality, and familiar inferential procedures, but they did so by imposing structure that was sometimes economically convenient rather than empirically credible. Semiparametric models offered an alternative path. They preserved finite-dimensional parameters of direct economic interest while allowing other components of the model to remain unspecified. In historical terms, this made semiparametrics more than an intermediate technical category. It became one of the principal ways in which nonparametric ideas were absorbed into mainstream econometrics.

This bridging role is central to the history of the field. Semiparametric methods reassured econometricians that relaxing parametric assumptions need not entail surrendering interpretable parameters or abandoning disciplined inference. At the same time, they showed that flexibility need not take the form of estimating everything nonparametrically. Instead, the guiding question became which parts of the model should be left open and which parts should remain structured. Seen from this perspective, semiparametric econometrics was not simply a compromise between two extremes; it was a constructive response to the profession's earlier skepticism about unrestricted flexibility. For broad accounts of this development, see Powell (1994), Härdle and Linton (1994), and Pagan and Ullah (1999).

11.6.1 Motivation for Semiparametric Models

The intellectual foundations of semiparametric econometrics drew on developments in both statistics and econometrics, and especially on the theory of semiparametric efficiency. A central question was how precisely a finite-dimensional parameter of interest could be estimated when the model also contained infinite-dimensional nuisance components. Chamberlain (1986) provided a systematic answer by deriving efficiency bounds for semiparametric models using the tangent-space approach. His framework characterized the best achievable asymptotic variance for estimators of structural parameters when parts of the data-generating process were left unspecified.

Historically, this mattered because it changed the way econometricians thought about model design. The question was no longer simply whether a nuisance component could be left nonparametric, but what price would be paid for doing so and under what conditions that price might be negligible. Semiparametric efficiency theory made this tradeoff precise. It showed when parametric-rate estimation could still be achieved and when the presence of nonparametric components necessarily reduced statistical precision. Just as importantly, it identified which features of the distribution carried information about the parameter of interest and which did not. In this way, it supplied a principled basis for deciding where to impose structure and where to remain agnostic.

The subsequent development of specific semiparametric estimators was guided by exactly this logic. Models such as the partially linear model, the single-index model, and average derivative estimators were not merely ad hoc constructions. They were concrete demonstrations that one could retain economically interpretable parameters while substantially weakening distributional or functional-form assumptions. This was one of the key reasons semiparametric econometrics became such an important bridge in the broader history of nonparametric methods.

11.6.2 Single-Index and Partially Linear Models

Two of the most influential semiparametric models in econometrics are the single-index model and the partially linear model. These frameworks became historically important because they provided clear, workable templates for combining parametric and nonparametric components in ways that preserved both flexibility and interpretability.

In the single-index model, the regression function depends on the explanatory variables only through a linear combination, so that the relevant covariate information is summarized by a single index. The link function remains unknown, allowing nonlinear responses to emerge from the data, while the coefficient vector continues to summarize the directional role of the covariates. Ichimura (1993) showed that the index coefficients in such models can be estimated consistently at the parametric rate without specifying the functional form of the link function, using a profile least-squares approach that concentrates out the unknown nonparametric component. This was historically important because it demonstrated that root-n estimation and flexible functional form need not be mutually exclusive.

The partially linear model provided a complementary route. In Robinson (1988), the covariates are divided into those that enter linearly with a finite-dimensional coefficient vector and those that enter through an unspecified smooth function. Robinson showed that the linear coefficients can be estimated at the parametric rate by partialling out conditional expectations, in a manner closely analogous to the Frisch–Waugh–Lovell logic of linear regression. This result was especially influential because it made the semiparametric idea operational in a very transparent way: the researcher could preserve a parameter of direct economic interest while allowing the remainder of the model to absorb nonlinearities flexibly.

At roughly the same time, however, Stock (1989) emphasized a somewhat different use of the same basic framework. Whereas Robinson's main contribution lay in showing how the parametric component could be estimated at the parametric rate in the presence of an unknown function, Stock's interest centered more directly on the nonparametric component itself and its value for policy analysis. This contrast is historically important. It reveals that the partially linear model did not enter econometrics with a single interpretation: from the outset, it supported both a semiparametric program focused on root-n inference for finite-dimensional parameters and a complementary program focused on recovering substantively meaningful nonparametric structure.

Historically, these models mattered not only because they were elegant, but because they were usable. They offered concrete blueprints for empirical work and became central tools in applied econometrics; see, for example, Powell (1994). More broadly, they helped establish a pattern that would recur throughout the semiparametric literature: the division between parametric and nonparametric components should be motivated by the economic question at hand rather than by statistical convenience alone.

11.6.3 Average Derivatives and Index Identification

Another influential development in semiparametric econometrics concerned the estimation of average derivatives. In many economic settings, the quantity of interest is not the entire regression function, but some interpretable functional summarizing how the outcome responds to changes in the regressors. Average derivatives provided a particularly important example of this idea, since they linked nonparametric estimation directly to economically meaningful marginal effects.

The seminal contribution of Powell, Stock and Stoker (1989) showed that average derivatives of regression functions can be estimated at the parametric rate even when the underlying function itself is estimated nonparametrically. This result was historically striking because it clarified that the slow pointwise convergence of nonparametric estimators need not govern the behavior of all quantities of interest. As later discussions emphasized, including Henderson and Parmeter (2015), averaging over the covariate distribution can eliminate leading bias terms and thereby restore root-n convergence for carefully chosen functionals.

This insight had implications beyond derivative estimation itself. Average derivatives also arise naturally in identifying index coefficients in semiparametric single-index models. Under suitable conditions, the direction of the index vector can be recovered from the average derivatives of the regression function with respect to the covariates. In this way, semiparametric estimation was shown not merely to soften parametric assumptions, but to provide a route to identification of interpretable structural parameters. Historically, this reinforced the broader legitimacy of semiparametric methods: they were not simply pragmatic compromises, but a coherent framework for recovering economically meaningful objects from flexible models.

11.6.4 Efficiency Bounds and Influence Functions

A central question in semiparametric econometrics concerns efficiency: how accurately can a parameter be estimated when parts of the model remain unspecified? In parametric settings, efficiency is characterized through the Fisher information matrix. In semiparametric settings, however, the presence of infinite-dimensional nuisance parameters requires a broader framework. The semiparametric efficiency bound gives the lowest achievable asymptotic variance for regular estimators of the parameter of interest, taking into account the information loss induced by leaving parts of the model unrestricted.

The development of this theory was one of the main achievements that turned semiparametrics into a mature field. Building on Chamberlain (1986), Newey (1994) developed systematic tools for deriving semiparametric efficiency bounds in models defined by conditional moment restrictions. These contributions showed that the central questions of parametric efficiency theory did not disappear once the model contained nonparametric components; they had to be reformulated in a way that accounted explicitly for nuisance tangent spaces, orthogonality, and the geometry of the parameter space.

A key tool in this analysis is the influence function, which describes the sensitivity of an estimator to local perturbations in the underlying distribution. Influence functions became central because they linked asymptotic linearity, efficiency, and estimator construction in a unified way. Historically, this mattered because it gave the semiparametric literature an organizing framework that was every bit as rigorous as classical parametric efficiency theory. Together, efficiency bounds and influence functions established not merely that semiparametric estimation could sometimes achieve root-n rates, but that such estimation could be evaluated against sharp information-theoretic benchmarks.

11.6.5 The Rise of Orthogonality and Robust Moments

An influential conceptual development in semiparametric econometrics was the use of orthogonality conditions to construct estimators that are locally insensitive to errors in the estimation of nuisance components. The underlying idea traces back to Neyman (1959), who introduced orthogonal score functions designed to reduce sensitivity to small perturbations in nuisance parameters. In the semiparametric setting, this idea took on a new role: it provided a way to preserve reliable inference on parameters of interest even when nuisance functions were estimated flexibly.

Within econometrics, these ideas were absorbed into the framework of conditional moment restrictions and efficiency theory. Newey (1994) formalized the role of orthogonality in the construction of efficient semiparametric estimators, showing how moment conditions could be chosen so that small first-step estimation errors would have only second-order effects on the parameter of interest. Historically, this was important because it further strengthened the bridge between flexibility and reliable inference. It showed that one could allow nuisance components to be estimated nonparametrically without necessarily sacrificing root-n behavior for the finite-dimensional target.

The importance of orthogonality became even more visible in later work linking semiparametrics to machine learning. Chernozhukov et al. (2018) showed that combining flexible nuisance estimation with orthogonal moment conditions and sample splitting permits valid inference on low-dimensional parameters even when the nuisance functions are learned using modern machine-learning methods. In this respect, one of the enduring legacies of classical semiparametric econometrics is that it supplied the conceptual foundation for much of what is now presented as modern debiased or orthogonal machine-learning inference. The historical continuity is important: these developments did not emerge from nowhere, but from decades of work on semiparametric efficiency, influence functions, and robust moments.

11.7 Beyond the Cross-Section

Much of the early development of nonparametric econometrics took place in cross-sectional settings under independence assumptions. This was natural both mathematically and historically. The i.i.d. framework provided the cleanest environment in which to establish consistency, derive rates of convergence, and study the behavior of smoothing estimators. It also aligned with the kinds of regression and density estimation problems through which nonparametric methods first entered econometric discussion. Yet many of the empirically important problems in economics do not arise in such settings. Time series data involve temporal dependence, persistence, and nonstationarity, while panel data combine repeated observations over time with unobserved heterogeneity across units. Extending nonparametric methods to these environments therefore required more than routine technical modification. It forced the literature to confront how dependence structures, dynamic behavior, and latent heterogeneity alter both the statistical and econometric content of flexible estimation; see, among others, Ullah (1988), Härdle and Linton (1994), Fan and Yao (2003), and Henderson and Parmeter (2015).

Historically, the move beyond the cross-section marked an important stage in the maturation of the field. So long as nonparametric methods remained tied mainly to i.i.d. settings, they could be viewed as elegant but limited tools, applicable only under relatively stylized conditions. Their extension to time series and panel data was therefore significant not only because it broadened the class of admissible data-generating processes, but because it tested whether nonparametric econometrics could adapt to the forms of dependence and heterogeneity that characterize much of empirical economics. As in earlier stages of the field's development, these extensions did not eliminate old concerns about dimensionality, inference, and interpretation. Rather, they reappeared in altered form and often with greater force.

11.7.1 Time Series

The extension of nonparametric methods to time series settings required econometricians to move beyond the independence assumptions that underpinned much of the early asymptotic theory. In a time series environment, consistency and asymptotic normality cannot simply be imported from the cross-sectional case, because serial dependence changes the behavior of local averages and kernel-weighted statistics. One of the earlier econometric contributions in this direction was Robinson (1983), who studied nonparametric regression with dependent observations and established conditions under which kernel estimators remain well behaved. This was an important step because it showed that nonparametric regression need not be confined to cross-sectional sampling schemes. More generally, the development of nonparametric time series methods required dependence conditions strong enough to support asymptotic theory while remaining flexible enough for economically relevant dynamics.

The early econometric significance of this extension is also visible in applications that embedded nonparametric components within time series models of substantive economic interest. Engle, Granger, Rice and Weiss (1986), for example, modeled electricity sales as a partially linear function of price, income, and a nonparametric component in temperature. Historically, this mattered because it demonstrated that semiparametric and nonparametric ideas could be used in time series applications without abandoning economic content. Flexible components were not confined to abstract smoothing exercises; they could be incorporated into empirically recognizable models.

As the literature developed, time series nonparametrics expanded beyond dependent-kernel regression to include models in which coefficients themselves vary over time or across states of the system. Functional-coefficient and varying-coefficient specifications allowed the regression relationship to evolve smoothly rather than remain fixed. Cai (2007) studied trending time-varying coefficient models with serially correlated errors, while Cai, Li and Park (2009) extended such methods to nonstationary environments involving integrated regressors. These contributions were historically important because they showed that nonparametric flexibility could be adapted not only to nonlinear conditional means, but also to smooth forms of parameter instability and persistence that arise naturally in macroeconomic and financial applications.

A parallel development involved the use of nonparametric methods for testing dependence structures that conventional linear tools could miss. Fan and Yao (2003) synthesized a broad range of techniques for nonlinear autoregressive models, conditional means, and volatility dynamics, helping consolidate the field. At the same time, work such as Hong and White (2005) developed nonparametric entropy-based measures of serial dependence, and Hong and Kao (2004) extended related ideas to tests for serial correlation of unknown form in panel regression settings. The historical significance of this literature lies in the fact that nonparametric time series methods were not used only for estimation; they also became tools for diagnosing forms of dependence that standard autocorrelation-based approaches were poorly equipped to detect.

