Econometric Theory, **0**, 2024, 1–40.
doi:10.1017/S0266466624000185

THE ET INTERVIEW: BENEDIKT M. PÖTSCHER

Interviewed by Manfred Deistler
University of Technology Vienna

Benedikt Pötscher in his office in 2024.

Benedikt Pötscher was born in Vienna, Austria, on 27 November 1955. After attending elementary school and high school (Gymnasium), he studied mathematics at the University of Vienna, finishing his studies with distinction with a doctoral thesis in the field of set-theoretic topology in the Spring of 1978.

Benedikt Pötscher received a scholarship from the Institute of Advanced Studies in Vienna in 1978, and in the Fall of 1979 joined the Institute of Econometrics and Operations Research at University of Technology Vienna (TU Vienna) as an assistant professor. He obtained his Habilitation in 1985 at TU Vienna. While at TU Vienna, he visited UC Santa Barbara, Yale University, and the University of Maryland (UMD) at College Park for extended periods of time. In 1989, he became associate professor at UMD where he was promoted to full professor in 1991. In the same year, he accepted an offer from the University of Vienna as a full professor of Statistics, where he still remains, despite several other offers.

Benedikt Pötscher has worked and published extensively in various areas of econometrics and statistics, resulting in numerous papers in top journals, such as the *Annals of Statistics*, *Econometrica*, *Econometric Theory*, *Journal of Econometrics*, and *Statistical Science*. He has coauthored two books, one with Ingmar Prucha on dynamic nonlinear econometric models and one with Immanuel Bomze on evolutionary game theory. He is one of the leading experts in econometric theory. He is or has been a member of the editorial boards of *Econometric Theory*, *Econometric Reviews*, *Journal of Econometrics*, and *Journal of Statistical Planning and Inference*.

The interview was conducted in the Fall of 2023 in Benedikt Pötscher's office. Helpful comments by Hannes Leeb, Richard Nickl, Maximilian Pötscher, Severin Pötscher, David Preinerstorfer, and Ingmar Prucha on a first draft are gratefully acknowledged.

1. FAMILY BACKGROUND AND EDUCATION

First, thank you for having agreed to do this interview. I would like to start by asking you about your family background and your early education.

I was born in 1955 pretty much around the time the last allied soldiers finally left Austria. I grew up in Vienna with my parents and two sisters, one senior by 2 years and the other one junior by 3 years. My father was a classical philologist who, after having obtained his doctoral degree from the University of Vienna, taught Latin and ancient Greek in Gymnasium (high-school) and pursued academic research in his spare time. This was not unusual at that time, although the university system was already in transition to a system as we know it now where academic researchers start their university career right after their PhD, without having to spend some time in purgatory teaching high-school students. He eventually obtained a "Habilitation" at the University of Salzburg, where, for many years after, he then taught as an external "Dozent" one day a week while still keeping the employment as a high-school teacher in Vienna. Later in his life, he became a full professor at the University of Graz. My mother was the homemaker and took care of the children. My parents had met at the university where my mother was studying to become a high-school teacher in Latin and History. After they had gotten married, she left university before finishing her studies, something that was not uncommon back then, but that she later regretted. Given her background, she had a keen interest in my father's research and success, and also was a person my father could talk to about these matters. She furthermore supported him by typing all of his manuscripts, including his books, which was quite a feat given that she had to take care of the household and the children. This production of articles and books happened in our living room; so from a young age on I was confronted with various aspects of academic research. I remember that around the age of 10, I volunteered, and was enlisted, to help with the compilation of the index to one of my father's books, which—given the absence of modern technology at that time—meant going through the entire text several times, watching out for the keywords and jotting down the page numbers. My mother was also very much interested in music, played the piano and organ, and later in her life worked as the organist at a local church. My older sister eventually studied psychology and worked in human resources for various companies. My younger sister became a medical doctor and opened her own practice as a general practitioner.

I went to the local elementary school around the corner from where we lived, starting in 1961. My teacher did not like the fact that I was left-handed and wanted me to switch, but my parents were strictly opposed. So I continued writing with my left hand, and my poor teacher—like generations of my students later on—had to put up with my terrible handwriting. I went to three different high schools, where I attended the humanistic branch which—besides 8 years of English—included 6 years of Latin and 4 years of ancient Greek. I graduated (Matura) in 1973. From a quite young age on, probably 11 or 12 years, I wanted to study mathematics. I remember only a period of about 1 year when I was 14, in which I wanted to study chemistry, the reason being that I had an excellent chemistry teacher that year. Also, after graduating from high school, I weighed the idea of simultaneously studying mathematics and classical philology, an idea I dropped rather quickly.

You graduated from high school in 1973 and afterwards began to study mathematics at the University of Vienna. Tell us about those years.

I began studying mathematics and physics at the University of Vienna in the Fall of 1973, but after three or four semesters, I stopped taking courses in physics as I found mathematics more appealing. I also took some courses in formal logic, which at that time was taught in a tiny department separate from mathematics. There was a café next to the math, physics, and chemistry building where my friends and I would meet everyday in the morning before deciding which classes to attend and which to skip. This was a group mostly composed of students of mathematics, physics, and chemistry with an occasional economist or statistician thrown in. We studied together, mostly in this or other cafés, and also spent quite a bit of our spare time together. Through one of these friends, I also came to participate in a reading group at the Statistics Department where she was studying. In this reading group, we went through several chapters of Sewastjanow's book on multitype branching processes (Sewastjanow, 1974), which would prove useful later on.

One of the courses I took in the second year was on set-theoretic topology which aroused my interest. I then started to read the book by Willard (1970) on said topic. This led me to talk to Hans-Christian Reichel, a professor specializing in set-theoretic topology, more about this subject. Eventually, he suggested that I write my thesis with him as the advisor, and so it happened. The thesis was about a class of topological spaces called ω_μ-metric spaces, where ω_μ denotes a regular (infinite) ordinal number. These are spaces that can be "metrized" by a "metric" that takes values not in the real line but in a totally ordered abelian group of cofinality ω_μ. Metric spaces correspond to the case $\mu = 0$. In the case $\mu > 0$, these spaces behave very much differently from metrizable spaces. I finished my thesis and passed the final exams ("Rigorosa") in the spring semester of 1978. Since I graduated with a special distinction ("promotio sub auspiciis praesidentis"), the graduation ceremony, which the president of Austria honors with his presence, took place only later in January 1979. Some parts of the thesis were published in Pötscher (1982).

While it was fun to work on the thesis, I soon realized that the topic was very far removed from any form of application and that the field of set-theoretic topology had matured quite a bit, with little dynamic left. So I decided that it was time for a change, and I applied to the Institut für Höhere Studien Wien (IHS) for a scholarship without having much of a plan what I would do there.[1] That said, I still kept some interest in set-theoretic topology for a while and eventually wrote another paper on the subject (Pötscher, 1987b).

2. FROM PURE MATHEMATICS TO ECONOMETRICS AND STATISTICS: INSTITUT FÜR HÖHERE STUDIEN AND TU VIENNA

What was the experience at IHS like?

In the Fall of 1978, I started my 2-year scholarship at IHS. As a scholar, you only needed to attend a few lectures, mostly given by visiting professors, and you had to hand in some paper at the end of each year. Other than that you were pretty much free to do what you pleased. To obtain the scholarship, you had to pass an exam in statistics and mathematics and also to write a paper on some subject relevant to social sciences. So I had to study an elementary statistics book. For the paper, I read Owen's book on game theory (Owen, 1968) and basically summarized some of the chapters in that book. One of the visiting professors in the first semester was Harry Kelejian who gave lectures on econometrics. Having no clue what this was, and given that the class was not scheduled early in the morning, I decided to attend Kelejian's class. This was my first contact with the field of econometrics. I also met Ingmar Prucha at IHS, who was an assistant professor there and who, together with Gerhard Munduch, was responsible for the economic forecast released by IHS every quarter. Later, I also participated in the econometrics research seminar at IHS that you were running together with Ingmar, and this is how we both met. Other visiting professors included James Coleman (sociology), Pradeep Dubey (game theory), Paul Holland (statistics), Paul Newbold (time series), and Jean-Francois Richard (Bayesian econometrics). There was also a class on austromarxism taught by Norbert Leser. From the cohort of scholars that year, I remember Klaus Neusser who later became a professor of Economics at the University of Bern, and more recently, interim director of IHS. Towards the end of the first year, you offered me a position as a "Vertragsassistent" at your department at TU Vienna, which I gladly accepted. The plan was that I would take up this position in the fall while simultaneously finishing the second year at IHS, which was not an unusual

[1] The IHS is a postgraduate institution in Vienna founded by Paul Lazarsfeld and Oskar Morgenstern in the 1960s with the goal to prop up research in the social sciences.

procedure. However, I then had a fall-out with Gerhard Schwödiauer, the director of IHS, and my contract with IHS was swiftly rescinded.

You took up your position in my department at TU Vienna in the Fall of 1979. What was the environment at TU like at that time?

The Department of Econometrics and Operations Research consisted of two groups: your group, the Econometrics group, and the Operations Research group headed by Gustav Feichtinger. The groups operated pretty much independently. The Econometrics group consisted of you as the head, Gerhard Tinter as emeritus professor, a few post-docs and PhD students. One could always tell when Tinter was in his office, because one would smell the cigars he was smoking. At that time, it was considered normal to smoke everywhere, including in your office, even if you shared it with someone else. One of the PhD students (and later post-doc) was Werner Ploberger with whom I shared an office for some time. Werner's desk was interesting: Imagine a small mountain of various pieces of paper of all origins (notes, old newspapers and magazines, copies of papers, empty cigarette packs, etc.) covering the entire desk. Werner would carve out a tiny flat space from this paper mountain where he would do his writing. Our desks were facing each other sharing a common boundary. Over time, the heap of papers on Werner's desk would start to encroach on my desk. When the encroachment became too much, I would push the overflow back towards Werner's desk. As a consequence, some pieces of paper would fall off the desk on Werner's side. He could not care less, and the next day, the cleaning lady would pick up whatever had landed on the floor and throw it into the trash. Occasionally, books that the library had deemed lost surfaced from Werner's paper covered desk. Later on, I shared an office with Erhard Reschenhofer, a PhD student of yours at that time, who much later would become a colleague of mine at my current department.

There was a weekly econometrics seminar in the department with presentations by visitors or by members of our group. Visitors I remember were Ted Hannan, Rudolf Kalman, Jorma Rissanen, Paul Newbold, and Jan Willems. The atmosphere was always relaxed and one could ask all kinds of questions without embarrassing oneself. This was sort of new to me when compared to the atmosphere I had come to know when studying in the Mathematics Department at the University of Vienna. You were also running a research seminar at IHS in which I participated. There I met Walter Krämer after he had joined IHS in the early 1980s. I never wrote a paper with Walter, but once, when he approached me with questions about testing for structural break, I directed him to Werner who I knew had developed some interest in this topic. The ensuing collaboration led to a stream of papers on this topic by Walter and Werner.

TU Vienna managed to hire Pál Révész in the mid 1980s. One of his PhD students was Kurt Hornik who also sometimes came to your seminar. After his PhD, he spent some time at UC San Diego and did some work with Halbert White on the approximation capability of neural nets which has become quite well-known.

At TU you began working on ARMA models. How did you get started in a subject completely new to you?

After I had taken up the position in your department, you suggested that I work on the problem of selecting the lag-orders in autoregressive moving average (ARMA) models. Since I had no formal training in statistics or econometrics apart from what I had learned at IHS, I had a lot of catching up to do. Eventually, I wrote a paper on how a sequence of Lagrange multiplier tests could be used to perform consistent order selection in ARMA models. In the end the paper was split into two parts, one appearing in the *Annals of Statistics*, the other one in *Metrika* (Pötscher, 1983, 1985a). The technical challenge was that, due to non-identifiability issues, standard asymptotic results for Lagrange multiplier tests do not apply. Furthermore, in order to achieve consistency of the resulting order selection procedure, one has to let the significance levels go to zero at a certain rate as the sample size increases, something I had not seen in the literature before. A spin-off of these papers is Bauer, Pötscher, and Hackl (1988), where consistency of model selection by multiple test procedures in general semi-parametric models is investigated.

