Measuring Quantity and Quality of the Chinese Industrial Workforce, 1949-2005

Harry X. Wu, Hitotsubashi University & The Hong Kong Polytechnic University

Ximing Yue, Renmin University, Beijing

Prepared for Workshop on “East Asian Economic Development in Historical Perspective”, Institute of Economic Research, Hitotsubashi University, July 3, 2009

Agenda

- Motivation and the previous studies
- Main problems in measuring labour input in the Chinese economy
- The labour input indexing methodology
- Procedures in constructing marginal labour employment and compensation matrices
- Procedures of constructing the full-dimensioned labour employment and compensation matrices
- Results and discussion
- Concluding remarks

Motivation

- Debates about China's productivity performance over both the planning and reform periods remain inconclusive largely due to data problems in measuring both inputs and outputs.
- To satisfy a homogenous production function analysis the measure of every input has to be homogenous (Jorgenson, 1990, p.33). This study deals with the measurement problem of one of the inputs in the industrial sector, that is labour.
- The core issue here is how to hold the quality of hours worked constant when there are actually changes in the quality of workforce due to changes in the composition of age, gender, education and occupation, as well as industrial structure
- In other words, how to convert heterogeneous hours worked into homogenous volume of labour input
- Ignoring this will force part of labour input (non-measured quality change, which is inputs related or non-Hick neutral) into TFP

The previous studies

- Labour input in the Chinese economy has never been properly measured.
- Most studies simply used the numbers employed as a proxy for labour input (Borensztein and Ostry, 1996; Chan et al., 1988; Chow, 1993; Chow and Li, 2001; Hu and Khan, 1997) and focused on the aggregate economy.
- Young (2003) tried to identify the conceptual and inconsistency problems with limited census and survey data and did not provide alternative estimates on labour input
- In terms of using the standard Jorgenson approach, Li et al (1993) was the pioneer and then followed by Cao et al (2009). However, both focused on the reform period only and Li et al covered a very short period; and both did not attempt to tackle conceptual problems in the official labour statistics.

Data Problems

- Serious inconsistency in industrial classification over time (1972, 1985, 1994 and 2002 CSICs) especially for the pre-1994 period but lack of information that is necessary for reconciliation or reconstruction
- Inclusion of service employees in industrial workforce who provide services in education units, medical clinics, child care centres, commercial outlets, and social and political organisations in industrial enterprises, yet lack of systematic information for adjustment
- The quantity of employment has never been measured in its natural unit, i.e. hours worked, despite different working hour system among industries and institutional changes of standard working hours in general

Data Problems …

- Until 1998, off-post (unemployed) workers in SOEs were kept in the payroll and counted into the workforce
- Regarding data on human capital of employees, both regular reporting system and census do not provide full cross-classification of different demographic, educational and occupational attributes of the industrial workforce
- Data on labor compensation are even more scant; income data from limited number of IO tables cannot be easily reconciled with the wage data from regular labour and industrial statistics

Labor Input Indexing

- Labour input indexing should be discussed coherently with a production function aggregating the services provided by different types of labour and capital.
- The essential idea of constructing labour input index roots in the heterogeneity of labour in the sense that different types of labour have different marginal products (Denison, 1962; Jorgenson and Griliches, 1967). (eq.1 and eq.2)

$$Y_t = A_t f(L_t, K_{1t}, \ldots, K_{jt})$$

$$L_t = \phi(H_{1t}, \ldots, H_{nt})$$

Labor Input Indexing…

- Following Chinloy (1980) and Jorgenson, Gollop and Fraumeni (1987), if assume efficient labour market and linear homogeneity of Φ, then we have: (eq.3)

$$\frac{\partial \ln L_t}{\partial t} = \sum_{i=1}^{n} s_{it} \frac{\partial \ln H_{it}}{\partial t}$$

- where s_{it} is the share of the ith type of labour in total labour compensation, which is equal to its logarithmic marginal output under the efficiency assumption (eq.4)

$$s_{it} = \frac{w_{it} H_{it}}{\sum_{i=1}^{n} w_{it} H_{it}} = \frac{\partial \ln \phi}{\partial \ln H_{it}}$$

Labor Input Indexing…

- In eq.4 the hourly wage of the *i*th type of labour is w_{it} and its compensation is $w_{it}H_{it}$. The growth rate of labour input is a convex combination of growth rates of total hours for each type of labour, with compensation shares as weights.
- The eq.4 also indicates that the necessary condition for producer equilibrium is given by equality between the share of the *i*th type of labour in the labour aggregate and the elasticity of the aggregate with respect to the *i*th type of labour.

Labor Input Indexing…

- Let total hours worked by all types of labour be $H_t = \sum_{i=1}^{n} H_{it}$ Then, the growth rate of H_t is the sum of the weighted growth rates of hours worked by each type of labour (eq.5)

$$\frac{\partial \ln H_t}{\partial t} = \sum_{i=1}^{n} b_{it} \frac{\partial \ln H_{it}}{\partial t}$$

- with $b_{it} = H_{it} / \sum_{i=1}^{n} H_{it}$ the weight of the *i*th labour type. Therefore, average labour quality per hour can be defined as labour input divided by hours worked: (eq.6, eq.7)

$$Q_t = L_t / H_t$$

$$\frac{\partial \ln Q_t}{\partial t} = \sum_{i=1}^{n} (s_{it} - b_{it}) \frac{\partial \ln H_{it}}{\partial t}$$

Labor Input Indexing…

- Now following Christensen, Jorgenson and Lau (1973), we specify the labour aggregate in the translog form: (eq.8)

$$\ln L_t = \alpha_0 + \sum_{i=1}^{n} \alpha_i \ln H_{it} + 1/2 \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij} \ln H_{it} \ln H_{jt}$$

- where α_0, α_i, $i = 1, \ldots, n$, and β_{ij}, $i, j = 1, \ldots, n$, are parameters and where $\beta_{ij}=\beta_{ji}$ to satisfy the required symmetry conditions. Under linear homogeneity, we have $\sum_{i=1}^{n} \alpha_i = 1$ and $\sum_{j=1}^{n} \beta_{ij} = 0, \quad i = 1, \ldots, n$
- With efficiency assumption, the share of the ith type of labour equals its logarithmic marginal product (eq.9) (s_{it} is as eq.4)

$$s_{it} = \alpha_i + \sum_{j=1}^{n} \beta_{ij} \ln H_{jt}$$

Labor Input Indexing…

- Equations (8) and (9) as well as the symmetry condition $\beta_{ij}=\beta_{ji}$ simply that the growth rate of the translog index of labour input l_t is: (eq.10)

$$l_t \equiv \Delta \ln L_t = \sum_{i=1}^{n} v_{it} \, \Delta \ln H_{it}$$

- Where $v_{it} = (s_{it} + s_{it-1})/2$
- From eq.6, we can have (eq.11)

$$q_t \equiv \Delta \ln Q_t = \Delta \ln L_t - \Delta \ln H_t = l_t - h_t$$

- Clearly, the growth rate of quality will be positive if hours worked by relatively high wage labour increase more rapidly than total hours worked.