The same pattern appeared in work on structural change and nonstationarity. Many classical time series procedures were designed for abrupt breaks or stationary environments, whereas nonparametric methods allowed more gradual forms of parameter drift and more general departures from stationarity to be studied. B. Chen and Hong (2012) proposed a test capable of detecting both smooth structural change and abrupt breaks, while Gao, King, Lu and Tjøstheim (2009) established nonparametric specification testing results for nonlinear autoregressions with nonstationary regressors. These developments reinforced a broader lesson that would recur throughout the field: once nonparametric methods were pushed into more realistic data environments, the main challenge was no longer simply how to smooth, but how to do so in the presence of persistence, dependence, and nonstandard limiting behavior.

11.7.2 Panel Data

The extension of nonparametric methods to panel data came later and proved, in many ways, even more difficult. Panel datasets combine repeated observations over time with cross-sectional heterogeneity, thereby offering richer variation than pure cross-sections while also introducing complications that are especially acute for flexible estimators. Chief among these is unobserved individual heterogeneity. In linear models, fixed effects can often be removed by within transformations or first differencing without fundamentally changing the object of interest. In nonlinear settings, however, it had long been understood that unobserved effects generate much deeper problems; see, for example, Chamberlain (1984). In nonparametric settings, those same operations can distort the regression function itself. This meant that the panel extension of nonparametric econometrics was not simply a matter of porting over familiar panel tricks.

Early attempts to move flexible methods into panel settings often proceeded through semiparametric or partially linear specifications rather than through fully nonparametric fixed-effects models. This was not accidental. Such models provided a way to incorporate heterogeneity and dynamics while keeping the dimensionality of the problem manageable. Contributions such as Li and Stengos (1996) and Kniesner and Li (2002) are useful in this regard, since they illustrate an early stage in which economists were already exploring panel smoothing and varying-coefficient ideas, but were doing so in forms disciplined enough to remain empirically tractable. Historically, this intermediate stage matters because it shows that the delay in the panel literature was not due to lack of interest, but to the genuine difficulty of the underlying econometric problem.

The random effects case was conceptually simpler and therefore developed earlier. Because the individual-specific effect need not be eliminated through transformation, the main issue is often efficiency rather than consistency. In this setting, Lin and Carroll (2000) obtained an unexpected result: under standard bandwidth shrinkage, generalized least squares adjustments for within-cluster correlation yield no asymptotic efficiency gains over working independence. Henderson and Ullah (2005) then developed feasible GLS procedures for nonparametric panel models with random effects, showing how correlation structures could still be incorporated in practice despite the asymptotic equivalence result. Historically, these contributions were important because they clarified that dependence in panels mattered differently in nonparametric settings than in familiar parametric ones.

The fixed effects case poses the more fundamental challenge. In a nonparametric model, subtracting within-unit means or first-differencing does not merely remove the fixed effect; it changes the nonparametric object being estimated and may induce substantial asymptotic bias.

Progress in nonparametric fixed effects panel econometrics emerged gradually. Henderson, Carroll and Li (2008) proposed estimators that account for unobserved heterogeneity while preserving the flexibility of kernel regression. Later work by Rodriguez-Poo and Soberon addressed the bias problem more directly. Rodriguez-Poo and Soberón (2013) considered semiparametric varying-coefficient panels, while Rodriguez-Poo and Soberón (2015b) showed both that naive within-type transformations lead to a non-negligible asymptotic bias and that a backfitting algorithm applied after the within transformation can restore optimal convergence rates while preserving oracle properties. Their Monte Carlo comparison study Rodriguez-Poo and Soberón (2015a) confirmed that these corrections substantially improve on naive differencing strategies. The historical importance of this line of work is that it made clear why panel nonparametrics had lagged: the obstacles were not computational alone, but conceptual.

Semiparametric panel models also provided an important bridge. Su and Ullah (2006) developed profile likelihood methods for partially linear fixed effects panels, offering a more tractable route to estimation when a fully nonparametric treatment would be too costly. This bridging role fits naturally with the broader semiparametric perspective emphasized in Powell (1994). Dynamic settings added another layer of difficulty, since lagged dependent variables enter the flexible component and generate new sources of dependence. Su and Lu (2013) showed that the first-differenced model in such cases can be recast in additive form and analyzed through iterative kernel methods, characterizing the estimator through a Fredholm integral equation and establishing its large-sample properties. These developments illustrate a broader historical pattern: when fully nonparametric panel estimation ran into severe obstacles, semiparametric structure often supplied the discipline needed to keep the problem manageable.

More recent work has explored alternatives to differencing altogether. An example is the correlated random-effects approach of Henderson, Henry and Soberón (2025), which embeds the Mundlak (1978) device within a nonparametric regression framework by augmenting the specification with within-unit averages of the regressors. This line of work is important because it points to a recurring lesson in the history of nonparametric econometrics: progress often came not from abandoning structure, but from finding economically interpretable forms of structure that made flexible estimation feasible.

An interesting historical feature of the panel literature is the speed with which it expanded once workable entry points had been established. For many years, panel nonparametrics appeared to be a difficult frontier with only scattered contributions. Yet once the basic problems of heterogeneity, transformation bias, and dimensionality began to admit tractable solutions, the literature grew rapidly enough to generate multiple review treatments within a relatively short period. This is evident in surveys such as Ai and Li (2008), Su and Ullah (2011), J. Chen, Li and Gao (2013), Rodriguez-Poo and Soberón (2017), and Parmeter and Racine (2019). That pattern is itself revealing. It suggests that the delay in the development of the field reflected the genuine difficulty of the underlying econometric problems rather than any lack of interest in panel applications.

The broader historical significance of panel nonparametrics lies in what it revealed about the limits of direct extension. Panel datasets offer richer information than cross-sections, but the combination of multiple regressors, time dependence, and latent heterogeneity intensifies the curse of dimensionality and complicates identification. For that reason, many successful panel contributions have relied on hybrid strategies that combine nonparametric components with semiparametric restrictions designed to control complexity. The eventual appearance of multiple review essays and chapter-length syntheses was itself evidence that the field had crossed an important threshold: what began as an uneasy attempt to carry smoothing methods beyond the cross-section had become a recognizable econometric subliterature in its own right.

11.8 Discrete Smoothing and Mixed Data

One of the most practically important developments in nonparametric econometrics was the extension of smoothing methods to settings involving both continuous and discrete data. Classical kernel methods were developed primarily for continuous variables, yet much of economic data is inherently mixed: regressors such as family size, occupation, treatment status, choices, market categories, and institutional indicators are discrete, while other covariates are continuous. This mismatch between the support conditions envisioned in early nonparametric theory and the support structure of actual economic data was not a minor technical inconvenience. It was one of the reasons nonparametric methods could appear less relevant to applied econometrics than their theoretical elegance suggested; see Li and Racine (2003) and J. Racine and Li (2004).

The historical significance of discrete smoothing lies precisely here. It helped close the gap between nonparametric theory and empirical practice. Conventional approaches to handling discrete regressors in a nonparametric framework often relied on frequency estimators or sample splitting, effectively partitioning the data into cells defined by the discrete variables and then applying smoothing only within those cells. While theoretically valid, this strategy quickly became impractical when the number of cells was large relative to the sample size, as is often the case in economics. The development of kernel methods that smooth discrete as well as continuous regressors therefore marked an important turning point: it made fully nonparametric and semiparametric methods more compatible with the support structure of real economic data.

11.8.1 From Discrete Smoothing in Statistics to Mixed Data in Econometrics

The statistical origins of this literature predate its econometric uptake. A foundational contribution was Aitchison and Aitken (1976), who proposed a kernel method for multivariate binary data in a discrimination setting. Their key insight was that smoothing need not be confined to continuous supports. Instead, one could smooth probability estimates over discrete outcomes in a way that avoided the brittleness of purely frequency-based procedures. Wang and van Ryzin (1981) extended this line of thought by developing smooth estimators for discrete distributions, helping broaden the statistical foundations for nonparametric methods on noncontinuous supports.

These contributions were historically important because they established that smoothing could be understood more generally as a way of borrowing strength locally, even when locality was defined over discrete rather than Euclidean structure. Yet for some time this remained largely a statistical development. The econometric importance of discrete smoothing became much clearer only when researchers began confronting mixed data settings directly, where the relevant problem was not merely smoothing a discrete distribution but estimating regression, density, and conditional distribution functions with both categorical and continuous covariates.

11.8.2 The Mixed-Data Turn in Econometrics

The econometric literature took a major step forward when mixed-data kernel methods were developed for regression and density estimation. Li and Racine (2003) proposed a nonparametric kernel estimator for joint distributions with mixed discrete and continuous variables, showing that such estimators could be equipped with cross-validated bandwidth selection and asymptotic normality. J. Racine and Li (2004) extended these ideas to regression functions, developing a nonparametric estimator that handles both continuous and categorical regressors in a natural way and demonstrating that it can outperform conventional frequency-based approaches in finite samples.

Historically, this was a major advance because it changed the practical scope of nonparametric econometrics. The issue was no longer how to apply continuous-data kernel methods to stylized settings, but how to build a genuinely mixed-data nonparametric toolkit suited to the kinds of data economists actually analyze. These contributions showed that the presence of discrete variables did not require an automatic retreat to semiparametric models or cell-based estimation. Instead, discrete regressors could themselves be smoothed, allowing interactions between qualitative and quantitative variables to be modeled without sample splitting and without the severe efficiency losses that such splitting often entails.

11.8.3 Cross-Validation, Irrelevant Regressors, and Automatic Dimension Reduction

A further development gave the mixed-data literature even greater econometric significance. One of the most persistent objections to nonparametric methods was that they suffer acutely from the curse of dimensionality and therefore become impractical when the regressor set is moderately rich. The mixed-data literature responded not only by avoiding sample splitting, but also by showing that data-driven smoothing methods could automatically reduce effective dimensionality when some regressors were irrelevant.

Hall, Li and Racine (2007) showed that least-squares cross-validation in mixed-data kernel regression can asymptotically smooth out irrelevant regressors, assigning them bandwidths that effectively remove them from the regression function. This was a striking result. Historically, it mattered because it suggested that nonparametric methods need not always take dimensionality as fixed ex ante; the data-driven smoothing procedure itself could partially determine which variables matter for the nonparametric fit. This gave the mixed-data literature an importance that went beyond support conditions alone. It connected directly to one of the field's deepest practical concerns: how to make nonparametric estimation feasible in multivariate empirical work.

Related developments extended this logic beyond regression. Ouyang, Li and Racine (2006) studied cross-validation and the estimation of probability distributions with categorical data, highlighting how smoothing can reduce variance relative to conventional frequency estimators and how vector-valued smoothing parameters can oversmooth variables that do not contribute useful information. The broader lesson was that discrete smoothing was not merely a way to avoid empty cells; it was also a way to let the data discipline the effective complexity of the problem.

11.8.4 Beyond Regression: Densities, Distributions, Quantiles, and Testing

Once established, mixed-data smoothing methods spread rapidly to other parts of nonparametric econometrics. Hall, Racine and Li (2004) considered conditional density estimation with mixed data, while J. Racine, Li and Zhu (2004) developed multivariate conditional distribution estimation in settings where both the conditioned and conditioning variables may be mixed. Li and Racine (2008) extended the framework to conditional distribution and quantile estimation, again emphasizing the practical value of mixed-data bandwidth selection. Hsiao, Li and Racine (2007) developed a specification test for models with mixed discrete and continuous regressors, showing that smoothing the discrete variables can yield substantial power gains relative to conventional frequency-based tests.

This broadening of scope is historically revealing. It shows that discrete smoothing was not an isolated trick attached to one corner of the literature. Rather, it became a general principle for making nonparametric econometrics operational in mixed-data environments. As the framework spread from regression to densities, distributions, quantiles, varying-coefficient models, and specification testing, it helped establish that support heterogeneity was not a peripheral issue but a central practical dimension of nonparametric econometric methodology.

11.8.5 Why Discrete Smoothing Mattered

The importance of discrete smoothing in the history of nonparametric econometrics lies in three related achievements. First, it made nonparametric methods compatible with the support structure of actual economic data. Second, it avoided the efficiency losses associated with conventional sample splitting and thereby expanded the range of empirical problems for which flexible estimation was feasible. Third, through cross-validation and automatic oversmoothing, it provided one of the most practically relevant responses to dimensionality concerns in applied work; see, in particular, J. Racine and Li (2004), Hall et al. (2007), and Hsiao et al. (2007).