Around that time the two of us also wrote a paper (Deistler and Pötscher, 1984) on properties of the (Gaussian quasi) likelihood of multivariate ARMA models, where we investigated questions of continuity of the likelihood function and of existence of a maximizer of this function. It may surprise the reader to learn that even for the scalar $ARMA(1,1)$ model there is—to my knowledge—no theoretical result ensuring that a maximizer of the likelihood indeed exists (when the natural parameter space without artificial restrictions is used). One problem here is that the likelihood is not even upper semi-continuous (the parameter space is also not compact, but this could be fixed); see the discussion surrounding Example 1 in Deistler and Pötscher (1984). While this feature of the likelihood does not pose a problem for, for example, consistency results (one simply works with approximate maximum likelihood estimators that maximizes the likelihood function up to a error ε_n that goes to zero as samples size $n \to \infty$), it certainly poses problems for numerical optimization routines. In the paper, we also provide sufficient conditions when a maximizer indeed exists. This paper has gone pretty much unnoticed, perhaps also because we buried it in "*Advances in Applied Probability*," a journal probably not typically read by econometricians and time series analysts.

You went on to write a paper on consistency of maximum likelihood in multivariate ARMA models. Can you tell us about this research?

The consistency result for (quasi)maximum likelihood estimation of multivariate ARMA models in Dunsmuir and Hannan (1976) has a number of problems as you know very well, being one of the coauthors of Deistler, Dunsmuir, and Hannan (1978), a paper which had already fixed some of the gaps. One problem remaining was that the theory in Dunsmuir and Hannan (1976) relies on a condition on the parameter space, called condition (B6) in that paper. While it is easy to show that condition (B6) is satisfied in the *univariate* case for the standard ARMA parameter spaces obtained by prescribing the degrees of the AR- as well as MA-polynomial, there was (and to my knowledge still is) no proof for the standard parameter spaces in the multivariate case. In 1984/85, I thus started to work on consistency proofs for maximum likelihood and related estimators for multivariate ARMA models in a very general setting, including misspecified cases, trying to improve the results in the literature and to close remaining gaps. In particular, I was able to dispense with condition (B6), at least for the case where the data are Gaussian. This was published in Pötscher (1987a). In a sequel, Dahlhaus and Pötscher (1989), we succeeded to achieve the same goal even without the Gaussianity assumption. Due to the attempted generality, unfortunately these two papers are a bit difficult to decipher.

A related paper, but written later in 1990 when I was already at the University of Maryland, is Pötscher (1991b) which reviews some of the results in Pötscher (1987a) in a simpler setting (thus making it more accessible) and then goes on to apply them to (quasi)maximum likelihood estimation of misspecified moving average models when the data generating process has spectral zeros. This research was prompted by reading Tanaka and Satchell (1989).

While at TU, you also started to work with Ingmar Prucha on estimation of nonlinear models in the early 1980s. How did this come about?

Ingmar had left the IHS in the Fall of 1979 for a visiting assistant professor position at the University of Maryland, to where he returned in 1981 as an assistant professor. The academic year 1980/81 he spent at NYU as a visitor. Somehow we had kept in touch after IHS, and I remember that I visited Ingmar when I came to New York during that period. We would also meet whenever Ingmar came to Vienna. At some point, probably in 1981 or 1982, we began to discuss Ingmar's work with Harry Kelejian on (quasi) full information maximum likelihood estimation (including IV estimation derived from approximations to the normal equations) when the errors follow a t-distribution (Prucha and Kelejian, 1984). In that paper, the degrees-of-freedom parameter was assumed to be given, and we began to think about simple ways to estimate this parameter. The motivation was that this would achieve some sort of "partial" adaptation to the tail-thickness of the error-distribution. Adaptive estimation was fashionable at that time (see Bickel, 1982; Manski, 1984). We wrote this up for the linear regression model and sent it to JASA, where it was promptly rejected. So we sent it on to the *Journal of Econometrics*. The referee did not like the paper either, but thankfully the Associate Editor Herman Bierens suggested that we try to generalize this to nonlinear regression and resubmit. This we did, and the paper eventually appeared in the *Journal of Econometrics* in 1986 (Pötscher and Prucha, 1986a).

In the process of generalizing the aforementioned paper to the nonlinear case, we read the pertinent literature on estimation in nonlinear models. We quickly realized that using the uniform law of large numbers (ULLN) by Hoadley (1971) as the basis for proving consistency of estimators in nonlinear econometric models,

FIGURE 1. Ingmar Prucha (left) and Benedikt Pötscher (right) at Ingmar's home in Germantown, MD in 1982.

as done, for example, in White (1980), Domowitz and White (1982), White and Domowitz (1984), and Bates and White (1985), is ill-suited as it rules out typical models and error distributions, something that seems to have gone unnoticed in these papers. We thus wrote a further paper in which we pointed out these limitations, and also came up with a new and more versatile ULLN.[2] We sent the paper to a few people including Peter Phillips for comments. Peter (and some others) wrote back that Donald Andrews had also noticed the problem with White's and coauthors' work and had presented an alternative ULLN. While Andrews' ULLN assumed a Lipschitz-type condition w.r.t. the parameters, our ULLN was based on equicontinuity-type conditions jointly w.r.t. the parameters and the data, so the two ULLNs nicely complemented each other. Both papers were eventually published in *Econometrica* (Andrews, 1987; Pötscher and Prucha, 1989). Newey (1991) then also chimed in with a ULLN closely related to ours. Further extensions and generalizations can be found in Andrews (1992) and Pötscher and Prucha (1994a, 1994b).

In the second half of the 1980s, Herman Bierens kindly approached us with the proposal to write a survey paper on estimation of nonlinear models for *Econometric*

[2] The working paper version Pötscher and Prucha (1986b) also contains a discussion of the fact that some proofs in White (1980), Domowitz and White (1982), and White and Domowitz (1984) are actually inconclusive.

Reviews. In reviewing the literature for this project, we noticed that, for example, different authors such as Bierens, Gallant, and White were using different notions of "weak dependence." We thought that comparing these concepts and providing a unifying approach could be helpful. We hence set out to write two longish papers (Pötscher and Prucha, 1991a, 1991b) that—inter alia—tried to accomplish this goal. These papers provide a self-contained and quite general framework for proving consistency and asymptotic normality of M-estimators in dynamic nonlinear models, including the misspecified case. Several years later, we revised these papers into a book (Pötscher and Prucha, 1997), adding also a new chapter on nonlinear (quasi) full information maximum likelihood estimation.

In the early 1980s, you also wrote a paper in the area of mathematical biology. How did you get involved in this topic?

A friend of mine was writing her PhD thesis with Peter Schuster, a theoretical chemist at the University of Vienna. Schuster had been a close collaborator of Manfred Eigen, a Nobel laureate in chemistry, who was one of the inventors of the theory of hypercycles, a theory of self-replicating molecules, which is important for understanding molecular evolution. My friend's dissertation was in this area, and we frequently had discussions related to the subject. That way I became aware of a paper by Schuster and Sigmund (1984), dealing with order statistics of extinction times of random selection models described by independent linear birth and death processes. These results are of some interest in mathematical biology and theoretical chemistry in the context of Kimura's theory of neutral selection (Kimura, 1982). The question then arose what one can say if the system was not governed by independent birth and death processes, but more generally by independent irreducible multitype branching processes. As mentioned earlier, I had accidentally acquired some knowledge concerning multitype branching processes, which now came in handy and which allowed me to answer the question (see Pötscher, 1985b). Additionally, this paper also extends known results concerning moments of extinction times of irreducible multitype branching processes to the reducible case.

In 1985, you obtained your Habilitation in Stochastics at TU Vienna, which implied that your assistant professor position became permanent.

A Habilitation provides you with the "venia legendi" meaning that you have the right to teach any course falling within the scope of the "venia" (stochastics in my case). At that time, it also gave you immediate tenure as an assistant professor, something that has changed since. Also, a Habilitation was typically expected when you applied for a full professor position in Austria or Germany. This is no longer the case nowadays and the importance of obtaining a Habilitation has significantly decreased. The procedure itself is somewhat similar to a tenure review, but unfortunately is often not overly selective.

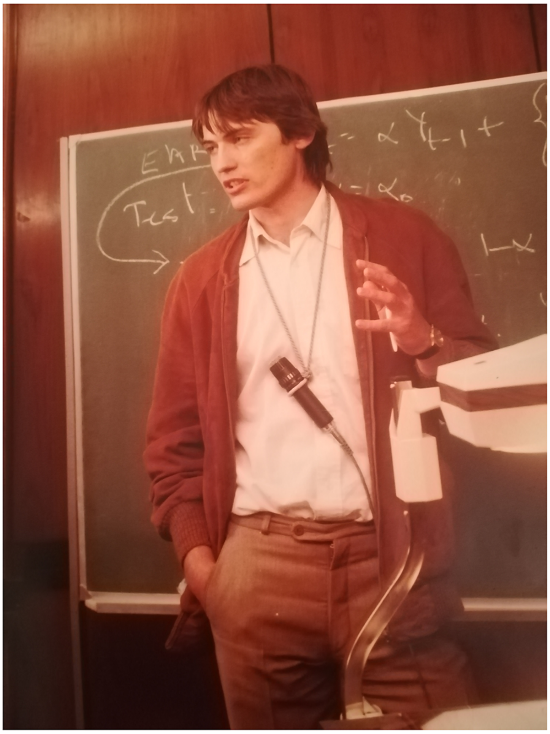

Figure 2. At the ISI-Meeting in Madrid, 1983.

3. U.S. ACADEMIC LIFE: UC SANTA BARBARA, YALE, AND UNIVERSITY OF MARYLAND

You visited the Statistics Department at UC Santa Barbara for a month in April 1986, followed by a 1-year stay as a Visiting Research Scientist in the Department of Statistics at Yale University from May 1986 till May 1987. Your impressions?

Joe Gani was head of the department at UC Santa Barbara at that time, and Ted Hannan had mentioned my name to him. So, thanks to Ted I got the invitation to come to UC Santa Barbara. It was an enjoyable visit, people in the department were friendly, Joe Gani took good care of his visitors, and the campus overlooking the ocean is beautiful. As far as I remember, I used the time during the visit to work on the ULLN-paper with Ingmar mentioned before. My visit to Yale was also facilitated by Ted Hannan, who knew the department quite well as he had earlier spent some time as a visitor there himself. The Statistics Department at Yale was rather small, and was located at Dana House, an old Victorian-style building, across the street from the Economics Department and the Cowles Foundation. The head of the department was Richard Savage, brother to Leonard Jimmy Savage. Other senior faculty members were Frank Anscombe, Joe Hartigan, and David Pollard. Assistant professors were Stephan Morgenthaler, a student of Tukey, who later became professor at EPFL, and Ross Ihaka, who later became very much involved with the R-project. Given the smallness of the department, most of the faculty and graduate students would typically have a brown-bag lunch together in the common room at Dana House. So you quickly got to know people quite well.

I also met Peter Phillips and Don Andrews for the first time in person during that period. I had no obligation to teach as my visit was funded by the Max-Kade Foundation, but nevertheless I gave a mini-course on multivariate ARMA-models. Apart from some graduate students from the Statistics Department, some of Peter's students also sat in on that course. Bruce Hansen, later on, told me that he was one of these students. To my amazement, Peter also came to some of my lectures.

During that period, partly motivated by the growing literature on unit roots, I studied conditions under which BIC-type model selection criteria lead to consistent selection of the model order of nonstationary autoregressive models, and more generally to consistent variable selection in stochastic linear regression models. This resulted in Pötscher (1989). I also continued to work on the survey papers with Ingmar mentioned before, often traveling down to UMD to visit Ingmar. Also, Bauer et al. (1988) dealing with consistency of model selection by multiple test procedures in general semi-parametric models mentioned earlier was finalized during my time at Yale.

My visit to Yale was enjoyable, and definitely was a major factor that I began to contemplate moving to the US for a longer period, which eventually came to pass 2 years later.

After you returned from Yale, what were you working on?

I continued to work on order selection in ARMA-models. All consistency results for order estimators in $ARMA(p,q)$ models obtained through minimizing a criterion such as Schwarz' BIC that were available at that time assumed that the minimization w.r.t. the order (p,q) is effected over the range $0 \leq p \leq P$, $0 \leq q \leq Q$, where P and Q are *given* and represent an upper bound for the minimal true model order (p_0,q_0).[3] This is a somewhat unsatisfactory assumption. I was able to show in Pötscher (1990) that considering *all* $ARMA(p,p)$ models with $0 \leq p < \infty$ and defining $\hat{p}$ as the first "local" minimizer of BIC, results in an estimator $\hat{p}$ that is consistent for $\max(p_0,q_0)$. Consistent estimators for the minimal true model order (p_0,q_0) can then easily be obtained in a second step by minimizing BIC over all model orders $\{(p,q) : 0 \leq p \leq \hat{p}, 0 \leq q \leq \hat{p}\}$. This gives an order selection procedure that does not require specifying an upper bound. The proof rests on a result from Pötscher (1983, Thm. 5.2), extended somewhat in Lemmata 6.1 and 6.2 in Pötscher (1990). Roughly speaking, this result states that when fitting a

[3]The results can be made to hold if P and Q are allowed to increase with sample size slowly enough. However, this does not relieve one from the need to come up with a specification for the upper bounds P and Q.

