居民用电需求与异质性

——基于有限混合模型进行分析

吴力波 周阳

**摘要：**

本文使用高频用电数据进行了用户用电行为与基本属性之间的关系，填补了以往电力需求分析中没有高频数据进行分析这一缺陷。本文在获取了上海市浦东新区居民用电日冻结电量数据，结合当地气象数据以及调研获取的地理位置信息、家庭基本特征数据等，构造了温度用电序列，并使用有限混合模型（finite mixture model）对居民用电需求进行了分析，并从收入、房价、家庭人口构成、地理位置等角度识别了居民在用电需求上的异质性。结果显示，在上海浦东地区，收入、房价的上升会导致居民成为高用电用户的可能性上升，并且以中环为分界线，上海市浦东新区存在明显的用电断层；同时居民的用电温度敏感性也会随收入、房价等变量的变动而发生变化。这样的异质性研究将有利于整个上海电网的配置优化，对电网系统平稳运转，缓解节点拥堵有着重要意义。

**关键字** 日用电量 有限混合模型 异质性 区域断层 温度敏感性

**一、研究背景及文献综述**

目前我国电力市场正在经历市场化改革，这意味着传统的垄断市场以及政府定价的市场机制将被打破，远期市场、现货市场的引入将大幅度改变电力价格，进而影响电力系统的供给与需求以及电力系统的均衡。为了适应这样的变革，需要对我国电力市场需求侧进行准确的评估与分析。

以往的研究中，最主要、也是最为常用的分析方法是面板数据分析。Reiss and White (2005)使用RECS数据库中1307户加州居民的年度数据进行家庭属性、用电器与用电量之间的面板研究，但是由于数据频度较低，其将年度数据转化为月度数据进行估计。Melkinov (2012)构建了PET一般均衡模型，对异质性居民长期人口增长、区块变化等造成的能源使用变化、CO2排放变化等进行了模拟研究。Bernard (2011)使用1989至2002年的电力面板数据对居民电力需求进行分析，其发现短期及长期，居民对于价格的敏感性都是显著的，但是收入弹性不显著。Alberini and Filippini (2011)使用美国48州的电力使用面板数据，使用改进方法分析了1995至2007年美国居民的电力价格弹性以及区域对于弹性的影响，并指出传统方法估计的有偏性。Blázquez（2013）也使用面板数据对西班牙47省2000至2008年的用电情况进行面板数据分析，其得出价格弹性为负且小于1，同时气象变量也将影响电力需求。Fell (2014)也曾使用面板数据进行分析，并得出了类似的结论。在进行用户需求分析时，面板数据的确是比较合适的分析方法，并且，面板数据分析可以在一定程度上分析用户的异质性。在国外研究中，对于价格的弹性是学者都比较关注的问题。但是在中国，由于比较特殊的年度或月度阶梯电价，并且在电力系统内价格政策往往多年保持不变。因此，价格信号对于居民用户来说十分不明确，在进行面板分析时很难使用价格变量或政策工具变量，从而也使面板分析出现了很多困难。

另一方面，上述研究过程中普遍使用的是月度、年度数据，这也导致电力数据与气象数据之间的匹配相当困难，传统的方法往往是使用HDD、CDD等来表示一个月里面的制冷制热天数，但是这样的数据对于电力需求来说无疑是不准确的，因为不同环境温度下，对于制冷制热的电力需求是不同的，能够精确地将电力使用数据与气温进行匹配将大幅度改善这方面的拟合效果。幸运的是，目前上海市已经完成了全样本用户的智能电表更换，因此，在智能电表的配合下，日用电数据的获取变成可能。同时，智能电表带来的信息也将改变人们的用电行为（Gans, 2012）。

在分析方法层面，面板数据分析无法将用户进行归类，从而无法明确用户的异质性，也无法了解用户用电需求异质性的驱动因素。正如上述文献中进行的研究一样，传统的经济学研究方法，如面板数据研究在进行电力市场分析或弹性测算时往往将需求侧的影响因素同质化或归一化，即各个因素对于用电行为的影响是相同的。但是在实际中，尤其在居民用电需求方面，异质性将在很大程度上影响模型的估计结果；此外，由于电力系统并不是市场的简单出清，还包括每个电网配送节点的正常运转，需要考虑网架、台区的拥堵等情况，因此对于不同区域、不同类型用户电力需求的异质性研究就十分重要。

针对样本群体的异质性，对于样本的用户类别分类以及分析影响用户类别的因素非常重要。有限混合模型恰好符合我们对于该数据的分析要求（Mclachlan and Peel, 2004）。有限混合模型是纯粹数据驱动的数据分析方式，并且目前已经在经济学研究中广泛应用(Masyn, 2013)。早在1997年，Deb (1997)就使用有限混合模型针对医疗需求进行了分析；Allenby (1998)也使用有限混合模型对市场中消费者异质性进行了研究；Gerdtham (2001)则针对瑞典医疗设别使用方面进行了有限混合探索，并修正了前人的研究成果；Trovato (2008)将有限混合模型拓展到索罗模型上，从而研究国家在经济增长上的异质性；Maitra (2015, 2016)针对用户耐用品拥有数量及其随时间变化的改变量研究印度贫困程度的改变，并发现了印度农村穷人在未来越来越穷的现象；Parmeter (2015)也通过有限混合模型对于国家碳排放量的时间序列进行了分析，并得出影响国家碳排放总量以及人均量变化的因素；Ruto (2012)使用隐分类模型对于消费者对环境管理的偏好进行研究。如前文总结，有限混合模型正应用于统计、医疗以及经济学等各个方面的研究之中，并且往往能对以往研究进行修正或改进。

据我们所知，本研究是首次使用日用电量这样的高频用电数据去分析挖掘居民用电需求与其他因素之间的关系的研究，弥补了以往文献在电力高频数据以及高频需求管理分析方面的缺失。此外，不同于传统经济学建模过程，本研究还是数据驱动（data-driven）的用电数据特征挖掘研究，这样的数据挖掘过程会大量减少经济学假设，从而使本研究的结论适应性更广泛，也将减少模型选择的误差。

**二、假设及实证模型**

**2.1基本假设及有限混合模型**

居民的用电需求是通过用电器实现的，也就是说居民的用电需求本质上是短期内其对不同用电器的使用与否以及使用时长。相比于高频用电数据（如15分钟一次的用电数据），居民的日用电需求几乎不存在相互关联，即今日的用电器使用需求并不会影响明日的用电器使用需求。因此本文假设，居民的日用电量之间是相互独立的。在这样的假设下，本文构建了温度用电序列，使用有限混合模型对居民的用电需求进行了一系列分析。

有限混合模型构建了一个由多个分布共同组成的混合分布。其基本假设为某一分布*F*由*K*个分布共同构成（*K*可以由模型实证求出），其中最为常用的为高斯混合分布，即由*K*个高斯分布共同组成的混合分布形式，其表达式如下：

 （1）

其中，*F(·)*为观测到的分布情况，在本文中代表居民的用电情况，为每一个高斯分布密度函数，为每一个独立分布出现的概率，其中>0，并且，在用于居民用电分类时，可以代表群体中居民群体不同用电需求出现的比例。Marron and Wand (1992)曾证明，任何一个连续的概率分布都可以通过有限个高斯分布进行拟合，因此，本研究的分析结果不会随成分的分布形式变化而变化。在使用有限混合模型进行用电分析时，为各类居民日用电分布的均值，为标准差。

**2.2带有伴随变量的有限混合模型**

然而在进行居民用电需求研究时，我们不仅需要研究用户的用电分类，还需要了解居民自身的属性，如收入、地理位置等对用电需求的影响，即用户不同用电需求概率如何随着伴随变量的变化而变化（Dayton and Macready, 1988）。因此我们引入带有伴随变量的有限混合模型，其中伴随变量将决定用户不同用电需求的分类依据，并通过伴随变量可以进行不同用户用电需求概率的计算。其基本形式如下：