For these reasons, the mixed-data literature deserves to be seen as more than a technical extension of kernel methods. It was one of the developments that made nonparametric econometrics genuinely usable in the empirical settings economists most often confront. In that sense, discrete smoothing occupies an important place in the field's history: it helped transform nonparametric methods from tools designed largely for continuous-support theory into a broader econometric framework capable of accommodating the discrete-continuous mixtures that characterize much of applied economics.

11.9 Nonparametric Instrumental Variables

Instrumental variables methods had long occupied a central place in econometrics before they became a nonparametric topic. In linear and parametric settings, IV offered a familiar route for addressing endogeneity while preserving interpretable structural parameters. But once functional-form assumptions were relaxed, the IV problem changed in a fundamental way. The issue was no longer simply how to isolate exogenous variation for a finite-dimensional coefficient vector, but how to recover an entire structural function in the presence of endogenous regressors. In that environment, identification, estimation, and stability became tightly intertwined. Nonparametric instrumental variables therefore marked another major stage in the maturation of nonparametric econometrics: it showed that flexibility in the presence of endogeneity required not just weaker assumptions, but a new conceptual apparatus built around operators, inverse problems, and regularization. For broad treatments of this development, see Florens (2003), Newey and Powell (2003), Newey (2013), Horowitz (2014), and Darolles et al. (2011).

Historically, NPIV was important not only because it extended IV methods into a flexible setting, but because it exposed more clearly than many earlier literatures the costs of abandoning parametric structure. In ordinary nonparametric regression, the curse of dimensionality and bandwidth choice already posed serious obstacles. With endogeneity present, the problem became still harder: instruments had to identify a function, not merely shift a regressor, and even when point identification could be established, the resulting estimation problem was often ill posed. In this sense, NPIV did not simply inherit the earlier tensions in nonparametric econometrics; it intensified them and gave them a sharper mathematical form.

11.9.1 From Linear IV to Nonlinear and Nonparametric IV

Extending IV methods beyond linear parametric models proved substantially more difficult than adapting other regression techniques to nonparametric settings. In linear IV, the identifying content of the instruments is summarized through moment conditions that isolate a finite-dimensional parameter vector. Once functional-form assumptions are relaxed, however, the relationship between the endogenous regressor and the outcome must be treated as an unknown function. The instruments must then identify that function itself. This shift transformed both the identification problem and the associated estimation task in ways that had no real analogue in conventional linear IV.

The move toward NPIV was not instantaneous. Intermediate semiparametric and nonlinear IV frameworks helped clarify the issues that would later become central. In particular, Newey, Powell and Vella (1999) showed how triangular systems and control-function methods could be used to address endogeneity in flexible settings under stronger structural conditions. This literature did not yet formulate the problem in terms of operator inversion, but it demonstrated that once endogeneity entered nonlinear and semiparametric models, identification and estimation became substantially more delicate. In that sense, it formed an important bridge between classical IV and the later NPIV literature.

A major step forward came with the development of the nonparametric instrumental variables framework in which the structural relationship is left unspecified and identified through conditional moment restrictions. Newey and Powell (2003) provided a foundational treatment of this problem, while Florens (2003) developed an operator-based formulation that made explicit the connection between NPIV estimation and inverse problems on function spaces. These papers, and later surveys such as Newey (2013), mattered historically because they gave the literature a common language. They showed that nonparametric IV was not simply 'IV with a smoother', but a distinct econometric problem requiring new ways of thinking about identification and estimation.

Once these ideas were in place, NPIV emerged as a recognizable research area rather than a loose collection of nonlinear IV attempts. In linear IV, endogeneity could often be discussed in terms of rank conditions and finite matrices. In NPIV, by contrast, the relevant objects were conditional expectation operators, completeness conditions, and the stability of functional inversion. That shift in language was itself part of the field's intellectual consolidation.

11.9.2 Integral Equations and Ill-Posed Inverse Problems

One of the most important conceptual advances in the NPIV literature was the realization that the identifying relation could be expressed as an integral equation. Specifically, the conditional expectation of the dependent variable given the instrument can be written as an operator applied to the unknown structural function. Recovering that function therefore requires inverting the operator defined by the conditional distribution of the endogenous regressor given the instrument. This recasting of the problem, emphasized particularly by Florens (2003), connected econometrics directly to the mathematical theory of inverse problems.

What made this formulation historically decisive was that it clarified why NPIV was so much more difficult than linear IV. The operator linking the structural function to observable moments is typically compact, and its inverse is often discontinuous. As a result, small perturbations in the estimated conditional expectation can lead to large changes in the recovered structural function. In econometric terms, this means that even when the model is point identified, the estimation problem may be extremely unstable in finite samples. The contribution of Carrasco et al. (2007) was especially important here, since it showed how spectral properties of the operator govern the degree of ill posedness and therefore the difficulty of estimation. See also Horowitz (2014) for a broader retrospective treatment of ill-posed inverse problems in economics.

This operator perspective also helped bring NPIV into closer connection with the broader identification issues discussed earlier in the chapter. It underscored that uniqueness of the structural function and reliable recovery of that function are distinct matters. A model can be point identified while still being so ill posed that estimation is highly fragile. In that sense, NPIV sharpened one of the central lessons of nonparametric econometrics: identification and estimation cannot always be treated as cleanly separable stages of analysis.

11.9.3 Regularization and Stability

Once NPIV was understood as an ill-posed inverse problem, regularization became unavoidable. Direct inversion of the relevant operator is typically too unstable to be useful in finite samples, so estimation must proceed by replacing the original problem with a nearby, well-behaved approximation. The essential idea is to trade some bias for a substantial reduction in variance and instability. In NPIV, this is not merely a technical convenience; it is part of what makes estimation possible at all.

Several distinct regularization strategies emerged. Hall and Horowitz (2005) proposed regularized estimators for NPIV models and derived minimax convergence rates, showing how the attainable rate depends jointly on smoothness and the degree of ill posedness. Darolles et al. (2011) developed Tikhonov regularization for kernel-based NPIV estimators and established asymptotic normality. A different route was pursued by Blundell, Chen and Kristensen (2007), who used sieve methods to approximate the unknown structural function with flexible basis expansions and derived corresponding convergence and inference results. Together these contributions established the main toolkit of NPIV estimation: spectral truncation, penalization, and sieve approximation. For broader overviews of how regularization became central to the field, see Horowitz (2011) and Horowitz (2014).

Historically, the importance of this literature lies in the fact that it moved the discussion beyond the binary question of whether NPIV was identified. It showed that even after identification had been secured, the econometrician still faced a nontrivial design problem: how should the inversion be stabilized, and what kind of approximation error is acceptable? This was a further step away from the older hope that nonparametric methods might simply relax assumptions and let the data speak for themselves. In NPIV, the data could speak only through an explicitly regularized filter.

11.9.4 Completeness Conditions and Their Critiques

The condition that came to occupy the most prominent place in NPIV identification was completeness. In this context, completeness ensures that the conditional expectation operator linking the structural function to observable moments is injective, so that distinct structural functions generate distinct conditional expectations given the instrument. In other words, it is the condition that turns observational equivalence into uniqueness. Newey and Powell (2003) made this role especially clear, and subsequent work treated completeness as one of the defining identifying assumptions of the NPIV framework.

At the same time, completeness also became a focal point of criticism. Its mathematical role is clean, but its empirical content is often difficult to assess, and in most applications it is not something that can be directly verified from the data. Canay, Santos and Shaikh (2013) showed that completeness is not testable in general, while Freyberger (2017) investigated the extent to which completeness itself can be probed empirically. Relatedly, Freyberger and Horowitz (2015) studied what can still be learned when completeness fails, thereby connecting NPIV identification to the broader partial identification perspective. These contributions were historically important because they brought NPIV back into contact with the central concern of nonparametric econometrics more generally: not merely whether an assumption is mathematically sufficient, but whether it is economically and empirically credible.

The response to this problem was not to abandon NPIV, but to become more explicit about alternatives. One possibility was to work with partial identification and bounds when completeness could not be justified. Another was to incorporate additional structure, such as monotonicity or concavity, that might help restore identifying power without reverting to full parametric specification. In this sense, the critique of completeness played a productive role. It reminded the literature that the formal elegance of NPIV identification conditions did not remove the need for economically and empirically credible assumptions.

11.9.5 Empirical Applications and Practical Limitations

Although the theoretical NPIV literature developed rapidly, empirical applications remained comparatively sparse. That asymmetry is historically revealing. It reflects not a lack of interest in endogeneity, which had always been central to econometrics, but the fact that NPIV methods impose severe practical demands. Large samples are often needed, instruments must be sufficiently informative in a functional sense, and regularization choices can materially affect the results. For many empirical settings, these hurdles made NPIV far less routine than its linear counterpart. Indeed, the question of how useful NPIV would prove in applied work became a topic of discussion in its own right; see Horowitz (2011).

Still, several influential applications demonstrated what the approach could offer. Blundell et al. (2007) applied NPIV methods to Engel curves, showing that nonlinear expenditure relationships could be estimated under endogeneity without relying on conventional parametric forms. Henderson, Papageorgiou and Parmeter (2013) applied flexible IV methods to the relationship between financial development and economic growth, allowing both endogeneity and nonlinear heterogeneity in the response. Their implementation used the three-stage local polynomial procedure of Su and Ullah (2008), which builds on the triangular-system control-function framework of Newey et al. (1999) and thereby avoids direct operator inversion under stronger structural conditions.

This last point is historically significant. One of the enduring lessons of the NPIV literature is that many empirical researchers preferred approaches that sidestepped the harshest features of ill-posed operator inversion when plausible additional structure was available. Control-function methods, triangular systems, and related semiparametric approaches often proved more attractive in practice precisely because they traded some generality for feasibility. In that respect, the empirical history of NPIV mirrors a broader pattern in nonparametric econometrics: the most theoretically general method is not always the one most readily absorbed into applied work.

The conceptual impact of NPIV, however, has been substantial regardless of the relative scarcity of applications. It revealed with unusual clarity how endogeneity, identification, and regularization interact once functional-form assumptions are relaxed. It also underscored a lesson that runs throughout the history of nonparametric econometrics: greater flexibility does not eliminate the need for structure. It changes the form that structure must take.

11.10 Shape Restrictions and Economic Theory

One of the most important ways in which nonparametric econometrics matured was through the incorporation of qualitative restrictions derived from economic theory. Monotonicity, convexity, concavity, homogeneity, and related shape restrictions offered a way to discipline flexible estimation without collapsing it back into a fully parametric model. This was historically significant because it helped clarify that the choice between parametric and nonparametric methods was never simply binary. Between unrestricted flexibility and fully specified functional form lay a broad class of intermediate approaches in which theory entered through inequality and curvature restrictions rather than through closed-form equations.

Seen in this light, shape restrictions did more than improve finite-sample behavior. They changed the conceptual role of theory in nonparametric work. Instead of treating theory as a source of exact parametric specification, the literature increasingly treated it as a source of qualitative information about the admissible set of functions. This proved especially valuable in settings where unrestricted nonparametric estimators were too variable to be useful, but parametric assumptions were stronger than the economics could justify. In that sense, shape-restricted estimation became one of the clearest embodiments of theory-guided nonparametrics. For general discussions of this perspective, see Matzkin (1994, 2007), and Henderson and Parmeter (2009). More broadly, the econometric literature built on an earlier statistical tradition of order-restricted inference and constrained estimation; see, for example, Barlow, Bartholomew, Bremner and Brunk (1972) and Robertson, Wright and Dykstra (1988).

11.10.1 Monotonicity, Convexity, and Concavity

Economic theory often implies qualitative restrictions on economic relationships even when it does not determine their exact functional form. Demand curves are typically downward sloping, production functions are often assumed to be monotone and concave in inputs, and cost functions are convex in output, reflecting increasing marginal costs. These restrictions matter because they reduce the class of economically admissible functions without requiring the econometrician to commit to a specific parametric family.

The idea that such restrictions could be imposed directly on flexible estimators has deep roots. Hildreth (1954) provided one of the earliest formal treatments of regression under concavity constraints, showing that qualitative shape information could enter estimation directly rather than only through parametric approximation. This work anticipated a broader statistical literature on order-restricted inference, later synthesized by Barlow et al. (1972) and Robertson et al. (1988). Historically, these contributions are important precursors to later econometric work, because they established that qualitative restrictions could themselves be an organizing principle for statistical estimation.

