



中西部地区农户收入及其差异 的影响因素分析

内容：

- 问题的提出
- 模型设定及数据说明
- 估计结果
- 讨论

问题的提出

- 农户收入一直以来是农户经济行为研究中的重点问题，同时，随着“三农”问题的备受关注，农民增收已成为各级政府工作的重中之重；与农民增收问题相伴而生的农户收入差异现象在中国也普遍存在。实证研究农户收入及其差异的影响因素，无论对于宏观层面的社会经济发展，还是对于微观层面的有效实现农民增收、缩小农户收入差距，均具有重要的现实意义。

问题的提出

- 综述相关文献可以发现，已有研究大都侧重于分析中国农户收入及差异的现状及其影响因素，但是，这些影响因素在农户收入增长和差异形成中起到怎样的作用？这些作用对地区间以及不同收入组间农户收入增长的影响存在怎样的差异？对这些更为根源性的问题的分析，在现有文献中并不多见。

问题的提出

- 本文在已有研究的基础上，通过进一步的改进和完善，依据农户收入方程，确定农户收入的决定因素，并从中西部地区间差异和高低收入组间差异两个角度，对农户收入差异进行实证研究。本文利用Blinder-Oaxaca分解（Blinder, 1973; Oaxaca, 1973）得到影响农户收入差异的原因，进而深入比较地区间以及高低收入组间上述影响因素对农户收入所产生的效果的异同，并对解决农户增收以及缩小农户收入差距问题提出相应的政策建议。
- 第二部分介绍模型设定和使用的数据，
- 第三部分汇报估计结果，
- 第四部分对估计结果进行讨论分析。

模型设定及数据说明

- 农户家庭经营收入函数表示为：

$$\ln(M) = f(\ln(X_1), \ln(X_2), \ln(X_3), X_4, X_5, X_6, X_8, X_9, X_{10}, X_{11}, D_i)$$

- 依据标准的Mincer（1974）模型，农户工资性收入方程表示为：

$$\ln(W) = f(X_7, X_7^2, X_7^3, X_{10}, X_{10}^2, X_{10}^3, X_5, X_5^2, X_5^3, X_8, X_3, D_i)$$

- 影响农户家庭总纯收入的变量包括上述两部分因素。本文建立农户收入方程的简化式：

$$\ln(I) = f(\ln(X_1), \ln(X_2), X_3, X_4, X_5, X_6, X_7, X_8, X_9, X_{10}, X_{11}, D_i)$$

表1：变量解释及其统计特征

变量名	变量定义	观察值数	均值	标准差	最小值	最大值
I	人均纯收入（元）	802	2087.14	1480.40	-3893.81	13056.04
M	人均家庭经营纯收入（元）	802	2234.93	1815.28	0	19422.27
W	人均工资收入（元）	802	568.43	909.72	0	11750
X ₁	家庭经营投入（元）	802	2799.03	2820.94	10	26624.80
X ₂	生产性固定资产原值（元）	802	4969.83	6888.81	0	77800
X ₃	耕地面积（亩）	802	12.21	16.62	0	94
X ₄	农业投入占家庭经营总投入比例（%）	802	0.97	0.13	0	1
X ₅	粮食播种面积占总播种面积比例（%）	802	0.75	0.25	0	1
X ₆	劳动力最高受教育年限（年）	802	9.22	3.36	0	16
X ₇	平均受教育年限（年）	802	7.26	2.62	0	14
X ₈	农户类型（纯农=0，兼业=1，非农=2）	802	0.46	0.60	0	2
X ₉	是否外出务工（是=1，否=0）	802	0.73	0.44	0	1
X ₁₀	家庭规模（人）	802	4.10	1.28	1	9
X ₁₁	劳动力负担系数	802	1.62	0.55	1	4
D ₁	省份虚变量：内蒙古	802	0.13	0.33	0	1
D ₂	省份虚变量：四川	802	0.13	0.33	0	1
D ₃	省份虚变量：贵州	802	0.16	0.37	0	1
D ₄	省份虚变量：甘肃	802	0.09	0.28	0	1
D ₅	省份虚变量：安徽	802	0.34	0.47	0	1

-
- 本文采用本人参与的对中国中西部地区共6个省（区）15个县（旗）802户农户2004年相关情况调查的数据，其中，西部省份包括内蒙古、四川、甘肃和贵州，中部省份包括吉林和安徽。
 - 对农户工资性收入回归过程中，因对收入变量取自然对数，回归样本仅为工资收入大于0的农户，即因变量是一个限值变量，所以，通过实际发生的农户工资性收入数据分析其影响因素时，必须对没有外出务工的农户观察样本进行审查，以保证对估计结果的无偏性。本文选用Tobit模型进行回归分析。