 （2）

其中*z*为伴随变量，即各不同用电需求分布的概率将随着*z*的变化而变化，并且，在本研究中，我们认为负荷多变量logit过程，即：

  （3）

上式中，*k*为非参考项时的概率，（2）其他所有变量与（1）式中相同。由于不同类别用电需求分布的概率服从多变量logit过程，因此需要设定参考项，本文将在下文中标明参考项类别。

**2.3有限混合回归及重复测量**

在进行有限混合分析模型时，一种常用的拓展方式是将有限混合分布中*F(·)*拓展为回归形式，即：

 （4）

其中y为被解释变量而x为解释变量，为每一类用户用电回归的参数。本研究使用（4）进行居民用户用电温度敏感性的异质性分析。在进行居民用电敏感性异质性分析时，等式（4）需要进行一些拓展，由于观测到的居民用电需求是由很多日期的用电数据共同组成的，也就是说每一个用户将存在多个用电观测值，即将替换为，其中*t*代表不同日期的用电量观测值。因此，用户用电的温度敏感性可以写成以下形式：

 （5）

其中和分别代表归属于*i*类别的*j*用户观测到的第*t*天的用电数据以及当日温度，为该类居民的用电-温度敏感性系数。进一步，为了明确不同用户类型对于其温度敏感性的影响，本文在有限混合回归的基础上引入伴随变量，其具体模型如下：

 （6）

其中，*i*代表用户归属的类别，总类别数由*K*决定；*j*为每一名用户的ID，共计1277名，其他有关于不同类型用户的概率的假设与（2）式中相同。

**三、数据来源及数据处理**

本研究与国网上海市电力公司合作，获取了上海市浦东新区居民2014年1月1日至2016年4月30日共计852天的日用电数据，并从上海气象局浦东监测站获取了对应时间每一天的气象数据，包括最低/最高温度、湿度、风速等。之后，本研究从用电用户中进行随机抽样[[1]](#footnote-1)进行入户调查，共计获得1932名分布于上海市内环、中环、外环以及郊环[[2]](#footnote-2)的用户用电以及基本属性信息。

然而由于上海市智能电表替换时间存在差异，这1932名用户的用电数据起始时间不同，因此在进行分析时需要进行一定的过滤清洗。本研究将每一名用户的用电时间序列作为一行，构建了1932\*852的用电数据矩阵[[3]](#footnote-3)。基于保留信息最大的原则，进行行与列的共同截取，最终保留后828天的用电数据。

此外，对于这些超过两年的用户用电数据，部分用户在2014年时尚未入住于当前住宅，或在这两年之中处于搬迁、租用、用户更替等现象，需要进一步对用户进行清洗。本文在筛选常驻用户时使用冰箱作为筛选条件，一般来说，小型冰箱的日耗电量为0.3kw.h/天[[4]](#footnote-4)，若用户长时间的用电数据低于0.3kw.h，则可以认为该房屋长时间处于空置状态，即该用户不满足研究条件，将予以剔除。在应用中，考虑到用户短时离家不会关闭冰箱的使用，本文取所有用电时长的5%进行筛选，若有超过44天用电低于0.3 kw.h，则可以认为该房产（用户）经历过空置状态，予以删除。

经过上述两步筛选，本研究共计剩余1277名有效用户。之后通过搜索房天下网站，我们获取了所有1277名用户住址对应的住房价格信息以及所处区域，并与样本进行匹配，最终生成本研究需要的数据库。

**四、分析结果**

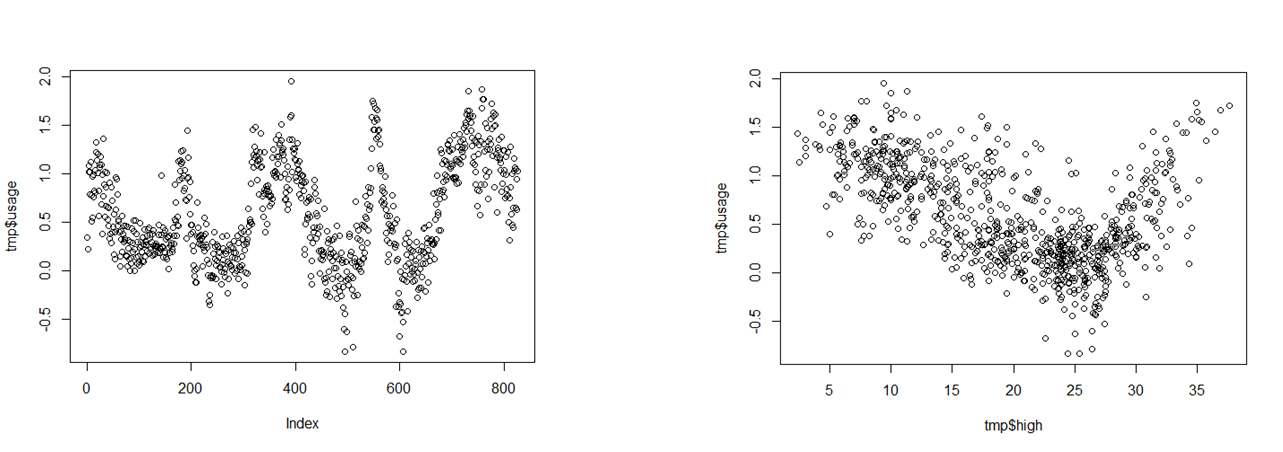
**4.1用电数据的处理**

基于前文对于日用电数据的假设，本文认为1277名用户的828天用电数据是相互独立的，并且每一天的用电数据可以与当日气象数据进行匹配，因此，本文构建了用电量-温度数据，即将每一名用户每一天的用电数据归纳为特定温度下的用电值。其中浦东地区在2014年至2016年828天日最高温和最低温如图1所示。

**图1 上海市浦东新区2014年至2016年日最低最高温度曲线图**

可以发现，上海市浦东新区的日最低最高温度存在极大的相关性。因此在进行温度选取时不会存在温度选取对于结果的影响。本文选取了每日最高温度作为用电温度序列的参考值。如图2所示，经过上述转换过程，可以将用户的用电时间序列转化为用电温度序列，其中左图为用户的用电时间序列，而右图为用户的用电温度序列，其中纵轴为用户家庭人均用电量的对数值，右侧横轴为每一日的最高温度的整合。

**图2 用电时间序列向用电温度序列的转化**



**4.2用户用电量的有限混合分析**

在进行单变量用户用电需求分析时，本文主要是用（1）式进行用户用电温度序列的分析。其中，*K*的选取十分重要，本文使用了纯数据驱动的筛选方法，使用BIC进行用户类别总量的筛选。表1为每个温度下不同分类数目对应的BIC值，其中同一温度下BIC值最小代表拟合程度最好的类别总数。可以发现，在12度以下时，居民的用电数据趋于统一，但是在12~31度以及高于36度时用户的用电需求出现了差异，我们在12至31度以及高于36度时发现了两组用电需求。

