RESERVE BANK OF INDIA

# OCCASIONAL PAPERS

Vol. 9 No. 4 DECEMBER 1988

**The Reserve Bank of India Occasional Papers** contains contributions by the staff of the Reserve Bank of India. Articles, notes, and reviews published in this journal do not necessarily reflect the views of the Editorial Committee or of the Reserve Bank of India. Responsibility for the views expressed and for the accuracy of statements contained in the contributions rests with the author(s).



It is proposed to publish four issues in the year 1988, the current issue being the third in the year.

**Subscription rates for 1988 are:**

| | | | |
|---|---|---|---|
| Inland | : | (i) | Rs. 100 (for four issues of the 1988 volume inclusive of postage) |
| | | (ii) | Rs. 25 (for a single issue of the 1988 volume inclusive of postage) |
| Foreign | : | (i) | $32 (for four issues of the 1988 volume inclusive of postage by surface mail) |
| | | (ii) | $8 (for a single issue of the 1988 volume inclusive of postage by surface mail) |

For foreign air mail, please add $4 per issue.

There is no special rate for any category of subscribers. Subscriptions and orders, accompanied by cheques/drafts drawn on Bombay in favour of Reserve Bank of India, should be sent to:

The Director,
DRRP (Sales Section),
Department of Economic Analysis & Policy,
Reserve Bank of India,
Amar Building (Ground Floor),
Post Box No. 1036,
Bombay - 400 001 (India).

All other correspondence may be addressed to the Editor, Reserve Bank of India Occasional Papers, Department of Economic Analysis & Policy, Reserve Bank of India, Post Box No. 1036, Bombay - 400 023.

RESERVE BANK OF INDIA

# *OCCASIONAL PAPERS*

VOL. 9 NO. 4 DECEMBER 1988

ARTICLES Page



BOOK REVIEWS

Published by S.L. Shetty for the Reserve Bank of India and printed by him at
Modern Arts & Industries, 151, A-Z Industrial Estate, Ganpatrao
Kadam Marg, Lower Parel, Bombay 400 013.
Editor: S.L. Shetty

*Reserve Bank of India Occasional Papers*
Vol. 9, No. 4, December 1988

# On Some Estimates of Rural Indebtedness

S. P. Gothoskar*

THE All-India Debt & Investment Survey (AIDIS), being conducted decennially since 1951-52, is a comprehensive enquiry into assets and liabilities of rural households. The first survey conducted in 1951-52 tried to estimate rural credit, both from the demand side and the supply side. The demand side estimates of liabilities were obtained from a sample survey of rural households, wherein data on sources and quantum of debt of rural households were captured through household enquiry. On the supply side, information was obtained from lending institutions in the organised and unorganised sectors. The second survey carried out in 1962 gave the estimates of assets and liabilities at the state level. Both these surveys were carried out by the Reserve Bank of India (RBI).

The third survey of 1971-72 was undertaken jointly by the RBI and the National Sample Survey Organisation (NSSO), Government of India, which has been carrying out extensive household surveys. The sampling design of the NSSO usually has two samples, known as the Central sample, carried out by the field agency of the NSSO and the second sample, known as the State sample, carried out by the NSS agencies with the State Governments. In the survey of 1971-72, a third matching sample was also canvassed and this was known as the RBI sample.

The fourth survey was undertaken entirely by the NSSO in its 37th round conducted in January-December 1982, with active participation of the RBI in the design of the survey. However, there was no matching RBI sample in this survey. The RBI undertook the entire

* We regret to record the sad demise of the author of this article, Shri S.P. Gothoskar, on April 28, 1989 at London, following a cardiac arrest, while he was proceeding from Washington to take up an IMF assignment at Muscat, Sultanate of Oman. He contributed this article as Principal Adviser, Department of Statistical Analysis and Computer Services (DESACS), Reserve Bank of India, before leaving for Washington. We are publishing this article without any alteration as a mark of respect to the departed colleague.
—Editor

computer processing of both the Central and State samples. The reference period for collecting information was July 1981-June 1982.

So far, four such surveys have, thus, been carried out and the results of these surveys are being used widely in policy formulation and in empirical estimates on physical capital formation and financial characteristics of the household sector. These results can also be used to evaluate how the rural economy is responding to various schemes designed for rural uplift, particularly in regard to rural credit policies.

The results of the survey of 1982 indicate marked shifts in the pattern and quantum of rural indebtedness, to such an extent as to give rise to some doubts on the reliability of the survey estimates. It is, therefore, proposed to review in this article[1] the shifts in the profiles of rural indebtedness during 1970-80 in the context of other evidence available on rural credit. Section I of the article gives the salient results of the Survey. In this presentation, we have confined only to the proportions of rural households indebted and the quantum of outstanding household debt. Changes in the pattern of institutional sources and purpose of debt are also examined. Section II of the article gives rural credit data as available directly from the institutions supplying credit to rural households. Other evidence available on the rural household economy is also presented. The problems and limitations of the Survey of rural indebtedness and suggestions for the next survey due in 1991-92 are given in Section III of the article.

For the sake of brevity, we refer to the four surveys as 1951 Survey, 1961 Survey, 1971 Survey and 1981 Survey. The field work of the 1961 Survey was carried out in the calendar year 1962, with 30th June 1961 and 30th June 1962 as the reference dates for the information on assets and liabilities of households. The field work of 1971 Survey was carried out during July 1971 to September 1972 with 30th June 1971 and 30th June 1972 as the reference dates. The 1981 Survey was carried out in the calendar year 1982 with 30th June 1981 and 30th June 1982 as the reference dates. For details of concepts and definitions used in the surveys, a reference may be made to the RBI publications listed at the end of this article.

## I

The two marked features of the results of the 1981 Survey are (i) a steep fall in the percentage of indebted households to all rural households, from about 41 per cent in 1971 to about 19 per cent in 1981, and (ii) a significant shift in the sources of supply of credit to rural households, in favour of institutional agencies whose share increased from about 29 per cent to about 61 per cent. It is the steep fall in the percentage of indebted households and possible reasons and implications thereof that needs a careful analysis.[2]. To recapitulate, these data are given in Table Nos. 1 to 5.

Table No. 1 shows the changing profile of rural indebtedness during the three decades 1951-1981 and we see marked changes in the characteristics of indebtedness, particularly during 1971-1981, which has raised many eyebrows. The increase in the share of commercial banks in the total credit along with the rise in institutional credit was as expected. The cultivator households do not appear to be an exceptionally favoured class. There was more than 50 per cent increase in average outstanding debt per household, from Rs. 406 in 1961 to Rs. 653 in 1981. If these amounts are adjusted for the rate of inflation, then in real terms, the credit assistance would seem

**Table 1: Changing Profile of Rural Indebtedness**

| | Item | 1951 | 1961 | 1971 | 1981 |
|---|---|---|---|---|---|
| 1. | Percentage of indebted households to total households | N.E. | 62.8 | 41.3 | 19.4 |
| 2. | Percentage of indebted cultivator households to total cultivator households | N.E. | 66.7 | 44.4 | 21.7 |
| 3. | Average amount of debt outstanding per household (in Rs.) | N.E. | 406<br>(647) | 487<br>(1,179) | 653<br>(3,366) |
| 4. | Institutional share in total rural credit (per cent) of which, | 7.2 | 17.3 | 29.2 | 61.2 |
| | a) share of cooperatives | 3.5 | 10.4 | 20.1 | 28.6 |
| | b) share of commercial banks | N.A. | 0.3 | 2.2 | 28.0 |

Note: Figures in brackets relate to amount per reporting household.
N.E. means not estimated at all-India level.
Amount figures are rounded.
N.A. means not available.

to have gone down. The amounts of debt outstanding per reporting[3] household are now much higher, consistent with the lower incidence of indebtedness.

Table No. 2 shows the aggregate estimates from the three Surveys. Here we see that the absolute number of borrowing households has declined from 3.18 crores in 1971 to 1.82 crores in 1981. This seems contradictory as compared to about 4 to 5 times increase in the number of borrowing accounts of the co- operative societies and commercial bank branches in rural areas. We have examined this point further in Section II of this article. The main share in this fall of about 14 million households can be attributed to households borrowing from the non-institutional sources, that too, for non-productive purposes.

Table No. 3 brings out the differentials in profiles of indebtedness amongst the cultivator households and other rural households, over the three surveys. Though the fall in the percentage of indebted households in the two classes is about the same, the gap between the average amounts of debt outstanding per household has widened. Whereas for the cultivator household, the average debt increased from Rs.473 in 1961 to Rs.793 in 1981, for the other (non-cultivator) household, it declined from Rs. 224 to Rs.202 during the

**Table 2: Aggregate Estimates from the AIDIS 1961, 1971 and 1981**

| | Item | 1961 | 1971 | 1981 |
|---|---|---|---|---|
| 1. | Total number of indebted rural households | N.A. | 3,18 | 1,82 |
| | of which, | | | |
| | a) indebted cultivator households | N.A. | 2,48 | 1,55 |
| | b) other indebted households | N.A. | 70 | 27 |
| 2. | Aggregate outstanding amount of debt | 2,789 | 3,752 | 6,111 |
| | of which, | | | |
| | a) against cultivator households | 2,380 | 3,290 | 5,662 |
| | b) against other households | 409 | 462 | 449 |
| 3. | Estimated aggregate amount of co-operative credit | 255 | 753 | 1,754 |
| 4. | Estimated aggregate amount of commercial bank credit | 12 | 82 | 1,717 |

Note: Number of households in lakhs; Amount in Rs. crores.

**Table 3: Differentials in Cultivator & Other Rural Household Indebtedness**

| Item | Cultivator Households | | | Other Households | | |
|---|---|---|---|---|---|---|
| | 1961 | 1971 | 1981 | 1961 | 1971 | 1981 |
| 1. Percentage of indebted households | 66.7 | 44.4 | 21.7 | 52.0 | 33.3 | 12.0 |
| 2. Average amount of debt outstanding per household (Rs.) | 473 (708) | 590 (1,330) | 793 (3,656) | 224 (430) | 217 (652) | 202 (1,678) |
| 3. Average amount of debt for productive purposes per household (Rs.) | 190 | 320 | 567 | 65 | 45 | 76 |
| 4. Average amount of institutional credit per household (Rs.) | 75 | 187 | 502 | 21 | 23 | 74 |

Note: Figures in brackets relate to amounts per reporting household.

**Table 4: Debt Owed to Different Credit Agencies by Rural Households**

(*Per cent*)

| | 1951 | 1961 | 1971 | 1981 |
|---|---|---|---|---|
| 1. Institutional | 7.2 | 17.3 | 29.2 | 61.2 |
| 1.1 Government | 3.7 | 6.6 | 6.7 | 4.0 |
| 1.2 Co-operative Society/Banks | 3.5 | 10.4 | 20.1 | 28.6 |
| 1.3 Commercial Banks | — | 0.3 | 2.2 | 28.0 |
| 1.4 Insurance | — | — | 0.1 | 0.3 |
| 1.5 Provident Fund | — | — | 0.1 | 0.3 |
| 2. Non-Institutional | 92.8 | 82.7 | 70.8 | 38.8 |
| 2.1 Landlords | 3.5 | 1.1 | 8.6 | 4.0 |
| 2.2 Agricultural moneylenders | 25.2 | 47.0 | 23.1 | 8.6 |
| 2.3 Professional moneylenders | 46.4 | 13.8 | 13.8 | 8.3 |
| 2.4 Traders | 5.1 | 7.5 | 8.7 | 3.4 |
| 2.5 Relatives/Friends | 11.5 | 5.8 | 13.8 | 9.0 |
| 2.6 Others | 1.1 | 7.5 | 2.8 | 5.5 |
| 3. Total | 100.0 | 100.0 | 100.0 | 100.0 |

same period. If the overall supply of credit to rural areas has increased along with larger number of indebted cultivating households, then this would mean that larger share of institutional credit has gone to cultivating households. However, if we see the estimates of debt incurred for productive purposes, we find that other (non-cultivator) households have incurred debt more for non-productive purposes, which is generally not favoured by the institutional agencies. It is also interesting to note that the gap between debt for productive purposes and the supply by institutional agencies has considerably narrowed down and about 90 per cent of borrowing needs of both classes of rural households for productive purposes are now being met by institutional agencies.

Table Nos. 4 & 5 give percentage distributions of outstanding debt classified by the supply agencies and by purposes of loan. These distributions relate to amount of debt outstanding. As the same household may borrow from more than one agency or for more than one purposes, it has not been possible to give distributions of borrowing households classified by sources of supply or purpose of credit. However, it was observed that the percentage of households borrowing from multiple sources has fallen during 1971-1981 and was of the order of 14 per cent of total indebted households in 1981. Agricultural and professional money lenders have lost their ground

**Table 5: Distribution of Outstanding Debt according to purpose of borrowing**

*(Per cent)*

| Purpose | 1971 | 1981 |
|---|---|---|
| 1. Capital expenditure in farm business | 31.2 | 42.4 |
| 2. Current expenditure in farm business | 13.5 | 17.6 |
| 3. Capital expenditure in non-farm business | 3.7 | 7.2 |
| 4. Current expenditure in non-farm business | 1.7 | 1.7 |
| Sub-total Productive purposes | 50.1 | 68.9 |
| 5. Household expenditure | 40.9 | 22.4 |
| 6. Expenditure on litigation | 0.8 | 0.2 |
| 7. Repayment of debt | 1.8 | 0.8 |
| 8. Financial investment expenditure | 0.7 | 0.9 |
| 9. Others | 5.1 | 6.6 |
| 10. More than one purpose | 0.3 | — |
| 11. Unspecified | 0.3 | 0.2 |
| Total | 100.0 | 100.0 |

considerably, from more than 70 per cent in 1951 to about 17 percent of outstanding debt in 1981. It may be recalled that it was the 1951 Survey finding on the high share of private money lenders in rural household debt that had led to far reaching policy formulations on rural banking and the 1981 survey findings would indicate that a good deal of success has been achieved in achieving the goals set. Classification of debt by purpose of loan shows that debt for productive purposes now accounts for nearly 70 percent of the outstanding loan as against 50 percent in 1971. These are the changes in right direction.

The average value of assets of a rural household was Rs. 36,090. But for a cultivator household, the average value was Rs. 44, 524 as against the average value of Rs. 8,974 for a non- cultivator household. Overall, the real rate of growth in total assets during the decade was estimated at about 4 per cent per annum. Land accounted for about 62 per cent of the total assets of rural household. There was a marginal decline in the share of land and buildings in the total assets from 84.6 percent to 82.8 percent over the decade. Thus, the average debt per household formed only 4 percent of assets in 1971 and 2 per cent in 1981 (Table No. 6).

According to the 1981 Survey about 19 percent of the rural households were in debt. Debt in kind was relatively negligible. The proportion of households reporting debt was 22 percent for the cultivator households and 12 percent for the non-cultivator households. The average debt per reporting[3] household was Rs. 3,656 for a cultivator household and Rs. 1,678 for a non-cultivator household.

## II

The decade 1970-1980 had seen a phenomenal expansion in the commercial banking network, that too in the rural areas.[4] In June 1970, there were only 1,692 rural centres having 1,790 commercial bank branches. By June 1980, the network of 12,889 branches had been spread in a total of 11, 968 rural centres. Most of these centres were single branch centres. In June 1980, the average amount of deposit per account was of the order of Rs. 1,100-Rs. 1,400, while the average amount of credit per account was about Rs.2,100 to

**Table 6: Assets & Liabilities of Rural Households**

| Item | Cultivator Households | | Other Households | | All Rural Households | |
|---|---|---|---|---|---|---|
| | 1971 | 1981 | 1971 | 1981 | 1971 | 1981 |
| 1. Average value of assets (in Rs.) | 14,627 (14,630) | 44,524 (44,529) | 2,613 (2,647) | 8,974 (9,106) | 11,311 (11,355) | 36,090 (36,133) |
| 2. Percentage share of land in assets | 68.6 | 64.1 | 31.3 | 30.8 | 66.2 | 62.1 |
| 3. Percentage of indebted households | 44.4 | 21.7 | 33.3 | 12.0 | 41.3 | 19.4 |
| 4. Average debt per household (in Rs.) | 590 (1,330) | 793 (3,656) | 217 (652) | 202 (1,678) | 487 (1,179) | 653 (3,366) |
| 5. Debt-Asset Ratio (per cent) | 4.1 | 1.8 | 8.5 | 2.3 | 4.4 | 1.8 |

Note: Figures in brackets relate to per reporting household.

Rs.2,800. Nearly 60 per cent of rural credit by commercial banks was extended for agricultural purposes.

By the end of 1981, commercial banks had more than one crore of credit accounts in rural areas with an aggregate outstanding amount of Rs. 3,438 crores (Table No.7). Comparative estimate for the year 1971 was Rs.159 crores. But at that time, commercial banks had not made any significant inroads into rural areas. In the co-operative sector, the primary agricultural credit societies (PACs) and primary land development banks (PLDBs) had together about 148 lakhs accounts with an aggregate outstanding credit of Rs. 1,233 crores in 1971. But by 1981, the number of accounts had shot up to 642 lakhs, with an outstanding credit of Rs. 3,607 crores. These data would roughly indicate that the outstanding rural credit at the end of 1981 by the two main institutional sources was of the order of Rs.7,000 crores. This could be treated as an estimate of rural credit from the supply side. On the other hand, the estimate of outstan-

**Table 7: Aggregate Demand and Supply of Credit to Rural Households**

| Item | | 1971 | 1981 |
|---|---|---|---|
| **I. Estimated from the Survey** | | | |
| a) No. of indebted households (in thousands) | i) Co-operatives | 58,93 | 60,70 |
| | ii) Commercial banks | 3,39 | 29,09 |
| | iii) Total | 62,32* | 89,79 |
| b) Outstanding loan from (in Rs. crores) | i) Co-operatives | 753 | 1,754 |
| | ii) Commercial banks | 82 | 1,717 |
| | iii) Total | 835 | 3,471 |
| **II. As available from Institutional Sources** | | | |
| a) No. of borrowal accounts of (in thousands) | i) Co-operatives | 147,70 | 642,19 |
| | ii) Commercial banks | + | 101,73 |
| | iii) Total | + | 743,92 |
| b) Outstanding credit (in Rs. crores) | i) Co-operatives | 1,233 | 3,607 |
| | ii) Commercial banks | 159 | 3,438 |
| | iii) Total | 1,392 | 7,045 |

* There may be some duplication in this total as the same household might have borrowed from both the sources.

\+ Not available.

ding cash liabilities to institutional sectors of rural households obtained from the AIDIS could be treated as the demand side estimates. How do these two sets of estimates compare? Table No. 7 gives the comparative data.

While presenting the 1971 Survey results, the RBI did attempt such a comparison. It showed that, as compared with the Population Census data of 1971, the total number of rural household in the country estimated from the Survey was lower by about 3 percent. But the extent of under-estimation in the estimated debt owed by households to co-operatives was much higher, the Survey estimate being about 61 per cent of the official figure of outstanding against primary agricultural credit societies and primary land development banks. The extent of under-estimation varied from about -67 per cent in Andhra Pradesh to over-estimation of + 38 per cent in Gujarat. A quick test check in Maharashtra and Uttar Pradesh showed that the households surveyed had under-reported their debt by over 17 per cent and 28 per cent, respectively. In states like Assam, Bihar, Haryana, Karnataka, Kerala, Madhya Pradesh, Orissa, Punjab, Rajasthan and Tamil Nadu, the under-reporting of co-operative debt in 1971 was of the order of about 45 per cent to 65 per cent. Such an under-reporting of debt even in the earlier Debt and Investment Survey of 1961 has been mentioned. It has been stated that the under-reporting at that time was lower at 17 per cent at the beginning of the survey period.

The demand side estimates of rural household credit by co-operative societies and commercial banks obtained from the Survey, were Rs. 1,754 crores and Rs. 1,717 crores, respectively. Thus, in comparison with the estimate of credit from the supply side, the demand side estimate was an under estimate by about 50 per cent in 1981 as against an under estimate of about 40 per cent in 1971.

The under estimation in total debt could arise either because of under estimation of total number of indebted households or because of under reporting of debt outstanding against a borrowing household or both. We had earlier pointed out that the absolute number of borrowing households showed a fall by 14 million households during 1971-1981, whereas the number of credit accounts with the co-operative institutions and commercial banks in rural areas had gone

up fourfold. On the other hand, the average amount of debt outstanding from commercial banks per borrowing (reporting) household, as estimated from 1981 Survey, was Rs.5,903 against the average amount of credit per account of about Rs.3,400, as estimated from the supply side. We have no information about the number of borrowing accounts per household but with some adjustment, the two estimates of amounts of credit would appear to be quite close and, therefore, it would be clear that the main reason for the overall under estimation of total rural credit from commercial banks and co-operative societies is on account of the under estimation of the total number of borrowing households, i.e., households reporting cash debt liabilities.

The accuracy of the estimate of percentage of indebted households can also be verified against some indirect evidence on

**Table 8: Statewise Profile of Rural Indebtedness in 1971**

| Sr. No. | Name of the State | Beginning of the 1971 survey period | | End of the 1971 survey period | |
|---|---|---|---|---|---|
| | | P | A | P | A |
| 1. | Andhra Pradesh | 47.4 | 637 | 50.8 | 734 |
| 2. | Assam | 25.1 | 177 | 24.3 | 157 |
| 3. | Bihar | 38.4 | 288 | 37.2 | 310 |
| 4. | Gujarat | 47.7 | 930 | 36.6 | 680 |
| 5. | Haryana | 35.3 | 921 | 43.5 | 1,263 |
| 6. | Himachal Pradesh | 37.5 | 573 | 41.4 | 701 |
| 7. | Jammu & Kashmir | 47.7 | 356 | 50.3 | 374 |
| 8. | Karnataka | 49.2 | 731 | 52.5 | 873 |
| 9. | Kerala | 34.3 | 366 | 38.5 | 403 |
| 10. | Madhya Pradesh | 40.3 | 395 | 43.8 | 493 |
| 11. | Maharashtra | 45.0 | 591 | 48.9 | 658 |
| 12. | Orissa | 29.2 | 163 | 29.8 | 181 |
| 13. | Punjab | 54.4 | 1,039 | 55.9 | 1,217 |
| 14. | Rajasthan | 55.9 | 893 | 61.8 | 1,165 |
| 15. | Tamil Nadu | 49.7 | 705 | 54.2 | 790 |
| 16. | Uttar Pradesh | 37.3 | 343 | 42.2 | 415 |
| 17. | West Bengal | 34.5 | 188 | 40.9 | 261 |
| | All-India | 41.3 | 487 | 44.0 | 552 |

Note: P denotes percentage of indebted households to total households.

A denotes average amount of cash dues outstanding for household (in Rs.).

the incidence of indebtedness among rural households. In the first place, information on assets and liabilities of the same set of households has been collected in the 1981 Survey at two points of time, viz., at the beginning of the year of Survey (as on 30th June 1981) and at the end of the reference period (i.e. as on 30th June 1982). Such information was also collected in the 1971 Survey. These estimates are given in Table Nos. 8 and 9. Secondly, the acquisition of additional assets/liabilities by these households during the period of Survey and their sources were also recorded in the two surveys.

In the 1971 Survey, 41.3 per cent of households were indebted

**Table 9: Statewise Profile of Rural Indebtedness in 1981**

| Sr. No. | Name of the State | Beginning of the 1981 survey period | | Gross capital expenditure during 1981-82 | | |
|---|---|---|---|---|---|---|
| | | P | A | PE | AE | BH |
| 1. | Andhra Pradesh | 26.3 | 934 | 57.7 | 622 | 1.1 |
| 2. | Assam | 4.3 | 50 | 46.0 | 330 | 0.1 |
| 3. | Bihar | 12.2 | 202 | 62.8 | 348 | 0.4 |
| 4. | Gujarat | 17.4 | 848 | 50.5 | 661 | 0.5 |
| 5. | Haryana | 10.5 | 962 | 71.2 | 1,345 | 0.1 |
| 6. | Himachal Pradesh | 12.3 | 353 | 55.6 | 742 | 0.5 |
| 7. | Jammu & Kashmir | 8.9 | 251 | 77.8 | 1,297 | 0.4 |
| 8. | Karnataka | 23.8 | 1,249 | 59.8 | 1,001 | 0.9 |
| 9. | Kerala | 28.2 | 951 | 63.3 | 1,576 | 1.0 |
| 10. | Madhya Pradesh | 19.7 | 568 | 82.1 | 554 | 0.1 |
| 11. | Maharashtra | 22.0 | 848 | 61.0 | 717 | 0.5 |
| 12. | Orissa | 19.2 | 346 | 71.5 | 395 | 0.3 |
| 13. | Punjab | 18.9 | 1,499 | 59.5 | 1,648 | 0.3 |
| 14. | Rajasthan | 24.2 | 1,157 | 41.4 | 890 | 0.4 |
| 15. | Tamil Nadu | 28.5 | 1,009 | 54.7 | 565 | 1.1 |
| 16. | Uttar Pradesh | 17.4 | 448 | 80.0 | 750 | 0.3 |
| 17. | West Bengal | 17.5 | 297 | 57.4 | 414 | 0.5 |
| | All-India | 19.4 | 653 | 63.6 | 685 | 0.5 |

Notes: P denotes percentage of indebted households to total households.
A denotes average amount of cash dues outstanding per household (in Rs.).
PE denotes percentage of households reporting capital expenditure.
AE denotes amounts (in Rs.) of gross capital expenditure.
BH denotes percentage of households reporting borrowings during the year.

at the beginning of the survey as against 44 per cent at the end of the survey. Liabilities per household amounted to Rs.487 and Rs.552 for the two periods, respectively. The two estimates are fairly consistent at the State level.

In the 1981 Survey, at the beginning of the year, 19.4 per cent of the households were indebted, with a liability of Rs. 653 per household. Similar data for the end of the period have yet not been published. But considering the rate of inflation and expansion in institutional credit, the average liability of Rs. 653 per household in 1981 appears to be quite low as compared to the average liability of Rs. 552 at the end of 1971.

The 1981 Survey results also show that about 64 per cent of rural households have incurred a gross capital expenditure of Rs. 685 per household during 1981-82. But the Survey has further estimated that only 0.5 per cent of these households have financed this expenditure from borrowings. Thus, the incidence of indebtedness as at the end of the period is unlikely to be significantly different from that at the beginning of the period. Despite the high proportions of households incurring capital expenditure during the year, the incidence of indebtedness at the beginning of the year and the shares of borrowings in meeting capital expenditure are quite low.

The estimates of proportions of indebted households and the average amount of borrowing per household can also be examined against the average value of assets per household. Since shares of borrowings for productive purposes in total borrowings are high, we also take into consideration the shares of land in total household assets. We have also worked out the debt to land assets ratios in the two periods. These data are given in Table Nos. 10 and 11 for different States in India.

In these Tables, an attempt has been made to juxtapose the average value of assets per household and the proportion of land in the assets against the indebtedness of households in different states. This cross-sectional data at two points of time does not indicate any definite relationship between assets and indebtedness.

