Predicting the past

Valedictory address

Delivered in abbreviated form
at Tilburg University
on June 14, 2013

ISBN: 978-94-6167-154-7
Homepage: www.janmagnus.nl

www.tilburguniversity.edu

Predicting the past

Pablo Picasso: The artist and his model. From ***Le Chef-d'Oeuvre Inconnu****, text by Honoré de Balzac, illustrations by Pablo Picasso, Paris: Ambroise Vollard, 1931.*

Prelude

Mr. Rector Magnificus, colleagues, former and current students, family, friends, and others (enemies perhaps) who have taken the trouble to come and attend this lecture.

Ever since I started teaching, now more than forty years ago, I realized — based on my own experience as a student — that a good lecture should have two ingredients. First, it should not assume that students remember anything from the previous lecture; and second, that it should take into account that students have an attention span of ten minutes maximum. After ten minutes there should be an anecdote or something else to make the students feel relaxed and keep them awake. In the beginning of my career I did not yet follow these rules and consequently many of my students did fall asleep, which is a little disturbing, especially when they sit on the front row and snore loudly.

The current lecture has the advantage that there was no previous lecture, so you don't have to remember anything. Regarding the second ingredient I have brought Pieter Nieuwint along. The locals know of course that Pieter was for many years the driving force behind the famous KUB cabaret. For today's visitors from outside Tilburg it may be useful to explain that Pieter was for many years the driving force behind the famous KUB cabaret. My lecture today consists of 16 little blocks, each of which is independent of the other, and Pieter will keep count on where we are, so that at each point you know how many more blocks there are to go.

Predicting the past. Surely that should be easy. The past has no secrets. It is the future that is dark and obscure. It may come as a shock to many of you to learn that an econometrician's job is to predict the past. To understand what happened in the past, why it happened, how it happened, which model could best describe it, this is all an essential — arguably *the* essential — part of the art of economic modeling and estimation.

For example, if I conduct a survey among one thousand academics asking them their income in the previous year and how much of this they consumed and how much they saved, then I can estimate the relationship between consumption and income for this group, the so-called consumption function. Then, if you give me the income of some other academic in the previous year, I can predict his or her consumption in the previous year and also the uncertainty of my prediction. This is predicting the past, and it is not easy.

But suppose you want me to predict this academic's *next* year's consumption, which is the sort of thing often asked from an econometrician. This is even more difficult, because in addition to predicting the past, I now also need two further ingredients, namely the academic's next year's income (which of course we don't know) and an assumption that the changes in the economy do not affect the consumption function. Our ignorance about these two further ingredients makes predicting the future a hazardous exercise.

Nevertheless, as Søren Kierkegaard wrote, while life can only be understood backwards, it must be lived forwards. So, this challenge needs to be faced. But not today. In today's lecture, I shall attempt to understand and describe my own past and tell you a little bit of the things that have occupied me during all these years. Only at the end of the lecture will I attempt to predict my future.

My academic life consisted of three main periods: my student days at the University of Amsterdam, my junior academic life at the London School of Economics, and my senior academic life here in Tilburg.

Matrix calculus

My first two scientific interests were maximum likelihood estimation and matrix calculus. My interest in matrix calculus was inspired by Heinz Neudecker, who taught a graduate class on the subject around 1973 which I attended. I was puzzled and fascinated, and subsequently Heinz and I worked together for fifteen years, culminating in many papers and eventually, in 1988, in a monograph, the publication of which is still one of the highlights of my academic life.

What is matrix calculus? Most people are familiar with high-school calculus: the derivative of x^2 is $2x$, for example. But suppose now that this function f depends on several variables, say $x_1, \ldots, x_n$. And suppose that you have not one function, but several, say $f_1, \ldots, f_m$. Then we move from high-school calculus to vector calculus. The derivative is now a matrix, that is an array of functions, in this case with m rows and n columns.

So far there are no new problems. Vector calculus is well established and well understood. But suppose next we want to know the second derivative, that is the derivative of the derivative, the so-called Hessian. Then we need to differentiate the first derivative, which is a matrix. How do we do that? Or suppose the functions that we start off with are arranged in a matrix, not in a vector, for example the inverse function $F(X)=X^{-1}$.Then we require a theory of matrix calculus.

The theory that Heinz and I developed is based on two pillars. First, we employ differentials, not derivatives, and this has the advantage that the dimensions do not grow as we differentiate. They always stay the same: the differential of a matrix is a matrix of the same dimension. Second, we transform matrix functions to vector functions (and in the end we transform them back), so that we can employ the existing theory of vector calculus.