**表1 不同温度下用电混合模型的BIC值**

|  |  |  |  |
| --- | --- | --- | --- |
| temperature | number of groups | | |
| 1 | 2 | 3 |
| 2 | **3625.053** | 3643.474 | 3663.251 |
| 3 | **3337.091** | 3357.654 | 3376.104 |
| 4 | **3425.249** | 3447.142 | 3464.545 |
| 5 | **3190.933** | 3208.521 | 3223.523 |
| 6 | **3146.738** | 3163.315 | 3178.276 |
| 7 | **3098.297** | 3104.706 | 3125.778 |
| 8 | **3065.174** | 3071.859 | 3089.64 |
| 9 | **3009.342** | 3013.537 | 3032.68 |
| 10 | **2951.186** | 2951.212 | 2971.618 |
| 11 | **2964.255** | 2968.047 | 2985.805 |
| 12 | 2858.357 | **2853.424** | 2872.637 |
| 13 | 2881.454 | **2878.177** | 2897.45 |
| 14 | 2784.906 | **2781.365** | 2794.431 |
| 15 | 2815.653 | **2811.167** | 2825.397 |
| 16 | 2714.77 | **2707.247** | 2721.47 |
| 17 | 2701.193 | **2697.156** | 2711.908 |
| 18 | 2658.476 | **2654.61** | 2669.389 |
| 19 | 2645.985 | **2643.851** | 2654.662 |
| 20 | 2651.598 | **2641.551** | 2659.739 |
| 21 | 2667.788 | **2664.329** | 2685.187 |
| 22 | 2654.322 | **2640.099** | 2658.613 |
| 23 | 2648.61 | **2635.981** | 2656.063 |
| 24 | 2647.815 | **2635.187** | 2655.996 |
| 25 | 2654.124 | **2637.41** | 2657.624 |
| 26 | 2663.637 | **2639.595** | 2660.97 |
| 27 | 2673.861 | **2652.978** | 2674.719 |
| 28 | 2737.077 | **2719.918** | 2741.568 |
| 29 | 2768.754 | **2761.872** | 2780.957 |
| 30 | 2785.188 | **2781.112** | 2800.325 |
| 31 | **2915.763** | 2925.258 | 2943.616 |
| 32 | **2952.862** | 2968.73 | 2986.384 |
| 33 | **2985.714** | 3006.315 | 3024.255 |
| 34 | **2899.592** | 2921.509 | 2930.945 |
| 35 | **3091.86** | 3102.565 | 3114.952 |
| 36 | 3218.543 | **3199.986** | 3221.518 |
| 37 | 3385.885 | **3354.448** | 3374.775 |

在之后的分析中，我们都使用BIC作为数据分类的评价指标

加粗的数字为比较后数值最小的BIC数，即最优选项

此外，我们在表2中汇总了各个温度下各个类别和样本总体的用电平均水平以及随着温度变化的各组用电需求变动情况。在表2中，我们汇总了每一组的平均的人均用电水平、每一组用户所占整体用户的比例以及组中心到整体平均用电水平的距离。

**表2 不同温度下各组用户用电分布及距离中心距离**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| temperature | high usage group | | |  | low usage group | | | sample mean |
| percentage | mean | relative distance |  | percentage | mean | relative distance |
| 2 |  |  |  |  |  |  |  | 0.782012 |
| 3 |  |  |  |  |  |  |  | 0.564902 |
| 4 |  |  |  |  |  |  |  | 0.633309 |
| 5 |  |  |  |  |  |  |  | 0.592874 |
| 6 |  |  |  |  |  |  |  | 0.509806 |
| 7 |  |  |  |  |  |  |  | 0.511528 |
| 8 |  |  |  |  |  |  |  | 0.461341 |
| 9 |  |  |  |  |  |  |  | 0.441187 |
| 10 |  |  |  |  |  |  |  | 0.43394 |
| 11 |  |  |  |  |  |  |  | 0.394281 |
| 12 | 0.750908 | 0.4578426 | 28.80% |  | 0.249092 | 0.0468694 | -86.81% | 0.355472 |
| 13 | 0.8175785 | 0.4125931 | 22.30% |  | 0.1824215 | 0.0001553 | -99.95% | 0.337356 |
| 14 | 0.7182254 | 0.3637827 | 37.33% |  | 0.2817746 | 0.0128169 | -95.16% | 0.264889 |
| 15 | 0.6945342 | 0.3795289 | 50.03% |  | 0.3054658 | -0.034805 | -113.76% | 0.252964 |
| 16 | 0.5641064 | 0.3509599 | 58.53% |  | 0.4358936 | 0.0536898 | -75.75% | 0.221382 |
| 17 | 0.6367678 | 0.324957 | 55.88% |  | 0.3632322 | 0.0042535 | -97.96% | 0.208467 |
| 18 | 0.5988291 | 0.3099376 | 69.01% |  | 0.4011709 | -0.005529 | -103.01% | 0.183382 |
| 19 | 0.8258895 | 0.2197672 | 54.93% |  | 0.1741105 | -0.227732 | -260.54% | 0.141853 |
| 20 | 0.5427142 | 0.2675277 | 135.80% |  | 0.4572858 | -0.069403 | -161.17% | 0.113454 |
| 21 | 0.5716917 | 0.2636315 | 79.57% |  | 0.4283083 | -0.009111 | -106.21% | 0.146814 |
| 22 | 0.5050591 | 0.2943383 | 156.18% |  | 0.4949409 | -0.068213 | -159.37% | 0.114897 |
| 23 | 0.4573299 | 0.3240923 | 224.77% |  | 0.5426701 | -0.089235 | -189.42% | 0.099792 |
| 24 | 0.4365823 | 0.3526137 | 199.00% |  | 0.5634177 | -0.063924 | -154.21% | 0.117929 |
| 25 | 0.4851051 | 0.3380451 | 155.41% |  | 0.5148949 | -0.061432 | -146.41% | 0.132356 |
| 26 | 0.4509932 | 0.3829038 | 166.49% |  | 0.5490068 | -0.052831 | -136.77% | 0.143682 |
| 27 | 0.4674743 | 0.4047312 | 128.00% |  | 0.5325257 | -0.021949 | -112.36% | 0.177513 |
| 28 | 0.4417733 | 0.5054131 | 88.43% |  | 0.5582267 | 0.0805229 | -69.98% | 0.268228 |
| 29 | 0.4390518 | 0.5953258 | 63.45% |  | 0.5609482 | 0.1833376 | -49.66% | 0.364222 |
| 30 | 0.53591 | 0.5755042 | 47.02% |  | 0.46409 | 0.178922 | -54.29% | 0.391454 |
| 31 |  |  |  |  |  |  |  | 0.560121 |
| 32 |  |  |  |  |  |  |  | 0.624336 |
| 33 |  |  |  |  |  |  |  | 0.772306 |
| 34 |  |  |  |  |  |  |  | 0.667901 |
| 35 |  |  |  |  |  |  |  | 1.059005 |
| 36 | 0.546555 | 1.3567933 | 23.58% |  | 0.453445 | 0.8830958 | -19.56% | 1.097892 |
| 37 | 0.461547 | 1.3801665 | 28.16% |  | 0.538453 | 0.8169292 | -24.14% | 1.07689 |

表中用电均值为人均用电量的对数

距离表示每组均值与样本均值之间的差异

从整体来看，结果表明在温度较低/较高时，居民用户的用电水平相对来说比较统一，即对于制热/制冷的需求都会使用户的用电需求趋于一致，但是在适温区范围内时，用户的用电需求开始分化。从表2中我们可以发现，从12至31度范围，用户应分为两组，即高用电群组以及低用电群组。随着温度的上升，高用电群组的比例呈现逐渐下降的趋势，并且两组用户的用电均值都呈现先下降后上升的趋势，这表明，高用电用户在12至30度内用电曲线呈现U型，这也和我们对于上海市居民用电需求的前期估计相符合，在冬季和夏季温度很低或很高时，上海市居民往往使用空调制热/制冷，从而造成电力负荷的大幅上升。

在适温区范围内，随着温度上升，两组用户的用电差异呈现先扩大再减小的过程，这意味着用户在不使用制冷制热设备时，日常用电需求或习惯有着较大差异。在36~38度时，上海市居民用户的统一化需求发生了变化。我们发现用户中再次出现了两组用电人群，但是这两组人群的用电量差异要小于在12至31度时的范围，并且比例都在50%左右。

**4.3带有伴随变量的有限混合分析**

我们进一步进行了带有伴随变量的用户用电异质性分析，研究用户属性对于用户用电需求的影响。这样的分析可以让我们基于用户各个特性，对于用户未来的用电需求进行预测；反过来，也可以根据用户的用电需求进一步进行用户属性的推断。

在带有伴随变量的有限混合分析中，我们考虑使用收入、房价、地理位置、家庭构成等因素进行分析。我们认为高收入群体、高房价地区以及市中心区域的用户会具有更高的平均用电量。同时根据经验，我们认为家庭成员中包含老人的家庭或以老人为主导的家庭，其用电行为往往更为节约，人均用电量也相对较低。