What changed later was the recognition that these restrictions were not merely statistical conveniences, but often carried clear economic content. Matzkin (1994) provided a systematic treatment of nonparametric and semiparametric estimation under economically motivated shape restrictions, showing how monotonicity, concavity, and homogeneity could be exploited across a range of models including production, demand, and discrete choice. This marked an important stage in the history of the field: theory no longer entered flexible estimation only through exclusion restrictions or maintained stochastic assumptions, but also through direct restrictions on the shape of the unknown function itself.

11.10.2 Nonparametric Regression Under Shape Constraints

Once the economic motivation for shape restrictions became clear, a substantial methodological literature developed around how to impose them in practice. The problem was not trivial. Standard nonparametric estimators are local and flexible, but precisely for that reason they may violate economically necessary qualitative properties in finite samples. A regression curve that is meant to be monotone may wiggle downward; a production surface that should be concave may exhibit local convexities. These were not merely cosmetic issues. If the estimated function violated the economic model's basic qualitative implications, its interpretability was compromised.

A range of estimation strategies emerged in response, including isotonic regression, kernel-based constrained estimators, and sieve methods in which shape restrictions are imposed directly on the approximating basis. The statistical foundations for such procedures lie in the broader order-restricted literature represented by Barlow et al.

(1972) and Robertson et al. (1988). Early theoretical work, such as Mammen (1991), established statistical properties of shape-constrained estimators and helped show that these procedures could recover the underlying function while respecting qualitative restrictions. Hall and Huang (2001) developed a kernel-based approach to imposing monotonicity on nonparametric regression estimators, thereby integrating smoothing and shape discipline within a unified procedure.

In econometric applications, the attraction of these methods lay precisely in their ability to reconcile theory and flexibility. Demand functions could be constrained to satisfy monotonicity or curvature conditions, production functions could be forced to exhibit diminishing marginal returns, and treatment-response functions could be restricted to move in one direction with treatment intensity when such monotonicity had substantive justification. Henderson and Parmeter (2009) surveyed the range of economic constraints that can be incorporated into nonparametric regression and discussed practical implementation in multivariate settings. Historically, this literature helped move shape restrictions from a marginal technical topic to a central tool in theory-guided empirical work.

11.10.3 Identification Gains from Economic Theory

Shape restrictions did more than stabilize estimation. They also became an important source of identifying power. In many econometric models, unrestricted nonparametric identification is weak or unavailable, while full parametric identification is achieved only by imposing assumptions whose economic content is unclear. Shape restrictions offered an intermediate route. By ruling out certain classes of functions on economic grounds, they could reduce ambiguity and in some cases restore point identification.

This role of economic theory in nonparametric identification was emphasized clearly by Matzkin (2007), who showed how qualitative restrictions such as monotonicity, concavity, and homogeneity can identify structural objects that would otherwise remain weakly identified or unidentified. Historically, this mattered because it extended the nonparametric identification literature beyond support conditions and completeness. It demonstrated that theory could aid identification not only by introducing instruments or exclusion restrictions, but also by constraining the geometry of the structural function itself.

Even when such restrictions were not sufficient for point identification, they often narrowed the identified set and thereby linked the shape-restrictions literature directly to the partial-identification perspective discussed earlier in the chapter. A particularly influential example is Manski and Pepper (2000), who showed how monotonicity restrictions could sharpen identification through the use of monotone instrumental variables even when conventional point identification was unavailable. In this sense, the identification gains from shape restrictions ranged across a spectrum: in some models they delivered full recovery of the object of interest, while in others they merely sharpened the set of admissible values. That flexibility is one reason the literature became so important methodologically. It showed that qualitative theory can matter even when it is too weak to generate full parametric structure.

11.10.4 Computational and Inferential Challenges

The appeal of shape-restricted methods came with costs of their own. Imposing inequality constraints often turns estimation into a constrained optimization problem, and the difficulty rises sharply when multiple shape restrictions must hold simultaneously in multivariate settings. This meant that practical progress in the literature depended not only on theoretical insight, but also on advances in numerical methods capable of enforcing economically meaningful constraints without destroying the flexibility that made nonparametric estimation attractive in the first place.

Inference proved even more subtle. Because shape restrictions may bind in some regions of the support and not in others, the asymptotic distribution of a constrained estimator often depends on which restrictions are active at the truth. This produces nonstandard limiting behavior and complicates the construction of confidence intervals and test procedures. In other words, the same restrictions that help stabilize estimation and sharpen identification also make inferential theory more difficult. For broader modern treatments of both computational and inferential issues in shape-constrained estimation, see Groeneboom and Jongbloed (2014).

These issues became especially visible in settings where shape restrictions yielded partial rather than point identification. Chernozhukov, Lee and Rosen (2013) developed inference procedures for intersection bounds, providing tools for empirically relevant problems in which economic theory narrows but does not fully determine the admissible parameter space. Historically, this is an important endpoint for the section because it makes clear that theory-guided nonparametrics did not eliminate uncertainty; rather, it changed the form in which uncertainty had to be analyzed. The resulting literature thus occupies a central place in the history of nonparametric econometrics: it showed how weak structure could simultaneously improve estimation, sharpen identification, and complicate inference.

11.11 Inequality, Welfare, and Distributional Analysis

One of the domains in which nonparametric econometrics proved especially natural was the analysis of income distributions, inequality, and welfare. In this literature, the object of interest is often not a finite-dimensional parameter or even a regression function, but the distribution itself, together with functionals derived from it such as Lorenz curves, dominance orderings, and welfare rankings. This made inequality and welfare analysis an especially important setting for the development of nonparametric methods. Where parametric approaches required researchers to impose a family such as the lognormal or Pareto on the income distribution, nonparametric methods allowed the data to determine the shape of the distribution more directly. In historical terms, this mattered because it helped normalize the idea that empirical economics could be organized around estimation and inference for function-valued objects rather than only around coefficients and summary statistics.

At the same time, this literature also revealed some of the inferential and conceptual challenges that accompany such flexibility. Distributional objects are function-valued, and economically meaningful comparisons often depend on the behavior of entire functions rather than on isolated points. Pointwise procedures therefore proved inadequate for many of the questions economists wanted to ask. More broadly, the literature showed that once one shifts from summary measures to full distributions, the boundary between statistical measurement and normative evaluation becomes more visible, a distinction emphasized clearly by Maasoumi (1998). This is one reason inequality and welfare analysis occupies a distinctive place in the history of nonparametric econometrics: it is not merely an application area, but a setting in which the field's emphasis on functional objects, partial orderings, and weak assumptions found particularly direct expression.

11.11.1 Nonparametric Estimation of Income Distributions

Empirical analysis of income and wealth inequality often focuses on the shape and evolution of entire distributions rather than on a small collection of summary measures. Kernel density estimation and related smoothing techniques made it possible to estimate these distributions directly from the data without imposing a particular parametric family. In that respect, the distributional literature drew very naturally on earlier statistical work in density estimation, including the foundational contributions of Rosenblatt (1956) and Parzen (1962).

Historically, this mattered because many of the salient features of income distributions are precisely the ones that simple parametric families have difficulty accommodating: skewness, heavy tails, multimodality, and changing shape over time or across populations. Nonparametric methods therefore did more than provide technical flexibility. They altered the empirical questions researchers could ask. Instead of comparing means or variances under maintained distributional forms, economists could study the evolution of the entire distribution itself.

A prominent illustration is Quah (1996), who argued that conventional parametric approaches were poorly suited to the study of the world income distribution and used kernel density methods to document the emergence of a 'twin peaks' pattern suggestive of polarization rather than convergence. That contribution was historically important because it showed how nonparametric methods could change the substantive interpretation of a major empirical question. Later work also made clear that such findings raised their own inferential issues. Henderson, Parmeter and Russell (2008), for example, examined whether apparent multimodality in worldwide productivity distributions was statistically credible or instead a byproduct of bandwidth choice and testing strategy. In this way, the literature on income distributions illustrated both the power and the fragility of nonparametric distributional analysis.

11.11.2 Lorenz Curves and Stochastic Dominance

Lorenz curves provide one of the most familiar graphical representations of inequality, describing the cumulative share of total income received by the bottom fraction of the population. From the standpoint of nonparametric econometrics, their importance lies in the fact that the Lorenz curve is itself a functional object derived from the income distribution. Estimating and comparing Lorenz curves therefore extends the logic of nonparametric distribution estimation into the analysis of inequality orderings. Early statistical work by Gastwirth (1971) studied the sampling properties of the Lorenz curve and methods for estimating it from observed data.

The broader significance of Lorenz curves in this context is that they connect statistical estimation to partial welfare orderings. Income distributions are often compared through stochastic dominance relations, which permit welfare rankings without requiring the analyst to commit to a specific parametric model of the income distribution. This made stochastic dominance a particularly congenial setting for nonparametric methods. Because the relevant comparisons depend on the behavior of entire distribution functions or their integrals, nonparametric estimators are well suited to the problem.

This connection was developed more fully in work such as Davidson and Duclos (2000), who proposed inference procedures for stochastic dominance that account for sampling variability in estimated distributions, and Linton, Maasoumi and Whang (2005), who developed consistent dominance tests valid under general dependence conditions and without parametric restrictions on the distributions being compared. Historically, these contributions are important because they show how nonparametric econometrics extended beyond flexible estimation into the construction of robust orderings and hypothesis tests for distributional comparisons.

11.11.3 Welfare Comparisons Without Functional Assumptions

Nonparametric methods also became important in welfare analysis, where the central question is not merely how unequal a distribution is, but how alternative distributions should be ranked from a social point of view. Traditional welfare comparisons often proceed by specifying a social welfare function or inequality index, thereby embedding strong normative assumptions in the form of the evaluator's objective function. This is not inherently problematic, but it makes the welfare ranking depend on a particular functional form as well as on the observed data. As Maasoumi (1998) emphasized, the value judgments implicit in particular welfare and inequality measures became increasingly well understood, together with the limitations of basing welfare statements on any single index.

A different route is to use dominance criteria that permit welfare comparisons across broad classes of admissible welfare functions. The classic contribution of Atkinson (1970) showed how second-order stochastic dominance can be linked to welfare rankings under inequality aversion, thereby separating the statistical problem of estimating distributions from the normative problem of selecting a particular welfare function. Historically, this separation is one of the most important reasons this literature belongs in a history of nonparametric econometrics. It illustrates a central aspiration of the field: to weaken arbitrary functional assumptions while making more transparent which conclusions are driven by the data and which depend on maintained evaluative structure.

11.11.4 Partial Rankings and Distributional Policy Evaluation

Distributional comparisons do not always yield complete rankings. Income distributions may intersect in ways that prevent clean dominance conclusions: one distribution may be more favorable at the bottom of the income scale and less favorable at the top, so that the welfare ranking depends on how strongly the evaluator values redistribution. In such cases, partial rankings become more informative than any single summary statistic, because they identify the range of welfare criteria under which the ranking is robust and the regions in which it is not.

This point was developed by Atkinson and Bourguignon (1987), who extended dominance analysis to settings in which welfare comparisons are only partial and may involve multidimensional criteria. From the perspective of nonparametric econometrics, the key feature is that these rankings depend on the full shape of the distribution rather than on a finite set of moments or parameters. The econometric problem is therefore one of estimating and comparing complex distributional objects under weak assumptions.

A related and influential development was the use of reweighting methods to construct counterfactual distributions. DiNardo, Fortin and Lemieux (1996) showed how one could decompose distributional changes by reweighting observations to reflect alternative institutional or labor-market environments. Although often discussed in the decomposition literature, this contribution also fits naturally into the present narrative. It showed that once economists began to treat entire distributions as empirical objects of interest, nonparametric tools could be used not only to describe distributions as they were, but also to study how they might have differed under alternative economic conditions. In this respect, the literature on inequality and welfare analysis helped push nonparametric econometrics beyond estimation of smooth curves and toward the broader analysis of distributional structure, partial orderings, and counterfactual functions.

11.12 Nonparametric Demand, Auctions, and Industrial Organization

Industrial organization provided some of the clearest examples of how economic theory could discipline nonparametric estimation without dictating a particular functional form. In many IO settings, the objects of interest are constrained by qualitative implications of economic theory rather than by fully specified stochastic structure; see, for example, Athey and Haile (2007) and Reiss and Wolak (2007). This made IO an especially important domain for theory-guided nonparametric work: rather than treating flexibility as a rejection of structure, economists in this literature used nonparametric and semiparametric methods to avoid functional-form assumptions that theory itself did not require.

