

中国家庭幸福感测量

社会
2013·6
CJS
第33卷

王广州 王 军

摘要:中国传统社会历来是“家本位”,测量家庭幸福感在中国具有特殊意义。为克服家庭幸福感测量中存在的自评异质性问题,本文采用全国范围有代表性的概率样本调查数据,引入“测量锚点”和“等比例方法”对家庭幸福感受评分进行标准化,在此基础上,进而分析家庭幸福感自评异质性的影响因素以及标准化与非标准化测量的评分差异,从而揭示以往主观评价变量非标准化测量可能存在的问题。最后,本文以社区融合状况对居民家庭幸福感的影响为例,通过分层线性回归模型比较了家庭幸福感标准化与非标准化测量在研究结论方面存在的显著差异。

关键词:家庭幸福感 主观测量 测量锚点 标准化

The Measurement of Chinese Family Happiness

WANG Guangzhou WANG Jun

Abstract: Subjective happiness, health and life satisfaction are usually measured by self-reported ratings. However, studies have shown that, in this type of measurement of subjective indices, there is a problem of “self-rating heterogeneity”, that is, the same score given by different respondents may very well represent quite different states in reality. This has been extensively evidenced in health measures. Considering the fact of family being the basic unit in the traditional Chinese society, measuring family happiness is, therefore, of unique significance. Taking the measurement of family happiness as an example, this study used the survey data from a representative national sample to examine if there was a problem of “self-rating heterogeneity” and if so, how to solve that problem. The findings revealed clear existence of this problem, including significant differences in the family happiness standards

* 作者1:王广州 中国社会科学院人口与劳动经济研究所(Author 1: WANG Guangzhou, Institute of Population and Labor Economics, CASS) E-mail: Wangguangzhou@126.com;作者2:王 军 中国社会科学院人口与劳动经济研究所(Author 2: WANG Jun, Institute of Population and Labor Economics, CASS)

used by the people in urban vs. rural areas, and between people of different ages. To address the problem of “self-rating heterogeneity,” this study employed measurement anchors and used the “equal ratio method” to standardize the family happiness scores. Then the factors that influenced the “self-rating heterogeneity” were analyzed and the standardized scores were compared with the raw scores to determine the difference. Finally, the hierarchical linear model was applied to the examination of the impact of community integration on residents’ family happiness and the standardized and raw scores were compared as well.

Keywords: family happiness, subjective measurement, measurement anchor, standardization

一、研究问题

随着经济的不断发展和人们对生活质量的日益重视,幸福成为政府执政理念和人民生活追求的目标。幸福感是测量人们幸福与否及幸福程度的重要指标。在西方学术界,well-being、happiness 和 life satisfaction 基本同义,均可翻译成“幸福感”。幸福感一般指人们对当前生活满意度的主观评价。从 20 世纪 60 年代开始,国际社会关于幸福感的研究逐渐增多。尤其进入 21 世纪以来,幸福感日益成为社会学、心理学、经济学、政治学等领域的研究热点。

国外幸福感研究一般可分为幸福感的相关概念与理论、幸福感测量、幸福感的影响因素等几个主要方面。迪纳(Diener,1984)对幸福感的概念界定、内在维度及相关理论进行了总结性回顾。维特斯和尼尔森(Vittersø and Nilsen,2002)分析了主观幸福感的概念及内在结构,认为主观幸福感包括生活满意度、积极性情绪和消极性情绪三个维度。迪纳等(Diener and Suh,1997)比较了经济指标、社会指标和幸福感主观评价指标,认为三种指标在描述生活质量方面的作用不可相互替代,政策制订者要重视幸福感评价等主观指标在政策制订中所能发挥的作用。

幸福感测量方面,主要集中于特定情境下所用测量量表的编制,不同量表之间测量效度和信度的比较,不同层次幸福感指标的测量等。吕波密斯基和莱珀(Lyubomirsky and Lepper,1999)提出了一个包含 4 个问题的全球幸福感量表,通过对所收集的美国和俄罗斯的数据进行

评估,表明该量表具有很好的内在效度。利姆(Lim,2008)利用已有研究文献,比较了用4分位、5分位、7分位和11分位李克特量表在幸福感评分方面的差异,并通过已有数据证明11分位李克特量表测量的幸福感评分要显著高于用4分位和7分位量表所得评分,并提出了改进建议。波比特等人(Bobbitt, *et al.*, 2005)则构造了县级层次的幸福感指数。

幸福感的影响因素一般可分为个体因素和家庭、社区或国家等中观及宏观层次的影响因素。个体因素方面,一般集中于探讨性别、年龄、婚姻、职业、受教育程度、经济收入、健康状况、宗教信仰等与幸福感的关系,如威尔森(Wilson,1967)详细分析了主观幸福感的个体差异性;伊斯特林(Easterlin,1974,2001)对主观幸福感与经济收入的关系进行了研究。在中观及宏观方面,哈勒和哈德勒(Haller and Hadler, 2006)利用分层模型(hierarchical linear model)研究了社会经济结构对个人幸福感的影响,认为经济不平等状况、政治民主进度、社会关系密切程度等均会影响个人幸福感。杨扬(Yang,2008)利用分层的年龄—时期—队列方法(hierarchical age-period-cohort)研究了社会不平等对个人幸福感的影响,认为社会性别、种族和教育方面的不平等对个人幸福感会产生显著影响。

近年来,国内对幸福感的研究也逐渐增多,一般集中于探讨户籍制度、就业状况、经济收入等对个人主观幸福感的影响。罗楚亮(2006)、陈钊等(2012)研究了户籍状况对人们主观幸福感的影响。王鹏(2011)、刘斌等(2012)则分别研究了居民收入差距、社区平均幸福感与个人幸福感的关系。

目前,在幸福感研究中主要存在以下问题:一是与国外相比,国内对幸福感的研究刚处于起步阶段,主要集中于对幸福感影响因素的探讨,对幸福感测量的研究很少。二是研究对象大多局限于个人幸福感,对家庭幸福感的研究很少涉及。中国传统社会历来是“家本位”,家庭生活也是社会的一个缩影,与其他国家相比,测量家庭幸福感在中国具有更加重要和特殊的意义。三是数据收集方面,国内有全国代表性的概率样本数据还比较少,尤其缺乏专门针对家庭幸福感的全国样本。四是无论对个人幸福感还是家庭幸福感等主观评价指标的测量都有非常大的改进空间,由于设计的缺陷或缺乏科学性,很难得出可靠的测量结果。针对现有研究的这些不足,本文将从测量方法和实际操作两方