Labor Input Indexing…

- The contribution of each attribute of labour to quality change can be decomposed into two types of effects: the main effect of the attribute and the interactive effects of the attribute with each of the rest attributes. The main effect of the *i*th attribute is defined as the difference between the growth rates of labour input due to the *i*th attribute and total hours worked, regardless the time (eq.12)

$$q_i = l_i - h$$

- In the case of $q_i > 0$, as noted in Jorgenson and Griliches (1967), labour input measured as total hours worked is biased downward and hence TFP is biased upward.

Labor Input Indexing…

- Suppose that there are two attributes of labour, *i* and *k*, as a proper subset from *n* factors, a first-order interactive effect is derived from the partial index growth rate l_{ik} for the two factors and the single factor indices l_i and l_k: (eq.13)

$$q_{ik} = (l_{ik} - h) - (l_i - h) - (l_k - h) = l_{ik} - h - q_i - q_k$$

- that is, the joint effect of *i* and *k* or (l_k-*h*) less the main effect of each. If there are only two factors, *i* and *k*, the growth rate of labour quality is defined as the summation of the main effects of two factors and their first-order interactive effect:

$$q = l_{ik} - h = q_i + q_k + q_{ik}$$

- As for labour input with *j* factors, *l =1…j*, interactive effects up to (*j* – 1)th order are obtainable following the same principle (Chinloy, 1980, p.111).

Procedures in Constructing Marginal Labor Employment Matrices (Time Series)

- Collecting, sorting and checking through all employment data available, published or kept in the state archives
- Reconciling different standards of industrial classification and making all industry data in line with the 1994 CSIC (=2002 CSIC for industry) at two-digit level, with the state sector as the "hard core", followed by the "township layer" and "outer layer"
- Filling gaps in all three layers (painstaking!)
- Removing service employees from the industrial workforce
- Converting numbers employed to hours worked
- The result used as "control totals":

$$\mathbf{N}^{M}_{s+++jot} = \sum_{g=1}^{2} \sum_{a=1}^{7} \sum_{e=1}^{5} \mathbf{N}_{sgaejot}$$

$$\mathbf{N}^{M}_{s+++jot} = 24 \times 3 \times 3 \times 57 = 12312$$

TABLE 1
CLASSIFICATION FOR SECTORAL AND HUMAN CAPITAL ATTRIBUTES OF LABOUR INPUTS

Industrial Sector (*s*)	Human Capital Attribute
1. Coal mining	Gender: (*g*)
2. Oil and gas extraction	1. Male
3. Metallic mineral mining	2. Female
4. Non-metallic minerals mining	Age Group: (*a*)
5. Food & kindred products	1. 15-19
6. Tobacco products	2. 20-24
7. Textiles	3. 25-29
8. Apparel	4. 30-39
9. Leather & leather products	5. 40-49
10. Saw mill products & furniture	6. 50-54
11. Paper products, printing & publishing	7. >54
12. Petroleum & coal products	Education Attainment: (*e*)
13. Chemicals & allied products	1. Illiteracy or semi-illiteracy
14. Rubber & plastics products	2. Primary school
15. Stone, clay & glass products	3. Junior high school
16. Metals smelting, pressing & rolling	4. Senior high school
17. Metal products	5. Tertiary education
18. Industrial machinery & equipment	Occupation: (*j*)
19. Transportation equipment	1. Managerial & administrative staff
20. Electrical equipment	2. Technicians & engineers
21. Electronic & telecommunication equip.	3. Production workers
22. Instruments and office equipment	Ownership Type: (*o*)
23. Miscellaneous manufacturing	1. State-owned enterprises (SOEs)
24. Power, steam, gas and tap water supply	2. Non-SOEs at/above township level
	3. Other status below township (village level and household/self-employed)

Reconciliation of Various CSICs

Wu-Yue Code[a]	1994 CISC		1985 CSIC		1972 CSIC	
	Code	Industry	Code	Industry	Code	Industry
02	⇦07	Oil and natural gas extraction	⇦0900	Oil and natural gas extraction	⇦0401 ⇦0402	Oil extraction Gas extraction
12	⇦25	Petroleum refinery and coking	⇦3400	Petroleum refinery	⇦0403	Petroleum refinery
			⇦3510 ⇦3520	Coking Coal gas	⇦0321 ⇦0322	Coking Coal gas
24[b]	⇦45	Coal gas			⇦0310	Coal mining
01	⇦06	Coal mining	⇦0800	Coal mining		
…	…	…	…	…	…	
03	⇦08	Ferrous metal ore mining	⇦1000	Ferrous metal ore mining	⇦0111	Ferrous metal ore mining
	⇦09	Non-ferrous metal ore mining	⇦1100	Non-ferrous metal ore mining	⇦0121	Non-ferrous metal ore mining
16	⇦32	Ferrous metal smelting and pressing	⇦4800	Ferrous metal smelting and pressing	⇦0112	Ferrous metal smelting and pressing
	⇦33	Non-ferrous metal smelting and pressing	⇦4900	Non-ferrous metal smelting and pressing	⇦0122	Non-ferrous metal smelting and pressing

Occupational structure of the state firms

	1963				1985				1995			
	W	T	A	N	W	T	A	N	W	T	A	N
	A: The State "Staff and Workers" (the "Hard Core")											
All	75.3	4.4	9.6	10.6	70.3	4.1	10.7	14.9	62.0	6.0	10.2	21.8
1	76.2	3.1	7.8	12.8	69.1	1.6	7.5	21.8	56.1	2.4	6.8	34.8
2	68.7	8.0	12.1	11.2	61.5	4.1	12.9	21.5	52.0	8.5	11.8	27.8
3	70.2	4.3	8.8	16.8	67.3	3.1	10.2	19.3	59.6	4.7	10.1	25.6
4	72.6	3.1	10.1	14.2	70.2	1.9	11.9	15.9	64.9	4.4	11.3	19.5
5	79.0	1.3	11.8	7.9	79.8	1.9	9.9	8.4	66.2	5.5	11.2	17.1
6	82.7	1.4	7.6	8.3	80.8	1.2	9.2	8.8	68.5	6.2	11.1	14.2
7	81.3	2.6	6.6	9.4	80.6	1.7	7.4	10.3	68.7	3.1	6.5	21.7
8	82.4	1.7	8.1	7.8	78.8	1.6	7.7	11.9	69.6	3.3	9.4	17.7
9	83.4	1.6	8.5	6.5	78.8	1.5	10.0	9.7	66.7	3.0	8.6	21.7
10	79.3	1.7	10.1	8.9	74.3	2.3	11.1	12.3	62.1	4.1	11.2	22.6
11	78.5	3.0	9.0	9.6	76.4	2.2	10.4	11.0	68.2	4.3	10.6	16.9
12	70.4	6.1	8.9	14.5	61.5	6.1	12.1	20.3	54.9	11.1	12.5	21.6
13	73.8	5.0	10.4	10.8	69.2	5.0	11.9	14.0	63.3	7.9	11.1	17.7
14	78.4	3.5	9.4	8.7	74.6	3.2	11.6	10.6	62.5	5.0	9.9	22.6
15	78.4	2.0	9.9	9.6	74.6	2.7	10.6	12.0	68.0	4.9	10.8	16.3
16	71.8	5.1	8.3	14.8	69.4	3.8	10.4	16.4	64.3	5.6	9.6	20.5
17	77.3	4.1	10.7	7.9	71.8	3.8	13.0	11.4	59.7	5.3	11.8	23.2
18	70.8	9.1	10.8	9.2	68.8	5.8	12.2	13.2	58.5	7.1	11.7	22.6
19	74.2	8.0	9.1	8.7	65.4	6.9	12.9	14.8	58.4	8.8	11.8	21.1
20	70.8	9.1	10.8	9.2	69.4	6.0	12.8	11.8	60.6	7.6	11.8	20.0
21	70.8	9.1	10.8	9.2	65.3	10.1	12.8	11.8	51.7	12.5	10.8	25.1
22	70.7	9.8	11.8	7.7	64.0	10.0	13.6	12.4	50.9	9.5	11.7	27.9
23	81.4	1.3	10.2	7.0	76.2	2.6	12.0	9.2	63.3	6.8	10.6	19.3
24	72.2	6.4	12.3	9.2	67.7	5.3	13.0	14.0	64.7	10.0	12.4	12.9