(misspecified) $ARMA(p,p)$ model to a higher order ARMA-process (or an even more complicated stationary process) by (Gaussian quasi) maximum likelihood, the pseudo-true values (i.e., the maximizers of the asymptotic counterpart to the likelihood) do *not* correspond to $ARMA(r,s)$ structures with both $r < p$ and $s < p$ (more precisely, are not observationally equivalent to such structures).[4] The results in Pötscher (1990) also imply that the (Gaussian quasi) likelihood ratio test for testing $ARMA(p,p)$ versus $ARMA(p+1,p+1)$ has power converging to 1 as $n \to \infty$ (n sample size) even when the data generating process can not be expressed by an $ARMA(p+1,p+1)$ model, that is, even when the alternative is misspecified.[5] For $p = 0$, this shows that the likelihood ratio test for testing $ARMA(0,0)$ versus $ARMA(1,1)$ gives rise to a consistent test of the null of white noise against stationary alternatives. Andrews and Ploberger (1996) further elaborate this point. It is perhaps worth noting that all the aforementioned results are not true in general for order estimation within the class of pure AR models. The reason is that when fitting a misspecified $AR(p)$-model, the pseudo-true value may actually correspond to $AR(r)$ models with $r < p$. A later paper, resting on Pötscher (1990) and exploring the topic further, is Pötscher and Srinivasan (1994).

On a completely different front, I started to work with Immanuel Bomze on evolutionary game theory. I had known Immanuel from the time we both studied mathematics and we frequently discussed mathematical questions. Immanuel was working on evolutionary game theory, and he wanted to discuss some unresolved issues in a paper he was in the process of writing. The overall goal of the paper was to provide a unifying framework for evolutionary game theory. The discussion turned into a collaboration. When we realized that the paper we were working on grew longer and longer, we decided to convert this into a book (Bomze and Pötscher, 1989).

Your work on the effects of model selection on inference also seems to date from this period.

Actually, it was roughly a year later that I started to think about the effects of model selection on subsequent inference. My work on order/model selection mentioned earlier had mainly been concerned with conditions under which the order/model selection procedure would consistently estimate the minimal true order/model, or at least would be conservative in the sense that it selects a true, but possibly non-minimal, order/model asymptotically. In the beginning of 1989, I came across a paper by Ensor and Newton in Biometrika (Ensor and Newton, 1988) where they considered estimating the peak frequency of a spectral density based on autoregressive models when the autoregressive model order is chosen by AIC.

[4] This result was independently also discovered by Kabaila (1983).

[5] While Theorem 5.2 in Pötscher (1983) actually is formulated more generally for fitting $ARMA(p,q)$ models, Lemmata 6.1 and 6.2 in Pötscher (1990) are formulated for $ARMA(p,p)$ models. They can easily be extended to $ARMA(p,q)$ models, using Theorem 2.1 in Pötscher (1991b) (and in the accompanying correction). As a consequence, the statement about the power of the (Gaussian quasi) likelihood ratio test in the main text immediately extends to the case of testing $ARMA(p,q)$ against $ARMA(p+1,q+1)$.

They showed that the so obtained estimator for the peak frequency is consistent. Their proof rested on the particular structure of their estimation problem. However, it was not difficult to see that the result actually has little to do with peak frequency estimation, easily generalizes far beyond, and has a very simple proof: Suppose you are given a set of models and a quantity you want to estimate that is the same for any two correct models. Furthermore, for each model you are given an estimator that is consistent for that quantity provided the model is correct (and fixed). Then using a model selection procedure that asymptotically never selects misspecified models (that is, a "conservative" model selection procedure), followed by the aforementioned estimator obtained from the selected model, results in a consistent estimator for the quantity of interest.[6] In other words, using a conservative model selection procedure does not destroy consistency under the conditions mentioned before. [Note that, in particular, this covers consistent model selection procedures.] But more important than being able to make this rather straightforward observation, was the fact that the paper by Ensor and Newton made me ask what would happen if one goes beyond consistency and looks at the asymptotic distribution of the estimator obtained from such a procedure ("post-model selection estimator"). Hence, I set out to derive the asymptotic distribution of M-estimators that follow a *conservative* (but *not consistent*) model selection procedure. This was set in the framework of nested semi-parametric models and when model selection is through a general-to-specific testing procedure. It turned out that the resulting limiting distributions can be highly non-normal, for example, can be bimodal. This resulted in a paper I sent to Biometrika where it was swiftly rejected. It was then published in *Econometric Theory* (Pötscher, 1991a). I should mention a related paper by Sen (1979) here, which unfortunately I was not aware of when writing my paper, and which thankfully was pointed out to me by Don Andrews.

The results in my paper are fixed-parameter asymptotic results (i.e., results where the data generating process is held fixed while sample size goes to infinity), and hence do not tell the full story. At this point, I had not fully appreciated that the proper set-up was to consider moving-parameter asymptotics to get the full picture of the behavior of post-model-selection estimators. This would come later in joint work with Hannes Leeb (Leeb and Pötscher, 2003, 2005), see also the master's thesis by Richard Nickl (Nickl, 2003). In any case, Pötscher (1991a) and the later work with Hannes Leeb seem to have helped in starting research on inference post model selection, an aspect of statistics which previously had not received the attention it deserved, something Breiman once called a "quiet scandal in statistics." Since then model selection has remained a very active research area until today.

I should add that in the case of *consistent* model selection, at first glance model selection does not seem to have any impact on the asymptotic distributional properties of the post-model-selection estimator: The post-model-selection estimator indeed has the same fixed-parameter asymptotic distribution as the estimator based

[6] See Lemma 2 in Pötscher (1991a) for a formal statement and proof.

on the true minimal model.[7] While this observation is mathematically correct (and is rather trivial), it is, however, totally misleading as already pointed out in Pötscher (1991a). Nevertheless, this observation would—years later—be baptized the "oracle property" by Fan and Li (2001), who marketed it as a desirable property of an estimator. I shall probably return to this later.

Starting August 1989 you became associate professor in the Department of Economics at UMD at College Park, and you were promoted to full professor in 1991.

I had been a visiting associate professor at UMD in the Spring of 1988. During that period, I was asked if I were interested in a permanent position at UMD, and I said yes. The formal recruiting process then took a while to complete (the university had to make sure that there were no better qualified minority candidates, etc.). So I eventually joined UMD as an associate professor in August 1989.

The Economics Department at UMD was a very friendly place at that time. Within the department, I mostly interacted with Harry Kelejian and Ingmar Prucha. Roger Betancourt was also someone I would talk to quite a bit, and my colleagues of Greek extraction (Harris Dellas, Michael Halliassos, and Sam Ouliaris) were also always available for a chat. The department did not have an econometrics seminar when I joined, so we decided to initiate one. I believe the first speaker we invited was Jim Heckman, who drew a considerable audience. From there on, the seminar was firmly established. As far as I know, the seminar is still in existence. I also attended the statistics seminar in the Mathematics Department. This is how I met Ching-Zong Wei, and also David Findley, who came to the seminar from the Census Bureau. With both of them, I would co-author two papers on multistep prediction error estimation (Findley, Pötscher, and Wei (2001, 2004)) years later when I was already back in Vienna. Other participants in that seminar I remember were Benjamin Kedem, Abraham Kagan, and Eric Slud.

With Ingmar I worked on the sequels to our 1989 ULLN paper already mentioned before as well as on the two survey papers on estimation in nonlinear dynamic econometric models (Pötscher and Prucha, 1991a, 1991b). Also the paper on pseudo-maximum likelihood estimation of misspecified moving average models (Pötscher (1991b)) mentioned earlier came into existence during that period at UMD. Towards the end of my period at UMD, Ingmar and I also slowly began to convert the two survey papers into a book (Pötscher and Prucha, 1997).

In summary, I enjoyed my time in the D.C. area quite a bit, both in and out of academia.

While at UMD you received an offer of a full professor position in the Department of Statistics of the University of Vienna, which you eventually accepted. Any thoughts on this decision?

[7] See Lemma 1 in Pötscher (1991a) for a general statement (but note the caveat in the discussion following the lemma). Special cases had been noted before in the literature, see Hannan and Quinn (1979) and Ensor and Newton (1988, Thm. 2.1).

This was a tough one. As I said, I was happy at UMD and saw little reason to move at this point in time. At the same time, it was quite clear that, should I decline the offer from Vienna, chances that another position in statistics/econometrics would open up in Vienna several years later were rather small. So I pondered the question for several weeks, drawing up lists with pros and cons for UMD and Vienna, only to end up with a tie every time. Eventually, I decided to accept the offer from Vienna, but I was not enthusiastic about this decision at all.

4. RETURNING TO VIENNA

What were your experiences after returning to Vienna and joining the University of Vienna in 1991?

Bluntly speaking, bad ones. Within a few weeks of taking up my position in Vienna, it became clear that the department and the school were dysfunctional. Returning from the US equipped with a meritocratic vision, my advent in Vienna unified all the factions in the department into a solid opposition in no time. This made bringing about necessary change very difficult, if not impossible. The atmosphere in the department was what would nowadays be called toxic.

As a consequence, I put most of my efforts into building my own group and isolating it from the rest of the department. Changes in the department and school could only be achieved slowly. A good ally in such undertakings was Egbert Dierker, an economic theorist in the Economics Department of my school, to whom I am most grateful for his steady support through all those years. The situation in my department improved a bit in 2004, when the opposition was weakened when a subgroup split off and joined the Computer Science School. A further important step towards changing the power dynamics in the department occurred in 2009 when I managed to get Hannes Leeb hired from Yale. The situation further improved in 2013 with the hiring of Nikolaus Hautsch. In 2019 and now in 2023, the remaining difficult people retired, and we were able to hire new and pleasant colleagues. It has been a long and thorny road, but finally, we are now an academically as well as socially well-functioning department.

While I had to live in a non-meritocratic and dysfunctional system at the University of Vienna for a considerably long period of time, a system, where no one cared if I did research or not, and where and if I published, this paradoxically also held its advantages: I did not have to churn out one paper after another to secure next year's pay-raise, but I could take time to think about what I believed were important and interesting questions without having to worry if this would result in publications or not.

Given what you just described, did you consider moving elsewhere?

I had some offers from places like the University of Bonn, University of Frankfurt, or TU Dortmund, which I seriously considered. In the end, the offers did not convince me, in particular, as I would have had to move my family, including my two, then young, sons. Now they are grown up, one is a medical doctor and the other one is studying mathematics.

I remember that in the early 1990s you also had an interest in the problem of estimating persistence of a time series which was sort of a hot topic back then. How did this come about?

Michael Hauser, a colleague from WU Vienna, and Erhard Reschenhofer, a colleague in my department, brought me to that topic. In the wake of Nelson and Plosser (1982), in the late 1980s and early 1990s, econometricians and macroeconomists were interested in quantifying persistence of the effect of a shock on the level of a time series y_t such as log real U.S. GDP. Measures of persistence used in the literature were $c_K = \sum_{i=0}^{K} a_i$, the sum of the impulse response coefficients a_i up to a given horizon K, often with $K = \infty$. [Here, a_i are the coefficients in an infinite order *MA* representation of $\Delta y_t - \mu$, where $\mu = E\Delta y_t$.] Various proposals for estimating the persistence measures were made in the literature, the pro and cons were religiously discussed, and the estimates obtained by the various authors often did not agree: Some papers reported estimates for c_∞ close to 0.5, others reported values as large as 1.5. The proposed procedures ranged from estimating the persistence measure based on *ARMA* models (Campbell and Mankiw, 1987a, 1987b) or *ARFIMA* models (Diebold and Rudebusch, 1989) to nonparametric procedures (Cochrane, 1988). Given that the spectral density of Δy_t (more precisely of $\Delta y_t - \mu$) at frequency zero is $f(0) = \sigma^2 c_\infty^2 / 2\pi$, where σ^2 is the variance of the white noise driving the infinite order *MA* representation of Δy_t, estimation of the persistence measure essentially boils down to estimation of the spectral density at frequency zero (and of σ^2), a well-studied problem in time series analysis. Why would different authors then come up with wildly different estimates? This question was the starting point for our paper.