面对中国家庭幸福感的测量进行探讨。

二、测量方法

在幸福感的研究过程中,测量是否科学是研究结果是否可靠的前提和基础。一方面,测量是以幸福感基本概念的准确界定、量化以及样本的代表性为前提,只有做到对测量对象及其范围和测量内容的准确界定,才能保证测量的信度和效度;另一方面,测量也是正确进行幸福感影响因素分析的必要条件,只有有效的测量才能保证幸福感研究的相关结果不发生严重偏误。

从基础数据收集的角度看,目前对幸福感的测量一般采用抽样调查的方法。国际上比较有代表性的调查有美国综合社会调查(General Social Survey)、世界价值观调查(World Values Survey)、英国住户追踪调查(British Household Panel Survey)、人民之声(Voice of the People)、跨国社会调查[International Social Survey Program:Religion II (1998)]、世界大都市幸福感和竞争力调查(Assessing Happiness and Competitiveness of World Major Metropolises 2006)、亚洲调查(Asia Barometer 2004)、生活质量调查(Quality of Life Survey 2006)、CBS民意调查(CBS News/60 Minutes/Vanity Fair National Poll 2011)等。国内主要有全国城乡居民住户调查(Chinese Household Income Project 2002)、中国综合社会调查(Chinese General Social Survey 2006)以及复旦大学“和谐社区与社会资本”调查(2006—2007年)等。

从幸福感测量的具体指标来看,目前的主流做法是借鉴1972年美国综合社会调查的测量方式:“总体来看,你最近的幸福感程度是:1.非常幸福,2.比较幸福,3.不是很幸福”。后续相关调查大多以此为基础,发展出种类繁多的关于幸福感测量的各种量表。总的来说,对幸福感的测量可分为三种主要类型。

第一种方法是使用单个问题的3分类、4分类或5分类的李克特量表,将幸福感按幸福程度分为少数几个类别,如美国综合社会调查、CBS民意调查将幸福感评价分为非常幸福、比较幸福、不是很幸福三类;世界价值观调查将幸福感分为非常幸福、比较幸福、较不幸福、很不幸福四类;生活质量调查、全国城乡居民住户调查和中国综合社会调查则将幸福感评价分为非常幸福、比较幸福、不好也不坏、较不幸福、很不

幸福五大类。

第二种方法是使用单个问题的 7 分类或 11 分类李克特量表进行测量。如英国住户追踪调查通过询问“你对自己的生活满意度如何？1. 非常不满意，……，7. 非常满意”来测量个体幸福感。“和谐社区与社会资本”调查通过“我对自己的生活感到满意：1. 强烈不同意，……，7. 强烈同意”的 7 分类量表来测量幸福感；香港岭南年度幸福感调查（the Lingnan Annual Happiness Surveys 2011）则采用 11 分类量表对幸福感进行测量：“总的来说，采用 0—10 的 11 分类量表，10 代表最幸福，你有多幸福？”

与前两种方法通过一个问题来测量个人幸福感不同，第三种方法采用复杂量表来测量幸福感。如香港家庭调查（the Household Survey 2009）将国外的主观幸福感量表（Subjective Happiness Scale）改编成中文版本，采用“一般而言，我觉得自己：1. 不是一个很快乐的人，……，7. 是一个很快乐的人”、“与我大部分同辈相比，我：1. 不比他们更快乐，……，7. 比他们更快乐”、“有些人一直很快乐，无论发生什么事情，他们都热爱生活并从中得到满足。用这些话来形容你有多贴切？1. 很不贴切，……，7. 十分贴切”、“有些人经常不快乐，虽然他们并不抑郁。用这些话来形容你有多贴切？1. 很不贴切，……，7. 十分贴切”共 4 个问题，每个问题均采用 7 分类的李克特量表进行测量，通过对 4 个问题得分的汇总得到幸福感最终评分。

对幸福感进行测量的这三种方法均存在一定问题。第一种方法仅将幸福感分为诸如“非常幸福、比较幸福、一般、较不幸福、很不幸福”几大类，不仅不同类别之间的真正差异无法测量，而且在具体定量分析中也存在很大局限性。如罗楚亮（2006）基于 2002 年全国城乡居民住户调查数据，以个人幸福感主观评价为因变量，运用 Logit 模型来分析个人幸福感的影响因素。在对因变量进行赋值时，将幸福感分类中的“非常幸福”和“比较幸福”赋值为 1，而将“不好也不坏、不太幸福、很不幸福”赋值为 0。这样就人为强制设定了“不好也不坏”这一选项为“不幸福”类别，与“不好也不坏”选项的本意并不一致，因此采用这种编码方法并运用 Logit 模型所得结果的有效性值得怀疑。¹刘斌等（2012）运用

1. 虽然有文献（Praag *et al.*, 2004）表明，就统计推断而言，采用 Logit 或 Probit 模型与采用 OLS 模型所得结论等价，但在未经更多数据检验的情况下，该结论需要谨慎使用。

2006年中国综合社会调查数据,采用OLS回归分析研究社区平均幸福感对个人幸福感的影响。该研究将“非常不幸福、不幸福、一般、幸福、非常幸福”分别赋值为“1、2、3、4、5”,然后计算社区的平均幸福感和个人幸福感。暂且不论社区层次变量和个人层次变量运用OLS回归分析的恰当性和有效性,单就对个人幸福感的编码方式而言,就存在强制设定不同幸福感相邻类别差距为1的主观假定。由于这种假设与现实是否一致本身就存在很大疑问,从而造成该研究结果的有效性也同样值得怀疑。

第二种方法通过7分类或11分类的利克特量表对幸福感进行测量,一定程度上克服了第一种方法进行量化分析时的局限性,但也存在一些问题。如陈钊等(2012)将生活满意度近似于幸福感,并将“我对自己的生活感到满意:1. 强烈不同意,……,7. 强烈同意”分别赋值为1到7。运用这种赋值方法进行定量分析时,可以部分避免第一种方法存在的局限性,更适合于使用统计模型进行分析。但是,这种方法与第一种方法一样,也容易受到自评异质性的影响,即由于每个人打分时的评判原则或标准不一样,从而使得人们给出的幸福感评分之间的可比性会存在一定问题。以香港岭南年度幸福感调查采用的11分类量表为例,假设甲乙两人同样给各自的幸福感评了7分,但在甲的幸福感受价体系中认为6分(含)以上就称得上幸福,而乙则认为8分(含)以上才是幸福的。因此,由于甲乙两人对个人幸福感评价的基准不一致,同样的7分,甲认为自己是幸福的,而乙则认为自己不幸福。因此,如果不对自评异质性问题进行统计控制,直接运用评分结果进行统计分析会存在一定问题。²

