CSCanada
Higher Education of Social Science
Vol. 5, No. 1, 2013, pp. 1-5
DOI:10.3968/j.hess.1927024020130501.1083

ISSN 1927-0232 [Print]
ISSN 1927-0240 [Online]
www.cscanada.net
www.cscanada.org

Education, Experience or Discrimination? A Path Analysis of Mobile Population Wage Gaps in China's Urban Labor Market

YANG Yanjun[a],*

[a]The Military Economic Academy, Xiangyang, Hubei, China.
*Corresponding author.

Received 15 December 2012; accepted 14 May 2013

Abstract

For the purpose of estimating the determinants of mobile population wage gaps in China's urban labor market, this thesis conducts a path analysis on the variables which have significant effects on China's labor market using the date provided by the CHIP. The empirical analysis shows that the main factors affecting labor income are workers' education, experience, gender, marital status, industry and occupation. Mobility does not directly affect personal income. But it can impact on the possibility of workers entering monopoly industry and formal occupation, and consequently have an indirect effect on income. Labor market discrimination against mobile population is a significant factor which causes the wage gap in China's urban labor market.

Key words: Education; Experience; Discrimination; Mobile Population; Income Disparities

YANG Yanjun (2013). Education, Experience or Discrimination? A Path Analysis of Mobile Population Wage Gaps in China's Urban Labor Market. *Higher Education of Social Science*, *5*(1), 1-5. Available from: URL: http://www.cscanada.net/index.php/hess/article/view/j.hess.1927024020130501.1083
DOI: http://dx.doi.org/10.3968/j.hess.1927024020130501.1083

INTRODUCTION

Since the reform and opening up, the operating mechanism of China's urban labor market has experienced a significant change and the size of the mobile population increased significantly. The mobile population has become a real part of the urban population. In this process, their employment situation has also been improved, but some problems, such as low quality of employment, low income levels remain widespread. This thesis aims to discuss the determinants of mobile population wage gaps in China's urban labor market.

There are different theoretical models which may be relevant for the analysis of the determinants of labor income. The most commonly used is the Mincer Model which presented by Mincer (1958, 1970). The standard equation of the model shown as follow:

$$\ln(Inc) = a + bSch + cEpr + dEpr^2 + u$$

Where the variable Inc represents the annual income of workers, Sch represents years of education, Epr represents the work experience of workers, a is the intercept of the equation, b is the return to human capital, c and d are two different rates of return to people's work experience, and u is the corresponding disturbance vector.

According to the current literature, we see that most of empirical researches about the worker income are based on multiple linear regression of the Mincer equation. However, this method has two deficiencies: on the one hand, such studies ignore other income-related variables, such as personal capacity and jobs, on the other hand, multiple linear regression can only reflect the direct relationship between the dependent and independent variables.

In fact, the variables that affect the income of workers in many cases are not independent of each other. The independent variable can not only affect the dependent variable by acting directly, but also can have an indirect effect on the dependent variable. To clarify the direct and indirect effects between these variables, we can use the path analysis.

1. THEORY FRAMEWORK

Path analysis is an extension of the regression model, used to test the fit of the correlation matrix against two

or more causal models which are being compared by the researcher. The model is usually depicted in a circle-and-arrow figure just as Figure 1.

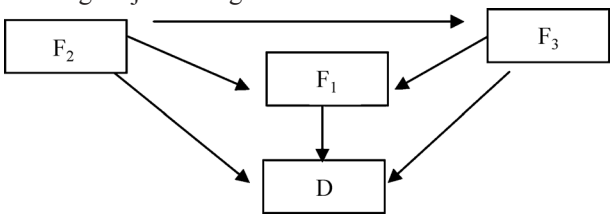

Figure 1
Path Analysis of the Causal Relationship Between the Variables

The single-headed arrows in the figure indicate causation. And the researcher will do a regression for each variable in the model as a dependent on others which the model indicates are causes. The regression weights predicted by the model are compared with the observed correlation matrix for the variables, and a goodness-of-fit statistic is calculated. The best-fitting of two or more models is selected by the researcher as the best model for advancement of theory.

A path model is a diagram relating independent, intermediary, and dependent variables. And the arrows between the variables are causal paths. Single arrows indicate causation between exogenous or intermediary variables and the dependent(s). Double arrows indicate correlation between pairs of exogenous variables.