The 1971 Survey reveals that against an average value of assets at Rs. 11,311 with 66 per cent of the assets in land, the average

**Table 10: Statewise percentage of indebtedness and average debt and assets per household in 1971**

| Sr. No. | State | Percentage of indebted households | Average debt (in Rs.) | Average assets (in Rs.) | Percentage share of land in assets | Debt to land assets ratio (per cent) |
|---|---|---|---|---|---|---|
| (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| 1. | Andhra Pradesh | 47.4 | 637 | 8,080 | 67.6 | 11.7 |
| 2. | Assam | 25.1 | 177 | 7,833 | 60.0 | 3.8 |
| 3. | Bihar | 38.4 | 288 | 12,828 | 74.3 | 3.0 |
| 4. | Gujarat | 47.7 | 930 | 12,874 | 57.3 | 12.6 |
| 5. | Haryana | 35.3 | 921 | 27,139 | 71.5 | 4.7 |
| 6. | Himachal Pradesh | 37.5 | 573 | 22,673 | 52.3 | 4.8 |
| 7. | Jammu & Kashmir | 47.7 | 356 | 15,260 | 55.3 | 4.2 |
| 8. | Karanataka | 49.2 | 731 | 10,032 | 65.3 | 10.4 |
| 9. | Kerala | 34.3 | 366 | 11,615 | 63.2 | 5.0 |
| 10. | Madhya Pradesh | 40.3 | 395 | 10,520 | 65.0 | 5.8 |
| 11. | Maharashtra | 45.0 | 591 | 11,682 | 71.1 | 7.1 |
| 12. | Orissa | 29.2 | 163 | 6,023 | 66.2 | 4.1 |
| 13. | Punjab | 54.4 | 1,039 | 31,833 | 72.3 | 4.5 |
| 14. | Rajasthan | 55.9 | 892 | 12,754 | 55.8 | 12.5 |
| 15. | Tamil Nadu | 49.7 | 705 | 6,827 | 64.7 | 16.0 |
| 16. | Uttar Pradesh | 37.3 | 343 | 13,531 | 66.4 | 3.8 |
| 17. | West Bengal | 34.5 | 188 | 7,331 | 60.3 | 4.3 |
| | All-India | 41.3 | 487 | 11,311 | 66.3 | 6.5 |

**Table 11: Statewise percentage of indebtedness and average debt and assets per household in 1981**

| Sr. No. | State | Percentage of indebted households | Average debt (in Rs.) | Average assets (in Rs.) | Percentage share of land in assets | Debt to land assets ratio (per cent) |
|---|---|---|---|---|---|---|
| (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| 1. | Andhra Pradesh | 26.3 | 934 | 26,247 | 60.7 | 5.9 |
| 2. | Assam | 4.3 | 50 | 20,503 | 50.0 | 0.5 |
| 3. | Bihar | 12.2 | 202 | 32,347 | 65.6 | 1.0 |
| 4. | Gujarat | 17.4 | 848 | 36,876 | 53.0 | 4.3 |
| 5. | Haryana | 10.5 | 962 | 90,950 | 63.9 | 1.7 |
| 6. | Himachal Pradesh | 12.3 | 353 | 62,558 | 53.5 | 1.1 |
| 7. | Jammu & Kashmir | 8.9 | 251 | 59,001 | 58.3 | 0.7 |
| 8. | Karnataka | 23.8 | 1,249 | 33,052 | 54.6 | 6.9 |
| 9. | Kerala | 28.2 | 951 | 76,479 | 68.2 | 1.8 |
| 10. | Madhya Pradesh | 19.7 | 568 | 29,725 | 61.2 | 3.1 |
| 11. | Maharashtra | 22.0 | 848 | 35,077 | 65.4 | 3.7 |
| 12. | Orissa | 19.2 | 346 | 17,630 | 64.3 | 3.0 |
| 13. | Punjab | 18.9 | 1,499 | 96,631 | 71.8 | 2.2 |
| 14. | Rajasthan | 24.2 | 1,157 | 40,888 | 53.2 | 5.3 |
| 15. | Tamil Nadu | 28.5 | 1,009 | 19,520 | 50.1 | 10.3 |
| 16. | Uttar Pradesh | 17.4 | 448 | 44,660 | 65.0 | 1.5 |
| 17. | West Bengal | 17.5 | 297 | 20,710 | 59.0 | 2.4 |
| | All-India | 19.4 | 653 | 36,090 | 62.1 | 2.9 |

outstanding debt was Rs. 487. The debt-land asset ratio was only 6.5 per cent. In Haryana and Punjab, where the average value of assets and proportion of land in it were quite high, the debt-land asset ratio was quite low. On the other hand, this ratio was very high in middle level states like Andhra Pradesh, Gujarat, Rajasthan and Tamil Nadu. In Assam, Bihar and Uttar Pradesh, the incidence of indebtedness and the average value of outstanding debt were low, even though Bihar and Uttar Pradesh had recorded high values of average assets per household.

In 1981 Survey, the picture has changed. The average value of assets had tripled but the average debt had increased by only about 30-40 per cent. As a result, the debt-land asset ratio fell sharply from 6.5 per cent to 2.9 per cent. In this case also, we do not see any significant relationship between assets and indebtedness.

There is another evidence available from the demand side, which corroborates the findings of the 1981 Survey. This is from the National Council of Applied Economic Research (NCAER) which has been carrying out periodic sample surveys on income, savings and consumption of rural households. Bhatty and Vashishtha[5] have compared the economic status of an average rural household for the years 1970-71 and 1981-82. These results for a common set of households surveyed in both the periods are also given.

About 72 per cent of the sample canvassed by the NCAER in 1981-82 was also covered in the Survey of 1970-71. It was observed that the savings rate of rural households had increased from 3.26 per cent in 1970-71 to 10.25 per cent in 1981-82. The corresponding increase in the rate of investment was from 9.11 per cent to 13.03 per cent. The increase in the savings was mainly contributed by the rise in the rate of financial saving from $-4.41$ per cent to $+2.97$ per cent. There was also a decline in the liabilities to gross financial saving ratio, from 1:0.246 to 1:2.07. In 1970-71, all categories of households had negative financial savings. In 1981-82, only the agricultural wage earners had negative financial savings. A reduction in the liabilities of the rural household sector to indigenous money- lenders and the like was also recorded. This has been attributed to rapid spread of bank branches in rural India and expansion of price support policies undertaken by the Food Corpora

tion of India. The gross financial saving rate had also increased from 1.44 per cent in 1970-71 to 5.75 per cent in 1980-81.

According to Bhatty and Vashishtha, there was only a marginal decline in the physical saving rate during the decade. In fact, the real gross fixed capital formation had increased by 18 per cent in this period. The gap between the saving and investment rate had narrowed down considerably during the decade. A steeper rise in the saving rate during the decade indicated greater reliance of the rural households on internal finance.

Considering the common sample between the NCAER Surveys, it was observed that the aggregate saving rate rose from 5.39 per cent in 1970-71 to 8.55 per cent in 1980-81. The change from negative financial savings to high positive financial savings was also observed in respect of this common sample. In this case also, only the agricultural wage earner had a negative financial saving in 1980-81.

It is known that the PACs and PLDBs do not have large deposits of households and the credit extended by them to households is mainly funded through their own borrowings from state/district level organisations. Thus, the cooperative lendings in rural areas bring in additional funds to rural areas. But in the case of commercial banks, there appears to be net out-flow of funds from rural areas. Whereas the rural deposits with commercial banks were of the order of Rs. 540 crores in 1972, they received by way of commercial bank credit only Rs. 257 crores. In 1981, rural areas received credit of Rs. 3,438 crores against deposits of Rs. 5,834 crores. More than 80 per cent of these deposits were held by households. Thus, as far as the commercial banks are concerned, at the aggregate rural level, households were net lenders, corroborating NCAER hypothesis of net financial surplus with rural households.

A study by Iqbal[6] based on the NCAER surveys on Indian farm households between the period 1968 and 1971 gives interesting results. The study showed that the adopters of High Yielding Varieties borrow far more than the non-adopters (Rs. 759 vs. Rs. 254 per household of Rs. 208 vs. Rs. 114 per hectare cultivator). It was also noticed that from the supply agency side, the adoption of the improved practices was an important consideration for lending. It was

noted that the bulk of the lending of official lending agencies goes to high yielding variety adopter households (81 per cent of direct government loans and 62 per cent of co-operative credit society loans). Small farmers constituted the single most important outlet for the private money lender finance, constituting about 70 per cent of the total credit disbursed by the money lenders and about 88 per cent of the money lenders clients. Thus, the indebtedness of rural households seems to have been influenced by a variety of factors.

The high income and financial saving has possibly given rise to low incidence of indebtedness. This hypothesis, supported by the NCAER surveys on the demand side, can be examined from the supply side also. As per the National Accounts Statistics (NAS) during 1970-71 to 1980-81, agricultural production index, based on triennium averages, increased at compound annual growth rate of 2.5 per cent. In real terms, the gross NDP from agriculture increased by 2 per cent per annum (compounded), but more importantly for our purpose, we should consider the nominal increase, which was of the order of 9.3 per cent per annum (compounded). Some portion of the additional incomes generated would be available with the households for investment purposes, reducing their reliance on borrowed funds as pointed out by NCAER survey.

The short-term borrowings of cultivator households could also be more influenced by agricultural performance in the periods immediately preceding the survey reference periods. Prior to 1971-72, the agricultural production index had increased by more than 33 per cent. In 1979-80, the index had fallen by 15 per cent but in the next two years, it rose by 20 per cent. It is a matter for further investigation as to how the agricultural production performance over short periods affect the demand for short term and long term loans and repayments by households.

Table Nos. 12 and 13 give data on credit operations of the PACs in rural areas. In 1981-82, the outstanding credit of the PACs was Rs. 2,966 crores of which about 72 per cent were short term loans. During the same year, they had disbursed credit of Rs. 2,110 crores, of which nearly 85 per cent were short term loans. Even accepting that some of the short-term loans might subsequently get converted into medium term loans, the annual figures of loans issued and loans

outstanding indicate a fast roll-over of loans. We have no data on exact periods in which loans are issued and are held by cultivators. Therefore, data on loans outstanding on a particular date could give an over-estimate or an under-estimate of indebtedness, depending on the period in which the reference date falls. Perhaps, 30th June (which is the reference date for AIDIS) may be a bad choice, considering the interlude between the end of earlier agricultural year and beginning of the next year.

To summarise the above discussion, it may be pointed out that compared to the supply side estimates, the incidence of rural indebtedness and the quantum of debt as estimated from the 1981 Survey appear to be grossly under-estimated. However, the cash debt outstanding per (reporting) borrowing household is close to the of-

**Table 12: Outstanding Loans given by PACS**

(Amount in Rs. Crores, June end)

| Year | Short-term | Medium-term | Total |
|---|---|---|---|
| 1969-70 | 596 | 115 | 711 |
| 1970-71 | 647 | 137 | 784 |
| 1971-72 | 696 | 162 | 858 |
| 1979-80 | 1,697 | 667 | 2,364 |
| 1980-81 | 1,908 | 700 | 2,608 |
| 1981-82 | 2,149 | 817 | 2,966 |
| 1982-83 | 2,225 | 882 | 3,017 |

**Table 13: Loans Issued by PACs during the Year**

*(in Rs. Crores)*

| Year | Short-term | Medium-term | Total |
|---|---|---|---|
| 1969-70 | 488 | 52 | 540 |
| 1970-71 | 519 | 59 | 578 |
| 1971-72 | 541 | 74 | 615 |
| 1979-80 | 1,358 | 266 | 1,624 |
| 1980-81 | 1,519 | 227 | 1,746 |
| 1981-82 | 1,796 | 314 | 2,110 |
| 1982-83 | 1,908 | 383 | 2,291 |
| 1983-84 | 2,158 | 341 | 2,499 |

ficial figure of outstanding credit per account, with suitable adjustments. That the percentage of indebted households at the beginning of the year could be an under-estimate is also corroborated by the very low proportion of households financing their capital expenditure from borrowings.

However, the NCAER survey has indicated that the gap between savings and investment of rural households has narrowed down during the Seventies and an average rural household has now a net financial surplus. The increases in agricultural incomes in the periods immediately preceding the survey reference period could also possibly lead to greater reliance by the rural households on internal sources of finance rather than on external borrowings.

Information on the movements in the debt outstanding on a particular reference date is not available. This could have also affected the survey estimate. Though this reference date has been adopted uniformly as 30th June in all the previous surveys, it may now be necessary to study the withdrawals and repayment periods in short term credit to rural households.

Our impressions on large scale expansion of bank credit in rural area, based on targets of priority sector and agricultural advances are really valid for post-1980 developments. Between 1970-1980, commercial banks were in the process of widening their net-work through establishment of bank branches in unbanked centres. The rural credit share in the total commercial bank credit had gone up from 3.1 per cent in 1971 to only 12.0 per cent in 1981. Thus, the 1981 Survey results should not be viewed with scepticism, simply because the incidence of indebtedness and the volumes of estimated debt looks small by today's standards.

Moreover, the rural areas referred to in the Survey relate to villages having population of less than 5,000 while the Banking Statistics treats villages with less than 10,000 population as rural. We do not readily have the information about the number of bank branches in villages with population less than 5,000 persons and as such, we do not know whether the expansion of banking facilities has really benefitted the rural areas as defined in the Survey.

## III

In this Section, we discuss some technical details of the sampling design and precision that can be attached to the survey estimates. The possible reasons for errors in the estimates of the 1981 Survey and a few suggestions for the next survey due in 1992 are also given.

The sampling design adopted for the 1971 and 1981 Surveys was a general two stage stratified sampling, usually adopted in the NSS. Villages in each state were grouped into agro-economic regions, which formed the strata and were selected according to the probability proportional to size, with replacement. Households in selected villages were grouped into four sub-strata, depending on the size of land possessed. The sample of villages selected for survey is divided into two sub-samples and allocated to NSSO under the Central Government and the NSS organisations with the State Governments. Each sub-sample is further divided into two sub-samples.

In the 1961 Survey, 80,000 households were selected from 2,000 villages all over the country. In the 1971 Survey, because of the matching RBI sample, the number of households surveyed was much larger at over one lakh households from about 12,500 villages. In the 1981 Survey, only 61,000 households were selected from about 7,700 villages, as there was no matching RBI sample. The reduction in sample size, either in terms of the number of villages or number of households surveyed need not by itself affected the survey estimates.

The survey sampling design was a general design adopted in most of the rounds of the NSS and cannot be considered as the best design for a survey of debt and investment. Ideally, villages should have been stratified as the banked centres and the unbanked centres, and then the required number of villages should have been allocated to the two strata in proportion to the number of households in them.

As it is, the 1981 Survey estimates appear to have large sampling errors. Separate estimates based on the Central sample and the State sample are available and we have given these two sets of estimates for different states in India, for two items, viz., the proportion of indebted households and the average amount of loan per household, in Table Nos. 14 and 15.

**Table 14: Sub-samplewise Estimates of Indebted Households in Selected States in 1981**

*(Per cent to total rural households)*

| Name of the State | Central sample | | | State sample | | | Central & State sample pooled |
|---|---|---|---|---|---|---|---|
| | SS-1 | SS-2 | combined | SS-1 | SS-2 | combined | |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| 1. Andhra Pradesh | 35.14 | 34.35 | 34.75 | 17.75 | 17.54 | 17.64 | 26.47 |
| 2. Assam | 3.05 | 6.57 | 4.83 | 6.45 | 3.06 | 4.79 | 4.81 |
| 3. Bihar | 16.08 | 19.71 | 17.81 | 8.04 | 8.84 | 8.46 | 13.24 |
| 4. Gujarat | 20.14 | 17.51 | 18.85 | 17.90 | 18.46 | 18.20 | 18.53 |
| 5. Haryana | 28.31 | 24.02 | 26.09 | 5.16 | 3.77 | 4.48 | 10.57 |
| 6. Himachal Pradesh | 12.04 | 13.65 | 12.90 | 12.44 | 10.21 | 11.43 | 12.32 |
| 7. Jammu & Kashmir | 7.38 | 8.77 | 8.21 | 9.60 | 9.91 | 9.75 | 8.91 |
| 8. Karnataka | 32.37 | 29.16 | 30.74 | 17.69 | 15.40 | 16.50 | 24.00 |
| 9. Kerala | 25.43 | 29.30 | 27.37 | 32.94 | 26.47 | 29.67 | 28.46 |
| 10. Madhya Pradesh | 21.18 | 18.80 | 20.21 | 22.02 | 20.48 | 21.24 | 20.47 |
| 11. Maharashtra | 23.76 | 21.16 | 22.48 | 23.85 | 20.96 | 22.45 | 22.47 |
| 12. Orissa | 30.48 | 22.90 | 26.90 | 15.47 | 12.83 | 14.06 | 20.45 |
| 13. Punjab | 29.83 | 21.04 | 25.52 | 12.55 | 13.96 | 13.27 | 19.53 |
| 14. Rajasthan | 30.70 | 34.56 | 32.65 | 19.27 | 14.04 | 16.76 | 24.74 |
| 15. Tamil Nadu | 39.68 | 36.19 | 37.92 | 20.84 | 17.97 | 19.46 | 28.69 |
| 16. Tripura | 9.09 | 17.91 | 12.92 | 5.21 | 7.54 | 6.35 | 9.19 |
| 17. Uttar Pradesh | 18.34 | 16.29 | 17.30 | 18.80 | 18.30 | 18.55 | 17.93 |
| 18. West Bengal | 23.45 | 19.26 | 21.35 | 16.20 | 13.84 | 15.00 | 18.21 |
| All-India | 24.21 | 22.84 | 23.54 | 17.33 | 15.75 | 16.56 | 19.97 |

**Table 15: Sub-samplewise Estimates of Average Amount of Debt in 1981**

*(in Rs. per reporting household)*

| Name of the State | Central sample | | | State sample | | | Central & State sample pooled |
|---|---|---|---|---|---|---|---|
| | SS-1 | SS-2 | combined | SS-1 | SS-2 | combined | |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| 1. Andhra Pradesh | 3,817 | 3,476 | 3,653 | 3,171 | 3,473 | 3,324 | 3,547 |
| 2. Assam | 1,198 | 1,436 | 1,362 | 939 | 524 | 808 | 1,070 |
| 3. Bihar | 1,444 | 1,732 | 1,596 | 1,475 | 1,572 | 1,528 | 1,575 |
| 4. Gujarat | 5,708 | 5,515 | 5,621 | 3,779 | 3,766 | 3,772 | 4,719 |
| 5. Haryana | 6,123 | 7,990 | 6,959 | 11,982 | 17,080 | 14,069 | 9,122 |
| 6. Himachal Pradesh | 2,335 | 3,126 | 2,826 | 3,462 | 2,422 | 3,043 | 2,867 |
| 7. Jammu & Kashmir | 2,043 | 4,842 | 3,648 | 1,596 | 2,319 | 1,947 | 2,813 |
| 8. Karnataka | 2,811 | 9,580 | 6,063 | 3,995 | 2,846 | 3,437 | 5,209 |
| 9. Kerala | 3,699 | 2,550 | 3,087 | 4,090 | 3,050 | 3,621 | 3,351 |
| 10. Madhya Pradesh | 2,531 | 2,688 | 2,660 | 3,186 | 2,927 | 3,060 | 2,836 |
| 11. Maharashtra | 3,157 | 3,756 | 3,433 | 3,903 | 4,589 | 4,213 | 3,806 |
| 12. Manipur | 3,004 | 1,007 | 2,740 | 897 | — | 897 | 1,412 |
| 13. Meghalaya | 561 | 4,621 | 1,955 | 690 | 2,428 | 1,797 | 1,883 |
| 14. Orissa | 1,611 | 1,610 | 1,611 | 1,925 | 1,944 | 1,934 | 1,723 |
| 15. Punjab | 5,953 | 7,841 | 6,715 | 11,289 | 9,477 | 10,314 | 7,911 |
| 16. Rajasthan | 5,338 | 4,041 | 4,643 | 5,306 | 4,420 | 4,966 | 4,753 |
| 17. Tamil Nadu | 3,130 | 3,407 | 3,264 | 4,249 | 3,801 | 4,845 | 3,528 |
| 18. Tripura | 2,895 | 1,583 | 2,068 | 891 | 1,886 | 1,473 | 1,877 |
| 19. Uttar Pradesh | 2,671 | 2,671 | 2,670 | 2,468 | 2,426 | 2,447 | 2,554 |
| 20. West Bengal | 1,816 | 1,899 | 1,853 | 1,358 | 1,418 | 1,386 | 1,663 |
| All-India | 3,139 | 3,624 | 3,374 | 3,356 | 3,156 | 3,257 | 3,311 |

The accuracy or the margin of error in the survey estimates can be judged from sampling errors or by comparing the survey estimate with the known control totals. The survey publications do not give explicit estimates of sampling errors as such. However, we can compute the same in a crude manner. The survey was conducted in two main samples, Central sample and the State sample and within each sample there were two sub-samples. Samplewise estimates are also published on a few items and the difference between the two sample estimates provide rough estimates of the sampling error.

Table No. 14 shows that, at all India level, the difference between the Central and State sample estimates of the percentage of indebted households is very high and the State sample estimates are generally on the lower side. In fact, in as many as eight states, the State estimates are about half of the Central sample estimates. But within the two sub-samples of each of the Central and State samples, the estimates appear to be somewhat closer. At the all-India level also, the Central sample gave an estimate of 23.5 per cent of indebted households whereas the State sample estimate was 16.5 per cent. This would indicate that there could be a major difference between the survey investigations carried out by the Central agency and the State agency.

In the case of the estimates of average liability per reporting household, the Central and State samples have given fairly close estimates at the all-India level and in six states, but, in this case, the State sample estimates are higher than the Central sample estimates in as many as nine states. If we consider the sub-samplewise estimates within the respective two samples, we notice larger differences, unlike as in the estimates of percentage of indebted households, where smaller differences were recorded in the sub-samplewise estimates.

The control totals for the Survey estimates are provided by the population count or the number of households in the country, both of which are available from the Population Census, 1981. The total number of rural households, as estimated from the 1981 Survey and the Population Census, 1981 shows an over-estimation in the Survey to the extent of 3.03 per cent (without adjustment for the small time difference) at the all-India level. But there were wide statewise differences, particularly, in bigger states. For instance, in Bihar, West

Bengal & Punjab, there was an over-estimation of 10 per cent and in Madhya Pradesh, Gujarat & Orissa, there was an under-estimation of 7 to 12 per cent.

Finally, some aspects of non-sampling errors of such large scale sample surveys also deserve special attention. As per the definition of "debt" adopted in the Survey, all current liabilities of households which are in the nature of bills payable for purchase of goods and services on credit are included as under "outstanding debt". Current liabilities of households are in the nature of informal credit and it is quite likely that some of the respondents have not included them under outstanding credit. Secondly, under-statement of actual debt by respondents is likely on account of two reasons. It could be that, with a consideration of obtaining debt waiver, households have repudiated their liabilities. Also, availment of credit and its value could have been under-stated to emphasize the inadequacy of loan facilities and secure more credit. In any case, in an atmosphere surcharged with politico-economic motives, it is difficult to assess how the respondents have reacted to the questions on debt put to them by the survey investigators.

The Central and State samples are canvassed by different agencies but the set of instructions, procedures of conducting the field work, etc. are all laid down by the NSSO. Extensive training camps for the field agency are held prior to the conduct of the Survey and there is a close supervision over the field work. On this background, the reasons as to how the state estimates of incidence of indebtedness are generally lower with wide margins while the estimates of average debt outstanding are closer in the two sets would need further probing by the NSSO.

On the above background, a few suggestions for the next AIDIS due in 1991-92 may be offered. First, the sampling design needs to be oriented towards improving efficiency of the estimates of debt and investment by households. Inquiry into these financial characteristics of households should not be clubbed with other socio-economic inquiries for the purposes of sample selection. Villages may be grouped into two strata, viz., villages having institutional credit facilities and villages not having credit facilities. Secondly, households in selected villages should further be stratified as those availing of loans from

institutional agencies and others. As regards the items of inquiry, along with the household schedules, information may also be collected from the agencies supplying credit in selected villages. Reference dates for collecting information on assets and liabilities may be chosen after studying the time-schedules of funds flow and their recycling in rural areas. Informal loans in the nature of trade credit and bills payable (current liabilities) may be treated separately from formal loans. The tabulation programme should bring out the multi-agency and multi-purpose availment of loans by rural households. The NSSO should also investigate into the reasons for a somewhat systematic under-estimation in the State samples as compared with the Central samples.

**Notes and References**

1. The article is based mainly on the RBI publications on the AIDIS. For the 1981 Survey, there are also two NSSO articles published in Sarvekshna. These publications are listed at the end. The author gratefully acknowledges the assistance given by K.B. Puthran under the guidance of A.S. Mhatre in the compilation of Tables in the article.
2. See, for instance, *P.K. Tandon, "A Profile on Rural Indebtedness"* in the *Social Scientist*, April 1988. In his note, Tandon has argued that the green revolution together with increased institutional credit has intensified the process of differentialisation in rural society. While noting the fact that the proportion of institutional credit given by all agencies to rural households has increased from 17.3 per cent in 1961 to 61.2 per cent in 1981, Tandon doubts the reliability of information on outstanding debt from non-institutional sources. According to him, the survey data grossly underestimates the actual indebtedness of the poorer sections, relying mainly on non-institutional sources for their supply of credit and totally fails to take a note of the implied forms of bondage when it talks of debts supplied without insisting on security or payment of interest.
   See also K. Seeta Prabhu & Others in "Rural Credit: Mystery of Missing Households", *Economic & Political Weekly*, december 10, 1988, wherein it has been pointed out that not only the percentage of rural indebted households in total rural households has declined between 1971-72 and 1981-82, but also the absolute number of households has declined from 32.98 million in 1971 to 18.70 million in 1981. This decline in the incidence of indebtedness is observed for all the categories of rural hosueholds, in better-off as well as poor states.
3. The difference between the estimates given as 'average per household' and 'average per reporting household' needs to be understood clearly. In the former estimate, the aggregate value is divided by the number of total households while in the latter case, the divisor is the number of households reporting the value of an item being estimated. Both the estimates can be used within their own contexts.

4. The definition of rural centre, according to the Population Census, is broadly, a village with a population of less than 5,000 persons. In the Banking Statistics, rural centre connotes a village with population of less than 10,000 persons. The definition adopted in the AIDIS conforms to Population Census, while in this article, banking statistics relate to the definition of rural as defined in banking statistics. Bank branches include branches of regional rural banks.
5. Bhatty I.Z. and Prem Vashishtha, "Seminar on Saving Estimates in India—Rural Household Saving and Investment Behaviour", National Institute of Public Finance & Policy, New Delhi (unpublished).
6. Iqbal Farrukh, Journal of Development Studies, April, 1988.

**Sources**

**1. Table No.1**

(a) Outstanding loans, borrowings and repayments of Rural Households, AIDIS, 1961-62, RBI Bulletin - September, 1965.
(b) All India Debt & Investment Survey, 1971-72. Indebtedness of Rural Households as on 30th June 1971 and Availability of Institutional Finance - Reserve Bank of India, 1977.
(c) All India Debt & Investment Survey, 1971-72. Statistical tables relating to Cash dues outstanding against Rural Households as on 30th June 1971 - Reserve Bank of India, 1976.
(d) All India Debt & Investment Survey, 1981-82. Assets & Liabilities of Households as on 30th June 1981 - Reserve Bank of India, 1987.

**2. Table No.2**

(a) All India Debt & Investment Survey, 1971-72. Statistical tables relating to Cash dues outstanding against Rural Households as on 30th June 1971 - Reserve Bank of India, 1976.
(b) Sarvekshana - Vol.XI No.1, July 1987.
(c) RBI Bulletin, September 1965.

**3. Table No.3**

(a) RBI Bulletin, September, 1965.
(b) All India Debt & Investment Survey, 1971-72. Statistical tables relating to Cash dues outstanding against Rural Households as on 30th June 1971 - Reserve Bank of India, 1976.
(c) All India Debt & Investment Survey, 1981-82. Assets & Liabilities of Households as on 30th June 1981 - Reserve Bank of India, 1987.