In our paper, Hauser, Pötscher, and Reschenhofer (1999), we made a number of points: One point was that (i) an argument in Cochrane (1988) leveled against the use of *ARMA* models for persistence estimation does not hold up to scrutiny, and (ii) that nonparametric estimators for $f(0)$ (used in Cochrane, 1988) are likely to be downward biased if Δy_t has to be demeaned (which usually is the case) and if, at the same time, the truncation lag parameter is chosen too large. The latter observation is also in line with Monte Carlo findings in Campbell and Mankiw (1987b).

A second point in our paper concerned methodological issues when fitting *ARMA* models that are related to nonidentifiability issues. In particular, we explained that *ARMA*(p,q)-based estimators of $f(0)$ can be heavily distorted if the order (p,q) is not judicially chosen: While it is not surprising that one can expect to suffer from misspecification bias if (p,q) are chosen too small, it is less obvious that choosing the order (p,q) too large (in the sense that both p and q are too large) can also lead to serious distortions in the *ARMA*(p,q)-based estimate $\hat{f}(0)$. As discussed in our paper, this is related to the fact that a pair of near-cancelling *AR*- and

MA-roots close to 1 (or −1) will then often be present in the estimated $ARMA(p,q)$ structure from which $\hat{f}(0)$ is calculated (in view of results in Hannan, 1982), which will lead to distorted estimates $\hat{f}(0)$. These observations help to explain findings in Campbell and Mankiw (1987a), namely that $ARMA(p,q)$-based estimates $\hat{f}(0)$ are close to zero for some choices of (p,q), while they are not so for other choices of (p,q), although both choices lead to essentially the same goodness-of-fit. We also used the methodological insights to explain some of the numerical findings in Christiano and Eichenbaum (1990).

A third point the paper made was that using *ARFIMA* models for estimating persistence measures as in Diebold and Rudebusch (1989) is fundamentally flawed: The estimated fractional differencing parameter $\hat{d}$ will almost always be either positive or negative (and not equal to zero), even in case $d = 0$. As a consequence, the *ARFIMA*-based estimated persistence $\hat{c}_\infty$ is almost always either ∞ (if $\hat{d} > 0$) or 0 (if $\hat{d} < 0$), even if c_∞ is finite and positive (which can only occur if $d = 0$). This obviously renders the *ARFIMA*-based estimate $\hat{c}_\infty$ pretty much useless in the case where c_∞ is finite and positive.[8] To be fair, Diebold and Rudebusch (1989) mainly consider the persistence measures c_K for finite K. However, the observation just made for $\hat{c}_\infty$ obviously has implications for the *ARFIMA* -based estimates $\hat{c}_K$ of the finite-horizon persistence measures c_K: The estimate $\hat{c}_K$ will be severely downward (upward) "biased" if $\hat{d} < 0$ ($\hat{d} > 0$) as the finite-horizon quantities converge to the infinite-horizon ones as $K \to \infty$. This undesirable effect, in fact, already shows up for relatively small values of K as can be seen in numerical examples.

Using what we had learned from the analysis, we then went on to properly reanalyze the data with *ARMA* models as well as with the nonparametric proposal, and we were able to reconcile the persistence estimates from both approaches to a large extent.

How was that paper received?

We tried some good journals, like *Journal of Econometrics*, *Journal of Applied Econometrics*, and *Review of Economic Studies*, but each time ran into heavy opposition. Especially, our criticism of the use of ARFIMA models for persistence estimation was apparently not appreciated by some of the referees at any of the journals we tried. Time was also against us, as in all the years, the paper was sitting at journals only awaiting to be rejected, the literature moved on and away from univariate persistence measures. In the end, we decided to send the paper to a lower ranked journal, where it eventually was published in 1999.

Your paper on ill-posed estimation problems in *Econometrica* in 2002 also seems to have grown out of your work on persistence estimation?

[8] If one believes that $d > 0$ ($d < 0$) holds, then one knows that $c_\infty = \infty$ ($c_\infty = 0$), and there is then actually no need to estimate this quantity.

This is definitely true. Already while working on Hauser et al. (1999), it became clear to me that a main reason why estimating $f(0)$ or c_∞ had caused so much controversy in the literature was that it is an ill-posed estimation problem. Except one is willing to make strong assumptions on the feasible set of spectral densities (such as imposing *quantitative* bounds on derivatives of f), I realized that any estimator of $f(0)$ or c_∞ necessarily has unbounded worst case risk (for the usual loss-functions), and that any confidence set for these quantities necessarily is uninformative (in a sense that can be made precise) with large probability. Also, tests of the hypothesis $f(0) = 0$ (or $f(0) = c$) necessarily have power not exceeding size; see also Blough (1988) and Faust (1996) for a similar point. In that sense, estimation of $f(0)$ is strictly speaking a futile exercise, and should better be replaced by estimating the integral of f over a frequency band around zero. Interestingly, even restricting the set of feasible spectral densities to a small set such as the set of all $ARMA(1,1)$ spectral densities turns out not to remove all of the ill-posedness. These results are given as an example of a more general theory of ill-posed estimation problems in Pötscher (2002). A short expository section sketching these issues can also be found in Hauser et al. (1999).

5. MODEL SELECTION AND INFERENCE

One of your major contributions has been your work on post-model-selection estimation, which brought new insights, some of which were surprising to many of us. As already described before, you began to look into this problem with your 1991 Econometric Theory paper. It then took several years until you took up this line of research again, this time together with Hannes Leeb. Can you comment on this?

The hiatus is easily explained: My work with Ingmar on the two survey articles, and subsequently on the book, consumed a lot of my time and energy after 1990/91. Also the move back to Vienna, coupled with the difficult situation in the department here in Vienna, kept me from returning to the problem of model selection and inference. Then in 1998, Hannes Leeb joined my group as a post-doc, and we started to collaborate on post-model-selection estimation, a collaboration that has been most fruitful and which I enjoyed quite a bit over the years. We started by considering a collection of nested linear regression models, where model selection is conducted through a general-to-specific testing procedure. In a first step (Leeb and Pötscher, 2003), we studied the finite-sample as well as the asymptotic distribution of the resulting post-model-selection estimator allowing for moving-parameter asymptotics (i.e., asymptotics where the parameters governing the data generating process are allowed to vary with the sample size); we considered unconditional as well as conditional distributions of the post-model-selection estimator (conditional on the selected model). These distributions are highly non-normal. We also derived the corresponding finite-sample distributions. Allowing for moving-parameter asymptotics was an important improvement over the results in Pötscher (1991a), where only fixed parameter asymptotics was considered, since this allows one to asymptotically capture the effects of underestimation of the model order, something that gets lost under a fixed parameter asymptotic regime.

Another important question is how to construct confidence sets post model selection: If one could find an estimator of the finite-sample distribution of the post-model-selection estimator (centered at the true parameter value and appropriately scaled) that is consistent *uniformly* w.r.t. the parameters, such an estimator could be used to easily generate a confidence set that is uniformly asymptotically valid. Now, while it is easy to construct consistent estimators for the finite-sample distribution, we soon realized that *uniformly* consistent estimators typically do *not* exist, not even locally uniform ones. Of course, this does not rule out the existence of valid confidence sets, but it shows that the usual line of reasoning does not work. This impossibility result was published in Leeb and Pötscher (2006a) for the conditional distribution function of the post-model-selection estimator, and for the unconditional one in Leeb and Pötscher (2008a). [Both these papers make use of a general framework for impossibility results laid out in Leeb and Pötscher (2006b).] Basically, this impossibility result seems to force one to construct valid confidence sets derived from the worst-case distribution: Suppose for simplicity that the parameter θ is scalar and that the finite sample distribution $F_{n,\theta}$ of $| n^{1/2}(\hat{\theta}_n - \theta) |$ can be determined, where $\hat{\theta}_n$ denotes the post-model-selection estimator (the scaling by $n^{1/2}$ being irrelevant here). If $c_n(\theta)$ denotes the $(1-\alpha)$-quantile of $F_{n,\theta}$, then the set of θ's satisfying $| n^{1/2}(\hat{\theta}_n - \theta) | \leq \sup_\theta c_n(\theta)$ is obviously a $(1-\alpha)$-confidence set for θ (see Leeb and Pötscher, 2017 for more comments on this issue). Of course, this is much more burdensome, as one needs to be able to find $c_n(\theta)$ and $\sup_\theta c_n(\theta)$. There is a uniformly valid asymptotic version of this approach (see Andrews and Guggenberger, 2009), but the main ideas and issues are the same.

An easily accessible overview of the consequences of model selection on properties of post-model-selection estimators and associated inference procedures is given in Leeb and Pötscher (2005). Both the case of consistent as well as of conservative model selection is discussed in a simple framework. In particular, the somewhat counterintuitive and—at that time—little appreciated fact, that textbook inference procedures (i.e., procedures acting as if the selected model were given prior to seeing the data) typically suffer even more under consistent model selection than under conservative model selection, is explained in this paper.

I probably should also add that the problems arising with inference post model selection similarly affect inference post model averaging (see Pötscher, 2006).

Penalized maximum likelihood estimators like the LASSO came to prominence in the late 1990s. There is obviously a close connection to model selection. Can you comment on this?

Conceptually, there is not much difference. Model selection based on *AIC* or *BIC* can be seen as penalized maximum likelihood using an l_0-penalty, whereas

FIGURE 3. Rudy Beran, Xu Cheng, and Benedikt Pötscher (left to right) at the Rathausplatz in Vienna after the 2008 Workshop "Current Trends and Challenges in Model Selection and Related Areas."

LASSO uses an l_1-penalty. A similar comment applies to methods such as SCAD or the adaptive LASSO, etc., which differ from LASSO only in the choice of the penalty function. Much of what we had done for the more classical forms of model selection can be carried over to such penalized maximum likelihood methods (see, e.g., Leeb and Pötscher, 2006b; Pötscher and Leeb, 2009; Pötscher and Schneider, 2009, 2010, 2011).

One of the methods that rose to prominence was the SCAD-estimator introduced by Fan and Li (2001). They showed, that—when properly tuned—the SCAD estimator performs consistent model selection (i.e., asymptotically finds the zeros in the true parameter vector) and that the SCAD-estimators of the nonzero components of the parameter vector have the same asymptotic distribution, as if maximum likelihood assuming knowledge of the position of the true zeros were performed. They dubbed this the "oracle property." [For consistent model selection followed by maximum likelihood type estimation, this "oracle property" had already been noticed in Lemma 1 in Pötscher (1991a), but already came equipped with a warning about the downsides of this property.] The paper by Fan and Li triggered a considerable flow of papers enthusiastically establishing the "oracle property" for a wide variety of models and penalized maximum likelihood procedures. While the "oracle property" is a mathematically correct statement, it is completely misleading: The "oracle property" seems to entail that given a (regular) model and a submodel defined by zero restrictions, one can construct an estimator, that is asymptotically as good as the restricted maximum likelihood estimator if the restrictions are satisfied, and as good as the unrestricted maximum likelihood estimator if the restrictions are violated. However, this is at odds with Hajék-LeCam style asymptotic efficiency of (unrestricted) maximum likelihood, and points to the fact that estimators satisfying the "oracle property" must have a dark side. The dark side is that such estimators are necessarily non-regular, and that their actual behavior is not at all well described by *pointwise* asymptotic results such as the "oracle property." In a sense, Hodges' estimator, invented in 1951 to serve as a counterexample and well-known to have undesirable properties, can be seen as the first example of an estimator possessing the "oracle property." Based on these insights, we showed in Leeb and Pötscher (2008b) (see also Appendix C of Leeb and Pötscher, 2005) that estimators satisfying the "oracle property" (more generally, "sparse" estimators) necessarily have dismal finite sample properties. For example, the ratio of such an estimator's maximal risk to the maximal risk of the maximum likelihood estimator computed from the largest model diverges to infinity as sample size becomes large. Also, confidence sets based on estimators having the "oracle property" (more generally, "sparse" estimators) have undesirable properties as shown in Pötscher (2009). In summary, the "oracle property" is in fact the opposite of a desirable property of an estimator, despite statements to the contrary in the literature.[9] The irony here is that this was already understood in the mid 20th century in mathematical statistics. Unfortunately, it did not prevent the hype surrounding the "oracle property", an entirely misguided concept, 50 years later. There is a nice quote from Hajék concerning the issues involved which is reproduced in Leeb and Pötscher (2005, 2008b).