This may sound easy (at least to the mathematically inclined amongst you), and in the end it *is* quite easy, but it took us fifteen years to put it all together in a satisfactory framework. Not only did we have to understand and develop calculus in the context of differentials, but we also had to fight with colleagues who had invented their own ad-hoc methods, not mathematically correct because not based on the underlying concept of derivative. Even today this battle is not completely won, and I recently called these other definitions and methods 'derisatives' (instead of derivatives), not causing much amusement in these circles.

At the end of this long period I had become an expert in matrices and matrix calculus, and I could have stayed with this subject for the rest of my career. Maybe I should have done that. But I did not see any major new challenges in this field. It was time to change subject.

Maximum likelihood and asymptotic theory

Maximum likelihood is a theory of estimation, probably the best-known theory of estimation. Estimation is at the heart of econometrics. Consider the following problem.

> Suppose I wish to measure a table. Being too lazy to do it myself I hire a research assistant. The assistant takes 100 measurements of the length and the width of the table. Then, because he thinks that I am interested in the *area* of the table, he multiplies for each observation the length and the width, and throws the underlying measurements away. This gives me 100 measurements of the area. But I am not interested in the area. I want to estimate the length and the width of the table. Can I do this from these 100 measurements of the area?

It was shown by Tom Rothenberg that indeed we can. He also showed something else, arguably even more interesting, namely this. If the research assistant is very sloppy, then clearly his measurements will be bad and we will get a lot of variation in the estimates of the area. To estimate the length and width from such noisy data is then difficult. If on the other hand the research assistant is perfect, then he makes no errors and all measurements are exact. This results in 100 identical measurements of the area, so in fact we have only one measurement. From this one measurement, we could not possibly recover the length and the width of the table. Hence, both too much and too little sloppiness is bad. There exists an optimal level of sloppiness of research assistants, and this optimum can be calculated.

What this story really illustrates is that measurement error is not necessarily a bad thing. Much can be learned from it if we are willing to put some structure on the error process. Then, measurement error reveals more than that it hides.

I was fortunate to work on maximum likelihood with Risto Heijmans of the University of Amsterdam, and I learned a lot from him. He was (and is) a better statistician than I am, but I am better in writing it up, and this was essentially our division of labor.

When Risto and I worked on maximum likelihood in the period 1981–1986, most of statistics had been developed for observations that are independently drawn and can be assumed to be identically distributed. This corresponds to situations where experiments can be repeated as in physics, medicine, or psychology. But in economics experiments can typically not be repeated and our observations are dependent and not identically distributed. For this more difficult class of observations we examined the behavior of the maximum likelihood estimators when the number of observations becomes larger and larger, the so-called asymptotic theory. My family and I remember with fondness Risto's frequent visits to our house in London during this period.

Aggregation

Before London, from 1979–1981, Eveline and I lived in Canada, where I worked at the University of British Columbia and Eveline obtained a masters in Communication Studies at Simon Fraser University, completing her degree with a thesis later published by Cambridge University Press under the title *The Bilingual Experience*. We moved to Canada with one child and left two years later with two.

At UBC I became interested in aggregation, possibly inspired by the growth in my family, and by Alan Woodland, a leading expert on aggregation, who was then also working at UBC. Alan and I worked together primarily on the following question. Suppose we wish to estimate or forecast something at the macro level, say the growth of a national economy. Underlying the macro economy are sectors, which together make up the macro economy. Since we wish to say something about the macro economy we could work at the macro level, that is the national level, only. But we could also model the underlying sectors (the micro level), using a much bigger and richer data set. The question then is: how much do we learn from this additional information. Is it worth the trouble?

At that point, I had moved to the London School of Economics and Alan had been appointed to a chair in Sydney. To make some progress, he kindly invited me for a three-week visit. I took the cheapest possible flight, spent two nights in the air, and arrived exhausted in Sydney. But the three weeks were highly enjoyable and very productive.

The answer to our question is that we don't learn that much from the additional information. In fact, it may be counterproductive. How can that be? Is it not true that the more data the better? Well, yes and no. More data are preferred over fewer data, but we also need to model these underlying sectors and specify how they make up the national economy. In this process we cannot avoid making modeling and specification errors, and the harm caused by these errors may be greater than the advantage of the richer data set.

More detail is not necessarily a good thing, a somewhat counterintuitive and therefore appealing idea, which I took up again in working with Jan van Tongeren on national accounts data, many years later.

Methodology

You probably know the Ten Commandments, or at least some of them. In many disciplines leading scientists have compiled the ten commandments of their discipline, attempting to formalize the most important rules that need to be obeyed.

We also have the 'ten commandments of econometrics,' compiled by Peter Kennedy in 2002. I was asked to comment on these commandments, which I did. But before doing so I asked myself why there are *two* tablets of stone? Surely, God could have written all ten commandments on one tablet, especially since the tablets were written on both sides (Exodus, 32:15). Personally, I like the following explanation, which offends almost everybody, an added bonus.