**4.3.1单一伴随变量的有限混合模型**

首先我们进行的是单一伴随变量的有限混合分析，在单一变量中我们主要考虑了收入以及房价对于人均用户用电需求的影响。在这两个单一变量的有限混合模型中，我们以高用电用户为参照组，得出了低用电用户的用电量以及相应的单一变量在多变量混合logit模型中的参数。结果请参照表3。

我们发现，在任意一组中，收入以及房价这两个单一变量都是十分显著的。以高用电量为参照组，我们发现，收入以及房价的上升会导致其成为低用电用户的概率下降。换言之，收入以及所处区域房价的上升会导致该用户用电量上升的可能性上升。

从平稳性的角度来看，在考虑收入这个伴随变量时，其参数相对平稳，其均值普遍分布在-2至-3范围内；反观房价这一伴随变量，在2至10度范围内，最优组数为三组，但是我们发现在温度非常低的时候，房价对于用电量分类的影响非常高，但是一旦超过5摄氏度，这样的影响将大幅下降，这意味着在温度较低时，高房价区域的用户具有更高的可能性成为高用电用户。

在10到27度范围内，我们发现用户的分类相对统一，仅在14以及23度两个温度段存在差异，并且在分类为4的温度域内，房价对与用电量的影响相对一致，而在温度高于27度时，房价对于用户用电的影响相对减小，这表明，不同房价区域的用户的用电需求差异苏子和温度升高在不断减小，也表明上海地区的制冷需求相对更为普遍。

**表3 单一伴随变量对于用户用电分类的影响**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| temperature | income | | K | price | | K |
| 2 | usage | 0.380202 | 2 | usage | 0.3951 | 3 |
| income | -2.3354(0.830648333)\*\*\* | price | -30.3505(5.7255)\*\*\* |
| 3 | usage | 0.244965 | 2 | usage | 0.2152 | 3 |
| income | -2.48203(0.970201621)\*\*\* | price | -27.0318(5.6619)\*\*\* |
| 4 | usage | 0.269011 | 2 | usage | 0.2461 | 3 |
| income | -2.71617(0.546103225)\*\*\* | price | -29.5499(6.8913)\*\*\* |
| 5 | usage | 0.211678 | 2 | usage | 0.2047 | 3 |
| income | -2.49496(0.749527545)\*\*\* | price | -9.0351(2.8465)\*\*\* |
| 6 | usage | 0.151825 | 2 | usage | 0.1478 | 3 |
| income | -2.6969(0.72847666)\*\*\* | price | -7.3378(1.5239)\*\*\* |
| 7 | usage | 0.159598 | 2 | usage | 0.134 | 3 |
| income | -2.53767(0.624413766)\*\*\* | price | -7.8654(1.7652)\*\*\* |
| 8 | usage | 0.121741 | 2 | usage | 0.0964 | 3 |
| income | -2.56758(0.488448122)\*\*\* | price | -7.6937(1.415)\*\*\* |
| 9 | usage | 0.106732 | 2 | usage | 0.0978 | 3 |
| income | -2.48589(0.570747555)\*\*\* | price | -7.4681(1.4263)\*\*\* |
| 10 | usage | 0.114361 | 2 | usage | 0.1406 | 4 |
| income | -2.62332(0.542288226)\*\*\* | price | -39.1244(7.5711)\*\*\* |
| 11 | usage | 0.056056 | 2 | usage | 0.0927 | 4 |
| income | -2.50452(0.581693153)\*\*\* | price | -40.3226(7.753)\*\*\* |
| 12 | usage | 0.04967 | 2 | usage | 0.0962 | 4 |
| income | -2.44256(0.640035756)\*\*\* | price | -51.3227(20.4705)\*\* |
| 13 | usage | 0.018901 | 2 | usage | 0.0629 | 4 |
| income | -2.54212(0.629740439)\*\*\* | price | -41.3135(7.8447)\*\*\* |
| 14 | usage | -0.02675 | 2 | usage | 0.0661 | 2 |
| income | -2.62742(0.665883757)\*\*\* | price | -4.6768(0.8755)\*\*\* |
| 15 | usage | -0.05207 | 2 | usage | -0.0036 | 4 |
| income | -2.47016(0.427331221)\*\*\* | price | -37.8032(6.8029)\*\*\* |
| 16 | usage | -0.058851565 | 2 | usage | -0.0199 | 4 |
| income | -2.707618(0.406950794)\*\*\* | price | -37.6448(7.1471)\*\*\* |
| 17 | usage | -0.06244 | 2 | usage | -0.0129 | 4 |
| income | -2.68582(0.437537928)\*\*\* | price | -45.5795(23.6586)\* |
| 18 | usage | -0.080465181 | 2 | usage | -0.0234 | 4 |
| income | -2.707329(0.524699408)\*\*\* | price | -48.6512(25.8784)\* |
| 19 | usage | -0.1237 | 2 | usage | -0.0679 | 4 |
| income | -2.68695(0.647010319)\*\*\* | price | -36.2391(9.3042)\*\*\* |
| 20 | usage | -0.148960902 | 2 | usage | -0.0766 | 4 |
| income | -2.629178(0.596029453)\*\*\* | price | -75.9728(20.3612)\*\*\* |
| 21 | usage | -0.114342229 | 2 | usage | -0.0599 | 4 |
| income | -2.691655(0.582149795)\*\*\* | price | -41.4765(10.4473)\*\*\* |
| 22 | usage | -0.136369339 | 2 | usage | -0.0728 | 4 |
| income | -2.531551(0.582907584)\*\*\* | price | -58.7499(14.1671)\*\*\* |
| 23 | usage | -0.13627 | 2 | usage | -0.071 | 2 |
| income | -2.54219(0.52806957)\*\*\* | price | -3.1163(0.6033)\*\*\* |
| 24 | usage | -0.12014 | 2 | usage | -0.0993 | 4 |
| income | -2.57277(0.652038892)\*\*\* | price | -40.8755(7.8493)\*\*\* |
| 25 | usage | -0.10247 | 2 | usage | -0.0986 | 3 |
| income | -2.49017(0.588546017)\*\*\* | price | -26.0933(5.8353)\*\*\* |
| 26 | usage | -0.08728 | 2 | usage | -0.0687 | 4 |
| income | -2.54292(0.640522376)\*\*\* | price | -39.4124(11.1304)\*\*\* |
| 27 | usage | -0.05824 | 2 | usage | 0.0125 | 2 |
| income | -2.43991(0.350522731)\*\*\* | price | -3.4181(0.6686)\*\*\* |
| 28 | usage | 0.040843 | 2 | usage | 0.0919 | 2 |
| income | -2.53599(0.368785062)\*\*\* | price | -3.5389(0.6495)\*\*\* |
| 29 | usage | 0.128922 | 2 | usage | 0.1111 | 3 |
| income | -2.59203(0.497783643)\*\*\* | price | -23.9795(4.9056)\*\*\* |
| 30 | usage | 0.146519 | 2 | usage | 0.2058 | 2 |
| income | -2.51645(0.544420941)\*\*\* | price | -4.0563(0.7648)\*\*\* |
| 31 | usage | 0.296219 | 2 | usage | 0.3351 | 2 |
| income | -2.62947(0.591353325)\*\*\* | price | -4.0691(0.7077)\*\*\* |
| 32 | usage | 0.355534 | 2 | usage | 0.4022 | 2 |
| income | -2.67358(0.533210007)\*\*\* | price | -4.1728(0.7366)\*\*\* |
| 33 | usage | 0.518091 | 2 | usage | 0.5628 | 2 |
| income | -2.70428(0.529567483)\*\*\* | price | -4.4181(0.8972)\*\*\* |
| 34 | usage | 0.447029 | 2 | usage | 0.4661 | 3 |
| income | -3.06391(0.498348443)\*\*\* | price | -5.5835(1.24)\*\*\* |
| 35 | usage | 0.82489 | 2 | usage | 0.6557 | 3 |
| income | -2.71733(0.530246713)\*\*\* | price | -4.4906(0.9448)\*\*\* |
| 36 | usage | 0.756023 | 3 | usage | 0.778 | 3 |
| income | -2.43189(0.5456286739)\*\*\* | price | -4.2822(1.0029)\*\*\* |
| 37 | usage | 0.713127 | 3 | usage | 0.6681 | 3 |
| income | -2.6875(0.462812957)\*\*\* | price | -4.3143(1.0608)\*\*\* |