Converting numbers to hours

	1949-1953	1954-1957	1958-1993	1994	1995-2005
State: A (the baseline)	54	48	48	44	40
B	54	48	42	38	34
B*	54	48	42 (-1992)	46 (1993)	46 (1994-)
C	54	48	36 (-1992)	40 (1993)	40 (1994-)

	1949-1953	1954-1984	1985-1992	1993-2005
"Township Layer" (all industrial sectors)	54	48	52	56

	1949-1957	1958-1959	1960-1984	1985-1992	1993-2005
"Outer Layer" (all industrial sectors)	31.2 (65% of 48)	48	31.2 (65% of 48)	38.4 (80% of 48)	48

- Results:

$$\mathbf{H}^{M}_{s+++jot}$$

Procedures in Constructing Marginal Labor Compensation Matrices (Time Series)

- Collecting, sorting and checking all wage data available by industry and ownership types, published and kept in the state archives
- Wage data are adjusted to include welfare payment using census data, 1985, 1995 and 2004, assuming information for 1985 is applicable to the planning period
- The results are reconciled with the income data from China's input-output tables beginning in 1987 aiming to capture any in-kind payment estimated by the I/Os
- The results used as "control totals":

$$\mathbf{C}^{M}_{s++++ot} = \sum_{g=1}^{2}\sum_{a=1}^{7}\sum_{e=1}^{5}\sum_{j=1}^{3}\mathbf{C}^{M}_{sgaejot}$$

$$\mathbf{C}^{M}_{s++++ot} = 24\times 3\times 57 = 4104$$

Procedures in Constructing Additional Marginal Matrices (Benchmarks)

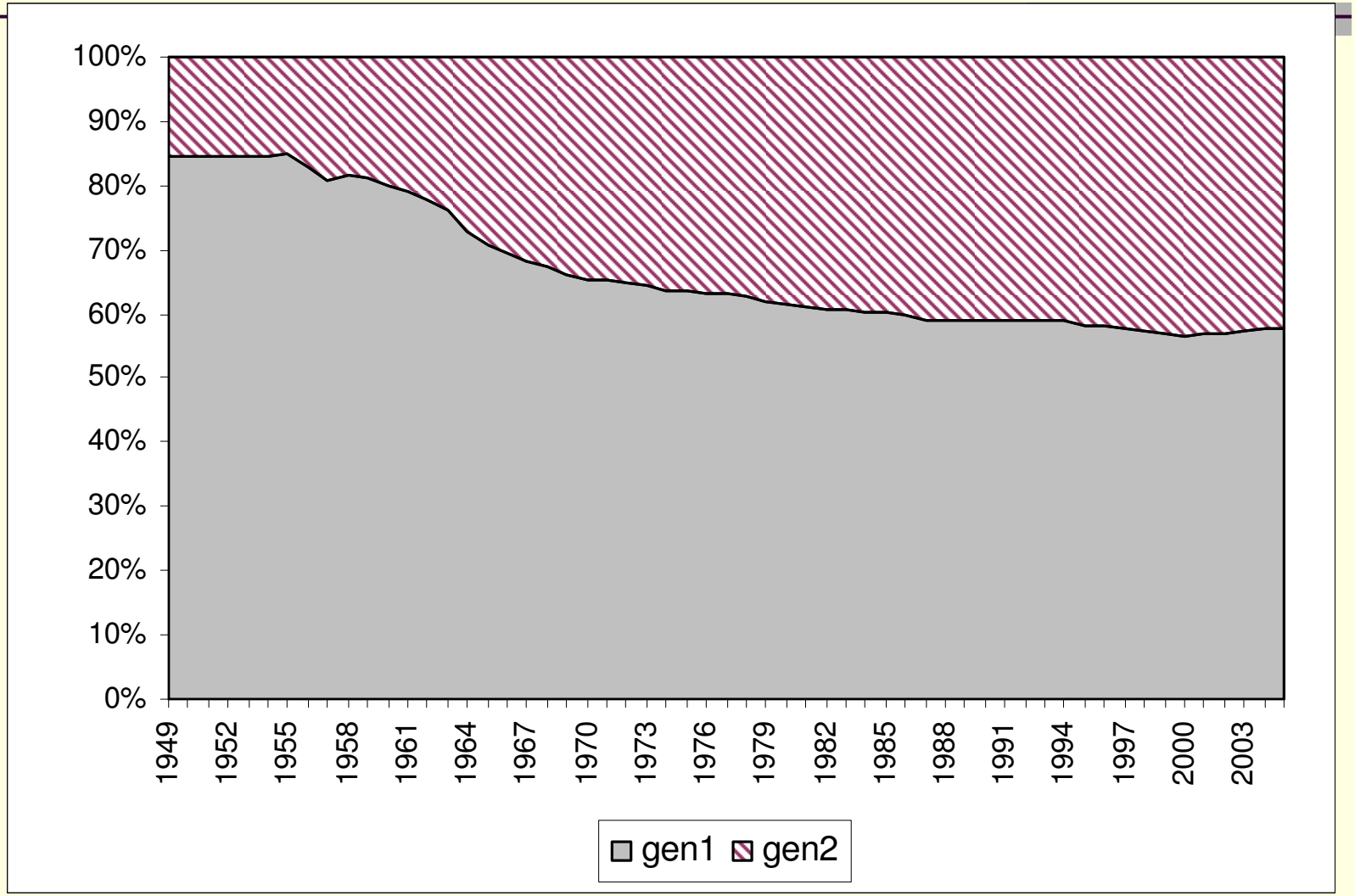
- All data used for constructing the time series marginal matrices are suitable for building up time trend and hence are at least systematic and consistent for certain period
- There are additional data that are only available for one or several isolated time points due to occasional surveys or censuses. The additional information can make our estimation more reliable for these time points
- Based on such data, we selected eight points as our benchmark years to add additional marginal matrices that can better anchor the constructed time series and make the estimation of full-dimensioned matrices more reliable
- The additional marginal matrices are reported in the following table:

Benchmarks	Additional Marginal Matrix		Source of Data	Marginal Matrix as in the "Control Totals"
	Employment	Compensation		
1955	$\tilde{\mathbf{N}}^{M,55}_{sg++j}$, $\tilde{\mathbf{N}}^{M,55}_{++a+j}$, $\tilde{\mathbf{N}}^{M,55}_{s++ej}$	$\tilde{\mathbf{C}}^{M,55}_{s+++j}$, $\tilde{\mathbf{c}}^{M,55}_{s+++j}$	NBS archives: DPES *Bulletin*, 1955, 1959, 1960, 1961 and 1963	$\mathbf{N}^{M,55}_{s+++jo}$, $\mathbf{H}^{M,55}_{s+++jo}$ $\mathbf{C}^{M,55}_{s++++o}$, $\mathbf{c}^{M,55}_{s++++o}$
1963	$\tilde{\mathbf{N}}^{M,63}_{sg++j}$, $\tilde{\mathbf{N}}^{M,63}_{++a+j}$, $\tilde{\mathbf{N}}^{M,63}_{s++ej}$	$\tilde{\mathbf{C}}^{M,63}_{s+++j}$, $\tilde{\mathbf{c}}^{M,63}_{s+++j}$	As above…	$\mathbf{N}^{M,63}_{s+++jo}$, $\mathbf{H}^{M,63}_{s+++jo}$ $\mathbf{C}^{M,63}_{s++++o}$, $\mathbf{c}^{M,63}_{s++++o}$
1982	$\tilde{\mathbf{N}}^{M,82}_{sga+++}$, $\tilde{\mathbf{N}}^{M,82}_{s++e++}$, $\tilde{\mathbf{N}}^{M,82}_{++aej+}$, $\tilde{\mathbf{N}}^{M,82}_{+gae++}$	--	1982 Population Census	$\mathbf{N}^{M,82}_{s+++jo}$, $\mathbf{H}^{M,82}_{s+++jo}$ $\mathbf{C}^{M,82}_{s++++o}$, $\mathbf{c}^{M,82}_{s++++o}$
1987	--	$\tilde{\mathbf{c}}^{M,87}_{+gaej+}$	1988 CASS Household Survey; 1987 1%-Population Survey	$\mathbf{N}^{M,87}_{s+++jo}$, $\mathbf{H}^{M,87}_{s+++jo}$ $\mathbf{C}^{M,87}_{s++++o}$, $\mathbf{c}^{M,87}_{s++++o}$
1990	--	$\tilde{\mathbf{c}}^{M,90}_{+gaej+}$	Based on assumptions, see text…	$\mathbf{N}^{M,90}_{s+++jo}$, $\mathbf{H}^{M,90}_{s+++jo}$ $\mathbf{C}^{M,90}_{s++++o}$, $\mathbf{c}^{M,90}_{s++++o}$
1995	--	$\tilde{\mathbf{c}}^{M,95}_{+gaej+}$	1995 CASS Household Survey	$\mathbf{N}^{M,95}_{s+++jo}$, $\mathbf{H}^{M,95}_{s+++jo}$ $\mathbf{C}^{M,95}_{s++++o}$, $\mathbf{c}^{M,95}_{s++++o}$
2000	$\tilde{\mathbf{N}}^{M,00}_{sg++++}$, $\tilde{\mathbf{N}}^{M,00}_{+gae++}$	$\tilde{\mathbf{c}}^{M,00}_{+gaej+}$	2000 Population Census; 2002 CHIP Household Survey	$\mathbf{N}^{M,00}_{s+++jo}$, $\mathbf{H}^{M,00}_{s+++jo}$ $\mathbf{C}^{M,00}_{s++++o}$, $\mathbf{c}^{M,00}_{s++++o}$
2005	Ximing???	Ximing???	Ximing???	$\mathbf{N}^{M,05}_{s+++jo}$, $\mathbf{H}^{M,05}_{s+++jo}$ $\mathbf{C}^{M,05}_{s++++o}$, $\mathbf{c}^{M,05}_{s++++o}$

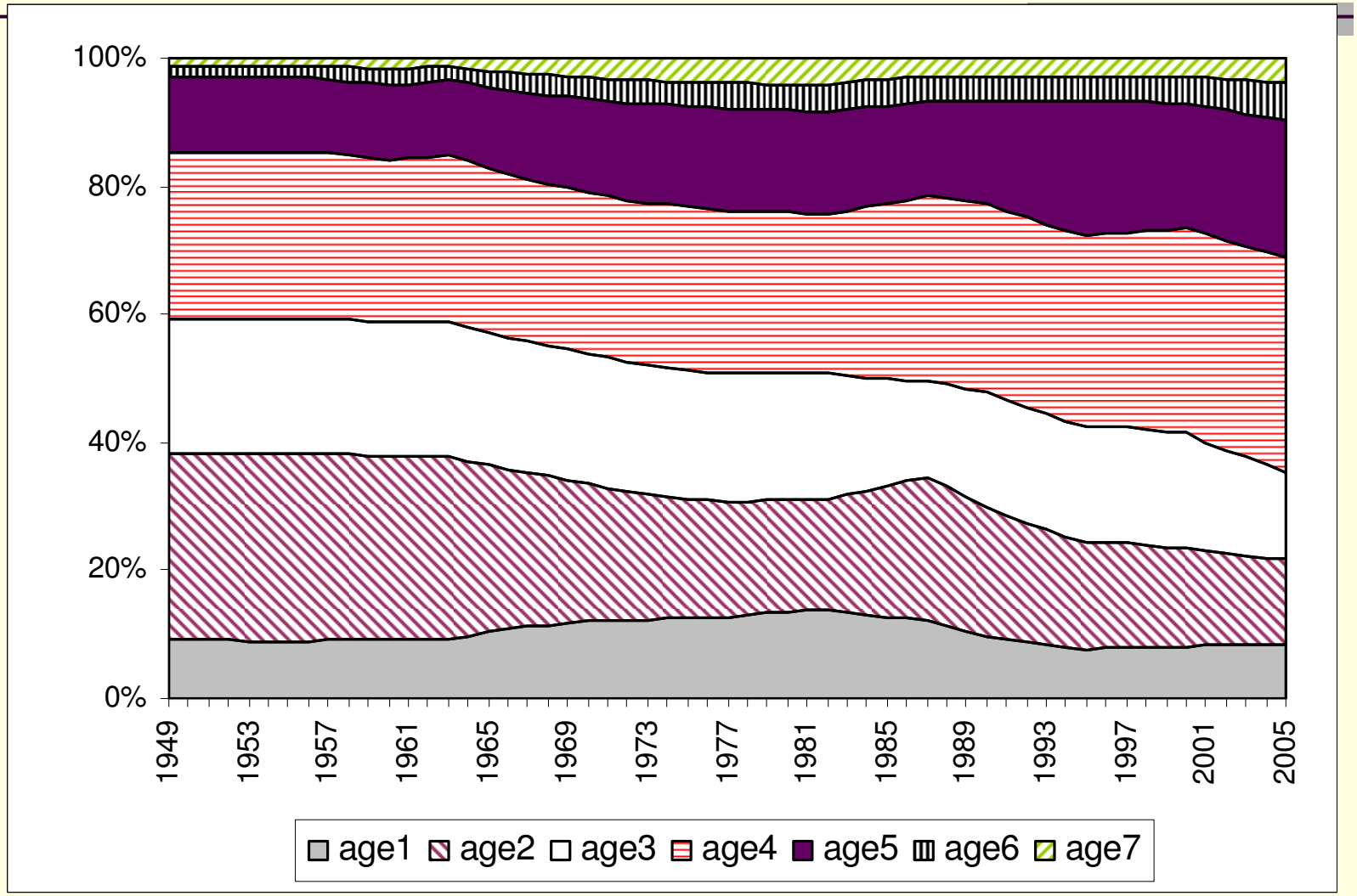
Procedures in Constructing Full-dimensioned Matrices (Time Series)