第三种方法,即采用复杂量表对幸福感进行测量也存在诸多问题:一是幸福感测量的复杂量表大多专门针对某一特定情境(如特定人群、特定地区、特定期限)而编制,由于对特定情境的强依赖性使得其外在效度和可推广性不高。虽然已有的复杂量表(如主观幸福感量表)的内部一致性容易评估,但其外在有效性却无法真正测量。二是虽然尚未发现运用一个简单问题对幸福感进行测量(第一种方法和第二种方法)

2. 虽然采用多元回归,通过控制个体特征会在某种程度上缓解自评异质性问题,但统计控制本身作用比较有限,做不到对自评异质性的较大改善。

与运用复杂量表进行测量两种方式之间存在显著差异,但从简约性原则和成本控制角度考虑,没有证据让我们舍弃第一种、第二种这种单问题测量方式而选择复杂量表进行测量。三是即便是复杂量表,如主观幸福感量表所包含的四个问题也均采用7分类李克特量表,其本身同样存在前两种方法中的自评异质性问题。

比较这三种方法,我们发现:第一种方法由于将幸福感划分为有限几个分类,因此定量分析时会出现较大问题;第二种方法和第三种方法与第一种方法一样,都存在自评异质性问题,但第三种方法过于复杂,且其外部有效性和可推广性也值得怀疑。因此,如果能找到克服自评异质性的方法,第二种测量方法无疑是一个比较理想的选择。

其实,不仅在幸福感测量中存在自评异质性问题,所有主观概念的测量,如个人健康状况、工作满意度等均存在自评异质性问题。亚历山大(Alexander, 1978)很早就探讨了调查研究中情境选择测量(Vignettes)的应用问题。芬奇(Finch, 1987)比较了美英两国在关于信念、规范等方面的情境化测量中存在的差异。21世纪以来,更是涌现出一批运用情境化测量的研究成果,如亨特(Hunter, 2009)探讨了情境化测量方法如何运用于研究患者对使用电子医疗记录的态度;金(King, 2004)、赖斯(Rice, 2010, 2011)、格罗尔—普罗科普兹克(Groß-Prokopczyk, 2011)使用引入“锚点”的情境测量法(Anchoring Vignettes)来评估不同群体间自我健康评价的差异;卡普坦(Kapteyn, 2007)、詹宁斯(Jennings, 2001)则分别探讨了工作称职状况等劳动经济领域情境选择测量方法的应用问题。此外,斯伯丁(Spalding, 2007)、帕卡涅拉(Paccagnella, 2011)还专门讨论了情境选择测量法的信度和效度、样本选择性问题。

通过以上文献分析可知,虽然情境选择测量法比较复杂,且在不同研究中的具体操作化方法也不一样,但其核心思想却比较一致,即将所有人的比较基准统一调整到同一起点。以自我健康评价为例,由于不同群体、不同个人对健康的定义、判断标准和期望水平等可能存在显著差异,相同的自我健康状况评分很可能代表不同的健康状况。已有研究(Idler, 1997; Benjamini, 1999)表明,不同性别、种族、年龄的群体关于健康的标准存在很大差异。国际上一般使用引入“锚点”的情境选择测量法来克服自我健康评价中存在的自评异质性问题,如坦顿

(Tandon, 2002)提出了健康测量中引入“锚点”的情境选择测量法,即首先请受访者就自己的健康状况做出主观评价(采用5分类李克特量表),然后再请受访者使用同样的5分类李克特量表对一组可以清楚排序的情境进行评价,通过受访者对这组情境的评价来修正自我健康评价中存在的自评异质性问题。具体修正方法是对各个情境进行排序并赋分,使用5分类李克特量表一般会设置5个标准化情境,并按照健康状况由好到坏分别赋值为4、3、2、1、0分。假设受访者对自我健康状况评价为3分,如果受访者自我健康评分与其对场景2的评分相同,而场景2的实际赋值分数为2分,则受访者对自我健康状况的评价得分便会被调整为2分,这样就在一定程度上克服了不同受访者因为对健康的定义和判断标准不同所造成的自评异质性问题。

与自我健康评价相比,家庭幸福感的测量更加主观。自我健康状况的自评异质性可以通过引入能够客观评价的标准化情境来加以克服,而家庭幸福感的标准化场景却不容易建立。鉴于目前国内外均没有家庭幸福感测量及幸福感自评异质性克服方法的相关研究,因此,本文主要借鉴健康测量中克服自评异质性的主要思路,使用问题1(11分类的李克特量表)来测量受访者家庭幸福感的自我评价,并采用问题2来校正问题1所得评分。³两个问题均采用相同的测量原点,这样可起到测量标准化的作用,使测量对象的主观感受可量化,且不同对象的测量结果具有可比性。

问题1:请给您目前的家庭幸福感评分(最高为10分,最低为0分);

问题2:在对幸福评分中您认为几分以上是幸福的(最高为10分,最低为0分)?

问题1要求受访者对自己的家庭幸福感进行自我评分,最高得分为10分,最低得分为0分,中间分数为5分。需要指出的是,以往许多

3. 本文数据均来源于“六普”及“2012年中国家庭幸福感热点问题调查”,后者主要包括被调查者个人情况、家庭基本情况、家庭和邻里关系、对家庭幸福感的评价、配偶及两性关系、子女情况6个部分。通过数据分析,我们发现,经过问题2调整后的问题1所得评分与个人和家庭特征、配偶和两性关系、父母和子女关系、邻里关系等各个分项的得分高度相关。即使我们使用比较复杂的量表,如包括上述各种特征和关系的加权得分,所得结果与使用单一问题仍基本一致。因此,从简约的角度考虑,本文主要使用问题1来测量家庭幸福感,使用问题2来克服家庭幸福感测量中的自评异质性问题。

研究(罗楚亮,2006;刘斌等,2012)均将李克特量表的中间值作为幸福的临界点,即高于中间值定义为幸福,低于中间值定义为不幸福,本文认为不同群体的幸福临界点很可能存在显著差异,从而将量表中间值作为幸福与否的临界点也同样可能存在严重问题。如甲乙两人均给各自家庭幸福感评了7分,但甲认为5分以上就称得上家庭幸福,乙则认为8分以上的家庭才是幸福的。因此,由于甲乙两人对家庭幸福感基准认知的不一致,同样的7分,甲认为自己家庭幸福,乙则认为自己家庭不幸福,从而说明家庭幸福感知基准的显著差异很可能会造成家庭幸福感测量中的自评异质性问题。