This interaction mattered because it forced a sharper reckoning with what economic theory actually contributes. In many empirical settings, parametric models had seemed well identified because functional-form assumptions quietly supplied much of the structure. Once those assumptions were relaxed, the literature had to confront which restrictions were genuinely implied by theory and which were merely convenient approximations. In that sense, the IO literature did not simply apply nonparametric methods to new problems. It helped clarify one of the central methodological lessons of the field: economic structure and functional-form structure are not the same thing.

11.12.1 Revealed Preference as Early Theory Without Functional Form

One of the earliest and most influential connections between nonparametric reasoning and economic theory appeared in the revealed preference literature. Rather than positing a parametric utility function, revealed preference analysis asks whether observed choices are consistent with utility maximization under weak behavioral assumptions. This made the literature inherently congenial to nonparametric thinking, since the goal was to test or recover economic structure without imposing a specific functional form.

The classic result here is Afriat's theorem, due to Afriat (1967), which provides necessary and sufficient conditions under which a finite set of observed choices can be rationalized by a utility-maximizing consumer. What made Afriat's theorem historically important for the present chapter is that it showed how economic theory could generate testable restrictions and recoverability results without relying on parametric specification. Varian (1982, 1983) translated these ideas into practical econometric procedures for testing revealed preference restrictions on finite datasets, thereby turning revealed preference into an operational empirical program.

In retrospect, this literature foreshadowed much of what later became central in theory-guided nonparametric econometrics. It demonstrated that monotonicity, convexity, and related qualitative implications of theory could substitute for functional-form assumptions while still yielding economically meaningful empirical content. In that sense, revealed preference analysis served as an early example of a broader lesson that would later become central in nonparametric demand estimation: the most useful role of economic theory may be to restrict the admissible shape of a function rather than to dictate its parametric form.

11.12.2 Nonparametric Demand Estimation Under Economic Restrictions

The estimation of consumer demand functions became a natural extension of these ideas. Traditional empirical demand systems, such as the translog model of Christensen, Jorgenson and Lau (1973) and the Almost Ideal Demand System of Deaton and Muellbauer (1980), offered tractable and influential parametric representations of consumer behavior. But they also imposed global structure that was often justified more by convenience than by theory. Historically, this created an opening for approaches that sought to preserve the restrictions genuinely implied by demand theory while relaxing auxiliary assumptions about functional form.

An intermediate step in this evolution was provided by more flexible parametric and semiparametric demand specifications, such as the quadratic Engel curve work of Banks, Blundell and Lewbel (1997), which made clear that empirical demand analysis often required more flexibility than conventional systems allowed. But the deeper shift came when researchers began to combine nonparametric estimation with revealed-preference and shape restrictions. In this setting, conditions such as homogeneity, symmetry, and Slutsky negativity do not pin down a functional form, but they do define a narrower class of economically admissible demand functions. Demand estimation therefore became one of the clearest cases in which nonparametric econometrics could be disciplined by theory rather than left fully unrestricted.

An important strand of this literature is represented by Blundell, Browning and Crawford (2003), who combined nonparametric Engel curve estimation with revealed preference bounds to recover demand responses without imposing parametric preferences. Blundell, Browning and Crawford (2008) sharpened this program by deriving best nonparametric bounds on demand responses under weak assumptions, thereby making even more explicit how revealed preference restrictions could substitute for parametric structure. Blundell, Horowitz and Parey (2012) extended this logic by developing a nonparametric estimator of gasoline demand that explicitly imposes the Slutsky inequality as a shape restriction. These contributions are historically important because they show that the role of theory in nonparametric work is not merely to motivate the model, but to constrain the estimator itself in economically meaningful ways.

11.12.3 Auctions and the Identification of Valuations

Auction models provided another setting in which theory-guided nonparametric methods flourished. Here the econometrician observes bids, but the object of interest is typically the distribution of latent private valuations. Recovering that distribution requires exploiting equilibrium restrictions from auction theory. This made auctions a particularly revealing setting for nonparametric econometrics, because identification depended not on flexible smoothing alone but on whether the equilibrium structure was strong enough to map observables back to economically meaningful primitives.

The landmark contribution of Guerre, Perrigne and Vuong (2000) showed how the equilibrium bidding strategy in first-price sealed-bid auctions can be used to transform bids into pseudo-valuations, from which the valuation distribution can then be recovered nonparametrically. This paper was historically important because it provided one of the clearest demonstrations that equilibrium theory could supply the identifying bridge from observed data to latent structural objects while leaving the distribution of those objects unspecified. In that respect, it became a model for later theory-guided nonparametric work in IO more broadly.

Subsequent work also made clear that practical implementation required further structural discipline. Henderson, List, Millimet, Parmeter and Price (2012) addressed the imposition of monotonicity on equilibrium bidding strategies and developed bandwidth-selection procedures tailored to auction data, illustrating again that economic restrictions often need to be built directly into flexible estimators. At the same time, Haile and Tamer (2003) showed that under weaker behavioral assumptions in ascending auctions, bids yield only bounds on valuations rather than point identification. This connected the auction literature directly to the partial-identification themes discussed earlier in the chapter and underscored a broader lesson: once strong structural assumptions are relaxed, nonparametric flexibility often reveals limits of learnability rather than overcoming them.

11.12.4 Structural IO Without Fully Parametric Likelihoods

A broader development in industrial organization was the move away from fully specified parametric likelihoods toward estimation based on weaker implications of optimizing behavior and equilibrium. In many structural models, economic theory implies inequalities, moment restrictions, monotonicity conditions, or comparative-static relations that are considerably weaker than a full stochastic specification of the data-generating process. This opened the door to empirical strategies that remained structural in ambition while being substantially less parametric in implementation.

Seen in this light, IO helped push nonparametric econometrics toward a richer understanding of what flexible structural modeling could mean. The goal was not to estimate without theory, but to rely on the qualitative restrictions that theory actually implies rather than on auxiliary assumptions that were only indirectly motivated. Pakes, Porter, Ho and Ishii (2015) is especially important in this respect, since it shows how inequality restrictions generated by optimizing behavior and equilibrium conditions can serve as the basis for estimation and inference in multi-agent problems without requiring a fully parametric likelihood.

Historically, this is why the IO literature fits so naturally beside the shape-restrictions literature. It was one of the places where the profession learned that nonparametric and semiparametric methods could make economic theory more precise rather than less relevant. The main lesson was not that flexibility should replace structure, but that structure should come from the economic model wherever possible and from arbitrary functional-form assumptions only where necessary.

11.13 Computational Advances and Practical Adoption

The rise of nonparametric econometrics was not only a methodological story but also a computational one. Many of the estimators developed in the theoretical literature of the 1960s and 1970s were demanding by the standards of the time, and the gap between formal possibility and routine implementation was a genuine obstacle to adoption. In this respect, early skepticism toward nonparametric methods reflected not only concerns about identification, interpretation, and the curse of dimensionality, but also the fact that many procedures were simply difficult to compute on the datasets and machines available to most empirical researchers. The history of nonparametric econometrics therefore cannot be written solely as a sequence of theoretical breakthroughs. It must also be understood as a history of when those methods became feasible enough to enter ordinary empirical practice.

This mattered because computation did not merely accelerate existing methods; it influenced which methods became practically relevant. Improvements in computing power, numerical optimization, and software environments gradually changed the kinds of flexible estimators that economists could use, the sample sizes they could handle, and the amount of tuning and diagnostic work that became routine. As a result, the practical meaning of nonparametric econometrics changed over time. Techniques that once appeared too demanding for all but specialists increasingly became part of the ordinary applied toolkit. For broader discussions of the transition from theoretical development to practical implementation, see Härdle (1990), Fan and Gijbels (1996), and Li and Racine (2007).

11.13.1 From Theoretical Curiosity to Applied Tool

Although the statistical foundations of nonparametric estimation were established relatively early, widespread empirical use in economics came much later. In the first phase of development, many nonparametric procedures remained closer to theoretical demonstrations than to routine applied methods. This was not simply because economists were intellectually conservative. It was also because implementation required repeated smoothing, tuning-parameter selection, and sometimes numerically intensive optimization, all of which were costly in the computing environments of the time.

The expansion of computing power in the 1980s and 1990s began to change this situation. Kernel estimators, local polynomial methods, and related smoothers became increasingly feasible on datasets of the size typically encountered in empirical economics. At the same time, a new generation of monographs translated the literature from asymptotic theory into applied methodology. Härdle (1990) played an important role in this transition by presenting nonparametric regression as something that could actually be implemented, while Fan and Gijbels (1996) helped make local polynomial methods practically accessible and conceptually unified. Historically, these works mattered not only because they synthesized theory, but because they lowered the barrier between methodological research and empirical practice.

11.13.2 Bandwidth Choice and Cross-Validation

A central practical issue in nonparametric estimation is bandwidth choice, which governs the bias–variance tradeoff. Theoretical development of data-driven bandwidth selection methods, including cross-validation and plug-in rules, was crucial for the credibility of nonparametric methods. But the historical significance of bandwidth selection is also computational. Choosing the bandwidth is often the most numerically burdensome step in implementation, especially when the selection rule requires repeated estimation across a grid of candidate values or repeated leave-one-out calculations.

This meant that some of the most principled bandwidth selectors were, for a time, among the least convenient to use. Leave-one-out cross-validation, for example, can be computationally expensive even in moderate samples, particularly in multivariate settings. Historically, this is one reason bandwidth selection remained as much a practical obstacle as a statistical one. The growth of computing power gradually made such procedures feasible on the sample sizes common in economics, while simpler selectors and rules of thumb provided faster alternatives. Silverman (1986) played an important role here by offering practical bandwidth rules that, even when not optimal, made density estimation more accessible to applied users. Implementation-oriented discussions in Härdle (1990) and the broader applied treatment in Li and Racine (2007) further reflect the extent to which bandwidth choice had become part of routine empirical workflow rather than a purely theoretical issue.

The broader historical point is that the adoption of nonparametric econometrics depended not only on having a theory of tuning-parameter choice, but on making that choice routine enough to become part of ordinary empirical workflow. Once bandwidth selection became something that could be automated, diagnosed, and replicated with relative ease, nonparametric methods became much easier to absorb into applied research culture.

11.13.3 Curse of Dimensionality in Practice

The curse of dimensionality was not merely a theoretical warning; it was a practical constraint that shaped the direction of the field. Applied researchers quickly confronted the fact that even a modest number of continuous regressors could make unrestricted nonparametric estimation unreliable on datasets of ordinary size. In this sense, the curse of dimensionality was one of the main forces pushing nonparametric econometrics toward more structured forms of flexibility.

One response was to restrict the interaction structure of the regression function. Additive models, in which the unknown function is represented as a sum of lower-dimensional components, offered an important way to retain flexibility while reducing dimensionality. Stone (1985) established theoretical results for additive regression, and Hastie and Tibshirani (1990) made generalized additive models widely accessible to applied researchers. Historically, additive models are important not only because they mitigated dimensionality, but because they exemplified a recurring compromise in the field: rather than estimate everything nonparametrically, researchers imposed just enough structure to make flexible estimation workable.

A second response was to let the smoothing procedure itself determine which variables mattered. Hall et al. (2007) showed that cross-validation can asymptotically assign very large bandwidths to irrelevant continuous regressors, effectively smoothing them out of the regression function. This was historically significant because it suggested that dimensionality could sometimes be reduced endogenously by the data rather than only by ex ante model restriction. While the curse of dimensionality remained a fundamental theoretical constraint, such results helped show that its practical effects could be mitigated in empirically relevant settings.

11.13.4 Software, Accessibility, and Replicability

The practical adoption of nonparametric econometrics depended heavily on software. For many years, applied work required researchers to write custom code for kernel estimation, bandwidth selection, specification testing, and inference. This raised the barrier to entry substantially and limited adoption to those with both econometric and computational expertise. In historical terms, software infrastructure mattered because it determined whether a method remained the province of specialists or could become part of ordinary empirical practice.

Early dedicated software systems reflected an awareness that implementation itself was a major bottleneck. Over time, however, the decisive change came from incorporation into more general statistical computing environments. A particularly important contribution was the `np` package for `R` developed by Hayfield and Racine (2008), which brought together kernel regression, density estimation, bandwidth selection, specification testing, and support for mixed data types in a unified software platform. This mattered historically because it lowered the cost of experimentation, made replication easier, and allowed a wider community of applied researchers to use nonparametric and semiparametric methods without building everything from scratch. The broader applied synthesis of Li and Racine (2007) likewise reflects a stage at which nonparametric econometrics had become sufficiently standardized to be organized around practical implementation rather than only around asymptotic theory.

