

关于外来人口分母效应的再讨论

郭志刚

【内容摘要】本文就梁秋生对京津沪的外来人口分母效应检验提出的反质疑做出说明,并补充了采用新的划分口径和采用其他数据进一步检验的结果。新划分口径的检验和更大数据样本的检验仍然表明京津沪三城市中外来人口分母效应并不像他所推算得那样显著,本地妇女的总和生育率仍显著低于 1.0。本文还指出了梁秋生在计算北京平均生育年龄变化时的错误、以及他在把握生育年龄变化方面的误解,并探讨了在理解当前总体生育水平方面时生育进度效应与外来人口分母效应的层面问题,以及外来人口分母效应能够得以体现的必要条件问题。本文将梁秋生对京津沪外来人口分母效应推算中应用的假设与五普 1% 样本数据的相应统计结果进行了具体比较,指出他在推导在校学生性别年龄分布时可能产生的漏洞及其影响。

关键词: 外来人口;总和生育率;生育进度

【作者简介】郭志刚,北京大学中国社会发展研究中心研究员、社会学系教授。北京:100871

本刊今年第 2 期发表了梁秋生教授“再论大城市超低总和生育率中外来流入人口的分母效应问题”(2005),对我在本刊今年第 1 期上发表的根据京津沪的五普抽样原始数据对总和生育率中外来人口效应的检验结果(2005)提出了一系列的反质疑。我想仅就以下几个方面做个回应,谈谈自己的看法,并补充一些新的研究分析结果。

1 对学术争鸣的态度

怀疑是学术发展的动力之一,科学的本质要求所有的观点和认识应该接受不断的检验。所以,怀疑是科学发展中的正常现象。学术前沿涉及的都是我们知之不多的方面,在探索中出现错误也是正常现象。所以,怀疑是深入探索的开始,我们应该报以欢迎态度。但是,大家在学术探讨中都是平等的,每个人都不要给别人划什么禁区,学术讨论中主要应陈述根据来支撑自己的观点,不要轻言“不容置疑”或“不可否认”等。

怀疑也应包括自己对自己以前研究结论的怀疑。事实上,我对外来人口分母效应的检验,同时也是对自己以前用北京市五普数据简单按户口状况划分来比较本市和外来人口生育水平研究结果的一种检验。受梁秋生教授(2004)观点的启发,北京大专院校在校学生数量很大,而他们又属于北京户口,因此我也很想知道在控制在校学生以后,以前的结论会不会有变化。

至于研究中采用的具体数据和具体方法上存在的问题,都可以质疑和讨论。大家可以根据不同数据和方法就外来人口分母效应进行新的检验,不断修正和提高自己对于问题的认识水平。

我欢迎梁教授或其他学者就此研究提出更好的方法,我很乐意继续对此问题进行研究。

2 对检验的进一步说明和补充

2.1 不同划分口径的检验

研究方法要服从研究目的。我对外来人口分母效应的检验中所采取的人口划分口径正是循着梁教授的两个基本观点的思路建立的,试图控制省际外来流入人口和高校在校人口。但是正如我在上一篇检验报告所指出的,这种划分在操作上具有一定复杂性。比如,外来流入人口的时间界限如何具体把握;又如发现京津沪“在校学生”中居然有接近 1/3 的人为本地出生且拥有本地家庭户籍,对这些情况应如何操作处理都是新问题。特别后一种情况是出我意料的,我想梁教授在其推算中可能也并

未考虑过。问题是,梁教授是否认为这些人也应作为外来流入人口?我在检验中没有把这些人从本地人口中排除出去。要是将这些人的效应也从本地人口中划分出去,那么最好不要笼统说是“外来人口分母效应”,因为她们实在是属于本地的“高等教育发展效应”。

先不提概念定义之争,为了搞清这种情况到底有多大影响,我在原来口径1的基础上稍做调整,将那些虽然户籍在本市家庭户但申报为“在校”的育龄妇女权且算做是“外来”的,形成了口径3。由于这一变化涉及到京津沪育龄妇女中的案例相对很少(案例数为395人,仅占原来口径1中本市育龄妇女的12.8%),其统计结果与口径1相比变化也很小。按照口径3统计,本市妇女的TFR从口径1的0.7024提高为0.7173,而外来妇女的TFR从口径1的0.7842降低为0.7801。因此,原有结论都仍维持不变。

梁教授此次质疑我检验时将五普前五年内新迁来的户籍人口全部当作外来流入人口会影响的检验结果的可信度。其实,这倒不是我的创造,原本也是按梁教授自己的定义做的。梁教授在其前一篇文章中明确地说,“这样定义的外来流入人口,既包括‘五普定义的外来人口,也包含户籍迁入人口、户口待定人口和原住本地、现在出国工作学习而暂无常驻户口的人口”。