“2012年中国家庭幸福感热点问题调查”通过引入问题2试图克服测量中由于比较基准的不一致而造成的自评异质性问题。问题2的主要目的就是要测量每个人家庭幸福感的比较基准(幸福临界点),高于这个比较基准则该受访者被认为家庭幸福,本文通过问题2测量出的比较基准来校正问题1所得家庭幸福感评分。

具体操作方面,本研究采用标准化方法来解决家庭幸福感测量中遇到的自评异质性问题。标准化方法主要有等比例标准化方法和等距离标准化方法,两者虽然在具体操作上不一样,但其核心思想是一样的,即将“在对幸福评分中您认为几分以上才是幸福的”中的得分(以下简称评分基准或幸福临界点)统一调整到6分,⁴从而使得不同个体的家庭幸福感评分之间具有可比性。

等比例标准化方法假定调整后的得分与评分基准(本文设定为6分)之比等于原始分数与幸福及格线之比。具体操作方面,以表1中的案例2和案例3为例说明。案例2和案例3的家庭幸福感评分分别为7.65分和6.66分,前者高于后者,但其幸福临界点分别为8分和6分。我们将每个案例的幸福基准统一固定在6.00分,调整系数定义为6分除以幸福临界点的所得结果,案例2和案例3的调整系数分别为0.75(即 $6/8$)和1(即 $6/6$)。这样,案例2和案例3的家庭幸福感最终得分分别为5.74分和6.66分。虽然案例2的家庭幸福感最初评分高于案例3,但由于两者幸福临界点的差异,导致其最终得分低于案例3,

4. 幸福基准本身是人为确定的,主要是为了有一个统一的调整参照或尺度,可以为6分,也可以为任意一个其他分数。本文将幸福基准设定为6分,主要是考虑到符合常规做法。

从中也可看到自评异质性对测量结果可能造成的扭曲作用。

表 1: 家庭幸福感等比例标准化方法

案例	幸福感评分	幸福临界点	调整基准	调整系数	最终得分
1	7.98	7.00	6.00	0.86	6.84
2	7.65	8.00	6.00	0.75	5.74
3	6.66	6.00	6.00	1.00	6.66
4	8.24	7.00	6.00	0.86	7.06
5	8.01	5.00	6.00	1.20	9.61
6	6.90	7.00	6.00	0.86	5.91

等距离标准化方法假定调整后的得分与评分基准(本文设定为 6 分)之差等于原始分数与幸福临界点之差。同样以表 1 中的案例 2 和案例 3 为例,案例 2 和案例 3 与幸福临界点的距离分别为 -0.35 分(即 $7.65 - 8.00$)和 0.66 分($6.66 - 6.00$)。将每个案例的幸福基准也统一固定在 6.00 分,则案例 2 和案例 3 的家庭幸福感最终得分分别为 5.65 分($6.00 - 0.35$)和 6.66 分($6.00 + 0.66$)。同样,案例 2 的家庭幸福感最终得分低于案例 3。

鉴于家庭幸福感等比例标准化方法与等距离标准化方法的最终调整结果相差不大,本研究采用等比例方法对家庭幸福感评分进行标准化处理。

三、数据收集及质量评估

本文数据主要来源于“六普”及“中国家庭幸福感热点问题调查”。“中国家庭幸福感热点问题调查”2011 年主要采用网络调查,2012 年主要采用实地抽样调查,同时进行了少量网络调查。不同的调查方式为比较数据收集方法的有效性和代表性提供了便利。

(一) 数据收集

1. 网络调查

由于互联网本身所具有的低成本和高效率等优势,网络调查在我国日益普及。与实地调查相比,网络调查一般面向自愿参加调查的上网群体,并通过后台数据库和逻辑结构对网络答题进行实时控制,调查所得结果也能第一时间传回数据库,因而比实地调查更加方便和快捷。

2011 年“中国家庭幸福感热点问题调查”主要采用网络调查方式,共获得网络样本 12 683 份。2012 年“中国家庭幸福感热点问题调查”

在进行实地调查的同时,也进行了网络调查,共获得网络样本 3 424 份。

2. 实地抽样调查

为获得对全国有代表性的概率样本,2012 年“中国家庭幸福感热点问题调查”主要采用实地抽样调查的方式收集数据。首先采用分层概率比例抽样方法,抽取了北京、浙江、安徽、甘肃、河南、四川、辽宁七个省份(直辖市)作为样本点。接着根据全国各县(市、区)最新的经济、社会和人口等发展状况,同样采用分层概率比例抽样方法选取抽中省份中接受调查的县(市、区),每省抽取 3 个县(市、区)。随后,采用系统抽样法抽取样本县(市、区)中的社区和住户,每个县抽取 6 个社区,每个社区抽取 50 个住户。最后,在每个抽中的住户中,通过入户抽样表随机抽取一名家庭成员作为接受调查者。⁵2012 年实地调查设计总样本量为 6 000 份,其中,北京地区样本量为 600 份,其他 6 个省份每省样本量为 900 份。本次调查共获得 18 周岁及以上成年人有效样本 5 547 个,有效样本为 92.45%。

(二) 数据质量评估

与实地调查相比,网络调查存在诸多缺陷。首先,网络样本由于网络群体的特殊性而可能存在一定的选择性偏差,很难有代表性;其次,网络调查如果缺乏必要的物质激励,应答率常常会比较低;另外,网络调查中被调查者身份和回答内容的真实性很难保证。2011 年网络调查由于有回答问题即可参与抽奖的物质激励,出现了小规模的重重复答人群。而 2012 年网络调查由于没有相应的物质激励,应答率很低。2012 年网络样本共有 3 424 人,而包含家庭幸福感评分的有效样本仅为 747 人,应答率仅 21.82%。下面以“六普”数据为评价标准,对网络调查和实地抽样调查的偏差进行分析。

年龄结构方面,2012 年实地调查数据对全国具有较好的代表性,而 2012 年网络调查样本中城镇人口、70 后和 80 后群体(70 后指 1970 年至 1979 年期间出生,80 后指 1980 年至 1989 年期间出生,以此类推)、男性、离婚人口、大学专科及以上群体、体制内职业人口占比明显偏多。

城乡结构方面,实地调查的农村样本量略高于城镇,而网络调查的

5. 本文所说的家庭幸福感,即是指被抽取家庭成员关于自己所在家庭的成员关系、日常生活、家庭经济状况、邻里关系的综合评价。

城镇样本量远大于农村,城镇占比高达 83.44%。从年龄分布来看,实地调查与“六普”分年龄段人口除了 90 后较少外,其他年龄段人口都比较接近(见表 2)。网络调查中 80 后和 70 后占比明显过高,反映了我国网民以 80 后和 70 后为主体的现状。