- Three steps towards the final full-dimensioned matrices:
  - Filling gaps or missing cells in additional marginal matrices, especially compensation matrices, using regression approach
  - Consistency adjustment to additional marginal matrices by scaling up them to the “control totals”
  - Estimating missing dimensions and making all dimensions fully cross-classified/fully compatible cross all dimensions using the IPF (iterative proportional filling) approach

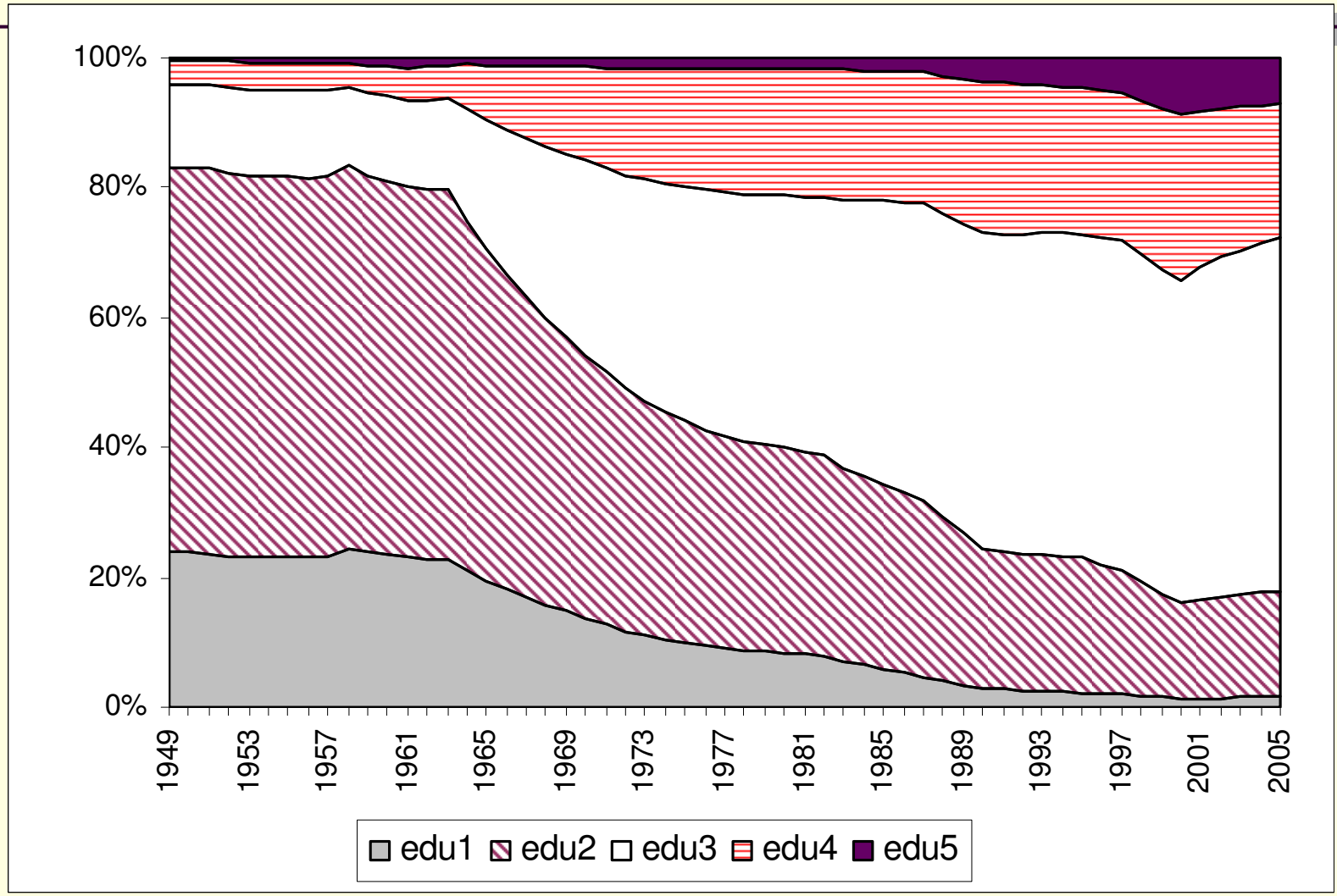
Changes of gender ratio of Chinese workforce, 1949-2005



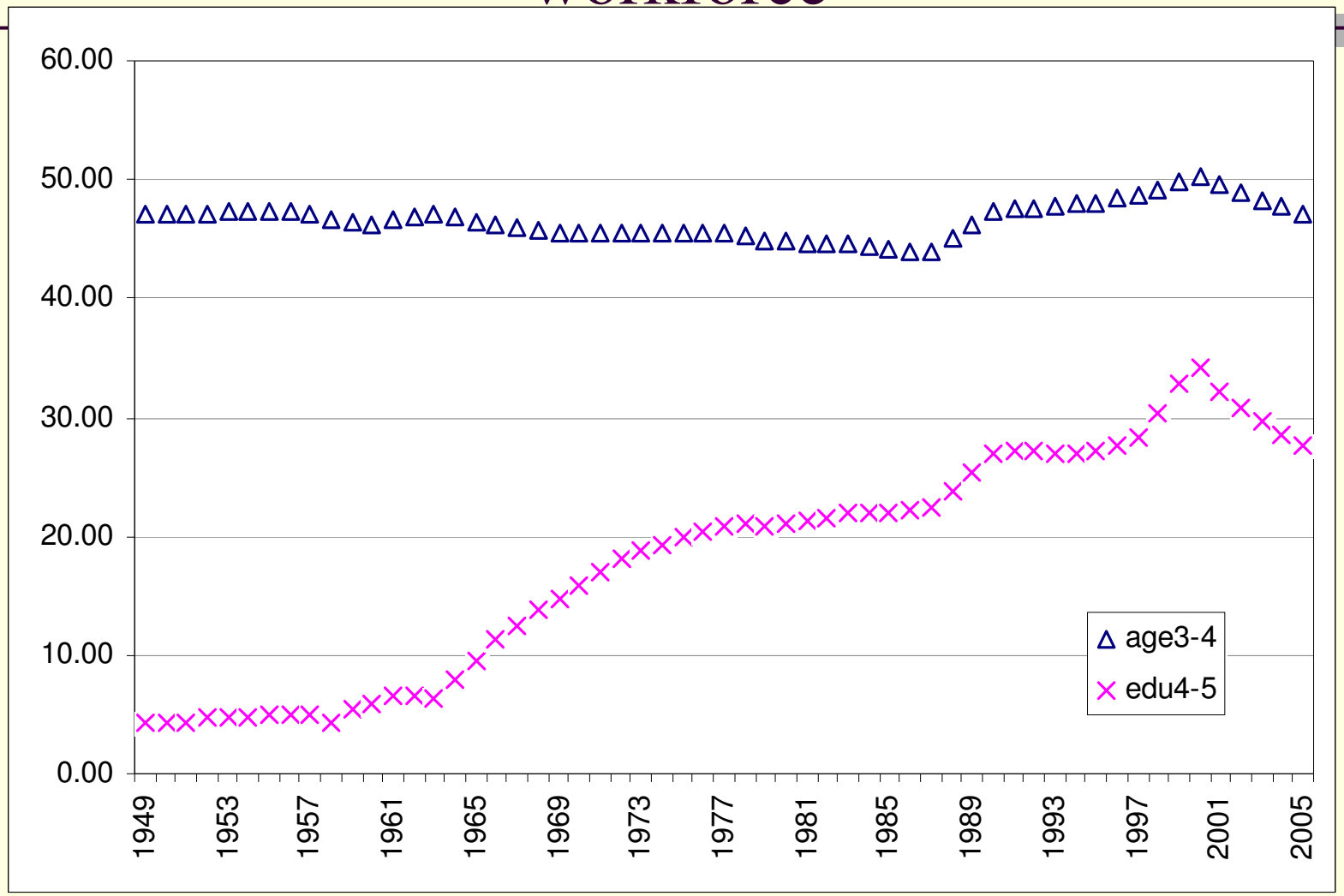
Changes of age structure of Chinese industrial workforce, 1949-2005



