Test for Unit Roots and the Initial Observation

Ulrich Müller

I am very happy and truly honoured to have been awarded the *Prix Latsis Universitaires 2002 de l'Université de St. Gall* for my dissertation in time series analysis. The Foundation makes proof of an extraordinary generosity, and I would like to express my deep gratitude. There are many people that have contributed to the quality of my work ; I would like to especially mention my advisors Professor Kirchgässner, Professor Keel and Professor Elliott. Without their generous support and invaluable suggestions my research project would not have been possible.

Basic Concepts

Prior to explaining my dissertation's contribution, it is necessary to familiarise the reader with a number of basic concepts.
Statisticians (and economists) classify data into two main categories : cross-section data and time series data. Cross-section data is data where the ordering of the observations is unimportant. As a typical example, consider data from a political poll, which is undertaken to estimate the approval ratings of the different political parties. In such a data set, it has no bearing whether M. Mueller was asked for his preferred party before or after Mrs. Smith. All responses are independent in the sense that the answer of M. Mueller does not influence what Mrs. Smith replies.
This is different in time series data, where data is naturally ordered in the time dimension. The crucial distinction to cross-section data is not so much that there is a natural ordering, but rather that the observations are not necessarily independent from one another. To stick with the example above, consider a time series of the average approval rating of a specific political party, sampled every month. If the party goes through some kind of crisis in a given month, then this will lower the approval ratings not only in that month, but also the following months. A low approval rating will typically be followed by another low rating, whereas high ratings are typically followed by another high rating. Changes in the time series take place only slowly. This dependence over time has to be taken into account when working with time series data.
This leads us to a further distinction for time series data that is key for my research : Most time series are either 'integrated' or 'mean reverting'. A 'mean reverting' series, as the name suggests, tends to revert to a constant mean, at least in the long run. While one large observation

(relative to this mean) might tend to be followed by another large observation, observations far in the future are distributed around the constant mean, independent of today's value. Shocks might have a prolonged impact, but not an everlasting one. A typical example might be the amount of sugar in human blood. At the beginning of a physical exercise, the sugar content falls. But the body reacts by transforming fat into sugar, or by inducing hunger, so that in the medium-term, the sugar level returns to its equilibrium value. Only in the short-run (say, minutes) a low sugar level is likely to be followed by another low sugar level reading. An observation of a low sugar level today does not increase the likelihood of a low sugar level in two months time.

In contrast, 'integrated' series do not revert to any mean. Any shock has an everlasting impact, and a large observation at a given time increases the likelihood of a large observation even in the distant future. As an example, consider the wealth of a Casino. When M. Smith hits it big one night, the Casino's wealth falls, and this decline in wealth has an everlasting impact. Even when the Casino does very well afterwards (i.e. the players are unlucky), then also in two years time the wealth of the Casino would be even greater had M. Smith not have had the lucky strike. The (negative) impact of M. Smith's winning is everlasting on the Casino's wealth.

While for these two examples it is pretty obvious whether they fall in the 'mean reverting' or 'integrated' category, this is not the case for many other series. A classical economic example is the Gross Domestic Product. Do recessions have an everlasting negative impact on the Gross Domestic Product, or does the occurrence of a recession lead to above average growth afterwards which, in the medium term, leaves the economy where it would have been without the recession ? This is of obvious interest to economists, but has also important policy implications : If the negative impact of recessions is everlasting, then economic policy should try very hard to avoid them. If recessions have impact over a limited time only, then in the long run, economic policy can make a contribution to a larger output by increasing the average only.

What is needed then is some kind of procedure that allows distinguishing mean reverting from integrated series. Procedures that aim at distinguishing two competing models are called 'hypothesis tests' in the statistical jargon. As a straightforward example for a hypothesis testing problem, imagine you have a dice, and you sense that the dice is unbalanced in the sense that is shows too many sixes. The two hypotheses are « the dice is regular » and « the dice is pronged ». Imagine further that the only thing you can do is to throw the dice 30 times. An obvious way to decide between the two hypothesis is to conclude « the dice is pronged » when you have observed more than, say, 10 throws of sixes, and to conclude « the dice is regular » otherwise. In doing so you can make two kinds of mistakes : On the one hand, by chance even a regular dice might have produced more than 10 sixes in 30 throws. On the other hand, by chance a dice that shows a six one in three times on average might have resulted in less than 10 sixes in the 30 throws. An optimal procedure is a procedure that makes the probability of such mistakes as small as possible (in fact, the above procedure is optimal in this sense).