本着认真求实的态度,这里按梁教授再次提出的要求对五普1%样本京津沪育龄妇女案例再附加了两种划分的检验。

前一种我们就称其为口径4,划分时将由外省流入妇女作为一类(即户口不在本乡镇街道而且也不在本市),其余都作为本地妇女(其中也包括市内流动者)。其结果是本地妇女的TFR为0.7390,外来流入妇女的TFR为0.8183。这种划分口径应该是严格对应着梁教授上篇论文表3的统计口径,但是结果并没有出现像梁教授推算那样三大城市一致出现了本市妇女比外来流入妇女生育水平高的情况,而是显示外来流入妇女略高于本地妇女。

后一种称其为口径5,操作方法即在口径4划分基础上再将本地妇女中的所有高校在校生(不论户口在学校集体户还是在家庭户)以及那些户口待定和出国人员的情况划入外来流入人口。统计结果为本地妇女的TFR为0.7750,另外一类妇女的TFR为0.7695。这种划分口径应该是严格对应着梁教授上篇论文表4的统计口径。按梁教授的推算结果,本市妇女的TFR已经都高于1.0了。但是,这里按五普样本汇总结果却不过是本地妇女比另外一类的生育水平略高一点,但仍然远远低于1.0。

其实与梁教授结果更具可比性的是看口径4和口径5的本地妇女TFR上的变化(反映在校生等类人口形成的外来人口效应),五普1%样本统计这一TFR变差仅为0.04,而梁教授推算出这一变化很大,北京、天津、上海分别为0.35、0.15和0.28。

以上这几种按新的划分口径重新汇总的结果,都没有改变我上篇论文中所做检验的基本结论:京津沪三大城市的外来人口分母效应不像梁教授推算的那样显著,京津沪本地妇女的总和生育率水平显著低于1.0的水平。对于这样的检验结果,我并不感到奇怪。

更明确地说,我认为梁教授与我之间的观点差异其实主要并不在于外来人口是否有很低生育率上,而是在于京津沪本地人口是否有较高的生育率上。梁教授在未控制在校生时推算的市内妇女TFR与外来妇女TFR差距本来就不太大,而进一步的调整就是从先推算的本地生育率的分母中排除在校女生等(因为她们生育概率极小),所以,实际上他经过两番调整推算出的京津沪TFR其实就是本地妇女(并排除在校女生等)的生育水平,并且按他的推算,京津沪本地育龄妇女的TFR应该在1.0以上。而我则根据自己以前的研究和新对外来人口分母效应的检验,不相信本地原有育龄妇女的生育水平有那么高。正因为如此,我一方面要看的是外来人口生育水平有多低,另一方面要看的是京津沪本地妇女的生育水平有多高。这就是为什么在我的检验中,总是清楚地将全部育龄妇女划分为本地与外来两个部分,分别计算其生育水平。另外,在我的检验中,只有划分口径不同,但并未涉及什么假设。

2.2 关于样本规模问题

梁教授发现我以前用北京五普 1%样本与上次全国五普 1%样本所计算的北京市总和生育率有所不同,因此就检验时样本规模小的问题对整个检验结果提出怀疑。

样本不同总会造成统计结果有些不同,因为各样本不可能对每个局部都有同样的抽样比。样本规模大时,统计结果的可靠性就大。正是由于样本偏小,我更看重的是京津沪的一揽子的检验结果,而将京津沪各自的检验结果列出来只是一个补充。其实,我最先是用北京市五普的 1%样本进行过相同方式的检验,已经可以表明北京的外来人口效应并不像梁教授推算得那样显著。之所以上次正式发表的是全国 1%样本检验结果,主要是想对应梁教授对京津沪的分别推算。并且,由于在全国普查办系列招标课题中有几十家单位都是用 1%样本做的,为了保持一致,并使其他学者有可能复制再现我的检验结果,或采用不同方法来进一步做检验,所以并未发表北京更大样本的检验结果。后来,我曾有偶然机会接触过全国五普的 1%抽样数据,也按上述方法做过检验,其主要结论并无什么大的变化。既然梁教授对于 1%样本的结果存有那么大疑义,我不妨在此将更大规模样本的检验结果公布出来供大家参考。

表 1 提供了分别按各城市所做的汇总结果。汇总的北京、天津和上海的总和生育率分别为 0.724、0.912 和 0.675,与五普全部长表统计结果(0.67、0.88 和 0.68)的误差均在 ± 0.06 以内。

按照梁教授的推算步骤,口径 4 已经将外省市的流入人口作为外来人口、口径 5 又进一步将大专及以上在校生、户籍待定人员和出国人员作为外来人口。因此从检验梁教授推算结果的角度,这里应当对比表 1 口径 4 和口径 5 的本地妇女 TFR 之差,那么京津沪的这个差均在 0.05 以内,远不像梁教授在排除在校生等效应前后 TFR 的变化值那样夸张。另外,表 1 中本地妇女 TFR 高于 1.0 的情况在任何一个地方或任何一种划分口径中都不存在。即使是在本地妇女 TFR 水平较高的口径 5 结果中,北京和上海的本地妇女 TFR 都低于 0.8,天津的本地妇女 TFR 虽然相对较高,但也不过为 0.95。因此,这一更大样本的检验仍然可以视为否证了梁教授对外来人口分母效应的过高估计。