More broadly, software changed the culture of the field. Once nonparametric methods were available in widely used platforms with reasonably stable defaults, documented routines, and reproducible workflows, they became easier to teach, easier to replicate, and easier to incorporate into empirical research programs. In that sense, software was not merely a vehicle for implementation. It was one of the institutions through which nonparametric econometrics became a routine part of applied econometric practice.

11.14 What Aged Well and What Did Not

Any historical assessment of a methodological field should ask not only which ideas were influential at the time, but which proved durable under later empirical and theoretical use. The relevant criterion is not simply citation counts or elegance in the original formulation. Some contributions age well because they remain routine tools in applied work. Others age well because they become lasting conceptual frameworks, even if their original implementations are superseded. Still others survive by being absorbed into adjacent literatures in altered form. By contrast, what ages less well is often not what was wrong, but what proved narrower in practical reach than early enthusiasm suggested. In nonparametric econometrics, this distinction is especially important. The field's history is full of methods that were theoretically profound but operationally demanding, and of ideas that found their most durable life in hybrid or reformulated form.

With that perspective in mind, several contributions stand out as having aged particularly well. They tend to share certain features: they made flexibility operational under realistic data constraints, clarified the informational role of assumptions, or supplied concepts that later econometrics could absorb and repurpose. What aged less well, by contrast, were usually ambitions that underestimated the force of dimensionality, the difficulty of inference, the practical burden of tuning and regularization, or the continuing value of economic structure for interpretation and identification.

11.14.1 What Aged Well

Several contributions from the nonparametric tradition remain durable either as working empirical tools, as conceptual frameworks, or as foundations later absorbed into neighboring literatures.

Kernel density estimation and local polynomial smoothing. The kernel density estimator introduced by Rosenblatt (1956) and extended by Parzen (1962) remains the canonical nonparametric tool for density estimation. Its basic logic has survived with remarkable stability. Nonparametric regression followed a more differentiated path. The Nadaraya–Watson estimator was a decisive conceptual advance, but in applied work local polynomial methods, especially local linear regression, came to dominate because they offer improved boundary behavior and better bias properties; see Fan and Gijbels (1996). In this sense, early kernel regression aged well less as a final implementation than as the foundation on which more practically successful smoothers were built.

Semiparametric bridge models. Some of the most durable contributions were those that used nonparametric ideas selectively rather than universally. The partially linear model of Robinson (1988) and the single-index framework of Ichimura (1993) remain central because they showed that flexibility could be introduced while preserving finite-dimensional parameters of direct economic interest. At the same time, the partly linear tradition aged well in two distinct ways. Robinson's formulation emphasized reliable inference for the parametric component, while Stock (1989) highlighted the substantive importance of the nonparametric component itself for policy analysis. Both ambitions survived. Semiparametric models endured not because they split the difference between parametrics and nonparametrics in an abstract sense, but because they solved the practical problem of how to relax functional-form assumptions without sacrificing interpretability.

Cross-validation and data-driven smoothing. The cross-validation principle introduced by Stone (1974) proved far more durable than one might have expected from the early bandwidth-selection debates. Its importance lies not only in replacing ad hoc tuning choices with a data-driven procedure, but also in the fact that it adapted to new settings as the field expanded. In mixed-data environments, cross-validation became one of the mechanisms through which nonparametric econometrics became genuinely usable on economic data rather than only on stylized continuous-support problems. The result of Hall et al. (2007) — that cross-validation can asymptotically smooth out irrelevant regressors — is especially striking in this respect. A procedure originally designed to balance bias and variance turned out also to perform automatic dimension reduction.

Discrete smoothing and mixed-data methods. The extension of nonparametric methods to mixed discrete–continuous data addressed a genuine practical gap. Classical kernel methods were developed largely for continuous variables, but economic data routinely involve binaries, categories, and counts. The discrete-smoothing literature, building on Aitchison and Aitken (1976) and Wang and van Ryzin (1981) and carried into econometrics by Li and Racine (2003) and J. Racine and Li (2004), reduced the need for sample splitting and made kernel methods more compatible with the support structures economists routinely face.

Semiparametric efficiency, influence functions, and orthogonality. The framework developed by Chamberlain (1986) and Newey (1994) aged well, not only within semiparametric econometrics narrowly understood, but across modern econometric theory more broadly. Semiparametric efficiency bounds provided a principled benchmark for what can be estimated well when nuisance components are infinite-dimensional, while influence-function logic supplied a general language for robust first-order approximations. These ideas did not remain confined to their original setting. They reappeared in doubly robust estimation, orthogonal scores, and modern machine-learning-assisted inference, a case in which a classical semiparametric contribution survived by being absorbed into later econometric practice rather than by remaining frozen in its original form.

Identification without functional forms and the legitimacy of identified sets. One of the durable intellectual contributions of the nonparametric turn was the insistence that identification must be analyzed before flexible estimation is celebrated. The nonparametric identification literature made explicit how much identifying power had often been supplied implicitly by functional-form assumptions. In that sense, what aged best here was not a particular estimator but a way of thinking. The partial-identification program associated with Manski (1990) and later work endured because it replaced a false choice between strong point-identifying assumptions and complete agnosticism with a disciplined account of what data and assumptions jointly determine. Modern identification arguments are now routinely formulated in nonparametric terms before further structure is imposed for estimation.

Shape restrictions and theory-guided flexibility. The insight that economic theory often implies qualitative restrictions — monotonicity, convexity, concavity, homogeneity — that can substitute for arbitrary functional-form assumptions has aged well. Work such as Matzkin (1994) showed that theory-guided restrictions could stabilize estimation, improve interpretability, and sometimes aid identification without collapsing flexible models back into rigid parametric forms. Applications in demand analysis, production, auctions, causal inference, and partial identification have continued to develop. Shape restrictions are a strong example of the chapter's broader lesson: what endured best were not attempts to abandon structure, but attempts to use structure more precisely.

Nonparametric ideas absorbed into machine learning. A durable legacy of nonparametric econometrics is that many of its central concerns reappeared, often in reformulated form, in the later interaction between econometrics and machine learning. The management of the bias–variance tradeoff through regularization, the use of flexible function approximators, the centrality of resampling and out-of-sample validation, and the distinction between predictive success and structural or causal interpretation were all familiar themes within nonparametric and semiparametric econometrics before the recent machine-learning wave. In this sense, the machine-learning turn did not represent a clean break with the earlier history traced in this chapter. Rather, it provided a new setting in which older nonparametric insights found renewed relevance. Early work such as White and Racine (2001) already illustrated how bootstrap inference and neural-network modeling could be brought into econometric analysis, while later contributions such as Athey and Imbens (2019) made explicit how modern machine-learning methods could be interpreted, adapted, and disciplined for econometric purposes.

What aged well, then, was not any single algorithm, but a broader methodological lesson: flexible statistical methods become most durable in economics when they are combined with credible identification strategies, transparent regularization, and interpretable inferential goals.

11.14.2 What Aged Less Well

What aged less well was usually not the underlying theory, but the scope of the original ambition. In several important areas, the literature's lasting contribution was real, yet its routine practical reach turned out to be narrower than early hopes implied.

Rules of thumb as general solutions to the tuning problem. Normal-reference bandwidth rules, of which Silverman (1986) remains the most influential example, were indispensable in making nonparametric estimation accessible. But their durable role turned out to be narrower than one might have hoped. Their limitations under skewness, heavy tails, multimodality, and regression settings are now well understood, and careful applied work typically prefers data-driven selectors when feasible. Still, this is not a story of outright failure. Rules of thumb aged reasonably well as starting points, benchmarks, and diagnostic devices. What did not age well was the hope that they could serve as broadly reliable substitutes for principled tuning-parameter selection.

Nonparametric instrumental variables in routine applied work. The NPIV literature made major contributions to identification theory and to the econometrics of ill-posed inverse problems. What aged less well was the expectation that these methods would become common empirical workhorses. Large sample requirements, sensitivity to regularization choices, and the practical difficulty of justifying conditions such as completeness kept NPIV estimators from becoming routine outside a relatively small set of applications. This gap between theoretical importance and empirical uptake is an example of a contribution whose conceptual value exceeds its day-to-day applied use.

Nonparametric specification tests as routine diagnostics for applied researchers. Consistent specification tests against nonparametric alternatives were an important conceptual achievement. They showed that one could test parametric models against very broad alternatives without specifying the alternative in advance. What proved harder was turning that asymptotic attractiveness into routine applied practice. Finite-sample power can be sensitive to tuning choices, bootstrap implementation is not always trivial, and the interpretation of a rejection is often less informative than applied researchers would like. A rejection establishes that the parametric model is inadequate, but not which alternative structure should replace it. These are practical rather than foundational limitations, but they help explain why such tests have remained more influential methodologically than as routine components of applied empirical workflows.

The terminology of flexible approximation methods. Some terminology aged less well than the underlying ideas. The label 'semi-nonparametric,' attached to flexible approximation methods based on increasingly rich functional expansions, was meant to signal a middle ground between rigid parametric specification and fully unrestricted nonparametric estimation. In practice, however, the term was never especially transparent, and it has largely fallen out of active use. The underlying ideas — sieve-like flexibility, approximation theory, and tractable departures from parametric form — proved useful and in some cases quite durable. What aged less well was the terminology rather than the methods themselves.

The expectation that nonparametric methods would displace parametric modeling. Early in the development of the field, it was natural to imagine that sufficiently flexible nonparametric regression might broadly displace conventional functional-form modeling. In practice, the curse of dimensionality, data limitations, interpretability concerns, and the need for economically meaningful restrictions meant that fully unrestricted estimation rarely became the default empirical strategy. Nonparametric methods became informal diagnostic tools, components of semiparametric estimators, nuisance-function technologies in causal inference, mixed-data workhorses in selected settings, and conceptual guides to identification and inference. The field moved toward additive structures, mixed-data methods, and theory-guided restrictions rather than toward unrestricted flexibility. That is not a smaller achievement than early advocates imagined. The durable lesson was that flexible methods needed help from structure to become routinely useful, and learning that lesson was itself one of the field's lasting contributions.

Taken as a whole, the history traced in this chapter suggests a clear lesson. What aged best were contributions that either made flexibility operational under real data constraints or clarified the role of assumptions in identification and inference. What aged less well were ambitions that underestimated the continuing force of dimensionality, regularization, interpretation, and economic structure. The long-run success of nonparametric econometrics therefore lies not in having overthrown parametric modeling, but in having taught econometricians how to relax structure more intelligently.

References

Afriat, S. N. (1967). The construction of utility functions from expenditure data. *International Economic Review*, *8*(1), 67–77.

Ai, C. & Li, Q. (2008). Semi-parametric and non-parametric methods in panel data models. In L. Mátyás & P. Sevestre (Eds.), *The econometrics of panel data: Fundamentals and recent developments in theory and practice* (Vol. 46, pp. 451–478). Berlin Heidelberg: Springer. doi: 10.1007/978-3-540-75892-1_14

Aitchison, J. & Aitken, C. G. G. (1976). Multivariate binary discrimination by the kernel method. *Biometrika*, *63*(3), 413–420.

Armstrong, T. B. & Kolesár, M. (2020). Simple and honest confidence intervals in nonparametric regression. *Quantitative Economics*, *11*(1), 1–39. doi: 10.3982/QE1199

Athey, S. & Haile, P. A. (2007). Nonparametric approaches to auctions. In J. J. Heckman & E. E. Leamer (Eds.), *Handbook of econometrics* (Vol. 6A, pp.