-
- 家庭经营收入的回归，通过对数据特征值的描述可以看到，观察样本的家庭经营投入均大于0，说明样本农户均从事家庭经营活动；因为有63户家庭耕地面积为0，在回归过程取自然对数时软件系统自动将这部分样本剔除，对数据进行相关统计得知，剩余739户样本中家庭经营收入小于或等于0的样本仅为1户，此时采用Tobit模型的意义不大，因此，对家庭经营收入的回归仍旧采用OLS方法。同理，由于802户总样本中总的纯收入小于或等于0的样本仅为4户，对总纯收入的估计也采用OLS方法，有效样本为798户。
 - 由于耕地面积取自然对数使得家庭经营收入回归中的实际样本仅为耕地面积不为0的农户，但在农户纯收入的回归中，须考虑耕地对外出务工时间和收入的影响，因此，将未取对数的耕地面积作为影响农户纯收入的解释变量，耕地面积为0的农户也被纳入回归样本。

- 农户收入差异分析，即分析收入高低的差距及其原因，可以从地区间差异和高低收入组差异两个角度来进行。本文借鉴Blinder-Oaxaca的分解方法对造成农户收入差异的因素进行分解。收入差异可以表示为：

$$\overline{\ln(I^A)} - \overline{\ln(I^B)} = (\alpha_0^A - \alpha_0^B) + \sum_{i=1}^n (\overline{X_i^A} - \overline{X_i^B}) \alpha_i^A + \sum_{i=1}^n \overline{X_i^B} (\alpha_i^A - \alpha_i^B)$$

Blinder-Oaxaca分解

$$\overline{\ln(Y^A)} - \overline{\ln(Y^B)} = \underbrace{\sum_{i=1}^n (\overline{\ln(X_i^A)} - \overline{\ln(X_i^B)}) \alpha_i^A}_{\text{A: 特征值可以解释的差异部分}} + \underbrace{\sum_{i=1}^n \overline{\ln(X_i^B)} (\alpha_i^A - \alpha_i^B) + (\alpha_0^A - \alpha_0^B)}_{\text{B: 特征值不能解释的部分, 由回归参数的差异造成}}$$

A: 特征值可以解释的差异部分

B: 特征值不能解释的部分, 由回归参数的差异造成

系数估计结果(I)

家庭经营收入 (1a)			工资性收入 (2a)		
解释变量		系数	解释变量		系数
家庭经营投入	$\text{Ln}(X_1)$	0.410	平均受教育年限	X_7	0.999
生产性固定资产原值	$\text{Ln}(X_2)$	0.280		X_7^2	-0.135
耕地面积	$\text{Ln}(X_3)$	0.204		X_7^3	0.009
农业投入比例	X_4	-0.717	家庭规模	X_{10}	0.716
粮食播种面积比例	X_5	-0.545		X_{10}^2	-0.077
最高受教育年限	X_6	0.007	粮食播种面积比例	X_5	8.255
农户类型	X_8	0.018		X_5^2	-6.076
是否外出务工	X_9	-0.088	农户类型	X_8	1.823
家庭规模	X_{10}	-0.211	耕地面积	X_3	-0.076
省份虚变量：内蒙古	D_1	-0.219	省份虚变量：内蒙古	D_1	-1.351
省份虚变量：四川	D_2	-0.151	省份虚变量：四川	D_2	-1.503
省份虚变量：贵州	D_3	0.052	省份虚变量：贵州	D_3	-1.350
省份虚变量：甘肃	D_4	-0.122	省份虚变量：甘肃	D_4	-1.650
省份虚变量：安徽	D_5	0.201	省份虚变量：安徽	D_5	-1.467
常数项	C	5.732	常数项	C	-0.748

系数估计结果(II)

解释变量		总体	按中西部地区回归		按高低收入分组回归	
			西部	中部	低收入组	高收入组
		R_1	R_2	R_3	R_4	R_5
家庭经营投入	$\ln(X_1)$	0.106	0.147	0.068	0.110	0.047
生产性固定资产原值	$\ln(X_2)$	0.002	0.032	-0.014	0.004	-0.014
耕地面积	X_3	0.010	0.003	0.018	-0.002	0.004
农业投入比例	X_4	-0.658	-0.669	-0.509	0.269	-0.445
粮食播种面积比例	X_5	-0.392	-0.462	-0.484	0.108	-0.300
最高受教育年限	X_6	0.017	0.020	0.012	-0.005	0.015
农户类型	X_8	0.082	0.105	0.050	0.068	0.014
是否外出务工	X_9	0.289	0.155	0.366	0.207	0.088
家庭规模	X_{10}	-0.176	-0.157	-0.181	-0.076	-0.094
省份虚变量：内蒙古	D_1	-0.220	0.093		0.040	-0.174
省份虚变量：四川	D_2	-0.356	-0.061		-0.276	-0.081
省份虚变量：贵州	D_3	-0.149	-0.012		0.016	-0.246
省份虚变量：甘肃	D_4	-0.115			-0.096	-0.106
省份虚变量：安徽	D_5	0.106		0.219	0.117	-0.062
常数项	C	7.843	7.239	8.036	6.013	8.503
调整后的 R^2		0.29	0.24	0.28	0.22	0.23

Blinder-Oaxaca 分解结果(I)

解释变量		中部地区为权数				西部地区为权数			
		可解释的部分	比重 (%)	不可解释部分	比重 (%)	可解释的部分	比重 (%)	不可解释部分	比重 (%)
家庭经营投入	$\text{Ln}(X_1)$	-0.0183	-5.56	-0.5518	-167.97	0.0348	-10.60	0.5353	-162.93
生产性固定资产原值	$\text{Ln}(X_2)$	-0.0007	-0.22	-0.3627	-110.41	-0.0030	0.91	0.3664	-111.54
耕地面积	X_3	-0.0733	-22.32	0.1451	44.18	0.0197	-6.01	-0.0916	27.87
农业投入比例	X_4	0.0024	0.73	0.2106	64.10	-0.0035	1.06	-0.2095	63.77
粮食播种面积比例	X_5	0.0306	9.33	0.0833	25.37	-0.0382	11.62	-0.0758	23.08
最高受教育年限	X_6	0.0059	1.80	-0.1162	-35.37	-0.0176	5.35	0.1279	-38.92
农户类型	X_8	0.0077	2.35	-0.0213	-6.49	-0.0164	5.00	0.0301	-9.15
是否外出务工	X_9	0.0277	8.44	0.1555	47.34	-0.0114	3.48	-0.1718	52.29
家庭规模	X_{10}	0.0394	12.00	0.0332	10.12	-0.0414	12.60	-0.0313	9.52
常数项	C	0.0000	0.00	0.7313	222.59	0.0000	0.00	-0.7313	222.59
总量		0.0215	6.55%	0.3070	93.45%	-0.0769	23.41%	-0.2516	76.59%

Blinder-Oaxaca 分解结果(II)

解释变量		高收入组为权数				低收入组为权数			
		可解释的部分	比重 (%)	不可解释部分	比重 (%)	可解释的部分	比重 (%)	不可解释部分	比重 (%)
家庭经营投入	$\ln(X_1)$	0.0061	0.62	-0.4701	-48.06	-0.0143	1.46	0.4783	-48.90
生产性固定资产原值	$\ln(X_2)$	-0.0036	-0.37	-0.1325	-13.54	-0.0009	0.10	0.1370	-14.01
耕地面积	X_3	0.0277	2.83	0.0621	6.34	0.0158	-1.61	-0.1056	10.79
农业投入比例	X_4	0.0105	1.08	-0.7005	-71.63	0.0064	-0.65	0.6836	-69.90
粮食播种面积比例	X_5	0.0215	2.20	-0.3228	-33.00	0.0078	-0.79	0.2935	-30.01
最高受教育年限	X_6	0.0098	1.00	0.1742	17.81	0.0031	-0.32	-0.1871	19.13
农户类型	X_8	0.0011	0.12	-0.0228	-2.33	-0.0056	0.57	0.0273	-2.79
是否外出务工	X_9	0.0044	0.45	-0.0840	-8.59	-0.0104	1.07	0.0900	-9.21
家庭规模	X_{10}	0.0583	5.96	-0.0795	-8.13	-0.0471	4.82	0.0684	-6.99
省份虚变量: 内蒙古	D_1	-0.0042	-0.43	-0.0123	-1.26	-0.0048	0.50	0.0213	-2.18
省份虚变量: 四川	D_2	-0.0023	-0.23	0.0441	4.51	-0.0180	1.84	-0.0239	2.44
省份虚变量: 贵州	D_3	0.0178	1.82	-0.0422	-4.31	0.0065	-0.66	0.0179	-1.83
省份虚变量: 甘肃	D_4	-0.0004	-0.04	0.0085	0.87	-0.0045	0.46	-0.0036	0.37
省份虚变量: 安徽	D_5	0.0053	0.54	-0.0287	-2.93	-0.0161	1.64	0.0395	-4.04
省份虚变量: 吉林	D_6	0.0116	1.18	0.0085	0.87	-0.0034	0.35	-0.0166	1.70
常数项	C	0.0000	0.00	2.4124	246.66	0.0000	0.00	-2.4124	246.66
总量		0.1637	16.74%	0.8143	83.26%	-0.0857	8.76%	-0.8923	91.24%