表 1 还提供了分别用五普 1%样本对京津沪三城市合计的各种口径结果。特别值得注意的是,与相应的各城市分别统计对比,可以看到口径 5 的三城市合计统计的 TFR 结果显得本地妇女与外来妇女的差异增大,两者之间差异达到 0.12,这一差异幅度超过了每个城市单独统计的差异(最大的不到 0.08)。这种“合计效应”很可能也属于路磊(1992,1996)发现并分析过的总和生育率“怪现象”。

为了便于比较,表 1 还同时列出了 1%样本的相同口径结果。可以看出,两个样本在三市合计结果上的确存在一些差异。第一,从全市合计 TFR 水平上看,1%样本略高一点。第二,各对应 TFR 值的水平之间存在差异,但差异幅度均不到 0.1。第三,从本地妇女和外来妇女的 TFR 之差来看,1%样本有时(口径 5)会相对较小,有时会相对较大(口径 1、口径 2),有时甚至会改变方向(口径 4)。但是对于在比较口径 4 和口径 5 的本地妇女 TFR 之差时,方向并未改变,差值有细微缩小。也就是说,1%所反映的在校生等类的效应还不及 1%样本反映的那样大。

总而言之,更大的普查数据样本的统计结论与五普 1%样本的基本结论一致,所以梁教授质疑的检验时样本规模问题的影响并不太大。

3 关于婚育年龄变化的测量

我在上篇检验论文中提到,我的怀疑是以生育进度效应理论为依据的。梁教授在“再论”中并未对我质疑他的推算假设做任何进一步的论证和验证,却转而提出婚育年龄推迟不大可能是京津沪低生育率的主要原因,因为他自己认为这方面的变化并不很大。但是,他的根据却存在明显的错误。

路磊发现,全部人口的 TFR 完全有可能同时小于或同时大于两个分人口的 TFR。他用图析法和解析法证明,导致这一现象的原因主要是两个分人口之间的相对年龄结构差异。路磊提到,上海 1986 年的全市 TFR 同时分别大于城镇 TFR 和农村的 TFR。笔者在分析北京五普数据时,也发现过全市 TFR 分别小于本市户籍和外来育龄妇女的 TFR。其实,就在本文的表 1 中也存在这种现象。

表 1 京津沪不同口径的本地与外来妇女的总和生育率

地区	口径 1:		口径 2:		口径 4:		口径 5:		合计
	本地	外来	本地	外来	本地	外来	本地	外来	
北京	0.770	0.715	0.595	0.864	0.732	0.733	0.767	0.693	0.724
天津	0.901	0.911	0.792	1.118	0.904	1.073	0.949	0.956	0.912
上海	0.649	0.683	0.573	0.770	0.673	0.694	0.711	0.653	0.675
总计	0.771	0.747	0.655	0.865	0.770	0.743	0.813	0.695	0.753
1%总计	0.702	0.784	0.572	0.946	0.739	0.818	0.775	0.770	0.765

他的第一个根据是,按马小红(2003)提供的北京 1990 年代末期的平均初育年龄与五普相应统计计算了变化量,然后折算年均增量,举证说明北京的初育年龄年均变化极小。但从行文中便可以看出他在计算中犯了一个低级错误。大概他是一时误将 1990 年代末期错当作 1990 年年末了,于是除以 10 来求年均增量。问题在于,他自己其实也不得而知这个变化到底是几年的变化,总之不由分说就除以 10 是不对的。其实马小红在本刊一年前所发表的另一篇论文(2004)中已经明确说明,上述被梁教授所引用的北京市平均初育年龄实际为 1999 年的统计数字。也就是说,梁教授求出的平均初育年龄变化其实本来就是一年的,而他再缩小 10 倍便是画蛇添足,那么在计算错误基础上继续做出的其他推论也就毫无意义了。

他的第二个根据来自于北京市前两次人口普查的年龄别生育率曲线图的比较。他发现,主要是生育水平(表现为曲线峰度)下降了,而生育率模式(表现为曲线偏度)并未明显变化,据此认为“足以说明,婚育年龄的推迟不是形成北京市超低 TFR 的主要原因”。对此推论我只能说,显然梁教授不太清楚应当如何准确测量生育年龄的变化。且不说就在他所绘图中生育率峰值年龄已经提高了一岁,即使平均生育年龄有所提前,也不能作为婚育年龄没有推迟的佐证。原因是,他这里并未划分孩次来分别进行平均生育年龄的比较,因此这种笼统的探查并不足以把握分孩次生育年龄的变化。