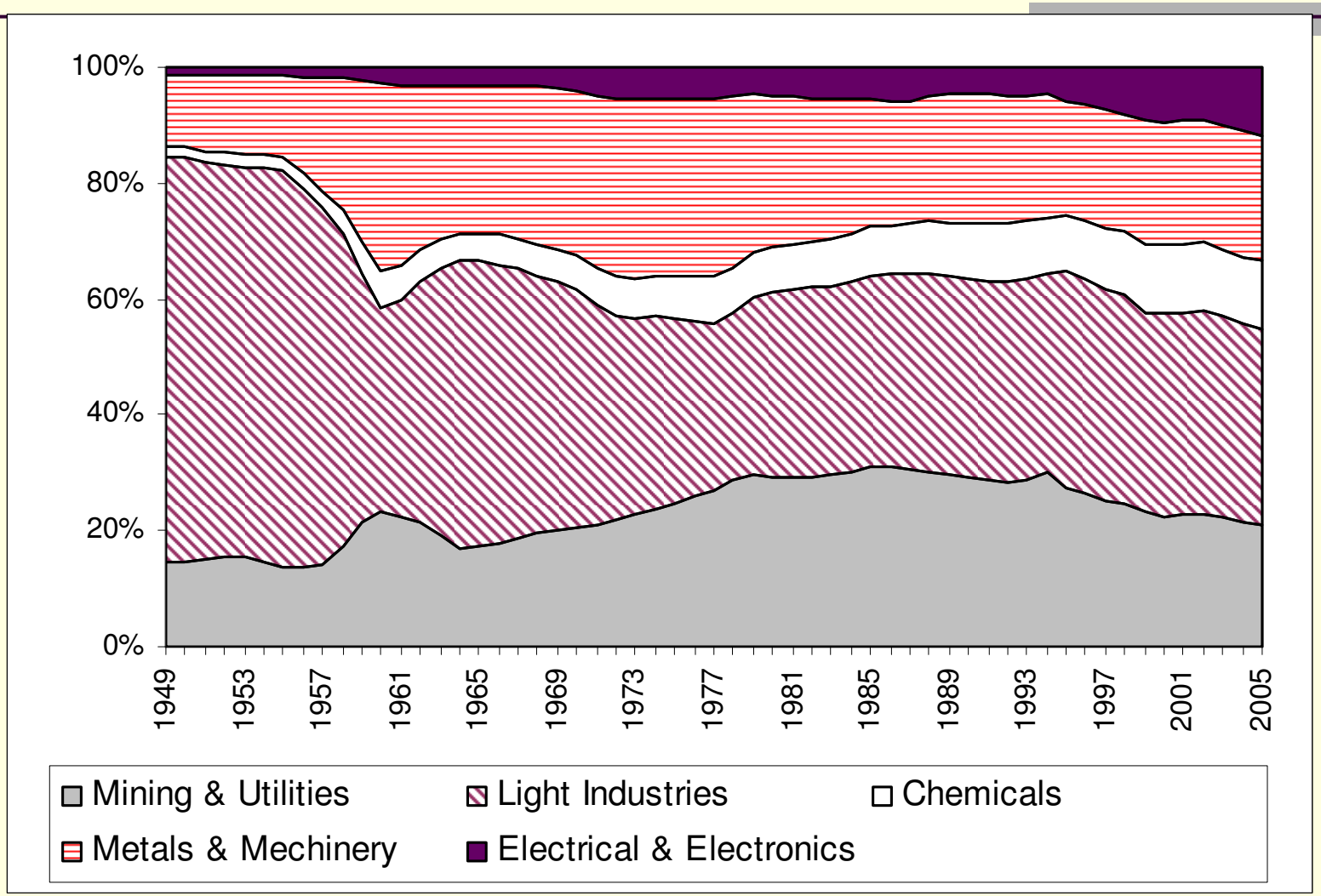
Changes of education attainment structure in Chinese industrial workforce, 1949-2005