2009-9-23

结论和评述

就总体而言，农户家庭经营投入、生产性固定资产原值、农业投入占农户家庭经营总投入比例、粮食播种面积占总播种面积比例、耕地面积、劳动力最高受教育水平、是否外出务工以及家庭规模等，均是农户家庭经营收入、工资性收入的重要影响因素，并且这些因素中除生产性固定资产原值外，均直接显著影响农户纯收入的差异。通过总体样本回归以及分地区和收入组回归的比较分析，可以看出农户家庭经营不是影响农户纯收入及其差异的重要原因，工资性收入的多寡已经成为影响农户收入和差异的重要因素，这也是生产性固定资产原值对农户纯收入影响不显著的原因。

结论和评述

Blinder-Oaxaca收入差异分解的结果表明，中国农村居民收入地区间和收入组间的差异绝大部分来自比较对象回归系数的差异。中西部地区差异的14.98%由地区资源禀赋造成，85.02%由回归系数差异造成；高低收入组差异的12.75%由资源禀赋造成，87.25%由回归系数差异造成。导致系数差异的原因很多，主要是地区特点。

结论和评述

中国长期以来存在西部地区农民收入低于中东部地区的现象，由此可以看到，按地区分组与按收入分组回归得到的估计结果存在一定的共性：家庭经营投入、农户类型以及家庭规模三个解释变量的估计系数在对比分析中体现了相似的特征趋势。家庭经营是农户收入的主要来源，家庭经营投入对低收入农户的收入增长比高收入农户产生更大的影响，结合对农户类型系数的分析，进一步说明非农产业是农户增收更有效的途径。家庭人口增多不利于农户人均收入的增加，主要是因为人口越多的家庭，其劳动力负担系数相对较高。

结论和评述

生产性固定资产原值、农业投入占家庭经营总投入比例、粮食播种面积占总播种面积比例、耕地面积、是否外出务工五个变量的回归系数在进行中西部对比和高低收入对比分析中存在差异。针对上述差异的分析表明：①对于中部和西部地区，家庭农业经营均不具有比较优势，农户家庭经营不是影响农户纯收入及其差异的重要原因，工资性收入的多寡已经成为影响农户收入增长及差异的重要原因。②从分地区的回归结果来看，中部和西部地区农户家庭农业经营中的粮食生产均不具有比较优势，而从分收入组的回归结果来看，高收入组农户家庭农业经营中的粮食生产也不具有比较优势，但是，低收入农户经济作物种植比例普遍较低，没有凸显出对农户收入的显著促进作用，粮食播种面积占总播种面积比例的估计系数不显著。③人力资本投入对农户收入增长具有显著的正向影响，受教育水平的提高对于增加农户工资性收入还存在相当大的空间，而对于以种植业为重要农业经营活动的低收入农户，由于受教育水平偏低，且对收入增长的影响并不显著，劳作经验相对于接受学校教育更有利于产出增长。

结论和评述

农户是农村经济活动的行为主体，是广大农村投资、生产和消费等经济活动的微观行为主体，是农业生产中最基本的决策单元。随着宏观经济形势和微观经营环境的变化，农户家庭经济收入增长问题更具有复杂性。本文实证研究表明，农户经济收入增长不仅受其自身物质资本和人力资本积累水平的影响，同时还受家庭经营结构、农业产业内部结构、外出务工、劳动力负担系数等因素的制约，延伸到宏观层面即是受农业结构调整、农村产业升级、劳务经济发展以及当地社会保障等外部环境的影响。由此，合理调整农业结构，促进农村产业升级，积极促进当地劳务经济发展，提高劳动力技能培训和受教育水平，加强农村社会保障力度以减轻劳动力负担等，都是有效促进农民增收、缩小农户收入差异的途径。