> After having written the commandments, God wonders what to do with them. He first turns to the British, who look at the commandments and say: 'Ah, we are not allowed to lie? This is not for us.' Next, he turns to the Germans, who say: 'Ach, no killing? Sorry, but no.' Then, he tries the French: 'Oh, we can't sleep with other women?' Finally, God turns to the people of Israel, who ask: 'What does it cost?' 'Nothing,' says God. 'Then, give us two!'

The econometric commandments formulated by Peter Kennedy are all concerned with methodology and relate to the question: how should I do econometrics? I followed the literature on methodology and participated in workshops, but my thoughts were confused until I started to work with Mary Morgan. This was in 1995, surprisingly late considering the fact that Mary and her husband Charles had been friends of ours since I started at the LSE in 1981. Mary and I organized an exciting experiment, where we gave a number of leading researchers or research groups one data set, fully documented, and a series of questions. They had to answer the questions while using only the data provided by us. They could however use their own models and methodology. After one year — I had then moved from LSE to Tilburg — we all came together for a workshop, here in Tilburg. Each group presented its findings and we tried to figure out why some results were close to each other, while other results were far apart. We were perhaps not completely successful in obtaining the answers. I expect that most groups went home thinking that their results were of course the best and that the others were incompetent fools. But Mary and I did not think that any of them were fools, and we used the material for a book, published in 1999, that is still close to my heart.

This book also contains a second experiment, on 'tacit knowledge'. Tacit knowledge is the sort of thing that is not written down but which the pupil learns by watching the master at work. For example, if I consult a cookbook, then I have a hard time to prepare a dish, because I lack experience and many things are taken for granted and are not written down. The collection of these things is called tacit knowledge. Our tacit knowledge experiment was particularly striking, because it produced results which were completely opposite to what we and everybody else had expected, namely that our colleagues who advertise their methods most forcefully turned out to be not the easiest but the most difficult to imitate.

Interlude on intuition

Research, one often hears, starts with asking the right question. When is a question the right question? What does 'right' mean? Here we enter the domain of intuition. We need intuition to formulate a question that is relevant, excites us, and will be of interest to others too, but in such a way that it is neither too easy (so that anybody could solve it) nor too difficult (so that you cannot solve it). We also need intuition to reduce the number of paths to solve a problem, once we have formulated it.

Intuition, however, is a slippery customer and it needs constant care and attention. Poor results or counterintuitive outcomes (and these occur all the time) should not simply be dismissed. Each time we have to ask: why do we find a result that is counterintuitive? Are the data wrong, is the model wrong, is the estimation method inappropriate? Or perhaps our intuition is wrong and the result is in fact a sensible outcome, just different from what we expected. Only then, I am convinced, can we sharpen our intuition and avoid the same mistake next time. Then we will have learned something and this will prove useful in our future research.

Consider the following case, inspired by Tversky and Kahneman's famous article, which appeared in *Science* in 1974.

> The town of Tilburg has two hospitals, a large one and a small one. In the larger hospital 45 babies are born each day, and in the smaller hospital 15 babies are born each day. About 50% of the babies are boys, but the exact percentage varies from day to day. Sometimes it is higher than 50%, sometimes lower. Over a period of one year, each hospital records the days on which more than 60% of the babies were boys. Which hospital do you think recorded more such days?

If you are like most people, then you would answer that the probability of obtaining more than 60% boys is about the same in each hospital. But this is not true. Sampling theory tells us that the expected number of days on which more than 60% of the babies are boys is much greater in the small hospital than in the large one. Why? Because with more observations the precision increases and we are less likely to stray from 50%.
This example is one of many where simple statistical thinking proves too difficult for the majority of people. My explanation for this fact (that people find statistics particularly difficult, much more difficult than say mathematics) has been for a long time that we are not (or hardly) exposed to statistics at school. We are exposed to mathematics from a very young age, but not to statistics. Hence, our intuition is not developed and we make stupid errors. This was my explanation, but Daniel Kahneman has convinced me that I am wrong by presenting the following two mathematical cases, taken from his recent book *Thinking, Fast and Slow*. First, this puzzle.

> A bat and ball cost together € 1.10.
> The bat costs one euro more than the ball.
> How much does the ball cost?

This is a mathematics example, there is no statistics in it, and a very simple one. Still, almost all of you will think that the answer is 10 cents. But it isn't. The answer is 5 cents. The ball costs 5 cents and the bat costs € 1.05, together € 1.10.

Here is another puzzle.

> All roses are flowers.
> Some flowers fade quickly.
> Therefore some roses fade quickly.