以往对平均生育年龄变化的研究文献已经多次用经验数据的统计图表清楚地说明(Bongaarts and Feeney, 1998; 郭志刚, 2000),不划分孩次的笼统的平均生育年龄即使没有变化、甚至在有所降低的情况下,分孩次计算的平均生育年龄仍然可能存在显著的提高。其原因是,笼统的平均生育年龄是各孩次平均生育年龄的加权平均数,在生育转变中各孩次平均生育年龄的提高往往同时伴随着高孩次生育比例的降低,而高孩次的平均生育年龄一般高于不分孩次的笼统平均生育年龄,这样一来总生育中高孩次比例减少产生的结构性影响便可能从权数的角度完全抵消掉、甚至超出各孩次平均生育年龄推迟的效应,导致笼统计算的平均生育年龄不变、甚至反而下降。打个简单的比方加以说明,比如,一孩和二孩的生育年龄分别为 23 岁和 25 岁,孩次比例各占 50%,那么笼统的平均生育年龄为 24 岁;如果分孩次的生育年龄都提高 0.1 岁,而一孩和二孩比例分别变为 80%和 20%,尽管孩次别生育年龄都提高了,笼统的平均生育年龄为 $0.8 \times 23.1 + 0.2 \times 25.1 = 23.5$ 岁,结果反而降低了。所以,要想细致地研究生育年龄的变化,必须划分孩次来统计,才能控制这种生育结构变化的影响,反映出各孩次平均生育年龄的真实变化。

梁教授在“再论”中还根据我上篇“检验”论文中本市和外来妇女的生育率图型猜测京津沪育龄妇女的生育年龄的提高可能主要是外来流入人口所致。这又是一个同样的误解,因为那个图型中提供的还是未分孩次的生育率,而且我们知道京津沪本地育龄妇女能生育二孩已属凤毛麟角,但外来人口生育中二孩比例则还相对较高(郭志刚, 2004a),当然其生育率曲线在对应较大年龄时相对较高。

4 关于婚育年龄推迟的进度效应和外来人口分母效应的进一步讨论

4.1 我为什么更强调婚育年龄推迟效应对时期生育水平的影响

梁教授在“再论”中认为“剔除外来流入育龄妇女的分母效应也能同时部分排除婚育年龄推迟对

TFR 的影响作用”。这是对的,但也只是“部分”排除,本地原有妇女的婚育年龄推迟效应就未被排除。我从习惯上总是要考虑这两种理论解释在总体上认识当前生育水平时所处的层次。我们应该清醒地知道,外来人口效应只适用于解释局部地区的低生育率问题。而婚育年龄推迟效应的适用性更强,在方法论角度处于更高层次,或者说更具有一般性。比如,它既可以部分地解释为什么当前总的生育水平很低,也可以应用于解释外来人口很少的地方的低生育水平,还可以应用于解释京津沪的原有人口的总和生育率甚至大大低于政策要求的水平。甚至,它也可以直接解释外来人口本身的低生育率。比如,京津沪外来人口的 TFR 都显著低于 1.0,大家当然不会认为这种水平就代表了这些人的真实终身生育水平吧?外来人口之所以有可能在所到地区产生降低当地生育率的分母效应,就是因为他们的时期生育率很低;而为什么他们的时期生育率那么低则与他们自己由于迁移流动而推迟婚育年龄有很大关系。所以,要想非常细致地研究迁移人口的时期生育水平为什么这么低,恐怕仍然离不开 Bongaarts 和 Feeney 提出的生育中的数量效应与进度效应的研究框架。

这两种观察生育率变化的方法的层次差别还可以从内因和外因的角度来理解。生育年龄推迟效应更接近于内因,而外来人口分母效应则更接近于外因。在京津沪超低总和生育率问题上,梁教授的观点只强调从外因的解释,否定了内因的解释(如他后一篇论文已经开始否定北京本地妇女婚育年龄的推迟);而我则坚持不肯放弃内因的解释,并且还认为外因本身也有其自己的内因解释。

关于总和生育率,我以前曾阐述过以下观点(2002):(1)以前总和生育率一直服务于两个功能,一是作为时期生育水平的测量,二是作为终身生育水平的测量估计;(2)总和生育率作为终身生育水平的测量估计存在缺陷,而 Bongaarts 和 Feeney 新提出的去进度效应总和生育率可以比较有效地克服这些缺陷,因此我建议用这一新指标替代常规总和生育率作为终身生育水平的测量估计。梁教授在其上篇论文中曾索引过这篇文章,对上述观点既没肯定,也没否定。但在“再论”中,则明确提出了相反的观点,即认为“在推断终身生育水平方面 TFR 具有不可替代的作用,而时期生育水平可以用一般生育率或其他指标来描述”。但是前一句只是断言而已,既无理论支撑,又无实证检验。实际工作需要了解终身生育率,而真实终身生育率测量总是马后炮,因此就需要根据时期数据加以估计,常规 TFR 和去进度效应的 TFR 都不过是一种估计,重要的是哪一种估计更好一些。在学术争鸣中仅靠一句“不可替代”是很难让人信服的。至于后一句,则不知所云。