The Fan and Li paper on the "oracle property" was published in JASA, your critique in *Journal of Econometrics*. Didn't you try JASA first?

Yes, indeed, we did and our paper was promptly rejected. We received two reports, one only a short paragraph stating that we had not understood Fan and Li's paper, the other one having 14 pages or so (more pages than our paper had) doing a lot of singing and dancing, including a Monte Carlo study, etc. We appealed, the appeal was sent to the same referees, and the outcome was the same. So we moved on to the *Journal of Econometrics* where we were successful. Ironically, a few years later Hannes and I received an e-mail from the then-editor of JASA to the effect that he had read some of our papers on model selection, found them interesting, and was wondering whether we had ever considered sending our work to JASA.

[9] Such "sparse" estimators typically also have uniform convergence rates *slower* than $n^{-1/2}$ (despite typically having pointwise convergence rates equal to $n^{-1/2}$), see Pötscher, 2009, Pötscher and Leeb, 2009, Pötscher and Schneider, 2009, 2011.

In the 2010s you seem to have slowed down in working on model selection. Were there any particular reasons for that?

I think there were a few factors involved. First, one of my weaknesses is that I lose interest in a topic rather fast once I have understood the basic structure. Second, the literature moved into directions I did not find that exciting. For example, Berk et al. (2013) changed the target of estimation from the parameters in the true model to something one could call surrogate parameters, namely to the "parameters" describing the best approximation of the true data generating process within the selected model. These surrogate parameters are data-dependent and also depend on the model selection method chosen. I have my reservations about the interpretability of this notion (see Section 2.3 in Bachoc, Leeb, and Pötscher, 2019 as well as Leeb, Pötscher, and Ewald, 2015 for more discussion). With this change of target, the problem of inference post model selection becomes much easier, essentially because bias is eliminated. Perhaps, this was the reason why subsequent to Berk et al. (2013) a considerable amount of literature followed their example. Another strand in the literature was concerned with "debiased" or "desparsified" versions of various post model selection estimators: In the context of selecting submodels from a given finite-dimensional standard linear model, this is (uniformly asymptotically) equivalent to doing ordinary least squares on the largest model. Intuitively, what happens here is the following: Suppose one is interested in the coefficient of the first regressor ("treatment effect"), the other regressors representing controls. Now, with the just-mentioned methods, controls are dropped only (i) if they do not significantly contribute to the explanation of the dependent variable *and* (ii) if they are (nearly) orthogonal to the first regressor (treatment).[10] But under the (near) orthogonality of treatment variable and the controls implied by (ii), dropping or not dropping of those controls does not (uniformly asymptotically) affect the estimator. In other words, these methods perform variable selection only if variable selection is essentially irrelevant as it (nearly) has no effect on the estimator anyway. These methods become nontrivial and more interesting when one allows for high-dimensional models, but then one has to impose all sorts of sparsity assumptions and the like, something I could never get excited with. Third, other projects had started to attract my interest.

6. NONPARAMETRIC INDIRECT INFERENCE, HETEROSKEDASTICITY AND AUTOCORRELATION ROBUST INFERENCE

You had some excellent PhD students. One of them was Richard Nickl, who now is a professor in the Department of Pure Mathematics and Mathematical Statistics at the University of Cambridge. Could you describe your collaboration with Nickl?

[10] For comparison, note that more classical model selection procedures drop controls whenever (i) applies.

Figure 4. Benedikt Pötscher, Manfred Deistler, and Wolfgang Scherrer (left to right) at a dinner celebrating Manfred's retirement in 2009.

I only had a handful of PhD students, but some of them were truly excellent. One of them was Richard Nickl, who started out as a student of economics and philosophy at the University of Vienna. At that time, bachelor students of economics had to take a class on linear models from me. He quickly got hooked, and eventually wrote his Master's thesis in economics (Nickl, 2003) with me, followed by a PhD in statistics under my supervision. To improve the mathematical education he had received while studying economics, I sent Richard to the Math Department to take courses in functional analysis, measure theory, topology, and so on. He picked up all this material amazingly fast. Furthermore, I also made him attend a summer school on the theory of empirical processes which featured Evarist Giné and Sara van de Geer as instructors. All this paid off handsomely.

While still doing his Master in economics, Richard visited the Philosophy Department at Duke University for a semester or so in 2003. I suggested to get in touch with Ronald Gallant and George Tauchen in the Economics Department, which he did. This brought Richard in contact with indirect inference. Essentially, all of the work on indirect inference at that time was set in a framework of a parametric auxiliary model. Now, for the indirect inference estimator to be efficient, a crucial assumption is that the auxiliary model is correctly specified, which is a quite strong and questionable assumption if one maintains a parametric auxiliary model. Upon his return to Vienna, Richard wanted to do his PhD work on indirect inference, but allowing for nonparametric auxiliary models in order to make the assumption of correct specification more palatable. After closely studying the nonparametric literature, it soon became clear that a number of technical tools had to be developed before the problem of nonparametric indirect inference could be successfully tackled in the generality intended. This resulted in two papers that would make up Richard's PhD thesis.

The first paper, Nickl and Pötscher (2007), provides (bracketing as well as sup-norm) metric entropy bounds for classes of "smooth" functions (such as bounded subsets of Besov or related spaces) when the functions are defined over potentially unbounded domains in Euclidean space. [For bounded domains, classical results can be found in Kolmogorov and Tihomirov (1961) and Birman and Solomjak (1967).] The metric entropy bounds obtained in our paper are then used to provide conditions when the function classes under consideration are (universal and uniform) Donsker. The paper seems to have filled a gap in the literature, and is now being used quite a bit in nonparametric statistics and theoretical machine learning.

The second paper, Nickl (2007), considers nonparametric density estimation from i.i.d. data by the maximum likelihood method when the "parameter" space is (an appropriate subset of) a Sobolev ball. That paper provides a Donsker-type result for the "nonparametric maximum likelihood" process (i.e., for the process obtained by replacing in the empirical process the empirical measure by the measure corresponding to the nonparametric density estimator) where the process is indexed by a suitable subset of a Sobolev space. This is a remarkable result which made quite a splash in the nonparametrics/empirical process community. It can be seen as an asymptotic normality result for the nonparametric maximum likelihood estimator.

Equipped with what we had learned in the process, we then were able to provide the desired generalization of indirect inference to the case where a nonparametric auxiliary model is employed, and to establish that the resulting indirect inference estimator is asymptotically normal with variance–covariance matrix equal to the Cramér–Rao bound. The auxiliary estimators used in this paper (Nickl and Pötscher, 2010) are spline projection estimators. Similar results can be obtained if nonparametric maximum likelihood estimators are used instead (Gach and Pötscher, 2011). The results are given for univariate i.i.d. data, and unfortunately we never got around to extending the results to multivariate dependent data, mainly because we both were distracted by other projects.

Another one of your excellent PhD students was David Preinerstorfer, with whom you have worked extensively on heteroskedasticity and autocorrelation robust testing. How did this come about?

David Preinerstorfer had come to the statistics program from psychology, and eventually finished with master degrees in both fields. In 2011, he joined my group as a PhD student. Like most of my PhD students, David also attended courses in the Math Department to further his mathematical education. Around that time, Martin Wagner drew my attention to a paper by Perron and Ren (2011), in which the authors declared impossibility results like the ones in my Econometrica 2002 paper as irrelevant. Naturally, I read their paper with great interest. It turned out that their arguments were unconvincing for several reasons; in particular, the Perron and Ren (2011) paper is concerned with autocorrelation robust testing, something my Econometrica 2002 paper does not touch on at all.[11] But to the credit of Perron and Ren (2011) I have to say that their paper started me to think about properties of autocorrelation robust tests. It became clear to me quite soon that these tests often suffer from severe size distortions and/or lack of power and that one could establish this theoretically. Since David had just joined my group as a PhD student at that time, I put him on this problem. Amazingly fast he came up with interesting results. That there could be size distortions had been noted in simulation studies before, and some fixes like fixed-bandwidth versions of autocorrelation robust tests had been suggested. While these fixes help to ameliorate the problem to some extent, they do not solve it. What we found is that—depending on the interaction of the regressor matrix, the hypothesis to be tested, and the assumed model for the correlation structure of the errors in the regression—the size of such tests can be one. In fact, size can be one in some situations, regardless of the choice of the critical value. And for many more situations, the usual critical value obtained from asymptotic theory is much too small and leads to a test with size much larger than the nominal significance level. These issues are discussed at length in the first of our papers, Preinerstorfer and Pötscher (2016). This paper, together with Preinerstorfer (2017) and Preinerstorfer and Pötscher (2017), formed the PhD thesis of David.

The before mentioned work laid the foundation for Pötscher and Preinerstorfer (2018) where we were able to provide sufficient conditions under which the size of autocorrelation robust tests can be bounded by the prescribed significance level α through an appropriate choice of the critical value (“size control”). We were also able to show that the sufficient conditions are generically satisfied. In a follow-up paper, Pötscher and Preinerstorfer (2019), we showed that for certain autocorrelation robust tests designed for trend testing, such size control is not possible and these test have size one. Similar phenomena as for autocorrelation robust tests also appear in the context of heteroskedasticity robust tests, and similar solutions for size control exist. This is discussed in Pötscher and Preinerstorfer (2021). That bootstrapping does not solve the problem is demonstrated theoretically and numerically in Pötscher and Preinerstorfer (2023). There are many open questions in this area, which we plan to pursue in the future.

Supervising Richard and David has been a privilege and a pleasure. In both cases, the supervision quickly turned into a collaboration from which I have learned a lot.

[11] Section 3.2.2 in Preinerstorfer and Pötscher (2016) contains a detailed discussion of why Perron and Ren (2011) are off-base.

FIGURE 5. Delivering the Laudatio at Manfred Deistler's promotion to Doctor honoris causa at TU Dortmund in 2016.

7. EDITORIAL ACTIVITY AND THE ACADEMIC PROCESS

You have served on the editorial board of a few journals, like *Econometric Theory*, *Econometric Reviews*, *Journal of Econometrics*, and *Journal of Statistical Planning and Inference*. Any comments?

I have been associated with *Econometric Theory* since 1988, first as an associate editor and then as a co-editor. I hope I could help the journal to grow and to maintain rigorous standards. I want to take this opportunity to thank Peter Phillips for the collaboration on this project, a collaboration I always enjoyed and in the course of which I have learned a lot about the publishing process. I have been an Associate Editor of the *Journal of Econometrics* from 1999 till 2013 when I left the editorial board because I was not aligned with the changes that occurred at that time. I have also been co-editor of *Econometric Reviews* from 2003 until 2005.

I joined the board of the *Journal of Statistical Planning and Inference* in 2012 upon an invitation by Holger Dette, who had taken over the journal as an editor and had plans to improve the quality of the journal. Unfortunately, the hoped for improvement did not materialize, and in 2022 I decided to pull out.

What I have learned from my work on editorial boards is that a considerable percentage of papers submitted are not really ready for submission, and would profit from a few more months of polishing. Unfortunately, our profession seems to incentivize quantity over quality. It also puts too much weight on the perceived status of the journal a paper is published in, when it is the quality of the paper's content that should matter. And the criteria for publication in so-called top journals do not always seem to relate to quality alone.

Talking about quality standards in scientific publishing, you have served the community also by critical commentary on papers and books, which I believe is important for keeping the profession healthy. How did you decide whether or not to comment on a publication?

I guess there are many contributing factors. It had to be a topic that interested me sufficiently, and the paper needed to have the potential to generate confusion or a wrong sense of security in the econometrics community. And I needed to have time to write up my criticism.

These critical commentaries were not always well-received, I suppose?

Unfortunately, but not unexpectedly, this was often the case. Criticizing someone's paper openly seems to be considered an unkind act by a considerable portion of the community for some reasons. I believe this is misguided, and seems to be the result of a false sense of camaraderie in the profession. Quite to the contrary, science is the "culture of doubt" as Feynman once put it, and critical evaluation of the literature is central to the healthiness of the profession. Unfortunately, publishing critical commentary is often an uphill battle, but it can be won, at least sometimes. In the process, one does not only make friends, but this can be survived, and is a price worth paying for. The defensive culture of stifling criticism and dissent in the academic profession is worrying, and can even lead to censorship.

You once told me a curious story about an experience you had with JASA. Can you share this story here?

Yes, indeed this was some story! I had been asked by Karen Kafadar, then associate book review editor at JASA, if I were willing to review Hamilton's time series book (Hamilton, 1994). I agreed, wrote the review, and sent it to Kafadar. She had only a few minor suggestions for grammatical changes, which I dutifully implemented, and the review was then accepted for publication in JASA. Some time later, I received the galley proofs in the mail. The galleys were accompanied by a copy of the manuscript. Looking at the manuscript, which had my name on it, I noticed that it was typeset differently than the original manuscript I had submitted.