Is the argument correct? Most people — according to Kahneman — think it is. But it isn't. The fact that some flowers fade quickly could mean that some roses fade quickly, but it could also mean that some flowers, the hibiscus for example, fade quickly but that roses do not fade quickly at all.

So, mathematics intuition seems not that well developed either, and my own intuition about the difference between mathematics and statistics appears to have been wrong.

Sensitivity analysis

All models are wrong but some are useful. In fact we can go one step further. Even if we knew the truth, that is if we knew the model that actually generates the data, we would not want to use it. More precisely, if I tell you the correct functional form of the model (but not the values of the parameters), then in general you should not use this model in estimation or forecasting.

Isn't that strange? The reason is as follows. If we delete a few variables from this large and complex, but true, model, then a bias will occur in the estimates. If the variables that we exclude are relatively unimportant, then the associated parameters will be close to zero, and the bias will be small. But, since we estimate fewer parameters, their precision will be better. Now, while the bias depends on the values of the parameters, the estimated variance does not. So, we can get a large reduction in variance even when the deleted parameters are close to zero. We will get better estimates using the smaller model. It is acceptable, even desirable, that the smaller model captures not everything, as long as it captures the essence. This is the art of modeling. As Einstein put it: 'As simple as possible but not simpler.'

After I came to Tilburg some twenty years ago, I became interested in sensitivity analysis, and I worked in this area with two PhD students: Anurag Banerjee and Andrey Vasnev. In sensitivity analysis we do not ask the question: is it true? Instead we ask: does it matter? A model is not true or false in this way of thinking, it is useful or not useful. And this depends on the question we wish to answer. A model can be useful when we wish to answer one question, but not when we wish to answer another question. Sensitivity analysis thus breaks with the standard way of thinking that there exists a truth which we need to approximate. Instead we now realize that this truth is not absolute and depends on the question that we ask.

Sensitivity analysis works like this. Suppose we have one parameter of particular interest. We estimate this parameter using various models. These models and the associated methods of estimation depend on assumptions. The estimate of the parameter may be sensitive or not sensitive to these underlying assumptions. If it is sensitive to a specific assumption, then we have apparently overlooked some essential aspect in the model. If it is not sensitive, then this assumption is harmless.

This is quite different from the traditional model selection methodology. The most important, and rather surprising, result that Andrey proved is that diagnostics (coming from the world of model selection) and sensitivity statistics are essentially independent. Hence, knowledge of one tells us *nothing* about the other. Clearly, this reinforces the need to take sensitivity analysis seriously.

I would have liked to go one step further and use sensitivity analysis to build a proper framework for specific-to-general modeling. I have an intuitive aversion against general-to-specific modeling, the most common modeling methodology. Instead I am a firm believer in specific-to-general modeling, because I believe that we should start simple and then extend, rather than start general and then shrink. But my many attempts in this direction have all led to nothing. Maybe Pieter can help me here.

Pieter's view on econometrics

Music: Vincenzo De Crescenzo (1875-1964)
Title: Tarantella sincera
Lyrics: Pieter Nieuwint

The prediction of the past, which is the theme of this oration
(Finding out what happened, and not only that, but how and why),
Is the art of economic modeling and estimation,
Which I find incomprehensible, no matter how I try.
Still, Jan Magnus has a wonderful profession,
And I'm simply overwhelmed by the idea of being present at this overwhelmingly important academic session.
When I hear that prediction of the past
Is now possible, I tend to think: at last!
But one question must be asked:
Is the method computationally fast?

I am much impressed by all the scientific terms he uses:
Matrix calculus and asymptotic theory and such.
I admit that quite a lot of it perplexes and confuses,
And it wouldn't make a difference if it had been in Dutch.
Intuition also gives me indigestion,
And the question whether models that you use in your research are right or wrong for me is certainly too difficult a question.
Oh, the orator casts a magic spell,
But the things he says don't even ring a bell.
No, as far as I can tell,
Magnus just does not explain it very well.

I have always been intrigued by the econometric sector,
And I can assure you that there's little that I wouldn't give
If they made me grasp the essence of a matrix or a vector
Or indeed the underlying concept of derivative.
It is also quite beyond my comprehension
That, whenever you employ a differential, then the differential of a matrix is a matrix of the same dimension.
I am no statistician and confess
That I cannot even make a learned guess.
And I may know even less
After Jan's valedictory address.

Pretesting and model averaging

> It is sad but nonetheless,
> I think I will progress.

Let me tell you a fairy-tale.