我曾经推算过 1990 年代全国的去进度效应的 TFR['](2004b),发现它与常规 TFR 之间存在着较大的差异,这说明在全国范围内 1990 年代婚育年龄的推迟导致常规总和生育率偏离终身生育水平的幅度很显著,因此这是一个普遍现象,那么三个直辖市的原有居民在更好的社会经济条件下比较明显的推迟婚育应该不是什么离经叛道的推论。并且,我对北京市 1990 年代分孩次平均生育年龄也的确反映出明显推迟,对去进度效应总和生育率的研究也的确反映出婚育年龄推迟效应的存在(2004a),正是因为梁教授对这三个城市本地妇女生育水平的较高推测值几乎留不下婚育推迟效应存在的余地,才直接导致我对他推算结果的怀疑。

4.2 外来人口分母效应产生的必要条件

我的本意只是质疑京津沪的外来人口分母效应是否像梁教授推算得那样高,并不一概否定局部地区存在这种效应,因为我知道其他一些研究人员早就发现广东的低生育水平是与外来人口相联系的。本期杂志中陈卫(2005)所发表的研究结果再次证明在广东省外来人口分母效应极为显著,而且更重要的是他还发现在全国城市中也普遍存在比较显著的外来人口分母效应。然而他的研究也发现京津沪是个例外,尽管外来人口很多,但分母效应反而不太显著。这些统计结果对我有很大启发,引起我思考这种结果背后的原因是什么。

从理论上我们可以推出,一个地方的外来人口能否产生分母效应,除了外来人口占本地现有人口的比例相对较大以外,一方面得要求外来人口的生育水平很低,另一方面还得要求本地原有人口生育水平相对较高。所以更准确地说,不管外来人口分母效应是降低还是提高了本地的生育水平,其本质

是外来人口与本地人口在生育率上存在着显著的差异,而不是仅仅取决于外来人口生育的绝对水平。

换句话说,在广东省或全国城市,外来人口的生育率与本地人口的生育率之间的差别更为显著,于是便能充分体现出外来人口的分母效应。而京津沪尽管外来人口较多、其生育率很低,但是本地原有人口的生育率也很低,因而两者之间差别不太显著,于是分母效应反而不太容易体现。从这个角度来看,虽然梁教授提出外来人口分母效应的研究命题很有学术价值和实际意义,但他选择京津沪为例来体现这种效应的存在则并不太恰当,因而其研究结论也相对更容易证伪。

梁教授在“再论”中对自己原文潜在含义总结为“像京津沪这样,不仅拥有大量外来人口,而且拥有众多高等院校在校大学生人口、无户籍人口和出国学习工作人口的大城市,其育龄妇女的超低 TFR 在很大程度上是受这些外来流入人口分母效应影响所致。事实上,就全国来说具有这种前提条件的小城市和镇来说,则不能得出这样的结论;对于全国和人口外流的农村地区而言,根本不存在流入人口的问题,更谈不上流入人口的效应问题。”根据以上相对生育率差异这一必要条件的来看,梁教授上述前半段话可能过于强调了大城市中的外来人口效应的幅度,而后半段话又有可能局限了外来流入人口分母效应发生的范围。

5 对梁教授的外来人口分母效应推算的质疑

迄今为止,我只是表示了对梁教授推算结果的怀疑,并且认为问题可能出在他推算中的假设上。正如梁教授所说,我是“另辟蹊径”来对其结果进行检验的。我自己并不认为这种作法有什么不妥,因为科学研究方法论认为,一个理论观点的可证伪性越强表明其科学性越强。要检验一个理论观点或推演结论,完全允许选择不同的环节或采用不同的手段,只要所选的环节和手段对于研究目的是有效的。这是科学发展史的通则。

但梁教授发表的“再论”已经表明,他无意自己进行检查,并对我未能直接论证其原文结论的可信性表示“遗憾”。这里,为了方便梁教授和其他有兴趣的读者,我用五普 1% 样本的京津沪案例做了一些简单汇总,来与梁教授的推算假设加以比较,探讨一下我与梁教授得到的不同结论究竟是从哪里分道扬镳的。

首先,梁教授假定京津沪在校大学生和研究生中 90% 来自外地,这与五普样本汇总结果不合。因为从原始样本数据来汇总,京津沪女性大专及以上学历在校中占有很大比例(五普 1% 样本中为 33.7%,而 1% 样本为 32.8%)为本地出生且户口就任在本市家庭户,且不说那些户口在学校集体户的学生中还可能有一部分就是本地原有人口。于是,他的假定会导致本地妇女生育率推算中对分母的扣除相对较多,当然会算出本地生育水平相对较高。

其次,鉴于梁教授自己在未调整在校大学生等其他情况时所推算的外来人口效应也不太大,下面仅集中讨论高校在校大学生等另类的分布假设。梁教授调整在校大学生等的分母效应时应用了一系列假设推论,我们且将其与五普 1% 样本数据汇总结果做一比较。

(1) 梁教授假定各类需要控制的人口的性别比为 100,即假设高校学生为男女均匀分布。从五普 1% 样本中反映出京津沪大专以上学历在校生的性别分布并不是这样,其中大学专科生 337 人,性别比为 90.4;大学本科生 501 人,性别比为 145.6;研究生为 83 人,性别比为 232.0。可见京津沪样本中高校在校大学生中大学本科及以上的学生所占比例较大,且男生远远多于女生。要是将大学专科与本科合计起来统计,性别比为 119.9。也就是说,梁教授按高校在校大学生性别均匀分布假设来推算时,育龄妇女分母中就会扣除更多,提高推算的生育率,体现出更显著的分母效应。