The model in Figure 1 is a standard path model which has correlated exogenous variables F_1, F_2 and F_3, and endogenous variables D. The causal paths relevant to variable D are the paths from F_1 to D, from F_2 to D, from F_3 to D which represent the direct causes, and the paths reflecting indirect causes which include the paths from F_2 to F_3 to D, from F_2 to F_1 to D. This model is specified by the following path equations:

Equation1: $D=a_{11}F_1+ a_{12}F_2+ a_{13}F_3+e_1$

Equation2: $F_1= a_{21}F_2+ a_{22}F_3+e_2$

Equation3: $F_3= a_{31}F_2 +e_3$

Where a_{ij} are the regression coefficients and their subscripts are the equation number and variable number, thus a_{12} is the coefficient in Equation 1 for variable 2, which is F_2.

2. EMPIRICAL ANALYSIS

2.1 The Date and Variables

The data we used in the paper is provided by the CHIP (Chinese Household Income Project). The purpose of this project was to measure and estimate the distribution of income in both rural and urban areas of the People's Republic of China. Data were collected through a series of questionnaire-based interviews conducted in rural and urban areas in 1988, 1995, and 2002. Individual respondents reported on their economic status, employment, level of education, sources of income, household composition, and household expenditures. To meet the needs of the analysis, we filter the survey samples based on the worker's age, job and residence before the empirical study. Then, we delete the samples which information are missing and finally get 9127 samples.

Table 1
Income Comparison Between Mobile Population and Local Population

Variables	Minimum	Maximum	Mean	Std.Dev.	Percentiles		
					25	50	75
Mobile population	0	100000	9004.77	9160.378	4407	6955	10575
Local population	0	130000	12087.84	8697.796	6400	10198	15000

Table 1 shows the income comparison between mobile population and local population of 9127 samples from CHIP, which means that the average income level of the mobile population is significantly lower than local residents. And even more importantly, the income disparity within mobile population is also significant.

Table 2
The Main Characteristics of the Selected Variables

Variables	Meaning	Type	Value
Income	Annual income of workers	Continuous	
Education	Years of education of workers	Continuous	
Experience	Work experience	Continuous	
Marriage	The marital status of workers	Dummy	Unmarried=1 Married=0
Gender	Workers' gender	Dummy	Male=1 Female=0
Mobility	Population movements	Dummy	Local=1 Mobile=0
Area	The area of population outflow	Dummy	Urban=1 Rural =0
Industry	Industry of workers	Dummy	Monopoly=1 Competitive=0
Occupation	Occupation of workers	Dummy	Formal =1 Informal =0

To estimate the determinants of the income disparities, we select several variables include workers' education, experience, marital status, industry, occupation, gender and area, which may have significant affect on workers' income. The main characteristics of the selected variables are shown in Table 2.

Multiple Linear Regression of the Mincer Equation

To estimate the direct relationship between the variables, we take stepwise multiple linear regression of the Mincer equation. The independent variable is worker's annual income, the dependent variables as shown in Table 2 which include worker's education, experience, industry, occupation, gender, marital status and mobility. The result shows as Table 3. Model 1 is the directly regression analysis results of Mincer equation and model 2 is the results of stepwise multiple linear regression of extended Mincer equation.

Table 3
Multiple Linear Regression of the Mincer Equation

Model		B	Std. E	Beta	t	Sig.	95% Confidence Interval for B	
							Lower	Upper
1	Constant	7.771	0.036		216.017	0.000	7.701	7.842
	Education	0.080	0.002	0.358	37.024	0.000	0.076	0.084
	Experience	0.030	0.003	0.423	11.679	0.000	0.025	0.035
	Experience^2	0.000	0.000	-0.112	-3.089	0.002	0.000	0.000
2	Constant	7.795	0.036		217.245	0.000	7.725	7.866
	Education	0.057	0.002	0.255	25.245	0.000	0.053	0.062
	Experience	0.022	0.003	0.309	7.577	0.000	0.016	0.027
	Experience^2	0.000	0.000	-0.082	-2.121	0.034	0.000	0.000
	Industry	0.265	0.014	0.193	19.285	0.000	0.238	0.292
	Occupation	0.158	0.015	0.106	10.665	0.000	0.129	0.187
	Gender	0.145	0.013	0.108	11.539	0.000	0.121	0.170
	Marriage	0.085	0.024	0.041	3.580	0.000	0.038	0.132

From statistical indicators of model 1 and model 2, we can see model 2 is much better than model 1. And model 2 shows that among the factors listed in Table 2, mobility and area have no direct impact on the annual income of workers. Industry, occupation and gender attributes of workers are the most significant factors that directly affect their income.

2.3 Determinants of Workers' Industry and Occupation

From the analysis above, we can get the conclusion that mobility of the population does not directly affect their income, which is inconsistent with the relevant data in Table 1. Accordingly, we believe there may be intermediate factors, through which the attribute of mobility can have an indirect effect on workers' income. However, industry and occupation can only be used as intermediate variables in all factors. So we analyze the determinants of workers entering monopoly industry and formal occupation respectively.