表中用电均值为人均用电量的对数

表中收入为家庭月收入的对数，房价为房价的对数形式

K为该温度范围内根据BIC选择的最优组数

\*，\*\*，\*\*\*分别表示在10%，5%以及1%的统计意义上显著

**4.3.2多伴随变量的有限混合模型**

接下来我们考虑多伴随变量的有限混合模型，在这里我们主要使用了家庭人口构成以及家庭地理位置信息进行分析。家庭人口构成以及住宅地理位置信息的分析结果请看表4、5。在进行家庭人口构成分析时，我们主要考虑了四种人口构成，分别是小于6岁的幼龄儿童、正在上学的少年或青年、已经有工作的成年人，以及大于60岁的退休老人。在进行家庭地理位置信息分析时，我们根据居民地理位置所处区域（内环、中环、外环），对其地理位置进行了划分，使用哑变量进行表示[[5]](#footnote-5)，其中作为参考项的是内环内区域。

在进行家庭人口结构对于用电需求的影响分析时，我们依旧使用高用电用户作为参考项。从表4中我们发现，大多数参数在所有温度范围内都是显著的，但是在29至31度，家庭人口结构对于用户的人均用电需求影响不显著，并且，相比于其他参数，6岁以下儿童对于家庭人均用电需求的影响具有很多不显著项，这也符合婴幼儿对于家庭用电需求影响较小的预期。

从总体上来看，家庭人口成员的增加（增加任意一名处于不同年龄的人员）都将导致用户的人均用电需求趋于减少。这就说明，家庭人数的增加可以增加用电的效率，从而有更多的电力使用被分摊了，如照明、制冷制热等。另一方面，我们发现，在所有温度范围内，增加一名老人相对于增加其他年龄的用户，对于人均用电需求的减弱效应更大，这也和我们的预期相符合，即老人相对于其他人群更为“节能”。这可能是由于老人自身能源需求较低，并且老人往往具有很强的节能习惯，并会对其他家庭成员进行督促导致的。并且，正在工作的成年人对于家庭能源消耗的影响最小，也就是说，有工作的成年人往往更“浪费”，增加一名成年人对于家庭整体用电效率的提升有限。

**表4 家庭成员对于用户人均用电情况的影响**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| temperature | usage | infant | teenager | work | old | K |
| 2 | 0.5541 | 1.1875(0.862) | 2.6223(1.2874)\*\* | 1.2197(0.6086)\*\* | 3.7618(1.7304)\*\* | 2 |
| 3 | 0.3479 | 1.2173(0.8097) | 2.4012(1.1874)\*\* | 2.0384(0.9644)\*\* | 4.2164(1.9002)\*\* | 2 |
| 4 | 0.4572 | 1.4596(1.4488) | 4.1823(2.4799)\* | 2.0819(1.2348)\* | 10.5904(6.5253) | 2 |
| 5 | 0.3884 | 1.5702(0.9026)\* | 2.2938(1.0595)\*\* | 1.8457(0.8074)\*\* | 4.3125(1.7545)\*\* | 2 |
| 6 | 0.3009 | 1.6904(0.8499)\*\* | 2.6269(0.9959)\*\*\* | 1.7045(0.6442)\*\*\* | 4.1712(1.4106)\*\*\* | 2 |
| 7 | 0.3053 | 1.4388(0.7652)\* | 2.2643(0.9539)\*\* | 1.6385(0.6669)\*\* | 3.8244(1.4316)\*\*\* | 2 |
| 8 | 0.2533 | 1.4845(0.7453)\*\* | 2.2304(0.9054)\*\* | 1.6136(0.6229)\*\*\* | 3.5913(1.3097)\*\*\* | 2 |
| 9 | 0.2293 | 1.5134(0.7109)\*\* | 2.0971(0.831)\*\* | 1.5503(0.5673)\*\*\* | 3.368(1.2144)\*\*\* | 2 |
| 10 | 0.2204 | 1.5784(0.7191)\*\* | 2.2191(0.8461)\*\*\* | 1.5284(0.5431)\*\*\* | 3.3415(1.1509)\*\*\* | 2 |
| 11 | 0.1824 | 1.5124(0.7028)\*\* | 2.1508(0.88)\*\* | 1.4624(0.5495)\*\*\* | 3.3431(1.2275)\*\*\* | 2 |
| 12 | 0.0433 | 0.8382(0.3501)\*\* | 0.9317(0.2718)\*\*\* | 0.6785(0.1792)\*\*\* | 1.33(0.3074)\*\*\* | 2 |
| 13 | 0.1091 | 1.3847(0.6625)\*\* | 1.5617(0.8438)\* | 1.1306(0.5081)\*\* | 2.3704(1.1276)\*\* | 2 |
| 14 | 0.0351 | 1.5883(0.6153)\*\*\* | 1.9764(0.691)\*\*\* | 1.3531(0.3862)\*\*\* | 2.7841(0.8207)\*\*\* | 2 |
| 15 | -0.0419 | 1.0314(0.5837)\* | 1.0381(0.4595)\*\* | 0.8653(0.3382)\*\* | 1.6094(0.6917)\*\* | 2 |
| 16 | -0.015 | 1.6676(0.5906)\*\*\* | 2.0055(0.6421)\*\*\* | 1.3607(0.3553)\*\*\* | 2.8448(0.7442)\*\*\* | 2 |
| 17 | -0.023 | 1.7489(0.6571)\*\*\* | 2.0564(0.6429)\*\*\* | 1.3331(0.3412)\*\*\* | 2.7862(0.7069)\*\*\* | 2 |
| 18 | -0.0513 | 1.8076(0.6717)\*\*\* | 2.0877(0.6078)\*\*\* | 1.3385(0.3159)\*\*\* | 2.7058(0.6471)\*\*\* | 2 |
| 19 | -0.0993 | 1.7556(0.5995)\*\*\* | 2.2293(0.5691)\*\*\* | 1.4478(0.3163)\*\*\* | 2.8909(0.6193)\*\*\* | 2 |
| 20 | -0.1275 | 1.8208(0.6435)\*\*\* | 2.1258(0.5714)\*\*\* | 1.4045(0.31)\*\*\* | 2.7263(0.604)\*\*\* | 2 |
| 21 | -0.0828 | 1.7249(0.6155)\*\*\* | 2.3294(0.6779)\*\*\* | 1.4323(0.3502)\*\*\* | 3.0976(0.7436)\*\*\* | 2 |
| 22 | -0.1288 | 9.4639(36.7131) | 2.4095(0.7357)\*\*\* | 1.6056(0.3608)\*\*\* | 2.8619(0.7288)\*\*\* | 3 |
| 23 | -0.2141 | 14.8931(486.2701) | 6.3249(2.3743)\*\*\* | 7.0696(3.5186)\*\* | 10.2736(4.0343)\*\* | 4 |
| 24 | -0.1918 | 13.7062(314.5686) | 6.3004(2.3645)\*\*\* | 7.4819(5.1185) | 10.7113(5.4117)\*\* | 4 |
| 25 | -0.2117 | 14.6243(373.8169) | 5.0186(1.8521)\*\*\* | 5.3563(1.6791)\*\*\* | 7.8965(2.4277)\*\*\* | 4 |
| 26 | -0.2013 | 14.063(319.8829) | 5.1298(1.6775)\*\*\* | 5.1267(1.4784)\*\*\* | 7.5591(2.0876)\*\*\* | 4 |
| 27 | -0.1552 | 1.4309(0.5868)\*\* | 3.3416(0.9807)\*\*\* | 2.1758(0.7649)\*\*\* | 4.3707(1.3171)\*\*\* | 3 |
| 28 | -0.0525 | 1.4176(0.6637)\*\* | 3.5699(1.2855)\*\*\* | 2.4297(1.0136)\*\* | 4.8453(1.7142)\*\*\* | 3 |
| 29 | 0.0658 | 1.393(1.3716) | 0.026(0.3034) | 0.1598(0.2452) | 0.3236(0.2709) | 3 |
| 30 | 0.0925 | 1.7792(1.175) | 0.0138(0.2616) | 0.1333(0.2397) | 0.3293(0.2536) | 3 |
| 31 | 0.3352 | 1.2858(0.5633)\*\* | 2.6099(0.8635)\*\*\* | 1.9582(0.6333)\*\*\* | 3.7628(1.1313)\*\*\* | 2 |
| 32 | 0.4062 | 0.9902(0.5808)\* | 3.0007(1.0753)\*\*\* | 2.2721(0.7735)\*\*\* | 4.3445(1.3219)\*\*\* | 2 |
| 33 | 0.5524 | 0.9736(0.5949) | 3.169(1.0479)\*\*\* | 2.3168(0.7599)\*\*\* | 4.5054(1.2803)\*\*\* | 2 |
| 34 | 0.4469 | 1.4011(0.6049)\*\* | 2.3192(0.6747)\*\*\* | 1.9471(0.4965)\*\*\* | 3.7646(0.891)\*\*\* | 2 |
| 35 | 0.8268 | 0.9597(0.4882)\*\* | 1.8157(0.4926)\*\*\* | 1.591(0.4069)\*\*\* | 3.052(0.6828)\*\*\* | 2 |
| 36 | 0.8439 | 1.0457(0.4623)\*\* | 1.8065(0.4744)\*\*\* | 1.3784(0.3566)\*\*\* | 2.7275(0.6248)\*\*\* | 2 |
| 37 | 0.5852 | 0.7008(0.405)\* | 1.604(0.4323)\*\*\* | 1.4023(0.3601)\*\*\* | 2.3083(0.5683)\*\*\* | 3 |