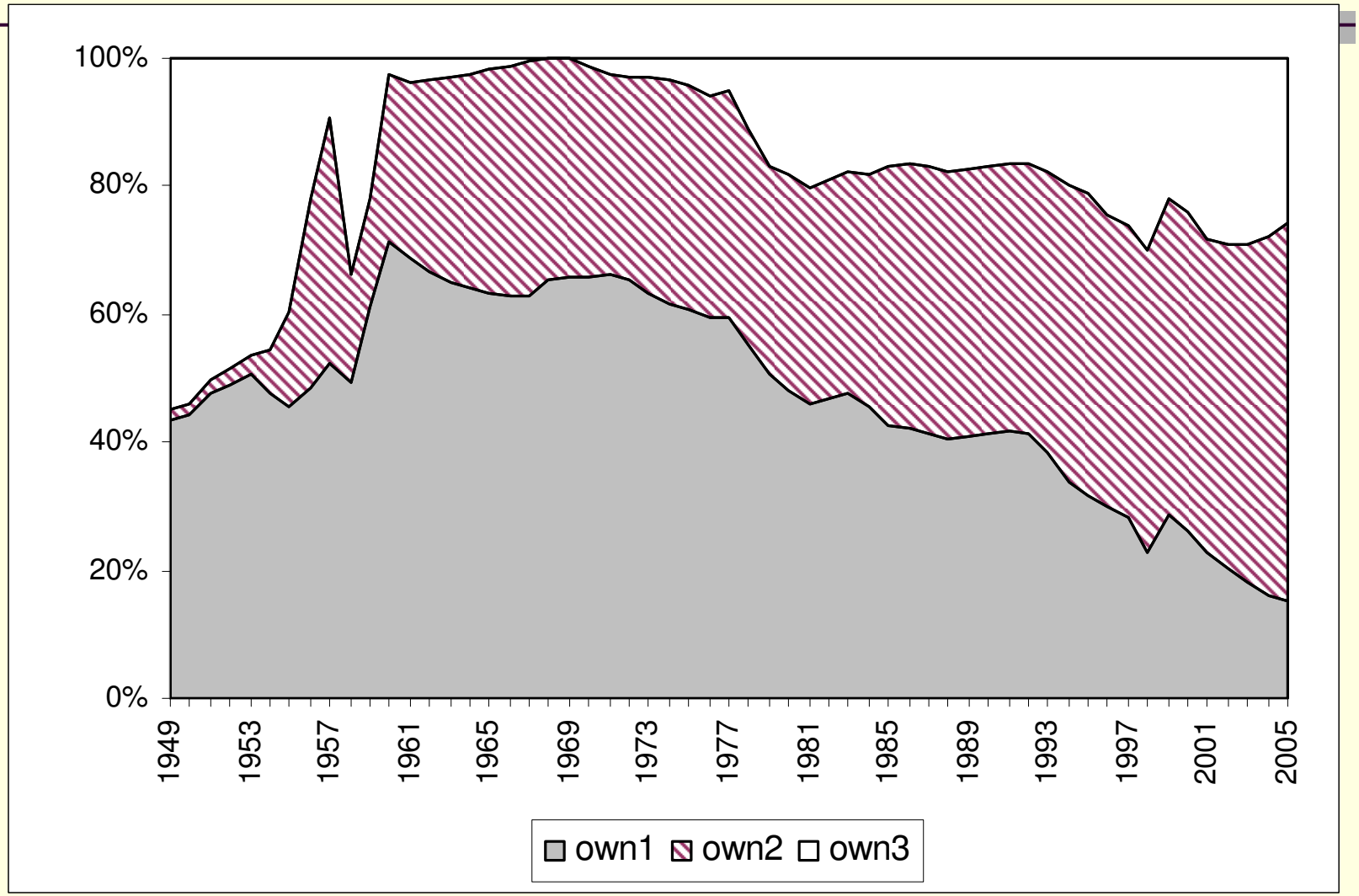
25-39 aged versus senior-high/above education attainment in Chinese industrial workforce



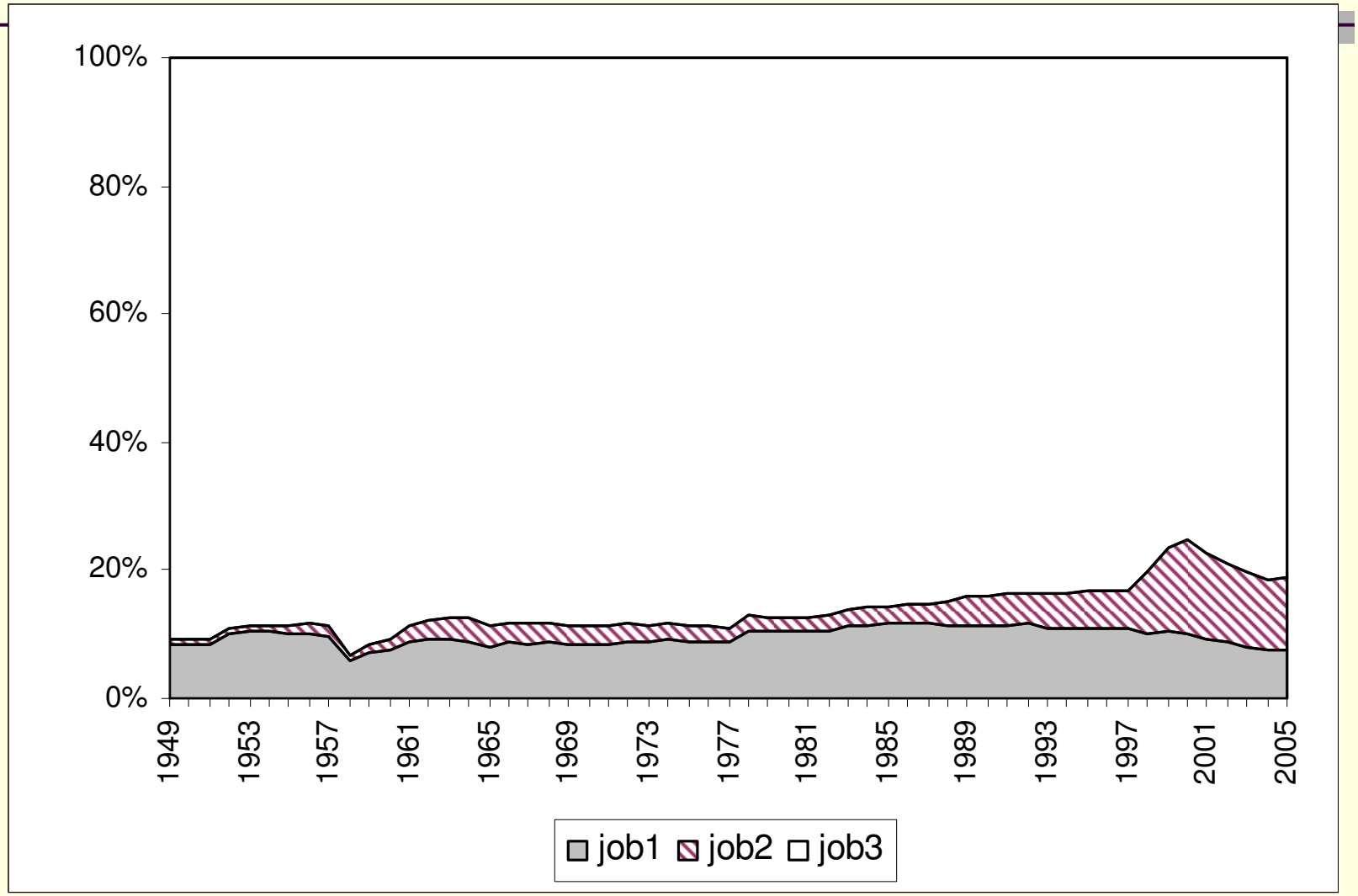
Changes of the industrial structure in Chinese industrial workforce, 1949-2005