(2) 梁教授对京津沪在校大学生的年龄分布假设则是根据他本人所在学校的入学新生的年龄分布

作为参照:该样本中全国大学专科生 3372 人,性别比为 101.4;大学本科生 3662 人,性别比为 153.6;研究生 230 人,性别比为 206.7。

进一步推出的。这里,如果入校新生的年龄别的性别分布假设不合实际,又会导致所有在校女生的年龄分布假设上出现不合实际。

(3) 谨依据梁教授提供的大专生和本科生的年龄分布比例(就是说女生也是如此),其中22岁及以上的合计比例为35.5%,然而用五普1%样本京津沪案例汇总的相应比例只有13.1%,可见两者之间差距悬殊。计算22岁以上合计比例是因为生育高峰期大致是从这个年龄开始。我们知道,分母效应主要体现在生育高峰的年龄段,而在其他年龄组影响较小。比如15岁女性生育率低到可忽略程度,缩小其分母规模对大体水平影响并不大,但在生育高峰年龄组就不同了。与梁教授的女生年龄分布假设相比,五普样本的京津沪大学女生并没有那么高的比例处于生育高峰年龄段。

此外,出于很想探讨一下梁教授所在学校的在校大学专科和本科22岁以上女学生所占比例为什么那么高的目的,我循着梁教授的思路复制了他的推算(表2)。看得出来,梁教授是从某一年的新生年龄分布简单递推出4年,然后得出了在校生的总体年龄分布。这种方法类似于计算一个静止人口生命表,只不过每个入学队列进入时的年龄便不相同而已。虽然梁教授没有提,其实这种推算方法背后至少还有另外4个假设:第一,各年入学新生年龄分布相同;第二,各年入学新生数量不变;第三,各年新生中大专与本科的分布比例不变;第四,大专与本科一样学制为4年。这4个假设其实风险也挺大的。比如,各年入学新生会不会出现年轻化、各年招生数量会不会扩大等等。如果存在这些情况,对应的便是一个稳定的或非稳定的增长型人口,根据人口统计原理可以推论在校生总的年龄分布将会年轻化,也就是说可能会导致推算的外来人口分母效应缩小。

能够明确提出异议的主要是大专学制也为4年的假设。按照这种假设计算就意味着大专学生的在校留存期比实际情况延长了一年,其后果会导致处于生育高峰期的年龄比例提高,因而加大对生育率的调整影响,即扩大分母效应。我在表2的试算中,将第3项假设维持在大专和大本各占50%(大体在全国平均水平)的基础上将第4项假设改变为大专学制3年,其模拟结果便可以导致推算的大专和本科女生中22岁及以上的合计比例从原来的35.5%降到27.4%,可见结果对这一假设很敏感。这个假设从统计上延长了大专学生的在校“寿命”,因此梁教授对所在学校在校大专和本科学生合计的年龄分布推算很可能已经过于“老龄化”了。

然后,梁教授将这些假设基础上的推算结果再移植到京津沪所有大专院校头上,又给其外来人口分母效应研究再加上一层风险。比较表2提供的五普样本对京津沪、以及全国的大专和本科在校女生汇总的年龄分布,可以看出梁教授推算的在校学生年龄分布结果与生育高峰年龄段更为相关,自然是很有利于增大外来人口分母效应的。

(4) 梁教授在其最后推算的原表4既排除了在校学生、也同时排除了户口待定人员和出国人员。这里且不说梁教授假定户口待定者一定是来自外省市是否属实,这两种人口必须要做性别年龄划分则是绕不过去的。梁教授并未明确说明他是否掌握这些人的性别信息,所以可能对他们也采用了性别均匀分布假设。另外,对比梁教授论文(2004)中的表3和表4,发现30岁以上各年龄组妇女数并未做任何调整,可见梁教授也没有这两种人口的年龄信息。而且,他并未提供任何证据就认为“在校研究生、户籍待定人口和出国学习工作的人口,年龄绝大多数分布在25~29岁之间”,所以很可能他将这两种人口的半数从25~29岁(注意这也是生育高峰段!)的生育率分母中排除了。尽管五普1%样本的京津沪案例汇总中这两种情况人数很少,但是可以看出与梁教授上述认识差别极大(表3)。首先,在这两种户口状况人口中,其实25~29岁的人口比例很小,尤其是户口待定类别的女性。梁教授认为“户籍待定人口中,绝大部分属于高校毕业后寻找工作的人口”,而汇总结果却显示出他们中大多数却是少儿人口,而且在15~29岁间的户口待定人员中还是男性远远多于女性。因此,尽管这两类