**4. Table No.4**

(a) All India Debt & Investment Survey, 1981-82. Assets & Liabilities of Households as on 30th June 1981 - Reserve Bank of India, 1987.
(b) All India Debt & Investment Survey, 1971-72. Indebtedness of Rural Households as on 30th June 1971 and Availability of Institutional Finance - Reserve Bank of India, 1977.

**5. Table No.5**

(a) All India Debt & Investment Survey, 1971-72. Indebtedness of Rural Households as on 30th June 1971 and Availability of Institutional Finance - Reserve Bank of India, 1977.

(b) All India Debt & Investment Survey, 1981-82. Assets & Liabilities of Households as on 30th June 1981 - Reserve Bank of India, 1987.

**6. Table No.6**

(a) All India Debt & Investment Survey, 1981-82. Assets & Liabilities of Households as on 30th June 1981 - Reserve Bank of India, 1987.

(b) Sarvekshana, Vol.X, No.1, July 1986.

(c) All India Debt & Investment Survey, 1971-72. Assets of Rural Households as on 30th June 1971, Reserve Bank of India, 1976.

(d) All India Debt & Investment Survey, 1971-72. Statistical tables relating to Cash dues outstanding against Rural Households as on 30th June 1971 - Reserve Bank of India, 1976.

**7. Table No.7**

(a) Statistical tables relating to Banks in India, 1973 - Reserve Bank of India, 1975.

(b) All India Debt & Investment Survey, 1971-72. Statistical tables relating to Cash dues outstanding against Rural Households as on 30th June 1971 - Reserve Bank of India, 1976.

(c) Banking Statistics - Basic Statistical Return Vol.1, December 1972 - Reserve Bank of India, 1973.

(d) Banking Statistics - Basic Statistical Return Vol.11 - Reserve Bank of India, 1987.

(e) Statistical statements relating to Co-operative Movement in India; 1971-72 - Part I - Credit Societies - Reserve Bank of India, 1973.

(f) Statistical statements relating to Co-operative Movement in India, 1981-82 - Part I - Credit Societies - NABARD 1986.

**8. Table No.8**

(a) All India Debt & Investment Survey, 1971-72. Statistical Tables relating to Cash dues outstanding against Rural Households as on 30th June 1971 - Reserve Bank of India, 1976.

(b) All India Debt & Investment Survey, 1971-72. Statistical Tables realting to Cash Borrowings and Repayments of Rural Households during July 1971 to June 1972 and Cash dues outstanding as on 30th June 1972 - Reserve Bank of India, 1978.

**9. Table No.9**

(a) All India Debt & Investment Survey, 1981-82 Assets & Liabilities as on 30th June 1981 - Reserve Bank of India, 1987.

(b) All India Debt & Investment Survey, 1981-82 Statistical Tables on Capital Expenditure & Capital Formation of Households during the year ended 30th June 1982 - Reserve Bank of India, 1988.

**10. Table No.10**

(a) All India Debt & Investment Survey, 1971-72. Statistical Tables relating to Cash dues outstanding against Rural Households as on 30th June 1971 - Reserve Bank of India, 1976.

(b) All India Debt & Investment Survey, 1971-72. Assets of Rural Households as on June 30, 1971 - Reserve Bank of India, 1976.

(c) All India Debt & Investment Survey, 1971-72. Assets & Liabilities of Households as on 30th June 1971 - Reserve Bank of India, 1975.

**11. Table No.11**

All India Debt & Investment Survey, 1981-82.Assets & Liabilities of Households as on 30th June 1981, Reserve Bank of India, 1987.

**12. Table No. 12 & 13**

Agricultural Credit Institution: India - Their Structure & Role in Development by Dr. P.D. Ojha, Reserve Bank of India Bulletin, February 1986.

**13. Table No. 14 & 15**

Sarvekshana - Vol.X, No.1, July 1986.

*Reserve Bank of India Occasional Papers*
Vol. 9, No. 4, December 1988

# Financial Innovation: Determinants and Implications for Monetary Policy— A Model

N. Nagarajan*

## 1. Introduction

IN recent years, the Indian financial system has been experiencing a process of rapid structural transformation. The trend is evident particularly in the banking sector. The commercial banks have diversified their activities and, through their subsidiaries, have began to offer numerous financial services to the public including the chance to participate in attractive mutual funds. Pertinently, they have started issuing money-substitutes such as credit cards. These activities, though at a low level now, are expected to gather momentum quickly. On the other hand, the capital market has emerged as a vibrant force. Also, indications are that the insulation enjoyed by it so far from the developments abroad will dissolve sooner since Indian corporate sector is expected to enter global markets to raise funds. This integration will impart greater dynamism by facilitating the absorp-

* Dr. N. Nagarajan is an Assistant Adviser in the Department of Economic Analysis and Policy. The study formed a part of the author's work at the Indiana University, Bloomington (USA) to which he was sent by the Reserve Bank of India under its scheme of deputing officers abroad for higher studies.

In the preparation of this study, the author has benefited from valuable comments of many faculty members, particularly Prof. W.E. Witte and Prof. C.J. Waller of the Department of Economics, and Prof. von Hogen of Business School, Indiana University as also officials of the Reserve Bank of India, Bombay, especially, Dr. S.L. Shetty, Shri R. Nagaraja Rao,Shri Y.S.R. Sarma and Dr. N.D. Jadhav. The author is grateful to the Reserve Bank of India for financial and other support. A preliminary draft of this paper was presented at the Seminar in Money at the Indiana University in September 1988.

tion of the developments abroad. These likely trends imply that, soon, financial innovation, in its varied form, will be a reality in India leading to the availability of a wide choice of well-structured financial instruments to the investors.

In this paper, an exploratory theoretical attempt is made to assess the determinants of a specific form of financial innovation, viz., interest-bearing-money, substituting non-interest-bearing-money. The model is then articulated to evaluate the likely effects on selected crucial macro-economic variables, viz., rates of interest, interest-bearing-money and non-interest-bearing- money, as also the implications of such innovation for the conduct of monetary policy through the regulation of supply of reserves. The exercise is done within the general framework of portfolio balance approach which is sufficiently flexible to permit a consideration of different types of money. The method adopted is comparative static analysis with equilibrium conditions imposed on all markets.

The desirability of the study stems from two quarters. From a purely academic angle, available literature have accepted financial innovation as a 'shift' or abberation in the conventional demand for money function and a few, which have had focussed attention on innovation, have done so at micro-level. This study, on the other hand, specifies the determinants of financial innovation and then traces its effects. Thus, adequate attention is paid to innovation and its integration with the financial system. Perhaps the importance lies more in its relevance to policy formulation which is the second rationale of this study. Developed countries, especially the USA and Canada, responded to the phenomenon through empirical analysis after its effects were felt; even then, as would be pointed out shortly, they have concentrated on obtaining a stable demand for money function and rarely on the effects of innovation which is more important. In contrast, this study provides a theoretical futuristic perspective which could be used for timely action. This is particularly relevant for India in view of the growing trend towards the rationalisation of administered interest rate structure including gradual withdrawal of rigidities and the large-scale introduction of mechanisation at the banks.

The results indicate that, under given conditions, financial in-

novation by way of introduction of money-substitute instruments, has a self-sustaining tendency. The trend is accentuated by interest rate spread for the banks, technology and restrictions imposed by the monetary authority including monetary targeting. The consequences of financial innovation are felt on the interest rates as also on interest-bearing and non-interest-bearing monies. While interest rates tend to rise, both types of monies, especially interest-bearing, also expand. Thus, during a period of financial innovation, neither targeting narrow money by way of supply of reserve nor any attempt to regulate interest rates is unlikely to yield expected results.

Following this introductory Section, available major studies are reviewed in Section 2. The basic model is developed in Section 3 and articulated in Section 4. Main findings are summarised, limitations noted and areas for further study identified in the last Section. A note on financial innovation including experiences of the USA and Canada and details of the Model are provided as Appendix I and II, respectively.

## 2. Review of Literature

Recent past witnessed a proliferation of literature on various aspects of financial innovation. The discussions ranged from descriptive international comparisons [Aktar (1983), BIS (1984), de Cecco (1987) and Suzuki and Yomo (1986)] to theoretical micro-level appraisals (Silber, 1975). Besides, several empirical studies on the effects of financial innovation were attempted. In fact, according to Offenbacker (1982), financial innovation is an empirical issue.

Basically, the empirical studies endevour to obtain a stable demand for money function. Available literature on this subject is so vast as to invite two surveys. [Jonson and Rankin (1986) and Judd and Scadding (1982)]. The method adopted is either to redefine the dependent variable so as to include a particular money-substitute-financial-asset or re-specify the explanatory variables with the inclusion of dummies to take into account financial innovation. The findings are not unambiguous. Generally, it is accepted that the USA and, to a less extent, Canada had experienced shifts in the demand for money functions. Simpson and Porter (1980) found that, in the USA, the functional relationship between narrow money and real

variables was deteriorating. According to Boughton (1981), the money-demand functions in respect of $M_1$ in the USA and Canada were unstable. Kabir and Mangla (1988) reported that, in Canada, innovation had displaced the money-demand function from its trend in the early 1980s. Similar conclusion was reached by Milbourne and Moore (1986). According to Miller (1986), financial innovation had affected the function in the USA but the effects were gradual. Atkinson *et al* (1984) maintained that financial innovation had not significantly destabilised the demand for money, especially in the Continent. Chirinko and Colloton (1984) found some instability in the function in the USA but could not attribute the instability to financial innovation. An inescapable conclusion from these studies, especially in view of sophisticated statistical tools employed by Baba *et al* (1988) and Broemeling and Tsurumi (1987), is that with sufficient treatment, any data could be made to confess to any hypothesis.

As a logical sequence to the search for a stable demand for money function, the implications of shifts in such function for the conduct of monetary policy was examined by Cagan and Schwartch (1975), Hall (1986) and Hester (1981).

At micro level, various hypothesis were advanced to explain financial innovation. Kane (1984) attributed it to "regulatory dialectics" in which the political process of regulation and the economic process of regulatee-avoidance interact by continuously adapting to each other with the latter always remaining one step ahead of the former. Desai and Low (1987) and Dufey and Giddy (1981) put forward the "opportunity-gap" theory. This suggested that the financial firms would float new securities to cover the gap, if any, in the maturity pattern of existing securities. Silber (1975) propounded the "constraint-induced-innovation" hypothesis. According to this, a profit - maximising financial firm would innovate, having regard to the cost of such innovation, when exogenous changes alter its constraints.

On the relationship between innovation and macro-economic activity, Podolski (1986) had presented an excellent perspective on the existing appraisals. Wenninger and Witte made significant contributions to the theoretical relationships. Wenninger (1984), *a priori*, hypothesised that innovation would result in (a) a decline in the interest-elasticity of demand for money since the new instrument

would bear market-related interest along with the conventional conveniences of chequing accounts, (b) an increase in the income-elasticity due to the same factors, and (c) an increase in the interest-elasticity of expenditure because loans would be converted/made on floating interest rates. He then tested the effects of innovation within the conventional IS-LM framework. Innovation was equated with shifts in both demand for and supply of money. The static solutions to the system provided complex and ambiguous sets of results. However, he could derive some broad conclusions. First, a shift in the demand for money led to a larger deviation of narrow money from the target, requiring larger compensatory action from the central banking authority if the target was to be achieved. Second, narrow money was found less sensitive to changes in the supply of reserves but income and interest rates became more sensitive. Finally, the effects of innovation on the real sector too, was found ambiguous implying increased uncertainty about the underlying economic relationships. Also, all these relationships were changing simultaneously. For policy, this had two implications : (i) uni-target was inadvisable and (ii) a discretionary rather than a fixed rule monetary policy would be preferable.

On the lines of portfolio approach of Tobin, Witte (1982) developed a model of the financial sector consisting of four assets that included two highly liquid assets, both usable as transaction balances but one of which paid market-determined rate of interest. The model was then used to examine the effects of shifts in demand and supply, separately, in the interest-bearing-transaction-asset on interest rate and monetary aggregates. The consequent impact on the effectiveness of the conventional monetary policy instruments was analysed and found to be rather ambiguous, especially that of targeting money supply.

Both these studies have treated financial innovation as shocks. This dilution is significant since, in the absence of an adequate specification of the phenomenon, its effects cannot reasonably be assessed. Wenninger's attempt suffers from two more drawbacks, First, the IS-LM framework assumes that the LM curve is relatively stable, which, in turn, assumes, *inter alia,* that its composition is unchanged. Since innovation does change this composition, the use of the IS-LM framework would appear inappropriate. Further, 'since

modern monetary analysis must take place with reference to the whole financial sector, the conventional LM analysis is inadequate' (Podolski, 1986. p. 227). Second, innovation and deregulation were treated synonymously, which is incorrect (he, however, isolated and omitted the effects of deregulation through *a priori* reasoning).

The foregoing brief review of available major studies reveals that the relationship between financial innovation and the financial system with adequate attention to the former has not been sufficiently explored. The present study is an exploratory theoretical attempt to assess the determinants of financial innovation, to analyse the interrelationship between financial innovation and financial sector and to evaluate the effects of selected exogenous variables and policy parameters on critical macro-economic variables.

## 3. Basic Model

The model begins with the innovation sector. Financial innovation (I), for the present purpose, is interpreted to mean the availability of an additional financial instrument, whether entirely new or modified version of the existing one, carrying a market-related-interest rate. Conceptually, the participants in the market could be distinguished as the banker who demand innovation and the innovation agency which supplies the innovation (in practice, they correspond to the retail banker and the merchant banker, respectively). Banks demand innovation because it would facilitate them to augment their resources. The demand will increase with the level of interest-bearing-money (IBM). The demand would also depend upon interest rate differentials. The specification of spread instead of absolute levels reflects the present trend of banks to work on margins, particularly after the emergence of liability management techniques. In the present instance, this would imply that the net earnings accruing to banks, i.e., their earnings from the deployment of money obtained *minus* interest payable on the interest-bearing-money, would induce banks to demand innovation. Assuming that the deployment of money is only on the purchase of bonds, this would mean that the spread between interest rate on bonds ($r^B$) and interest rate on interest-bearing-money ($r^{IBM}$) would be a determinant in the demand for innovation. Bonds refer to all financial assets except interest-bearing-money (IBM) and non-interest-bearing-money (NIBM).

Savings account with chequing facility is an example of IBM. Demand deposits are an example of NIBM. The demand function would be positively related to restrictions imposed by the monetary authority (RE) since an important objective of innovation has been to circumvent such restrictions. Advancement in technology (T) would enable the rapid diffusion of innovation and hence a positive relationship is prescribed. 'Technological advances deployed in modern finance have been a common denominator of all major financial innovations. They have had a strong impact on the macro-economic demand for narrowly depicted money, largely by reducing transactions costs' (Podolski, 1986, p. 173). These specifications would indicate a positive sloping demand function. This apparently unusual phenomenon is due to the fact that the banks are, in this context, only intermediaries and this function, in reality, refers to the supply of innovation to the public who are the ultimate users. Symbolically,

$$I_b^d = K\left[\underset{+}{IBM},\ \underset{+}{RE},\ \underset{+}{T},\ \underset{+}{(r^B - r^{IBM})}\right] \qquad (1)$$

As mentioned earlier, I refers to financial innovation. Superscripts on the left hand side variable refer to the type of function - d for demand, for instance, and subscripts to the sector - b for banks, for example. On the right hand side, IBM stands for interest-bearing-money, r for the rate of interest and the superscripts on it for the financial asset to which the respective r relates - $r^B$ for rate of interest on bonds, for example. RE refers to various restrictions, except reserve ratios, imposed by the monetary authority on the banks with regard to the creation of deposits. T stands for technology, i.e., advancement in telecommunication, data processing, data storing and data retrieval. The sign above a variable denotes the hypothesised partial for that variable.

Supply of innovation is assumed to adjust passively to demand. As explained in detail in Appendix I, most of the innovative instruments are in effect slightly modified versions of the older ones; only the rapid diffusion, facilitated by technology, has been responsible for the present concern. Further, the requirements of the public (as reflected through the banks) vary necessitating appropriate instrument to suit specific needs. On the other hand, given the resources including technology, innovating new instruments would be easier. Thus, the supply function is

$$I^{s}_{INN} = I^{d}_{b} \qquad (2)$$

Innovation sector will be in equilibrium when supply matches demand. Thus

$$I^{s}_{INN} = I^{d}_{b} = I \qquad (3)$$

The formulation of the model relating to the financial system is based on the conventional approach. There are four financial assets. The non-bank (public) holds two assets - non- interest bearing money (NIBM) and interest-bearing money (IBM). Both are bank deposits and hence liabilities of the banks. NIBM is similar to demand deposits and carries no explicit return, but IBM is similar to savings deposits and carrys a positive market-determined-return. *The distinguishing feature of NIBM and IBM is thus the latter earns on explicit interest while the former does not.* Hence, the introduction of any new instrument bearing a rate of interest and usable as money will be included in IBM. Both NIBM and IBM are highly liquid. The third asset is bonds (B) which is a financial asset used to finance Government deficits and private investment. Thus, bonds are substitutes for IBM. For the present purpose, bonds on private investment account is taken as zero. Thus, in this model, B stands for bonds of all bypes issued by the Government. The fourth asset is reserve (R) which is a liability of the central bank and an asset of the banks. Their presence in the banks' portfolio reflects the statutory reserve requirement established by the central bank.

The financial system has three sectors : the non-bank (public), banks and the central bank. In this model, the term 'banks' include all financial institutions. In addition, the innovation sector interacts with the financial sectors. There are five markets - one for innovation and one each for four financial assets.

The public holds non-interest-bearing money, interest- bearing-money and bonds in their portfolios. The respective demand functions are:

$$NIBM^{d}_{p} = N\,(r^{IBM}, r^{B}, \bar{I}, \dot{W}) \qquad (4)$$

$$IBM^{d}_{p} = M\,(\dot{r}^{IBM}, \dot{r}^{I\,NF}, \dot{I}, \dot{W}) \qquad (5)$$

$$B_p^d = B\,(r^{\bar{I}BM}, r^{\dot{+}B}, \bar{I}, \dot{W}) \qquad (6)$$

subject to the wealth constraint

$$W = NIBM_p^d + IBM_p^d + B_p^d \qquad (7)$$

Four observations on the specification are in order. First, rate of inflation ($r^{INF}$) is taken as an explanatory variable for $IBM_p^d$ and the relationship hypothesised is positive. This is because of the observed (in the USA) public's preference to have IBM as a hedge against inflation. Second, in view of imperfect substitution, the effect of $r^B$ in the IBM market is likely to be weak. Third, the specification is based on the levels of interest rates because IBM is a substitute mainly for NIBM which bears zero rate of interest. Finally, the effect of I on $B_p^d$ could be positive in the unlikely event of IBM turning unattractive to the public.

Banks issue two liabilities, NIBM and IBM and hold two assets, reserves (R) and bonds (B). The supply of NIBM by the banks adjusts passively to the demand of the public. This follows from the earlier stipulation, viz., NIBM is non-interest bearing and hence profitable to banks but not so to the public. Banks, therefore, issue all NIBM that the public is willing to hold at the prevailing interest rates structure and would be prepared to issue more. Then,

$$NIBM_b^s = NIBM_p^d = NIBM \qquad (8)$$

In respect of the other liability of the banks, IBM, one could assume that (a) the banks are profit maximisers and (b) their choice would depend upon the yield differential between the asset of banks, viz., bonds, and interest payable on IBM. Then

$$IBM_b^s = G\,[\,(r^B \dot{-} r^{IBM}), \dot{I}\,] \qquad (9)$$

Obviously, the spread would have a positive relationship. Innovation would also have a positive relation because the banks would obtain more monies with innovation; also, the new instrument, among other things, would facilitate the banks to economise on cash.

The other asset of the banks is reserves. This includes statutory requirements and excess, if any. One could assume that they would maintain the minimum, both on NIBM and IBM. On NIBM, even if they refer exclusively to currency, it is reasonable to assume that the banks would maintain a specified fixed proportion in cash - either to meet the legal reserve requirement and/or to meet their daily cash drain. Both on NIBM and IBM, the banks are statutorily required to maintain a fixed proportion as reserves. If the statutory ratios on NIBM and IBM are denoted as $\alpha$ and $\beta$, respectively, the demand for reserve function is

$$R_b^d = \alpha NIBM + \beta IBM \tag{10}$$

Similarly, the banks' demand for bonds will depend positively upon the rate on bonds. Thus,

$$B_b^d = C(r^{+B}) \tag{11}$$

The balance sheet constraint of the banks is

$$R_b^d + B_b^d = NIBM_b^s + IBM_b^s \tag{12}$$

The central bank is assumed to control the level of reserves by means of open market operations. Then,

$$R_c^s = R_b^d = B_c \tag{13}$$

In this framework, the equilibrium conditions are

$$\text{Innovation} = K(.) = I_{INN}^s = I \tag{15}$$

$$\text{Non-interest-bearing-money} = N(.) = NIBM_b^s = NIBM \tag{16}$$

$$\text{Interest-bearing-money} = M(.) = G(.) \tag{17}$$

$$\text{Bonds} = [B(.) + C(.)] = B\text{-}B_c \tag{18}$$

$$\text{Reserves} = \alpha NIBM + \beta IBM = B_c \tag{19}$$

$$\text{where } B_p^d + B_b^d + B_c = B \tag{14}$$

For simplicity, let the interest rate differential ($r^B - r^{IBM}$) be $a$.

## 4. Articulation of the Model

The basic model is articulated in this Section with the twin objective of assessing the relative roles of selected exogenous variables

in determining innovation and evaluating the impact of changes in selected exogenous variables and policy variables on the crucial endogenised macro-economic variables. For simplicity, linear relationship is assumed. Technical considerations enable the omission of an equation but keeping in view the number of variables involved, all equations are retained. Of the variables, innovation, interest rates on money and bonds, the size of non-interest-bearing and interest-bearing monies are taken as endogenous. The system is evaluated to obtain comparative static solutions with equilibrium conditions imposed on all the markets.

*Determinant of the matrix*: The determinant of the coefficient matrix is:

$$\Delta = -m_1 \left\{ \beta(b_2 + c_1) + \alpha k_1 [n_3(b_2 + c_1) + n_2 b_3] \right\} \quad (25)$$

The overall effect of the determinant is unambiguously negative.[1]

*Determinants of Innovation*: The numerator for calculating a change in innovation is:

$$\dot{I} = -(k_2 R\dot{E} + k_3 \dot{T} + k_4 \dot{a}) [\beta m_1 (b_2 + c_1)] + k_1 \left( (-m_2 i^{INF} + g_1 \dot{a}) \right.$$
$$\left. \{\alpha [n_1(b_2 + c_1) - n_2 b_1]\} + (\dot{B} - \dot{B}_c) [\alpha m_1 n_2] - \dot{B}_c [m_1(b_2 + c_1)] \right) \quad (26)$$

The demand response for IBM in innovation market ($k_i$) plays an extensive role in the determination of changes in innovation. In the process, it affects even the exogenous variables. It accentuates interest rate spread and holdings of bonds by the public and banks: meanwhile, it undermines the rate of inflation and supply of reserves. In view of the complexities involved in respect of others, it is preferable to assess the relative roles of exogenous and policy variables separately.

The contribution of *interest rates spread* is

$$-k_4 \beta m_1 (b_2 + c_1) + g_1 \left\{ \alpha k_1 [n_1 (b_2 + c_1) - n_2 b_1] \right\}$$

The first term is unambiguously negative. The positive demand response from the public and banks for bonds, together with the pro-

1 Empirical studies relating to the USA also lend support to this effect.

duct of reserve requirement on IBM and positive demand response of the public for IBM, reacts with the positive demand for innovation from the banks due to interest rate differential; and the term is preceded by a negative sign. The second term also would be negative in the most likely event[2] of $n_2 b_1 < n_1(b_2 + c_1)$ in absolute terms. In this instance the product of the negative demand response for NIBM in NIBM market due to a change in the rate of interest on bonds ($n_2$) and the negative demand response for bonds in bond market due to a change in the rate of interest or IBM ($b_1$) would be less than the product of the negative demand response for NIBM in NIBM market due to a change in the rate of interest on IBM ($n_1$) and the positive demand responses of banks and the public in the bond market for bonds due to a change in bond rate ($b_2 + c_1$). The net sign of the term is negative and hence the overall effect on innovation is positive.

The share of *technology* is relatively simple:

$$-k_3 \beta m_1 (b_2 + c_1)$$

The term is certainly negative since all the elements are positive but preceded by a negative sign. Since the determinant of the matrix is negative, the result is positive i.e., technological advances lead to greater innovation, through the positive demand response of banks for innovation in innovation market due to a change in technology, reserve requirement on IBM, positive demand response of the public for IBM due to a change in the interest rate on IBM and the combined positive response of the public and banks for bonds due to a change in the bond rate. The total effect on innovation is promotional.

The role of *inflation* is provided by the expression

$$m_2 \alpha \left[n_1 (b_2 + c_1) - n_2 b_1\right]$$

This could be negative only in the unlikely event of $n_2 b_1 > n_1 (b_2 + c_1)$ in absolute terms and hence it is reasonable to take it as positive. In this content, the earlier condition that $n_2 b_1 < 0$

2 This is observed partially in the empirical studies relating to the USA.

may also be recalled. This would imply that changes in the rate of inflation tend to depress innovation through public responses in the bond, NIBM and IBM markets, along with the reserve requirement on NIBM. This result is in contrast with the general belief.

Among policy variable, the impact of *supply of reserves* or *monetary targeting* on innovation is $m_1 (b_2 + c_1)$. All the multipliers are positive. However, the policy variable $B_c$ has a negative sign. Thus, the total effect is negative implying a promotional effect on innovation. The extent of the effect is provided by the multipliers viz., the demand response of the public and banks for bonds due to a change in bond rate and the demand response of the public for IBM in the IBM market due to a change in the rate of interest on IBM. Also, restrictions (other than reserve requirements) imposed by the monetary authorities have the following influence:

$$- \beta m_1 (b_2 + c_1)$$

The expression is negative since all the elements are positive and the product is preceded by a negative sign. Reserve requirement on IBM, the public's positive demand response for IBM in IBM market due to a change in the rate of interest on IBM as also the public's and banks' positive demand response for bonds lead to the effect. The impact on innovation is positive.

In short, *change in interest rate spread, an improvement in technology, monetary restraints other than reserve requirements and supply of reserves tend to promote further innovation. Change in the rate of inflation appears to moderate the tendency.*

The effects of selected exogenous and policy variables on the endogenised variables are furnished in the following table.

