Duvvuri Subbarao: Statistics in the world of the Reserve Bank of India

Inaugural address by Dr Duvvuri Subbarao, Governor of the Reserve Bank of India, at the Statistics Day Conference of the Reserve Bank of India, Mumbai, 5 July 2011.

* * *

We are gathered here today for the 5th Annual Statistics Day Conference of the Reserve Bank to honour the signal contribution of late Prof. P.C. Mahalanobis to the Indian statistical system. Let me start by noting that the United Nations declared October 20, 2010 as the first World Statistics Day. In recognizing the importance of statistics and statistical analysis to public policy formulation, we in India have been ahead of the world.

Statistics day in RBI

I note with satisfaction, indeed some pride, that the celebration of the Statistics Day in the Reserve Bank is much more than an annual calendar event. We are deeply conscious of the fact that our policy calibration, be it monetary policy, regulatory actions or decisions in pursuit of financial stability, is improved by the quality of data at our command and our ability to analyze and interpret that data. The decisions that we in the Reserve Bank make have a profound impact on the macroeconomy, and errors can be costly. Our policy judgement should therefore be based not only on state of the art skills in data analysis and interpretation but also on an intellectual value system of ruthlessly honest validation and peer review.

While on the subject of intellectual integrity, I must pay a tribute to Prof. Suresh Tendulkar who passed away a couple of weeks ago. Prof. Tendulkar, as many of you know, was a director on the Central Board of the Reserve Bank. He was a distinguished economist and was widely respected for his teaching and scholarship as well as for his authoritative comments on Indian macroeconomic issues. Among Prof. Tendulkar's many strengths was his unwavering commitment to base whatever he said on data analysis. You could rely on what he said because you knew that it came from painstaking empirical work and a sharp mind that could relate what the data said to the real world situation. It is that value system and that judgemental capability that we must pursue and promote in the Reserve Bank.

This year's conference

I gather this year's conference has two special themes: (i) time series analysis; and (ii) financial statistics. I can hardly overemphasize the relevance of the two themes. Time series analysis is important for understanding the underlying trends and for identifying inflexion points. In interpreting time series analysis, we need to apply judgement on whether the future would be different from the past, and if so in what way. In other words, we must be able to say if the underlying trend has turned non-linear. The second conference theme, Financial Statistics, has always been important; indeed, its importance has grown in the rapidly globalizing world. In a globalizing world, financial systems are interconnected both within and across jurisdictions. In analyzing and interpreting Financial Statistics, we must keep these interconnections in view.

We are privileged to have with us today two eminent guest speakers – Prof. Sastry G. Pantula from the University of North Carolina, USA and Prof. Rajeeva Karandikar, Director of the Chennai Mathematical Institute – two statisticians in the forefront of their respective fields of specialization. They bring together enormous intellectual acumen, wisdom and expertise in two vital branches of statistics: time series analysis and financial statistics. On behalf of all of us, I want to extend a hearty welcome to both of them.

Global crisis and statistical analysis

Let me make another observation on a larger canvas before turning to specifics. As we emerge out of the global crisis, the deepest financial crisis of our time, there are many lessons being drawn. Among the "big picture" lessons is that the real world is not deterministic but that the real world is stochastic. Our analysis and policy judgements should not lose sight of this truism.

Let me explain this by drawing an analogy from physics. In the Newtonian world view, we have a clockwork universe. All we need to know are the basic laws of physics and the initial conditions, and it would be possible to determine the outcome of the universe at any and every point of time in the future with absolute precision. Newtonian physics requires a God, but a God who is no more than a clockmaker.

With the benefit of hindsight, we now know that the mistake economists made was to extend the Newtonian world view, applicable to the world of machines, to the real world inhabited by capricious human beings. In line with the "deterministic reductionism" that was the defining feature of Newtonian mechanics, economists deluded themselves into thinking that the laws of economics can be reduced to precise and deterministic mathematical equations. We should have known better. It took a devastating crisis to tell us that this was a serious mistake. How the economy evolves, both at the macro level and the micro level, is determined by human behaviour which can scarcely be fitted into an analytical mathematical equation. The sub-prime crisis and the financial engineering that caused it are a powerful illustration of the fact that it is human behaviour that drives the destiny of finance and economics.