3847–3965). Amsterdam: Elsevier.
Athey, S. & Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, *11*, 685–725. doi: 10.1146/annurev-economics-080217-053433
Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory*, *2*(3), 244–263.
Atkinson, A. B. & Bourguignon, F. (1987). Income distribution and differences in needs. In G. R. Feiwel (Ed.), *Arrow and the foundations of the theory of economic policy* (pp. 350–370). London: Macmillan.
Banks, J., Blundell, R. & Lewbel, A. (1997). Quadratic Engel curves and consumer demand. *The Review of Economics and Statistics*, *79*(4), 527–539.
Barlow, R. E., Bartholomew, D. J., Bremner, J. M. & Brunk, H. D. (1972). *Statistical inference under order restrictions*. London: John Wiley & Sons.
Bickel, P. J. & Rosenblatt, M. (1973). On some global measures of the deviations of density function estimates. *Annals of Statistics*, *1*(6), 1071–1095.
Bierens, H. J. (1982). Consistent model specification tests. *Journal of Econometrics*, *20*(1), 105–134.
Bierens, H. J. (1983). Uniform consistency of kernel estimators of a regression function under generalized conditions. *Journal of the American Statistical Association*, *78*(383), 699–707.
Bierens, H. J. (1990). A consistent conditional moment test of functional form. *Econometrica*, *58*(6), 1443–1458.
Blundell, R., Browning, M. & Crawford, I. (2003). Nonparametric Engel curves and revealed preference. *Econometrica*, *71*(1), 205–240.
Blundell, R., Browning, M. & Crawford, I. (2008). Best nonparametric bounds on demand responses. *Econometrica*, *76*(6), 1227–1262.
Blundell, R., Chen, X. & Kristensen, D. (2007). Semi-nonparametric IV estimation of shape-invariant Engel curves. *Econometrica*, *75*(6), 1613–1669.
Blundell, R., Horowitz, J. L. & Parey, M. (2012). Measuring the price responsiveness of gasoline demand: Economic shape restrictions and nonparametric demand estimation. *Quantitative Economics*, *3*(1), 29–51.
Bowman, A. W. (1984). An alternative method of cross-validation for the smoothing of density estimates. *Biometrika*, *71*(2), 353–360.
Cai, Z. (2007). Trending time-varying coefficient time series models with serially correlated errors. *Journal of Econometrics*, *136*(1), 163–188.
Cai, Z., Li, Q. & Park, J. Y. (2009). Functional-coefficient models for nonstationary time series data. *Journal of Econometrics*, *148*(2), 101–113.
Canay, I. A., Santos, A. & Shaikh, A. M. (2013). On the testability of identification in some nonparametric models with endogeneity. *Econometrica*, *81*(6), 2535–2559.
Cantelli, F. P. (1933). Sulla determinazione empirica delle leggi di probabilità. *Giornale dell'Istituto Italiano degli Attuari*, *4*, 421–424.
Carrasco, M., Florens, J.-P. & Renault, E. (2007). Linear inverse problems in structural econometrics. In J. J. Heckman & E. E. Leamer (Eds.), *Handbook of econometrics* (Vol. 6B, pp. 5633–5751). Amsterdam: North-Holland.

Chamberlain, G. (1984). Panel data. In (Vol. 2, p. 1247-1318). Elsevier. doi: https://doi.org/10.1016/S1573-4412(84)02014-6

Chamberlain, G. (1986). Asymptotic efficiency in semiparametric models with conditional moment restrictions. *Journal of Econometrics*, *32*, 189–218.

Chen, B. & Hong, Y. (2012). Testing for smooth structural changes in time series models via nonparametric regression. *Econometrica*, *80*(3), 1157–1183.

Chen, J., Li, D. & Gao, J. (2013). *Non- and semi-parametric panel data models: A selective review* (Monash Econometrics and Business Statistics Working Papers No. 18/13). Monash University, Department of Econometrics and Business Statistics. Retrieved from https://ideas.repec.org/p/msh/ebswps/2013-18.html

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W. & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *Econometrics Journal*, *21*(1), C1–C68.

Chernozhukov, V., Lee, S. & Rosen, A. M. (2013). Intersection bounds: Estimation and inference. *Econometrica*, *81*(2), 667–737.

Chesher, A. (2003). Identification in nonseparable models. *Econometrica*, *71*(5), 1405–1441.

Christ, C. F. (1994). The Cowles Commission's contributions to econometrics at Chicago, 1939–1955. *Journal of Economic Literature*, *32*(1), 30–59.

Christensen, L. R., Jorgenson, D. W. & Lau, L. J. (1973). Transcendental logarithmic production frontiers. *Review of Economics and Statistics*, *55*(1), 28–45.

Darolles, S., Fan, Y., Florens, J.-P. & Renault, E. (2011). Nonparametric instrumental regression. *Econometrica*, *79*(5), 1541–1565.

Davidson, R. & Duclos, J.-Y. (2000). Statistical inference for stochastic dominance and for the measurement of poverty and inequality. *Econometrica*, *68*(6), 1435–1464.

Deaton, A. & Muellbauer, J. (1980). An almost ideal demand system. *American Economic Review*, *70*(3), 312–326.

DiNardo, J., Fortin, N. M. & Lemieux, T. (1996). Labor market institutions and the distribution of wages, 1973–1992: A semiparametric approach. *Econometrica*, *64*(5), 1001–1044.

Donoho, D. L. (1994). Statistical estimation and optimal recovery. *Annals of Statistics*, *22*(1), 238–270.

Dvoretzky, A., Kiefer, J. & Wolfowitz, J. (1956). Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *Annals of Mathematical Statistics*, *27*(3), 642–669.

Engle, R. F., Granger, C. W. J., Rice, J. & Weiss, A. (1986). Semiparametric estimates of the relation between weather and electricity sales. *Journal of the American Statistical Association*, *81*(394), 310–320.

Fan, J. & Gijbels, I. (1996). *Local polynomial modelling and its applications*. London: Chapman & Hall.

Fan, J. & Yao, Q. (2003). *Nonlinear time series: Nonparametric and parametric methods*. New York: Springer.

Florens, J.-P. (2003). Inverse problems and structural econometrics: The example of instrumental variables. In M. Dewatripont, L. P. Hansen & S. J. Turnovsky

(Eds.), *Advances in economics and econometrics: Theory and applications* (Vol. 2, pp. 284–311). Cambridge: Cambridge University Press.

Freyberger, J. (2017). On completeness and consistency in nonparametric instrumental variable models. *Econometrica*, *85*(5), 1629–1644.

Freyberger, J. & Horowitz, J. L. (2015). Identification and shape restrictions in nonparametric instrumental variables estimation. *Journal of Econometrics*, *189*(1), 41–53.

Gao, J., King, M., Lu, Z. & Tjøstheim, D. (2009). Nonparametric specification testing for nonlinear time series with nonstationarity. *Econometric Theory*, *25*(6), 1869–1892.

Gastwirth, J. L. (1971). A general definition of the Lorenz curve. *Econometrica*, *39*(6), 1037–1039.

Glivenko, V. (1933). Sulla determinazione empirica delle leggi di probabilità. *Giornale dell'Istituto Italiano degli Attuari*, *4*, 92–99.

Groeneboom, P. & Jongbloed, G. (2014). *Nonparametric estimation under shape constraints*. Cambridge: Cambridge University Press.

Guerre, E., Perrigne, I. & Vuong, Q. (2000). Optimal nonparametric estimation of first-price auctions. *Econometrica*, *68*(3), 525–574.

Haile, P. A. & Tamer, E. (2003). Inference with an incomplete model of English auctions. *Journal of Political Economy*, *111*(1), 1–51.

Hájek, J. & Šidák, Z. (1967). *Theory of rank tests*. New York: Academic Press.

Hall, P. (1992). Effect of bias estimation on coverage accuracy of bootstrap confidence intervals for a probability density. *Annals of Statistics*, *20*(2), 675–694.

Hall, P. (1994). Methodology and theory for the bootstrap. In R. F. Engle & D. L. McFadden (Eds.), *Handbook of econometrics* (Vol. 4, pp. 2341–2381). Amsterdam: North-Holland.

Hall, P. & Horowitz, J. L. (2005). Nonparametric methods for inference in the presence of instrumental variables. *Annals of Statistics*, *33*(6), 2904–2929.

Hall, P. & Horowitz, J. L. (2013). A simple bootstrap method for constructing nonparametric confidence bands for functions. *The Annals of Statistics*, *41*(4), 1892–1921. doi: 10.1214/13-AOS1137

Hall, P. & Huang, L.-S. (2001). Nonparametric kernel regression subject to monotonicity constraints. *Statistica Sinica*, *11*(3), 563–586.

Hall, P., Li, Q. & Racine, J. S. (2007). Nonparametric estimation of regression functions in the presence of irrelevant regressors. *Review of Economics and Statistics*, *89*(4), 784–789.

Hall, P., Racine, J. & Li, Q. (2004). Cross-validation and the estimation of conditional probability densities. *Journal of the American Statistical Association*, *99*(468), 1015–1026.

Härdle, W. (1990). *Applied nonparametric regression*. Cambridge: Cambridge University Press.

Härdle, W. & Linton, O. (1994). Applied nonparametric methods. In R. F. Engle & D. McFadden (Eds.), *Handbook of econometrics* (1st ed., Vol. 4, p. 2295-2339). Elsevier.

Härdle, W. & Mammen, E. (1993). Comparing nonparametric versus parametric regression fits. *Annals of Statistics*, *21*(4), 1926–1947.

Hastie, T. J. & Tibshirani, R. J. (1990). *Generalized additive models*. London: Chapman and Hall.

Hayfield, T. & Racine, J. S. (2008). Nonparametric econometrics: The np package. *Journal of Statistical Software*, *27*(5), 1–32.

Heckman, J. J. (2001). Micro data, heterogeneity, and the evaluation of public policy. *Journal of Political Economy*, *109*(4), 673–748.

Heckman, J. J., Ichimura, H. & Todd, P. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *Review of Economic Studies*, *64*(4), 605–654.

Henderson, D. J., Carroll, R. J. & Li, Q. (2008). Nonparametric estimation and testing of fixed effects panel data models. *Journal of Econometrics*, *144*(1), 257–275.

Henderson, D. J., Henry, E. K. & Soberón, A. (2025). Nonparametric correlated random-effects models. In *Advanced studies in theoretical and applied econometrics* (Vol. 57, pp. 287–305). Springer.

Henderson, D. J., List, J. A., Millimet, D. L., Parmeter, C. F. & Price, M. K. (2012). Empirical implementation of nonparametric first-price auction models. *Journal of Econometrics*, *168*(1), 17–28.

Henderson, D. J., Papageorgiou, C. & Parmeter, C. F. (2013). Who benefits from financial development? New methods, new evidence. *European Economic Review*, *63*, 47–67.

Henderson, D. J. & Parmeter, C. F. (2009). Imposing economic constraints on nonparametric regression: Survey, implementation and extensions. In Q. Li & J. S. Racine (Eds.), *Advances in econometrics: Nonparametric econometric methods* (Vol. 25, pp. 433–469). Bingley: Emerald.

Henderson, D. J. & Parmeter, C. F. (2015). *Applied nonparametric econometrics*. Cambridge: Cambridge University Press.

Henderson, D. J., Parmeter, C. F. & Russell, R. R. (2008). Modes, weighted modes, and calibrated modes: Evidence of clustering using modality tests. *Journal of Applied Econometrics*, *23*(5), 607–638.

Henderson, D. J. & Ullah, A. (2005). A nonparametric random effects estimator. *Economics Letters*, *88*(3), 403–407.

Hettmansperger, T. P. (1984). *Statistical inference based on ranks*. New York: Wiley.

Hildreth, C. (1954). Point estimates of ordinates of concave functions. *Journal of the American Statistical Association*, *49*(267), 598–619.

Hirano, K., Imbens, G. W. & Ridder, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, *71*(4), 1161–1189.

Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Annals of Mathematical Statistics*, *19*(3), 293–325.

Hong, Y. & Kao, C. (2004). Wavelet-based testing for serial correlation of unknown form in panel models. *Econometrica*, *72*(5), 1519–1563.

Hong, Y. & White, H. (2005). Asymptotic distribution theory for nonparametric entropy measures of serial dependence. *Econometrica*, *73*(3), 837–901.

Horowitz, J. L. (2001). The bootstrap. In J. J. Heckman & E. E. Leamer (Eds.), *Handbook of econometrics* (Vol. 5, pp. 3159–3228). Amsterdam: North-Holland.

Horowitz, J. L. (2003). The bootstrap in econometrics. *Statistical Science*, *18*(2), 211–218.

Horowitz, J. L. (2011). Applied nonparametric instrumental variables estimation. *Econometrica*, *79*(2), 347–394. doi: 10.3982/ECTA8662

Horowitz, J. L. (2014). Ill-posed inverse problems in economics. *Annual Review of Economics*, *6*(1), 21–51. doi: 10.1146/annurev-economics-080213-041213

Horowitz, J. L. & Manski, C. F. (2000). Nonparametric analysis of randomized experiments with missing covariate and outcome data. *Journal of the American Statistical Association*, *95*(449), 77–84.

