India KLEMS

Labour Input- Quantity and Quality Measurement in India

Suresh Chand Aggarwal
Satyawati College,
University of Delhi, India
&
External Consultant ICRIER, New Delhi, India

RIETI/G-COE Hi- Stat International Workshop
Tokyo
22nd October 2010

Research assistance by Gunajit Kalita in creating the India KLEMS Labour Input dataset

Major tasks for Data Base on Labour

I. Make a Time series of Employment [number of persons and person hours] from 1980 to 2004

II. Prepare a Labour Quality Index from 1980 to 2004

III. Make a Time series of Labour Input from 1980 to 2004[I*II]

(Extended recently to 2008)

Major Contributions

- Efforts have been made for the first time to estimate employment in Hours
- Average number of Hours worked in a day have been estimated for the first time
- Both the Quinquennial and the Annual rounds have been used, for the first time for constructing the time series of employment
- A separate decomposition of Labour Quality into indices of age, sex and education has been attempted

Major Sources of Data Used

- For all sectors of the economy Employment and Unemployment Surveys (EUS) by **National Sample Survey Organization** (NSSO) and Population Census

  The two are Household/Individual specific

- Manufacturing Sector:
  - Organized Manufacturing industries-**Annual Survey of Industries(**ASI) by Central Statistical Organization (CSO)
  - Unorganized Manufacturing industries- Residual

Task I

Methodology for Constructing the Time Series of Employment

Time Series of employment requires estimation of:
a) Number of persons, and

b) Total days and hours worked by each person.

- **Time Series of Labour Input- Number of persons employed**
  - OECD(2001) and EU KLEMS have estimated Labour-productivity in terms of **output per labour hour worked**. OECD does not favour using count of jobs

So for international comparisons we must make efforts to do the same

Then the issues are i) How to measure number of persons employed?; and ii) How to measure number of total hours worked?

Number of persons employed

- In India, the number of employed may be estimated from Census and/or from Employment and Unemployment Survey (EUS)
- While Census has been held every ten years, NSSO has conducted both major (or Quinquennial) and thin (or annual) rounds of EUS

- Census gives us the population every ten years since 1951 and also number of Main workers, Marginal workers and non workers. It also gave main workers with other work (MWOW) for 1981 and 1991 census

Employment and Unemployment Survey (EUS)

- Major (Quinquennial) Rounds of EUS since 1980: 38th (1983), 43rd(1987-88), 50th(1993-94), 55th (1999-00) and 61st(2004-05)
- Thin (Annual) Rounds: 45th to 60th
- EUS uses Usual Status [Usual Principal Status(UPS) and Usual Principal & Subsidiary Status (UPSS)], Current Weekly Status(CWS) and Current Daily Status (CDS) measures for Quinquennial (or major) rounds and Usual Status & CWS for annual (thin) rounds
- While UPS, UPSS and CWS measure number of persons, the CDS gives number of jobs

ISSUE IS CHOICE OF AN APPROPRIATE MEASURE and COMPARABILITY OF DIFFERENT ROUNDS

Definition of UPSS, etc.

- The ***usual principal status*** gives the number of persons who worked for a relatively longer part of the reference period of 365 days preceding the date of survey
- While the ***usual principal status*** and the ***subsidiary status,*** includes the persons who (a) either worked for a relatively longer part of the 365 days preceding the date of survey or (b) who had worked some time (minimum 30 days since 61st round) during the reference period of 365 days preceding the date of survey
- The ***current weekly status*** provides the number of persons worked for at least 1 hour on any day during the 7 days preceding the date of survey, and
- The ***current daily status*** gives the picture of the person-days worked during the reference week of the survey period

Contd….

- UPSS is the most liberal and widely used of these concepts. It includes all workers who have worked for a longer time of the preceding 365 days in either the principal or in one or more subsidiary economic activity
- **Advantages of using UPSS**
  - It provides more consistent and long term trend
  - More comparable over the different EUS rounds
  - When adjusted for population distribution, it provides the count of jobs
  - Wider agreement on its use for measuring employment [Visaria(1996), Bosworth; Collins & Virmani (2007), Sundaram (2008), Rangarajan (2009)]
  - It can also be calculated for thin rounds

Contd….