*Changes in interest rate spread*: The effect on $r^{IBM}$ is ambiguous. The reserve requirement on IBM is certainly positive resulting in a depressing effect. On the other hand, the reserve requirement on NIBM, the negative demand response for NIBM and bonds in NIBM and bond markets, respectively, due to a change in innovation and the negative demand response for NIBM effect of changes in the rates in NIBM market due to a change in the rate of interest on bonds

**Table: Effects of Exogenous and Policy Variables on Endogenous Variables**

| Exogenous/ Policy Variable | Effect on | | | |
|---|---|---|---|---|
| | IBM $\dot{r}$ | B $\dot{r}$ | NIBM | IBM |
| $\dot{a}$ | $g_1\{\beta(b_2+c_1)-\alpha k_1[n_3(b_2+c_1)+n_2b_3]\}$ | $g_1[\beta b_1-\alpha[n_3b_1k_1-b_3n_1]-k_4\beta b_3$ | $g_1\{\beta[-n_1(b_2+c_1)+n_2b_1]\}+k_4[\beta\{-n_2(b_2+c_1)+b_3n_2\}]$ | $g_1\{-\alpha[n_1(b_2+c_1)+n_2b_1]\}+k_4(m_1[\alpha\{n_3(b_2+c_1)+b_3n_2\})$ |
| $\dot{T}$ | o | $k_3\beta m_1b_3$ | $k_3[\beta\{n_3(b_2+c_1)+n_2b_3\}]$ | $-k_3\{\alpha[n_3(b_2+c_1)+n_2b_3]\}$ |
| $\dot{B}_c$ | o | $k_1b_3$ | $-k_1[-n_3(b_2+c_1)-n_2b_3]$ | $-m_1(b_2+c_1)$ |

will positively influence the rate. The precise effect then would depend upon the relative values of these two terms also the values of $n_3(b_2+c_1)$ and $n_2\ b_3$.

The effect on $r^B$ is an unambiguous negative. Of the multipliers, $b_1$, $b_3$, $n_1$ and $n_3$ are negative; $g_1$, $\beta$, $\alpha$ and $k_4$ are positive. But the interaction leaves the expression negative. The implication is that innovation raises the rates on bonds which is logical since bonds are competitive assets albeit imperfect to IBM.

As for the effects on NIBM, the expression is positive. Multipliers with negative signs such as $n_1$, $n_2$, $b_1$, $b_3$ are present but in totality the expression is an unambiguous positive. The implication is that, the interest rate spread would depress NIBM, which is in line with the past experience. On the other hand, interest rate spread promotes IBM. Some of the multipliers in the expression such as $k_4$ $m_1$, $\alpha$ and $(b_2+c_1)$ are positive but the overall sign is negative. In other words, the interest rate spread increases IBM. This is also in line with the received postulate.

On the whole, the effects of interest rate spread on the endogenous variables are logically consistent. It raises the interest rates on bonds; promotes IBM; depresses NIBM; and the effect on the

interest rate on IBM is uncertain.

*Technology:* The effects on the endogenous variables are unambiguously negative indicating a promotional overall effect. The effect on $r^{IBM}$ is, however, nil. Technology influences $r^{B}$ by a factor of reserve requirement on IBM, the public's demand for IBM in IBM market due to a change in $r^{IBM}$ and the public's negative demand response for bonds in bond market due to innovation. These results conform to the generally accepted propositions.

Technology also increases both NIBM and IBM. The extent of the effect depends upon, in addition to those affecting $r^{IBM}$, the positive demand response of the public and banks for bonds due to a change in $r^{B}$, the negative demand response of the public for NIBM in NIBM market due to changes in bond rate and innovation as also the positive reserve requirement. These conform the common belief.

*Supply of reserves*: The impact of this policy parameter or monetary targeting, is an unambiguous increase in $r^{B}$ and expansion in NIBM and IBM. This policy variable has no effect on $r^{IBM}$. It affects $r^{B}$ by a factor of the positive banks' demand response for innovation due to a change in $r^{IBM}$ and the negative demand response of the public for bonds due to a change in innovation. The impact on NIBM depends, besides these factors, on the public's and the bank's demand response for bonds in bond market due to a change in $r^{B}$ and the public's demand response for bonds and NIBM due to changes in $r^{IBM}$ in bond and NIBM markets, respectively. Pertinently, supply of reserves would increase IBM by a factor of the positive demand response of the public and banks for bonds in bond market due to a change in $r^{B}$ and the public's demand response for IBM in IBM market due to a change in $r^{IBM}$. If the coefficients are lower than unity, this would indicate that this impact would be larger than all others. A reasonable inference then is that during a period of financial innovation, tarteging narrow money would result in a substantial increase in IBM. This result conforms the *a priori* proposition of earlier writers such as Tobin and Wenninger. The contribution lies in that this study explains the manner in which the result is obtained and the extent of the effect. In this sense, the present endeavour provides the missing link in the earlier works.

## 5. Summary and Conclusions

An attempt was made in this exploratory theoretical study to explain financial innovation in terms of selected, commonly accepted contributory factors. The analysis was within the general framework of portfolio balance approach and the method adopted to solve the system was comparative static analysis. The objective was to identify the parameters through which financial innovation is determined as also those through which its effects are transmitted. So far this area had not been sufficiently explored and hence the present study is expected to supplement the existing knowledge on the macroeconomic aspects of financial innovation with special reference to its interrelationship with the financial sector. By presenting a futuristic view, this would also enable the policy makers to take necessary anticipatory actions.

The model, when articulated, yielded the following results:

When financial innovation is sought to be explained in terms of technology, interest rate spread, rate of inflation and restrictions imposed by the monetary authorities such as monetary targeting, it is found that technology and restrictions unambiguously tend to promote innovation and, most probably interest rate differential also intensified innovation. On the other hand, rate of inflation appears to have a depressing impact.

The effects of a change in the interest rate spread on $r^{IBM}$, is relatively uncertain. The effect on $r^{B}$, and IBM are promotional and, on NIBM, depressing. Technology has no effect on $r^{IBM}$ and positive effect on $r^{B}$, NIBM and IBM. Supply of reserves also has the same effects as technology. For policy, this would imply that, during periods of financial innovation, adherence to targeting narrow money nor to any set of interest rates may not achieve desirable results.

A major limitation of the study may be noted. Financial innovation has far reaching implications for the solvency of the financial system. The generally held belief is that the system would become more fragile, necessitating extensive use of the lender of last resort

function by the central bank. This aspect, though important, is not considered so as to retain the area of enquiry within manageable limits.

Being an exploratory attempt, the present work opens up new areas for investigation, of which two theoretical and one empirical possibilities may be indicated. First, financial innovation sector *per se* may be studied in greater detail. A separate model could be built for this sector. This would supplement the present study in that the specification of the first three equations in this study could be done with greater precision. It may be recalled that, in this paper, a passive supply function is assumed for innovation. The equation on innovation sector contains three exogenous variables (though one is policy and another is implicitly endogenised variable) with accompanying technical limitations. The second probable theoretical study would be to segregate the effects of interest- bearing-money from other interest-bearing-financial assets. In Indian context, this could be of greater use partly because more financial innovations are occuring in this area and partly because for the purpose of monetary policy, the Reserve Bank targets the broad money. At empirical level, the system could be estimated. Such attempt would assess the verocity of the model.

## Appendix I: A Note on Financial Innovation

In this note, selected aspects of financial innovation and the experience of the USA and Canada are discussed.

### A. Facets of Financial Innovation

Financial innovation refers to the introduction of new financial assets which serve as money as also to the process in which interest-bearing-financial assets could be converted into the conventional money and vice-versa. Thus, the distinction between product and process innovation, which is thin even in the real sector, is more nebulous in finance. Moreover, financial innovation could take the form not only of new instruments but also of new markets, institutions and practices. Usually, one form of financial innovation is associated with another. Recent history reveals that the appearance of new financial instruments and their diffusion has been in clusters due probably to the inter-dependence between the process and product as also instruments/markets/practices *inter se.*

In finance, appearance of innovation *per se* is not as important as the pace of its diffusion. If the instrument/market/practice remains in vogue for a long period and is slowly absorbed in the system, the effects are gradual and unlikely to cause anxieties. On the other hand, rapid diffusion of innovation i.e., faster adoptive response of the public along with the creative response of innovative spirit, leads to large scale distortions in economic relationships.

Financial innovation has varied characteristics and hence presents considerable taxonomic problems. However, applying certain norms, a meaningful categorisation could be obtained to meet a specific objective. The present endeavour is basically to assess the impact of innovation on the measurement and controllability of money. Accordingly, financial innovation may broadly be divided into two categories: those having a direct effect on monetary aggregates and those having either indirect or negligible effect on such aggregates. The first category includes new financial instruments serving as money substitute, i.e., they are financial assets, usable as means of payments but do not correspond to the theoretical construct of money

conceived as a zero interest rate asset. Most of the new instruments in the USA and Canada fall under this group. They include the Money Market Mutual Funds, Money Market Mutual Accounts, Negotiable Order of Withdrawal, Super-NOW, Automatic Transfer Service, and Certificate of Deposits. The effect of these instruments is felt on the parameters of conventional money demand function. They change the size of fiscal and money multipliers by influencing the interest sensitivity of the variables.

The second category consists of different instruments, institutions and markets which do not directly affect the monetary aggregates but influence factors shaping the structure of the aggregates. They include the establishment of holding companies of banks, new institutions providing financial services including equity capital and risk-ameliorating techniques such as the adoption of variable rate of interest on the asset, maturity shortening of securities, index-linked bonds and financial futures. To keep the enquiry within manageable limits, we would be concerned with the first category.

Based on the experience of the advanced countries, factors conducive to financial innovation may be grouped under three heads, viz., economic, institutional and attitudinal. Economic factors comprise higher opportunity cost of holding idle cash, wide amplitude in the rates of inflation and/or interest and large and growing Government deficits as also the Governmental attempt to non-monetise such deficits. The institutional aspect includes policy measures. Some of them were harsh inducing the institutions to circumvent. Some of them were liberal (with or without incentives) thereby positively encouraging innovation. Technological developments, particularly in telecommunication, data processing, data storing and data retrival systems and extent of specialisation of the financial institutions are some other factors. Specific mention, however, should be made to market expansion and the emergence of new markets, particularly globalisation of financial markets, facilitating quick transfer of impulses from one country to another. Perhaps the most important is the attitudinal factors. These include the growing tendency to find profit in every activity - 'emergence of unabashed financial greed as financial *dernier cri*' the existing firm's deterent and turf protection attempts and the banks' willingness to circumvent rules and regulations to satisfy the financial needs of credit-worthy customers.

An international comparison shows that these aspects, particularly the attitudinal ones, explained both the large scale financial innovation in the USA and its relative absence in the Continent.

## B. International Experience

### The U.S.A.

The USA experienced extensive financial innovation in wholesale banking, retail banking and corporate sector which used the emerging instruments in the first two sub-sectors. Chronologically, the system went through three distinct phases: upto around second quarter of 1976 when the new instruments made their appearance and diffusion was very gradual having virtually no macro-effects; between the second quarter of 1976 to 1980 when new instruments were relatively less but the diffusion was faster thereby throwing an effective challenge to monetary orthodoxy; and since 1980 which is characterised by relatively less number of new instruments domestically (but greater number internationally) accompanied by a process of accelerated diffusion.

The three major innovative instruments in the *wholesale banking* (money market) are the Certificate of Deposits (CD), Repurchase Agreements (RPs or Repos) and the Money Market Mutual Funds (MMMFs). A CD is a negotiable instrument certifying that a sum of money has been deposited with a bank issuing it and that, on the stated maturity date, the deposit will be repaid with interest by the issuing bank. CDs have been in existence since 1950s but gained importance since 1961 when a secondary market for this instrument was established. CDs enable the banks to lend even during a period of credit squeeze by the central bank thus ensuring reliable credit lines to the large borrowers undisturbed by credit shortages. Thus, the tightening of the squeeze in 1966 resulted in large scale adoption of this instrument by the banks and other financial institutions.

RPs is an arrangement in which a security dealer sells a federal fund's instrument with a commitment to repurchase that security in a short period of time at an agreed rate of interest. This benefits the lender with a yield on a balance which would otherwise be idle and it enables the borrower to meet pressing credit obligations. Just

as CDs, RPs have been in existence since 1950s but grew rapidly especially after 1970, became an important source of non-deposit funds to banks and other financial institutions and an important element in liability management by financial firms. From the view point of present study, two of its effects are important; (a) in effect, it enables the lenders to receive interest on demand deposits and (b) it permits the banks to expand larger deposits on the basis of a given supply of reserves.

MMMFs collect household deposits in small amounts and invest them in short-term money market instruments thus giving small investors some benefits from high market interest rates. MMMFs, which began in 1972, has grown since 1978. They promised a rate of return plus facility for withdrawal on demand. Moreover, only relatively small amounts are required to hold these funds. These features enabled the households to participate in the scheme in large numbers resulting in a significant shift away from the conventional demand deposits. This trend diluted the relationship between narrow money and economic activity. While CDs and RPs undermined the effectiveness of monetary policy, MMMFs contributed its mite to the break-down of traditional demand for money function.

Three major instruments in the *retail banking* market also contributed to the break-down of the demand for money function. The Auotmatic Transfer Service (ATS) came into being in 1978. By providing a facility of automatic transfer of funds from interest bearing savings accounts into demand deposit accounts whenever such transfer was necessary to make transaction from the latter, ATS reduced the demand for non-interest bearing money.

The Negotiable Order of Withdrawal (NOW) accounts refer to interest bearing savings accounts in depository institutions from which withdrawals can be made by negotiable drafts rather than demand over the counter. The NOW began in 1970s and gained wider usage since 1976; it initiated a move towards interest bearing accounts with chequeing facilities, blurring the distinction between demand and time or savings deposits. Since 1986, super-NOW came into existence and this offered high interest rates and full chequeing facilities with the requirement of a minimum balance.

The Money Market Deposit Accounts (MMDAs) offered limited transaction service and market related rate of interest.

The *corporate sector* utilised developments in information technology and the rapid growth in new financial instruments in its cash management techniques. Improved information flows and forecasting procedures reduced uncertainty about cash flows and lowered the required level of precautionary balances; new financial instruments provided new profit opportunities and electronic devices enabled switches of funds into higher interest rate media; the process of transfer was quick and at low costs. From the inventory theoretic approach, these developments tend to reduce the overall demand for money.

**Canada**

In Canada, financial innovation occured in three rounds. During the first, which happened in mid-1970s, corporations adopted new cash management techniques. A main form of such innovation was the centralised accounting, which facilitated an integrated bookkeeping for several accounts. Similar methods were extended later to small business enterprises. While enabling the corporate bodies to economise on their cash holding, this wave of innovation had no significant macro-effect.

The second round started towards the end of 1979 when the chartered banks began to offer daily interest saving account (DISA) which entitled the account-holder to earn interest on his daily balance; this was in contrast with the earlier standard saving account envisaging interest entitlement on minimum balance during a specified period, payable at specific intervals. By providing interest on daily balances, the DISA encouraged closer management of demand deposits (and greater competition among banks). At aggregate level, demand deposits showed very little growth between 1980-82.

The third round was more important. Beginning late 1981, the banks introduced the daily interest chequable saving account (DICA) which incorporated both chequeing facilities and saving features. DICA attracted funds from the chequeing account portion of demand deposits. This had important repercussion at macro-level. Since DICA

were not included in the definition of narrow money, this period witnessed a break in the traditional demand for money function and the Bank of Canada temporarily abandoned using $M_1$ as the target.

## Appendix II: The Model

*Innovation Sector*

Demand for innovation $= I_b^d = K[IBM, RE, T, (r^B - r^{IBM})^{+\ +\ +\ +}]$ (1)

Supply of innovation $= I_{INN}^s = I_b^d$. (2)

Equilibrium $= I_b^d = I_{INN}^s = I$ (3)

Let $(r^B - r^{IBM}) = a$.

*Non-banking Sector*

Demand for non-interest bearing money (NIBM) $= NIBM_p^d = N(\bar{r}^{IBM}, \bar{r}^{B}, \bar{I}, \overset{+}{W})$ (4)

Demand for interest-bearing money (IBM) $= IBM_p^d = M(\overset{+}{r}^{IBM}, \overset{+}{r}^{INF}, \overset{+}{I}, \overset{+}{W})$ (5)

Demand for bonds (B) $= B_p^d = B(\bar{r}^{IBM}, \overset{+}{r}^{B}, \bar{I}, W)$ (6)

Wealth constraint $= W = NIBM_p^d + IBM_p^d + B_p^d$ (7)

*Banking Sector*

Supply of non-interest bearing money (NIBM) $= NIBM_b^s = NIBM_p^d = NIBM$ (8)

Suply of interest-bearing money (IBM) $= IBM_b^s = G(\overset{+}{a}, \overset{+}{I})$ (9)

Demand for reserves $= R_b^d = \alpha NIBM + \beta IBM$ (10)

Demand for bonds $= B_b^d = C(\overset{+}{r}^{B})$ (11)

Budget constraint $= NIBM_b^s + IBM_b^s = R_b^d + B_b^d$ (12)

*Central banking sector*

Reserves supplied $= R_c^s = R_b^d = B_c$ (13)

Bond constraint $= B_p^d + B_b^d + B_c = B$ (14)

*Equilibrium conditions*

$K(.) = I_{INN}^s = I$ (15)

$N(.) = NIBM_b^s = NIBM$ (16)

$M(.) = G(.)$ (17)

$[B(.) + C(.)] = (B - B_c)$ (18)

$\alpha NIBM + \beta\ IBM = B_c$ (19)

*Totally differentiated with respect to time, denoting derivatives with . and rearranging the terms*

$$k_1 \dot{IBM} + k_2\dot{RE} + k_3\dot{T} + k_4\dot{a} - \dot{I} = 0 \quad (20)$$

$$-n_1\dot{r}^{IBM} - n_2\dot{r}^{B} - n_3\dot{I} + n_4\dot{W} - \dot{NIBM} = 0 \quad (21)$$

$$m_1\dot{r}^{IBM} + m_2\dot{r}^{INF} + m_3\dot{I} + m_4\dot{W} - g_1\dot{a} - g_2\dot{I} = 0 \quad (22)$$

$$-b_1\dot{r}^{IBM} + b_2\dot{r}^{B} - b_3\dot{I} + b_4\dot{W} + c_1\dot{r}^{B} - (\dot{B} - \dot{B}_c) = 0 \quad (23)$$

$$\alpha\dot{NIBM} + \beta\dot{IBM} - \dot{B}_c = 0 \quad (24)$$

Let $m_3\dot{I} = g_2\dot{I}$ and $\dot{W} = 0$

*The System in Matrix*

$$\begin{bmatrix} -1 & 0 & 0 & 0 & k_1 \\ -n_3 & -n_2 & -n_1 & -1 & 0 \\ 0 & 0 & m_1 & 0 & 0 \\ -b_3 & (b_2+c_1) & -b_1 & 0 & 0 \\ 0 & 0 & 0 & \alpha & \beta \end{bmatrix} \begin{bmatrix} \dot{I} \\ \dot{r}^{B} \\ \dot{r}^{IBM} \\ \dot{NIBM} \\ \dot{IBM} \end{bmatrix} = \begin{bmatrix} -k_2\dot{RE} - k_3\dot{T} - k_4\dot{a} \\ 0 \\ -m_2\dot{r}^{INF} + g_1\dot{a} \\ (\dot{B} - \dot{B}_c) \\ \dot{B}_c \end{bmatrix}$$

*The determinant*

$$\Delta = -m_1\{\beta(b_2 + c_1) + \alpha k_1[n_3(b_2+c_1) + n_2 b_3]\} \quad (25)$$

*Solution for $\dot{I}$*

$$\Delta\dot{I} = (-k_2\dot{RE} - k_3\dot{T} - k_4\dot{a})[\beta m_1(b_2+c_1)] + k_1((-m_2\dot{r}^{INF} + g_1\dot{a}) \{\alpha[n_1(b_2+c_1) - n_2 b_1]\} + (\dot{B} - \dot{B}_c)[\alpha m_1 n_2] - \dot{B}_c[m_1(b_2+c_1)]) \quad (26)$$

*Effects of $\dot{a}$ on*

$$\dot{r}^{IBM} = g_1\{\beta(b_2+c_1) - \alpha k_1[n_3(b_2+c_1) + n_2 b_3]\} \quad (27)$$

$$\dot{r}^{B} = g_1[\beta b_1 - \alpha[n_3 b_1 k_1 - b_3 n_1] - k_4\beta b_3 \quad (28)$$

$$\dot{NIBM} = g_1\{\beta[-n_1(b_2+c_1) + n_2 b_1]\} + k_4[\beta\{-n_2(b_2+c_1) + b_3 n_2\}] \quad (29)$$

$$\dot{IBM} = g_1\{-\alpha[n_1(b_2+c_1) + n_2 b_1]\} + k_4(m_1[\alpha\{n_3(b_2+c_1) + b_3 n_2\}]) \quad (30)$$

*Effects of $\dot{T}$ on*

$$\dot{r}^{IBM} = 0 \quad (31)$$

$$\dot{r}^{B} = k_3\beta m_1 b_3 \quad (32)$$

$$\dot{NIBM} = k_3[\beta\{n_3(b_2+c_1) + n_2 b_3\}] \quad (33)$$

$$\dot{IBM} = -k_3\{\alpha[n_3(b_2+c_1) + n_2 b_3]\} \quad (34)$$

*Effects of $\dot{B}c$ on*

$$\dot{r}^{IBM} = 0 \quad (35)$$

$$\dot{r}^{B} = k_1 b_3 \quad (36)$$

$$\dot{NIBM} = -k_1[-n_3(b_2 + c_1) - n_2 b_3] \quad (37)$$

$$\dot{IBM} = -m_1(b_2 + c_1) \quad (38)$$

**References**

1. Aktar, M.A. (1983): *Financial Innovation and their Implication for Monetary Policy: An International Perspective,* BIS Economic Papers, No.9, Basle

2. Atkinson, P., Blundell-Wignall, A., Rondoni, M. and Ziegelschmidt, H. (1984): The Efficiency of Monetary Targeting: The Stability of Demand for Money in Major OECD Countries, *OECD Economic Studies,* No.3.(Autumn)

3. Baba, Y., Hendry, D.F., and Starr, R.M. (1988): *U.S. Money Demand: 1960-1984, Discussion Papers in Economics, No.27,* Nuffield College, Oxford

4. B.I.S. (1984): *Financial Innovation and Monetary Policy,* Basle

5. Boughton, J.M. (1981): Recent Instability in the Demand for Money : An International Perspective, *Southern Economic Journal,* Vol. 47.No.3 (January)

6. Broemeling, L.D. and Tsurumi, H. (1987): *Econometrics and Structural Change,* Marcel Dekker Inc., New York

7. Cagan, P. and Schwartch, A.J. (1975): Has the Growth of Money Substitutes Hindered Monetary Policy?, *Journal of Money, Banking and Credit,* Vol.7.No. 2 (May)

8. Chirinko, R.S. and Colloton Jr, J.E. (1984): *Money Demand Instability and Financial Innovation, Working Paper, No.319,* Cornell University (July)

9. de Cecco, M. (ed) (1987): *Changing Money: Financial Innovation in Developed Countries,* Basil Blackwell, New York

10. Desai, M. and Low, W. (1987): Measuring the Opportunity for Product Innovation, in *Changing Money: Financial Innovation in Developed Countries,* (Edited by Cecco, M.) Basil Blackwell, New York

11. Dufey, G. and Giddy, I.H. (1981): Innovation in the International Financial Markets, *Journal of International Business Studies* (Fall)

12. Hall, R.E. (1986): Monetary Policy under Financial Innovation and Deregulation in *Financial Innovation and Monetary Policy: Asia and the West,* (Edited by Suzuki, Y. and Yomo, H.) University of Tokyo Press, Tokyo

13. Hester, D.D. (1981) : Innovation and Monetary Control, *Brookings Papers on Economic Activity* : 1

14. Jonson, P.D. and Rankin, R.W. (1986): On Some Recent Developments in Monetary Economics, *The Economic Record* (September)

15. Judd, J. and Scadding, J.(1982): The Search for a Stable Money Demand Function: A Survey of Post-1973 Literature, *Journal of Economic Literature,* 20

16. Kabir, M. and Mangla, I. (1988): Effects of Financial Innovations on the Money Demand Function: A Canadian Evidence,*Applied Economics,* Vol. 20.No.9. (September)

17. Kane, E. (1984): Microeconomic and Macroeconomic Origins of Financial Innovation in *Financial Innovation,* Federal Reserve Bank of St. Louis, St. Louis

18. Milbourne, R. and Moore, H.A. (1986): Some Statistcal Evidence on the Effects of Financial Innovation, *Review of Economics and Statistics* (August)

19. Miller, S.M. (1986) : Financial Innovation, Depository- Institution Deregulation and the Demand for Money, *Journal of Macroeconomics* (Summer)

20 Offenbacker, E.K. (1982) : Effects of Substitution Among Assets and Technological Innovation on Money Demand, *Proceeding of the Business and Economic Section of the American Statistical Association.*

21. Podolski, T.M. (1986) : *Financial Innovation and the Money Supply,* Basil Blackwell, Oxford.
22. Silber, W. (1975) : Toward a Theory of Financial Innovation in *Financial Innovation,* Silber, W.(ed), Lexington Books, Toronto
23. Simpson, T.D. and Porter, R.D. (1980): Some Issues Involving the Definition and Interpretation of the Monetary Aggregates, *Federal Reserve Bank of Boston*
24. Suzuki, Y. and Yomo, H. (1986) : *Financial Innovation and Monetary Policy : Asia and the West,* University of Tokyo Press, Tokyo
25. Wenninger, J. (1984): Financial Innovation-A Complex Problem Even in a Simple Framework, *Federal Reserve Bank of New York Quarterly Economic Review* (Summer)
26. Witte, W.E. (1982) : A Short-Run Analysis of the Effects Of Portfolio Realignments Due to Money Market Innovations, *Journal of Economics and Business,* 34.

*Reserve Bank of India Occasional Papers*
Vol. 9, No. 4, December 1988

# Detection of Influential Observations and Outliers in Linear Regression Analysis

P.C. Sarkar*

## Introduction

IN the regression analysis where a multiple linear regression model: $Y_{Nxl} = X_{Nxp}B_{pxl} + E_{Nxl}$ is fitted to data, a fitted model can be viewed as a smoothed representation that captures global and essential features of the data but this view is not always appropriate. Key features of a fitted model can be dominated by a single observation or a few observations and conclusions in such situations tend to depend critically on the model. However, "models are always approximate descriptions of more complicated processes" @and, therefore, are not always exact. "Because of this inexactness, consideration of model adequacy is extremely important."@ Therefore, different diagnostic approaches have been made to detect the influential observations and outliers by different statisticians and hence to develop better regression models. Recently, Samprit Chatterjee and Ali S. Hadi (Statistical Science, 1986 Vol. I, No.3) have discussed various measures for detection of influential observations and outliers. Out of these various measures, measures based on 'Prediction Matrix' and those based on 'Residuals' are chosen and used in this study.

The main objectives of this study are (i) to identify 'Influential

---

* P.C. Sarker is a Research Officer in the Division of Statistical Analysis, Department of Statistical Analysis and Computer Services, Reserve Bank of India, Bombay. He is grateful to (late) Shri S.P. Gothoskar, former Principal Adviser and Shri R. Nagaraja Rao, Adviser for their encouragement and suggestions. He is indebted to Professor B.K. Kale, Department of Statistics, University of Poona, Pune for his valuable suggestions. Statistical Assistance rendered by Smt. L.A. Koparkar and Smt. P.P. Gavankar is sincerely acknowledged.

@ Comments made by R. Dennis Cook on the article of Samprit Chatterjee and Ali S. Hadi.

Observations' and 'Outliers' in fitting regression models and (ii) to find out the effects of deletion of influential observations/outliers on various measures of goodness of fit and significance of fitted parameters of regression models.

This introductory Section is followed by Section-I dealing with methodology for detecting 'Influential Observations' and 'Outliers'. A brief description of the data is given in Section-II. The empirical findings of the study are discussed in Section-III and conclusions are presented in Section-IV.