表中用电均值为人均用电量的对数

K为该文度范围内的BIC选择的最优组数

\*，\*\*，\*\*\*分别表示在10%，5%以及1%的统计意义上显著

表5是居民地理位置对于其用电需求影响的分析结果。我们使用低用电群组作为参考项。首先，所有的截距项都是显著的，并且数值相对较大，普遍在8左右，这说明内环内居民更容易成为用电高的用户。并且，所有中环及以外的位置变量的参数也都是显著的，全部为负且随着距离中心区域距离的扩大而不断扩大，这说明在居民用户用电需求方面确实存在严重的地理位置效应，即相比于中心区域用户，居民的地理位置随着距离中心区域越远，其用电相对越低。

相比于中环及以外的用户，中环及以内用户在各个温度端内的影响显著性明显低于中环及以外的位置变量，在22~31度范围内，我们可以发现相比于内环内区域，内环至中环范围内的用户成为高用电群组的概率要小很多，但是这样的影响要远远小于中环及以外的位置变量的影响。并且在部分温度下，内换上用户相比于内环内用户成为高用电类型用户的比例会升高，实际计算来看，二者的概率都接近于1。

这样的结果表明，在地理位置上，上海市浦东新区的居民用电需求存在明显的“断层”，并且这样的断层从中环开始。中环及以外的用户相比于中环以内的用户，其成为高用电用户的概率要下降非常多，而在各个温度段，中环以内的用户成为高用电类型的概率几乎都为1。

**表5 家庭地理位置对于用户人均用电情况的影响[[6]](#footnote-6)**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 温度 | 内环内（截距） | 内环上 | 中环内 | 中环上 | 外环内 | 外环及以外 |
| 2 | 6.987(0.096)\*\*\* | 2.724(0.243)\*\*\* | 0.622(0.986) | -5.466(0.086)\*\*\* | -7.105(0.089)\*\*\* | -13.633(0.124)\*\*\* |
| 3 | 8.121(0.153)\*\*\* | 0.72(0.366)\*\* | -0.17(0.29) | -6.296(0.153)\*\*\* | -8.436(0.153)\*\*\* | -10.631(0.153)\*\*\* |
| 4 | 8.067(0.166)\*\*\* | 0.901(0.329)\*\*\* | -0.102(1.656) | -6.961(0.176)\*\*\* | -8.955(0.178)\*\*\* | -12.932(0.24)\*\*\* |
| 5 | 8.489(0.332)\*\*\* | -0.03(0.48) | -0.273(0.394) | -7.125(0.332)\*\*\* | -9.206(0.332)\*\*\* | -11.185(0.332)\*\*\* |
| 6 | 8.59(0.285)\*\*\* | 1.315(0.807) | -6.308(2.664)\*\*\* | -7.249(0.258)\*\*\* | -9.416(0.266)\*\*\* | -11.157(0.27)\*\*\* |
| 7 | 8.603(0.286)\*\*\* | -0.074(0.429) | -0.2(0.388) | -7.173(0.286)\*\*\* | -9.3(0.286)\*\*\* | -11.2(0.286)\*\*\* |
| 8 | 8.845(0.348)\*\*\* | 0.251(0.503) | 0.091(0.467) | -7.382(0.348)\*\*\* | -9.476(0.348)\*\*\* | -11.397(0.348)\*\*\* |
| 9 | 8.661(0.209)\*\*\* | 1.347(0.619)\*\* | -6.062(0.818)\*\*\* | -7.351(0.194)\*\*\* | -9.398(0.201)\*\*\* | -11.17(0.203)\*\*\* |
| 10 | 8.808(0.988)\*\*\* | -0.326(0.455) | -0.797(0.406)\*\* | -7.273(0.835)\*\*\* | -9.455(1.053)\*\*\* | -11.275(1.237)\*\*\* |
| 11 | 8.775(0.315)\*\*\* | -0.415(0.436) | -0.799(0.4)\*\* | -7.146(0.315)\*\*\* | -9.471(0.315)\*\*\* | -11.293(0.315)\*\*\* |
| 12 | 9.102(0.289)\*\*\* | -0.349(0.716) | -1.39(2.616) | -7.678(0.279)\*\*\* | -9.768(0.282)\*\*\* | -11.468(0.284)\*\*\* |
| 13 | 9.092(0.261)\*\*\* | -0.461(0.473) | -0.299(0.383) | -7.566(0.261)\*\*\* | -9.899(0.261)\*\*\* | -11.53(0.261)\*\*\* |
| 14 | 9.063(0.245)\*\*\* | -0.682(0.417) | -0.357(0.346) | -7.545(0.245)\*\*\* | -9.855(0.245)\*\*\* | -11.459(0.245)\*\*\* |
| 15 | 8.543(0.278)\*\*\* | 0.194(0.362) | 0.254(0.371) | -7.105(0.278)\*\*\* | -9.365(0.278)\*\*\* | -10.916(0.278)\*\*\* |
| 16 | 9.105(0.291)\*\*\* | -0.23(0.372) | 0.128(0.354) | -7.341(0.291)\*\*\* | -9.948(0.291)\*\*\* | -11.522(0.291)\*\*\* |
| 17 | 8.837(0.386)\*\*\* | -0.004(0.539) | 0.425(0.526) | -7.532(0.386)\*\*\* | -9.713(0.386)\*\*\* | -11.185(0.386)\*\*\* |
| 18 | 8.971(0.341)\*\*\* | -0.057(0.43) | 0.039(0.43) | -7.765(0.341)\*\*\* | -9.707(0.341)\*\*\* | -11.227(0.341)\*\*\* |
| 19 | 8.661(0.276)\*\*\* | -0.168(0.334) | 0.199(0.337) | -7.407(0.276)\*\*\* | -9.415(0.276)\*\*\* | -10.777(0.276)\*\*\* |
| 20 | 8.687(0.277)\*\*\* | -0.326(0.42) | 0.537(0.413) | -7.568(0.277)\*\*\* | -9.598(0.277)\*\*\* | -10.9(0.277)\*\*\* |
| 21 | 8.817(0.32)\*\*\* | -0.166(0.419) | 0.345(0.396) | -7.586(0.32)\*\*\* | -9.753(0.32)\*\*\* | -11.094(0.32)\*\*\* |
| 22 | 9.707(0.253)\*\*\* | -1.874(0.354)\*\*\* | -1.223(0.336)\*\*\* | -8.776(0.253)\*\*\* | -10.613(0.253)\*\*\* | -11.936(0.253)\*\*\* |
| 23 | 10.148(0.317)\*\*\* | -1.985(0.361)\*\*\* | -1.465(0.363)\*\*\* | -9.304(0.317)\*\*\* | -11.032(0.317)\*\*\* | -12.386(0.317)\*\*\* |
| 24 | 10.147(0.403)\*\*\* | -2.209(0.443)\*\*\* | -1.901(0.451)\*\*\* | -9.154(0.403)\*\*\* | -11.023(0.403)\*\*\* | -12.385(0.403)\*\*\* |
| 25 | 9.79(0.251)\*\*\* | -1.787(0.271)\*\*\* | -1.559(0.282)\*\*\* | -8.903(0.251)\*\*\* | -10.625(0.251)\*\*\* | -12.045(0.251)\*\*\* |
| 26 | 10.02(0.398)\*\*\* | -2.056(0.467)\*\*\* | -1.653(0.468)\*\*\* | -9.299(0.398)\*\*\* | -10.931(0.398)\*\*\* | -12.315(0.398)\*\*\* |
| 27 | 10.109(0.29)\*\*\* | -2.054(0.249)\*\*\* | -1.89(0.248)\*\*\* | -9.392(0.29)\*\*\* | -11.024(0.29)\*\*\* | -12.487(0.29)\*\*\* |
| 28 | 10.072(0.281)\*\*\* | -1.798(0.347)\*\*\* | -1.495(0.363)\*\*\* | -9.19(0.281)\*\*\* | -10.877(0.281)\*\*\* | -12.387(0.281)\*\*\* |
| 29 | 9.436(0.43)\*\*\* | -1.357(0.448)\*\*\* | -1.505(1.802) | -8.227(0.384)\*\*\* | -10.168(0.408)\*\*\* | -11.753(0.413)\*\*\* |
| 30 | 9.948(0.289)\*\*\* | -1.789(0.357)\*\*\* | -2.042(0.391)\*\*\* | -8.749(0.289)\*\*\* | -10.66(0.289)\*\*\* | -12.301(0.289)\*\*\* |
| 31 | 8.735(0.313)\*\*\* | -0.396(0.348) | -0.864(0.339)\*\*\* | -7.501(0.313)\*\*\* | -9.263(0.313)\*\*\* | -11.129(0.313)\*\*\* |
| 32 | 8.919(0.333)\*\*\* | -0.743(0.455) | -0.555(0.457) | -7.779(0.334)\*\*\* | -9.388(0.333)\*\*\* | -11.385(0.333)\*\*\* |
| 33 | 8.557(0.297)\*\*\* | -0.207(0.459) | -0.354(0.328) | -7.192(0.297)\*\*\* | -8.866(0.297)\*\*\* | -11.263(0.297)\*\*\* |
| 34 | 8.206(0.206)\*\*\* | 0.313(0.744) | -6.272(0.206)\*\*\* | -7.644(0.206)\*\*\* | -9.404(0.206)\*\*\* | -17.387(0.26)\*\*\* |
| 35 | 6.747(0.154)\*\*\* | 1.996(0.533)\*\*\* | 2.629(0.131)\*\*\* | -5.41(0.153)\*\*\* | -7.036(0.153)\*\*\* | -14.362(0.174)\*\*\* |
| 36 | 7.844(0.196)\*\*\* | 2.633(1.111)\*\*\* | -6.72(0.196)\*\*\* | -7.312(0.196)\*\*\* | -8.601(0.196)\*\*\* | -15.61(0.205)\*\*\* |
| 37 | 8.515(0.276)\*\*\* | -1.535(0.425)\*\*\* | -7.819(0.175)\*\*\* | -7.473(1.287)\*\*\* | -8.379(0.876)\*\*\* | -13.783(0.653)\*\*\* |