I was wondering why someone had taken the pain to retype the manuscript, and started to read the galleys. After a few sentences, I noticed that this was not the review I had submitted. So I checked the enclosed manuscript and realized that this was not a retyped version of my manuscript, but that someone had drastically changed and shortened my review, in a way that portions of the new review did not even make sense. That person never had asked for my consent, but had been bold enough to put my name on this concoction. Not surprisingly, this illegally-tampered-with review was more lenient on Hamilton's book when compared to my review.

Naturally, I was upset. I wrote an e-mail to the then book review editor explaining the situation, that this was censorship, and that I would take legal action if the censored version of the review and not my original version would appear. I copied this mail also to the upper echelons of the American Statistical Association. The book review editor wrote back that he was traveling and that he would check on this matter once he returned home. In his subsequent mail, he then admitted to having rewritten the review himself, because it supposedly was too long, and that supposedly he had forgotten to ask for my permission. An e-mail exchange ensued where in every mail I told him that this was outright censorship and that I demanded publication of my review in its original form. He came up with 1,001 reasons why he had to rewrite my review, at some point essentially admitting that a reason was the negative tone of my report. I told him that he could have sent my review out to referees if he thought that I was erring in my evaluation, but that he could not alter my review without my permission, especially after it had been accepted for publication. I insisted that this was outright censorship and that my demands had to be met, otherwise legal action would be the consequence. The last line of defense by the book review editor towards the end then was that my review supposedly was too long. While this obviously was an excuse, and he should have told me that before accepting the review for publication, I decided to take out two or three rather irrelevant sentences from my review in order to break the stalemate and to bring this matter to an end. The so-shortened review then finally was published in JASA in 1996.

In this case, the attempt to stifle criticism has not been successful. However, we do not know how often similar attempts are successful, and how often self-censorship is at work. Also, sometimes more subtle forms of suppressing unwanted views are at work, for example, in the refereeing process. Given that even in low-stake situations, like the ones described above, attempts to stifle criticism and dissent occur, what then really becomes worrying is to imagine what happens in high-stake cases where the research in question has monetary or political implications, for example, in pharmaceutical or climate related research.

8. MISCELLANEA

Can you talk about what the major influences shaping your scientific career have been?

I believe growing up in a family where academic work was highly regarded was an important factor in shaping my views of academia and my value system. Of course, you yourself were an important influence. Much of what I know about time series and econometrics I learned from you. You also made me read Ted Anderson's time series book (Anderson 1971) when I joined TU Vienna, which was, and still is, a masterpiece, and which instilled quite some admiration for Anderson. I believe one can hardly overrate his contributions to econometrics and statistics.

In my early years, Ted Hannan was also a guiding star and a helping hand. I owe much to Ingmar Prucha, the collaboration with him taught me a lot, and he was instrumental in bringing me to UMD. I am also grateful to Hannes Leeb for our long-time collaboration from which I have profited enormously. My excellent PhD students Richard Nickl and David Preinerstorfer also deserve mention here; advising them was a true privilege and taught me a lot. Over the years, I also benefitted from numerous discussions with Rudy Beran, whom I had met at the University of Heidelberg in 1993 where we shared an office for several weeks.

You have been in academia now for more than 45 years. Looking back, what has changed?

The field of econometrics has become much more diverse and more technical. Back in the days, it was mostly static or dynamic linear models. Technology has changed, of course, in my early years, revising a paper was done by cutting and pasting in the literal sense. Top5-itis was less of an issue back then, I believe. The political climate at universities was not as intolerant as, unfortunately, it is nowadays. Freedom of speech is definitely under attack in academia and elsewhere.

Bibliography of Benedikt M. Pötscher

Books

1989

1. *"Game Theoretical Foundations of Evolutionary Stability"* (with I.M. Bomze), Lecture Notes in Economics and Mathematical Systems, Vol. 324, Springer-Verlag, 1989.

1997

2. *"Dynamic Nonlinear Econometric Models: Asymptotic Theory"* (with I.R. Prucha), Springer-Verlag, 1997.

Articles

1982

3. "Some Results on ω_μ-Metric Spaces," *Annales Universitatis Scientiarum Budapestinensis*, 25, 1982, 3–18. (Corrections, ibid., 28, 1985, 283.)

4. "Identifiability and Inference in ARMA-Systems" (with M. Deistler and W. Ploberger), in: O.D. Anderson (ed.), *Time Series Analysis: Theory and Practice 2*, Amsterdam: North-Holland, 1982, 43–60.

1983

5. "Order Estimation in ARMA-Models by Lagrangian Multiplier Tests," *Annals of Statistics*, 11, 1983, 872–885. (Corrections, ibid., 12, 1984, 785.)

1984

6. "Uniqueness of the Transfer Function of Linear Systems from Input-Output Data" (with M. Deistler and J. Schrader), *Metrika*, 31, 1984, 157–181.

7. "The Behaviour of the Likelihood Function for ARMA-Models" (with M. Deistler), *Advances in Applied Probability*, 16, 1984, 843–865.

1985

8. "The Behaviour of the Lagrangian Multiplier Test in Testing the Orders of an ARMA-Model," *Metrika*, 32, 1985, 129–150.

9. "Moments and Order Statistics of Extinction Times in Multitype Branching Processes and Their Relation to Random Selection Models," *Bulletin of Mathematical Biology*, 47, 1985, 263–272.

10. "Ordnungsschätzung in ARMA-Modellen," *Österreichische Zeitschrift für Statistik und Informatik*, 15, 1985, 62–74.

1986

11. "A Class of Partially Adaptive One-Step M-Estimators for the Nonlinear Regression Model with Dependent Observations" (with I.R. Prucha), *Journal of Econometrics*, 32, 1986, 219–251.

12. "Introduction to Special Issue 'Modeling Problems in Econometrics'" (with M. Deistler), *Applied Mathematics and Computation*, 20, 1986, 197–199.

1987

13. "A Generalization of Urysohn's Metrization Theorem and its Set-Theoretic Consequences," *Studia Scientiarum Mathematicarum Hungarica*, 22, 1987, 457–461.

14. "Convergence Results for Maximum Likelihood Type Estimators in Multivariable ARMA-Models," *Journal of Multivariate Analysis*, 21, 1987, 29–52.

15. "Comments on 'Towards a Theory of Point Optimal Testing' by M.L. King," *Econometric Reviews*, 6, 1987–88, 235–239.

1988

16. "Model Selection by Multiple Test Procedures" (with P. Bauer and P. Hackl), *Statistics*, 19, 1988, 39–44.

17. "Discriminating between Two Spectral Densities in Case of Replicated Observations" (with E. Reschenhofer), *Journal of Time Series Analysis*, 9, 1988, 221–224.

1989

18. “A Uniform Law of Large Numbers for Dependent and Heterogeneous Data Processes” (with I.R. Prucha), *Econometrica*, 57, 1989, 675–683.

19. “Convergence Results for Maximum Likelihood Type Estimators in Multivariable ARMA Models II” (with R. Dahlhaus), *Journal of Multivariate Analysis*, 30, 1989, 241–244.

20. “Distribution of the Coates-Diggle Test Statistic in Case of Replicated Observations” (with E. Reschenhofer), *Statistics*, 20, 1989, 417–421.

21. “Model Selection under Nonstationarity: Autoregressive Models and Linear Stochastic Regression Models,” *Annals of Statistics*, 17, 1989, 1257–1274.

22. “Solution to Problem 88.3.1: Pseudo Orthogonality and Granger Causality in Dynamic Data,” *Econometric Theory*, 5, 1989, 456–458.

23. “Generic Uniform Laws of Large Numbers” (with I.R. Prucha), in: S. Kotz and N.L. Johnson (eds.), *Encyclopedia of Statistical Sciences*, Suppl.Vol., J. Wiley & Sons, 1989, 67–69.

1990

24. “Reply to Comments on ‘Solution to Problem 88.3.1: Pseudo Orthogonality and Granger Causality in Dynamic Data,”’ *Econometric Theory*, 6, 1990, 291–292.

25. “Estimation of Autoregressive Moving Average Order Given an Infinite Number of Models and Approximation of Spectral Densities,” *Journal of Time Series Analysis*, 11, 1990, 165–179.

1991

26. “Effects of Model Selection on Inference,” *Econometric Theory*, 7, 1991, 163–185.

27. “Letter to the Editor on ‘The Impact of Model Selection on Inference in Linear Regression’ by C.M. Hurvich and C.L. Tsai,” *American Statistican*, 45, 1991, 171–172.

28. “Basic Structure of the Asymptotic Theory in Dynamic Nonlinear Econometric Models, Part I: Consistency and Approximation Concepts” (with I.R. Prucha), *Econometric Reviews*, 10, 1991, 125–216. (Reprinted in: *Nonlinear Models*, Vol. I, H. Bierens and A.R. Gallant (eds.), The International Library of Critical Writings in Econometrics, Edward Elgar Publ. Ltd., Cheltenham, UK.)

29. “Basic Structure of the Asymptotic Theory in Dynamic Nonlinear Econometric Models, Part II: Asymptotic Normality” (with I.R. Prucha), *Econometric Reviews*, 10, 1991, 253–326. (Reprinted in: *Nonlinear Models*, Vol. I, H. Bierens and A.R. Gallant (eds.), The International Library of Critical Writings in Econometrics, Edward Elgar Publ. Ltd., Cheltenham, UK.)

30. “Reply to Discussion of ‘Basic Structure of the Asymptotic Theory in Dynamic Nonlinear Econometric Models, Part I: Consistency and Approximation Concepts’ and ‘Basic Structure of the Asymptotic Theory in Dynamic Nonlinear Econometric Models, Part II: Asymptotic Normality”’ (with I.R. Prucha), *Econometric Reviews*, 10, 1991, 349–358.

31. “Noninvertibility and Pseudo Maximum Likelihood Estimation of Misspecified ARMA Models,” *Econometric Theory*, 7, 1991, 435–449. (Corrections, ibid., 10, 1994, 811.)

1992

32. “Batch Sequential Design for a Nonlinear Estimation Problem” (with W. Müller), in: V. Fedorov et al. (eds.), *Model Oriented Data-Analysis*, Physica-Verlag, 1992, 77–87.

1993

33. “Letter to the Editor on ‘On the Definition of an Evolutionarily Stable Strategy in the Playing the Field Model’ by V.P. Crawford” (with I.M. Bomze), *Journal of Theoretical Biology*, 161, 1993, 405.

34. “Solution to Problem 92.3.5: Efficiency of Maximum Likelihood,” *Econometric Theory*, 9, 1993, 535–536.

1994

35. “Generic Uniform Convergence and Equicontinuity Concepts for Random Functions: An Exploration of the Basic Structure” (with I.R. Prucha), *Journal of Econometrics*, 60, 1994, 23–63.

36. “A Comparison of Order Estimation Procedures for ARMA Models” (with S. Srinivasan), *Statistica Sinica*, 4, 1994, 29–50.

37. “On the Formulation of Uniform Laws of Large Numbers: A Truncation Approach” (with I.R. Prucha), *Statistics*, 25, 1994, 343–360.

38. “On Gagnon’s Critique of ARMA Models for Real GNP Growth” (with M.A. Hauser and E. Reschenhofer), *Economic Notes*, 23, 1994, 125–129.

1995

39. “Comment on ‘Adaptive Estimation in Time Series Regression Models’ by D.G. Steigerwald,” *Journal of Econometrics*, 66, 1995, 123–129.

40. “Comment on ‘Assessment and Propagation of Model Uncertainty’ by D. Draper,” *Journal of the Royal Statistical Society*, 57, Series B, 1995, 86.

41. “Comment on ‘The Effect of Model Selection on Confidence Regions and Prediction Regions’ by P. Kabaila,” *Econometric Theory*, 11, 1995, 550–559.

42 “Comment on ‘Model Uncertainty, Data Mining and Statistical Inference’ by C. Chatfield,” *Journal of the Royal Statistical Society*, 158, Series A, 1995, 461.

1996

43. “Optional Skipping of Martingale Differences and Related Sequences,” *Metron*, 54, 1996, 31–36.

1998

44. “The Distribution of Estimators after Model Selection: Large and Small Sample Results” (with A.J. Novak), *Journal of Statistical Computation and Simulation*, 60, 1998, 19–56.