In that sense, the crisis reinforced the importance of empirical analysis in interpreting the real world. People inhabiting the real world act, react and respond to the evolving situations in inexplicable, unpredictable and often capricious ways, and the only way to capture that real world sociology and psychology as it translates to economic trends is through collection of reliable and relevant data and analyzing and interpreting that data. Statistical analysis therefore has to be an increasingly important support system for our policy formulation.

Statistics in RBI's policy making

Let me now move on from the big picture to the world of RBI. Needless to say, statistical analysis is critical to RBI's policy formulation. We face many challenges and dilemmas in this area. I will address the more important of these under the following heads: (i) Quality of data; (ii) Data revisions; (iii) Data interpretation; and (iv) Data gaps.

Quality of data

At the outset, I must say that India has a rich tradition of statistics. Given the size, diversity and as yet low levels of literacy, this is a tribute to the diligence of our public statistical agencies. Surely, some of the basic statistics such as, for example, the poverty ratio are highly contentious. Both the methodology used for estimation of the poverty ratio and the quality of data used have frequently been called into question. Actually, I see this as a positive sign. I would read this as an indication of the intellectual watchdog function that is built into our statistical tradition. What is put out by a public agency should after all be contestable.

Having said that, in the Reserve Bank we are handicapped by the reliability of some of the basic data that we need to use in policy calculations. In particular, the data we get on unemployment and wages do not inspire confidence as regards quality. The recently put out data on employment throw up a paradox as they simultaneously indicate fewer jobs created in the five year period to 2010 along with a decline in the long-term unemployment rate. On the issue of wage statistics, the upward pressure on wages in the unorganized sector is

inconsistent with what are believed to be high rates of unemployment and underemployment in the informal economy.

The Index of Industrial Production (IIP) is another statistic that has shown counter-intuitive trends. During the period when the global financial crisis was at its peak – December 2008 to June 2009 – IIP growth was positive according to the then available IIP series. This was contrary to our assessment of the underlying trend of some deceleration on account of the crisis. The new IIP series, revised with 2004/05 as the base, now shows that IIP growth was, in fact, negative during that period vindicating our intuition. Again, the old IIP series indicated that industrial activity slowed in the second half of last year (2010/11) relative to the first half but the revised IIP series shows that industrial growth maintained roughly the same pace between the two halves of the fiscal year.

Another problem with IIP has been its volatility, with the volatility being even larger in the capital goods sector. This is analytically bewildering. The volatility persists in new series too. It is important for policy purposes to determine whether the root cause of such behaviour is the production decisions in the wake of uncertainty or whether it is due to the compilation process. Several explanations have been offered for this including that IIP captures the product manufactured in the month and not the product' under production'. Hence, items, particularly with a manufacturing cycle of longer than a month, can lead to sharp month-to-month spikes/dips, not necessarily consistent with the underlying demand in the economy.

The above illustrations demonstrate how poor quality data could potentially mislead policy calculations.

Data Revision

The Reserve Bank's policy formulation is also handicapped by frequent revisions to data. We make policies in real time and if the provisional data that these are based on are inaccurate, the resultant policies can turn out to be sub-optimal choices. Take estimates of GDP growth. For the year 2009–10, for example, the Advance Estimate of GDP growth rate at market prices from the expenditure side, that came out in February 2010, was 6.8 per cent. That was changed to 7.7 per cent in the Revised Estimate in May 2010 and further to 9.1 per cent in the Quick Estimate in February 2011. Therefore, policy that per force had to use information on Advance Estimate of GDP was fraught with the risk of underestimating the growth momentum.

The more critical data on wholesale price index (WPI) inflation too has been subject to large revisions. For example, initial estimates of WPI inflation were 8.2 per cent for January 2011 and 8.3 per cent for February 2011. Both these numbers were substantially revised upwards by 120 basis points each. Oftentimes, it is not clear if the revisions are occasioned by one-off factors or systemic factors. Nevertheless, each time we have to make an assessment of the inflation situation, we are left to double guessing how the provisional number might be revised.

I think it was Groucho Marx who said, "never make a forecast, especially about the future". The Reserve Bank does not enjoy the freedom of choice in this regard. As a matter of policy guidance for stakeholders, we are obliged to give our projection of inflation. Such a projection is evidently based on data available at that point in time. But if the provisional data that we feed into the econometric model is off-track and does not exhibit any systematic pattern, our projections of inflation too gets off-track.