> A king has twelve advisors. He wishes to forecast next year's inflation and calls each of the advisors in for his or her opinion. He knows his advisors and obviously has more faith in some than in others. All twelve deliver their forecast, and the king is left with twelve numbers. How to choose from these twelve numbers? Let me provide two ways of thinking (there are more). The king could argue: which advisor do I trust most, who do I believe is most competent? Then I take his or her advice. The king could also argue: *all* advisors have something useful to say, although not in the same degree. Some are cleverer and more informed than others and their forecast should get a higher weight. Which way of thinking is better?

Intuitively most people, and I also, prefer the second method, where all pieces of advice are taken into account. In standard econometrics, however, it is the first method which dominates. This method is called pretesting, while the second method is called model averaging.

Let me explain. In practice, econometricians use not one model but many models. Among these models, one is the largest and one is the smallest. Neither is probably the most suitable for the question at hand. If we use diagnostic tests to search for the best-fitting model, then we need to take into account not only the uncertainty of the estimates in the selected model, but also the fact that we have used the data to select a model. In other words, model selection and estimation should be seen as a combined effort, not as two separate efforts. Failure to do so leads to the misleadingly precise estimates, and the harm that this causes was studied by my PhD student Dmitry Danilov.

Returning to my little story, the pretest estimator corresponds to the king's argument: which advisor do I trust most, who do I believe is most competent? Then I take his or her advice. The king's second line of argument: all advisors have something useful to say, although not in the same degree, corresponds to a method called model averaging. In pretesting we still select one model, but in model averaging we use all models in a class. Each model tells us something about the parameter of interest, and in model averaging we use all this information in a continuous manner, not in the discrete manner of pretesting.

It took me a long time to realize that the machinery we had developed for pretesting could be applied to model averaging, and that in fact model averaging is easier than pretesting. This resulted in work with two other PhD students, Patricia Prüfer and Wendun Wang. In model averaging the question 'which is the correct model' is meaningless, and some colleagues find this unappealing. But for me this is not the main issue. The main issue is the optimal estimation of the parameters of interest using all available information.

With Patricia and Wendun we developed a method called 'weighted-average least squares', which takes full account of the fact that model selection and estimation are a joint effort. The method is computationally fast and is half-Bayesian, being based on a consistent theory of ignorance.

In developing this theory two peculiar problems arose. The first is this.

> Suppose we have one observation, say x, from a normal distribution with unknown mean ϑ and variance one. How would you estimate ϑ?

Almost anybody would reply: by x of course. After all, that is all we have. This seems reasonable, but the so-called 'equivalence theorem' proved in a joint paper with the late Jim Durbin shows that in an econometric context this would imply always selecting the unrestricted (that is, the largest) model. And this is not what we do in practice. In practice we search for smaller models and for good reason. A more appropriate estimator for ϑ is then given by a fraction, say $\lambda(x)\ x$, where λ depends on x. This leads to the next question: how do we find λ? This question puzzled me for a number of years and the answer owes a lot to Aart de Vos and John Einmahl who helped me solve statistical problems that I could not have solved by myself.

The second peculiar problem that arose in this context is this.

> Suppose we have not one but two pieces of information on the unknown parameter ϑ, say x and y. Both are normally distributed with common mean ϑ and some known variance, not necessarily equal. How do we estimate ϑ?

If the two variances are the same and x and y are uncorrelated, then the answer is simply the average of x and y. If the variances are not the same, but x and y are still uncorrelated, then the answer is some weighted average of x and y, where the weights depend on the variances such that the more precise observation gets the higher weight.

But suppose that x and y are correlated. Then the answer is that we still should estimate ϑ as a weighted average of x and y, except that the weights do no longer necessarily lie between zero and one. This seems very puzzling.

Let's return to our king, and suppose that he has reduced the number of advisors from twelve to two. He asks his two advisors about next year's inflation. One says 2%, the other says 4%. The first estimate has variance 1, the second has variance 4. If the two advisors were uncorrelated (they don't know each other and they base their forecasts on different data sets), then the king would weigh the two estimates, giving a higher weight to the first advisor because she is more accurate (has a lower variance). The answer then is 2.5%, which is in-between 2% and 4%, but closer to two than to four, as expected. But it is more likely that the advisors are correlated, that they do talk together, and that they use the same or similar data sets. If their correlation is 3/4 (which is not that high), then the king should estimate the inflation by 1.5%, and this lies outside the range indicated by the two advisors (2% and 4%).

Now we have a problem. We have a result which makes mathematical and statistical sense, but if the king goes outside the range indicated by his advisors he will surely be heavily criticized. In practice, therefore, policy recommendations do not follow the correct mathematical procedure. Maybe this is one reason why they are so often incorrect.

Interlude on integrity

What is proper scientific conduct, in general and for an econometrician in particular? Some breaches of scientific conduct, such as plagiarism or inventing your own data, are obviously wrong. But how about deleting data that give unsatisfactory results? Or adjusting your model again and again until it gives the results you want? This is also wrong, but some of it we all do and it seems to be necessary and accepted.