## I

## Methodology

Before reviewing the methodology, it is useful to define what is meant by 'influence'. The definition, which seems most appropriate, is given by Belsley, Kuh, and Welsch (1980): 'An influential observation is one which, either individually or together with several other observations, has a demonstrably larger impact on the calculated values of various estimates..... than is the case for most of the other observations'. An observation, however, may not have the same impact on all regression outputs. An observation may have influence on estimates of B, the estimated variance of B and/or goodness of fit statistics. The primary goal of the analysis should determine which influence to consider as most relevant. Considering that prediction is the primary goal or estimation of B is primary concern, the details of following two methods chosen for this purpose are set out.

(i) *Measures based on the Prediction Matrix:*

Consider a multiple linear regression model:

$$Y = XB + E \quad \ldots\ldots\ldots\ldots \quad (1)$$

where Y is an NX1 vector of values of the response (dependent) variable, X is an NXp full-column rank matrix of known predictors (regressors or explanatory variables), B is a pxl vector of unknown coefficients (parameters) to be estimated, and E is an Nxl vector of independent random variables each with zero mean and unknown variance $\sigma^2$. Then by the method of least squares we get the following results:

$$\hat{B} = (X^TX)^{-1} X^TY, \quad \ldots\ldots\ldots\ldots \quad (2)$$

$$\hat{Y} = X\hat{B} = PY \quad \ldots\ldots\ldots\ldots \quad (3)$$

where

$$P = X (X^TX)^{-1} X^T, \quad \ldots\ldots\ldots\ldots \quad (4)$$

$$\text{Var.}(B) = \sigma^2 (X^TX)^{-1}, \quad \ldots\ldots\ldots\ldots \quad (5)$$

$$\text{Var.}(Y) = \sigma^2 P, \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (6)$$

$$e = Y - \hat{Y} = (I - P)\, Y. \quad \ldots\ldots\ldots \quad (7)$$

$$V(e) = \sigma^2 (I - P)\, Y, \quad \ldots\ldots\ldots\ldots \quad (8)$$

$$\text{and } \sigma^2 = \frac{e^Te}{N-P} \quad (9)$$

the residual mean square estimate of Var (Ei) = $\sigma^2$. The matrix P defined in (4) plays an important role in determining $\hat{Y}$, e and their covariance matrices. Now $p_i$, the ith diagonal element of P is

$$P_i = x_i (X^TX)^{-1} x_i^T \quad \ldots\ldots\ldots\ldots \quad (10)$$

which can be thought of as the amount of leverage of the response value $Y_i$ on the corresponding value $\hat{Y}_i$. P is called sometimes, as the Hat matrix because it maps Y into $\hat{Y}$ i.e. $\hat{Y} = PY$. It is also called as a projection matrix because it generates the perpendicular projection of Y (an n-dimensional vector) into a p-dimensional sub-space. Here it can be called as the prediction matrix, becasue applying it to Y produces the predicted values. Hoaglin and Welsch (1978) recommended examination of $p_i$ for high leverage design points and suggested using 2p/N as a calibration point for $p_i$. A high leverage point in the factor space can be defined as a point $x_i$ which has large $p_i$. Points which are isolated in the factor space will have high leverage.

(ii) *Analysis of Residuals*:

One of the early methods of detecting model failures is examining the least squares residual

$$e_i = y_i - x_i \hat{B} \quad \ldots\ldots\ldots\ldots\ldots \quad (11)$$

where $x_i$ is the ith row of X, or preferably, examining scaled version of $e_i$, that is

$$e_i (\sigma) = \frac{e_i}{\sigma\sqrt{1 - p_i}} \quad \ldots\ldots\ldots\ldots\ldots \quad (12)$$

where $p_i$ is the diagonal element of p. Two special cases of (12) are:

$$t_i = e_i (\hat{\sigma}) = \frac{e_i}{\hat{\sigma}\sqrt{1 - p_i}} \quad \ldots\ldots\ldots\ldots\ldots \quad (13)$$

where $\hat{\sigma}$ is defined in (9), and

$$t_i^* = ei(\hat{\sigma}(i)) = \frac{e_i}{\hat{\sigma}(i)\sqrt{1-p_i}} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (14)$$

where $\hat{\sigma}(i) = \frac{Y^T(i)\,(I-P_{(i)})\,Y(i)}{(N-p-1)}$

$$= \frac{(N-p)\,\hat{\sigma}^2}{(N-p-1)} - \frac{e_i^2}{(N-p-1)\,(1-p_i)} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (15)$$

is the residual mean square when the ith observation is omitted. Identity (15) was given by Beckman and Trussell (1974). Equation (14), $t_i^*$ is called 'Jackknife' residuals by Atkinson (1981a). 'RSTUDENT' by Belsley, Kuh and Welsch (1980), and 'Studentized' residuals by Velleman and Welsch (1981). Atkinson (1981a) derived relationship between $t_i$ and $t_i^*$ as follows:

$$t_i^* = t_i\sqrt{(N-p-1)/(N-p-t_i^2)}$$

from which it is observed that $t_i^*$ is a monotonic transformation of $t_i$ and that $t_i^{*2} \to \alpha$ as $t_i^2 \to (N-p)$. Therefore, $t_i^*$ reflects large deviations more dramatically than does $t_i$. Detection of 'Influential Observations' has been performed with these three statistics $p_i$, $t_i$ and $t_i^*$. Again, $t_i$ and $t_i^*$ are mainly used for detection of 'Outliers' taking the standard tabulated values of

$r_i = \frac{t_i^2}{N-p} \sim B\left(\frac{1}{2}, \frac{N-p-1}{2}\right)$ for finding the calibration point for $t_i$ and $t_i^* \sim t(N-p-1)$ as a calibration point for $t_i^*$ at 5% level of significance.

## II

## Data and Period

The above procedure has been applied to two sets of data which have been considered for this study. The first set of data comprises data on Value of Production (VP) of the Company as dependent variable and Raw Material consumed (RM), Power & Fuel consumed (PF), Stores and Spares consumed (SS), Salaries (including wages

& bonus) paid (SL) and Plant and Machinery in use (PM) as independent variables to build up a regression model. All these data have been collected from the Company Finance Division of the Department of Statistical Analysis and Computer Services, Reserve Bank of India, Bombay. The industry group 'Cotton Textiles - Spinning' has been taken for the study and the data pertained to the period 1984-85. Again, only those companies are selected which have reported foreign earnings. A total of 26 companies with six variables have been taken for the study. For the second data set, the monthly-data on Index Numbers of Industrial Production (IIP) are treated as dependent variable and indices of productions of six infrastructure industries (Coal, Cement, Crude Petroleum, Electricity, Petroleum Products and Saleable Steel) are considered as independent variables in fitting a multiple linear regression model. Data on IIP are those published by the CSO and indices on infrastructure industries have been compiled on the basis of production data obtained from Ministry of Industries. Seven variables having 20 observations (from April 1981 to November 1983) for each variable are taken for the study. All these data are presented in Statement-I and Statement-2 respectively.

## III
## Empirical Findings

Taking all the observations for each set, parameters of the linear regression equations were estimated and are presented in Statement-3. For detecting the 'Influential Observations' and 'Outliers', the values of the different statistics ($p_i$, $t_i$ and $t_i^*$) were calculated and are presented in the Statement-4 and Statement-5 for first set and second set respectively. The calibration points were noted/calculated as follows:

*First set (Cotton Textiles)*:

For $p_i$ : $p_i = \frac{2p}{N} = \frac{2\times6}{26} = 0.46$

For $t_i$: $\frac{t_i^2}{N-p} = B\ (0.5,\ 9.5) = 0.1874$ at 5% level of significance.

or $t_i = 1.94$

For $t_i^*$ : $t^* \sim t\,(N-p-1) = 2.093$ at 5% level of significance with 19 degrees of freedom.

For convenience of calculations of $t_i^*$ the relationship between $t_i$ and $t_i^*$ could be written as

$$t_i^* = t_i \sqrt{\frac{N-p-1}{N-p-t_i^2}} = \frac{4.359\ t_i}{\sqrt{20 - t_i^2}}$$

*Second set (Industrial Production)*:

For $p_i$ : $p_i = \frac{2p}{N} = \frac{2\times 7}{20} = 0.70$

For $t_i$ : $\frac{t_i^2}{N-p} = B\ (0.5,\ 6.0) = 0.2835$ at 5% level of significance.

or $t_i = 1.92$

For $t_i^*$ : $t_i^* \sim t(N-p-1) = 2.179$ at 5% level of significance with 12 d.f.

The relationship between $t_i^*$ and $t_i$ which could be written as

$$t_i^* = \frac{3.464\ t_i}{\sqrt{13 - t_i^2}}$$

were used for calculating the value for $t_i^*$.

(A) *Analysis of First Set*:

From Statement-4, the values of $p_i$, $t_i$ and $t_i^*$ for the first set of data were compared with the calibration points of respective statistics and found that six observation: (serial numbers 1, 2, 3, 14, 23 and 24) could be considered as 'Influential Observations' under $p_i$ — statistic scheme. Actually, those were the companies which were having larger value of production (Y) as well as having larger values for inputs (i.e. raw materials, salaries, plant and machinery, etc.) during the year 1984-85. So, in estimating the values for regression coefficients as well as for correlation coefficient, these six observations have played the major role. Again under the scheme $t_i$ and $t_i^*$ 2nd, 6th, 10th and 24th observations might be considered as 'Outliers' as suggested by Hoglin and Welsch (1978). Similarly, 2nd and 24th observations could be considered as highly 'Influential Observations' under the scheme $p_i$, and also as 'outliers' under the

schemes $t_i$ and $t_i^*$. Deleting all the four outlier observations (2nd, 6th, 10th and 24th) regression analysis was performed and the parameters of the regression equation were estimated. These are presented in Statement-6. On comparing the value of X-coefficients, it was observed that all the coefficients were significant before and after deletion of outliers. The value of squared multiple correlation coefficient (i.e. $R^2$) nominally increased even after deletion of 4 observations out of total 26 observations. The value of standard error of estimation (SEE) decreased substantially from 4998.48 to 3102.25 indicating that the model obtained after deletion of the observations is superior to the model obtained by taking all observations. The effect of deletion of influential observations could be visualised from Statement-7. The six influential observations were mainly those companies which were having comparatively larger outputs (i.e. value of production) during the year 1984-85. After deleting these six influential observations, regression analysis was performed and it was observed that deletion of six observations did not improve the model in the context of 'fitness of the model' compared with the model obtained after deletion of four influential observations. The standard error of estimates reduced from 4998.48 to 4229.51 only and value of $R^2$ also reduced marginally. Again, after checking the original data for 2nd and 24th observations which were influentials as well as outliers, it was observed that the 2nd observation had the largest output and 24th observation had moderately larger output during the year 1984-85. Therefore, retaining these two outliers (2nd and 24th observations) and deleting the other two outliers (6th and 10th observations) regression analysis was performed. The results are also presented in Statement-7. The model obtained after deleting these two observations was somewhat better than model obtained after deleting six influential observations but was inferior to the model obtained after deleting all four outliers. This result shows that deletion of outliers yields a better model in terms of our selected criteria in comparison to other modes of data adjustment as described above.

(B) *Analysis of Second Set*

Similar to the exercise done for the first set of data, the values of $p_i$, $t_i$ and $t_i^*$ from Statement-5 were compared with the respective calibration points. It was found that none of the observations could be considered as 'Influential Observations' under $p_i$ statistic.

This might have happened due to the homogeneity of the data unlike in the first set of data on value of production. But under the scheme $t_i$ and $t_i^*$,observations numbered 3rd, 7th and 19th might be considered as 'Outliers' following the procedure suggested by Hoaglin and Welsch (1978). After deleting these three observations, regression analysis was performed taking remaining observations and it was observed that deletion of three observations improved the fitness of regression equation considerably. The value of $R^2$ increased from 0.82 to 0.94 and the value of SEE reduced substantially from 2.91 to 1.89. Here also reduction in residual sum of squares was prominent and proportionately less reduction was reflected in degrees of freedom (reduced from 13 to 10). Similarly, the unexplained portion of the equation (i.e. coefficient of constant) reduced from 46.93 to 12.31 and subsequently, the coefficients of the explanatory variables like electricity, coal, and cement have gone up. So, there was an overall improvement of the equation after deleting three outlier observations.

## IV
## Conclusions

For getting a better representative regression model, a large number of statistical techniques have been proposed to study 'outliers' and influence of individual observation in regression analysis. In this article, we have studied the interrelationships which exist among the three measures ($p_i$, $t_i$ and $t_i^*$) only. These diagnostic checks have been applied to two different sets of data for getting a better view of their capabilities to identifying 'Influential Observations' as well as 'Outliers'. Generally, "each influential measure is designed to detect a specific phenomenon in the data. They are closely related and they all are functions of basic building blocks used in model construction (e.g. residuals $e_i$, standard error, and ith element of the prediction matrix $p_i$). In any particular application, an analyst does not have a look at all the measures since there is a great deal of redundancy."£ Therefore, considering the nature of data and purpose of the model, a suitable equation could be built up for giving a better representation that captures global and essential features of the data.

Outliers should also be scrutinised carefully. If no unusual cir-

£ In summary of the article of Samprit Chatterjee and Ali S Hadi, 1986 (p. 387).

cumstances are found, these observations should not be deleted in a routine manner. To get an idea of the sensitivity of the data, the model should be fitted without the offending observations and the resulting fit should be examined carefully. Deletion of observations always involves loss of information and the degrees of freedom of SEE is also reduced. Therefore, different approaches for replacing these outliers (which may be called as ill-conditioned data) may be found out suitably.One approach is to use a method of fitting that gives less weight to high leverage points (robust estimation) as suggested by Rousseeuw and Leroy (1987) and Carroll and Ruppert (1980). Other types of corrected measures as discussed by Belsley, Kuh and Welsch (1980) for replacing ill-conditioned data are-(1) Introduction of new data, (2) Bayesian type techniques-(a) a pure Bayesian technique and (b) a mixed-estimation technique, and (3) techniques of Ridge Regression. In the case of first type, there is no guarantee that new data will be consistent with the original data. In the Bayesian type, it relies on subjective information and requires a rather exact statement of prior distribution. Ridge regression is also similar to mixed-estimation which imposes some prior restrictions. Therefore, using of those types of techniques depend on the nature of data and those were not attempted in this study.

This exercise has established that application of these three diagnostic checks along with $R^2$ (or $\bar{R}^2$) and SEE, and then re-estimation of the parameters can lead to evolvement of better regression models. In most of our exercises, at present, we use ony $R^2$ and SEE to have better linear regression models, in practice, without eliminating outliers. The packages/programs available in mainframe/PC are not provided with these types of diagnostic checks. Therefore, this study provides a very useful tool in applying better regression models with larger predictive powers and this type of exercise can be executed on PC easily.

## References

1. Belsley, D. A., Kuh, E. and Welsch, R. E. (1980), 'Regression Diagnostics—Identifying Influential Data and Sources of Collinearity'.
2. Chatterjee, S. and Hadi, A. S. (1986), "Influential Observations, High Leverage Points and Outliers in Linear Regression", Statistical Science, Vol. 1, No. 3, pp. 379-416.

3. Chatterjee, S. and Hadi, A. S. (1988), 'Sensitivity Analysis in Linear Regression', John Wiley & Sons, New York.
4. Kale, B. K. (1979), 'Outliers—A Review', Journal of the Indian Statistical Institute, Vol. 17.
5. Kale, B. K. (1976), 'Detection of Outliers', Sankhya: The Indian Journal of Statistics, Vol. 38, Series-B, Pt. 4, pp. 356-363.

**Statement 1: Data for 26 "Cotton Textile - Spinning" Companies which were having Foreign Earnings during 1984-85**

*(Rs. '000)*

| Sr.No. | Value of Production (VP) Y | Raw Materials (RM) $X_1$ | Power & Fuel (PF) $X_2$ | Stores & Spares (SS) $X_3$ | Salaries (SL) $X_4$ | Plant & Machinery (RM) $X_5$ |
|---|---|---|---|---|---|---|
| (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| 1 | 403410 | 233110 | 11012 | 19397 | 71340 | 169875 |
| 2 | 652823 | 343491 | 59441 | 25052 | 81929 | 289472 |
| 3 | 244714 | 136928 | 20818 | 26739 | 22117 | 81649 |
| 4 | 192066 | 126683 | 8955 | 9116 | 16701 | 83182 |
| 5 | 148917 | 72230 | 110999 | 16729 | 19180 | 56267 |
| 6 | 120648 | 69320 | 9359 | 9158 | 19628 | 76967 |
| 7 | 52404 | 36039 | 2684 | 1382 | 7619 | 32288 |
| 8 | 46426 | 27236 | 2104 | 1465 | 7105 | 17717 |
| 9 | 87840 | 36450 | 14546 | 16047 | 6740 | 57586 |
| 10 | 106839 | 60950 | 9426 | 3829 | 17332 | 60878 |
| 11 | 63720 | 34228 | 5722 | 173 | 8432 | 42447 |
| 12 | 271643 | 158928 | 16605 | 9129 | 39128 | 53605 |
| 13 | 36314 | 23015 | 2965 | 1503 | 4024 | 42426 |
| 14 | 536961 | 363606 | 28199 | 21399 | 35071 | 165427 |
| 15 | 97253 | 57420 | 5409 | 3539 | 10559 | 72407 |
| 16 | 52322 | 29848 | 4851 | 2752 | 5158 | 21058 |
| 17 | 59585 | 34504 | 5092 | 2389 | 8911 | 29603 |
| 18 | 68868 | 43780 | 3904 | 1425 | 6506 | 42476 |
| 19 | 92203 | 58844 | 6461 | 2662 | 7115 | 49311 |
| 20 | 71055 | 45331 | 5125 | 2467 | 9387 | 32684 |
| 21 | 79046 | 47110 | 8259 | 1794 | 5357 | 48290 |
| 22 | 54561 | 29358 | 4430 | 1220 | 6515 | 35843 |
| 23 | 588630 | 249676 | 79293 | 31356 | 78107 | 470862 |
| 24 | 471434 | 294651 | 21017 | 17750 | 41082 | 242187 |
| 25 | 63727 | 35247 | 4864 | 1262 | 9145 | 22905 |
| 26 | 180850 | 116252 | 11907 | 3585 | 16713 | 91536 |

**Statement 2: Data on Index Numbers of Industrial Production (IIP) and Indices of Production of Infrastructure Industries (Base: 1980-81=100)**

| Sr.No. | IIP Y | Electricity (EL) $X_1$ | Coal (CO) $X_2$ | Saleable Steel (SSL) $X_3$ | Crude Petroleum (CP) $X_4$ | Petroleum Refinery(PR) $X_5$ | Cement (CE) $X_6$ |
|---|---|---|---|---|---|---|---|
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| 1 | 100.2 | 107.91 | 93.99 | 114.19 | 141.74 | 102.57 | 102.29 |
| 2 | 99.3 | 109.73 | 93.69 | 125.29 | 141.40 | 120.83 | 109.97 |
| 3 | 107.9 | 105.35 | 96.49 | 130.09 | 151.95 | 115.55 | 98.40 |
| 4 | 106.3 | 107.86 | 97.39 | 128.04 | 145.64 | 125.85 | 113.12 |
| 5 | 107.8 | 110.38 | 94.34 | 137.27 | 131.65 | 118.74 | 110.67 |
| 6 | 105.6 | 109.30 | 97.66 | 124.14 | 147.36 | 107.45 | 110.80 |
| 7 | 104.6 | 112.61 | 107.07 | 136.56 | 149.54 | 121.08 | 118.20 |
| 8 | 109.9 | 107.78 | 112.92 | 138.0 9 | 151.03 | 113.96 | 118.97 |
| 9 | 119.5 | 117.22 | 126.64 | 141.95 | 164.33 | 123.51 | 128.55 |
| 10 | 116.5 | 113.92 | 129.22 | 120.71 | 177.18 | 113.57 | 111.00 |
| 11 | 110.7 | 104.22 | 124.80 | 114.82 | 162.61 | 111.72 | 104.18 |
| 12 | 123.4 | 115.25 | 141.32 | 138.07 | 191.86 | 125.85 | 119.81 |
| 13 | 106.8 | 113.28 | 102.28 | 95.65 | 177.52 | 99.04 | 120.71 |
| 14 | 108.2 | 113.92 | 103.23 | 103.56 | 191.28 | 124.31 | 126.56 |
| 15 | 106.5 | 111.66 | 106.39 | 98.93 | 185.90 | 125.56 | 117.56 |
| 16 | 107.8 | 118.31 | 96.65 | 106.96 | 195.53 | 129.09 | 127.01 |
| 17 | 108.2 | 117.57 | 94.39 | 109.63 | 191.06 | 134.46 | 124.24 |
| 18 | 108.7 | 117.76 | 103.23 | 118.42 | 193.81 | 130.96 | 114.98 |
| 19 | 107.1 | 119.83 | 107.97 | 122.81 | 199.89 | 125.60 | 123.79 |
| 20 | 111.6 | 112.44 | 110.91 | 123.95 | 186.58 | 130.53 | 128.94 |

**Statement 3: Results of Regression Analysis taking all Observations**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **A.** | **First set-(Cotton Textiles)** | | | | | | |
| | Constant | : | −4170.14 | | | | |
| | Std.Err.of Y Est. | : | 4998.475 | | | | |
| | R−Squared | : | 0.999419 | | | | |
| | No. of Observations | : | 26 | | | | |
| | Degrees of Freedom | : | 20 | | | | |
| | X-Coefficient(s) | : | 1.162795 | 1.744702 | 0.667827 | 1.265417 | 0.099354 |
| | Std.Err. of Coef. | : | 0.020117 | 0.168347 | 0.208313 | 0.108092 | 0.031838 |
| | Equation: VP | = | −4170.14 + | 1.16 (RM) + | 1.74 (PF) + | 0.67 (SS) + | 1.27(SL)+0.10(PM) |
| **B.** | **Second Set - (Industrial Production)** | | | | | | |
| | Constant | : | 46.93050 | | | | |
| | Std. Err. of Y Est. | : | 2.913437 | | | | |
| | R-Squared | : | 0.820179 | | | | |
| | No. of Observations | : | 20 | | | | |
| | Degrees of Freedom | : | 13 | | | | |
| | X-Coefficient(s) | : | 0.011238 | 0.288906 | 0.106696 | 0.063904 | 0.008507 0.042433 |
| | Std. Err. of Coef. | : | 0.298953 | 0.081441 | 0.100923 | 0.088396 | 0.113820 0.119929 |
| | Equation: IIP | = | 46.93 + 0.01 (EL)+ | 0.29(CO)+0.11(SSL) | + 0.06(CP) + | 0.01(PR)+0.04(CE) | |

**Statement 4: Calculations of $p_i$, $t_i$, and $t_i^*$ for 1st Set of Data relating to Companies (Cotton Textiles)**

| Sr.No. | $p_i$ | $\hat{\sigma}\sqrt{1-P_i}$ | $e_i$ | $t_i$ | $t_i^*$ |
|---|---|---|---|---|---|
| 1 | 0.795270 | 2261.663 | -819.190 | -0.36220 | -0.36340 |
| 2 | 0.557270 | 3325.882 | 6580.359 | 1.978530 | 2.206182 |
| 3 | 0.497186 | 3544.386 | -1060.48 | -0.29920 | -0.29987 |
| 4 | 0.062253 | 4880.390 | -3652.98 | -0.75468 | -0.76566 |
| 5 | 0.201092 | 4467.718 | 7124.756 | 1.594719 | 1.706930 |
| 6 | 0.046989 | 4879.624 | -12885.8 | -2.640740 | -3.27210 |
| 7 | 0.009458 | 4974.781 | -7195.41 | -1.44637 | -1.52852 |
| 8 | 0.005047 | 4985.844 | -89.3206 | -0.01791 | -0.01791 |
| 9 | 0.246635 | 4338.504 | -3278.01 | -0.75556 | -0.76658 |
| 10 | 0.020341 | 4947.375 | -9803.44 | -1.98154 | -2.21036 |
| 11 | 0.018766 | 4951.351 | -527.390 | -0.10651 | -0.10654 |
| 12 | 0.320461 | 4120.449 | -592.567 | -0.14381 | -0.14388 |
| 13 | 0.023835 | 4938.545 | -5254.53 | -1.06398 | -1.09543 |
| 14 | 0.623934 | 3065.276 | -3895.67 | -1.27090 | -1.32555 |
| 15 | 0.051028 | 4869.273 | -407.335 | -0.08365 | -0.08366 |
| 16 | 0.004301 | 4987.714 | -2023.12 | -0.40562 | -0.40729 |
| 17 | 0.004777 | 4986.520 | -4694.88 | -0.94151 | -0.96310 |
| 18 | 0.016729 | 4956.487 | -1390.62 | -0.28056 | -0.28111 |
| 19 | 0.019093 | 4950.526 | -2117.72 | -0.42777 | -0.42974 |
| 20 | 0.006088 | 4983.234 | -6543.35 | -1.31307 | -1.37361 |
| 21 | 0.024913 | 4935.819 | -2221.24 | -0.45002 | -0.45232 |
| 22 | 0.006853 | 4981.318 | 656.1462 | 0.13172 | 0.131778 |
| 23 | 0.839160 | 2004.625 | -499.911 | -0.249370 | -0.24976 |
| 24 | 0.511513 | 3493.526 | 10703.34 | 3.06376 | 4.205761 |
| 25 | 0.013558 | 4964.474 | 80.91020 | 0.016297 | 0.016297 |
| 26 | 0.073436 | 4811.441 | -5787.840 | -1.202930 | -1.248960 |

**Statement 5: Calculations of $p_i$, $t_i$ and $t_i^*$ for the 2nd Set of Data on Indices of Industrial Productions**

| Sr.No. | $P_i$ | $a\sqrt{i-p_i}$ | $e_i$ | $t_i$ | $t_i^*$ |
|---|---|---|---|---|---|
| 1 | 0.299968 | 2.437611 | 0.891965 | 0.365917 | 0.354349 |
| 2 | 0.229952 | 2.556610 | 4.405890 | 1.723332 | 1.871133 |
| 3 | 0.398377 | 2.259789 | −5.62192 | −2.48781 | −3.20993 |
| 4 | 0.240098 | 2.539711 | −1.98262 | −0.78064 | −0.76927 |
| 5 | 0.269370 | 2.490315 | −2.64174 | −1.06080 | −1.06609 |
| 6 | 0.181790 | 2.635348 | −1.76572 | −0.67001 | −0.65633 |
| 7 | 0.141215 | 2.699901 | 5.813511 | 2.153230 | 2.537435 |
| 8 | 0.311030 | 2.418275 | −0.46317 | −0.19153 | −0.18482 |
| 9 | 0.263316 | 2.500611 | 0.288169 | 0.115239 | 0.111114 |
| 10 | 0.327407 | 2.389361 | 0.394317 | 0.165030 | 0.159201 |
| 11 | 0.365565 | 2.320594 | −0.75342 | −0.32466 | −0.31408 |
| 12 | 0.411361 | 2.235270 | −0.10067 | −0.04504 | −0.04341 |
| 13 | 0.593877 | 1.856668 | −2.39550 | −1.29021 | −1.32468 |
| 14 | 0.212252 | 2.585825 | −1.44024 | −0.55697 | −0.54283 |
| 15 | 0.340770 | 2.365505 | 0.154421 | 0.065280 | 0.062923 |
| 16 | 0.208362 | 2.592202 | −0.50557 | −0.19503 | −0.18821 |
| 17 | 0.242486 | 2.535717 | −1.45489 | −0.57375 | −0.55957 |
| 18 | 0.301133 | 2.435582 | 1.721776 | 0.706925 | 0.693791 |
| 19 | 0.330374 | 2.384085 | 5.802330 | 2.433775 | 3.088079 |
| 20 | 0.331289 | 2.382456 | −1.16780 | −0.49016 | −0.47650 |