The Reserve Bank's inflation projections systematically under-predicted year-end inflation during 2010–11. The under-prediction is owed to a variety of factors: (i) the larger-than-expected increase in prices of international crude oil and other commodities; (ii) the less-than-expected decline in food prices, despite a normal monsoon, which in turn reflected structural changes in food consumption and more demand for protein-based food

items; (iii) erroneous signals from the then available IIP data which suggested some moderation in growth and demand pressures in the second half of 2010–11, whereas the revised IIP series, now available, shows no such deceleration in growth and demand; and (iv) larger than usual upward revisions to the past inflation data. In this context, persistently high inflation during 2010–11 and the continuation of this trend through the first half of 2011–12 suggest that we need to revisit our estimates of the potential growth rate of the economy.

Admittedly, data revisions are unavoidable in real world, real time statistics but we must make a conscious effort to minimize the quantum as well as the number of revisions in order to make policy choices better informed and more effective.

Interpretation of inflation data

This is possibly an appropriate context for me to address some criticism against the Reserve Bank's use and analysis of inflation data. Abstracting from all the details and nuances, the criticism can be summarized as follows:

(i) The use of WPI as the headline inflation index is flawed as it does not capture the final goods prices that consumers actually experience in the market. RBI should be guided more by the CPI which more accurately reflects demand pressures because it is demand pressures that monetary policy action can influence.

(ii) RBI's interpretation of non-food manufacturing inflation as a proxy for demand pressures is flawed as it reflects more supply side price adjustments and cost push factors than demand side pressures.

At the outset, I must admit that this criticism has some merit. Conceptually, the CPI is a better indicator of demand side pressures than the WPI. An increase in wholesale prices, if sustained, either results in an eventual increase in prices by retailers or a squeeze in their margins. If demand is strong, retailers may exercise pricing power and pass on the increase in wholesale prices to consumers. In case demand is weak, retailers will be forced to partly absorb the increase in wholesale prices in their margins. Thus, there is no denying that consumer prices better reflect demand side pressures than wholesale prices.

Even so, in India, we have opted for WPI over CPI as a second best choice for a number of reasons. First and most importantly, we do not have a single CPI that is representative of the whole country. Until recently we had four, and of those, currently we have three CPIs representing different segments of the population. While WPI is computed on an all-India basis, CPIs are constructed for specific centres and then aggregated to an all-India index. Second, WPI is available with a slightly shorter lag than the CPIs. Third, WPI has a broader coverage than the CPIs in terms of the number of commodities, number of quotations, inclusion of non-agricultural products and tradeable items.

Finally, the revision of the basket for CPI series lags that of the WPI series. Last year, the WPI series was revised to the base of 2004–05 whereas the existing CPIs continue with the old base – for CPI-RL (1986–87), CPI-AL (1986–87) and CPI-IW (2001) which makes CPIs ill-equipped to capture the price behaviour caused by the rapid structural changes in the economy. The revision of the base year in the case of WPI also led to some increase in the weights towards fuel and power and manufactured products away from primary articles. Although changes in the weights for manufactured products were not substantial for the group as a whole, there has been a tilt in the weights towards non-food manufactured products reflecting changes in the production pattern over the decade. The number of commodities included in the new WPI series increased from 435 to 676, even after dropping/revising 176 items from the old series.

Given the limited efficacy of monetary policy to deal with food and fuel inflation, and the limits on using core CPI inflation measures, we have focused our attention on non-food manufactured products inflation as an indicator of demand-side pressures in the economy.

Generally, core inflation is a derived inflation measure from the headline. The transitory component of the headline, food and fuel, is excluded. This is the standard practice though there are many statistical methods of exclusion. The short point is that core inflation is designed to capture the persistent component of inflation, which is amenable to policy. While it is true that commodity prices influence the non-food manufactured products component of WPI, it is also true that the pass-through from higher commodity prices to WPI depends critically upon the underlying demand conditions in the economy.