45. "Solution to Problem 97.4.2: Asymptotic Properties of the Least-Squares Estimator of the Variance in a Linear Model," *Econometric Theory*, 14, 1998, 527–533.

1999

46. "Measuring Persistence in Aggregate Output: ARMA Models, Fractionally Integrated ARMA Models and Nonparametric Procedures" (with M.A. Hauser and E. Reschenhofer), *Empirical Economics*, 24, 1999, 243–269.

47. "Solution to Problem 98.1.2: Lower Eigenbound for AR(1) Disturbance Covariance Matrix," *Econometric Theory*, 15, 1999, 156.

2000

48. "Basic Elements of Asymptotic Theory" (with I.R. Prucha), in: B. Baltagi (ed.), *A Companion to Theoretical Econometrics*, Basil Blackwell, 2000, 201–229.

2001

49. "The Variance of an Integrated Process Need Not Diverge to Infinity, and Related Results on Partial Sums of Stationary Processes" (with H. Leeb), *Econometric Theory*, 17, 2001, 671–685.

50. "Uniform Convergence of Sample Second Moments of Families of Time Series Arrays" (with D.F. Findley and C.Z. Wei), *Annals of Statistics*, 29, 2001, 815–838.

2002

51. "Lower Risk Bounds and Properties of Confidence Sets for Ill-Posed Estimation Problems with Applications to Spectral Density and Persistence Estimation, Unit Roots, and Estimation of Long Memory Parameters," *Econometrica*, 70, 2002, 1035–1065.

2003

52. "The Finite-Sample Distribution of Post-Model-Selection Estimators, and Uniform Versus Non-Uniform Approximations" (with H. Leeb), *Econometric Theory*, 19, 2003, 100–142.

53. "Some Problems in Statistical Inference Following Model Selection," in: *Proceedings of the 13th IFAC Symposium on System Identification*, SYSID-2003, Rotterdam, August 27–29, 2003.

2004

54. "Contributions to Econometrics, Time Series Analysis, and Systems Identification: A Festschrift in Honor of Manfred Deistler," Guest Editorial, (with I.R. Prucha), *Journal of Econometrics*, 118, 2004, 1–5.

55. "Modelling of Time Series Arrays by Multistep Prediction or Likelihood Methodes" (with D.F. Findley and C.Z. Wei), *Journal of Econometrics*, 118, 2004, 151–187.

56. “Nonlinear Functionals and Convergence to Brownian Motion: Beyond the Continuous Mapping Theorem,” *Econometric Theory*, 20, 2004, 1–22.

2005

57. “Model Selection and Inference: Facts and Fiction” (with H. Leeb), *Econometric Theory*, 21, 2005, 21–59.

2006

58. “Performance Limits for Estimators of the Risk or Distribution of Shrinkage-Type Estimators, and Some General Lower Risk-Bound Results” (with H. Leeb), *Econometric Theory*, 22, 2006, 69–97. (Corrections, ibid., 24, 581–583.)

59. “Can One Estimate the Conditional Distribution of Post-Model-Selection Estimators?” (with H. Leeb), *Annals of Statistics*, 34, 2006, 2554–2591.

60. “The Distribution of Model Averaging Estimators and an Impossibility Result Regarding Its Estimation,” in: H.-C. Ho, C.-K. Ing, and T.-L. Lai (eds.), *Time Series and Related Topics: In Memory of Ching-Zong Wei. IMS Lecture Notes and Monograph Series*, Vol. 52, 2006, 113–129.

2007

61. “Bracketing Metric Entropy Rates and Empirical Central Limit Theorems for Function Classes of Besov- and Sobolev-Type” (with R. Nickl), *Journal of Theoretical Probability*, 20, 2007, 177–199.

62. “The ET-Interview: Professor Manfred Deistler,” *Econometric Theory*, 23, 2007, 709–746.

2008

63. “Recent Developments in Model Selection and Related Areas” (with H. Leeb), *Econometric Theory*, 24, 2008, 319-322.

64. “Can One Estimate the Unconditional Distribution of Post-Model-Selection Estimators?” (with H. Leeb), *Econometric Theory*, 24, 2008, 338–376.

65. “Sparse Estimators and the Oracle Property, or the Return of Hodges’ Estimator” (with H. Leeb), *Journal of Econometrics*, 142, 2008, 201–211.

66. “Comment on ‘Approximating Data’ by P.L. Davies,” *Journal of the Korean Statistical Society*, 37, 2008, 227–229.

2009

67. “On the Distribution of the Adaptive LASSO Estimator” (with U. Schneider), *Journal of Statistical Planning and Inference*, 139, 2009, 2775–2790.

68. “On the Distribution of Penalized Maximum Likelihood Estimators: The LASSO, SCAD, and Thresholding” (with H. Leeb), *Journal of Multivariate Analysis*, 100, 2009, 2065–2082.

69. “Confidence Sets Based on Sparse Estimators are Necessarily Large,” *Sankhya*, 71-A, 2009, 1–18.

70. “Model Selection” (with H. Leeb), in: T.G. Andersen, R.A. Davis, J.-P. Kreiss, and T. Mikosch (eds.), *Handbook of Financial Time Series*, Springer-Verlag, 2009, 889–925.

2010

71. “Confidence Sets Based on Penalized Maximum Likelihood Estimators” (with U. Schneider), *Electronic Journal of Statistics*, 4, 2010, 334–360.

72. “Efficient Simulation-Based Minimum Distance Estimation and Indirect Inference” (with R. Nickl), *Mathematical Methods of Statistics*, 19, 2010, 327–364.

2011

73. “Non-Parametric Maximum Likelihood Density Estimation and Simulation-Based Minimum Distance Estimators” (with F. Gach), *Mathematical Methods of Statistics*, 20, 2011, 288–326.

74. “Distributional Results for Thresholding Estimators in High-Dimensional Gaussian Regression Models” (with U. Schneider), *Electronic Journal of Statistics*, 5, 2011, 1876–1934.

2013

75. “On the Order of Magnitude of Sums of Negative Powers of Integrated Processes,” *Econometric Theory*, 28, 2013, 642–658.

2015

76. “On Various Confidence Intervals Post-Model-Selection” (with Hannes Leeb and Karl Ewald), *Statistical Science*, 30, 2015, 216–227.

2016

77. “On Size and Power of Heteroskedasticity and Autocorrelation Robust Tests” (with David Preinerstorfer), *Econometric Theory*, 32, 2016, 261–358.

2017

78. “On the Power of Invariant Tests for Hypotheses on a Covariance Matrix” (with David Preinerstorfer), *Econometric Theory*, 33, 2017, 1–68.

79. “Testing in the Presence of Nuisance Parameters: Some Comments on Tests Post-Model-Selection and Random Critical Values” (with Hannes Leeb), in: S. Ejaz Ahmed (ed.), *Big and Complex Data Analysis: Methodology and Applications*, Springer, 2017, 69–82.

2018

80. “Controlling the Size of Autocorrelation Robust Tests” (with David Preinerstorfer), *Journal of Econometrics*, 207, 2018, 406–431.

2019

81. “Valid Confidence Intervals for Post-Model-Selection Predictors” (with Francois Bachoc and Hannes Leeb), *Annals of Statistics*, 47, 2019, 1475–1504.

82. “Discussion on ‘Model Confidence Bounds for Variable Selection’ by Yang Li, Yuetian Luo, Davide Ferrari, Xiaonan Hu, and Yichen Qin” (with Hannes Leeb and Danijel Kivaranovic), *Biometrics*, 75, 2019, 407–410.

83. “Further Results on Size and Power of Heteroskedasticity and Autocorrelation Robust Tests, with an Application to Trend Testing” (with David Preinerstorfer), *Electronic Journal of Statistics*, 13, 2019, 3893–3942.

2023

84. “How Reliable are Bootstrap-based Heteroskedasticity Robust Tests?” (with David Preinerstorfer), *Econometric Theory*, 39, 2023, 789–847.

2024+

85. “Valid Heteroskedasticity Robust Testing” (with David Preinerstorfer), April 2021. *Econometric Theory*. Published online by Cambridge University Press: 11 September 2023, pp. 1–53.

86. “Introduction to the Special Issue ‘High-Dimensional Time Series in Macroeconomics and Finance’” (with L. Sögner and M. Wagner), *Econometrics*, 12(1), 2024, 6.

87. “A Comment on: ‘A Modern Gauss-Markov Theorem’” (with David Preinerstorfer), May 2022. *Econometrica*, 92, 2024, 913–924.

Unpublished Working Papers

88. “Non-numerical Distance Functions in Topology, a Survey and Bibliography Furnished with Short Notes and Comments on Most of the Papers Listed Herein” (with H.-C. Reichel and R. Schmid-Zartner), Dept. of Mathematics Working Paper, University of Vienna, 1985.

89. “Consistency in Nonlinear Econometrics: A Generic Uniform Law of Large Numbers and Some Comments on Recent Results” (with I.R. Prucha), Dept. of Economics Working Paper No. 86-9, University of Maryland, 1986.

90. “Solution to Problem 93.1.3: Efficiency as Correlation,” manuscript, 1993.

91. “Learning Trigonometric Polynomials from Random Samples and Exponential Inequalities for Eigenvalues of Random Matrices” (with K. Gröchenig and H. Rauhut), manuscript, 2007, https://arxiv.org/abs/math/0701781.

92. “A Modern Gauss-Markov Theorem? Really?” (with David Preinerstorfer), February 2022, https://arxiv.org/abs/2203.01425.

REFERENCES

Anderson, T. W. (1971). *The statistical analysis of time series*. Wiley.

Andrews, D. W. K. (1987). Consistency in nonlinear econometric models: A generic uniform law of large numbers. *Econometrica*, 55, 1465–1471.

Andrews, D. W. K. (1992). Generic uniform convergence. *Econometric Theory*, 8, 241–257.

Andrews, D. W. K., & Guggenberger, P. (2009). Hybrid and size-corrected subsampling methods. *Econometrica*, 77, 721–762.

Andrews, D. W. K., & Ploberger, W. (1996). Testing for serial correlation against an ARMA(1, 1) process. *Journal of the American Statistical Association*, 91, 1331–1342.

Bachoc, F., Leeb, H., & Pötscher, B. M. (2019). Valid confidence intervals for post-model-selection predictors. *Annals of Statistics*, 47, 1475–1504.
Bates, C. E., & White, H. (1985). A unified theory of consistent estimation for parametric models. *Econometric Theory*, 1, 151–178.
Bauer, P., Pötscher, B. M., & Hackl, P. (1988). Model selection by multiple test procedures. *Statistics*, 19, 39–44.
Berk, R., Brown, L., Buja, A., Zhang, K., & Zhao, L. (2013). Valid post-selection inference. *Annals of Statistics*, 41, 802–837.
Bickel, P. J. (1982). On adaptive estimation. *Annals of Statistics*, 10, 647–671.
Birman, M., & Solomjak, M. Z. (1967). Piecewise polynomial approximations of functions of classes W_p^α. *Matematicheskii Sbornik*, 73(115), 331–355.
Blough, S. R. (1988). On the impossibility of testing for unit roots and cointegration in finite samples. *Working Paper*. Johns Hopkins University.
Bomze, I. M., & Pötscher, B. M. (1989). *Game theoretical foundations of evolutionary stability*. Lecture Notes in Economics and Mathematical Systems, 324. Springer-Verlag.
Campbell, J. Y., & Mankiw, N. G. (1987a). Are output fluctuations transitory? *Quarterly Journal of Economics*, 102, 857–880.
Campbell, J. Y., & Mankiw, N. G. (1987b). Permanent and transitory components in macroeconomic fluctuations. *American Economic Review Papers and Proceeedings*, 77, 111–117.
Christiano, L. J., & Eichenbaum, M. (1990). Unit roots in real GNP: Do we know and do we care? *Carnegie-Rochester Conference Series on Public Policy*, 32, 7–62.
Cochrane, J. H. (1988). How big is the random walk in GNP? *Journal of Political Economy*, 96, 893–920.
Dahlhaus, R., & Pötscher, B. M. (1989). Convergence results for maximum likelihood type estimators in multivariable ARMA models II. *Journal of Multivariate Analysis*, 30, 241–244.
Deistler, M., Dunsmuir, W., & Hannan, E. J. (1978). Vector linear time series models: Corrections and extensions. *Advances in Applied Probability*, 10, 360–372.
Deistler, M., & Pötscher, B. M. (1984). The behaviour of the likelihood function for ARMA models. *Advances in Applied Probability*, 16, 843–866.
Diebold, F., & Rudebusch, G. (1989). Long memory and persistence in aggregate output. *Journal of Monetary Economics*, 24, 189–209.
Domowitz, I., & White, H. (1982). Misspecified models with dependent observations. *Journal of Econometrics*, 20, 35–58.
Dunsmuir, W., & Hannan, E. J. (1976). Vector linear time series models. *Advances in Applied Probability*, 8, 339–364.
Ensor, K. B., & Newton, H. J. (1988). The effect of order estimation on estimating the peak frequency of an autoregressive spectral density. *Biometrika*, 75, 587–589.
Fan, J., & Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*, 96, 1348–1360.
Faust, J. (1996). Near observational equivalence and theoretical size problems with unit root tests. *Econometric Theory*, 12, 724–731.
Findley, D. F., Pötscher, B. M., & Wei, C.-Z. (2001). Uniform convergence of sample second moments of families of time series arrays. *Annals of Statistics*, 29, 815–838.
Findley, D. F., Pötscher, B. M., & Wei, C.-Z. (2004). Modeling of time series arrays by multistep prediction or likelihood methods. *Journal of Econometrics*, 118, 151–187.
Gach, F., & Pötscher, B. M. (2011). Nonparametric maximum likelihood density estimation and simulation-based minimum distance estimators. *Mathematical Methods of Statistics*, 20, 288–326.
Hamilton, J. D. (1994). *Time series analysis*. Princeton University Press.
Hannan, E. J. (1982). Testing for autocorrelation and Akaike's criterion. *Journal of Applied Probability*, 19, 403–412.
Hannan, E. J., & Quinn, B. G. (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society, Series B*, 41, 190–195.