表 2:样本年龄分布情况(%)

出生年代	2010 年人口普查	实地调查	网络调查
40 后	10.05	13.23	0.56
50 后	20.06	18.32	4.63
60 后	20.45	25.49	16.61
70 后	19.85	21.91	31.75
80 后	14.60	16.40	40.34
90 后	14.99	4.64	6.10

婚姻状况方面,实地调查的离婚人口占比、丧偶人口占比与“六普”很接近。由于本次调查的成人问卷只调查 18 周岁及以上人口,因此,实地调查的未婚人口占比明显低于“六普”数据。网络调查除了未婚人口占比与“六普”比较接近外,离婚人口占比明显偏高,丧偶人口占比明显偏低。

从受教育程度来看,实地调查中各受教育程度人群占比与“六普”非常接近,而网络调查中大学专科及以上人群占比明显过高(见表 3)。“六普”中 16 周岁及以上人口的大学本科占比为 4.20%,实地调查为 5.50%,网络调查则为 39.89%。同样,“六普”中 16 周岁及以上人口的大学专科、研究生占比分别为 6.30%、0.38%,实地调查分别为 7.52%、0.29%,网络调查分别为 29.18%、9.62%,明显偏高。文盲、小学、初中在网络调查中的占比则明显过低,分别为 0%、0.55%和 4.04%,而实地调查分别为 5.53%、19.32%和 40.48%,这与“六普”结果(分别为 5.09%、23.88%和 43.54%)非常接近。

表 3:样本受教育程度分布情况(%)

受教育程度	2010 年人口普查	实地调查	网络调查
未上过学	5.09	5.53	0.00
小学	23.88	19.32	0.55
初中	43.54	40.48	4.04
中专或高中	16.61	21.37	16.72
大学专科	6.30	7.52	29.18
大学本科	4.20	5.50	39.89
硕士及以上	0.38	0.29	9.62

职业类别方面,与实地调查相比,网络调查中农业劳动者占比过低。实地调查中农业劳动者占比 41.67%,与“六普”16 周岁及以上的农村地区人口比例比较接近(48.20%),而网络调查中农业劳动者占比仅为 0.33%,明显过低。网络调查中的事业单位、国企、私企、公务员、学生占比要高于实地调查,而个体工商户、离休和退休人员、下岗失业人员等占比则低于实地调查。

性别分布方面,实地调查与“六普”数据差异较大。“六普”16 岁及以上人口男性占比 50.57%,女性占比 49.43%,男性略高于女性。实地调查女性在占比方面稍高于男性,网络调查男性占比高达 60.02%,严重偏离“六普”数据。

总体来说,2012 年实地调查数据对全国具有较好的代表性,应答率、数据质量均较高,因此,本研究主要使用 2012 年实地调查所得数据。数据中有些偏差可能会严重影响测量结果,有些可能影响不大,例如,如果将全国、城镇和农村分别按照男女性别各占 50%(接近“六普”数据的性别分布)进行加权处理,则全国人口、城镇人口和农村人口的家庭幸福感变为 6.90 分、6.83 分和 6.94 分,与现在的 6.90 分、6.84 分和 6.95 分相比,全国人口的家庭幸福感基本没有变化,城镇和农村的家庭幸福感均降低 0.01 分,差异非常细微。鉴于对性别加权与否对最终结果基本没有影响,因此本研究不对性别进行加权处理。

四、标准化测量的结果及意义

(一) 自评异质性的影响因素

自评异质性方面,全国家庭“几分以上是幸福的”平均分为 7.27 分,显著高于通常认为的及格线(6 分)。其中,农村家庭平均分为 7.13 分,要低于城镇家庭的 7.47 分(见表 4)。t 检验结果表明,城乡之间这种差异统计性显著(显著度为 0.000)。因此,城乡居民家庭之间存在自评异质性问题,即城乡居民对家庭幸福感的评价基准不一致。

表 4:分城乡的自评异质性

分城乡	几分以上是幸福的	人数
农村人口	7.13	3 250
城镇人口	7.47	2 297
全国人口	7.27	5 547

除了城乡外还存在分年龄的自评异质性问题。如表 5 所示,40 后“几分以上是幸福的”平均分为 7.07 分,50 后为 7.18 分,60 后为 7.29 分,70 后为 7.32 分,80 后为 7.35 分,90 后为 7.56 分,各年龄段之间均存在差异。方差分析表明,不同年龄组之间在“几分以上是幸福的”平均分上至少在两组之间存在显著差异。

表 5:分年龄的自评异质性

出生年代	几分以上是幸福的	人数
40 后	7.07	735
50 后	7.18	1 020
60 后	7.29	1 420
70 后	7.32	1 223
80 后	7.35	912
90 后	7.56	256
合计	7.27	5 566

运用 OLS 模型,可以分析自评异质性的影响因素。因变量为被调查者“几分以上是幸福的”所给出的分数,自变量包括被调查者的户籍(农村赋值为 1,城镇赋值为 0)、出生年代(以 90 后为参照,分为 40 后、50 后、60 后、70 后和 80 后)、性别(女性赋值为 1,男性赋值为 0)、受教育程度(以小学及以下为参照,分为初中、高中、大专及以上)、婚姻状况(以未婚为参照,分为初婚、再婚、离异、丧偶)、子女结构(以男女双全户⁶为参照,分为纯男户和纯女户)、家庭收入状况(以低收入家庭为参照,分为中等收入户、高收入户)和住房状况(以住房状况一般家庭为参照,分为住房状况良好家庭、住房状况较差家庭)等。

最终结果如表 6 所示。在控制其他因素的情况下,农村“几分以上是幸福的”评分要显著低于城镇,农村居民的评分基准要比城镇居民低 0.385 分。年龄方面,40 后、50 后、60 后、70 后和 80 后的“几分以上是幸福的”评分要显著低于 90 后,比 90 后分别低 0.521 分、0.414 分、0.297 分、0.286 分和 0.223 分。另外,住房状况较差家庭“几分以上是幸福的”评分也要比住房状况一般家庭低 0.332 分。