**Statement 6: Results of Regression Analysis after Deleting Outliers**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **A.** | **First set-(Cotton Textiles)** | | | | | | |
| | Constant | : | -2192.34 | | | | |
| | Std. Err. of Y Est. | : | 3102.252 | | | | |
| | R-Squared | : | 0.999711 | | | | |
| | No. of Observations | : | 22 | | | | |
| | Degrees of Freedom | : | 16 | | | | |
| | X-Coefficient(s) | : | 1.131065 | 1.861177 | 0.739758 | 1.354265 | 0.067367 |
| | Std. Err. of Coef. | : | 0.013830 | 0.162083 | 0.144215 | 0.082814 | 0.030694 |
| | Equation | : | VP=-2192.34 | +1.13(RM) + | 1.86(PF) + | 0.74(SS)+1.35 | (SL)+0.07(PM) |
| **B.** | **Second set - (Industrial Production)** | | | | | | |
| | Constant | : | 12.31064 | | | | |
| | Std. Err. of Y Est. | : | 1.893475 | | | | |
| | R-Squared | : | 0.939140 | | | | |
| | No.of Observations | : | 17 | | | | |
| | Degrees of Freedom | : | 10 | | | | |
| | X-Coefficient(s) | : | 0.342175 | 0.315339 | 0.092424 | 0.010311 | -0.00763 | 0.108214 |
| | Std. Err. of Coef. | : | 0.219205 | 0.070504 | 0.090539 | 0.082315 | 0.083941 | 0.088779 |
| | Equation | : | IIP=12.31064 | +0.34(EL) + | 0.32(C)+0.09 | (SSL)+0.01(CP) | +(-0.01)(PR) | +0.11(CE) |

**Statement 7**

**A. Results of Regression Analysis after deleting six 'Influential Observations'**

*First Set - (Cotton Textiles):*

| | | | | | | |
|---|---|---|---|---|---|---|
| Constant | : | 1091.182 | | | | |
| Std. Err. of Y Est. | | 4229.505 | | | | |
| R-Squared | : | 0.996287 | | | | |
| No.of Observations | : | 20 | | | | |
| Degrees of Freedom | : | 14 | | | | |
| X-Coefficient(s) | : | 1.250254 | 1.651511 | 1.046930 | 0.953104 | -0.06083 |
| Std. Err. of Coef. | : | 0.070987 | 0.519499 | 0.318893 | 0.283110 | 0.073596 |
| Equation | : | VP=1091.18 | +1.25(RM) + | 1.65(PF)+1.05(SS) | + 0.95 (SL) | -0.06 (PM) |

**B. Results of Regression Analysis after deleting two observations which are only 'Outliers'**

*First Set - (Cotton Textiles):*

| | | | | | | |
|---|---|---|---|---|---|---|
| Constant | : | -3186.73 | | | | |
| Std. Err. of Y Est. | : | 4211.861 | | | | |
| R-Squared | : | 0.999623 | | | | |
| No. of Observations | : | 24 | | | | |
| Degrees of Freedom | : | 18 | | | | |
| X-Coefficient(s) | : | 1.154228 | 1.708710 | 0.691147 | 1.296957 | 0.1025569 |
| Std. Err. of Coef. | : | 0.017172 | 1.142945 | 0.177586 | 0.091700 | 0.026907 |
| Equation | : | VP=-3186.73 | + 1.15(RM) + | 1.71 (PF) + | 0.69(SS)+1.30(SL)+0.10(PM) | |

*Reserve Bank of India Occasional Papers*
Vol. 9, No. 4, December 1988

# Kharif Production and Monsoon Rainfall
## An Empirical Study@

R. R. Arif*

### Introduction

UNDOUBTEDLY the extent of monsoon rainfall still is the primary factor influencing the volume of agricultural output in India. The crucial question is how significant a role the monsoon plays and how its influence can be empirically quantified. The main objective of this paper is to construct behavioural functions that estimate the quantum of kharif foodgrains produced in India in relation to the spatial and temporal distribution of monsoon rainfall. A second objective is to examine the predictability of kharif output based on the progress of the monsoon month by month.

The study was prompted by the severe drought of 1987 which again under-scored the rain-dependence of agricultural production in India. Uncertainty regarding the onset of the monsoon and subsequently its vagaries during the main stages of crop growth affect cultivation in different agro-climatic zones and ultimately the aggregate foodgrains output of the kharif season. It was, therefore, necessary to first examine the pattern of monsoon precipitation in various parts of the country during the past decade. Monthly rainfall data of 35 representative meteorological regions for the years 1979 to 1986 were analysed and a monsoon rainfall index (MRI) was devised to obtain an appropriately weighted aggregate measure of rainfall. Several MRI series were then constructed for 19 states and for the country as a whole. These details appear in Section I.

---

* At the time of writing this paper Smt. R. R. Arif was an Assistant Adviser in the Department of Statistical Analysis and Computer Services.

@ The study was undertaken under the guidance of Dr. C. Rangarajan. The author acknowledges with thanks the encouragement received from Dr. D.K. Bhatia and the data made available by Shri V.H. Motwani. Shri A.K. Mohapatro and Shri S.B. Wavikar provided competent computational assistance. The study was presented at the 26th Indian Econometric Conference, 1989.

Attention was next focused on establishing a relationship between the series of kharif foodgrain production and the MRI. Section II discusses the various alternatives considered in the specification of aggregate output functions and the empirical results covering the sample period 1979 to 1986.

Attempts to forecast one and two years beyond the sample period were also made. This exercise prompted further investigation in two directions: (i) modifications in the form of the all-India supply function and (ii) state-wise disaggregated estimation. Section III deals with these aspects. Section IV concludes with a summary. The Appendices contain references, tables and graphs.

## I
## Analysis of Rainfall

The basic data analysed in this exercise pertains to the monsoon rainfall recorded in 35 meteorological regions (please see Appendix 1) during the years 1979 to 1986. On the average, June gets about 20% of total monsoon rain. By July end 50% and by August 80% of the monsoon precipitation is generally recorded. On the whole, July and August are the most crucial months of the monsoon. There are noticeable regional differences in the pattern of rainfall from month to month.

The simple correlation coefficients between monthly rainfall were

| May & June - July | July-August | August-September |
|---|---|---|
| 0.51 | 0.86 | 0.92 |

Cross-section regressions across the 35 meteorological regions fitted for a 'good' monsoon year (1983) are given in Appendix 2. A multiple regression to measure the relative contribution of each month's rainfall to total monsoon precipitation indicates the following:

| Period: | May to June | July | August | May to August |
|---|---|---|---|---|
| Mean Elasticity w.r.t. total rain | 0.07 | 0.46 | 0.38 | 0.91 |

Information on rainfall upto June-end is inadequate to estimate the overall monsoon precipitation. Though the explanatory variable is significant, the predictive power of the regression is poor. At the end of July, the variation in total monsoon rainfall could be better

explained. The margin of error, however, would be too high for predictive purposes. By August end, the predictions improve. The regression explained on an average 91 to 97 per cent of total monsoon rainfall.

## Monsoon Rainfall Indices (MRI)

Two types of aggregate rainfall indices related to (i) total foodgrains production ($MRI^f$) and (ii) rice cultivation ($MRI^r$) were constructed as described below:

For kharif foodgrains, the region's volume of production relative to total foodgrains production in 1982

$$\text{i.e., } W_i^f = \frac{f_i}{\Sigma f_i} \times 100$$

was considered as the weight. Accordingly, the all-India rainfall index associated with kharif foodgrain production would be

$$MRI^f = \sum_{i=1}^{35} W_i^f R_i$$

Where $R_i$ = observed precipitation in region i;
$f_i$ = production of kharif foodgrains in region i.

For the purpose of rice production, an alternate assumption viz., that the importance of rainfall in a region is proportionate to the area under rice cultivation ($a_i$) was made. Here, the weight for region i is the percentage of $a_i$ to total area cultivated in 1985.

$$\text{i.e., } W^r_i = \frac{a_i}{\Sigma a_i} \times 100$$

The all-India rainfall index for rice production was computed as

$$MRI^r = \sum_{i=1}^{35} W_i^r R_i$$

A preliminary analysis indicates that between the two weighting patterns, production weights seem to be more appropriate. Testing of this assumption by constructing production-weighted indices for rice and comparing them with the area-weighted MRI is proposed

to be taken up in a separate study. Instead of a single year's production one might also consider using the average of 3 to 4 years as a more suitable weight.

The monsoon was the worst in the first year (1979) of the sample with deficient rain in 17 out of 35 meteorological regions (please see Appendix 3). The best year in the sample period was 1983, with near-normal rainfall throughout the country. While the overall monsoon in 1987 was poor in 20 out of 35 meteorological regions, it was less deficient in states like Assam, Bihar and West Bengal. In 1988, the recorded rainfall was above normal in all the regions and the highest observed in many years. Compared to the sample average, the cumulative MRI of 1987 and 1988 reveals the following percentage deviations.

| Year | MRI | |
|---|---|---|
| | August | September |
| 1987 | −11.2% | −12.7% |
| 1988 | 30.1% | 32.1% |
| Average 1979-1986 | 682.2 | 840.8 |

Series of all-India indices for cumulative rainfall as at the end of June, July, August and September computed for the years 1979 to 1988 are given in Appendix 4. In the short series under review, no clear cyclical pattern could be discerned. The MRI varied by about 18 per cent around the sample mean.

The above are all-India aggregates. Similarly, MRI series were constructed for 19 states to be used separately for the state-wise analysis.

## II
## Production Functions

The basic premise is that major determinants of kharif production in India are the weather and technological change. This study focuses mainly on the influence of the former. The monsoon rainfall index is used in different forms as the explanatory variable to measure the response of crop growth to weather conditions. The latter, namely

technological change, is a catch-all term representing diverse factors such as the extent of irrigation, consumption of fertilisers, HYV and other farm inputs, mechanisation etc. Assuming that the technological factor[2] changes gradually at a uniform rate over the years, time is used as its proxy variable.

The function $Q_t = \alpha B^t$ fitted for data from 1970 to 1986 estimated that the long-term trend growth in all India kharif foodgrains production is 2 per cent per annum. The observed annual production figures deviated considerably from the trend line, clearly indicating that one must look to the main factor (viz., weather) to explain these annual fluctuations. Appendix 5 shows that the year-to-year variations in production closely follow the MRI. Accordingly, attention was focused on establishing a behavioural relationship between kharif production and the MRI.

Several alternative specifications were considered for the all-India output function. The sample comprised the following data for the period 1979 to 1986: annual production of kharif foodgrains in million tonnes (Q), the monsoon rainfall index (R) as at the end of May, June, July, August and September and lastly the trend variable (t).

The first formulation represents a behavioural relationship of output with time and the monsoon rainfall index covering total precipitation during June through September.

$$\text{Ln } Q = a + bt + c \text{ Ln } R \qquad \ldots(1)$$

In a second variant, four-monthly rainfall indices (r) replaced the cumulative MRI to give the specification:

$$\text{Ln } Q_t = a + bt + \sum_{\text{M: June}}^{\text{Sept.}} C_m \text{ Ln}^r m \qquad \ldots(2)$$

The assumption here is that the monthly pattern of rainfall has greater relevance than a single aggregate monsoon variable.

In the third case, the variables appear as ratios, with the underlying assumption that there is a positive association between the annual variations in rainfall and production, and that the technological factor does not make a significant impact in a short span of one year. Accordingly the specification was:

$$\text{Ln } (Q_{t+1}/Q_t) = a + b \text{ Ln } (R_{t+1}/R_t) \qquad \ldots(3)$$

Following another approach the supply model was estimated in two stages, assuming that deviations from the trend line are weather-induced.

$$\hat{Q}_t = \alpha \beta^t \qquad \ldots(4.1)$$

$$\text{Ln } (Q_t - \hat{Q}_t) = a + b \text{ Ln } R_t \qquad \ldots(4.2)$$

Here, the residual term of the first equation becomes the dependent variable at the second stage. Though this formulation gave a statistically adequate fit, it did not perform as well compared to the other 3 alternatives.

The regression estimates of the various alternative functions are listed in Appendix 6. The corresponding production estimates appear in Appendix 7. Adequate fits indicated by meaningful coefficients, high $\bar{R}^2$ (0.83 – 0.96%) and low RMSE (1.3 – 4.6) were obtained. The presence of auto-correlation was indicated in some equations.

Alternative 2, which has individual month's rainfall as explanatory variable appears to be a good specification. It also had the lowest RMSE of prediction (1.3). The intutive assumption that rainfall at different stages of crop growth explains better the influence of weather is empirically justified. Not only do the months of July and August together account for the bulk of monsoon precipitation (about 60 per cent), but, also have the most significant coefficients. This model estimates the quantum of kharif foodgrains production within 1 to 20 million tonnes during the sample years.

To pursue our second objective namely, obtaining early predictions of kharif output, the following method was employed.

$$\hat{R} : \text{Sept.} = f(R : \text{July})$$
$$\hat{R} : \text{Sept.} = f(R : \text{Aug.})$$

It was found that the progress of the monsoon upto July end, though indicative, is inadequate to generate a reliable forecast of production. By the end of August, a tentative estimate could be generated using the second of the above equations in the formulation.

$$\text{Ln } Qt = a + bt + c \text{ Ln } \hat{R}t \qquad \ldots(5)$$

In comparison with the corresponding model that uses the actual rainfall upto September end (i.e., alternative 1), the advance estimates had a wider error margin of $\pm$ 5 million tonnes. Nevertheless, alternative 5, based on rainfall data upto August end, performs better than the naive trend projection of 2 per cent per annum.

## III
## Forecasts

1987 was a year of drought and the MRI was 12.7 per cent below the sample average. In 1979, also a year of drought, the MRI was 17.5 per cent below the average and production was 14.6% below trend. Comparing the monthly rainfall figures, one finds that in 1987 the monsoon during June and July was more deficient than in 1979 and showed slight improvement in August. By that time it was perhaps too late to salvage the kharif crops. For 1987, the government estimated a sharp decline in output. Our model forecasts turn out to be 5 to 9 million tonnes above the official figure of 73 million tonnes. Since the behaviour of the monsoon in the initial weeks influences the farmers' decisions regarding areas to be sown, seeds and other inputs to be used and so on, it may be necessary to quantify these aspects to improve the predictability. A second look at the rainfall data also revealed that in some major rice producing regions, the total monsoon was not deficient. Some degree of geographical disaggregation in the analysis seems to be necessary. The current year (1988) witnessed unprecedented rains throughout the monsoon months. Excessive rainfall was recorded in all the 35 meteorological regions. Thus, the MRI for 1988 turns out to be an outlier—about 32 per cent above the sample mean. The 3 alternative forecasts of kharif foodgrain output in 1988 were on the higher side—exceeding the official figure of 92.5 million tonnes by 3 to 15 million tonnes. Apparently the functional specification should quantify the diminishing supply response to monsoon rain after a certain amount of high precipitation. Another approach would be to relate the geographical distribution of rainfall explicitly to the region-wise kharif crop outputs. Preliminary investigations were made in both these directions.

Assuming that excessive rainfall has an unfavourable effect on agricultural production, an excess rainfall variable $\overset{2}{Z}$ with *a priori* negative regression coefficient was sought.

The function

$$\text{Ln } Q_t = b_0 + b_1 t + b_2 Z + b_3 Z^2 \quad (b_3 < 0) \qquad \ldots(6)$$

Where Z = LnR, gave a satisfactory fit, with RMSE = 2.98. With the further assumption that rainfall more than 19-20 per cent above the sample average does not increase production, the value of Z and $Z^2$ were scaled down. The estimate for 1987 remains unaffected while for 1988 the forecast is 101.8 million tonnes.

**Forecasts of All India Kharif Foodgrain Output**
(Million tonnes)

| Basis | 1987 | 1988 | RMSE |
|---|---|---|---|
| Official figure | 73.0 | 92.5 | — |
| Trend Projection (2% per annum) | 86.6 | 88.2 | 8.81 |
| Advance estimation at August end (Equation 5) | 83.7 | 100.2 | 5.37 |
| Revised Model (Equation 6) | 80.9 | 101.8 | 2.91 |
| Aggregate of 19 State-wise forecasts (Equation 7) | 81.5 | 100.2 | 2.59 |

For each of the 19 states, the following function was estimated over the sample period 1979 to 1986.

$$\text{Ln } Q_s = a + bt + c \text{ Ln } R_s \qquad \ldots(7)$$

All the regressions were not uniformly good. The estimates aggregated over the 19 states, however, compared well with the observed all-India data (RMSE = 2.59). For 1987 and 1988 the state-wise forecasts add up to 81.5 and 96.2 million tonnes respectively.

## IV
## Conclusion

The study proceeded initially with the hypothesis that the level of rainfall is the crucial factor determining the quantum of kharif production in India and that its influence can be quantified through behavioural relationships. Other equally important, but slow-

changing influences, broadly labelled as technological factors, were represented by a single catch-all trend variable. The production-weighted rainfall index (MRI) constructed to obtain an aggregate all-India measure of the monsoon seems to be better than an area-weighted index. The MRI series was positively and significantly associated with the quantum of kharif output produced during the years 1979 to 1986. Analysis of the pattern of precipitation during June-September substantiates that July and August are the most important months for crop growth: together they account for about 60 per cent of total monsoon rainfall.

The estimated functions were good fits, judged by the usual criteria. Rainfall and technological change together explain over 90 per cent of the variability in foodgrain production. The predictions of the quantum of foodgrain output during the sample period were within $\pm$ 2 million tonnes of the observed values. Early prediction of output based on the observed monsoon upto August was possible with a $\pm$ 5 million tonnes margin. This method proved to be better than the naive projections of 2 per cent per annum obtained from the long-term growth curve.

A preliminary investigation at the disaggregated level was also done. The highly divergent growth rates and the varying reliability of the estimated functions from state to state indicated that a more detailed analysis beyond the scope of the present study is called for. Nevertheless, the sums of the 19 States' estimates of production were reasonably close to the actual all-India aggregates.

The unusually heavy monsoon in 1988 prompted us to build into the model a diminishing or negative supply response to higher values of the MRI. Forecasts were made one and two years beyond the sample period. Considering the simple aggregative functional approach used, they compared satisfactorily with the official estimates.

A brief mention may be made of some issues that could not be examined in the present exercise. An obvious improvement to the MRI is to use a 'normal' year's production or the average production of the latest 3 years as weights. A fairly adequate level of irrigation could mitigate the adverse effects of a poor monsoon. Accordingly, the proportion of irrigated area to the total area under kharif cultivation would be a relevant factor in the analysis. This information could be taken into account in the construction of the MRI itself. Another aspect that bears investigation is the influence of the previous

year's monsoon which affects the availability of water stored in catchment areas. A series on area under kharif foodgrains cultivation could also be included as an explanatory variable.

The study raises certain questions that might be probed through empirical work in future. Is there an optimum level of monthly rainfall beneficial to kharif crops? Would this ideal pattern differ significantly from region to region? To what extent will the short-term prediction of aggregate kharif foodgrains output improve by disaggregated analysis of individual crops like rice, coarse grains and pulses?

Setting aside several complex issues, this paper concentrated on the quantification of the elusive 'weather' factor as the main explanatory variable in an agricultural supply relationship; and to the choice of an appropriate model from a variety of alternatives. The econometric analysis may be of interest as it provides empirical evidence in support of a familiar phenomenon: namely, the rain-dependence of Indian agriculture. These preliminary findings also indicate the direction in which further investigations could be fruitfully pursued.

---

1 Weekly Weather Report (30th September 1987), Indian Meteorological Department.

2 The possibility of structural breaks in the supply function and the non-uniform adaptation to new technology are not considered.

## Appendix 1: Map of Meteorological Regions with States

| No. | State | Met. No. | Region | Weight@ |
|---|---|---|---|---|
| 1. | 2. | 3. | 4. | 5. |
| 1. | Andhra Pradesh | 27 | Coastal A.P. | 5.7 |
| | | 28 | Telengana | 4.4 |
| | | 29 | Rayalseema | 1.0 |
| 2. | Assam, Meghalaya | 3 | Assam& Meghalaya | 4.1 |
| 3. | Bihar | 8 | Bihar Plateau | 2.2 |
| | | 9 | Bihar plains | 4.0 |
| 4. | Gujarat, Dadra | 21 | Gujarat | 3.0 |
| | & Nagar Haveli | 22 | Saurashtra & Kutch | 0.8 |
| 5. | Haryana, Delhi | 13 | Haryana | 2.8 |
| 6. | Himachal Pradesh | 15 | Himachal Pradesh | 0.7 |
| 7. | Jammu & Kashmir | 16 | Jammu & Kashmir | 1.5 |
| 8. | Karnataka | 31 | Coastal Karnataka | 0.2 |
| | | 32 | North Interior-Karnataka | 2.1 |
| | | 33 | South Interior Karnataka | 4.9 |
| 9. | Kerala | 34 | Kerala | 1.7 |
| | | 35 | Lakshadweep | — |
| 10. | Madhya Pradesh | 19 | West M.P. | 3.6 |
| | | 20 | East M.P. | 5.6 |
| 11. | Maharashtra & Goa | 23 | Konkan & Goa | 1.5 |
| | Daman & Diu | 24 | Madhya-Maharashtra | 2.4 |
| | | 25 | Vidarbha | 2.6 |
| | | 26 | Marathawada | 3.2 |
| 12. | Manipur, Nagaland Mizoram & Tripura | 4 | Nagaland & Others | 1.1 |
| 13. | Orissa | 7 | Orissa | 4.9 |
| 14. | Punjab | 14 | Punjab | 7.2 |
| 15. | Rajasthan | 17 | West Rajasthan | 1.1 |
| | | 18 | East Rajasthan | 2.6 |
| 16. | Tamil Nadu & Pondicherry | 30 | Tamil Nadu | 6.6 |
| 17. | Uttar Pradesh | 10 | East U.P. | 6.9 |
| | | 11 | Plains of West U.P. | 3.4 |
| | | 12 | Hills of West U.P. | 1.9 |
| 18. | West Bengal & | 5 | Sub Himalayan West Bengal | 1.1 |
| | A.N. Islands | 6 | Gangetic W.B. | 5.1 |
| | | 1 | Andaman & N.Island | — |
| 19. | Arunachal Pradesh | 2 | Arunachal Pradesh | 0.2 |

@ Percentage Production of Kharif foodgrains in 1982.

## Appendix 2: Rainfall Regressions@

| | $R^2$ | D.W. | SEE | Mean |
|---|---|---|---|---|
| R = 102.12882* + 1.23876 R* aug.<br>(2.661) (34.217)<br>0.0866 0.913 | 0.97 | 1.50 | 129.92 | 1179.46 |
| R = 242.32765* + 1.96628 R* july<br>(4.080) (20.025)<br>0.205 0.795 | 0.92 | 1.69 | 216.37 | 1179.46 |
| R = 687.75694* + 1.44145 R* june<br>(3.885) (3.595)<br>0.583 0.417 | 0.26 | 2.05 | 665.13 | 1179.46 |
| R = 97.95914 + 1.80673 R july + 1.14376 R aug.<br>(1.805) (7.355) (5.462)<br>0.083 0.461 0.381<br>+ 0.25744 R (may + june)<br>(2.252)<br>0.074 | 0.96 | 1.05 | 160.66 | 1179.46 |

@ Cross section analysis of 1983 rainfall in 35 met. divisions.
* Indicates significance.

| | | | | |
|---|---|---|---|---|
| $R^{£}$ = 109.18388 + 1.07249* R aug.<br>t : (0.541) (3.643)<br>0.130 0.870 | 0.64 | 2.82 | 58.96 | 840.78 |

£ Regression of MRI over the sample period 1979 to 1986.

**Appendix 3: Deficient Rainfall Pattern: 1979-1988**

**Months: June-August**

| Year | Meteorological Divisions | No. of Deficient Divisions |
|---|---|---|
| 1979* | 1,5,8,10,11,12,13,14,15,16,18,19,20,21,23,27,28 | 17 |
| 1980* | 4,12,17,29,30,35 | 6 |
| 1981 | 4,12,17,25 | 4 |
| 1982 | 1,7,8,12,14,15,17,22,26,27,29,30,34 | 13 |
| 1983 | 9,10,12,15 | 4 |
| 1984 | 1,12,15,17,18,22,25,26,27,28,33 | 11 |
| 1985 | 8,21,22,25,26,34 | 6 |
| 1986 | 1,2,3,4,6,8,9,17,22,23,25,30,35 | 13 |
| 1987 | 1,4,7,10,11,12,13,14,15,16,17,18,20,21,22,27,31,32,34,35 | 20 |
| 1988 | | Nil |

Inclusive of September rainfall.

Out of 35 Divisions, those with rainfall.

20% below "normal" are listed.

**Appendix 4: All India Monsoon Rainfall Indices**

| Year | MRI: FOODGRAINS | | | |
|---|---|---|---|---|
| | June | July | August | September |
| 1979 | 159.73 | 379.69 | 542.21 | 693.61 |
| 1980 | 243.73 | 519.30 | 808.94 | 958.61 |
| 1981 | 169.63 | 443.46 | 689.21 | 865.02 |
| 1982 | 133.39 | 316.56 | 665.82 | 761.93 |
| 1983 | 127.39 | 368.07 | 712.86 | 988.69 |
| 1984 | 191.23 | 450.54 | 688.13 | 847.13 |
| 1985 | 112.51 | 419.35 | 714.04 | 817.92 |
| 1986 | 166.08 | 377.75 | 635.96 | 793.12 |
| Sample mean | 162.96 | 409.34 | 682.15 | 840.78 |
| | (19.4) | (48.7) | (81.1) | (100.0) |
| 1987 | 121.39 | 293.43 | 605.70 | 734.11 |
| | (16.5) | (40.0) | (82.5) | (100.0) |
| 1988 | 172.73 | 530.76 | 887.17 | 1110.52 |
| | (16.3) | (50.0) | (83.6) | (100.0) |
| | (163) | (500) | (836) | (1000)* |

Figures in brackets are cumulative percentage to total monsoon rainfall.

* Ceiling of 19 per cent above the sample mean. The earlier 3 months' MRI is proportionately scaled down.