Data are often used the way a drunkard uses a lamppost: for support rather than for illumination. If the lamppost does not give enough support, we fix it or we look for another lamppost. This is not right. However, to define an ethical line which should not be crossed seems impossible. Instead I recommend that every published paper should have a logbook appendix, either in the paper or on a website, where the author tells us the path that he or she has traveled, including the mistakes, the changes of model, and the data massaging. This will be useful for students and colleagues.

I have made this recommendation repeatedly, but as yet no econometric journal requires it and few authors take the trouble and the time to do it. In my view, we all should.

National accounts estimation

I believe it was Rick van der Ploeg who brought Jan van Tongeren and me together, around 1997. At the time, Jan was head of the Statistics Division at the United Nations in New York, and a world expert on national accounting, a subject I knew very little about.

Our common interest was (and is) based on the fact that data do not simply fall to Earth like manna falls from Heaven. Instead, data are constructed by some agency like Statistics Netherlands or the United Nations Statistical Division. Applied economists seldom see 'raw' data. The data that they receive as their starting point have already gone through a process. This process could involve imputing missing values or balancing items coming from different sources. The researcher should of course be aware of these facts, but very often he or she isn't. This shows a disrespect for the data, which in my view is only too common in econometrics.

Jan and I worked together for more than ten years on establishing new frameworks for national accounts estimation. In these frameworks, the data are the parameters that we need to estimate. But we know something about these data. For example, we know their approximate values in previous years and we know bounds for many ratios of the data. All this information can serve as priors in a Bayesian set-up. This is where Aart de Vos, my Bayesian friend, enters. Later, Dmitry Danilov also became a partner in this team, as master programmer.

In this ambitious project we developed a new theoretical Bayesian framework for national accounts estimation, and also the software required to estimate very large data sets. Jan successfully applied these ideas in various countries, and obtained his PhD (after retiring from the United Nations) based on these studies.

Warming and dimming

We are all seeing rather less of the Sun. Scientists looking at five decades of sunlight measurements have reached the disturbing conclusion that the amount of solar energy reaching the Earth's surface has gradually been falling. The effect was first spotted by Gerry Stanhill, an English scientist working in Israel, and he called the phenomenon 'global dimming'. His research, published in 2001, met however with a skeptical response from other scientists. It was only recently, when his conclusions were confirmed by Australian scientists, that climate scientists woke up to the reality of global dimming.

Horizon, BBC's celebrated science program, broadcast a film about dimming in 2005, which Eveline saw and urged me to see as well. As so often before and since, Eveline kept me grounded and interested in real-life affairs. Trying to find out a little more, I noted that the people involved in this debate were all physicists. No statisticians were involved, which struck me as strange. I wondered what the statistical evidence was. These vague thoughts started a project with Chris Muris and Bertrand Melenberg, where we tried to disentangle global warming and local dimming, using a rich and unique data set given to us by Martin Wild from the ETH in Zürich.

Not one, but two effects largely determine global warming: the well-known greenhouse effect and the less well-known solar radiation effect. An increase in concentrations of carbon dioxide and other greenhouse gases contributes to global warming: the greenhouse effect. In addition, small particles, called aerosols, reflect and absorb sunlight in the atmosphere. More pollution causes an increase in aerosols, so that less sunlight reaches the Earth (global dimming). Despite its name, global dimming is primarily a local (or regional) effect. Because of the dimming the Earth becomes cooler: the solar radiation effect. Global warming thus consists of two components: the (global) greenhouse effect and the (local) solar radiation effect, which work in opposite directions. Only the sum of the greenhouse effect and the solar radiation effect is observed, not the two effects separately.

What Chris, Bertrand and I tried to do is to identify the two effects. Why? Because the existence of the solar radiation effect obscures the magnitude of the greenhouse effect. What did we find? During the 43 years 1960–2002 temperature increased by an estimated 0.73 °C, which we were able to decompose as

$$0.73 = 1.87 - 1.09 - 0.05,$$

namely a greenhouse effect of 1.87 °C, a solar radiation effect of 1.09 °C, and a remainder term of 0.05 °C. Hence, if aerosols and solar radiation had remained at the 1959 level, then the expected global average temperature would have been 1.09 °C higher. The solar radiation effect is therefore important, masking 58% of the increase due to the greenhouse effect. Ignoring dimming thus causes a serious underestimation of the greenhouse effect.

Our results are alarming. Fortunately, we are now in an economic crisis: less production, less pollution, less global warming.