6. 男女双全户,指子女中既有男孩又有女孩的家庭,纯女户指子女中只有女孩的家庭,纯男户指子女中只有男孩的家庭。

表 6: 自评异质性的影响因素

变量	系数	标准误	标准化系数	t 值	显著度
农村	-0.385	0.050	-0.121	-7.743	0.000
生于 1940—1949 年	-0.521	0.140	-0.112	-3.728	0.000
生于 1950—1959 年	-0.414	0.132	-0.102	-3.130	0.002
生于 1960—1969 年	-0.297	0.128	-0.082	-2.313	0.021
生于 1970—1979 年	-0.286	0.128	-0.075	-2.230	0.026
生于 1980—1989 年	-0.223	0.122	-0.052	-1.818	0.069
女性	0.014	0.043	0.004	0.323	0.746
初中	0.086	0.058	0.027	1.484	0.138
高中	0.013	0.070	0.003	0.187	0.852
大专及以上	-0.119	0.088	-0.026	-1.349	0.177
初婚	0.003	0.092	0.001	0.038	0.970
再婚	0.155	0.160	0.015	0.971	0.332
离异	-0.089	0.202	-0.007	-0.440	0.660
丧偶	-0.047	0.136	-0.007	-0.346	0.729
纯女户	-0.053	0.072	-0.010	-0.743	0.457
纯男户	-0.104	0.062	-0.023	-1.675	0.094
中等收入家庭	-0.021	0.067	-0.005	-0.309	0.758
高收入家庭	0.094	0.186	0.007	0.507	0.612
住房状况良好家庭	0.081	0.044	0.025	1.843	0.065
住房状况较差家庭	-0.332	0.088	-0.053	-3.781	0.000
截距	7.803	0.138		56.628	0.000

注: $R^2=0.023$ 。

(二) 测量结果

城乡家庭幸福感评分结果如表 7 所示。标准化后的全国家庭幸福感评分为 6.91 分,显著低于未经标准化的 8.16 分。以城乡家庭幸福感状况差异为例可以进一步说明标准化对家庭幸福感的影响。未经标准化的农村家庭幸福感得分为 8.06 分,城镇为 8.30 分,且统计检验显著(t 检验显著度为 0.000),但由于农村家庭“几分以上是幸福的”平均

分(7.13分)低于城镇家庭(7.47分),因此,标准化后的农村家庭幸福感得分为6.96分,高于城镇家庭的6.84分。t检验结果表明,标准化后的农村家庭幸福感高于城镇,且统计检验显著(显著度为0.009)。

表 7:城乡评分情况

分城乡	标准化得分	未经标准化的得分
农村	6.96	8.06
城镇	6.84	8.30
全国	6.91	8.16

从家庭幸福感评分分布来看,标准化后的家庭幸福感评分集中于6分到9分的区间范围,其中6分到7分(不含)占比41.33%,7分到9分(不含)占比26.97%,二者合计68.3%。未经标准化的家庭幸福感得分集中于7分到10分的区间范围,其中7分到9分(不含)占比42.57%,9分到10分占比43.98%,二者合计86.55%。标准化和未经标准化的家庭幸福感评分,无论是总体均值还是具体分值分布,差异都非常明显(参见图1)。

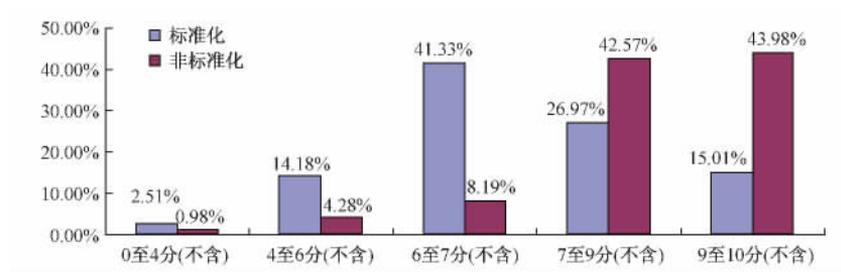


图 1:评分分布情况

(三) 一个研究示例:社区融合状况对家庭幸福感的影响

下文将以社区融合状况对家庭幸福感的影响为例来检验家庭幸福感标准化与否对研究结论的影响。国内外学者在社区融合对居民影响方面做了大量研究(Axinn, et al., 1996, 2001; 崔岩, 2012)。目前国内社区融合研究一般局限于理论和政策层面的探讨,定量研究很少,其中一个重要原因就是无法获得社区层面的定量数据,从而无法获得社区融合状况的定量化指标。2012年“中国家庭幸福感热点问题调查”为

研究社区融合与家庭幸福感的关系提供了可能。⁷

鉴于社区层面数据和个人层面数据具有明显的分层结构,本研究利用分层模型来研究社区融合状况对居民家庭幸福感的影响。本研究的因变量为家庭幸福感评分,分别就标准化的家庭幸福感评分和未经标准化的家庭幸福感评分做两个模型,以便进行对比分析。自变量为社区层次的社区融合指标。限于数据可得性,本研究的社区融合指标主要包括经济融合、社会融合、融合总体状况评价三部分。经济融合指标主要包括社区平均家庭年收入和社区家庭收入不平等两个变量。其中社区平均家庭收入不平等用社区居民家庭年收入的方差⁸表示。社会融合指标主要包括社区邻里互动指数和邻里关系评价两个变量。社区邻里互动指数是用每个社区中被调查居民在“与邻居见面互打招呼”、“帮邻居家或请邻居帮自家修理家具、设备等”、“受邀到邻居家做客或者邀请邻居到自己家做客”、“向邻居借钱或借给邻居钱”、“邻居不在家时,为邻居看门”共5个问题中每题有过相关行为比例的均值表示。对融合总体状况的评价本文使用每个社区中的被调查者对所在社区的总体评价的均值来近似表示。

控制变量为个人或家庭层次⁹的特征变量,包括家庭收入、住房状况、子女结构、被调查者年龄、性别、户籍、受教育程度、婚姻状况等(参见表8)。

表9为分层模型的最终结果。在标准化的家庭幸福感作为因变量的模型结果中,社区平均家庭年收入、社区家庭收入不平等状况、社区邻里互动指数、邻里关系评价、所在社区整体评价均取得了显著性统计结果。具体而言,经济融合方面,在控制层一变量的情况下,社区平均家庭年收入每提高1万元,居民家庭幸福感提高0.1493分;社区家庭收入不平等状况每提高1个单位,居民家庭幸福感将降低0.0120分。

7. 2012年“中国家庭幸福热点问题调查”问卷包括了家庭居民去年家庭总收入、邻里关系评价、与邻里日常互动情况、对居住社区的总体评价等内容。我们可以通过计算分社区的家庭总收入均值获得社区平均家庭收入,通过邻里关系评价及与邻居互动情况获得社区的总体邻里关系情况,通过每个家庭对所在社区的总体评价获得所在社区的总体情况。