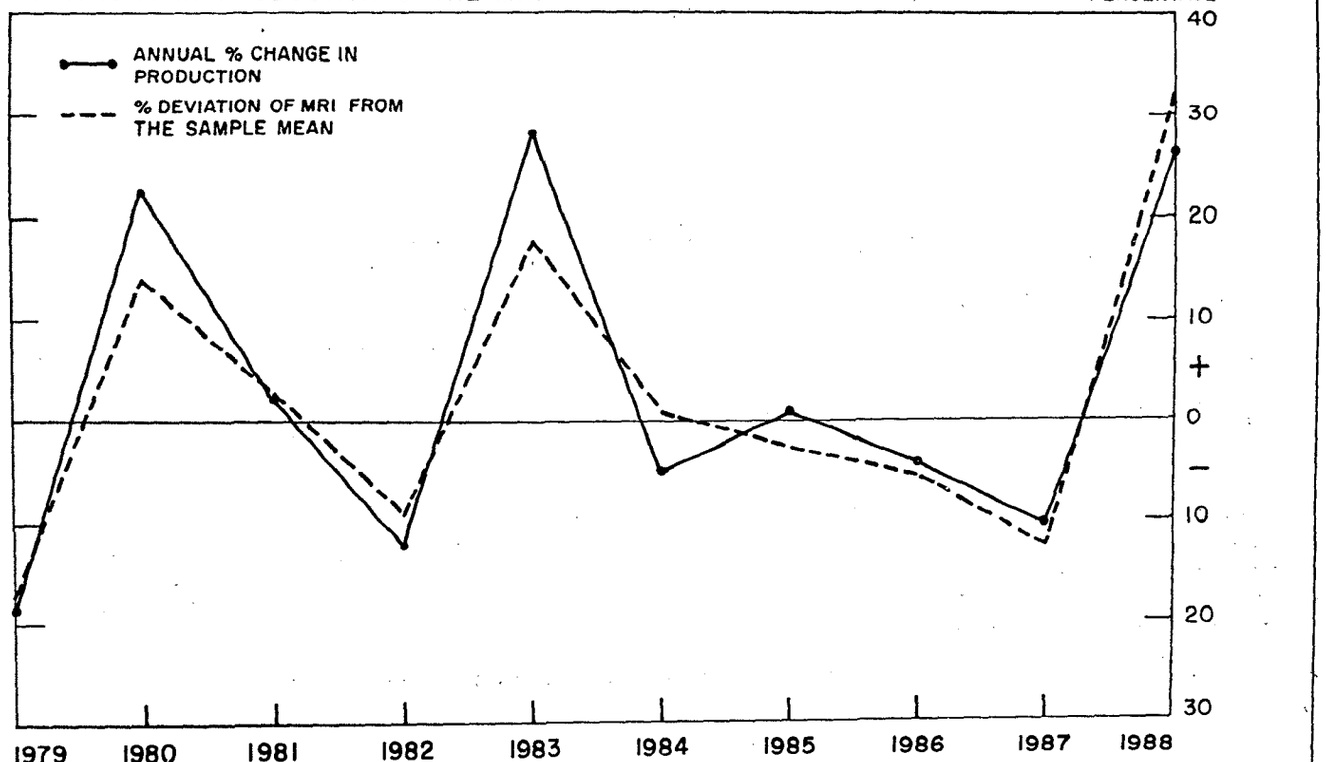
**Appendix 5: Variations in Production & the MRI**

| Year | Annual % Change in Production | Deviation of Production@ from trend | % Deviation of MRI from the sample mean |
|---|---|---|---|
| 1979 | -18.99 | -14.56 | -17.50 |
| 1980 | 22.76 | 1.54 | 14.01 |
| 1981 | 2.23 | 3.12 | 2.88 |
| 1982 | -12.46 | -11.48 | - 9.38 |
| 1983 | 28.22 | 11.51 | 17.59 |
| 1984 | - 5.14 | 3.53 | 0.76 |
| 1985 | 0.85 | 2.39 | - 2.72 |
| 1986 | - 4.35 | - 3.97 | - 5.67 |
| 1987 | -10.46* | -15.68* | -12.69 |
| 1988 | 26.71* | 4.78* | 32.08 |

@ Kharif foodgrains in million tonnes: rice + coarse grains + pulses.
* Provisional data.

APPENDIX 5 A

## VARIATIONS IN PRODUCTION & THE MRI



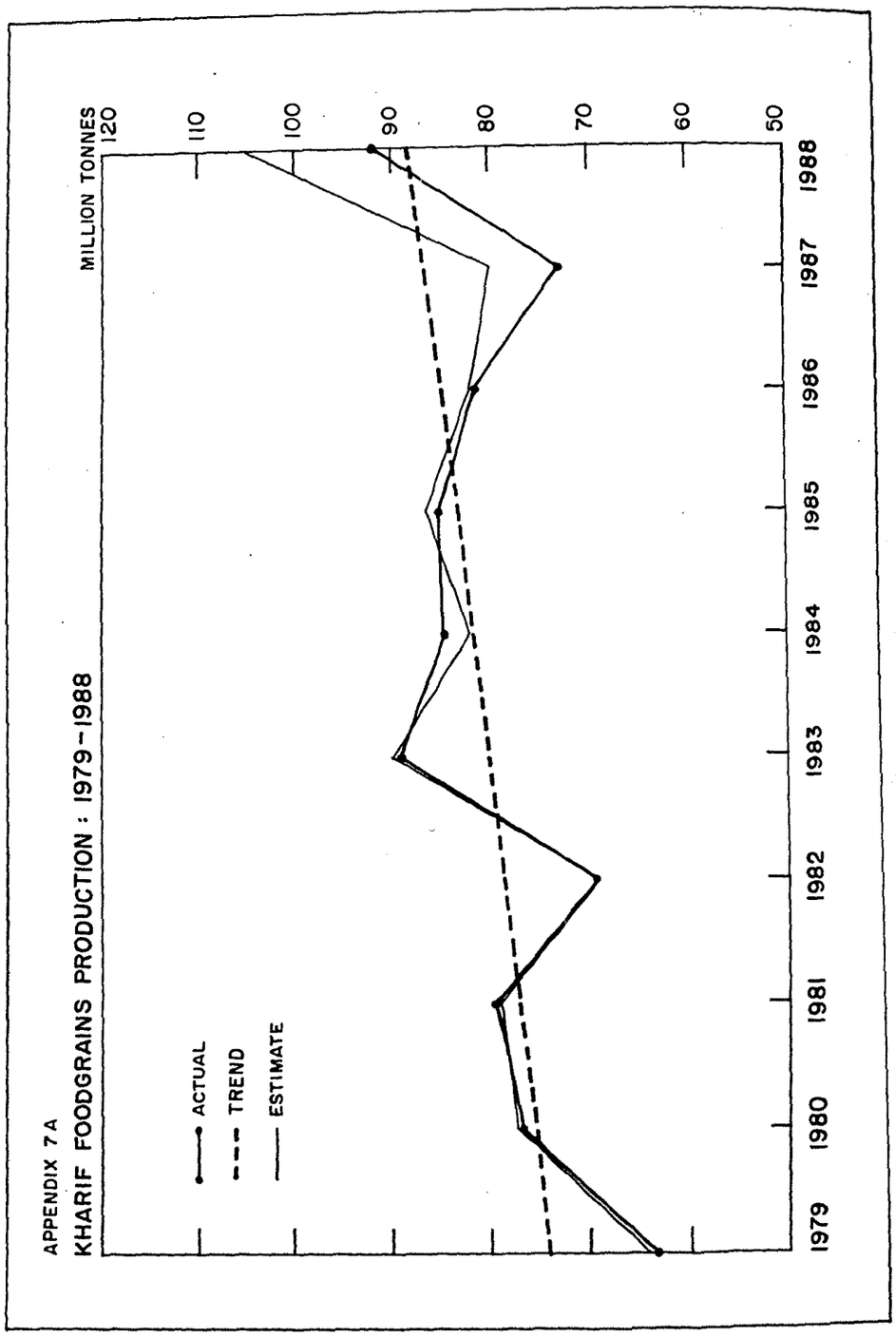
## Appendix 6: Alternative Regression Functions

| Sr.No. | Equation | $R^2$ | D.W. | SEE | Mean |
|---|---|---|---|---|---|
| 1. | $\ln Q = 6.60203 + 0.67407 \ln R^* + 0.02918\ t^*$ | 0.91 | 2.83 | 0.03 | 11.27 |
| 2. | $\ln Q = 7.77004 + 0.14446 \ln^* r_{sep.} + 0.16258 \ln r^*_{Aug.} + 0.31853 \ln r^*_{July} + 0.02567\ t^*$ | 0.96 | 2.83 | 0.02 | 11.27 |
| 3. | $\ln (Q_{t+1}/Q_t) = 0.02339 + 0.67421 \ln (R_{t+1}/R_t)^*$ | 0.83 | 2.94 | 0.06 | 0.04 |
| 4.1 | $\ln Q = 11.01644 + 0.01957\ t^*$ | 0.55 | 2.70 | 0.09 | 11.19 |
| 4.2 | $\ln (Q - \hat{Q}) = -64.54348 + 10.89052 \ln R^*$ | 0.65 | 0.84 | 0.89 | 8.73 |
| 5. | $\ln Q = 6.87154 + 0.63433 \ln \hat{R}^* + 0.02846\ t^*$ | 0.67 | 3.13 | 0.07 | 11.27 |
| 6. | $\ln Q = -16.41064 - 0.50817 (\ln R)^2 + 7.51557 \ln R + 0.02736\ t^*$ | 0.89 | 2.83 | 0.04 | 11.27 |

* Indicates significance 5% level
R = Cumulative MRI
r = Month's MRI
Q = Quantum of production
t = Time

**Appendix 7: Estimates of Annual Kharif Production**

| Year | Actual | Trend | Alt. 1 | Alt. 2 | Alt. 3 | Alt. 5 | Alt. 6 | Alt. 7 |
|---|---|---|---|---|---|---|---|---|
| 1979 | 63.25 | 74.03 | 62.39 | 64.01 | — | 62.76 | 62.03 | 64.63 |
| 1980 | 76.65 | 75.49 | 79.87 | 77.36 | 80.51 | 80.45 | 79.99 | 76.47 |
| 1981 | 79.49 | 76.98 | 76.74 | 78.94 | 74.16 | 75.69 | 77.37 | 76.24 |
| 1982 | 69.38 | 78.50 | 72.53 | 69.15 | 74.59 | 76.41 | 72.70 | 72.27 |
| 1983 | 89.10 | 80.05 | 89.01 | 89.93 | 84.79 | 81.63 | 88.24 | 86.99 |
| 1984 | 84.52 | 81.64 | 82.58 | 82.14 | 82.18 | 82.37 | 82.85 | 83.91 |
| 1985 | 85.24 | 83.25 | 83.04 | 86.59 | 84.50 | 86.50 | 83.14 | 83.75 |
| 1986 | 81.53 | 84.90 | 83.74 | 82.01 | 85.46 | 83.49 | 83.58 | 83.41 |
| 1987 | 73.0 | 86.57 | 81.85 | 79.91 | 79.51 | 83.65 | 80.93 | 81.54 |
| 1988 | 92.5 | 88.28 | 103.79 | 105.14 | 92.04 | 102.91 | 101.75 | 96.16 |
| RMSE | | 8.81 | 2.73 | 1.31 | 4.55 | 5.37 | 2.91 | 2.59 |


APPENDIX 7 A
KHARIF FOODGRAINS PRODUCTION : 1979-1988
MILLION TONNES
120
110
100
90
80
70
60
50
ACTUAL
TREND
ESTIMATE
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988

## Appendix 8

### References


1. Johnson J.: Econometric Methods: Mc Graw - Hill (1972).
2. Sastry V.K.: "The Use of Macro-Economic Regression Models of Developing Countries - A Comment": Oxford Economic Paper, Vol. 27 (1975).
3. Theil H.: Principles of Econometrics: Wiley (1970).
4. Thompson L.M.: "Multiple Regression Techniques in the Evaluation of Weather and Technology in Crop Production", Weather and Our Food Supply CAED Report 10: Iowa State University (1969).
5. Abraham T.P.: "Isolation of Effects of Weather on Productivity Including Other Risks such as Damage by Pests and Diseases", Indian Society of Agricultural Statistics, Vol. 2 (1965).
6 Heady E.O.: (ed), Economic Models and Quantitative Methods and Planning in Agriculture: Iowa State University (1971).
7. Herdt R.W.: "A Disaggregate Approach to Aggregate Supply", American Journal of Agricultural Economics, vol. 52 (1979).
8. —: "The Impact of Rainfall and Irrigation of Crops in Punjab", Vol. 27 (1972).
9. Lele U.J. and Mellor J.W.: "Estimates of Changes and Causes of Change in Foodgrains Production, India": Cornell University, New York.
10. Oury B.: "Allowing for Weather in Crop Production Model Building", Journal of Farm Economics, vol. 47 (1965).
11. Bapna S.L.: Aggregate Supply Response of Crops in a Developing Region (1980).
12. Stallings J.L.: "Weather Indexes", Journal of Farm Economics, vol. 42 (1960).
13. ICRISAT "Rainfall Probability Estimates for Selected Locations of Semi-Arid India, Research Report No. 1 (1978).
14. Government of India, Ministry of Food & Agriculture: Various Publications and Reports.
15. Weather Reports, Indian Meteorological Department.
16. Government of India, Economic Survey 1987-88 (Regional Rainfall Indices).

*Reserve Bank of India Occasional Papers*
Vol. 9, No. 4, December 1988

# Character Conversion from EBCDIC to ASCII and Vice Versa

**M.M. Khan** *

THE method of coding, adopted in the normal course, on the system Honeywell Bull 64/60, which is installed in DESACS, is EBCDIC (Extended Binary Coded Decimal Interchange Code). That is, all the programs as well as data are maintained in EBCDIC.

2. Need for conversion of data/programs from EBCDIC to ASCII (American National Standard Code for Information Interchange) and vice versa arises whenever data or programs are required to be transported from/to our system to/from other systems or to/from Personal Computers (PCs). Most of such systems are using ASCII form; hence, we need to convert the data/programs to and from ASCII. Also, outside users of our data-research *and* Government and other institutions—come up with requests for data *via* tape files, and in ASCII form. On the other hand, we receive data from them in ASCII form *via* tape files. Investigations were, therefore, carried out for getting to know all facilities available with our system, for conversion from one code to another and also to develop an independent technique for such conversion.

3. In other words, the probe was simultaneously conducted on two fronts. On one hand, the system's utilities and program facilities were tested for ASCII conversion, on the basis of whatever has been stated in the respective manuals. On the other hand, efforts were put in, to evolve a program for handling the conversion. The findings of these investigations are quite interesting and hence, are recorded in this paper, with full particulars of each stage of the experiments.

---

Shri M.M. Khan is an Assistant Adviser in the Systems and Programming Division of Department of Statistical Analysis and Computer Services (DESACS).

4. The system utilities allow certain special features to be declared by DEFINE parameters. According to the manual, there is a provision to declare the DATACODE as one of the following, BCD, H200, ASCII and EBCDIC; the last one is the default value as explained earlier. The ASCII option was tried for copying a file, by CREATE utility. For the said job, the translation was successful, i.e., JTRA accepted the declaration; the job also got completed; but the desired results were not obtained. In the course of different variations, which.were tried out, it was observed that whatever the configuration declared for input and output, the data were being copied in the same configuration as that of the input. In other words, the conversion from EBCDIC to ASCII or *vice versa*, was *not* effected even when opted for. Similar behaviour was noticed in case of PRINT utility also.

5. The SORT utility had a different story. The declaration DATACODE was disallowed by JTRA, hence, the job got aborted at translation stage itself. The only manipulation allowed in connection with ASCII, under SORT is COLLATE=ASCII to be declared in FUNCTION paragraph of COMFILE. This causes the file to be sorted as per the collating sequence followed by ASCII. Incidentally, the difference in the ASCII and EBCDIC collating sequences, is that in ASCII it is space, 0 –9, A-Z *and* in EBCDIC it is space, A-Z, 0 –9.

Here, when we say COLLATE=ASCII, only the arrangement of records, that is their sequence on the file, gets changed; but the data configuration of the records continues to be as it has been. This facility, thus, did not serve the purpose.

6. The alternative thought of, was, handling the character conversion from EBCDIC to ASCII in a specific program to be developed for the purpose. If the correspondence between the whole sets of EBCDIC and ASCII could be stored in a program, in the form of a table, it would be possible to search each EBCDIC character of input file in the table (using binary search, which is superior to linear search), find the corresponding value therefor in ASCII, *and* put the latter in the place of the former to write the output file. This exercise would transform each character of the data from EBCDIC to ASCII.

7. The correspondence between the EBCDIC and ASCII combinations is available in the manuals. The main problem was of storing the same in the program. This was done by using the COBOL numbers. Before proceeding with the details of the conversion processes, it might be worthwhile, at this stage, to discuss the concept of the COBOL numbers.

8. Each byte (formed of 8 bits) stored in the computer memory may be viewed as a pair of two sets of four bits each. Each 4-bit set corresponds to a hexadecimal number between 0 to F (Hexadecimal number system is an extension of decimal system, having the radix (base) of *16*. It is represented by numbers 0 to 9 and letters A to F, where A to F stand for numbers 10 to 15, respectively). The whole byte, thus, can have a value within a specific range of 256 (16 x 16) values as expressed in the matrix below-

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | A | B | C | D | E | F |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 00 | 01 | . | . | . | . | . | . | . | . | . | . | . | . | 0E | 0F |
| 1 | 10 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | 1F |
| 2 | 20 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | 2F |
| 3 | 30 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | 3F |
| 4 | 40 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | 4F |
| 5 | 50 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | 5F |
| 6 | 60 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | 6F |
| 7 | 70 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | 7F |
| 8 | 80 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | 8F |
| 9 | 90 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | 9F |
| A | A0 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | Af |
| B | B0 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | BF |
| C | C0 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | CF |
| D | D0 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | DF |
| E | E0 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | EF |
| F | F0 | F1 | . | . | . | . | . | . | . | . | . | . | . | . | FE | FF |

In other words, it should be possible to trace each and every byte in the memory, to this matrix. This is a known factor.

9. All these 256 combinations do not have corresponding graphic character presentations. That is, some of these, cannot be coded/punched/printed/displayed. This fact would pose a problem, when such combinations need to be stored/referred to in the program/s, as in case of conversion program mentioned in para 6 above. The con-

cept of COBOL numbers, introduced in the system, seeks to overcome this difficulty. They are numbers associated with each position on the matrix explained in the previous para. These numbers represent the requisite byte's occurrence number on the matrix, which can be considered as a string of 256 combinations. The Hexa combinations start from "00", whereas the COBOL numbers start from ""001"". Thus, the value 'COBOL number *less 1*' will be the decimal value of the Hexa combination corresponding to the said COBOL number. This COBOL number has been harnessed to link the EBCDIC characters to the corresponding ASCII characters.

10. The relation between the Hexa combination and the COBOL numbers, is a factor independent of the coding methods (say, EBCDIC or ASCII). What changes, is, the corresponding position of the characters along the string of 256 characters. In other words, the COBOL number, which occur in natural order under EBCDIC code, would occur in a changed order under ASCII coding method. To explain by an example-space is represented in EBCDIC by Hexa combination "40" occurring at *65th* position in the 256-character string and has the COBOL number ""065"" associated with it. In ASCII, space is expressed by Hexa combination "20" which would occur at 33rd position (as per EBCDIC collating sequence) in the 256-character string and the associated COBOL number is ""033"". Thus, for replacing an EBCDIC space by an ASCII space one needs to pick up the COBOL number ""065"" and replace it by COBOL number ""033""

11. Having said this much about the COBOL numbers, the discussion of conversion processes mentioned in para 6 above, may be resumed.

12. At first, the simple method of conversion as explained in para 6 above, was tried out in a program named as R2. This program was run on a data file to process first *4,721* records, where the record size was 2,292 characters. This job involving conversion of *1,08,20,532* characters (4,721 x 2,292) took *98.223 minutes* of CPU.

13. Despite the fact that efficiency techniques like use of binary search in preference to linear search *and* same record area for input and output files, had been applied in the program, the CPU time

taken was enormous. Obviously, the search operation was taking considerable CPU time. In the second variation, it was sought to eliminate the search operation by devising a procedure for converting the EBCDIC characters into numbers so as to serve as subscriptors for accessing the ASCII table and picking up the corresponding combination, therefrom. A subscriptor has to be a non-zero positive integer with value lying within the range of 1 to 256, for accessing the 256- chracter string described earlier. The idea, therefore, was to try to express each Hexa combination of EBCDIC as a decimal number, satisfying the aforesaid criteria.

14. This is easier said than done. None of the books and manuals are giving any clue as to this type of conversion. Various different methods were, therefore, tried, as described hereunder.

15. As already-mentioned in para 8 above, each byte can be viewed as a pair of 4 bits each. It may also be seen from the table given in the same paragraph that each group of four bits can be represented by a Hexa number between 0 to F. If we try to move a byte (an alphanumeric character) to a Numeric or Computational (Comp) or Computational-3 (Comp-3) field, in the first place, the compiler would give a warning "Moving Non- numeric to Numeric". When the actual movement is attempted, the first group of four is ignored, and the next group only, is considered. The movement would result in numbers, as far as the group to be moved corresponds to Hexa numbers 0 to 9. But for A to F, this would give rise to illegal decimal data. Even if it were possible to convert all these Hexa numbers into decimal numbers, the resultant numbers obtained from 256 Hexa combinations would be between 0 to 9 only; for instance, all combinations 00 to F0 would, on conversion, give rise to 0, all combinations 01 to F1 would give rise to 1, and so on. In effect, no unique numbers can be had from the 256 combinations, and these repetitive numbers would defeat the idea of using them as subscriptors. This method was, therefore, ruled out.

16. Another possibility considered, was that of movement of each byte to a Computational-1 (Comp-1) field, because all the operations on Comp-1 field are binary operations. A Comp-1 character is formed of *16* bits, of which one bit goes to represent sign (0 stands for -ve and 1 for +ve). The remaining *15* bits represent binary numbers.

Thus, only 15 bits are available for expressing a number. The maximum number that can be represented by these 15 bits would be *32,767* ($2^{15}$ - 1). In other words, the maximum subscriptor value obtainable thereby, would be *32,767*. Our experiments, however, revealed that movement of the 8-bit alphanumeric character to Comp-1, does not automatically effect conversion in numbers. Therefore, this alternative was also ruled out.

17. At this stage, a very ingenious method of conversion was hit upon. The facility of redefining a data field, was employed for the purpose, in the following manner. A group was described in the WORKING-STORAGE SECTION of DATA DIVISION as given below.

```
03 C.
05 FILLER PIC X VALUE LOW-VALUE.
05 H PIC X.
03 J REDEFINES C COMP-1.
```

(LOW-VALUE is the lowest configuration, where all the 8 bits are 0)

The EBCDIC character is to be moved to H. The first 8 bits of C would always contain zeros by virtue of the LOW-VALUE stored there; and the next 8 bits would give us the binary digits ( 0 or 1) corresponding to the Hexa combination of the EBCDIC character moved there. When referred as J - a Comp-1-this field gives us the desired conversion into a decimal number. The first bit of J being 0, this would be a +ve number. It is then moved to the index associated with the ASCII table by SET. We set this index number up by 1 and get the subscriptor, which is nothing but the COBOL number associated with the EBCDIC character. This procedure, thus, proved to be the key to the problem of conversion. Once the subscriptor is obtained, the ASCII equivalent thereof, can be picked up from the table to replace the original EBCDIC character, which was moved to H.

18. This method was used in a program named as R3. It was run on the same file, as for R2, in the manner described in para 12 above. This process brought down the *CPU time* drastically to *19.865 minutes* from *98.223 minutes* of earlier.

19. It may be seen from the conversion process used in R3 that setting up of the index by 1 every time before using it as a subscriptor, is necessitated by the fact that the index value obtained from the EBCDIC character LOW-VALUE would be zero, whereas the subscriptor has to be minimum 1. Such setting up operation for each and every character, in itself, would take considerable CPU time.

20. This was sought to be reduced by checking the LOW- VALUE separately. If the EBCDIC character from the input is LOW- VALUE, the ASCII equivalent of the same, which is also LOW-VALUE, is moved to the output. The other characters would be subjected to conversion process and the index number obtained thereby, would be used as subscriptor without setting it up by 1. In order to make it possible to access the ASCII table directly, in this manner, the ASCII table is also truncated by eliminating the first element thereof. The said procedure was adopted in a program named as R4. When run on the same data file as earlier, it did reduce the *CPU time* further to *16.869 minutes*, from *19.865 minutes* taken by R3.

21. Parallel investigations were carried out in respect of facilities available in system's COBOL. In COBOL, under file description in FILE SECTION of DATA DIVISION, there is a provision to declare

```
                (NATIVE      )
                (EBCDIC      )
                (STANDARD-1)
                (ASCII       )
CODE-SET IS     (GBCD        )
                (HBCD        )
                (IBCD        )
                (JIS         )
                (alphabet name).
```

Since the code-set native to our system is EBCDIC, first two values signify one and the same code-set. Likewise, STANDARD-1 and ASCII both signify the ASCII code-set. GBCD is the code-set meant for accommodating systems (series 100/400/600) managed earlier by General Electrics Corporation (GEC), since the said corporation merged with the Honeywell Company to form the present Honeywell Information System (HIS). On the other hand, HBCD is suppor-

ting earlier Honeywell Systems (series 200/2000). IBCD is the IBM commercial code-set and JIS is the *J*apanese *I*ndustry *S*tandard charter set. Besides all these, there is one more option described as "alphabet name". This option can be used in an indirect manner. That is, declare the code-set to be some data name, say, SAM. Then under SPECIAL- NAMES paragraph in CONFIGURATION SECTION of ENVIRONMENT DIVISION declare SAM IS ASCII or any one of the options mentioned earlier. In this connection, it may be mentioned that the code conversion occurs during the execution of Input/Output Operations. The conversion process is applicable to sequential files only.

22. A program named as R1 was developed using this facility, and it was run on the same file as used earlier. And the CPU time came to be just *0.920 minutes.* Even for the whole file of *2,51,06,568* characters (10,954 records X 2,292 characters per record), it took only *2.077 minutes of CPU.* It means that system's algorithm for interplay of bits employed in their own package(s), works amazingly faster.

23. There is one important point to note, in the context of such conversion. If any data having Comp or Comp-3 type of storage or a signed numeric field having sign overpunched, are directly subjected to the conversion process, they would not give valid results. Such a contingency is not uncommon. For instance, the amounts in Company Finances(CF) data are kept in Comp type, for saving on storage space so also as an efficiency technique, as it makes for faster computation. The amount expressed in the format S9(11) Comp (as for CF) is stored in memory in 6 characters. The conversion process described earlier, would be converting these 6 characters into ASCII and it would certainly not represent the input amount any more. It, therefore, follows that before subjecting such a file to the conversion program, it would be necessary to first enlarge the aforesaid computational amounts into 12 characters, where the sign would also appear as a separate character.

24. It can, thus, be seen that with the aforesaid experimentations, fruitful results were achieved on both the fronts, described in para 3 above. The facilities offered by the system, were employed to get the conversion effected, on the one hand. And more significantly, independent technique was developed to handle such conversion even without using the system's facilities meant for conversion.

*Reserve Bank of India Occasional Papers*
Vol. 9, No. 4, December 1988

# Economic Theory and New World Order

(Eds) H.W. Singer, Neelambar Hatti and Rameshwar Tandon
(Ashish Publishing House, 1987; price Rs. 400, pp. 637)

THE book is a collection of twenty-seven different papers, comprising over six hundred pages. Together, these papers throw some very useful light on many of the crucial issues and facts concerning the establishment of New International Economic Order (NIEO).

The specific fields in which fundamental changes and reforms are called for by the NIEO programme relate, among other things, to trade (between North and South) resource transfer from North to South, stabilisation of primary commodity prices, regulation of transnational corporations' restrictive business practices, international monetary system, intra LDC cooperation, and so on.

As regards trade, exponents of free trade policy have a great deal to be disheartened with, when R.H. Green's paper reports that 50 per cent of world trade is under *de facto* quantitative restrictions. However, developing countries have a special grievance with such an overall trade regime in that their gain from the Tokyo Round of Multilateral Trade Negotiations has been extremely limited: tariff reductions on primary products of export interest to LDCs were far below average for all products; several tropical products, in which exports of developing countries compete with those of developed ones, did not figure in tariff reduction at all (pp. 318-319). The same discrimination (tariff reduction far below average) also bedevils manufactures of export interest to LDCs with the difference for the worse: reduction in *MFN rates* has eroded the *preferential margin* of *GSP* without any compensation in terms of enlarged product coverage or enhancement of ceiling under GSP (p. 319).