情况在总人口中数量相对很少,但是如果梁教授真的是将这两类人口总数取半仅从 25~29 岁一个女性年龄组的生育率分母中排除的话,其影响就可想而知。

表 2 在校大专和大本学生年龄推算表与五普 1%样本汇总结果

年龄	第 1 年	第 2 年	第 3 年	第 4 年 (1)	合计 (1)	列 % (1)	第 4 年 (2)	合计 (2)	列 % (2)	京津沪 样本 %	全国 样本 %
< 18										2.9	6.4
18	8				8	2.0		8	2.3	20.7	18.4
19	60	8			68	17.0		68	19.4	30.2	25.8
20	18	60	8		86	21.5		86	24.6	20.2	20.7
21	10	18	60	8	96	24.0	4	92	26.3	12.9	16.1
22	4	10	18	60	92	23.0	30	62	17.7	8.7	7.2
23		4	10	18	32	8.0	9	23	6.6	2.1	2.6
24			4	10	14	3.5	5	9	2.6	1.0	0.7
25				4	4	1.0	2	2	0.6	0.3	0.5
> 25										1.0	1.6
总计					400	100.0		350	100.0	100.0	100.0
22+ 合计					142	35.5		96	27.4	13.1	12.7

注:列标题中,用(1)代表“大专学制 4 年”假设,用(2)代表“大专学制 3 年”假设。最后两列为五普 1%样本中京津沪和全国在校大专和大本在校女生年龄分布。

表 3 五普 1%样本户口状况与年龄性别分布的概要统计

统计指标	京津沪				全国			
	户口待定		出国		户口待定		出国	
	男	女	男	女	男	女	男	女
案例数	137	139	41	50	3750	3903	269	231
15 岁以下 (%)	60.6	82.0	9.8	8.0	79.0	79.0	3.3	5.2
15~29 岁 (%)	27.0	11.5	24.4	32.0	13.7	14.7	39.8	45.0
25~29 岁 (%)	2.2	2.9	12.2	24.0	2.4	3.2	18.2	20.8

最后,我们再来综合性地比较一下,梁教授在从上一轮的本地育龄妇女中将大学在校生、户籍待定及出国人员划拨到外来流入人口中去时到底应用的是多大的比例,他的推算比例与五普样本汇总出的相应比例有多大差别。表 4 中提供了用五普 1%样本数据按口径 4 和口径 5 汇总的本地育龄妇女数的变化量占口径 4 本地育龄妇女的比例,同时提供了根据梁教授论文(2004)中的表 3 和表 4 结果计算的相应比例。可以看出,与五普汇总结果相比,梁教授的调整比例特别突出了处于生育高峰期的 20~24 岁和 25~29 岁两个年龄组,于是我们就可以知道为什么梁教授的这一步调整产生了那么大的外来人口分母效应。同时也可以看出,正如前面已经逐项分析过的那样,梁教授的大多数假设并不像他所说的那样既可能产生正影响也可能产生负影响,当它们的影响累加起来时,并没有太多的抵消,而是集中地体现在生育高峰段上了。并且,梁教授施加在生育高峰年龄组的生育率分母调整量中的还有相当部分是借助其假设从其他女性年龄组、甚至是从男性那里“挪”过来的。

表 4 中的比例是分母调整人数占原来分母的比例。有了这些信息我们还可以比较,在梁教授上一步排除了外来流入人口的基础上,如果将梁教授那些对大学在校生、户籍待定人员和出国人员的性别年龄分布假设均换用五普 1%样本汇总出的相应分布比例的话,到底调整的 TFR 结果有多大差别。

表 4 在校生、户籍待定及出国人员在本市户籍人口中的比例

地区		年龄组						
		15~19	20~24	25~29	30~34	35~39	40~44	45~49
北京	五普汇总	15.9	23.8	2.8	1.6	1.0	0.0	0.4
	梁秋生推算	9.8	48.8	28.6	0.0	0.0	0.0	0.0
天津	五普汇总	8.7	18.5	1.4	0.6	0.0	0.2	0.2
	梁秋生推算	3.4	18.0	12.9	0.0	0.0	0.0	0.0
上海	五普汇总	17.0	13.6	2.4	1.7	0.8	0.3	0.3
	梁秋生推算	6.9	32.9	26.2	0.0	0.0	0.0	0.0
京津沪	五普汇总	14.3	18.4	2.3	1.3	0.7	0.2	0.3
	梁秋生推算	6.8	34.1	23.2	0.0	0.0	0.0	0.0

表 5 提供这一比较。按照梁教授对在校生等类计算出分母效应很显著,导致北京和上海的新的 TFR 的变化幅度均在 0.3 左右,而天津也有 0.15,并且三个城市的调整 TFR 都显著超过了 1.0。但是如果将这些调整对象的分布均按照五普样本汇总统计,那么北京和上海 TFR 的调整变化幅度均不到 0.1,调整后 TFR 均不过略高于 0.8。而天津 TFR 的调整变化幅度稍大,有 0.11。而天津调整后的 TFR 高过 1.0 其实还受到梁教授推算的基数 TFR 已经很接近 1.0 的影响,因为调整变化绝对值不仅受变化倍数影响,同时也受基数大小的影响。从五普 1%样本对天津的汇总情况看,这一步调整不过使天津从 0.904 提高到 0.949。