Tennis

Let us turn to a happier subject. It must have been in June 1992, during the period that I was working both in London and in Tilburg, that Eveline and I were watching the tennis at Wimbledon on TV and listening to the BBC commentators. We discussed and subsequently kept a record of the statements made by the commentators, for example that serving with new balls provides an advantage, that a 'winning mood' exists, or that top players must grow into the tournament and that they perform particularly well at the 'big' points. We wondered whether these commonly accepted ideas were true or not, so I contacted IBM to give me their data, which they did.

Two years later, in 1994, I was joined by Franc Klaassen, then a student here in Tilburg, now full professor at the University of Amsterdam. For almost twenty years we did research on tennis, first on these accepted ideas, later also on other — statistically more demanding — aspects of tennis, such as the existence of winning mood, how efficient top players are, and how we can best predict the outcome of a match, while the match is in progress. In the end Franc and I wrote 15 papers together, most but not all about tennis.

Tennis research, like all sports research, may be considered a frivolous activity, and you may well wonder why the taxpayer should support scientists engaged in such activities. You are right, but only in part. Sports data have some characteristics that make them ideally suited for studying human behavior. They are clean — there are few errors in the data — and the data collection is transparent and can be checked. Data in economics, psychology, and many other sciences are often messy and ambiguous. For economists, psychologists and other scientists to work with clean data provides a great opportunity and a welcome change from normal practice. If results do not come out the way they 'should', then this cannot be blamed on the data. There must be a 'real' explanation and we have to find it. Maybe our preconceived idea is wrong or maybe we have not applied the correct statistical method.

Given the abundance of clean sports data we can try and study human behavior in an *indirect* way. Let me give two examples. Suppose we wish to study whether judges and juries in our legal system are influenced by social pressure. Useful data from the law courts are not available, so we cannot directly study the possibility of favoritism in the courts of law under social pressure. But we can study it indirectly by considering football matches and asking whether referees favor the home team, for example by shortening matches in which the home team is one goal ahead and lengthening matches in which the home team is one goal behind at the end of regular time. It turns out that referees tend to do this, thus favoring the home team.

My second example is the question whether people become more cautious when pressure mounts. Franc and I analyzed this question using tennis data. Some points are more important than others. Do players behave differently at the key points? They do: they play safer at important points. This teaches us something about human behavior, and may have a broader significance, for example in economics. If salaries of agents working in the financial sector contain not only a bonus but also a substantial malus component, then the consequences of their activities matter in both directions (like winning *or* losing a tennis match). The behavior of professional tennis players suggests that financial agents will then pursue safer actions, reducing the possibility of a banking crisis.

A few months ago Franc and I completed our book *Analyzing Wimbledon*, which attempts to put everything that we found together in one framework and at one level. This book will appear later this year and I recommend it to you.

The Philip Swallow prize

When I started as an academic, I never thought about a 'research plan' or a 'career strategy'. I just studied what caught my interest, without plan or strategy. This, apparently, is all wrong. One should have a plan and stick to it, so that one works efficiently and maximizes the number of published papers.

Not just the number of papers, but also — and in particular — the points associated with these publications. These points were quite amusing when Arie Kapteyn and Tom Wansbeek first compiled their celebrated list in 1980 under the name 'A.D.S. de Schuite', an anagram for 'Dutch disease'. The list and subsequent lists generated a revolution in counting in Dutch economics. No doubt, some of this was much needed and beneficial. But too much counting is — in my view — counterproductive. Young scientists should study what they find interesting and publish their findings in the most appropriate outlet. They should not think about the points associated with a specific publication.

To put some power behind my minority view, I instituted the Philip Swallow Prize in Tilburg. What is the Philip Swallow Prize? Who is Philip Swallow in the first place? Philip Swallow is a character in David Lodge's 1975 novel *Changing Places*.

Among a group of American English literature scholars, an academic from England called Philip Swallow, himself a scholar in English literature, proposes to play a game, which goes like this. Each person in turn calls out a book or play that he or she has *not* read. The more people have read it, the more points you get. The purpose of the game is therefore not to show off and mention obscure books that nobody has read, but just the opposite. A very English game. In Lodge's story, an untenured American academic, called Howard Ringbaum, eventually understands the game and admits that he has never read *Hamlet*. Of course he wins the game (but loses his job). Apparently, Lodge himself has never read *War and Peace*, which gives him a solid chance of winning.

In the same spirit, I proposed last year to present a prize (which I named the Philip Swallow Prize) to the member or members of the econometrics group at Tilburg whose paper had been rejected by the lowest-ranked journal. What happened? I won the prize myself for my rejected publication (jointly with Mars Cramer) submitted and rejected by *Economisch Statistische Berichten*. Apparently I am the only one in the department whose papers get rejected by low-ranked journals. Naturally it fills me with pride to be associated with such a department.