Hsiao, C., Li, Q. & Racine, J. S. (2007). A consistent model specification test with mixed discrete and continuous data. *Journal of Econometrics*, *140*, 802–826. doi: 10.1016/j.jeconom.2006.07.015

Hurwicz, L. (1950). Generalization of the concept of identification. In T. C. Koopmans (Ed.), *Statistical inference in dynamic economic models* (pp. 417–439). New York: Wiley. (Cowles Commission Monograph 10)

Ichimura, H. (1993). Semiparametric least squares (SLS) and weighted SLS estimation of single-index models. *Journal of Econometrics*, *58*(1-2), 71–120.

Kniesner, T. J. & Li, Q. (2002). Nonlinearity in dynamic adjustment: Semiparametric estimation of panel labor supply. *Empirical Economics*, *27*(1), 131–148. doi: 10.1007/s181-002-8363-1

Koopmans, T. C. (1947). Measurement without theory. *Review of Economics and Statistics*, *29*(3), 161–172.

Leamer, E. E. (1983). Let's take the con out of econometrics. *American Economic Review Papers and Proceedings*, *73*(2), 31–43.

Li, Q. & Racine, J. (2003). Nonparametric estimation of distributions with categorical and continuous data. *Journal of Multivariate Analysis*, *86*, 266–292. doi: 10.1016/S0047-259X(02)00025-8

Li, Q. & Racine, J. S. (2007). *Nonparametric econometrics: Theory and practice*. Princeton: Princeton University Press.

Li, Q. & Racine, J. S. (2008). Nonparametric estimation of conditional CDF and quantile functions with mixed categorical and continuous data. *Journal of Business & Economic Statistics*, *26*(4), 423–434. doi: 10.1198/073500107000000250

Li, Q. & Stengos, T. (1996, None). Semiparametric estimation of partially linear panel data models. *Journal of Econometrics*, *71*(1-2), 389-397.

Lin, X. & Carroll, R. J. (2000). Nonparametric function estimation for clustered data when the predictor is measured without/with error. *Journal of the American Statistical Association*, *95*(450), 520–534.

Linton, O., Maasoumi, E. & Whang, Y.-J. (2005). Consistent testing for stochastic dominance under general sampling schemes. *Review of Economic Studies*,

72(3), 735–765.
Low, M. G. (1997). On nonparametric confidence intervals. *Annals of Statistics*, *25*(6), 2547–2554.
Maasoumi, E. (1998). Empirical analyses of inequality and welfare. In M. H. Pesaran & P. Schmidt (Eds.), *Handbook of applied econometrics, volume II: Microeconomics* (pp. 202–245). Oxford: Blackwell.
Mammen, E. (1991). Estimating a smooth monotone regression function. *Annals of Statistics*, *19*(2), 724–740.
Mann, H. B. & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics*, *18*(1), 50–60.
Manski, C. F. (1990). Nonparametric bounds on treatment effects. *American Economic Review Papers and Proceedings*, *80*(2), 319–323.
Manski, C. F. (1995). *Identification problems in the social sciences*. Cambridge: Harvard University Press.
Manski, C. F. (2003). *Partial identification of probability distributions*. New York: Springer.
Manski, C. F. & Pepper, J. V. (2000). Monotone instrumental variables: With an application to the returns to schooling. *Econometrica*, *68*(4), 997–1010.
Matzkin, R. L. (1991). Semiparametric estimation of monotone and concave utility functions for polychotomous choice models. *Econometrica*, *59*(5), 1315–1327.
Matzkin, R. L. (1993). Nonparametric identification and estimation of polychotomous choice models. *Journal of Econometrics*, *58*(1), 137-168.
Matzkin, R. L. (1994). Restrictions of economic theory in nonparametric methods. In R. F. Engle & D. L. McFadden (Eds.), *Handbook of econometrics* (Vol. 4, pp. 2523–2558). Amsterdam: North-Holland.
Matzkin, R. L. (2007). Nonparametric identification. In J. J. Heckman & E. E. Leamer (Eds.), *Handbook of econometrics* (Vol. 6B, pp. 5307–5368). Amsterdam: North-Holland.
Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica*, *46*(1), 69–85.
Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability and Its Applications*, *9*(1), 141–142.
Newey, W. K. (1994). The asymptotic variance of semiparametric estimators. *Econometrica*, *62*(6), 1349–1382.
Newey, W. K. (2013). Nonparametric instrumental variables estimation. *American Economic Review*, *103*(3), 550–556. doi: 10.1257/aer.103.3.550
Newey, W. K. & Powell, J. L. (2003). Instrumental variable estimation of nonparametric models. *Econometrica*, *71*(5), 1565–1578.
Newey, W. K., Powell, J. L. & Vella, F. (1999). Nonparametric estimation of triangular simultaneous equations models. *Econometrica*, *67*(3), 565–603.
Neyman, J. (1959). Optimal asymptotic tests of composite statistical hypotheses. In U. Grenander (Ed.), *Probability and statistics: The harald Cramér volume* (pp. 213–234). Stockholm: Almqvist & Wiksell.

Neyman, J. (1990). On the application of probability theory to agricultural experiments: Essay on principles. *Statistical Science*, *5*(4), 465–472. (Originally published in Polish in 1923; translated by D. M. Dabrowska and T. P. Speed)

Ouyang, D., Li, Q. & Racine, J. S. (2006). Cross-validation and the estimation of probability distributions with categorical data. *Nonparametric Statistics*, *18*(1), 69–100. doi: 10.1080/10485250600569002

Pagan, A. & Ullah, A. (1999). *Nonparametric econometrics*. Cambridge: Cambridge University Press.

Pakes, A., Porter, J., Ho, K. & Ishii, J. (2015). Moment inequalities and their application. *Econometrica*, *83*(1), 315–334.

Parmeter, C. F. & Racine, J. S. (2019). Nonparametric estimation and inference for panel data models. In M. Tsionas (Ed.), *Panel data econometrics* (p. 97-129). Academic Press.

Parzen, E. (1962). On estimation of a probability density function and mode. *Annals of Mathematical Statistics*, *33*(3), 1065–1076.

Powell, J. L. (1994). Estimation of semiparametric models. In R. F. Engle & D. L. McFadden (Eds.), *Handbook of econometrics* (Vol. 4, pp. 2443–2521). Amsterdam: Elsevier.

Powell, J. L., Stock, J. H. & Stoker, T. M. (1989). Semiparametric estimation of index coefficients. *Econometrica*, *57*(6), 1403–1430.

Quah, D. T. (1996). Twin peaks: Growth and convergence in models of distribution dynamics. *Economic Journal*, *106*(437), 1045–1055.

Racine, J. & Li, Q. (2004). Nonparametric estimation of regression functions with both categorical and continuous data. *Journal of Econometrics*, *119*, 99–130. doi: 10.1016/S0304-4076(03)00157-X

Racine, J., Li, Q. & Zhu, X. (2004). Kernel estimation of multivariate conditional distributions. *Annals of Economics and Finance*, *5*, 211–235.

Racine, J. S., Su, L. & Ullah, A. (Eds.). (2014). *The oxford handbook of applied nonparametric and semiparametric econometrics and statistics*. Oxford: Oxford University Press.

Ramsey, J. B. (1969). Tests for specification errors in classical linear least squares regression analysis. *Journal of the Royal Statistical Society, Series B*, *31*(2), 350–371.

Randles, R. H. & Wolfe, D. A. (1979). *Introduction to the theory of nonparametric statistics*. New York: John Wiley & Sons.

Reiss, P. C. & Wolak, F. A. (2007). Structural econometric modeling: Rationales and examples from industrial organization. In J. J. Heckman & E. E. Leamer (Eds.), *Handbook of econometrics* (Vol. 6A, pp. 4277–4415). Amsterdam: Elsevier.

Robertson, T., Wright, F. T. & Dykstra, R. L. (1988). *Order restricted statistical inference*. Chichester: John Wiley & Sons.

Robinson, P. M. (1983). Nonparametric estimators for time series. *Journal of Time Series Analysis*, *4*(3), 185–207.

Robinson, P. M. (1988). Root-n-consistent semiparametric regression. *Econometrica*, *56*(4), 931–954.

Rodriguez-Poo, J. M. & Soberón, A. (2013). Direct semiparametric estimation of fixed effects panel data varying coefficient models. *Econometrics Journal*, *17*(1), 107–138.

Rodriguez-Poo, J. M. & Soberón, A. (2015a). Differencing techniques in semi-parametric panel data varying coefficient models with fixed effects: A Monte Carlo study. *Computational Statistics*, *30*, 885–906.

Rodriguez-Poo, J. M. & Soberón, A. (2015b). Nonparametric estimation of fixed effects panel data varying coefficient models. *Journal of Multivariate Analysis*, *133*, 95–122.

Rodriguez-Poo, J. M. & Soberón, A. (2017). Nonparametric and semiparametric panel data models: Recent developments. *Journal of Economic Surveys*, *31*(4), 923–960. doi: 10.1111/joes.12177

Roehrig, C. S. (1988). Conditions for identification in nonparametric and parametric models. *Econometrica*, *56*(2), 433-447.

Rosenbaum, P. R. & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55.

Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *Annals of Mathematical Statistics*, *27*(3), 832–837.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, *66*(5), 688–701.

Rudemo, M. (1982). Empirical choice of histograms and kernel density estimators. *Scandinavian Journal of Statistics*, *9*(2), 65–78.

Sheather, S. J. & Jones, M. C. (1991). A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society, Series B*, *53*(3), 683–690.

Silverman, B. W. (1986). *Density estimation for statistics and data analysis*. London: Chapman & Hall.

Spearman, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology*, *15*, 72–101.

Sperlich, S. (2014). On the choice of regularization parameters in specification testing: A critical discussion. *Empirical Economics*, *47*(2), 427–450.

Stock, J. H. (1989). Nonparametric policy analysis. *Journal of the American Statistical Association*, *84*(406), 567–575.

Stone, C. J. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society, Series B*, *36*(2), 111–147.

Stone, C. J. (1977). Consistent nonparametric regression. *Annals of Statistics*, *5*(4), 595–620.

Stone, C. J. (1980). Optimal rates of convergence for nonparametric estimators. *Annals of Statistics*, *8*(6), 1348–1360.

Stone, C. J. (1982). Optimal global rates of convergence for nonparametric regression. *Annals of Statistics*, *10*(4), 1040–1053.

Stone, C. J. (1985). Additive regression and other nonparametric models. *Annals of Statistics*, *13*(2), 689–705.

Su, L. & Lu, X. (2013). Nonparametric dynamic panel data models: Kernel estimation and specification testing. *Journal of Econometrics*, *176*(2), 112–133.

Su, L. & Ullah, A. (2006). Profile likelihood estimation of partially linear panel data models with fixed effects. *Economics Letters*, *92*(1), 75–81.

Su, L. & Ullah, A. (2008). Local polynomial estimation of nonparametric simultaneous equations models. *Journal of Econometrics*, *144*(1), 193–218.

Su, L. & Ullah, A. (2011). Nonparametric and semiparametric panel econometric models: Estimation and testing. In D. E. A. Giles & A. Ullah (Eds.), *Handbook of empirical economics and finance* (pp. 455–497). Boca Raton, FL: Chapman & Hall/CRC.

Tamer, E. (2003). Incomplete simultaneous discrete response model with multiple equilibria. *Review of Economic Studies*, *70*(1), 147–165.

Ullah, A. (1988). Nonparametric estimation and hypothesis testing in econometric models. *Empirical Economics*, *13*(3), 223–249.

van der Vaart, A. W. & Wellner, J. A. (1996). *Weak convergence and empirical processes*. New York: Springer.

Varian, H. R. (1982). The nonparametric approach to demand analysis. *Econometrica*, *50*(4), 945–973.

Varian, H. R. (1983). Non-parametric tests of consumer behaviour. *The Review of Economic Studies*, *50*(1), 99–110.

Wang, M.-C. & van Ryzin, J. (1981). A class of smooth estimators for discrete distributions. *Biometrika*, *68*(1), 301–309.

Watson, G. S. (1964). Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics, Series A*, *26*, 359–372.

White, H. & Racine, J. (2001). Statistical inference, the bootstrap, and neural-network modeling with application to foreign exchange rates. *IEEE Transactions on Neural Networks*, *12*(4), 657–673.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, *1*(6), 80–83.

Yatchew, A. (1998). Nonparametric regression techniques in economics. *Journal of Economic Literature*, *36*(2), 669–721.