- **Some problems in using UPSS**
  - Seeks to place as many persons as possible under the employed
  - No single long-term activity status for many due to movement to many jobs
  - Requires a recall of one year
- Though, UPSS has some limitations, but this is the best measure to use given the data

For India KLEMS we have used UPSS to estimate employment

EUS- Rounds Comparability

Rounds NOT COMPARABLE

a) 43rd round because of severe drought

b) **Conceptual differences in different rounds-** esp. between first three (3-way classification of 'Employed', 'Unemployed' and 'Out of labour Force') and last three (2-stage classification; first of 'in labour force' and 'out of labour force', and second of 'employed' or 'unemployed')

Rounds are COMPARABLE

-First three and last three; Himanshu (2007); NSSO (SDRD Team; 2008), especially UPSS estimates

Suggestion: For KLEMS all rounds may be used

Time Series of Labour Input- Numbers

- But if only major rounds of EUS are used for estimating Employment, then we have data only on selected five points
- So the issue was of constructing a time series from these data points

- **Alternatives were:**
  I. Interpolation from These Five Points
  II. Since early nineties, annual (thin) round data is also available. Combine it with major rounds

Time Series of Labour Input –Numbers

- If thin rounds are also used, then the issues are:

  I. Comparing major rounds with thin rounds; and

  II. Obtaining three digits data through thin rounds

Accepting the suggestions of the experts, we made use of both the major and the thin rounds

Survey Rounds

Round	Survey period Month/Year	Mid Period	Forecasting Period
38	1/83 to 12/83	01-Jul-1983	1
43	7/87 to 6/88	01-Jan-1988	19
45	7/89 to 6/90	01-Jan-1990	27
46	7/90 to 6/91	01-Jan-1991	31
47	7/91 to 12/91	01-Oct-1991	34
48	1/92 to 12/92	01-Jul-1992	37
49	1/93 to 6/93	01-Apr-1993	40
50	7/93 to 6/94	01-Jan-1994	43
51	7/94 to 6/95	01-Jan-1995	47
52	7/95 to 6/96	01-Jan-1996	51
53	1/97 to 12/97	01-Jul-1997	57
54	1/98 to 6/98	01-Apr-1998	60
55	7/99 to 6/00	01-Jan-2000	67
56	7/00 to 6/01	01-Jan-2001	71
57	7/01 to 6/02	01-Jan-2002	75
58	7/02 to 12/02	01-Oct-2002	78
59	1/03 to 12/03	01-Jul-2003	81
60	01/04 to 6/04	01-Apr-2004	84
61	7/04 to 6/05	01-Jan-2005	87
62	**7/05 to 6/06**	**01-Jan-2006**	91
64	**7/07 to 6/08**	**01-Jan-2008**	99

Contd…

- Since different rounds of EUS use different National Industrial Classification (NIC), so a Concordance between India KLEMS, NIC-1970, 1987 and 1998 required for all the 31 sectors was done
- Some sector had to be sub-divided for the concordance
- The interpolation from the major rounds was done for the period 1980-81 to 2004-05. The interpolated numbers were then constrained by the numbers obtained from the industrial distribution of the thin rounds

Total days worked

- Once the numbers were obtained ,efforts were made to obtain *'total days worked'* estimates from:
  - EUS - Time Disposition during the week
  - CDS and use intensity of work –

  Full time (≥4hours) and Part time (<4hours)

- Information on man-days workers and man-days employees at all India for all organized manufacturing industries was taken from ASI

Estimation of Employment

- Employment has been computed as follows:
  - I. Used, like all the previous studies, the Work Participation Rates (WPRs) by UPSS from EUS and applied them to the corresponding period's Census population of Rural Male, Rural Female, Urban Male and Urban Female to find out the number of workers in the four segments
  - II. Use the 31-industry distribution of Employment from EUS and used these to the number of workers in step I and obtained L_{ij} *for each industry where i=1 for rural and 2 for urban sectors, and j=1 for male and 2 for female*

Contd….

III. Find out the average number of days worked per week 'd_{ij}' for each industry from the intensity of employment as given in the CDS schedule

IV. Assuming average 48 hours work week for regular workers and 8 hours per day for self employed and casual workers, find out the expected number of hours 'h_{ij}' worked per day from the status-wise distribution, in each industry for rural male, rural female, urban male and urban female

Contd….