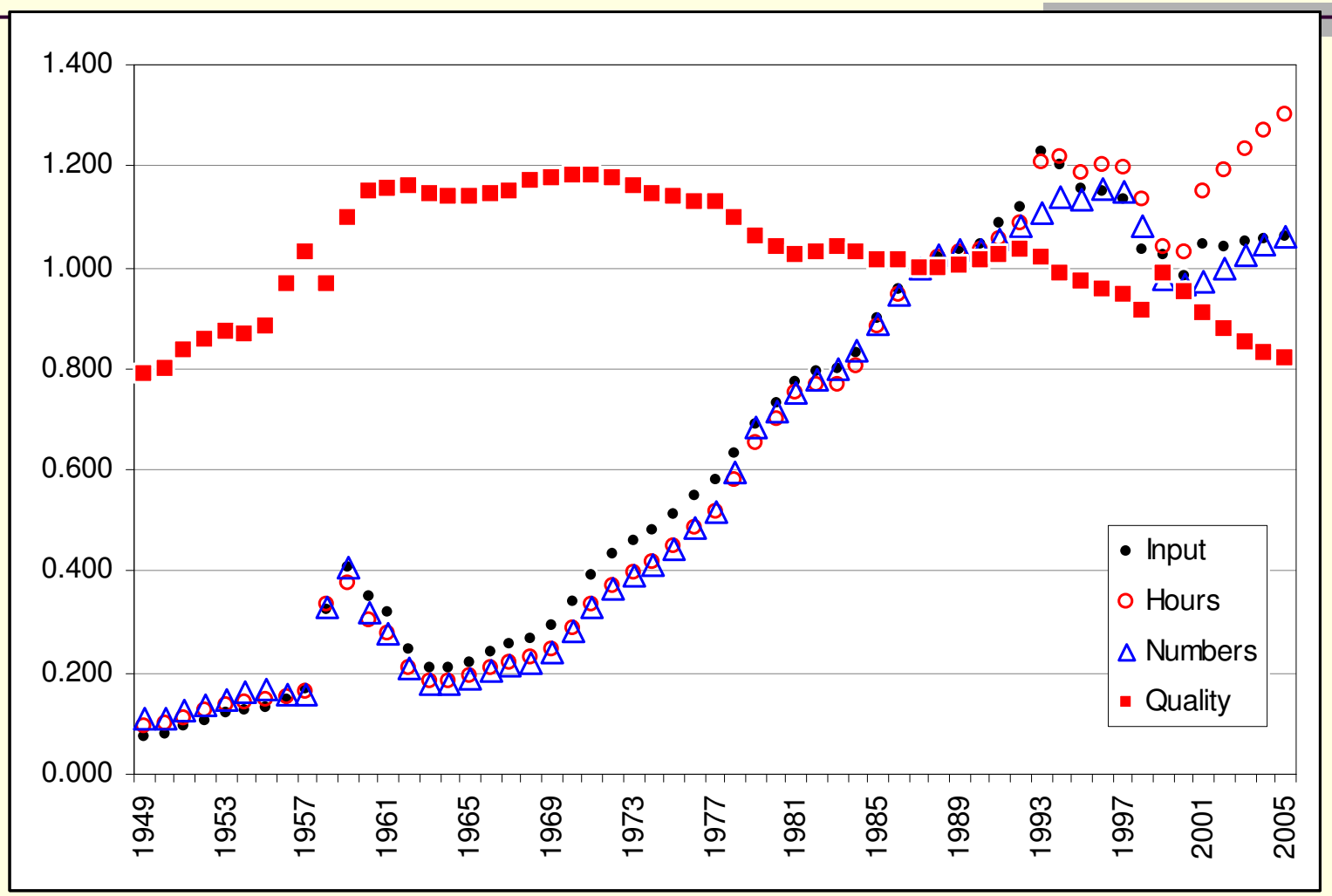
Changes of ownership structure in Chinese industrial workforce, 1949-2005



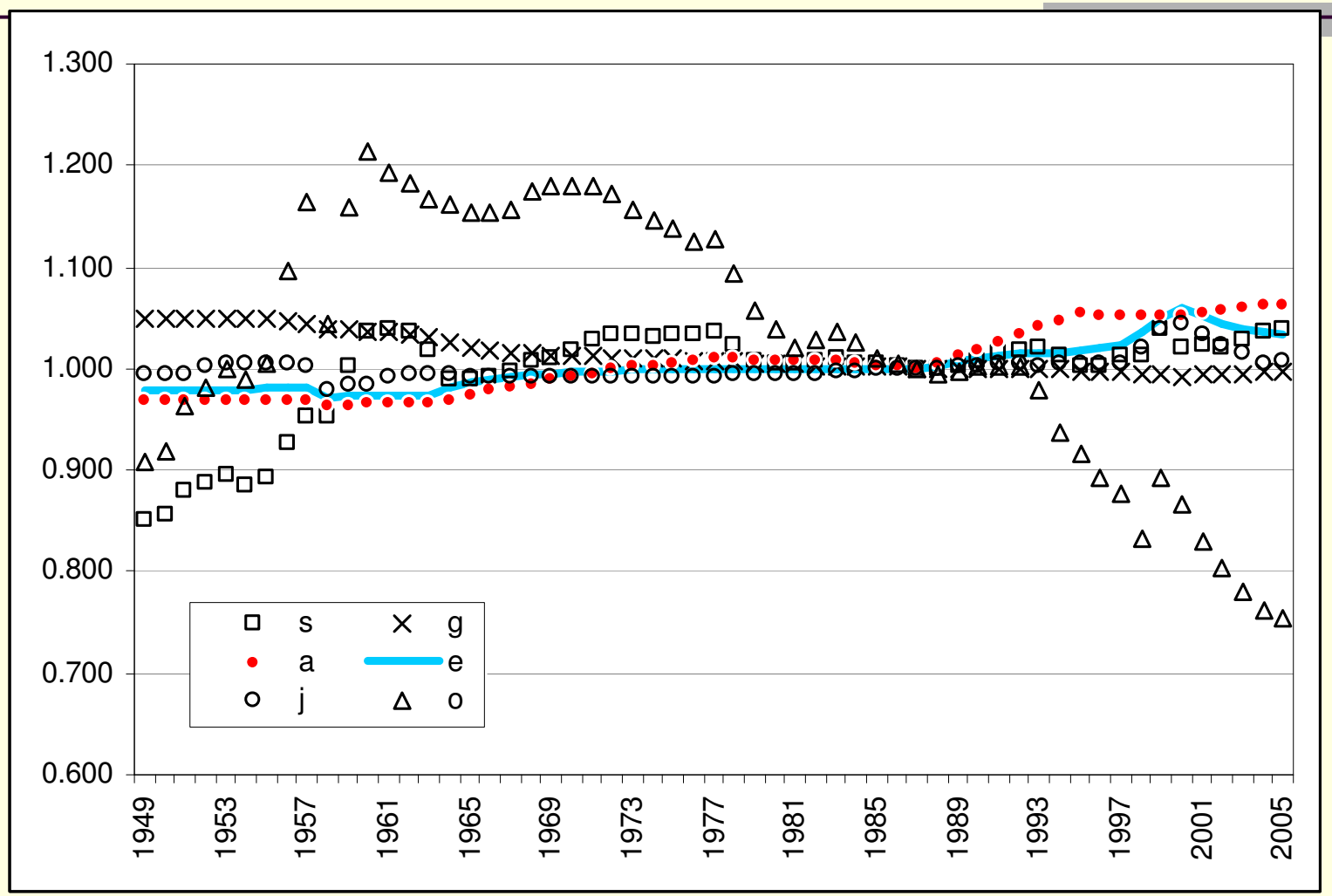
Changes of occupational distribution in Chinese industrial workforce, 1949-2005



Changes of numbers, hours, quality and labour input, 1987=100



Changes of main effects on quality, 1987=100



Education as the main effect and interactive effects with other attributes