8. 社区居民家庭年收入的方差表示社区居民与社区平均家庭年收入的差异性,可以在某种程度上代表收入不平等状况。

9. 根据本研究的抽样方法,每个家庭只抽取1人调查其家庭幸福感程度,因此从分层模型的数据层次定义规则看,家庭状况变量与个人变量同处于微观层次。

表 8: 分层模型所用变量信息

变量名称	均值(比例)	标准差	最小值	最大值
因变量				
未经标准化的家庭幸福感评分	8.15	1.51	0	10
标准化后的家庭幸福感评分	6.92	1.64	0	10
层二自变量(120 个案例)				
社区平均家庭年收入(万元)	4.59	2.71	1.80	17.14
社区邻里互动指数	0.47	0.12	0.24	0.81
邻里关系评价	8.66	0.69	6.85	9.91
所在社区整体评价	7.56	0.93	4.56	9.48
社区家庭收入不平等	11.78	19.56	0.47	124.12
层一控制变量(5296 个案例)				
中等收入家庭	0.87	0.34	0	1
高收入家庭	0.02	0.12	0	1
住房状况良好家庭	0.43	0.49	0	1
住房状况较差家庭	0.07	0.25	0	1

注: 限于篇幅, 本表只列出因变量、层二自变量以及部分层一控制变量的信息。

表 9: 分层模型最终结果

效应估计	家庭幸福感标准化的模型			家庭幸福感未经标准化的模型		
	系数	标准误	显著度	系数	标准误	显著度
层一截距: B0						
层二截距: G00	3.0982	0.603	0.000	1.1375	0.351	0.002
社区平均家庭年收入: G01	0.1493	0.041	0.001	0.0942	0.031	0.003
社区家庭收入不平等: G02	-0.0120	0.004	0.004	-0.0079	0.005	0.114
社区邻里互动指数: G03	1.4007	0.548	0.012	-0.1658	0.308	0.590
邻里关系评价: G04	0.1746	0.075	0.022	0.6142	0.048	0.000
所在社区整体评价: G05	0.1464	0.070	0.037	0.1908	0.039	0.000

注: 限于篇幅, 本表只列出层二自变量的最终结果。层二变量的标准化模型 $R^2 = 0.0445$, 估计可靠性 = 0.865, 非标准化模型 $R^2 = 0.1488$, 估计可靠性 = 0.592。

社会融合方面, 在控制层一变量的情况下, 社区邻里互动指数每提高 1 个单位, 居民家庭幸福感将提高 1.4007 分; 邻里关系评价每提高 1 分, 居民家庭幸福感将提高 0.1746 分。融合总体状况评价方面, 居民对所在社区整体评价每提高 1 分, 家庭幸福感将提高 0.1464 分。

在未经标准化的家庭幸福感作为因变量的模型结果中, 社区平均家庭收入、邻里关系评价、所在社区整体评价三个变量也取得了显著性结果, 但影响系数与采用标准化家庭幸福感作为因变量的模型有很大差异。社区平均家庭年收入的影响系数(0.0942)要小于标准化的家庭

幸福感作为因变量的模型(0.1493),而邻里关系评价和所在社区整体评价的影响系数(0.6142、0.1908)则要高于标准化的家庭幸福感作为因变量的模型(0.1746、0.1464)。

值得注意的是,未经标准化的家庭幸福感作为因变量的模型中,社区家庭收入不平等和社区邻里互动指数均未取得统计显著性结果,且邻里互动指数的系数方向(-0.1658)与标准化的家庭幸福感作为因变量的模型相反(1.4007),即未经标准化的家庭幸福感作为因变量的模型中,邻里互动指数每提高1个单位,居民家庭幸福感将降低0.1658分,这与实际经验明显不符。

五、结语

主观幸福感、健康状况、工作满意度等自我评价指标测量中普遍存在自评异质性问题,即每个人打分时使用不同的评判原则或标准,因此导致人们给出的主观评分不具有可比性。如果不消除自评异质性问题,将会影响测量结果的有效性和准确性。鉴于国内外幸福感等主观感受测量中还未考虑到自评异质性问题,本研究借鉴国外健康和满意度研究的最新成果,在测量中首次引入“测量锚点”,并采用“等比例方法”对家庭幸福感评分进行了标准化处理,以此达到控制自评异质性的目的。

与国内外许多关于幸福感的研究不同,本研究非常重视样本对全国的代表性,采用分层概率比例抽样、系统抽样、入户抽样等方法依次选取了接受调查的省份、县(市、区)、社区、住户及住户中接受调查者。数据质量评估结果表明,本研究所用实地调查数据对全国具有较好的代表性。而网络调查数据则明显有偏,城镇人口、70后和80后群体、男性、离婚人口、大学专科及以上群体、体制内职业人口占比明显偏多。

利用2012年中国家庭幸福感测量结果来评估家庭幸福感自评异质性的影响,结果表明,通过使用OLS模型,在控制其他因素的情况下,农村“几分以上是幸福的”评分要显著低于城镇。年龄方面,40后、50后、60后和70后的“几分以上是幸福的”评分要显著低于90后。因此,家庭幸福感测量中的城乡结构和年龄结构都存在自评异质性问题,这也说明了引入“测量锚点”并采用“等比例标准化方法”来控制自评异质性的必要性。

测量结果方面,标准化与未经标准化的家庭幸福感评分存在显著差异。标准化后的家庭幸福感评分为 6.91 分,显著低于未经标准化的 8.16 分。从评分分布来看,标准化后的家庭幸福感评分集中于 6 分到 9 分的区间范围,而未经标准化的家庭幸福感得分集中于 7 分到 10 分的区间范围。

本研究还以社区融合状况对居民家庭幸福感的影响为例,利用分层模型比较了家庭幸福感标准化处理与未经标准化处理在研究结论方面存在的显著差异。标准化后的家庭幸福感作为因变量的模型中,社区平均家庭年收入、社区家庭收入不平等状况、社区邻里互动指数、邻里关系评价、所在社区整体评价均取得了显著性统计结果。而未经标准化的家庭幸福感作为因变量的模型中,社区家庭收入不平等和社区邻里互动指数均未取得显著性统计结果,且邻里互动指数的系数方向与标准化的家庭幸福感作为因变量的模型正相反。这也进一步说明了家庭幸福感测量中引入“测量锚点”并进行标准化处理的必要性。

由于国内外还没有关于家庭幸福感的相关研究,因此本文的一些结论还需要后续研究进行验证。在后续研究中,需要进一步改进控制自评异质性的方法,尝试制订家庭幸福感的标准化场景和相应赋分方法,从而更好地克服家庭幸福感测量中的自评异质性问题。同时,通过引入“复合测量”¹⁰(multiple measurements)等技术手段,也可以对自评异质性问题进行更好的控制。此外,由于社区层次变量的有限性,本文探讨的社区融合状况对家庭幸福感的影响也存在一定局限性,需要在后续调查中不断补充和完善。

参考文献 (References)

- 陈钊、徐彤、刘晓峰. 2012. 户籍身份、示范效应与居民幸福感:来自上海和深圳社区的证据[J]. 世界经济(4): 79—101.
- 崔岩. 2012. 流动人口心理层面的社会融入和身份认同问题研究[J]. 社会学研究(5): 141—160.
- 刘斌、李磊、莫骄. 2012. 幸福感是否会传染[J]. 世界经济(6): 132—154.
- 罗楚亮. 2006. 城乡分割、就业状况与主观幸福感差异[J]. 经济学(季刊)(2): 817—840.