Hauser, M., Pötscher, B. M., & Reschenhofer, E. (1999). Measuring persistence in aggregate output: ARMA models, fractionally integrated ARMA models and nonparametric procedures. *Empirical Economics*, 24, 243–269.

Hoadley, B. (1971). Asymptotic properties of maximum likelihood estimators for the independent not identically distributed case. *Annals of Mathematical Statistics*, 42, 1977–1991.

Kabaila, P. (1983). Parameter values of ARMA models minimising the one-step-ahead prediction error when the true system is not in the model set. *Journal of Applied Probability*, 20, 405–408.

Kimura, M. (1982). *Molecular evolution, protein polymorphism and the neutral theory*. Springer-Verlag.

Kolmogorov, A. N., & Tihomirov, V. M. (1961). ε-entropy and ε-capacity of sets in functional spaces. *American Mathematical Society Translations*, 17, 277–364.

Leeb, H., & Pötscher, B. M. (2003). The finite-sample distribution of post-model-selection estimators, and uniform versus non-uniform approximations. *Econometric Theory*, 19, 100–142.

Leeb, H., & Pötscher, B. M. (2005). Model selection and inference: Facts and fiction. *Econometric Theory*, 21, 21–59.

Leeb, H., & Pötscher, B. M. (2006a). Can one estimate the conditional distribution of post-model-selection estimators? *Annals of Statistics*, 34, 2554–2591.

Leeb, H., & Pötscher, B. M. (2006b). Performance limits for estimators of the risk or distribution of shrinkage-type estimators, and some general lower risk-bound results. *Econometric Theory*, 22, 69–97. (Correction, *ibid*., 24, 581–583).

Leeb, H., & Pötscher, B. M. (2008a). Can one estimate the unconditional distribution of post-model-selection estimators? *Econometric Theory*, 24, 338–376.

Leeb, H., & Pötscher, B. M. (2008b). Sparse estimators and the oracle property, or the return of Hodges' estimator. *Journal of Econometrics*, 142, 201–211.

Leeb, H., & Pötscher, B. M. (2017). Testing in the presence of nuisance parameters: Some comments on tests post-model-selection and random critical values. In S. Ahmed (Eds.), *Big and complex data analysis* (pp. 69–82). Contributions to Statistics. Springer.

Leeb, H., Pötscher, B. M., & Ewald, K. (2015). On various confidence intervals post-model-selection. *Statistical Science*, 30, 216–227.

Manski, C. F. (1984). Adaptive estimation of nonlinear regression models. *Econometric Reviews*, 3, 145–210.

Nelson, C. R., & Plosser, C. I. (1982). Trends and random walks in macroeconomic time series: Some evidence and implications. *Journal of Monetary Economics*, 10, 139–162.

Newey, W. K. (1991). Uniform convergence in probability and stochastic equicontinuity. *Econometrica*, 59, 1161–1167.

Nickl, R. (2003). *Asymptotic distribution theory of post-model-selection maximum likelihood estimators*. Master's thesis, University of Vienna.

Nickl, R. (2007). Donsker-type theorems for nonparametric maximum likelihood estimators. *Probability Theory and Related Fields*, 138, 411–449. (Erratum, *ibid*., 141, 331–332).

Nickl, R., & Pötscher, B. M. (2007). Bracketing metric entropy rates and empirical central limit theorems for function classes of Besov- and Sobolev-type. *Journal of Theoretical Probability*, 20, 177–199.

Nickl, R., & Pötscher, B. M. (2010). Efficient simulation-based minimum distance estimation and indirect inference. *Mathematical Methods of Statistics*, 19, 327–364.

Owen, G. (1968). *Game theory*. W. B. Saunders Co.

Perron, P., & Ren, L. (2011). On the irrelevance of impossibility theorems: The case of the long-run variance. *Journal of Time Series Econometrics*, 3, 3(3).

Pötscher, B. M. (1982). Some results on ω_μ-metric spaces. *Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös Nominatae. Sectio Mathematica*, 25, 3–18. (Correction, *ibid.*, 28, 283).

Pötscher, B. M. (1983). Order estimation in ARMA-models by Lagrangian multiplier tests. *Annals of Statistics*, 11, 872–885. (Correction, *ibid*., 12, 785).

Pötscher, B. M. (1985a). The behaviour of the Lagrangian multiplier test in testing the orders of an ARMA-model. *Metrika*, 32, 129–150.

Pötscher, B. M. (1985b). Moments and order statistics of extinction times in multitype branching processes and their relation to random selection models. *Bulletin of Mathematical Biology*, 47, 263–272.

Pötscher, B. M. (1987a). Convergence results for maximum likelihood type estimators in multivariable ARMA models. *Journal of Multivariate Analysis*, 21, 29–52.

Pötscher, B. M. (1987b). A generalization of Urysohn's metrization theorem and its set-theoretic consequences. *Studia Scientiarum Mathematicarum Hungarica*, 22, 457–461.

Pötscher, B. M. (1989). Model selection under nonstationarity: Autoregressive models and stochastic linear regression models. *Annals of Statistics*, 17, 1257–1274.

Pötscher, B. M. (1990). Estimation of autoregressive moving-average order given an infinite number of models and approximation of spectral densities. *Journal of Time Series Analysis*, 11, 165–179.

Pötscher, B. M. (1991a). Effects of model selection on inference. *Econometric Theory*, 7, 163–185.

Pötscher, B. M. (1991b). Noninvertibility and pseudo-maximum likelihood estimation of misspecified ARMA models. *Econometric Theory*, 7, 435–449. (Corrigendum, *ibid*., 10, 811).

Pötscher, B. M. (2002). Lower risk bounds and properties of confidence sets for ill-posed estimation problems with applications to spectral density and persistence estimation, unit roots, and estimation of long memory parameters. *Econometrica*, 70, 1035–1065.

Pötscher, B. M. (2006). The distribution of model averaging estimators and an impossibility result regarding its estimation. *IMS Lecture Notes - Monograph Series*, 52, 113–129.

Pötscher, B. M. (2009). Confidence sets based on sparse estimators are necessarily large. *Sankhya*, 71-A, 1–18.

Pötscher, B. M., & Leeb, H. (2009). On the distribution of penalized maximum likelihood estimators: The LASSO, SCAD, and thresholding. *Journal of Multivariate Analysis*, 100, 2065–2082.

Pötscher, B. M., & Preinerstorfer, D. (2018). Controlling the size of autocorrelation robust tests. *Journal of Econometrics*, 207, 406–431.

Pötscher, B. M., & Preinerstorfer, D. (2019). Further results on size and power of heteroskedasticity and autocorrelation robust tests, with an application to trend testing. *Electronic Journal of Statistics*, 13, 3893–3942.

Pötscher, B. M., & Preinerstorfer, D. (2021). Valid heteroskedasticity robust testing. *Econometric Theory*. Published online by Cambridge University Press: 11 September 2023, 1–53.

Pötscher, B. M., & Preinerstorfer, D. (2023). How reliable are bootstrap-based heteroskedasticity robust tests? *Econometric Theory*, 39, 789–847.

Pötscher, B. M., & Prucha, I. R. (1986a). A class of partially adaptive one-step *M*-estimators for the nonlinear regression model with dependent observations. *Journal of Econometrics*, 32, 219–251.

Pötscher, B. M., & Prucha, I. R. (1986b). Consistency in nonlinear econometrics: A generic uniform law of large numbers and some comments on recent results. *Working Paper No 86-9.* Department of Economics, University of Maryland.

Pötscher, B. M., & Prucha, I. R. (1989). A uniform law of large numbers for dependent and heterogeneous data processes. *Econometrica*, 57, 675–683.

Pötscher, B. M., & Prucha, I. R. (1991a). Basic structure of the asymptotic theory in dynamic nonlinear econometric models. I. Consistency and approximation concepts. *Econometric Reviews*, 10, 125–216.

Pötscher, B. M., & Prucha, I. R. (1991b). Basic structure of the asymptotic theory in dynamic nonlinear econometric models II. Asymptotic normality.. *Econometric Reviews*, 10, 253–357.

Pötscher, B. M., & Prucha, I. R. (1994a). Generic uniform convergence and equicontinuity concepts for random functions: An exploration of the basic structure. *Journal of Econometrics*, 60, 23–63.

Pötscher, B. M., & Prucha, I. R. (1994b). On the formulation of uniform laws of large numbers: A truncation approach. *Statistics*, 25, 343–360.

Pötscher, B. M., & Prucha, I. R. (1997). *Dynamic nonlinear econometric models: Asymptotic theory*. Springer-Verlag.

Pötscher, B. M., & Schneider, U. (2009). On the distribution of the adaptive LASSO estimator. *Journal of Statistical Planning and Inference*, 139, 2775–2790.
Pötscher, B. M., & Schneider, U. (2010). Confidence sets based on penalized maximum likelihood estimators in Gaussian regression. *Electronic Journal of Statistics*, 4, 334–360.
Pötscher, B. M., & Schneider, U. (2011). Distributional results for thresholding estimators in high-dimensional Gaussian regression models. *Electronic Journal of Statistics*, 5, 1876–1934.
Pötscher, B. M., & Srinivasan, S. (1994). A comparison of order estimation procedures for ARMA models. *Statistica Sinica*, 4, 29–50.
Preinerstorfer, D. (2017). Finite sample properties of tests based on prewhitened nonparametric covariance estimators. *Electronic Journal of Statistics*, 11, 2097–2167.
Preinerstorfer, D., & Pötscher, B. M. (2016). On size and power of heteroskedasticity and autocorrelation robust tests. *Econometric Theory*, 32, 261–358.
Preinerstorfer, D., & Pötscher, B. M. (2017). On the power of invariant tests for hypotheses on a covariance matrix. *Econometric Theory*, 33, 1–68.
Prucha, I. R., & Kelejian, H. H. (1984). The structure of simultaneous equation estimators: A generalization towcards nonnormal disturbances. *Econometrica*, 52, 721–736.
Schuster, P., & Sigmund, K. (1984). Random selection – a simple model based on linear birth and death processes. *Bulletin of Mathematical Biology*, 46, 11–17.
Sen, P. K. (1979). Asymptotic properties of maximum likelihood estimators based on conditional specification. *Annals of Statistics*, 7, 1019–1033.
Sewastjanow, B. A. (1974). *Verzweigungsprozesse*, vol. 34 of *Mathematische Lehrbücher und Monographien, II. Abteilung: Mathematische Monographien*. Akademie-Verlag, Berlin.
Tanaka, K., & Satchell, S. E. (1989). Asymptotic properties of the maximum-likelihood and nonlinear least-squares estimators for noninvertible moving average models. *Econometric Theory*, 5, 333–353.
White, H. (1980). Nonlinear regression on cross-section data. *Econometrica*, 48, 721–746.
White, H., & Domowitz, I. (1984). Nonlinear regression with dependent observations. *Econometrica*, 52, 143–161.
Willard, S. (1970). *General topology*. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont.