

# 收入不平等会对消费和支出产生影响吗？

## ——来自中国城镇住户调查的证据

### 摘 要：

收入不平等是如何影响消费和其他支出的？这是近来有关收入不平等的研究中比较新颖的话题。本文在以往研究的基础上，以中国城镇住户调查的数据作为材料，从实证的角度探讨了收入不平等向消费的传导机制。研究有以下发现：在 1988 年到 2006 年期间，中国城镇住户收入和消费不平等水平上升非常明显；中国城镇住户的收入和消费不平等相合度很高，这与国外文献中的模型和数据有出入；我们探讨了收入不平等的来源，对收入不平等进行了分解，并发现近来组间收入不平等的来源主要是教育和地区的差异；收入和收入不平等对非耐用品消费、教育投资、银行储蓄和非储蓄型投资的影响各不相同。收入不平等的存在削弱了城镇居民，尤其是低薪阶层的消费意愿。低收入群体有很高的意愿通过储蓄、投资等方式来防御收入波动风险，减小与高收入群体的收入差距。

**关键词：**收入不平等 消费不平等 不平等分解 家庭支出

# 目 录

一、引言 .....	1
二、文献综述 .....	1
三、收入与消费的统计性描述 .....	2
3.1、收入与消费的年度基本情况.....	3
3.2、收入与消费不平等.....	4
四、实证研究 .....	6
4.1、收入不平等分解 .....	6
4.2、基于组间不平等的样本分组 .....	7
4.3、收入及收入不平等对消费的影响.....	11
4.4、收入不平等与储蓄、投资 .....	13
4.5、分位数回归比较 .....	14
五、结论与讨论.....	16



# 一、引言

自经济体制改革以来，我国连续 30 多年都保持着高速的经济增长。在所有的综合指标上，如 GDP、FDI、进出口总额上，中国一直保持着惊人的增长。但与此同时，许多统计机构和学术研究发现：伴随着经济规模的高速增长，中国家庭收入正在经历着严重的分化，不平等水平日益加剧。有研究表明，中国在过去 20 年内都是世界上收入不平等现象最严重的国家，Li Gan（2013）基于中国家庭金融调查（China Household Finance Survey）的研究显示中国总体的基尼系数已经高达 0.61。从地域空间上看，这种收入不平等体现在多种层次上，如城镇与农村之间、内陆与沿海之间，并且从时间维度上来看也呈现出明显的波动性（Kanbur and Zhang 1999, 2005；Wu and Perloff, 2005）。有基于国家统计局公布调查数据的研究表明，自 1992 年以来，我国的总体基尼系数始终不低于 0.4，其中城镇内部的基尼系数是最重要的影响全国基尼系数的因素（程永宏，2004）。

有关收入不平等的讨论在学术界已经有着非常丰富的成果。对于导致收入不平等的原因，学术界一直以来有许多富有见地的讨论：有学者认为中国的收入不平等是市场高效率配置的结果，可以看做是市场化经济成熟的必然结果（Gan, 2013）；另有研究发现城镇内部收入不平等的加剧很大程度上是由 1992 到 1997 年期间的一系列经济结构性转型，特别是国有企业改革、城镇化、全球化等因素导致的（Cai, Chen and Zhou, 2010）。但与此同时，有关消费、储蓄、投资等支出项目的不平等水平却在很长一段时间内没有引起足够的重视。近年来，研究消费不平等、收入不平等向消费不平等的传导机制的文章也逐渐丰富了起来，补充了这方面的空白。本文在这些研究的基础上，进一步做了如下的工作：在第二章我们对以往的文献进行了综述；在第三章我们呈现了中国城镇住户调查数据下中国的收入和消费分布情况，计算了收入和消费历年的不平等水平变化；在第四章我们对收入不平等指标进行了解构，进一步通过回归分析检验了收入、收入不平等对消费、储蓄、投资等各项支出的影响；在第五章我们总结了全文的主要发现和需要进一步探索的问题。

# 二、文献综述

目前国内外都有非常丰富的关于探讨中国收入不平等的学术成果，无论是对其的统计性分析还是机理性探讨。为了衡量收入不平等水平，研究者提出了许多的度量方法（这些度量的指标同时也被应用到了对消费不平等的讨论上）：其中使用广泛的有基尼系数，百分之五十分位点与百分之十分位点（50/10），这两种度量方法对于中等收入层次的结构变化敏感，对于两端（高收入和低收入群体）的收入变化更加稳健；此外，90/10 分位点比值，赛尔指数则是更好地侦测到了两端收入不平等的变化，并且赛尔指数在数学上具有组内组间的可分解性，这对于此类研究中的分组研究方法带来了极大的便利（Conceição, 2000）。除此之外，还有学者开始从其他维度研究收入分配是否均衡的问题，如 Yingqiang Zhang 和 Tor Eriksson（2010）将研究的对象转移到了机会公平上，他们认为中国的家庭财富不平等累计主要是由机会不平等（上一辈的财富和工作）引起的。

相较而言，对于消费不平等的研究起步就略显延迟。为什么我们需要研究消费的特征和分布？首先，消费比收入更能反映家庭的真实福利，从而可以对住户的真实生活水平有更加细致的刻画（邹红，李奥蕾，喻开志，2013）；其次，收入不平等并不会完全等价地转化为消费上的不平等，社会上信用、借贷市场的存在，会对个体之间、个人在整个生命周期内的消费行为产生广泛的影响（Krueger, 2006）。国外对于消费不平等的研究起步较早，Deaton（1994）和 Paxson 在之前发表的文章中首次在研究区域组经济福利时，同时使用收入和在生命周期理论假设下的消费不平等，在生命周期理论假设的框架下，来作为衡量福利的标准。

在这之后国外学术界在这个领域的钻研越来越深刻。在实证研究领域比较流行的一种研究范式是将不平等分解为组间不平等效应和组内不平等效应,这种方法在本部分接下来综述的论文中也多有应用。曾经也有学者以美国的数据为材料,讨论了近年来一直提高的不平等水平究竟是由组内(外部特征相似的家庭组)还是组间效应引起的,并强调了组间效应在扩大福利差距的过程中所起到的作用(Johnson,Shipp,1997)。也有学者尝试通过收入不平等来解释消费不平等,探索从收入到消费的传导机制。Kocherlakota 和 Narayana (1996) 曾经的研究建立了一套简单的双人数学模型,来解释在时间截面水平下,收入不平等到消费不平等的传导机制:他们通过同一时期内劳动性收入的方差来刻画收入不平等。不同收入层次的群体通过社会保险借贷体系来进行富人和穷人之间的消费转移,形成一种在每一期都成立的协作契约,从而使得消费会比收入更加的平滑。Dirk Krueger 和 Fabrizio Perri (2006) 在后来进一步的研究中将这种模型扩大到了多人市场中,并分别推导了在约束性债务市场假设下和标准非完全市场假设下模型的约束条件和均衡条件,并用 PSID 数据对其进行了检验。Blundell (2008) 等在之后研究了通过对个人收入的研究,认为收入中永久性冲击的变化是导致收入与消费不平等产生分歧的原因。同时,他们在对 PSID 的研究中发现部分人群,尤其是高学历者和临近退休者,对于永久性收入会采取部分保险的策略;而贫困人群往往对临时性收入采取完全保险分摊的策略。

当前的中国作为世界第二大经济体,收入和消费的不平等一直以来也是广大学者关注的话题。围绕收入不平等水平展开的话题有很多:王洪亮和徐翔(2006)的研究发现,中国收入不平等的来源主要是源自城乡间的收入差距,而不是地区间的不同。城乡间收入不平等的变动在收入不平等变动的各个阶段都占据主导地位。李国正,艾小青和李晨曦(2017)的研究认为,中国的劳动力市场上还存在性别歧视,导致女性的就业率和求职率均普遍低于男性,并且可以解释大部分未能被观测数据所解释的男女收入差距。除了收入不平等的话题外,在微观层面上研究支出、以及消费对支出的影响方面,如消费,储蓄,投资等问题上,近年来也有了令人瞩目的发展。通常引起关注最多的是“中国高储蓄率谜题”,有学者认为这是因为中国在 90 年代以来国有企业改革、下岗潮等因素的影响下,居民有较高的预防性储蓄动机,使他们在面对经济冲击时能保持平稳的消费(Meng, 2003)。杨俊(2008)等在研究中指出收入不平等会使得教育不平等扩大,而且教育不平等的改进并没有显著抑制住收入不平等的扩大。在研究消费方面,Hongbin Li (2011) 等通过截面数据实证了收入差距的存在会通过作用于人们寻求更高社会地位的心理,从而对他们在教育支出、显性商品(如珠宝首饰等奢侈品)的购买行为上产生影响,进而影响消费。

### 三、收入与消费的统计性描述

在这一部分,我们通过对观测得到的数据进行统计分析,来初步呈现我国城镇住户内部不平等水平的概貌。首先,我们介绍一下这个部分以及接下来的实证分析部分所使用的数据。我们所使用的数据来自中国城镇住户调查(UHS),顾名思义,调查的对象仅限于城镇居民,因此本文的研究对象无法涉及到城乡收入不平等话题。这项调查由中国国家统计局进行,截止 2006 年一共进行了五轮调查。该调查自 1986 年起,覆盖了全国绝大部分的省市地区,通过统计抽样和多步骤选取样本的方式抽取了家庭样本。该调查提供了中国城镇家庭在收入、支出方面详实的数据。该调查所产生的数据是一个轮换的面板结构,每年会有三分之一的样本更新,并且每一轮之间也会完全更新样本。但由于提供的调查数据中缺少能跨期识别家庭个体的变量,所以该我们本文采用横截面数据的处理方式。在本文中有权使用的数据是来自第二到第五轮(1988 至 2006 年)的调查数据。

#### 3.1 收入与消费的年度基本情况

我们总结了从 1997 年到 2006 年各家庭收入各项收入、消费的历年平均值、中位数、偏度和峰度水平，从而可以大致了解历年城镇家庭收入消费水平和分布特征。为了保证历年数据的可比性，我们对原始数据做了如下的调整：一、根据国家统计局公布的年度 CPI 指数，我们对每年的各项收入、支出的数据，以 1997 年为基准年份进行了调整。二、中国城镇住户调查是以城镇家庭为单位进行的数据搜集工作，家庭的收入、支出规模显然会受到家庭规模和家庭人口结构的影响，因此根据国际上流行的等值因子对家庭人口规模进行调整。在本文中户主的权重为 1，非户主的成年人权重为 0.7，未成年人权重为 0.5，该种调整方法应当和其他的家庭规模调整方法会得出相似的估计（Atkinson A 等，1995）。以该家庭等值规模的计算结果为基准，求出了家庭的各项人均数据的统计指标。

表3-1 历年收入消费统计

年份	总收入	总消费	年份	总收入	总消费
1997	8951.555	7436.906	2002	10442.96	7859.036
N=3206	7537.441	6142.485	N=20534	8905.334	6440.743
	2.813471	2.575008		3.125952	5.382221
	17.07987	14.17365		24.68666	65.4678
1998	9336.822	7786.909	2003	11260.15	8177.897
N=3109	7801.197	6279.274	N=23476	9435.554	6703.833
	2.545589	3.73988		4.342206	5.920172
	13.99882	29.59604		46.40536	78.35149
1999	10255.19	8347.404	2004	12973.04	9127.865
N=3026	8511.525	6825.65	N=24865	10553.54	7340.482
	2.352676	3.520189		7.293258	5.630606
	12.48502	23.38247		148.7336	62.34031
2000	11002.65	8973.019	2005	14342.64	9918.135
N=3115	9142.065	7379.899	N=25664	11644.03	7954.172
	2.841642	7.132999		7.973447	6.382155
	19.30983	118.0665		167.0907	95.44494
2001	12085.98	9427.379	2006	15995.07	10843.39
N=2790	9822.772	7726.653	N=23946	12997.2	8659.426
	2.584963	5.089319		4.868836	5.992274
	15.32451	64.94043		56.58218	80.25378

注：所有的名义值都根据历年CPI，以1997年为基准年调整。

由表 3-1 我们可以看出：1、无论是从收入还是消费的角度讲，中国城镇居民的均值水平和中位数水平都是在不断提高的。这间接印证了 Hongbin Cai（2010）等的研究结论：城镇居民的福利都在 21 世纪初左右这段时间内真实地增长了。2、历年来无论是收入还是消费，均值水平都大于中位数水平，分布的偏度也都大于零，并且随时间有增长的趋势，这代表收入、消费分布呈现右偏的特点，这很可能是由于极端富有家庭的。3、历年来无论是收入还是消费，峰度的值都显著大于 3，收入和消费分布呈现出尖峰厚尾的特点。通过对以上统计量的解读，我们大致总结出：过去的这段时间内，居民的真实福利水平都在提高，但收入分配、消费支出的对称性和集中度越来越差，由峰度的特征可以看出来分布具有厚尾的特征，并且这很有可能是非常富有的群体造成的。

## 3.2 收入与消费不平等

在本节，我们将对城镇住户的不平等水平（主要是收入和消费不平等）做更细致的统计刻画。首先，我们阐释一下本文中对收入和消费的定义。虽然中国城镇住户调查具有滚动的短面板结构，但自从 2002 年（第五轮调查）开始才在公布的数据中提供了能跨期识别的住户代号，这不利于我们分析更长时间段内的不平等水平演化发展，因此我们接下来的分析会将数据当作截面数据使用。

基于上述的模型，为了尽可能地避免在下一章实证研究部分可能出现的内生性问题，我们接下来对收入的定义如下：与 Dirk K 和 Fabrizio P（2006）的研究相似，为了尽可能地避免在下一章实证研究部分可能出现的内生性问题，我们衡量所有收入中对消费、储蓄决策外生的那一部分，我们将劳动性收入定义为：调查数据中各种所有制职工工资，个体经营者净收入，个体被雇者收入，离退休再就业人员工资，其他劳动收入，离退休金收入，转移性收入的总和。相应的，为了对应我们文章采用的截面数据分析方法，我们将本文中重点研究的消费定义为流动性消费，其内容包括：食品消费，衣着消费，非耐用家庭设备、用品及服务，医疗保健，交通通讯消费，娱乐文化消费和其他非耐用品消费。本文可以避免了对耐用品消费（不平等）的研究，原因如下：个体家庭的耐用品购买策略会面临更长周期的预算约束，需要考虑到家庭的收入波动和风险的度量问题。曾经有学者根据持久性收入假说，尝试用过去几年的收入的加权平均值来估计家庭的持久性收入，使用 probit 模型来度量家庭的事业风险，一次来衡量一个家庭的长期预算约束（Meng, 2003）。本文所采用的 UHS 数据由于其截面数据的结构特征，我们难以对收入的冲击和风险、以及家庭面临这种收入波动时的消费避险行为做出准确的描述（Blundell, 2008）。在本文接下来的部分，如果没有特殊情况，以上定义的劳动性收入和流动性消费将会简称为收入和消费。

根据以上对收入、消费的定义，我们使用 1988 年到 2006 年的每年的年度数据，通过基尼系数、赛尔指数、90 分位数与 10 分位数比值、中位数与 10 分位数比值来刻画每年城镇居民的收入、消费不平等水平。

首先，我们来简单介绍一下基尼系数和赛尔指数的计算原理。

基尼系数是意大利统计、社会学家 Corrado Gini 在 1912 年提出的一种衡量国家居民收入和财富分散程度的一种估计量，也是目前在研究不平等问题中使用最广泛的指标。基尼系数一般来讲是通过洛伦兹曲线定义的，其值等于 45 度直线和累计收入曲线之间围成的面积除以两交点向坐标轴垂直作线形成的三角形的面积。根据 Amartya Sen 的研究（1973），基尼系数的等价计算方法（与通过洛伦兹曲线定义等价）是，计算平均绝对差距的一半。若将居民的收入分布看作是离散的，则

$$Gini = \frac{1}{2\mu} \sum_{i=1}^n \sum_{j=1}^n p(x_i)p(x_j) |x_i - x_j| \quad (1)$$

$$\mu = \sum_{i=1}^n x_i p(x_i) \quad (2)$$

其中， $n$  表示收入组的数量， $p(x)$  是居民收入分布的概率质量函数， $x_i$  表示第  $i$  个单位个体的收入或财富。若将居民收入的分布看作是连续的，则

$$Gini = \frac{1}{2\mu} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x)f(y) |x - y| dx dy \quad (3)$$

$$\mu = \int_{-\infty}^{+\infty} x f(x) dx \quad (4)$$

其中， $f(x)$  是收入分布的概率密度函数<sup>1</sup>。这种等价的计算方法在计算机算法编写中更为常用。

赛尔指数是荷兰计量经济学家 **Henri Theil** 提出的另一种衡量经济不平等水平的广义熵指数。由于这种指数直接在数学上具有组间和组内的可分解性，目前在学术论文中的应用也越来越广泛。赛尔指数有 t 型指数和 l 型指数两种，其中 t 型指数对高层次水平，即分布右端，的变化更为敏感<sup>2</sup>。这更符合我们在上一节中的数据统计特征，因此本文将主要采用 Theil-t 指数作为衡量经济不平水平的指标，若将居民的收入分布看作离散的，则其计算方法如下：

$$T_t = \sum_{i=1}^n p(x_i) \frac{x_i}{\mu} \ln\left(\frac{x_i}{\mu}\right) \tag{5}$$

其中， $n$  表示收入组的数量， $p(x)$  是居民收入分布的概率质量函数， $x_i$  表示第  $i$  个单位个体的收入或财富。 $\mu$  的定义与基尼系数的计算相同，是样本平均值。若将居民收入的分布看作是连续的，则

$$T_t = \int_{-\infty}^{+\infty} f(x) \frac{x}{\mu} \ln\left(\frac{x}{\mu}\right) dx \tag{6}$$

其中， $f(x)$  是收入分布的概率密度函数， $\mu$  的定义与基尼系数的计算相同，是分布的数学期望。

根据以上的指标定义，我们计算了从 1988 到 2006 的收入、消费不平等水平<sup>3</sup>：

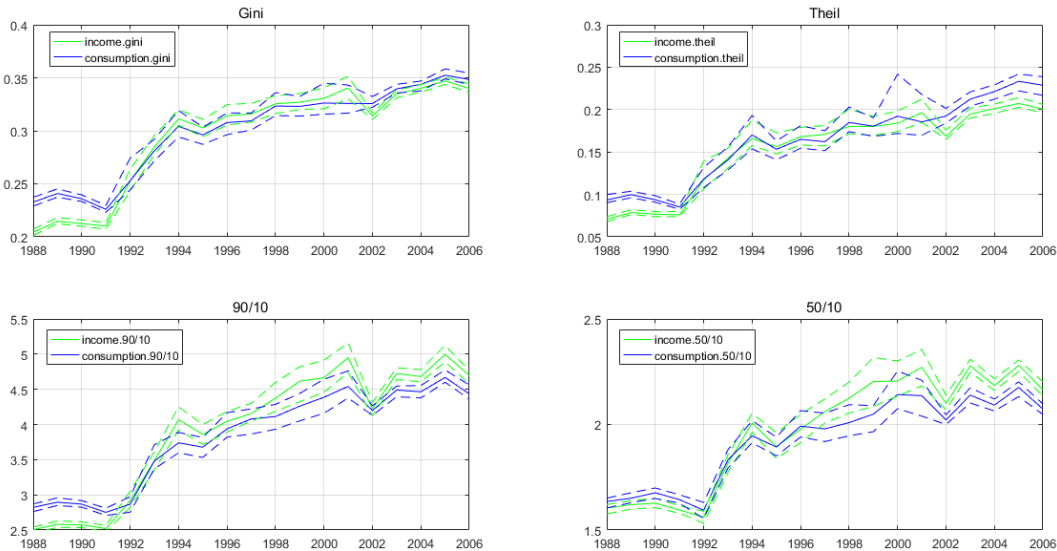


图 3-1 收入与消费不平等趋势

图 3-1 展示了我们通过四种指标，由 1988 年到 2006 年截面数据计算而得的居民不平等水平。左上角的是基尼系数，右上角是赛尔指数，左下角是 90 分位数比 10 分位数，右下角

<sup>1</sup> 我们此处将负债等情况视作负资产，因此参与积分的集合在这里定义为整个实数轴。以下赛尔指数计算的处理与此类似。  
<sup>2</sup> “Inequality Measures”. [www.urban.org](http://www.urban.org). Urban Institute. Retrieved 20th February 2018.  
<sup>3</sup> 所有不平等指标的计算均通过 Matlab 软件完成，部分计算的包是引自 Liber Eleutherios 编写的 “Inequality Package”。



是中位数比 10 分位数。绿色、蓝色的实线分别是代表收入、消费不平等水平；绿色、蓝色的虚线分别是代表收入、消费不平等水平的百分之 95 的置信区间<sup>4</sup>。从图中我们可以看出：

一、无论是用哪一种不平等水平指标来衡量，中国城镇居民的不平等水平都发生了巨大的增长：以基尼系数为例，劳动性收入不平等水平从 1988 年的 0.21 上升到了 2006 年的 0.34，增长率高达 61.9%，消费不平等水平也从 1988 年的 0.23 增长到了 2006 年的 0.35，增长率也达到了 60.9%。

二、观察收入和消费的整个变化过程，尽管在每一年都会略有起伏，总体都保持着上升的态势。其中从 1991 年到 1994 年发生了跳跃式的增长：以基尼系数的变化为例，三年内收入基尼系数从 0.21 变化到了 0.31，增长了 47.6%；消费基尼系数从 0.22 变化到了 0.3，增长了 36.4%。据我们推测可能的原因是 90 年代初国企改革导致的下岗潮，使得在岗员工与下岗员工之间劳动性收入的差距变大（Meng, 2003）。

三、UHS 的数据显示，中国城镇居民的收入和消费不平等水平的变化从 1988 年到 2006 年都是高度一致的。值得注意的是，根据 Kocherlakota 等（1996）和 Dirk Krueger 等（2006）提出的收入消费传导模型，截面水平上的消费不平等水平应当是低于收入的不平等水平，这种假设在对西方国家的调查数据进行实证研究时得到了很好的印证，如 Dirk Krueger（2006）等对美国上世纪八九十年代数据的研究，Pendakur（1998）对加拿大 1978 年到 1992 年数据的研究。然而，我们对数据的研究并没有发现这种特征，从基尼系数的演化过程来看，中国城镇居民收入和消费的变化几乎是完全同调的。从赛尔指数来看，1988 年到 1991 年、2002 年到 2006 年这两短时间内，收入不平等水平反而低于消费不平等。只有 90/10、50/10 这两种不平等测量指标显示消费不平等是略微低于收入不平等的水平。这与国内其他的研究成果是吻合的，Hongbin Cai（2011）等通过对“城镇居民收入支出调查”（UHIES）数据的研究，甚至发现了相似时间段内消费不平等显著高于收入不平等水平的现象。为什么收入不平等和消费不平等的水平如此接近呢？我们猜测这可能是由于在这段时间内我国的信贷市场不完全不成熟所导致，以至于我国不同收入层次家庭之间无法实现消费的转移。另一种可能的远因如同 Hongbin Cai（2011）等在他们的研究中提出的猜测，收入的不平等水平可能由于测量误差被低估了。这种测量误差很有可能是系统性的，因为高收入人群越有可能低报自己的实际收入、漏报一些边缘的收入（如收受礼品、节日奖金等）。

## 四、实证研究

### 4.1 收入不平等的分解

收入不平等指数的分解对于研究收入不平等的来源、制定相对应的降低不平等水平的政策都起到非常重要的作用。关于相应的分解方法，一直以来都有探讨，其中应用最广泛的是组内组间分解。组间不平等是指由一切我们可以观察到的、固定的外部人口特征（如居住地、户主年龄等）所决定的，及不同人口组之间是具有可观测到的异质性，Dirk Krueger（2006）等认为组间的收入差异是难以通过储蓄保险行为来进行风险防控的。组内不平等是总体的不平等水平中只能被以上因素所解释的部分，这种不平等一般被认为是由个体（各个组）的、特质的收入冲击所引起，如第三章的模型所述，这也是收入影响消费、储蓄等行为的传导机制的重要部分。该种组建组内分解方法简单易操作，并且具有鲜明的经济学意义，因此在研究此类问题的文章中得到了大量的应用。

---

<sup>4</sup> 置信区间的计算是采用 bootstrap 方法，经过 100 次自助重复抽样计算而得。

接下来，我们仿照 Katz 和 Artor（1998）的做法，以收入作为被回归的变量，可观测的家庭外部特征作为回归变量，其中包括：户主年龄和年龄的平方、互助的性别、家庭等职人口规模、16 周岁以下人口数量、户主教育水平、户主职业、户主行业、调查年份和省份的固定效应控制变量<sup>5</sup>。被回归变量由回归变量所解释的部分则是总的收入不平等水平中组间的部分，残差则代表的是组内的部分。我们以城镇居民收入的基尼系数为例，对收入不平等水平做了组间组内分解。系数的估计是通过 OLS 方法进行，方差的估计采用 White 异方差稳健估计。结果如图 4-1 所示：

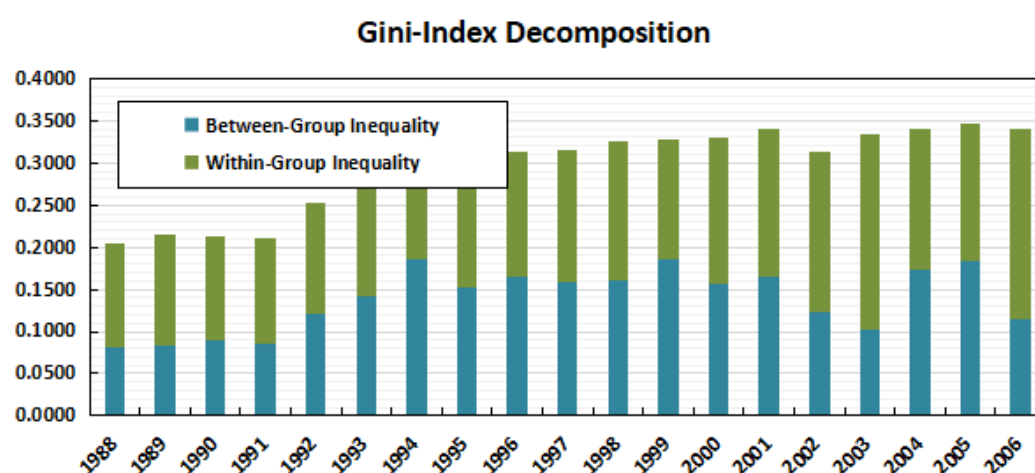


图 4-1 基尼系数组间组内分解

每根柱状图整体的高度代表了那一年总体的收入基尼系数，其中蓝色的部分是组间不平等水平，绿色的部分是组内不平等水平。从图中我们可以看出：除了个别年份，收入不平等的来源总体上是来自组内的部分更多，最多的一年是 2003 年，占到了总体的 70%；组间收入不平等水平起伏比较大，1988 年到 2006 年这段时间内没有明显的增长，而组内不平等水平总的看来保持着增长的趋势，从 1988 年的 0.12 增长到了 2006 年的 0.23，由此我们大致可以看出中国城镇居民收入不平等的主要来源是组内自身的收入分化，即自身的收入冲击、波动更为剧烈，而非来自不同人口特征所决定的组间差异。因此，除了研究总体的收入不平等对支出活动的影响，专门研究组内不平等的影响作用也有其重要性和必要性。

## 4.2 基于组间不平等的样本分组

为了从家庭微观级别的数据来研究收入不平等对消费的影响，将家庭按照一定的标准进行分组是很有必要的，因为基尼系数等数据是总体数据，要作为以家庭层面的微观数据作为被回归变量的话必须对其进行分解。在研究中常用的分解方法有出生组分解、年龄组分解，这种分组方式考虑了出生时家庭财富的禀赋效应和随年龄增长的年龄财富累积效应，对于耐用品消费等涉及到较长时间的消费约束等活动的研究具有重要的意义，并且在面板结构的调查数据的研究中也体现出了很好的性质和效果（邹红等，2013）。在对于横截面数据的研究中，在笔者有限的阅读中，还没有文献关于分组方法做出详细的探讨。Hongbin Li（2011）等在他们对于横截面数据的研究中通过构造年龄-地区的交互项来将样本分组，研究了每个“参照组”组内不平等对消费的影响。可见以前的研究都认为户主的年龄会是分组的一个重要标准。我们认为，分组的标准首先要体现出组内不平等的经济含

<sup>5</sup> 户主的教育水平、户主行业（户主所从事工作的行业，如制造业、农业等）、户主职业（户主具体的工作的内容，如技术人员、办事管理人员等）、时间、地域控制变量等均为一列虚拟变量。

义，即在本章上一小节所提及的组内收入的风险性波动。因此，我们分组一定要能尽可能地反映出组间的差异，使得每个组内的不平等的内容更“纯粹”。年龄是否会是决定组间不平等的一个重要因素？我们计算汇总了1988年到2006年的数据，计算了不同年代的家庭（以户主年龄为准）的收入和消费情况，如图4-2所示：

蓝色、绿色、红色、青色和粉色的线分别是20年代、30年代、40年代、50年代、60年代的家庭。左上的图和右上的图分别代表了不同年代家庭组的人均收入和消费情况，从中我们可以看出：稍后年代的家庭比更早年代的人财富积累得更快，如60年代的家庭比20年代的家庭更早达到相同的财富积累，家庭的消费也出相似的特征：收入和消费水平都在1990年后的一段时间内得到了大幅度的提升，这得益于宏观层面的改革措施使得国民福利有很大的提高。其中，消费的提升比收入更快，这可能一方面是经济市场化改革的结果，使得曾经制度经济下的消费转化到市场上，体现在了调查数据搜集的支出情况中，另一方面由于供给侧消费品升级使得出现了新的消费敞口，促进了人们的消费。

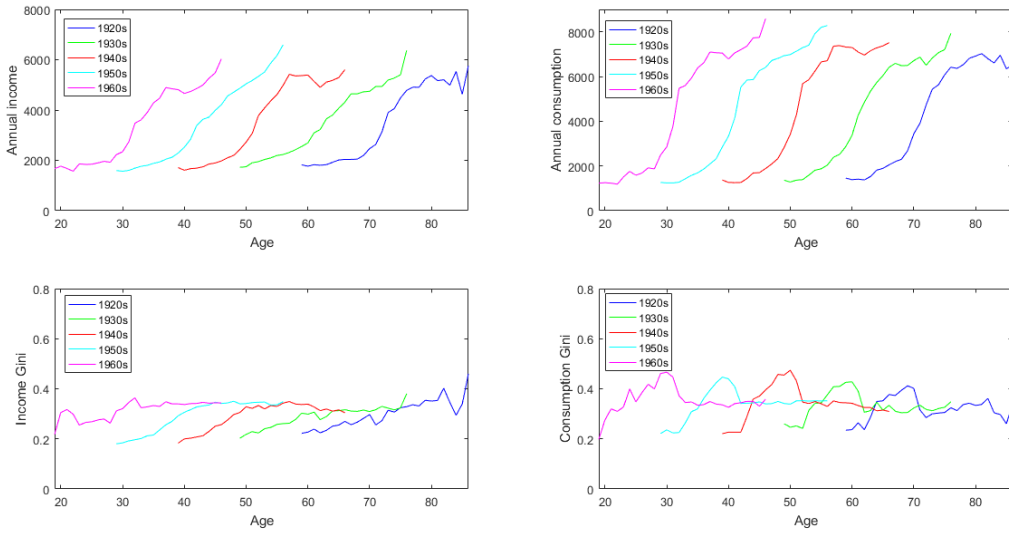


图 4-2 不同年代家庭组的收入、消费情况

左下和右下的图分别刻画了不同年代家庭组的收入和消费不平等情况。我们可以看见不同年代的家庭组收入和消费基尼系数都体现出相似的结构，这某种程度上说明年龄因素并不是区分组间不平等的有效标准，而是其他的一些因素，尤其是与时间相关的因素更加完整地刻画出了组建不平等水平。为了确定这些因素，我们有必要进一步对组间不平等做出分解。

首先，我们介绍一下不平等是如何根据来源进行分解，即在收入决定方程里，不同的变量对收入不平等的影响有多大。Fields（2002）在他的研究里提出了一种使用广泛、操作简便的方法。假设收入决定方程是线性形式的：

$$\ln(y) = \mathbf{X}'\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (7)$$

即

$$\begin{pmatrix} \ln(y_1) \\ \ln(y_2) \\ \dots \\ \ln(y_n) \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \dots \\ \beta_k \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_n \end{pmatrix} \quad (8)$$

。其中  $\ln(y)$  是家庭收入的对数值， $x_l$  是决定收入的第  $l$  个因素（在本文中，即为上一节中家庭外部特征等变量）， $\beta_l$  为其对应的系数， $\varepsilon$  为误差项（即组内不平等的来源）。对于基尼系数来说，可通过公式  $\beta_l \cdot \sigma(x_l) \cdot \text{corr}(x_l, \ln(y)) / \sigma(\ln(y))$  计算每一项外部特征变量对不平等水平的贡献度，若对  $x$  项再除以回归的  $R^2$ ，则可以得到不同家庭外部特征对组间不平等的贡献度。据此我们进一步分解了家庭收入的组间不平等，并得到了如图 4-3 所示的结果：

从图中我们可以看见五种因素对组间不平等的贡献度<sup>6</sup>：户主年龄一直以来都不是决定组间不平等的主要因素，而且是非常次要的因素；一直以来对决定组间不平等的重要因素是地区间的差异，其对组间不平等的贡献一直不小于 40%，且在 1993 年达到了 84%。在 90 年代地区间的收入差异一直都非常得明显，我们认为这是改革开放初期沿海地区受到的经济优惠政策导致的；另外一个值得注意的因素是户主的受教育程度，从 1988 年的 4% 增长到了 2006 年的 33.29%，且期间一直保持着增长的态势。这说明受教育程度成为了导致家庭人均收入水平分化的一个重要原因。自 1992 年开始，地域和教育因素一起来看，几乎决定了组间收入不平等的 70% 以上。为了进一步佐证我们的观点，我们仿照上面的做法，以省份和受教育程度为标准分组，绘制了描述收入、消费水平以及不平等水平的组图：

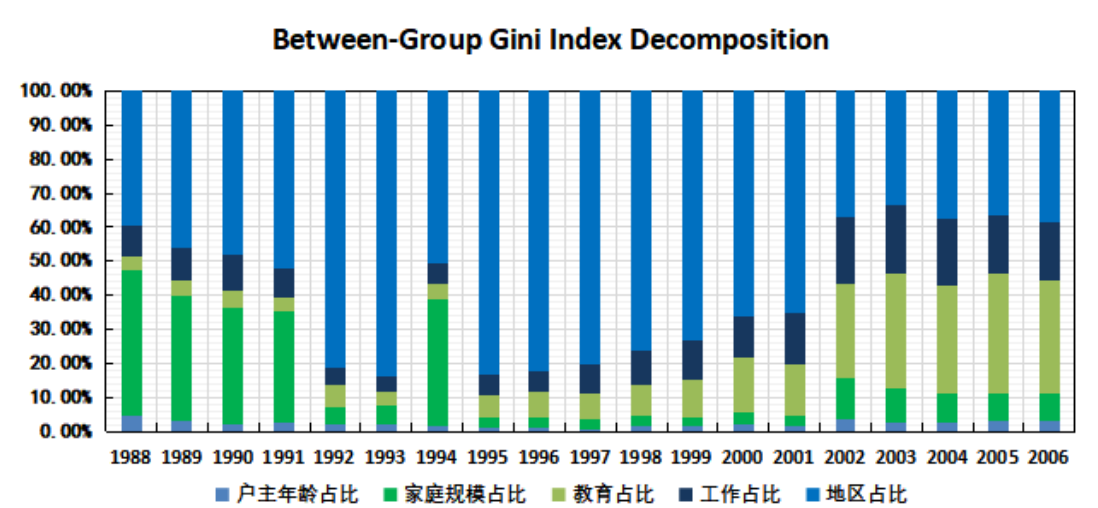


图 4-3 组间不平等分解

<sup>6</sup> 其中户主年龄占比是指户主年龄及其平方、出生年代的对组间不平等贡献度之和，工作占比是指户主工作行业、职业的贡献度之和。

图 4-4 选取了北京、辽宁、广东、四川四个省份的数据来进行说明：发现北京的城镇居民福利水平，无论是收入还是消费，在这段时间内的增速都是遥遥领先，在 2002 年左右超过了之前一直处于领先水平的广东省；调查的这段时间内，各个省份的不平等系数曲线都是“参差不齐”的，这进一步证实了地域是影响组间收入不平等的重要因素。

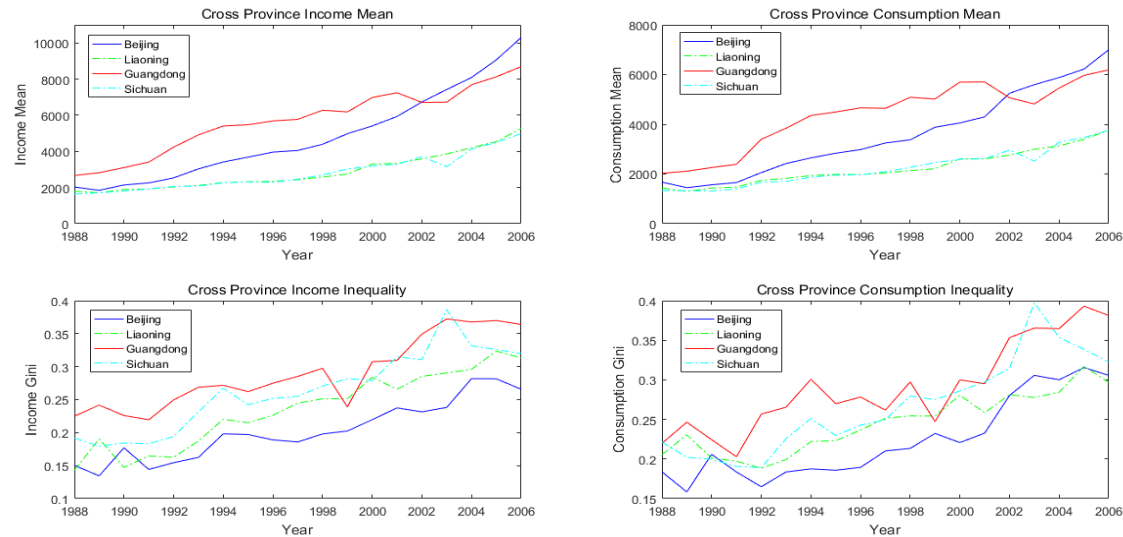


图 4-4 不同省份家庭的收入、消费情况

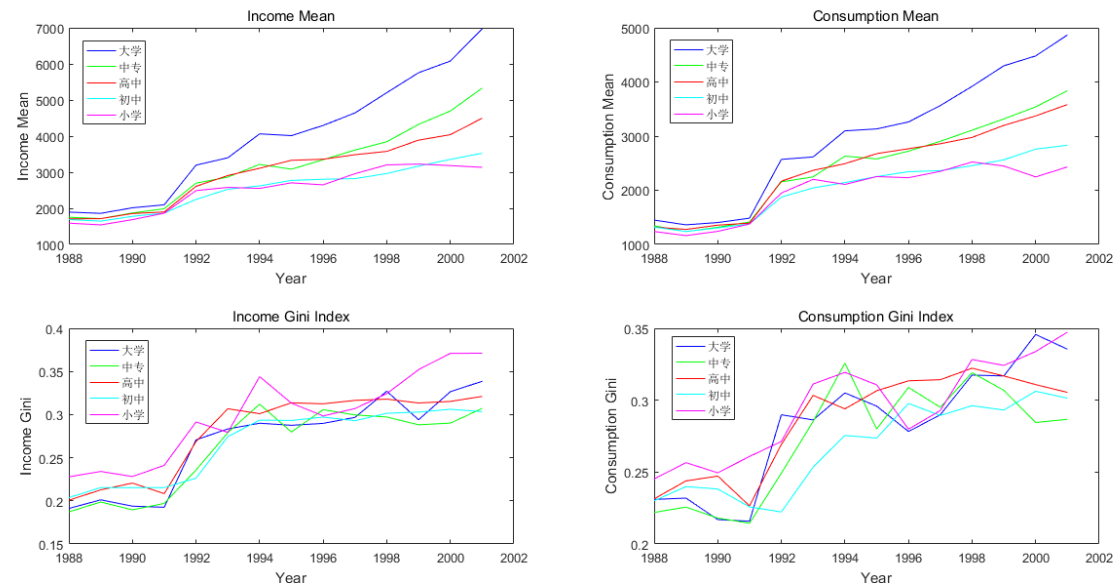


图 4-5 不同受教育程度家庭的收入、消费情况

图 4-5 刻画了不同教育程度（以户主为代表）的家庭的收入、消费情况：从图 4-4 和 4-5 我们明显可以看出，在 80 年代末 90 年代初，不同教育程度家庭的收入、消费情况都是极其接近的，而教育对家庭收入、消费的影响（主要是分化作用）随着时间的推移愈加得明显；从下面两张图同样可以看出，不同教育程度的家庭内部的不平等水平有明显的差异。这些组间的差异使得我们接下来的回归分析的回归变量有足够大的变化，使得收入不平等对消费的影响能被我们的回归模型显著地识别出来。

### 4.3 收入及收入不平等对消费的影响

首先，我们需要研究收入以及收入不平等对消费的影响。为此我们构造了如下计量模型：

$$\ln(con) = \alpha + \beta \ln(inc) + \gamma Gini + \delta \cdot X + \varepsilon \quad (9)$$

$\ln(con)$  和  $\ln(inc)$  表示流动性消费和劳动性收入的自然对数， $\beta$  测量了消费对于收入的弹性。 $Gini$  是按照省份和教育程度划分的组别计算的组内基尼系数<sup>7</sup>。 $X$  表示其他的一些解释变量和控制变量，我们会根据模型的选择在  $X$  中加入不同的变量。样本的选取我们限定在了户主为 18 周岁到 65 周岁的家庭。根据邹红等（2013）、Hongbin Li 等（2011）的研究的结果，我们在所有的模型都加入了调查年份和分组的固定效应（一系列 0-1 虚拟变量）作为控制变量。回归分析的结果如表 4-3-1 所示：

在模型（1）中我们额外加入的回归变量有：劳动性收入、户主年龄及其平方、户主性别<sup>8</sup>、家庭规模等。大部分的系数符号都与我们的预期相一致：消费对收入的弹性是 0.7717，这代表收入每增加百分之一，消费会增加 0.77%，这说明对于劳动性收入，城镇居民不会采取完全消费的策略，这符合我们对人们消费行为的认知；户主年龄的一次项是正数，二次项为负且绝对值远小于一次项，这说明在我们选取的年龄范围内随着户主年龄的增长，家庭人均消费会逐渐增长，但增长率逐渐降低。但总体来讲互助年龄对消费的影响不算明显；让我们比较意外的是户主为男性的家庭的消费比女性户主的家庭消费低 3% 左右。我们统计了调查数据中以女性为户主的家庭有 5016，其中 3635 是三口之家。由于社会传统的原因，家庭的户主一般是男性，于是我们猜测女性户主的情况可能是该户主工作能力和消费能力在家庭中比较突出；家庭规模的系数是 -0.0421，这说明家庭人口越多的家庭消费能力越低；而 16 周岁以下人口的系数是 0.0126，这可能是因为在城镇中选择二胎或多胎的家庭本身消费能力更强。所有感兴趣的回归变量的系数估计都以百分之一的置信水平显著。

在模型（2）中我们加入了组内不平等作为消费的解釋变量。与我们之前的理论相契合，组内收入不平等的系数是 -0.5586，即组内收入基尼系数每增加 0.1，消费会减少 5.6%，并且该系数以百分之一的置信水平显著。而其他的系数变化不显著。为了测试基尼系数在不同时间对家庭消费是否会有不同的影响，在模型（5）中我们建立了基尼系数和时间的交互项，结果显示这些交互项的系数估计普遍都不显著。

在保留这些回归变量的基础上，我们测试了收入和组内基尼系数是否会对消费有非线性的影响。我们将家庭收入划分为三个层次，百分之二十五分位点以下称为低层收入家庭，百分之二十五分位点到百分之七十五分位点之间的称为中层收入家庭，之上的称为上层收入家庭。在模型（3）中我们对家庭收入和收入层次做了交互项，可以看出三种层次的家庭消费对收入的弹性都有显著的区别，但这种区别不大。收入越高，消费对收入的弹性会略微变小，这符合边际消费倾向递减的假设。在模型（4）中我们对组内基尼系数和收入层次也做了交互项，我们发现不同收入层次的家庭消费对基尼系数的敏感程度有较大的区别。对于低层收入的家庭来说，基尼系数每增加 0.1，消费会减少 9%，对于中层收入的家庭而言，消费仅减少 5%，对于上层收入而言这个数字则更小，只会下降 3%。这充分说明了城镇家庭组内收入不平等不仅会降低人们的消费倾向，也一定程度上决定了消费不平等的程度。这些估计均在 1% 的置信水平上显著。

<sup>7</sup> 数据中省份包括北京、山西、辽宁等总计 18 个省份，学历分为本科及以上、大专、中专、高中、初中、小学和其他 7 种情况，将他们的交互项作为标准共分成 126 组。

<sup>8</sup> 户主性别是虚拟变量，性别=1 代表户主是男性。

表4-1 对消费的回归分析

被解释变量	OLS				
	流动性消费				
	(1)	(2)	(3)	(4)	(5)
劳动性收入	0.7717*** (0.0027)	0.7541*** (0.0026)	0.7993*** (0.0050)	0.8034*** (0.0050)	0.7535*** (0.0026)
基尼系数		-0.5586*** (0.0567)	-0.5591*** (0.0599)	-0.9013*** (0.0816)	-0.3808** (0.1531)
户主年龄	0.0212*** (0.0011)	0.0069*** (0.0010)	0.0208*** (0.0011)	0.0208*** (0.0011)	0.0069*** (0.0010)
户主年龄 <sup>2</sup>	-0.0002*** (0.0000)	-0.0001*** (0.0000)	-0.0003*** (0.0000)	-0.0003*** (0.0000)	-0.0001*** (0.0000)
户主性别	-0.0311*** (0.0051)	-0.0270*** (0.0052)	-0.0304*** (0.0051)	-0.0314*** (0.0051)	-0.0270*** (0.0052)
家庭规模	-0.0421*** (0.0026)	-0.0673*** (0.0026)	-0.0435*** (0.0026)	-0.0437*** (0.0026)	-0.0674*** (0.0026)
16周岁以下人口	0.0126*** (0.0024)	0.0126*** (0.0023)	0.0123*** (0.0023)	0.0121*** (0.0023)	0.0125*** (0.0023)
劳动性收入*中层收入			-0.0015*** (0.0004)	-0.0144*** (0.0022)	
劳动性收入*上层收入			-0.0063*** (0.0007)	-0.0238*** (0.0027)	
基尼系数*中层收入				0.4100*** (0.0713)	
基尼系数*上层收入				0.5964*** (0.0907)	
基尼系数*1998					0.1550 (0.2416)
基尼系数*1999					0.0287 (0.2324)
基尼系数*2000					0.5722** (0.2337)
基尼系数*2001					0.1276 (0.2425)
基尼系数*2002					-0.0725 (0.1675)
基尼系数*2003					-0.3451** (0.1590)
基尼系数*2004					-0.0632 (0.1690)
基尼系数*2005					-0.0026** (0.1685)
基尼系数*2006					-0.0794 (0.1746)
时间固定效应	Y	Y	Y	Y	Y
分组固定效应	Y	Y	Y	Y	Y
R-Square	0.6804	0.6808	0.6813	0.6815	0.6813
N	100,494	100,494	100,494	100,494	100,494

注：括号内为标准差；\*\*\*、\*\*和\*分别表示结果在1%、5%和10%的水平下显著。



#### 4.4 收入不平等与储蓄、投资

除了消费以外，我们还关心收入不平等对储蓄、投资行为的影响。我们在本节分析的对象包括家庭教育支出、银行储蓄和非储蓄型投资。Hongbin Li (2011) 等通过实证分析研究了收入不平等是如何在人们追求更高的社会地位时影响人们的储蓄、投资等行为。Weiss Freshtman (1998) 等认为社会人群根据某些外在的特质，比如财富多少、受教育程度、身份背景等划分为不同的社会群体，并且这种群体在社会印象中是有高低之分的。当不同社会群体之间的差距拉大时，往往伴随的群体性福利也会更加吸引人们去为向“上”流动而努力 (Corneo, Jeanne, 1998)。在教育上投资是改变社会群体的重要方式之一。Hongbin Li (2011) 等的研究显示了收入差距会促使家庭在教育上付出更多的投资，并且相关的度量显示几乎所有收入层次的家庭的教育支出弹性都相似。他们的研究对分组没有进行详细的探讨，而是以地区和年龄的双重标准进行分组，因此收入不平等的的内容不“纯粹”。根据上述对社会地位追求动机理论的探讨，我们认为是组间不平等对教育产生的影响。仿照上一节的模型设定，我们对教育支出也进行了类似的研究，结果见表 4-2：

在模型 (1) 中，我们同样了收入、基尼系数等变量作为解释变量。我们可以看见，收入对教育支出的影响是显著的，对于低层收入家庭来说，教育支出对收入的弹性是 1.0152，并且在不同收入层次内弹性的差别不大。而消费对收入的弹性是 0.7717，由此我们可以发现人们投资教育的倾向更强。与 Hongbin Li (2011) 等的研究不同，我们发现基尼系数对教育支出的影响是负的，-0.4587，在 2% 的置信水平下显著。这个“矛盾”其实是符合我们预期的，因为在以前的研究中参与回归分析的基尼系数中既包含“组内”的成分，也包含“组间”的成分。我们认为按照社会地位追求动机理论，不同家庭组之间的收入差距会是促进人们进行教育支出的原因，组内的收入差距仍然会是降低人们在教育上投资的意愿，因为对教育的投资是长期的，其回报也是间接、隐性的，因此人们并不会选择教育支出在对冲组内的收入冲击等风险。除此之外，与消费不同，家庭规模对教育支出的影响是正的，这可能是因为多出来的人口一般还处于受教育阶段。在模型 (2) 中我们对基尼系数和家庭收入层次也做了交互项，发现在本模型中基尼系数对教育支出的影响并不显著。

在模型 (3) 中，我们研究了收入不平等水平对银行账户储蓄的影响。由于在储蓄和投资上有很多缺失数据，我们的样本容量因此也减少了许多。我们可以发现无论是哪个收入层次的家庭的在银行的储蓄意愿都是非常强的。低层收入家庭的储蓄对于收入的弹性系数是 1.2924，中层收入家庭是 1.3595，而上层收入家庭是 1.3711。收入层次越高的家庭的边际储蓄意愿越强，但差别并不明显。组内基尼系数对储蓄的影响变化比较大：对于低层收入家庭来说，组内基尼系数每增加 0.1，银行储蓄会增加 27.34%，而对于中层收入的家庭而言这个数字是 3.65%，上层收入家庭是 0.86%。由此我们可以看出在这段时间内收入差距的存在，相比较于中产阶级和富裕阶级的家庭，会使得低层收入家庭住户会更多地以储蓄来作为防范收入波动的手段。除此之外，家庭人口规模对银行储蓄也有显著的正面的影响。家里每多一单位等值人口，储蓄就会增加 44.62%。这也符合我们的经济直觉，因为家里人越多，为了保障正常的生活需求得到满足，一般来说人们防范风险的意识和动机就会更强。在模型 (4) 中，我们进一步研究了收入不平等对非储蓄投资的影响。这里的投资包括：储蓄性支出保险、购买有价证券和其他投资支出。通过回归分析我们可以看出：不同收入层次的家庭也都具有很强的投资意愿。低层收入家庭投资对收入的弹性是 0.9744，中层收入家庭是 1.037，上层收入家庭是 1.0456。由此可知，我国城镇家庭住户在收入增加时的边际投资意愿还是比较强，但和储蓄相比，城镇居民普遍还是更倾向于储蓄。再来看收入不平等对投资的影响，与对储蓄的影响类似，不同收入层次的家庭的边际投资倾向也不尽相同。基尼系数每增加 0.1，低层收入家庭的非储蓄型投资会增加 13.76%，而上层收



入会减少 8.29%，中层收入的系数估计不显著。这说明在收入波动大的情况下，我国城镇住户中低层收入家庭会积极参与非储蓄型投资尝试实现财富增值，而高层收入住户则会规避这些投资所带来的进一步风险。

表4-2 对投资的回归分析

被解释变量	OLS			
	投资			
	(1) 教育	(2) 教育	(3) 储蓄	(4) 投资
劳动性收入	1.0152*** (0.0161)	1.0123*** (0.0167)	1.2924*** (0.0182)	0.9744*** (0.0430)
基尼系数	-0.4587** (0.1968)	-0.4403 (0.2869)	2.7335*** (0.3227)	1.3759*** (0.8936)
户主年龄	0.2535*** (0.0060)	0.2535*** (0.0060)	-0.0433*** (0.0034)	-0.0187* (0.0096)
户主年龄 <sup>2</sup>	-0.0030*** 0	-0.0030*** (0.0000)	0.0005*** (0.0000)	0.0001* (0.0001)
户主性别	-0.0695*** (0.0203)	-0.0698*** (0.0203)	0.0522** (0.0254)	-0.0795** (0.0334)
家庭规模	0.0989*** (0.0102)	0.0991*** (0.0102)	0.4462*** (0.0088)	0.2272*** (0.0224)
16周岁以下人口	0.3035*** (0.0088)	0.3034*** (0.0088)	-0.0043 (0.0084)	-0.0424** (0.0201)
劳动性收入*中层收入	-0.0036** (0.0017)	0.0037 (0.0087)	0.0671*** (0.0101)	0.0626** (0.0263)
劳动性收入*上层收入	-0.0062** (0.0026)	-0.0108 (0.0094)	0.0787*** (0.0105)	0.0712*** (0.0273)
基尼系数*中层收入		-0.2362 (0.2731)	-2.3684*** (0.3093)	-1.9122 (0.8402)
基尼系数*上层收入		0.1898 (0.3084)	-2.6471*** (0.3293)	-2.2047** (0.8834)
R-Square	0.2669	0.2670	0.5178	0.3821
N	97,164	97,164	63,742	24,672

注：括号内为标准差；\*\*\*、\*\*和\*分别表示结果在1%、5%和10%的水平下显著。

## 4.5 分位数回归比较

上两节的研究的模型分析中我们可以发现不同收入层次的家庭的消费、支出性质有所不同，我们还关心从消费层次的角度来看，收入和收入不平等会如何影响家庭的消费等支出行为。在此基础上，我们还对消费、教育支出、银行储蓄和非储蓄性投资做了分位数回归分析，我们研究了中国城镇住户不同支出水平受收入和收入不平等的影响。在计量模型的构造上，假设  $y$  是被解释的变量， $x$  是我们所有的解释变量和控制变量，包括劳动性收入、组内基尼系数、户主年龄及其平方、家庭等值规模、16 周岁以下人口。我们都假设条件分布  $y/x$  的分位数  $y_q(x)$  是关于  $x$  的线性函数，即：

$$y_q(x) = x' \beta_q \quad (10)$$

我们通过线性规划的方法来求解以下最优化问题，从而估计  $\beta_q$  的值：

$$\max_{\beta_q} \sum_{i: y_i \geq x_i' \beta_q} q |y_i - x_i' \beta_q| + \sum_{i: y_i < x_i' \beta_q} (1 - q) |y_i - x_i' \beta_q| \quad (11)$$

其中  $n$  是样本量总数， $i$  是对样本点的编号。渐进协方差矩阵的估计方法是 bootstrap 自助抽样法，接下来的模型都抽样 150 次。我们研究了  $q$  为 0.1、0.25、0.5、0.75、0.9 时候的情况，结果如表 4-3 所示：

表4-3 分位数回归分析

被解释变量		(1) 10%	(2) 25%	(3) 50%	(4) 75%	(5) 90%
非耐用品消费	劳动性收入	1.1247*** (0.0507)	1.1243*** (0.0306)	0.9892*** (0.0225)	0.9274*** (0.0247)	0.9730*** (0.0350)
	基尼系数	-2.5405** (1.2841)	-1.4195** (0.7114)	-0.3656 (0.4711)	0.3963 (0.4357)	0.8228 (0.5749)
教育	劳动性收入	1.1247*** (0.0507)	1.1243*** (0.0306)	0.9892*** (0.0225)	0.9274*** (0.0247)	0.9730*** (0.0350)
	基尼系数	-2.5405** (1.2841)	-1.4195** (0.7114)	-0.3656 (0.4711)	0.3963 (0.4357)	0.8228 (0.5749)
储蓄	劳动性收入	1.4835*** (0.0387)	1.4459*** (0.0214)	1.3732*** (0.0178)	1.272*** (0.0139)	1.1603*** (0.0128)
	基尼系数	0.4621 (0.5178)	0.5996* (0.3378)	0.5286*** (0.1931)	0.6874*** (0.2111)	0.5547*** (0.1754)
投资	劳动性收入	0.9775*** (0.0753)	1.0629*** (0.0419)	0.9677*** (0.0412)	0.9532*** (0.0369)	1.0696*** (0.0529)
	基尼系数	-2.1580* (1.1091)	-1.3809** (0.6496)	-0.4694 (0.4603)	0.3048 (0.4248)	0.5800 (0.5241)

注：括号内为标准差；\*\*\*、\*\*和\*分别表示结果在1%、5%和10%的水平下显著。

我们还是先关注非耐用消费的情况：劳动性收入对各个层次非耐用品消费的影响都是显著的。劳动性收入每增加 1%，非耐用品消费的 10%分位点就会上升 1.12%，中位数会上升 0.99%，百分之 90 分位点会上升 0.97%。由此可见收入的提升对各个消费层次的水平提升也都是明显的，即收入的提高可以真实地转化为效用的提升，随着消费层次的提高，其边际影响会略有减弱。再来看组内基尼系数对非耐用品消费的影响，我们可以发现基尼系数对不同消费层次的影响是大相径庭的：组内基尼系数每增加 0.1，非耐用品消费的 10%分位点会下降 25.4%，中位数会下降 3.7%，而百分之 90 分位点却会上升 8.2%。我们的分位数回归分析充分说明了收入不平等会拉大消费的差距。从统计有效性的角度看，基尼系数对低层次消费的影响是显著，对 10%和 25%分位点的边际影响估计值都在 2%的置信水平下显著。而对高层次消费的影响是否真实还不尽明确。

类似的，我们还研究了收入和收入不平等对不同银行储蓄层次、非储蓄型投资层次水平的影响。收入对这两者的影响都是非常显著的，所有的估计值都在 1%的置信水平下显著。与上一节的分析结果相似，收入的增加促使人们更多地在银行存款，这种影响随着储蓄层次的上升递减。这可能意味着我国城镇住户均将储蓄作为防范风险的重要手段，因为相较于非耐用品消费，收入的增加对储蓄的影响明显偏大，并且，储蓄水平越低的住户在收入增加时比高储蓄住户更倾向于向银行存款，借以对抗未来可能延续的收入波动。基尼系数的增加对储蓄的影响也是正向的，这也印证了我们的猜想。但我们还发现基尼系数的影响仅仅在中位

数、75%和 90%分位点处的影响是显著的，我们猜测这可能是因为原本较高的储蓄水平意味着住户对抗收入风险的意识更强，因此对于基尼系数的变化会更加敏感。收入的增加对教育投资和证券保险类投资的边际影响也是显著的，且与我们上一节对不同收入层次的住户的研究结果类似，我国不同投资层次的城镇居民的边际投资大致相当，并且均低于边际储蓄意愿。基尼系数对投资的影响估计不显著，在此不做过多探讨。

## 五、结论与讨论

本文着重研究了收入不平等对消费的影响，因为相较于收入，消费更加精确地反应了每个人享受到的真实福利水平；Kocherlakota (1996), Kehoe (2001)和 Kruger, D(2006)等人的研究通过数学模型模拟了收入不平等向消费不平等的传导机制，并指出者之间的差异是决定信贷市场均衡的有效性，Blundell 等（2008）认为个人所面临的永久性收入冲击是决定个人生命周期内消费策略的重要影响因素。本文通过回顾以往的这些研究，从实证分析的角度出发，结合中国城镇住户调查的横截面数据，剖析了我国城镇居民收入、消费不平等的现状，探讨了城镇家庭短期内收入及收入不平等对家庭支出规模、方式的影响，主要得出了以下的几点结论：

1. 我们通过对数据的统计分析发现无论是从收入还是消费的角度来看，我国城镇家庭住户的收入和消费分布的对称性都很差，并且在可能存在高收入、消费的异常值。不平等水平都在 1988 年到 2006 年这段时间内逐年升高。并且收入不平等水平和消费不平等水平非常得接近，变化几乎是同调的。

2. 我们在以往的研究基础上，对收入不平等进行了组内和组间分解。结果发现不同家庭组之间的收入差距有起伏但是平均水平不大，而目前增长的不平等水平来源主要是组内不平等扩大造成的。我们进一步对组间不平等进行了解，发现影响收入不平等的重要因素有受教育程度和居住地区，并且受教育程度的影响还在不断扩大。

3. 我们进一步以消费、储蓄、投资等作为被解释变量，分析了收入及收入不平等对他们的影响。研究发现消费对收入的弹性小于教育、储蓄和投资等；基尼系数也对消费有负面的影响，对储蓄、投资有正向的影响，并且我们对不同收入水平的家庭做了非线性影响的假设，结果显著。这说明我国城镇居民对抗收入风险的意识较强，并且储蓄是對抗风险的重要手段。

4. 通过对不同水平的消费、储蓄和投资水平的分位数回归分析，我们发现收入水平对各个层次的家庭消费、教育、储蓄和投资都有显著的影响。而组内收入不平等则对低消费人群和高储蓄人群有显著的影响。

最后，我们提出针对本文的不足提出需要进一步探索的方向：

1. 如第四章第二节所说，我们有足够的理由怀疑我们在家庭收入数据的搜集上存在测量误差，这会导致我们的估计出现内生性问题。如果我们假设收入越高的家庭漏报收入越多，则我们对收入影响的估计很可能偏高。

2. 我们的模型和实证分析都着重于横截面数据的分析，并且假设同一时间内的收入差距能刻画出收入波动的程度。我们还可以模仿 Xin Meng (2003) 和 Hongbin Cai (2011) 的做法，讲多个数据集同时进行分析，因为其他的数据集，如 UHIEE、UHIES 等，提供了失业率，个人跨时间周期多年收入信息等数据，可以进一步帮我们刻画出收入不平等和收入风险水平，并且估计个人跨周期的消费平滑程度。

## 参考文献

- 李国正, 艾小青, 李晨曦 (2017). 流动人口家庭束缚和收入不平等影响因素分析. 统计与决策 (08), 90-93
- 王洪亮, 徐翔 (2006). 收入不平等孰甚: 地区间抑或城乡间. 管理世界 (11), 41-50
- 杨俊, 黄潇, 李晓羽 (2008). 教育不平等与收入分配差距: 中国的实证分析. 管理世界 (01), 38-47
- 邹红, 李奥蕾, 喻开志 (2013). 消费不平等的度量、出生组分解和形成机制——兼与收入不平等比较. 经济学 (季刊) (04), 1231-1254
- Atkinson, A. B., Rainwater, L., Smeeding, T. M., & Development, O. F. E. C. (1995). *Income Distribution in OECD Countries: Evidence from the Luxembourg Income Study*. Organisation for Economic Co-operation and Development.
- Blundell, R., Pistaferri, L., & Preston, I. (2008). Consumption Inequality and Partial Insurance. *American Economic Review*, 5(98), 1887-1921.
- Cai, Hongbin Y. C. A. L. (2011). Income and Consumption Inequality in Urban China: 1992-2003. *Economic Development and Cultural Change*. 58, 385-413.
- Conceição, P., & Ferreira, P. (2000). *A Young Person's Guide to the Theil Index: Suggestive Intuitive Interpretations and Exploring Analytical Applications*.
- Corneo, G., & Jeanne, O. (1998). Social organization, status, and savings behavior. *Journal of Public Economics*, 70(1), 37-51.
- Deaton, A., & Paxson, C. (1994). Intertemporal Choice and Inequality. *Journal of Political Economy*, 102(3), 437-467
- Fields, G. S. (2002). Accounting for Income Inequality and its change.
- Gan, L. (2013). Income Inequality and Consumption in China.
- Jin, Y., Li, H., & Wu, B. (2011). Income inequality, consumption, and social-status seeking. *Journal of Comparative Economics*, 39(2), 191-204.
- Johnson, D., & Shipp, S. (1997). Trends in inequality using consumption-expenditures the U.S. from 1960 to 1993. *Review of Income and Wealth*(43), 133-152.
- Kanbur, R., & Zhang, X. (1999). Which Regional Inequality? The Evolution of Rural - Urban and Inland - Coastal Inequality in China from 1983 to 1995. *Journal of Comparative Economics*, 27(4), 686-701.
- Kanbur, R., & Zhang, X. (2005). Fifty Years of Regional Inequality in China: A Journey through Revolution, Reform and Openness. *Review of Development Economics*, 1(9), 87-106.
- Katz, L., & Autor, D. (1999). Changes in the wage structure and earnings inequality. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics* (3, Part A, pp. 1463-1555). 26: Elsevier. (Reprinted.
- Kehoe, T. J., & Levine, D. K. (2001). Liquidity Constrained Markets Versus Debt Constrained Markets. *Econometrica*, 69(3), 575-598.
- Kocherlakota, N. R. (1996). Implications of Efficient Risk Sharing without Commitment. *Review of Economic Studies*, 63(4), 595-609.
- Krueger, Dirk, & Perri, F. (2006). Does income inequality lead to consumption inequality Evidence and theory. *Review of Economic Studies*, 1(73), 163-193
- Pendakur, K. (1998). Changes in Canadian Family Income and Family Consumption Inequality between 1978 and 1992. *Review of Income and Wealth*, 44(2), 259-282.
- Sen, A. (1973). *On Economic Inequality*. Oxford: Clarendon Press.

Weiss, Y., & Fershtman, C. (1998). Social status and economic performance:: A survey. *European Economic Review*, 42(3), 801-820.

Wu, X., & Perloff, J. M. (2005). China's Income Distribution, 1985 - 2001. *The Review of Economics and Statistics*, 87(4), 763-775.

Zhang, Y., & Eriksson, T. (2010). Inequality of opportunity and income inequality in nine Chinese provinces, 1989-2006. *China Economic Review*, 21(4), 607-616.