表 5 应用普查汇总的在校生等在本市户籍人口中的比例所计算的总和生育率

	北京	天津	上海
梁秋生第一步推算结果	0.73	0.97	0.76
梁秋生第二步推算结果	1.08	1.12	1.04
如按五普汇总分布推算*	0.81	1.08	0.83
五普 1%样本口径 4 统计	0.732	0.904	0.673
五普 1%样本口径 5 统计	0.767	0.949	0.711

* 按五普汇总分布重新推算的方法:定义表 4 中五普样本汇总的比例为 d ,那么分母变化率为 $a=1-d$,年龄别生育率 f 实际变化率为 $1/a$ 。以梁教授前一步(原表 3)推算的市内生育率 f 为基础,按五普汇总分布计算的新的调整生育率的公式为 $f^*=f/a$,然后再计算出重新调整的 TFR 来。

比较梁教授第一步调整结果和本文口径 4 的对应结果,可以看到他对外来流入人口进行调整后,北京本地妇女的 TFR 与五普汇总结果很接近,但天津和上海的结果已经开始与五普 1%样本直接汇总结果出现了差距。这一差距来自于梁教授关于外来流入人口的性别年龄结构与年龄别妇女生育数的推算,以及五普样本中相应的误差。而梁教授最终结果与五普样本口径 5 汇总结果之间的差别,仍包含着上一步推算的差别,此外除了上述谈到的生育率分母的调整数总量及分布影响及五普样本误差外,其实还有生育率分子的影响。梁教授假设在校生、户口待定及出国人员都没有生育,然而我在计算口径 4 和口径 5 生育率时发现,生育数其实也有微小变化,也就是说,被排除出分母的部分也有少量生育,主要发生于户口待定人口和出国人口中。

6 结语

本文根据五普抽样数据提供了对京津沪外来人口分母效应更多的数据分析结果。不可否认,抽样数据总会存在一定误差。梁教授的推算虽然立足于更大规模数据,但是其推算中涉及了很多假设,

作为同行,我完全理解梁教授当时的推算必须借助于一些假设才能完成。但是,由于研究中尚不具备所有信息,研究人员有时不得不基于一些假定先做出初步结果,这些结果发表出来对于其他同行也有启发和帮助。不过,研究人员应当十分清楚自己研究中的软肋,一旦得到新的信息,就得随时重做自己的研究。特别是当不同的研究结论出现时,既要积极探讨别人研究的有效性和可靠性,同时也应将不同结论视为一个警示,回过头来重新认真探讨自己以前研究中可能存在的不足,力所能及地做一些收集新材料、验证假设、检查计算的工作,以便多取得一些证据和把握再来回应不同意见。使我感到特别不解的是,梁教授在看到不同研究结论时,几乎没有对自己的研究产生任何怀疑,也没有对自己的那些假设做进一步的检查验证或敏感性分析、提出新的论证,就急于宣称自己的结论不容置疑或不可否认,而且甚至在概念不清、材料不准的情况下又莽撞地对自己尚不太熟悉的领域做出更多的新断言。说实在的,这些本来都是科学研究工作者应十分忌讳的事。

参考文献:

- 1 梁秋生.再论大城市超低总和生育率中外来流入人口的分母效应 与郭志刚教授商榷.人口研究,2005;3
- 2 郭志刚.关于京津沪超低生育率中外来人口效应的检验.人口研究,2005;1
- 3 梁秋生.外来流入人口的分母效应与大城市育龄妇女的总和生育率 以京、津、沪为例.人口研究,2004;5
- 4 路磊.论总和生育率的一个特性.人口研究,1992;4
- 5 路磊.论生育的度量分析.中国人民大学博士学位论文(人口研究所),1996-04
- 6 马小红.低生育率水平下的人口生育变动探析.全面建设小康社会中的人口与发展研讨会论文,2003-09
- 7 马小红.“双独政策”影响下北京市人口生育水平变动分析.人口研究,2004;1
- 8 Bongaarts and Feeney. 1998. On the quantum and tempo of fertility. Population and Development Review 24(2): 271 - 291/ 中文译文:生育的数量与进度[邦加茨、费尼 著].人口研究,2000;1
- 9 郭志刚.从近年来的时期生育行为看终身生育水平——中国生育数据的去进度效应总和生育率研究.人口研究,2000;1
- 10 郭志刚.北京市第五次人口普查生育状况分析.载北京市第五次人口普查办公室和北京市统计局编.北京市2000年第五次全国人口普查课题论文集.中国人口出版社,2004
- 11 郭志刚.总和生育率的内在缺陷及其改进.人口研究,2002;5
- 12 郭志刚.对中国1990年代生育水平的研究与讨论.人口研究,2004;2
- 13 陈卫.外来人口与我国城市低生育率.人口研究,2005;4

(责任编辑:段成荣 沈 铭 收稿时间:2005-05)