表中分析使用的用电数据为人均用电量的对数

\*，\*\*，\*\*\*分别表示在10%，5%以及1%的统计意义上显著

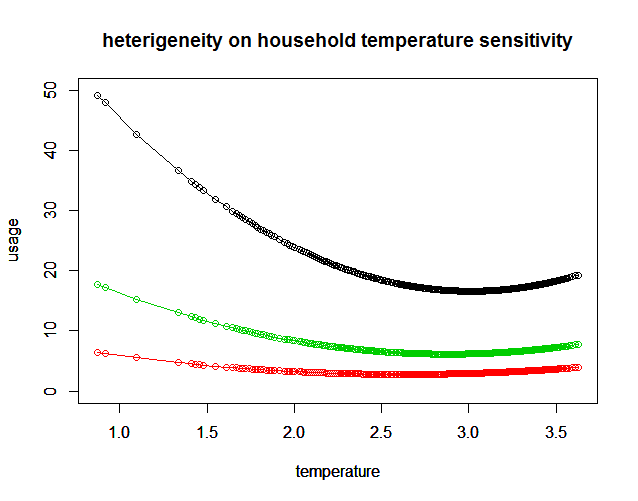
**4.4温度敏感性的有限混合分析**

我们最后以一名用户的全部用电数据为基准，对全样本数据进行分析。从图2中我们可以发现，在转化为温度用电序列后，居民的用电-温度曲线呈现明显的U型，这意味着单纯的一次项回归无法完成用户用电需求的拟合，添加温度的二次项是相对比较适合的方法。因此，在接下来的分析中，我们使用的回归方程如下：

 （7）

其中*y*代表一名用户*i*在第*t*天的用电量，*temperature*代表当天的最高温度，为其他因素对于当日用电的影响。在这个过程中，每一名用户*i*都具有两个参数和，二者组合将代表用户*i*的温度敏感性以及适温区间，对这两个参数进行分类可以将用户在温度敏感性层面上进行区分。因此我们使用（6）式对每一名居民的用电温度敏感性进行了分类

**图3 三种用电敏感性用户及其特征曲线**



根据回归结果，我们可以在居民用户用电敏感性层面上将其分为三类：高敏感性、中敏感性以及低敏感性，结果见表6。三类用户的用电曲线如图3所示，横轴为当日最高温度的对数值，纵轴为当日家庭用电总量。我们可以发现相比于其他两类用户，高敏感性用户具有很强的U型特征，并且，在温度较低时其曲线弧度相对较大，这说明高敏感性用户对于低温更加敏感，低温会迫使高敏感性用户是用更多的制热设备，从而导致用电量快速上升。

而红色的低敏感性用户的用电曲线几乎水平，在温度较低时虽然存在一定的上升，但是程度非常小，与平时用电需求差异不大。绿色的用电用户的温度敏感性介于二者之间。

**表6 用户属性与温度敏感性有限混合分析结果[[7]](#footnote-7)**

|  |  |  |  |
| --- | --- | --- | --- |
|  | component 1 | component 2 | component 3 |
| component variables | | | |
| Intercept | 81.429(0.8997)\*\*\* | 30.090(0.1412)\*\*\* | 10.859(0.1284)\*\*\* |
| temperature | -43.037(0.6963)\*\*\* | -16.634(0.1075)\*\*\* | -6.145(0.1033)\*\*\* |
| square temperature | 7.144(0.1316)\*\*\* | 2.888(0.02019)\*\*\* | 1.167(0.0192)\*\*\* |
| Number of household | 165 | 618 | 494 |
| concomitant variables | | | |
| Intercept | - | 31.68(2.97)\*\*\* | 21.91(2.91)\*\*\* |
| price | - | -2.27(0.26)\*\*\* | -1.83(0.25)\*\*\* |
| income | - | -0.40(0.20)\*\* | -0.05(0.19) |
| square foot | - | -0.43(0.11)\*\*\* | -0.40(0.11)\* |
| number of household | - | -0.30(0.08)\*\*\* | -0.06(0.07) |