10. 复合测量是对自我健康评价、幸福感等主观变量进行多次独立测量的方法。如在家庭幸福感测量中,同时使用 7 分类李克特量表和 11 分类李克特量表进行测量。

- 王鹏. 2011. 收入差距对中国居民主观幸福感的影响分析——基于中国综合社会调查数据的实证研究[J]. 中国人口科学(3): 93—101.
- Alexander, Cheryl S. and Henry Jay Becker. 1978. "The Use of Vignettes in Survey Research." *Public Opinion Quarterly* 42(1): 93—104.
- Axinn, William G. and Tom Fricke. 1996. "Community Context, Women's Natal Kin Ties, and Demand for Children: Macro-Micro Linkages in Social Demography." *Rural Sociology* 61 (2): 249—271.
- Axinn, William G. and Scott T. Yabiku. 2001. "Social Change, the Social Organization of Families, and Fertility Limitation." *American Journal of Sociology* 106 (5): 1219—1261.
- Benyamini, Yael and Ellen L. Idler. 1999. "Community Studies Reporting Association between Self-Rated Health and Mortality: Additional Studies, 1995 to 1998." *Research on Aging* 21(3): 392—401.
- Bobbitt, Linda, Sheridan Green, Leanne Candura, and George A. Morgan. 2005. "The Development of a County Level Index of Well-Being." *Social Indicators Research* 73 (1): 19—42.
- Diener, Ed. 1984. "Subjective Well-Being." *Psychological Bulletin* 95(3): 542—575.
- Diener, Ed and Eunkook Suh. 1997. "Measuring Quality of Life: Economic, Social, and Subjective Indicators." *Social Indicators Research* 40(1/2): 189—216.
- Easterlin, Richard A. 1974. "Does Economic Growth Improve the Human Lot? Some Empirical Evidence." In *Nations and Households in Economic Growth: Essays in Honour of Moses Abramovitz*, edited by Paul A. David and Melvin W. Reder. N. Y.: Academic Press: 89—125.
- Easterlin, Richard A. 2001. "Income and Happiness: Towards a Unified Theory." *The Economic Journal* 111(473): 465—484.
- Finch, Janet. 1987. "The Vignette Technique in Survey Research." *Sociology* 21(1): 105—114.
- Grol-Prokopczyk, Hanna, Jeremy Freese, and Robert M. Hauser. 2011. "Using Anchoring Vignettes to Assess Group Differences in General Self-Rated Health." *Journal of Health and Social Behavior* 52(2): 246—261.
- Haller, Max and Markus Hadler. 2006. "How Social Relations and Structures Can Produce Happiness and Unhappiness: An International Comparative Analysis." *Social Indicators Research* 75(2): 169—216.
- Hunter, Inga M., Richard J. Whiddett, Anthony C. Norris, Barry W. McDonald, and John A. Waldon. 2009. "New Zealanders' Attitudes towards Access to Their Electronic Health Records: Preliminary Results from a National Study Using Vignettes." *Health Informatics Journal* 15(3): 212—228.
- Idler, Ellen L. and Yael Benyamini. 1997. "Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies." *Journal of Health and Social Behavior* 38 (1): 21—37.
- Jennings, Ann. 2001. "Social Constructions of Measurement: Three Vignettes from Recent Events and Labor Economics." *Journal of Economic Issues* 35(2): 365—371.
- Kapteyn, Arie, James P. Smith, and Arthur van Soest. 2007. "Vignettes and Self-Reports of Work Disability in the United States and the Netherlands." *The American Economic Review* 97(1): 461—473.
- King, Gary, Christopher J. L. Murray, Joshua A. Salomon, and Ajay Tandon. 2004. "Enhancing the Validity and Cross-Cultural Comparability of Measurement in Survey Research." *American Political Science Review* 98(1): 191—207.

- Lim, Hock-Eam. 2008. "The Use of Different Happiness Rating Scales: Bias and Comparison Problem?" *Social Indicators Research* 87(2): 259–267.
- Lyubomirsky, Sonja and Heidi S. Lepper. 1999. "A Measure of Subjective Happiness: Preliminary Reliability and Construct Validation." *Social Indicators Research* 46(2): 137–155.
- Paccagnella, Omar. 2011. "Anchoring Vignettes with Sample Selection Due to Non-Response." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 174(3): 665–687.
- Praag, Bernard van and Ada Ferrer-i-Carbonell. 2004. *Happiness Quantified: A Satisfaction Calculus Approach*. N. Y.: Oxford University Press.
- Rice, Nigel, Silvana Robone, and Peter C. Smith. 2010. "International Comparison of Public Sector Performance: The Use of Anchoring Vignettes to Adjust Self-Reported Data." *Evaluation* (16) 1: 81–101.
- Rice, Nigel, Silvana Robone, and Peter Smith. 2011. "Analysis of the Validity of the Vignette Approach to Correct for Heterogeneity in Reporting Health System Responsiveness." *The European Journal of Health Economics* 12(2): 141–162.
- Spalding, Nicola J. and Terry Phillips. 2007. "Exploring the Use of Vignettes: From Validity to Trustworthiness." *Qualitative Health Research* 17(7): 954–962.
- Tandon, Ajay, Christopher Murray, Joshua Salomon and Gary King. 2002. "Statistical Models for Enhancing Cross-Population Comparability." *Global Programme on Evidence for Health Policy Discussion Paper No. 42*. Geneva: World Health Organization.
- Vittersø, Joar and Fredrik Nilsen. 2002. "The Conceptual and Relational Structure of Subjective Well-Being, Neuroticism, and Extraversion: Once Again, Neuroticism is the Important Predictor of Happiness." *Social Indicators Research* 57(1): 89–118.
- Wilson, Warner R. 1967. "Correlates of Avowed Happiness." *Psychological Bulletin* 67(4): 294–306.
- Yang, Yang. 2008. "Social Inequalities in Happiness in the United States, 1972 to 2004: An Age-Period-Cohort Analysis." *American Sociology Review* 73(2): 204–226.

责任编辑:田青