

基于 cgss2003数据关于住房面积影响因素的研究报告

李越

论文摘要：本篇文章利用“中国综合社会调查”2003年数据，^① 研究分析了影响受访者住房差异的因素，选取住房使用面积这一变量作为本次研究的自变量，探讨收入水平、婚姻结构和家庭人口数对于它的影响，通过描述、方差分析和多元回归模型的建立，研究发现：三个自变量对因变量都有着显著的影响，但是影响的程度远小于研究之前的设想，尤其是据结果显示收入竟然不是影响住房面积差异的最主要因素，家庭人口数对其影响高于收入因素的影响。对这一结果的讨论和反思，我认为出现原因是由于没有控制地区和城乡差异的影响因素，启示我在以后的研究中要格外注意控制影响因素，尝试和使用嵌套模型。

关键词：住房面积 收入水平 婚姻状况 家庭人口数

一、导言

“人的一生有三分之二的时间是生活在自己的住房里”，住房是人们自我存在的意识的映射和物化，住房消费也反映了人们的需求的特点。^② 尤其是在中国传统观念影响下，住房受到人们的格外重视，大众普遍不愿意长久租房居住，而会为买房努力打拼，将购买住房视为工作生活的重要目标。自21世纪初我国迎来第一批住房贷款热潮，房奴^③已经成为一个新的社会化问题，由此足以看出住房对于中国居民的重要意义。

当前我国面临市场经济转型的新形势，自20世纪80年代以来，市场化改革带来中国社会结构的巨大变迁成为国内外学者关注的热点。围绕中国市场转型与社会分层的关系，学者们进行了卓有成效的研究。其中，大部分研究通过收入决定模型探讨社会分层结构，以及市场机制如何对其产生影响。^④实际上，早期对社会主义体制下的社会不平等研究并非关注收入，而是一些特殊的集体消费品，例如医疗服务、教育和住房。

作为经济体制改革的一个重要部分，中国住房改革已经进行了30多年在改革过程中，住房政策的调整和变革不仅改变了城市住房体系的结构和本质，而且对居民生活产生深远影响。而且，已有研究普遍认为其中一个重要表现就是城镇居民住房不平等不断扩大。基于当前我国经济发展的不平等现状，笔者也认同已有研究中，学者们对于住房状况不平等的判断。而且不平等影响因素分析和探讨资料已有很多，因此本文尝试选取新的角度，研究对居民住房面积的影响因素。

住房面积的大小能够反映住房水平和居住条件的差异，那么住房面积的差异又是由于什么因素造成的呢？本文的研究试图回答：收入水平是不是影响住房面积差异的主要因素？住房面积差异与被访者婚姻状况有没有关系？住房面积是否受到家庭人口数的影响？哪一个变量对于住房的面积差异有最显著的影响？

①“中国综合社会调查”(CGSS2003项目)。该项目由香港特区政府研究资助局资助(CA03P04.SS01),中国人民大学社会学系与香港科技大学社会科学部执行,项目主持人为李路路教授、边燕杰教授。

② 中国城市住宅问题研究会、住宅社会学学术委员会主编,《住宅社会学导论》安徽出版社 1991年

③ “房奴”意思为房屋的奴隶,是指城镇居民抵押贷款购房,在生命黄金时期中的20到30年,每年用占可支配收入的40%至50%甚至更高的比例偿还贷款本息,从而造成居民家庭生活的长期压力,影响正常消费。购房影响到自己教育支出、医药费支出和抚养老人等,使得家庭生活质量下降,甚至让人感到奴役般的压抑。

④ 转引自 胡蓉《市场化转型下的住房不平等基于CGSS 2006调查数据》社会 2012-1 C J S 第3 2 卷

二、文献综述

关于住房差异的已有研究(定量)中,李喜梅就湖北省的普查数据(湖北省“五普”数据)以社会分层视角进行了住房差异状况的研究,主要选取职业分层来透视住房水平的差异,结果表明不同的职业对应着不同的住房水平:国家机关、党群组织、企业、事业单位负责人住房水平最高,其次是专业技术人员、办事人员和相关人员,商业、服务业人员住房水平最低,从而以这种差异揭示了社会制度存在的问题,国家部门人员享受特权的现象严重,其他居民相较而言处在弱势地位^①。

胡蓉在对市场转型下住房不平等的研究中认为:在市场转型过程中,地区市场化的推进有效提高了居民的整体住房水平,但在有房者和无房者之间,市场化的促进作用存在显著分化;与此同时,住房资源的贫富差距会随着地区市场化的推进呈现先扩大后缩小的趋势。^②刘玉亭等人研究市场转型下南京市住房差异发现:2000年以来南京市住房的社会分异现象十分显著,计划经济体制的惯性作用仍然存在,某些制度性因素弱化的同时,户口、单位等制度性因素的作用持续而重要;市场因素对于住房分异的决定作用同样显著,主要体现在不同受教育程度、不同收入群体之间的住房差异。

余腾飞、霍冉在《商场现代化》杂志上发表的文章认为收入水平差异对于我国城市家庭住房状况有着重要影响。^③以上对于住房差异研究资料无一例外显示社会分层中不平等因素(教育程度、职业、收入等)对于住房差异有重要影响,研究成果说服力很强再做后续研究的空间很小,所以我选取住房面积作为因变量来进行具体话的研究和探讨。

三、研究假设

本文在参考关于住房状况研究资料基础上,选取住房面积做具体话研究,结合“中国综合社会调查”2003年数据初步设想了收入水平、婚姻结构和家庭人口数三个可能会具有明显影响的自变量,初步提出以下研究假设以等待数据分析结果的检验:

假设1:收入水平是影响居民住房面积差异的主要因素

假设2:住房面积差异会受到婚姻结构的影响,婚姻状况与住房面积之间有显著关系。

假设3:家庭人口数是影响居民住房面积差异的因素之一,影响程度不好预测,有待于分析成果展示。

四、数据、变量与方法

(一) 数据

本研究的定量分析数据来源于“中国综合社会调查”2003年数据(CGSS2003),样本容量为3500。

(二) 变量

1、因变量、住房面积,选取能够直观反映住房状况的住房使用面积作为因变量,按照惯例对因变量计算对数值,以便使其更接近正态分布。

2、收入水平,这一变量在cgss2003数据中用被访者“全年家庭总收入”这一数据表示。

3、婚姻状况, cgss2003中婚姻状况:未婚=1,已婚=2,离婚后未再婚=3,离婚后再婚=4,丧偶后未再婚=5,丧偶后再婚=6,研究中选取有效值^④并将其转变为定类变量。在描述分析依照发现的规律又转化生成新的二分变量。

4、家庭人口数,使用cgss2003中家庭人口变量,选取有效值重新编码为不同变量。

① 李喜梅《从社会分层看住房差异——对湖北省“五普”资料的分析》社会 2003.7

② 胡蓉《市场化转型下的住房不平等基于CGSS2006调查数据》社会 2012·1 CJS 第32卷

③ 余腾飞、霍冉《我国城市家庭住房状况差异和影响因素分析》《商场现代化》2011年5月(上旬刊)总第646期

④ 转换-重新编码为不同变量去掉无效值。

（三）分析策略

本文研究的数据分析将分为三部分，第一部分是变量转换与描述，将需要用到的变量进行整理和转换，去掉无效值并进行描述展示。第二部分变量间关系的数据描述与方差分析，根据变量性质主要进行婚姻结构与因变量的关系展示和方差分析。第三部分是建立回归模型进行影响结果的展示和讨论。

五、分析结果

本部分将展示数据分析的结果并验证之前提出的假设。

（一）变量转换与描述

1、本文因变量为受访者的住房使用面积，^①去掉未回答、不清楚、不适合之后的无效值，描述统计如表一所示：

描述统计量						
	N	全距	极小值	极大值	均值	标准差
使用面积	3326	1294	6	1300	65.47	59.097
有效的 N（列表状态）	3326					

表一

取对数转化为住房面积后的结果如表二所示：

描述统计量						
	N	全距	极小值	极大值	均值	标准差
住房面积	3326	5.38	1.79	7.17	3.9610	.63981
有效的 N（列表状态）	3326					

表二

可见，自变量的有效样本数为3326，家庭住房使用面积最小只有6平方米，最大达到1300平方米，差异极大，经过取对数转换后，变量分布为近似正太分布，对应数值在1.79到7.17之间。

2、全年家庭总收入：

描述统计量						
	N	全距	极小值	极大值	均值	标准差
收入水平	3252	799900.00	100.00	800000.00	25006.5243	44334.84269
有效的 N（列表状态）	3252					

表三

^① 在此暂不牵扯房屋产权归属问题

有效样本为 3252 个，数值差异也是极大，平均全年家庭总收入为 25006.5 元。

3、婚姻结构，首先要将此变量变为定类变量，再去掉无效值，基本频率分布情况如表四：

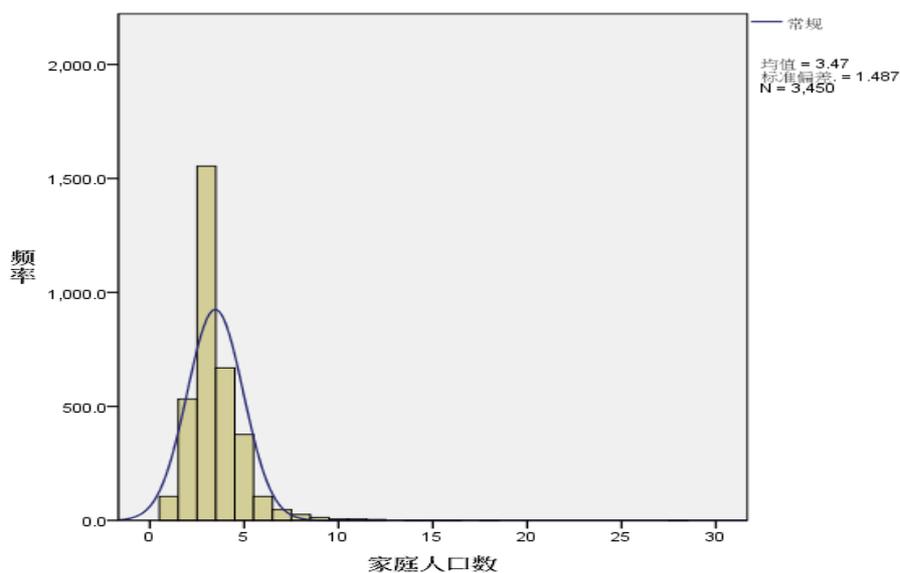
统计量		
婚姻状况		
N	有效	3486
	缺失	14

婚姻状况					
		频率	百分比	有效百分比	累积百分比
有效	1	350	10.0	10.0	10.0
	2	2893	82.7	83.0	93.0
	3	65	1.9	1.9	94.9
	4	52	1.5	1.5	96.4
	5	109	3.1	3.1	99.5
	6	17	.5	.5	100.0
	合计	3486	99.6	100.0	
缺失	系统	14	.4		
合计		3500	100.0		

未婚=1，已婚=2，离婚后未再婚=3，离婚后再婚=4，丧偶后未再婚=5，丧偶后再婚=6

表四

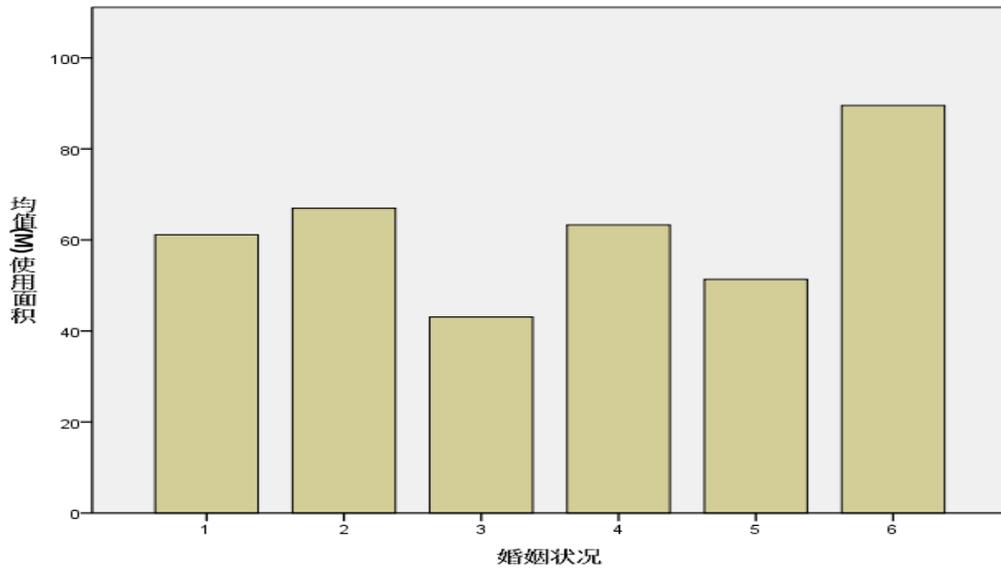
4、家庭人口数的转换与描述结果如下图：



根据家庭人口数的直方图描述可以看出家庭人口数比较集中的分布在 1 至 7 人之间，以 3 为众数。

(二)变量间关系的数据描述与方差分析

根据变量性质，在自变量中只有婚姻状况是定类变量，在数据描述分析中首先选取婚姻状况这一变量与因变量住房面积的关系进行初步直观的分析，做直方图如下：



在上图中，并不容易看出有什么显著差异，但是又可以发现存在着一定的规律：1、3、5表示的三种情况的单身，2、4、6代表三种情况的非单身^①，这提示我依据单身与否将婚姻状况转换为二分变量之后做 T 检验^②后得知 524 人单身和 2962 人非单身的统计情况下 F 值 4.218，sig 值为 0.040<0.05，所以是否结婚对于住房面积有显著影响。进一步研究细分婚姻状况与因变量之间的关系可以做方差分析，原假设婚姻状况对住房面积没有影响，备选假设为有影响，如表五所示：

ANOVA					
住房面积					
	平方和	df	均方	F	显著性
组间	19.695	5	3.939	9.735	.000
组内	1337.618	3306	.405		
总数	1357.313	3311			

表五中可以看出：组间离差平方和仅为 19.695，在总离差平方和中所占比例很小，相关比率为 0.01451 因此婚姻状况对于对于住房面积差异有直接影响，但是影响并不大，不构成住房面积差异的主要影响因素。

表五

（三）回归模型结果与多元回归模型建立

1、单变量回归模型分析^③：

（1）研究收入因素对于住房面积的影响时，由于原统计数据采用收入单位为元，家庭全年总收入数额较大，为保证回归预测准确性将收入单位转化成万元^④，建立线性回归模型。

① 即存在事实婚姻关系

② 因为论文篇幅的限制，在此不对 T 检验的表格进行详细呈现

③ 由于论文篇幅的限制，单变量的回归模型图表先不做展示，只描述结果。

④ 计算变量中将收入水平除以 10000

从回归模型的检验来看，收入水平与因变量住房面积的线性回归关系中，相关系数为 0.114，表示了低于 0.01 水平（双侧）上显著相关^①程度为 11.4%，但是决定系数（R 方）值只有 0.013，也就是说这个回归关系只能解释因变量 1.3% 的变异。方差分析结果 sig 值 0.00 显示存在显著线性关系。最后的检验结果显示，全年家庭总收入每增加 1 万元，住房面积增加 0.017 平方米，次检验结果与原假设差异很大，其原因将在下文中进行讨论。

(2) 家庭人口数与住房面积建立回归模型。从回归模型的检验来看，收入水平与因变量住房面积的线性回归关系中，相关系数为 0.286，表示在低于 0.01 水平（双侧）上显著相关^②程度为 28.6%，决定系数（R 方）值为 0.082，也就是说这个回归关系能解释因变量 8.2% 的变异。方差分析结果 sig 值 0.00 显示存在显著线性关系。最后的检验结果显示，家庭人口数每增加 1 人，住房面积增加 0.123 平方米。

2、多元回归模型的建立与结果展示：

输入 / 移去的变量 ^b			
模型	输入的变量	移去的变量	方法
1	家庭人口数, 收入, 婚姻		输入

a. 已输入所有请求的变量。
b. 因变量: 住房面积

模型汇总				
模型	R	R 方	调整 R 方	标准 估计的 误差
1	.293 ^a	.086	.085	.60534

a. 预测变量: (常量), 家庭人口数, 收入, 婚姻。

Anova ^b						
模型		平方和	df	均方	F	Sig.
1	回归	104.475	3	34.825	95.038	.000 ^a
	残差	1116.528	3047	.366		
	总计	1221.004	3050			

a. 预测变量: (常量), 家庭人口数, 收入, 婚姻。
b. 因变量: 住房面积

系数 ^a					
模型	非标准化系数		标准系数	t	Sig.
	B	标准 误差	试用版		

① 双变量相关测得的结果，P 值小于 0.001，相关性是显著的

② 双变量相关测得的结果，P 值小于 0.001，相关性是显著的

1	(常量)	3.542	.030		118.347	.000
	收入	.014	.003	.095	5.450	.000
	婚姻	-.108	.032	-.059	-3.372	.001
	家庭人口数	.114	.008	.256	14.645	.000
a. 因变量: 住房面积						

结果显示,三个进入自变量都没有被剔除,进一步验证是哪个自变量对因变量有显著影响,三个自变量影响相关系数为 0.293,能解释因变量 8.6%的变异,控制了婚姻和家庭人口数两个变量,全年家庭收入每增加 1 万住房面积增加 0.014 平米,单身比非单身情况的住房面积少 0.108 平米,控制其他两个变量,家庭人口增加 1 人住房面积增加 0.14 平米。

总体上看,所选取的三个自变量收入水平、婚姻结构和家庭人口数对于因变量(住房面积)都有着显著的影响,影响程度有所差异,其中家庭人口数对于住房面积的影响最大,其次才是收入,影响最小的是婚姻状况^①。

六、讨论与反思

经过以上分析结果的展示,验证了三个自变量收入水平、婚姻结构和家庭人口数对于因变量(住房面积)都有着显著的影响的同时比较意外的是结果与研究假设出入较大,从研究结果回应研究假设来看:

假设1:收入水平是影响居民住房面积差异的主要因素。得以验证收入的确影响住房面积,收入提高有助于住房面积的增大,但是收入并不是影响的主要因素,这一假设被推翻。

假设2:住房面积差异会受到婚姻结构的影响,婚姻状况与住房面积之间有显著关系。这一假设在数据分析时经过方差分析被证实,有现实婚姻关系住房面积要大。

假设3:家庭人口数是影响居民住房面积差异的因素之一,影响程度不好预测,有待于分析成果展示。在经过分析后显然这一假设也被证实,而且比较意外的在三个自变量中家庭人口数的影响力最大。

由于对研究结果的意外,我经过反思认为:在结果中收入影响力出乎意料的小主要是因为我在进行数据研究和回归模型建立时,没有考虑到地区差异和城乡很大的因素影响,而这两个因素又是对于住房面积差异有极其重要影响的,尤其是城乡差异!我们都知道城乡收入水平差异巨大,城市人均收入明显高于农村居民人均收入,但是,在住房状况中“从事农业生产的农村居民人均住房面积和人均房间数上普遍高于其他职业”^②,农村丰富的土地资源和廉价的土地价格是这种现象的根源。

我没有控制这样重要的城乡差别影响因素是造成意外结果的原因,也是本次研究的一个很大的不足,这一点有待老师的批评指导同时启示我在以后的研究中要格外注意控制影响因素,尝试和使用嵌套模型。

① 婚姻状况单变量回归模型中 R 方值只有 0.008

② 李喜梅 《从社会分层看住房差异——对湖北省“五普”资料的分析》 社会 2003.7

【参考文献】

- 1、中国城市住宅问题研究会、住宅社会学学术委员会主编，《住宅社会学导论》 安徽出版社 1991 年
- 2、胡蓉《市场化转型下的住房不平等基于 C G S S 2006 调查数据》社会 2012·1 CJS 第32卷
- 3、李喜梅 《从社会分层看住房差异——对湖北省“五普”资料的分析》 社会 2003.7
- 4、余腾飞、霍冉《我国城市家庭住房状况差异和影响因素分析》 《商场现代化》2011 年 5 月（上旬刊）总第 646 期
- 5、刘军民(2006)《对中国住房制度改革若干政策的反思》《中国经济时报》 7 月 17 日

文章来源：作者投稿