Let me reiterate what we have said several times in the past. For an assessment of the inflation situation, in the Reserve Bank, we look at all the measures of inflation, both overall and disaggregated components, in conjunction with other economic and financial indicators. In the context of monetary policy formulation, it is important to have a robust primary measure of inflation at the national level. In this direction, the compilation and dissemination of CPI (Urban), CPI (Rural) and the CPI for the country by CSO is an important step forward. However, long time series data, especially for the back period are not available for these new indices making them unsuitable for policy analysis. There is also a need to augment the price indices with appropriate coverage of the service prices to improve their overall representativeness. Moreover, regular and more frequent updating of the base year for price indices is also important to reflect the structural changes in the economy and to enhance their representativeness.

Data interpretation is also important in areas other than monetary policy. One of the critical building blocks of the post-crisis Basel III package is the build up of countercyclical buffers during the upswing of a business cycle and their use during the downturn of the business cycle. This means we need to be able to make a judgement on the inflexion point in the business cycle in real time. Quite evidently, both pre-mature action or delayed action can be costly in macroeconomic terms. In order to deploy countercyclical buffers, we need to hone our skills in business cycle analysis which again is based on time series analysis – one of the two focus areas of this conference.

Data gaps

Let me under this heading address some of the areas where we have to made improvements in statistical measurement.

Potential output

First, we need a more reliable measure of the potential output of the economy. Potential output, as we all know, is the maximum output that the economy can produce without putting pressure on the trend/average inflation rate. When the economy is operating at the potential level, aggregate demand and supply in the economy are balanced so that inflation tends to its long run expected value. A measure of potential output is therefore important not only to capture output data but to assess the inflation dynamics in the economy. The measurement of potential output depends on reliable data on employment and wages. I commented on both these statistics earlier while speaking about the quality of data. I am happy to note that the National Statistical Commission and the official statistical agencies are engaged in improving the quality of both employment and wage data. As that happens, we can look forward to a statistically more robust potential output measurement of the economy, and that in turn should improve our inflation watch.

In the aftermath of the global financial crisis, growth rates in major advanced economies remain sluggish despite large output gaps. Past experience with financial crises episodes suggests that growth and investment rates typically remain below the pre-crisis levels and potential growth rates take a hit. Similar dynamics are at play in the aftermath of the recent financial crisis. There is a perception that potential growth rates of advanced economies are likely to be lower than their pre-crisis trends. If so, given the growing trade and financial

globalization, potential growth rates of the fast-growing emerging market economies can be expected to be lower than the pre-crisis trends. Against this backdrop, it would be useful to re-assess India's potential growth rate consistent with our objective of low and stable inflation.

Services sector production and price index

A defining feature of the Indian economy is the large share of the services sector, unique at this level of per capita income. In order to assess our growth and inflation dynamics, we need reliable indicators of services sector production and prices. Admittedly, the output of services sector is less tangible than that of the traditional industrial sector, and that makes measurement of output and price levels of services intellectually challenging. I understand our official statistical agencies are engaged in addressing this challenge. I urge them to expedite their efforts.

Financial stability

Financial stability has come centre stage post crisis. As much as we now know that price stability and macroeconomic stability do not guarantee financial stability, we need to develop additional metrics to monitor financial stability. The IMF initiative on financial soundness indicators offers a model in this regard. Its "monitoring grid" focuses on four major areas, the first of which is financial market surveillance. For evaluation of the risk arising from imbalances or shocks, we need data on asset prices, as also on early warning indicators. The second area of focus of the IMF grid is macro-prudential surveillance, which concentrates on the impact of shocks on the financial sector. The third area consists of the analysis of macro-financial linkages, particularly on the credit spreads, credit to the private sector and balance sheet data for different sectors of the economy. The fourth area of focus is the surveillance of macroeconomic conditions.

To get started on adopting the IMF framework, we need to take stock of what data we have and of what quality. Off-hand, I can say that we need to improve our data and data analysis skills relating to real estate, capital flows, and with reference to banks and non-banks, data relating to credit, liquidity and market risk indicators.

Summing up

Let me now conclude. To sum up very briefly indeed, I have tried to communicate to you the importance of statistics and statistical analysis to the quality of RBI's policy formulation and emphasized areas where we need to improve the quality of data, areas where we need to learn to better interpret data and areas where we need to explore new frontiers. My exposition is, by no means, intended to be comprehensive, but I do hope it conveys the importance the Reserve Bank, as an institution, should attach to statistical analysis.

I wish this conference all success.