Predicting the future

Let me return to the title of my lecture: predicting the past. I have led you through a few of the highlights of my past. Is this a true account? Have I predicted the past well? Only up to a point. I have deleted a lot and mentioned only a fraction of the people I collaborated with.

As all historians know, predicting the past is not easy. Of course, if we built a model based on a given data set, and then predicted how well the model behaves using the same data, then we would get good — possibly even perfect — results. This, however, is ridiculous, which does not mean that some of my distinguished colleagues do not engage in such practices.

Predicting the past is not easy, but it is a necessary ingredient for predicting the future, that is for forecasting. Forecasting is a bit like driving a car blindfolded, following directions given by someone looking out through the back window. Take my own case. If you had wished to forecast last year what I would be doing the current year, then you would have been fairly accurate. But how about forecasting now what I shall be doing next year? This is more difficult, because a structural break is about to occur, namely my retirement. My situation changes from being employed to being unemployed. When a structural break occurs, econometricians are assumed to help making forecasts and policy recommendations. Unfortunately, we have a hard time doing so. Maybe I will do the same as I am doing now, maybe something completely different. I'll tell you next year.

Coda

Mr. Rector Magnificus, ladies and gentlemen.

Coming to the end of my lecture, I think of the people who have inspired me. My mother died young and did not even see me complete high school. She never met Eveline or her grandchildren. Still, fifty years after her death, my memories of her are alive and important to me. My father, and Ank, his amazing second wife, both died before my inaugural lecture. They would have enjoyed being here and I miss them.

I thank my teachers at the University of Amsterdam, my colleagues, and my students for keeping me alert, and I thank my two principal employers — the London School of Economics and Tilburg University — for providing truly inspirational working environments. I couldn't have been more fortunate.

Within Tilburg I moved office several times, being first located among the CentER professors, which was much fun, and finally among the econometrics group on the sixth floor, also much fun. This last move proved particularly fruitful. It is much underestimated how important the location of your room is within a group. If you are located in such a way that you meet casually at the coffee machine or the printer, then the probability of joint work is much increased. In my own case, I wrote joint papers with almost everybody on the sixth floor, which would never have happened if I had been located on a different floor.

I have been fortunate to be able to travel a lot. I spent much time in Moscow, helping establish the New Economic School, and recently in Japan, China, Australia, and New Zealand. I thoroughly enjoyed these experiences and I learned a lot from being in different institutions with different customs and different priorities.

I don't think I am a particularly easy person to work with. Still, I have had many coauthors:

James, Peter, and Giuseppe
Jan, Asraul, and Ashoke

Henk, Roger, and Anatoly
Jim, Bertrand, and Dmitry
Mars, Karim, and Hiroaki

John, Bahram, and Alberto
Kan, Owen, and Hugo
Flip, Karen, and Risto
Franc, Victor, and Masako
Chris, Arthur, and Franco

Pavel, Erricos, and Alexei
Kamlesh, Ronald, and Andrey
David, Anurag, and Mary

Heinz, Alan, and Otilia
Tom, Wendun, and Salima
Geert, Martin, and Patricia
Aart, Xhinyu, and Ursula

In joint scientific work, my belief is that compromises are unacceptable. What this means is not that what I think should be done irrespective of what the other person thinks. No, what it means is that authors should only collaborate if they are clear about a common goal. Then, if there is a disagreement, the question is not to win or lose the argument but rather to find out how it is possible that there is a disagreement in the first place. Maybe one has not fully grasped the implications of some assumption or is confused about some aspect of the data. In figuring out such disagreements much progress can be made, and it strengthens the partnership and improves the project. A compromise may be good in business, but not in science.

My family — Eveline, Gideon and Joyce, Hedda and Ralph, and my two grandsons Dorian and Midas — have often been taken for granted, yet they are and always have been the most important to me. My two grandsons have added a new sparkle of joy and tenderness to my and Eveline's life, Gideon and Hedda (and their splendiferous partners) have made my life more intense and more complete, and it is wonderful to see them as grown-ups and thriving in their work and life.

My wife Eveline has now been with me for more than forty years. When we met, I had after 4 1/2 years of study almost passed my 'propedeuse', almost completed my bachelor's, and had just started on my masters — a combination of half-finished degrees that would not be possible now. Meeting you provided a new and much-needed focus, which has been essential to me ever since. It was the luckiest thing that ever happened to me. You have been my partner emotionally, intellectually, and musically. You continue to be the music in my life.

We discussed predicting the past and predicting the future. How about predicting the present? The present is that this lecture has come to an end and that drinks await you outside.

I thank you for your attention.

Colofon

vormgeving
Beelenkamp Ontwerpers, Tilburg
foto omslag
Ton Toemen
druk
PrismaPrint, Tilburg University