Contribution of the Dissertation

Essentially my thesis is an application of the theory of statistics to the problem of correctly classifying a time series as either mean reverting or integrated. In the jargon developed above, I examine a hypothesis test for time series data with the two hypotheses being « the series is integrated » and « the series is mean reverting ».

This hypothesis test has received a great deal of attention. For reasons that I will discuss further below, it is often of great interest to know whether a series is integrated or mean reverting. In contrast to the examples of the blood sugar level and the casino wealth, for most series the answer is not obvious. One hence must to rely on the observed data for the classification. This is analogous to the problem of the potentially pronged dice, where the decision is taken on the grounds of the outcome of 30 throws. But in the dice example, it was pretty obvious how a good (in fact, optimal) procedure looks like : Count the number of sixes in the 30 throws and conclude « the dice is pronged » when this number is too big. But on which basis should one decide whether a time series is integrated or mean reverting ?

The literature contains many, many different procedures which might be used. At the same time, it was not very clear what the advantages and disadvantages of these procedures were. Even worse, it was generally unknown whether the suggested procedures were optimal in the sense that they minimise the two types of errors one can commit : (1) Classification of an integrated series as mean reverting and (2) classification of a mean reverting series as integrated. It is fair to say that the literature did not contain any clear guidance which procedure should be used in which circumstances, and whether it was a fruitful enterprise to try to somehow develop a better procedure.

My dissertation goes a long way towards clarifying the issues surrounding the problem of classifying series as mean reverting or integrated. The main idea which allows this clarification is the concept of optimality : rather than suggesting procedures in an ad hoc fashion, I derive optimal procedures from the start. This is done by applying optimality results for general hypothesis tests to the specific statistical situation for the problem of classifying series as integrated or mean reverting.

A derivation of optimal procedures has several advantages : First, optimal procedures are necessarily relatively good. This does not preclude that even optimal procedures commit the classification errors described above ; but the probability of such mistaken classifications is as low as possible. Second, by their very optimality, optimal procedures cannot be improved upon. It is unnecessary (even pointless) to continue to search for better procedures. Third, it is interesting to relate existing procedures to optimal procedures. The optimal procedures are optimal only for a specific set of additional assumptions. When an existing procedure turns out to be equivalent to an optimal procedure for a specific set of assumptions, then this implies that the existing procedure should only be used when these assumptions are likely to be true. In this indirect way it is possible to understand the merits and disadvantages of the existing procedures.

It turns out that most of the existing procedures that were suggested for the classification of time series as integrated or mean reverting are in fact very close to optimal, but for very different assumptions on the beginning of the time series. While certain procedures are optimal when the mean reverting series start far off the equilibrium value, others are optimal for a much smaller initial condition. The implicit assumption made by these procedures on the beginning of the time series is key to understanding their behaviour and their merits for specific circumstances. Consider German Gross Domestic Product data, for instance, where the observation period starts in 1946 (which is not unlikely, since much data was not collected or destroyed during the war). Even if the series is mean reverting, it will start far off (below) its equilibrium value. For such a series, tests that assume a small initial condition should not be used.

Practical Relevance

It is important to deeper understand procedures that discriminate mean reverting from integrated series for three main reasons. First, the knowledge that a series is integrated or mean reverting might have direct implications for the best course of action. A prime example is stock prices (or related series like price-earning ratios and the like). If stock prices are mean reverting, then at least in the medium run, low prices will be followed by prices closer to the equilibrium. One should hence invest when prices are low and wait that prices rise before selling the stocks again. If stock mprices are integrated, however, no such equilibrium exists and there is no possibility to 'beat' the market. In the more scientific discourse, some economic theories predict the existence or absence of integrated series, so a test of this prediction is an empirical test of the validity of the theory.

Second, mean reverting and integrated series have very different properties in a statistical sense. This implies that statistical tests and methods that were developed for one kind of series are usually not valid for the other kind. Reliable discrimination between mean reverting and integrated series is hence important for the appropriate statistical treatment of time series data.

Finally, it makes a big difference to forecasting whether a series is mean reverting or integrated. Mean reverting series return to their mean value in the medium term, so good forecasts should estimate this mean and predict a return to it. This would be a very bad strategy for integrated series, however, since they do not revert to anything. The best forecast of an integrated series is simply the last observed value.