V. From the major rounds separate interpolation of L_{ij} ; d_{ij}; and h_{ij} was done for rural male, rural female, urban male and urban female to obtain the respective time series

VI. Broad Industrial distribution(Primary, Secondary and Tertiary) from annual rounds was used as a control total on the corresponding interpolated L_{ij} and revised numbers were obtained

VII. Total person hours in a year were obtained for each industry as the sum of the products of revised persons with days; hours and 52 over gender and sectors $\Sigma_i\Sigma_j L_{ij}*d_{ij}*h_{ij}*52$

Task II

Methodology for Constructing the Time Series of Labour Quality Index

- Quality Index has been constructed using Jorgenson, et al (1987) methodology which uses the Tornqvist translog index

They have expressed the *volume of labour input*, L; as a *translog index* of its individual components and the weights are given by the average shares of the components in the value of labour compensation. The growth rate of the aggregate labour volume index is defined as:

$$\Delta \ln L_w = \Sigma_l v_{ll} \Delta \ln L_l$$

$$v_{ll} = \tfrac{1}{2} [v_l(t) + v_l(t-1)]$$

and $v_l = w_l^L L_l / \Sigma_l w_l^L L_l$

where L_w is the weight adjusted aggregate labour

L_l is labour of a particular education class

l= 1,2,…..,n i.e. the number of education categories

v_l is the value share of labour for the lth education category

w_l^L is the wage rate of labour for the lth education category

Σ_l is the summation over all education categories

Contd….

- Growth of labour volume L incorporates both growth in hours worked and improvement in labour quality
- Since data on hours worked for each educational category of labour is not easily available, we assume that labour input for each category is proportional to hours worked and the proportion is same for all categories
- It follows from this that the growth rate of the quality index Q^L can be expressed in the form:

  $$\Delta \ln Q^L = \Sigma_l v_{ll} \Delta \ln L_l - \Delta \ln L$$

  where $L = \Sigma_l L_l$

  Q^L is the quality index of labour

  L is the total number of labour (unadjusted) of all education categories
- This is the difference between the percentage change in quality-adjusted labour and the percentage change in actual labour, summed over all categories

Time Series of Labour Quality Index

- Analogously, other first order contributions by gender, age and education, Q^s , Q^a, and Q^e , have also been computed

a) Employment by sex by age by education by industry
b) Earnings for each of these cells

Data required for Quality Index is:

Broad classifications for the series

- Gender: Males/Females
- Age : <29; 30-49; and 50+
- Education: Up to Primary(5 years) ; From Primary to Higher Secondary(12 years); and above Higher Secondary
- Sectors : 31 sectors

So, the total cells are 2*3*3*31

Time Series of Labour Quality Index

- Since the required labour composition data is available only from major rounds of EUS, so
  - Only **Major rounds** have been used for estimating the indices and the indices have been interpolated to get the time series for the entire period
  - Only for **aggregate 31 sectors**- not for organized and unorganized separately

Earnings Data

- NSSO's EUS relates earnings to only regular- salaried workers and casual workers

- The issue was how to estimate earnings of self employed

Earnings of Self Employed is required for quality index and labour compensation.

Earnings of Self Employed by KLEMS

- OECD assumes that labour characteristics of both employees and self employed is same within an industry. So average compensation per hour of a self employed person is taken to be equal to that of a wage earner
- EU KLEMS has followed the OCED procedure for most of EU countries, but on the basis of some surveys in few places they have estimated it to be 0.80 for some sectors, especially agriculture and 1.20 for sectors like business services

Earnings of Self Employed in India

- Two alternatives were considered :

I. Use earnings of **Self employed** to be equal to that of **Casual labour** as the labour market for the two is comparable

II. To fit an earning function to earnings of casual and regular employees and use it to find the corresponding earnings of the self employed

India KLEMS preferred to use the second option and has used the Mincer Wage equation for the same and sample selection bias has been corrected for by using Heckman's two step procedure

Heckman Model

- The Heckman model is formulated in terms of two equations:
  - a selection equation – usually a Probit estimation (takes a value of 1 if a person is working, 0 otherwise) to explain the decision of whether to participate in the labour market and
  - a regression equation to explain days of actual labour market participation, observable only for those for whom the selection equation takes a value of 1
- This technique helps us overcome the problem of not being able to observe the wage for those who are not employed in the reference period
- The function has been used to the earnings of casual and regular employees
  - *The earnings have been regressed on the dummies of age, sex, education, location, marital status, social exclusion and industry*
  - *The identification factors used in the first stage are age, sex, marital status, and type of household / size of households*

Summary Results to be presented in the Workshop

Thank You