A circumstantial explanation of Northern protectionism towards the South is contained in the deteriorating employment situation brought out by A.G. Frank's paper: Volume of unemployment in OECD countries rose from 5 million in the 1967 recession to 10 million in that of 1969-70, again to 15 million during 1973-75; and even after post 1975 recovery lasting for four years, it stood at 17 million at the beginning of 1979 recession, reaching the figure of 32 million in the midst of this recession in 1982 (pp. 507-8), However, this can only be a matter-of-fact explanation of the North's restric-

tionist proclivities towards the South; it can by no means constitute a justification *in terms of economic rationale.* Indeed, Alfred Maizel's paper provides a telling rebuttal to the body of economic opinion that attributes Northern protectionism to North's prevailing economic sluggishness, when he points out that the long term arrangement for the limitation of low cost cotton textile exports of LDCs was instituted in 1960s (1962) when the developed countries experienced most rapid economic expansion (p. 297).

Even when developed countries do take up the issue of liberalising their imports from developing countries, there appears to be some built-in factor of impediment in their import regime towards such a process. Sunanda Sen's paper provides a typical illustration: tariff structures and levels differ considerably among developed countries, especially between U.S.A. and E.E.C. At the time of proposals for tariff reduction on LDC exports, U.S.A. and E.E.C. often disagree as to the precise formula to be adopted for tariff cuts; this stalemate between these two economic powers has the effect of delaying and often even stalling tariff cuts in favour of the LDCs. (pp. 258-9). The author confines the treatment to tariff alone. She could have brought out the point at issue even more vividly and potently, had she also considered non-tariff barriers where intra-Northern divergence is greater in degree and more complex in nature, especially those that hinder LDC's exports.

On the question of resource transfer to the South, the book throws some useful light on certain facts and issues concerning both official and private channels of the transfer.

Paul Streeten's paper discloses the deplorable aspect of the trend in Official Development Assistance (ODA): it fell from 0.42 per cent of the GNP during 1964-66 to 0.35 per cent during 1978 against the target of 0.7 per cent accepted by developed countries themselves at the UN (p. 20). In the light of such a poor showing of ODA, Willy Brandt's observations are very pertinent: in the context of socialist countries' disclaimer of any obligation to extend more than a modest measure of aid to LDCs on the plea that LDC's problems are the product of colonial exploitation of the past, he derisively observes that LDCs need help not in interpreting history, but, in shaping their future, and they need this help from both the West and East (p. 337). In further proof of the practical irrelevance of the issue of past colonialism, he points out that countries such as Sweden and Norway

are already commiting themselves to Official Development Assistance well in excess of 0.7 per cent of GDP, instead of emphasizing their non-colonial past. In this connection, he also makes a very thought provoking suggestion regarding ODA in that it should be measured against consumption rather than GNP per se (p. 337). This would mean that a rich country with a higher ratio of national consumption to national income should be called upon to provide higher ratio of its GNP in ODA than another rich country, say with the same GNP, but with smaller ratio of national consumption. This prescription makes good economic sense: a rich country with a higher ratio of national saving will perforce tend to contribute more (directly or indirectly) to South's development, since it will be in a position to do so either through private investment abroad of increased national savings or through balance of payments surplus which is generally invested abroad through official channels - a country with a lower rate of national consumption tends to generate surplus in its current account, other things being equal.

As regards resource transfer to South through private channels, especially in the form of direct investment, Paul Streeten makes a very instructive point in his paper through his allusion to a dilemma that confronts both the host developing countries and the investing firms from the developed countries. Fear of expropriation has raised the required rate of return to induce the Northern firms to invest in the developing countries. At the same time, actual realisation of high rate of return provokes expropriation on the ground that the companies are taking out more than what they put in. The dilemma for the developing countries is that repatriation entails a drain on foreign exchange, at the same time nonrepatriated profits lead to reinvestments which put increasing portion of the national stock of capital under foreign ownership. He advocates exploration of well designed measures to reduce uncertainty, as this would reduce both the rate of required return for the investing firms and the incentive for the host governments to expropriate, thus benefiting both sides. However, this resolves only the dilemma facing the investing firms. The dilemma for the host government in the form of trade-off between foreign exchange loss and alienation of national assets remains. One way to resolve this, which Streeten fails to bring out in his otherwise rigorous discussion of the matter, might be for the host government to stipulate reinvestment to be in the form of portfolio like bonds

and debentures either in the same foreign owned enterprise or in some other enterprises, since portfolio investment does not confer ownership on the investor. In the case of foreign investment being in export oriented activities, the outflow of foreign exchange via repatriation may be offset by inflow of foreign exchange via export proceeds; in this case, the host government can look upon repatriation with relative equanimity - again a point not touched upon by Streeten. Of course, in any case, Streeten's observations serve to underscrore the validity of South's *predilection for Official transfers* (as against private ones), especially from multilateral institutions like IMF and IBRD.

Indeed, the book has a wholesome message for multilateral credit institutions too, in the paper by Sukhamoy Chakravarty. The author's report (with specific statistics) of accelerating inflation in the North even in a setting of sustained deceleration in its economic growth during 1970's has a policy implication for IMF & IBRD in particular. Since for most developing countries (especially those on the lower rung of the GDP scale), imports from developed countries exceed their exports to them, the Northern inflation is bound to spill over into Southern countries via the latter's import prices. What is particularly debilitating for South's economies in this sequence is that this North-transmitted inflation is of cost push variety. If it were of a demand full variety, it may at least have the redeeming feature for developing countries in the form of increased export earnings with the concomitant propitious effect on economic growth and development. But, in reality, being cost push in its modus operandi, this spill-over merely raises South's general price level above what it would otherwise be; its effect on economic growth is in fact negative, as it severely curtails the availability of vitally needed imports. The IMF would do well to take full account of this North-transmitted aspect of South's inflation before blaming South's balance of payments problems on domestic inflation and then enjoining on the developing countries, strict austerity measures as a condition for its balance of payments assistance.

As regards stabilisation of primary commodity prices, Alfred Maizel brings into bold relief the great potential benefits which North also stands to reap from a regime of stable world prices for primary commodities, since developed countries too lose considerably from volatile primary commodity markets, in at least three respects. Firstly,

sudden jumps in commodity prices entail inflation in DC's economies and the attendant loss of real output and employment through disinflationary macro-economic policies. Secondly, uncertainty of supply, coupled with price rise, compels enterprises in the developed countries to hold much greater volume of inventories of certain primary commodities than what is otherwise necessary, with resultant waste of capital for the enterprises concerned. Thirdly, price uncertainties in the international market discourages long-run investment in capacity and in productivity improvement in the exporting LDCs, with consequent reduction in *long-run* supply of the primary commodities below what would otherwise be possible. Hence developed countries are better advised to espouse the Integrated Programme for Commodities (conceived at the UNCTAD IV) in their own interest. However, Kiyoshi Kojima misconceives the whole point when he sees in the Integrated Programme for Commodities (IPC) the elements of cartels similar to that of OPEC (p. 199). Cartelisation is the essential characteristic of a system of export quotas; this system operates through concerted restriction of supplies by the members of the scheme, the aim (or at least the effect) being extraction of maximum possible price, rather than stability per se. However, IPC seeks to stabilise world primary commodity markets through a system of buffer stocks for each of the eighteen commodities. A system of buffer stock operates through buying at pre-determined floor price, the commodities that are in excess supply, and releasing the same from stock during periods of excess demand as soon as the prices approach a pre-determined ceiling. It needs to be emphasised here that buffer stock, by preventing both *unnatural* rise and *unnatural* fall in the market price of the commodity concerned, tends to stabilise (not freeze) the marekt *along a dynamic path* by ensuring steady augmentation in *long run* supply, in line with *long run* demand; by the very logic of the maximum involved, a well formulated buffer stock scheme would have provisions for periodic revisions of both floor and ceiling prices so as to allow for both temporary and secular changes in supply, demand and cost conditions surrounding the commodity concerned.

Manmohan Singh's paper enlightens the debate on the commodity question by placing the inadequacy of the Common Fund (agreed to be set up in 1980 with a capital far below what was arrived at in UNCTAD IV) in the perspective of market structure, in so

far as there is *oligopsony* in the world markets for primary commodities (p. 318). The policy implication is that the inadequacy of the Common Fund is rendered particularly conspicuous since a *buyers' market* (which oligopsony constitutes) involves more instances of price fall than those of price rise, so that a fund that would be sufficient to support the floor price in such a market has to be necessarily larger than that in a *market of competitive buying.*

As regards regulation of transnational corporations, Oleg Bogomolov startles the reader by his disclosure that they control over one third of world trade, and over 70% of raw materials and major part of manufactures exported by developing countries (p. 573). He even goes on to argue that the nonrenewable natural resources of the world are now being threatened with depletion and exhaustion not so much by genuine needs of mankind as by the rapacity and cupidity of the TNCs.

The sway of the TNCs has already penetrated deep into the world financial system as well. Osvaldo Sunkel's paper provides a typical instance: 34 large transnational banks control a vast amount of international fianance, obtaining considerable portion of their funds from hundreds of small and medium sized banks through the system of syndicated loans (p.95). This has had the undesirable effect of channeling most of the private capital flow to the better-off semi-industrialized countries of Latin America, since the poorer countries did not constitute an attractive outlet for *commercial* loans.

In the context of the fact that Northerners often spurn NIEO proposals on the score of free trade maxims, Manmohan Singh in his paper highlights the widespread exercise of considerable market power by TNCs in world trade (especially in manufactures and processed primary products), and then exposes the moral and economic bankruptcy of such a posture on the part of the North (p. 323).

Gerald Helleiner in his paper, while pointing out that market concentration (both oligopoly and oligopsony) vitiates the assumptions behind free trade policy conclusions, provides the interesting (but little recognised in literature) information regarding the inclusion in the Havana Charter of ITO of the provisions for regulating restrictive business practices, whereas there is no such provision in the present GATT framework (p. 68).

It might also be of some interest to note Paul Streeten's exposure of North's unmistakable guilt of double standards in that USA pro-

scribes cartels in its domestic operations, but specifically exempts export cartels from any legal regulation — with repercussions on developing countries who often have to pay higher prices for their imports than those charged to Northern buyers of the same goods (p. 45).

On the question of international monetary reform, Alfred Maizel's advocacy of progressive reduction in the reserve function of the dollar and enthronement of SDR as the Central reserve asset of the world monetary system as a prerequisite for preserving the *real* value of South's international reserves and even for orderly increase in international liquidity (pp. 302-3), is of some analytical significance. As is common knowledge by now, US economy has been showing a tendency to periodic bouts of internal inflation ever since mid-1960's. This entails commensurate reduction in the *import-purchasing power* of developing countries' international reserves, in so far as dollar is the form of such reserves. Similarly, any substantial *net* increase in international liquidity has been dependent on creation of appropriate deficits on the US current account, to the extent dollar is the major component of international reserves. Such a mechanism of reserve creation confers an *unfair* advantage on the USA (or any reserve-currency country for that matter), since USA is in a position to draw upon the *real resources* of the rest of the world in the form of imported goods and services without giving in exchange *corresponding volume* of *its own resources*. Progressive reduction in the role of the dollar and simultaneous corresponding increase in that of SDR in the world monetary system would go a long way towards obviating both the aforesaid inequities in the system.

The book contains a great deal to demonstrate the existence of considerable harmony of interest between North and South, which by itself should be a sufficient inducement for the North to set in motion in right earnest the implementation of the NIEO programme—at least in its essentials if not in toto.

H.W. Singer points out (p. 7) in his paper that developed countries alone lose about $ 400 billion a year due to the growth-inhibiting effect of stagflation (which he quite rightly regards as the Northern manifestation of the 'disorder' in the contemporary world economy). Hence , the cost for developed countries of implementing many of the NIEO proposals, such as provision of 0.7 per cent of GNP in ODA and financing of the common fund for primary commodities,

would be ridiculously small in comparison with the saving of this $ 400 billion loss, since the NIEO measures would raise LDC incomes, which by augmenting world demand for the goods of developed countries, would add to the growth stimulus for the latter.

Paul Streeten brings out the harmony of interests between North and South by highlighting the commercial interdependence between the two. In this connection, he argues that importance (and hence indispensability) of commercial links between nations does not consist in the volume of trade per se but in the utility of goods imported, together with the difficulty of their domestic substitution in production on the one hand and in the domestic super flowity of the goods exported, on the other. Thus, much of the intra Northern trade consists of exchange of similar consumer goods catering for slight differentiation in taste, while the trade of North with South, though relatively much smaller in volume, is based on exchange of the relatively surplus manufactures of the North with vital foods and raw materials of the South (p. 35). Oleg Bogomolov provides a very instructive factual illustration of the same theme by showing how the North depended on the South for the supply of several vital raw materials. Even a country so richly endowed with natural resources as U.S.A. depended in mid-70s on imports for 68 out of 95 raw materials considered essential for a modern industrial economy, with complete dependence in the case of 15 of them (pp. 567-8).

Regarding Northern liberalisation of manufactured imports from the South, Alfred Maizel makes a very novel point by alluding to the fact that the Northern industries, into which such liberalisation would divert present as well as future productive resources, are of a kind in which the developed countries not only have a *natural* comparative advantage, but also in which their productivity growth is *above the national average* (pp, 296-7). This implies a double gain for the developed countries: shifting resources from activities with below average productivity (such as those that are now shielded from LDC competition) to those with average one, would itself be a net economic gain; shifting them to those with above average productivity would be doubly so.

Surendra J. Patel's observation that the stimulus provided to Northern economies by foreign trade in the fifties and sixties through intra Northern liberalisations, especially via "customs union" and "free trade areas", has run out of steam, implies the same harmony

of interest in that North should now increasingly look for expansionary stimulus from import demands of the South. Since South's import demands at present are severely circumscribed by their debt burden, IMF conditionality and above all by their insufficient export opportunities, NIEO programme, which has several built-in features to augment South's international purchasing power in a durable fashion, is something that North should espouse in its own economic interest.

The book advances the cause of informed debate on the third world's problems by focusing on the heterogenous nature of third world's economic conjunctures, and hence a New International Economic Order, if it is to be beneficial on a global scale, cannot be a uniform package of measures by all developed countries alike in favour of all developing countries alike. What the least developed countries need most is financial and technical aid in liberal measure while what the relatively advanced developing countries need is free access to market for their manufactured exports, to international capital market and to modern technology. The author (Paul Streeten) makes still further minute analysis of this aspect of heterogeneity: a country like Saudi Arabia should contribute to loans because of its high foreign exchange earnings, to aid because of its high per capita income, but should receive preferential treatment in trade because of its low level of industrialisation, and so on. This focus on the scenario of heterogeneity is very significant for the success of NIEO negotiations. Developed countries do reckon (even if wrongly) with the budgetary implications while giving aid, and with the displacing of output and employment at home while giving trade preferences. Hence, they will be better disposed to give aid to a country with low per capita income than to one with high per capita income. Similarly, they will be less averse to giving trade preference to the exports of a country with better industrial sophistication than to the country of opposite type, since in the former case, the industrial cost structures would be closer to the international ones and would hence involve smaller margin of preference. Indeed, at a meticulously practical level of consideration, a whole range of varying degrees and combinations of trade, aid, financial and technical measures suggest themselves that could be geared to developing countries at varying levels of *absolute* poverty, per capita income and industrial development and/or diversification. Asking for all sorts of things for all sorts

of Southern countries may result in none of them being attained for any of them, as experience to date sufficiently demonstrates.

Finally the book has a very persuasive counsel for Southern countries to redress their domestic economic disparity. The fundamental basis of the South's demand for NIEO is the egalitarian consideration which calls for a drastic reduction in economic inequality among national societies, even if it involves contravention of market principles and of individualism in economic decision making. But Northern negotiators found in this egalitarian percept—often overtly emphasised by Southerners and their sympathisers in the North— a convenient handle to turn against the South: from mid-1970s they began to underscore the gross inequality in income and wealth within developing countries, and increasingly took the position that without redressing internal disparity, the South had no moral case for demanding parity at the international level (see especially pp. 255-6). Southerners may legitimately see in this Northern position a devious design to find convenient alibi for evading their share of responsibility and obligation. However, the point at issue here is that the internal disparity of the South gives the North a scoring point in the debate, and more importantly, such an approach on the part of the North constitutes an obstacle to the implementation of NIEO proposals. Thus, even if the policy makers in the Southern countries choose to disregard the demands of moral imperative and of logical consistency involved in the whole issue, at least the sheer pragmatic consideration of the need to retain initiative and ascendency in negotiating exercises with the North should urge them to do everything possible to redress internal economic disparity.

**Y. Kalyanraman***

---

* Dr. Y.Kalyanraman is a Research Officer in the Special Studies Unit of the Department of Economic Analysis and Policy (DEAP).

*Reserve Bank of India Occasional Papers*
Vol. 9, No. 4, December 1988

# Quality Programming-Developing and Testing Software with Statistical Quality Control

by-Chin-Kuei Cho, Ph.D., Computa, Inc. (John Wiley & Sons, Inc. New York, 1987; price $43.95, pp.473)

STATISTICAL Quality Control, introduced in the United States during World War II for procurement of quality military hardware, has been a proven, powerful, and widely used tool in manufacturing industries to ensure product quality for over 40 years in the world over. The Japanese have been, and now Koreans are, using this technology so effectively that their products dominate the international competitive market. In some of the manufacturing products such as aircrafts, where even a minor defect of a component could cost human lives, quality control is mandatory, while in some other cases, it is necessary so that a company maintains its competitiveness.

In the present world, increasing computer-orientation of the multidisciplinary activities makes it necessary to think about the quality of a computer itself, hardware as well as software. Being a manufacturing item, the quality of computer hardware may be controlled by statistical approach and, therefore, warranties have always been offered by the manufacturing industries. But, the scenario is different in software industry. There has never been a statistical approach to control the quality of a piece of software and, hence, no warranties were being offered till the time Chin-Kuei Cho's book "An Introduction to Software Quality Control", was published in 1980. The book on hand for review is, in fact, second book by the author in which he refines his earlier techniques, reports many actual applications, and introduces a number of new techniques. The most appealing feature of the book under review is that unlike the other books which mainly come out of an author's desk exercise, this publication is the essence of the author's practical experiences gained through conducting more than 50 in-house training sessions and short-term seminars, courses, presentations, and tutorials for thousands of professionals throughout the world since 1980. Work on this fascinating area of

quality programming is scattered throughout the computer literature and the author has taken support of about 200 references to put forth his approach. The book contains 15 chapters and 7 appendices of statistical tables, useful for software developers in applying various suggested procedures for quality software.

A recent survey conducted by Thayer, revealed that there are 20 problems existing in the software industry. These problems are in the areas of: requirements, success, project, cost, schedule, design, test, maintainability, warranty, control, organizing type, accountability, project management, viability, reliability, goodness, programming, *and* tracing in software engineering and development. Although, modern software technologies such as structured programming, structured analysis, structured design, and structured testing are being practised, these problems persist. What could then, be the root of these problems and how could these problems be solved? The questions like these are answered by the author by introducing a new software methodology called "Quality Programming", incorporating statistical quality control for development of high quality and cost effective software, in greater detail. Emphases are placed on deployment of statistical quality approach during every stage of software life cycle from modelling, requirements, specifications, concurrent software design and test design, test and integration, to software acceptance that leads to the offering of meaningful software warranty.

In the very first chapter, the author has provided a back- ground for clear understanding of the subsequent contents. For applying statistical quality control technique, the author has examined whether software development was comparable with a manufacturing industry. Quoting from Webster's New World dictionary, the author has asserted that the word "manufacturing" is defined as working into usable form and, thus, software development is one of the manufacturing industries that works some raw material into usable form. The raw material in software development is the data input to a piece of software. The usable material is the data output by the program and the software itself is the factory that works the raw data into usable form. Here, the author's approach of viewing the software development as a manufacturing activity is quite innovative. The current practice in software industry, of considering the software as a final product and treating an error-free software as the quality software, is not recommended by the author. Further, in the same chapter,

the author briefly states how statistical techniques can be adopted during various phases of software development cycle.

The book is divided into three major parts: Statistical background for quality control (Chapters 2 through 6), good quality software through prevention (Chapters 7 through 10), and good quality software through promotion (Chapters 11 through 15).

Chapter 2 is concerned with basic probability concepts required for understanding statistical distributions. Major topics include repetitive operations, sample space, events, probability, random variable, randomization and random numbers.

Chapter 3 deals with important statistical distributions essential to the derivation of sampling plans for estimating software defectiveness and for software acceptance. The discussion is focused on the hypergeometric, binomial, poisson, and normal distributions with their inter-relationships. The normal distribution can be used as a convenient vehicle for statistical inference while poisson distribution is used for formulating acceptance sampling plans.

Chapter 4 examines methods of generating the random numbers by a computer which plays an important role in statistics and software quality control. Out of various methods available, viz., the middle square method, the multiplication method, and the congruence (or power residue) method, the author has focused his attention on the power residue method, which is commonly used. FORTRAN and COBOL generators are given as examples. For testing the randomness of the generated numbers, frequency, serial, poker, and gap methods are discussed in brief. Yet, the given illustrations facilitate practical understanding of the methods.

Chapter 5 describes two sampling techniques suitable for software testing and acceptance, simple random sampling and sequential random sampling. Statistical inference principles including estimation and testing the hypotheses relating to the defective rate of a product-unit-population and determination of the sample size, are discussed in detail.

Chapter 6 broadly deals with the formulation of product-unit-population acceptance sampling plans. There are many types of acceptance sampling such as single sampling, double sampling, multiple sampling, and sequential sampling. Each has advantages and disadvantages. The final choice of method may depend on the agreement between the producer and user of the product. For discussion

purpose, the author has selected only single sampling and acceptance sampling. Numerical examples make it easier to follow what are exactly the producer's risk, user's risk, and O.C. curve.

The modelling activities, which are required in order to apply statistical quality control (as discussed in Chapters 2 to 6), so as to enable the software user to demand a warranty and the software developer to deliver one, are discussed in Chapter 7. Modelling is the activity of building a model which is a representation of an existing or a conceptual object, an abstraction of a real world phenomenon that will be the basis for development of a piece of software. Modelling often includes devising solution for the problem under consideration, hence, the author has studied all aspects of input, processing and output. After discussing the modelling in manufacturing industry, the author points out its resemblance with modelling in software industry Explaining the product concept formulation, the author regrets that software industry has conventionally treated the piece of software itself as the end product for the user. The entire modelling cycle is illustrated by examples suitable for various phases. At the end of the chapter, the author has brought out the current situation in software industry and the critique offered by him is very useful for software developers.

The material presented in Chapter 8 pertains to the software requirement specification, which is supposed to serve as a framework of requirement document for a piece of software, during various developmental stages such as incorporation of principles and goals of software engineering, development, testing and acceptance which can be used for communicating with the software developer, designer, test designer, user and all other concerned parties. A new concept of SIAD (Symbolic Input Attribute Decomposition) tree to represent input domain of a piece of software for systematic generation of software test input units is given in this chapter. Four types of SIAD trees, viz., regular, weighted, ruled and network, have been developed by the author for the sake of handling different types of software applications. Five test methods, viz., regular, weighted, boundary, invalid, and special tests, available for assessing the software quality have been discussed at this point. It is stressed that the user and developer should, as a final step in developing system test requirements, identify any special feature that could be included in the tests.

Chapter 9 recommends concurrent development of software design and test design on the basis of requirement documents discussed earlier in Chapter 8. Software design methodologies, viz., function-oriented and object-oriented, are critically examined. Function-oriented design is associated with conventional programming languages such as FORTRAN and COBOL, while object-oriented design is appropriate for software written in newly developed 'Ada' language. Since both the methodologies are having their limitations, the author has not, categorically, favoured either of them. In general, there is no evidence that a piece of software developed in 'Ada' using object-oriented design is more reliable than that developed by other methods, though 'Ada' appears to be a more efficient language for software understandability. Thus, there is scope for carrying out further research in 'Ada' for improved quality. The software test design discussed in this chapter is based on the application of statistical quality control using four types of SIAD trees for the five test methods discussed earlier in Chapter 8. If the software design and test design cannot be developed concurrently, it means that the modelling and requirement specification documents have not been adequately prepared. And to avoid costly consequences, these may be re-examined at this stage.

Chapter 10 recommends concurrent implementation of software design and test design. Human factors, software portability, programming support environment, selection of programming language, and strategies for software implementation are discussed here. In test design implementation other than error sources and implementation strategies, the random number generators are also discussed. In accordance with the test design, test input units can be constructed efficiently as software modules and the system is ready to be tested. Cross- verification of the implementation serves as a crucial quality check.

Chapter 11 abounds with discussions of software testing integration, independent verification and validation, debugging *and* comparison of many softwares. All activities are guided by the principles of statistical quality control, discussed earlier and supported by many illustrations to make the procedure understandable. Software success criterion recommended is the defective rate of software product-unit- population; the traditional goal of software engineering being subjective in nature, is not suitable to give exact quality level of a

piece of software. Further, sampling processes discussed in Chapter 6, software module, system test procedures, and debugging effectiveness are also demonstrated by using the defective rate.

Once the software development, testing, integration, verification and debugging phases are over, software delivery takes place. Poor quality software that may result in damage to the user and possible law suits, is too costly for any business to accept and has to be avoided. Chapter 12 shows a way to accept a piece of software by using statistical quality control principles. A software acceptance procedure is a user's tool to determine whether to accept or to reject the software. The examples given make the procedure more clear.

Chapter 13 elaborates the quality programming methodology to develop quality programs by an example that follows each step in the process and shows how the methodology incorporates the principles of software engineering and statistical quality control. It also describes how to use the defective rate of software product-unit-population to assess software quality, and how to calculate, statistically, the measures of software quality. An example of software warranty based on statistical quality measures is also given.

After discussing about quality of a software in the preceding chapters, Chapter 14 has been devoted to software reliability. Currently, many software reliability models, based on hardware reliability models, are being proposed for measuring software reliability. Some of these models, viz., Mean time to failure, Jelinsk-Moranda, Schick-Wolverton, Musa, and Nelson models are discussed in this chapter. Despite the attractions of very strongly built mathematics, the applicability of hardware theory to software is in general questionable, because, a piece of software is in general not subject to deterioration. The author has brought out that most of these models cannot be applied even to a very simple and small piece of software. The deficiencies of these models are pointed out from the perspective of statistical quality control and a reliability measure based on defective rate of a software product-unit-population is proposed.

Last chapter of the book spells out the role of top management of a software development organization. Dr W.E. Deming, considered to be the father of the wave of industrial revolution, introduced new principles of management into Japanese industry and revolutionised Japanese product quality and productivity through his work in statistical quality control. He has identified 14 obligations that the

top management of any institution must fulfil in order to be responsible to stockholders and consumers. As many as ten of these can be applicable to the software industry. The details of these obligations and the way of implementing them are discussed in this chapter.

The book is written to serve as (i) a text book for software engineering and methodology courses at undergraduate and graduate levels, (ii) a reference book for practising professionals and (iii) a software development and quality assurance plan for both the users and the developers. It has fulfilled these objectives very well.

The usefulness of this book can also be judged from the fact that numerous software companies have successfully used statistical quality control tool in software development or are in the process of applying the tool to their software development environment.

**S.K. Sharma***

---

* Shri S.K. Sharma is a Research Officer in the Systems and Programming Division of the Department of Statistical Analysis and Computer Services (DESACS).