Table 4
Logistic Model on the Determinants of Workers Entering Monopoly Industry

Variable	Model 1	Model 2	Model 3	Model 4
Constant	-4.471***	-4.890***	-4.891***	-5.088***
	1088.420	600.158	600.280	125.167
Educaiton	0.302***	0.304***	0.303***	0.302***
	1053.902	1062.637	1056.151	1042.567
Experience	0.025***	0.011**	0.011**	0.011**
	98.228	4.397	4.326	3.894
Gender			0.015	
			0.047	
Mobility				9.471**
Mobility 1				0.229**
				4.145
Mobility 2				0.089
				0.488
Mobility 3				-1.377
				0.307
Overall Percentage	69.1	69.0	69.0	69.1
Chi-square	1308.721	1317.038	1317.145	1330.633
-2 Log likelihood	10917.602	10909.285	10909.178	10895.690
Cox & Snell R^2	0.134	0.134	0.134	0.136
Nagelkerke R^2	0.181	0.182	0.182	0.184
HL Test	26.405	21.201	23.290	22.138

Note: The first line of each cell is the estimated parameter of corresponding variable, the second line is Wald statistic, and symbol ***, **, * respectively indicates that the parameter is significant at 1%, 5% and 10% level.

Mobility is a multi-categorical variable which representative of different types of population which includes local rural population, local urban population, mobile urban population and mobile rural population. Therefore, for this variable, we set up three dummy variables include moblie1, mobile 2 and mobile 3, which representative of local rural population, local urban population and mobile urban population respectively.

Table 4 shows the results of Logistic model on the determinants of workers entering monopoly industry. From the statistical indicators of each model, we can learn that Model 4 is the best model fitting the data relationships, which means mobility have significant effects on workers entering monopoly industry. From the analysis above, we believe that workers' industry directly affect the level of income. Consequently, mobility can have an indirect effect on personal income by this way.

Table 5
Logistic Model on the Determinants of Workers Access to Formal Occupation

Variable	Model 1	Model 2	Model 3	Model 4	Model 5
Constant	-2.697***	-4.348***	-4.457***	-4.336***	-3.494***
	445.919	96.336	100.612	95.478	53.751
Educaiton	0.068***	0.065***	0.062***	0.054***	0.075***
	598.481	544.374	487.299	286.359	192.358
Experience	0.207***	0.198***	0.196***	0.199***	0.194***
	525.336	474.700	459.468	469.672	439.139
Mobility		81.890***	82.322***	89.341***	87.189***
Mobility 1		1.834***	1.866***	1.936***	1.908***
		18.119	18.683	20.183	19.328
Mobility 2		0.299	0.347	0.338	0.341
		0.354	0.474	0.453	0.454
Mobility 3		-0.408	-0.400	-0.390	-0.451
		0.480	0.458	0.439	0.583
Gender			0.299***	0.321***	0.332***
			35.541	40.561	43.152
Marriage				0.489***	0.598***
				36.762	51.153
Overall Percentage	74.3	74.6	74.5	74.6	74.7
Chi-square	1022.570	1126.692	1162.212	1198.615	1223.318
-2 log likelihood	9904.257	9800.136	9764.671	9728.214	9703.510
Cox & Snell R^2	0.106	0.116	0.120	0.123	0.125
Nagelkerke R^2	0.152	0.166	0.171	0.176	0.180

Note: The first line of each cell is the estimated parameter of corresponding variable, The second line is Wald statistic, and symbol ***, **, * respectively indicates that the parameter is significant at 1%, 5% and 10% level.

Table 5 shows the results of Logistic model on the determinants of workers access to formal occupation and Model 5 is the best model fitting the data relationships. Model 5 shows that mobility have significant effects on workers access to formal occupation, and affect personal income in the same way of industry.

2.4 The Path Analysis of the Mobile Population Wage Gaps

The preceding analysis means: some of the variables which influence personal income diversity are direct influence factors while another part of them are indirect factors. To clarify the direct and indirect effects between these variables, we can use path analysis.