\*，\*\*，\*\*\*分别表示在10%，5%以及1%的统计意义上显著

温度表示当日最高温度的对数值

表6是对于用电温度敏感性分析的分析结果。首先我们发现，用户的绝对用电量以及温度敏感性可以分为三类，并且所有的温度敏感性都是显著的，其中高敏感性用户相对较低少，而低敏感性以及中敏感性用户相对较多，大概占总体的90%。

在分析伴随变量时，我们考虑了房价、家庭收入、房间面积以及家庭人口数。我们发现，在以高用电用户为参考项时，房价、收入、住房面积以及家庭人口数[[8]](#footnote-8)的增加都会导致成为高用电用户的概率上升。从前文结论可以推断，在中环及以外的用户敏感性相比于中环以内用户的，成为高用电敏感性的概率要低。

**五、结论及政策建议**

本文获取了上海市浦东新区居民的日用电数据，并通过随机抽样的方式从中选取2000名用户进行了入户调研，深入了解了家庭收入、房价、人口构成等相关信息，进一步结合气象数据，对用电数据进行了分析。通过使用有限混合模型，本文深入挖掘了用户用电需求上的异质性，以及异质性与用户基本属性之间的关系。

本文发现，用户的基础用电在高温、低温区域会趋于统一，但是在适温区域，用户的用电需求会发生较大的差异，这表明，从总体上看，用户的制冷制热需求是趋于统一的；但是在日常生活中却会出现较大差异。进一步本文挖掘了用户属性与用电需求之间的关系，如我们的预期，用户的用电需求与用户收入与房价关系显著，房价以及收入较高的用户将更有可能成为高用电用户，并且在所有温度范围内，这样的影响都是显著的。另一方面，在家庭人口构成方面，家庭人口数较多的用户，其单位用电水平会降低，这意味着用电效率的提升。而由于老人相比于其他用户更加“节约”，相比于儿童和成年人，具有老人的家庭会更容易成为低用电用户；此外，在地理位置方面，我们发现了上海市居民用电会随着环线附近出现明显的断层现象，从中环向外拓展的区域出现低用电需求的概率将大幅上升，而中环以内的用户成为高用电用户的概率几乎为1，这意味着用电集聚性在上海浦东区域非常明显，即用电需求在地理位置的分布上出现了明显的断层。

接下来，通过使用重复测量的有限混合模型，本文对用户的总体用电曲线进行了研究，本文发现通过使用二次项拟合，可以将上海市居民的用电需求敏感性分为三类，即高温度敏感、中温度敏感以及低温度敏感型用户。并且，用户类型与用户的基本属性相关，同前文结论类似，收入、房价较高以及靠近中心区域的用户的将具有更大的可能性成为高用电用户，而随着家庭人口数的上升，家庭用电总量以及温度敏感性也会有所上升。

基于这样的结论，我们发现了居民用户的异质性以及异质性与居民属性的关系。居民在上海市用电需求中占有重要比重，并且随着经济的发展在不断上升。对居民用电需求进行挖掘，可以有效增加对于居民用户用电量预测的准确性，这对于整个电力系统的优化、设计、缓解拥堵有着重要的作用。本文是作者所知第一个使用高频用电数据进行用户需求分析的研究，这样的高频数据可以让我们直接将用户的用电需求和气象数据等匹配，从而避免使用HDD或CDD等哑变量代表气候变化，从而提供相比于月度、年度数据更有效的提供用户的温度敏感性测算。基于1277名用户有效数据的分析，使用有限混合的分析方式，我们可以准确识别用户的异质性，从而避免了传统经济学较为严格的一致性假设，这样的分析结果将更有效、准确。

**参考文献**

Alberini, A., Filippini, M., 2011. Response of residential electricity demand to price: The effect of measurement error. Energy Economics 33, 889-895.

Alberini, A., Gans, W., Velez-Lopez, D., 2011. Residential consumption of gas and electricity in the US: The role of prices and income. Energy Economics 33, 870-881.

Alfo, M., Trovato, G., Waldmann, R.J., 2008. Testing for country heterogeneity in growth models using a finite mixture approach. Journal of applied Econometrics 23, 487-514.

Allenby, G.M., Rossi, P.E., 1998. Marketing models of consumer heterogeneity. J ECONOMETRICS 89, 57-78.

Bernard, J., Bolduc, D., Yameogo, N., 2011. A pseudo-panel data model of household electricity demand. Resource and Energy Economics 33, 315-325.

Blázquez, L., Boogen, N., Filippini, M., 2013. Residential electricity demand in Spain: new empirical evidence using aggregate data. Energy economics 36, 648-657.

Dayton, C.M., 2008. An introduction to latent class analysis. Handbook of longitudinal research: Design, measurement, and analysis, 357-371.

Dayton, C.M., Macready, G.B., 1988. Concomitant-variable latent-class models. J AM STAT ASSOC 83, 173-178.

Deb, P., Trivedi, P.K., 1997. Demand for medical care by the elderly: a finite mixture approach. Journal of applied Econometrics 12, 313-336.

Fell, H., Li, S., Paul, A., 2014. A new look at residential electricity demand using household expenditure data. International Journal of Industrial Organization 33, 37-47.

Gans, W., Alberini, A., Longo, A., 2013. Smart meter devices and the effect of feedback on residential electricity consumption: Evidence from a natural experiment in Northern Ireland. Energy Economics 36, 729-743.

Garrod, G., Ruto, E., Willis, K., Powe, N., 2012. Heterogeneity of preferences for the benefits of environmental stewardship: a latent-class approach. ECOL ECON 76, 104-111.

Gerdtham, U.G., Trivedi, P.K., 2001. Equity in Swedish health care reconsidered: new results based on the finite mixture model. HEALTH ECON 10, 565-572.

Maitra, S., 2016. The poor get poorer: Tracking relative poverty in India using a durables-based mixture model. Journal of Development Economics 119, 110-120.

Maitra, S., 2016. Class matters: tracking urban inequality in post-liberalization India using a durables-based mixture model. Applied Economics Letters, 1-5.

Masyn, K.E., 2013. Latent class analysis and finite mixture modeling.

McLachlan, G., Peel, D., 2004. Finite mixture models. John Wiley & Sons.

Melnikov, N.B., O'Neill, B.C., Dalton, M.G., 2012. Accounting for household heterogeneity in general equilibrium economic growth models. Energy economics 34, 1475-1483.

Reiss, P.C., White, M.W., 2005. Household electricity demand, revisited. The Review of Economic Studies 72, 853-883.

1. 根据用户在电力公司的用户ID，通过随机抽样的方法从中进行抽样。 [↑](#footnote-ref-1)
2. 上海市根据环城高架桥公路被分为几个同心圆区域，对应分别为内环，中环以及外环。 [↑](#footnote-ref-2)
3. 在这里，该矩阵共有1932行，852列数据，每一行代表一名用户的用电时间序列，每一列代表一天内用户用电数据的截面。因为安装时间不同，部分数据存在空值，因此将数据进行一定的筛选。 [↑](#footnote-ref-3)
4. 本文查找了京东商城单柜门冰箱，其日耗电量为0.31kw.h，考虑到家庭冰箱应至少为双开门冰箱，容积应远大于单开门冰箱，故此0.3这一阈值可以有效筛选空置与否的用户。 [↑](#footnote-ref-4)
5. 由于部分居民所处地区处于环线边缘，因此我们在设定哑变量时会额外考虑环线附近的用户，因此会出现“内环内”或“内环上”的情况，其中内环内距离市中心相比于内换上更近。 [↑](#footnote-ref-5)
6. 由于使用的是哑变量，本研究使用内环内作为参考标准，因此截距即代表了内环内用户成为低用电用户的参数。 [↑](#footnote-ref-6)
7. 此处分析时使用的是家庭总用电量。因为用电的温度敏感性具有家庭决策性，对于家庭层面的分析更具有意义。 [↑](#footnote-ref-7)
8. 不同于人均用电量，由于这里使用的是家庭总用电量，因此家庭人口数上升一定会导致用电需求的上升。 [↑](#footnote-ref-8)