Table 6
The Result of Path Analysis Model on Workers' Income

	Varaibles	Edu	Exp	Gen	Mar	Occu	Ind	Mob
Direct effect	Ind	A11	A12					A17
	Occu	A21	A22	A23	A24			A27
	Income	A31	A32	A33	A34	A35	A36	
Indirect effect	Income	A11，A36 A21，A35	A12，A36 A22，A35	A23, A35	A24,A35			A17，A36 A27，A35

According to the models in Table 3, Table 4 and Table 5, we draw a table which clearly shows the relationship between the variables just as Table 6. The model in Table 6 is a path model which has correlated exogenous variables Education, Experience, Gender, Marrage, Occupation, Industry, Mobility, and endogenous variables Income. The causal paths relevant to variable Income are the paths from Education, Experience, Gender, Marriage, Occupation, and Industry to Income respectively which represent the direct causes, and the paths reflecting indirect causes which include the paths from Mobility to Industy to Income, from Mobility to Occupation to Income. This model is specified by the following path equations:

Equation1: $Ind=A_{11}Edu+ A_{12}Exp+ A_{17}Mob$

Equation2: $Occu=A_{21}Edu+ A_{22}Exp+ A_{23}Gen + A_{27}Mob$

Equation3: $Income=A_{31}Edu+ A_{32}Exp+A_{33}Gen+$

A_{34}Mar+ A_{35}Occu+ A_{36}Ind

Where A_{ij} are the regression coefficients and their subscripts are the equation number and variable number, thus A_{11} is the coefficient in Equation 1 for Edu. From the model, we can see: Although mobility does not directly affect personal income, it can impact on the possibility of workers entering monopoly industry and formal occupation, and consequently have an indirect effect on income.

CONCLUSION

As we noted at the outset, although the mobile population has become a real part of the urban population, some problems, such as low quality of employment, low income levels remain widespread. According to the current literature, most empirical researches about the worker income are based on multiple linear regression of the Mincer equation. In fact, the variables that affect the income of workers in many cases are not independent of each other. To clarify the direct and indirect effects between these variables, we use the path analysis using the date provided by the CHIP.

The empirical analysis results indicate that the main factors affecting labor income are workers' education, experience, gender, marital status, industry and occupation. Mobility does not directly affect personal income. But it can impact on the possibility of workers entering monopoly industry and formal occupation, and consequently have an indirect effect on income. Labor market discrimination against mobile population is a significant factor which causes the wage gap in China's urban labor market.

REFERENCES

Alicia Adsera, & Barry R. Chiswick (2006). Are there gender and country of origin differences in immigrant labor market outcomes across European destinations? *Journal of Population Economics,* (3), 495-526.

Alwin, Duane F., & Robert M. Hauser (1975). The decomposition of effects in path analysis. *American Sociological Review, 40*(Feb.), 37-47.

Amelie Constant, & Douglas S Massey (2005). Labor market segmentation and the earnings of German guestworkers. *Population Research and Policy Review,* (5), 489-512.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social sociological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182

Becker, Gary S., & Chiswick, Barry R. (1966). Education and the distribution of earnings. *The American Economic Review. 56*(1/2), 358-369.

Ferrer, Ana, & Riddell, W. Craig (2008). Education, credentials, and immigrant earnings. *Canadian Journal of Economics, 41*(1), 186-216.

Lynch, L. M. (1989). The youth labor market in the eighties: determinants of re-employment probabilities for young men and women. *The Review of Economics and Statistics, 71*(1), 37-45.

Nina Martin, Sandra Morales, & Nik Theodore (2007). Migrant worker centers: contending with downgrading in the low-wage labor market. *Geo Journal,* (2), 155-165.

Mincer, Jacob (1958). Investment in human capital and personal income distribution. *Journal of Political Economy, 66*(4), 281-302.

Mincer, Jacob (1970). The distribution of labor incomes: a survey with special reference to the human capital approach. *Journal of Economic Literature, 8*(1), 1-26.

Helena Skyt Nielsen, Michael Rosholm, Nina Smith, Leif Husted (2004). Qualifications, discrimination, or assimilation? An extended framework for analysing immigrant wage gaps. *Empirical Economics, (4)*, 855-883.

Prentice, Ross (1976). Use of the logistic model in retrospective studies. *Biometrics, 32*(3), 599-606.

Schultz, T. Paul (1982). Lifetime migration within educational strata in Venezuela: estimates of a logistic model. *Economic Development and Cultural Change, 30*(3), 559-593.

Smith, Peter M., & Mustard, Cameron A. (2010). The unequal distribution of occupational health and safety risks among immigrants to Canada compared to Canadian-born labour market participants: 1993–2005. *Safety Science, 48*(10), 1296-1303.

Solomon William, Polachek (1981). Occupational self-selection: A human capital approach to sex differences in occupational structure. *The Review of Economics and Statistics, 63*(1), 60-69.

Wright, S. (1921). Correlation and causation. *Journal of Agricultural Research*, *20,* 557-585.

Wright, S. (1934). The method of path coefficients. *Annals of Mathematical Statistics*, *5,* 161-215. The seminal article.