

人口迁移研究

省际人口迁移迁入地选择的影响因素分析

段成荣

【内容摘要】 当一个人作出迁移的决定之后, 还需要作出向何处迁移的决定。在以往的研究中, 对于人口迁移的迁入地选择过程缺乏研究。本文利用 1990 年人口普查资料, 通过建立人口迁移分析矩阵, 探讨了对人口迁移迁入地选择过程进行分析的实现方法。

【作者简介】 段成荣 博士, 中国人民大学人口所副教授。邮编: 100872

1 背景

作为一个完整的过程, 人口迁移至少包括两个方面: 决定是否迁移和决定向何处迁移。在目前的人口迁移研究中, 有关人口迁移是否发生的影响因素的研究比较多, 比如杨云彦有关我国人口迁移选择性的研究 (1994, 第 261 页至 300 页)。从地区之间的差异来考察地区之间的人口迁移流量和流向的研究也已开始, 比如王桂新对我国人口迁移与区域经济发展之间关系的研究 (1997, 第 373~430 页)。

但是, 有关“向何处迁移”, 即有关迁入地选择的研究则很少。已有的少量研究也主要是从迁入地与迁出地之间的流量和流向角度展开研究的。地区之间的流量和流向是无数个作为个体的迁移者的迁移信息的总和和概括, 它忽略了不同迁移个体之间的差异。从资料运用的角度讲, 这种研究属于“整合”型研究, 它存在信息浪费的问题 (段成荣 a, 2000), 也不能真正反映迁入地和迁出地的社会经济环境对作为个体的迁移者的迁移方向选择的影响作用。

正如水自高处流向低处一样, 作为人类活动的人口迁移也是人们由社会生活中的某种位置向另一位置的移动, 只是这时的移动方向不是由高处向低处移动, 而是由低处往高处移动。这就是通常所说的“人往高处走, 水往低处流”。判断高低的标准本身会因时而异、因地而异和因人而异。但“高”与“低”之间的核心问题是存在于两处条件之间的差异。从根本上讲, 人们主动进行迁移的目的就是要利用地区之间的形形色色的差异来实现自己利益的最大化。

由此可见, 地区之间的差异对于潜在人口迁移向实际人口迁移转化具有十分重要的意义。为此, 本文将考察地区之间差异对人口迁移迁入地选择的影响。考虑到获取资料的方便, 本文集中考察 1985~1990 年间我国省际人口迁移的迁入地选择过程。其他类型人口迁移的迁入地选择过程可以用同样的方法进行分析。将地区之间的差异与个体的迁入地选择过程结合在一起进行分析, 在以往研究文献中还不曾见到。如能建立某种分析框架用以准确测度地区间差异对个体迁入地选择过程的影响机制, 将为人口迁移研究提供新的分析工具。

2 方法

为了实现上述分析目的, 特提出如下分析思路: 设想有一个人, 受到自身特征和原居住地环境因素的影响, 作出了要进行省际人口迁移的决定。那么, 这个人可能迁往什么地方呢? 实际上又会迁往什么地方

呢?利用 1990 年全国人口普查资料,我们可以对这个问题作出比较准确的回答。

假定这个人的原居住地在北京市,而且他最终迁移到了上海市。下面我们来看一看这个过程是如何实现的。在他作出进行省际人口迁移的决定时,他可以迁往的省份多达 28 个,他可以迁往除北京市以外的其余 28 个省份(本文分析未包括西藏自治区和台、港、澳地区,当时重庆市尚未从四川省中分出,故全国总省份数量为 29 个)中的任何一个。但他的实际迁移活动只能是从这 28 个省份中挑选一个,一旦他挑选了其中的某一个省份,这一次就不能迁往其它 27 个省份中的任何一个了。我们的研究目的就是要弄清楚他为什么选择上海市而不选择任何其它省份。当然,对于某一个具体的人来讲,其迁移过程的最终决定因素或多或少具有某种偶然性。但许许多多的迁移者的迁移过程的决定因素有没有一定的规律性呢?如果有规律性,那么这种规律性又是什么呢?这是我们希望回答的问题。

在 1990 年全国人口普查原始资料中,我们可以根据被调查者在 1990 年 7 月 1 日常住地和 1985 年 7 月 1 日常住地之间的异同来判别一个被调查者是否在此五年间发生过省际迁移。如果一个被调查者在此期间发生过省际迁移,我们就把该被调查者称为一个省际迁移者。在 1990 年全国人口普查原始资料中,我们可以提取出所有的省际迁移者。为了分析的简便,我们从国家统计局提供的 1%原始资料中抽取了一个 5%的样本。换句话说,本文所用分析资料是一个 1990 年全国人口普查原始资料的十万分之五样本。在这个样本中,有 614 个被调查者曾经在 1985 年 7 月 1 日至 1990 年 6 月 30 日期间发生过省际迁移。本文以下的分析就是根据这 614 个省际迁移者的信息进行的。

为了达到我们的分析目的,需要对这 614 个省际迁移者的信息进行全面的加工。

首先,我们建立了一个 614×29 的人口迁移分析矩阵。在原始资料中,每一个迁移者对应着一条记录(或者说一行)。这条记录包含有该省际迁移者的“迁出省份”、“迁入省份”等信息。在新建立的人口迁移分析矩阵中,每一个迁移者将对应 29 条记录(行),即每个省际迁移者对应的记录将在原来一条的基础上再增加 28 条。在这些新增的 28 条记录中,每一条记录都将保留原有记录中的“迁出省份”和“迁入省份”等信息。比如,在前面所讲例子中,对于那个由北京迁移到上海的迁移者,其“迁出省份”将始终为北京市;其“迁入省份”则始终为上海市。当然,如果另一个迁移者是由四川省迁移到海南省,则该迁移者所对应的 29 条记录的“迁出省份”都是四川省,“迁入省份”都是海南省;

与此同时,我们为全部 29 条记录新增设一个变量:“潜在迁入省份”。对于每一个省际迁移者所对应的 29 条记录,其在变量“潜在迁入省份”上的取值将依次为北京、天津、……、上海、……、宁夏、新疆。它表示:每一个被抽中的省际迁移者都对应着全国 29 个省、市、自治区,或者说有 29 个可供选择的潜在迁入省份。当然,在这 29 个“潜在迁入省份”中,每一个省际迁移者必然来自其中的某一个省份,这样,他就不可能由该省份迁移到同一个省份。因此,我们自上述 614×29 的人口迁移分析矩阵中,删除掉“迁出省份”与“潜在迁入省份”相同的那一条记录(或者行)。关于迁移分析矩阵的建立和变量“迁出省份”、“迁入省份”与“潜在迁入省份”的取值,可以参看表 1 所示例子,该例子展示了如何将前述由北京市到上海市的迁移(一条记录)扩展为表示各种迁移可能性的 28 条记录。

这样,上述 614×29 的人口迁移分析矩阵就变成了 614×28 的人口迁移分析矩阵,它表示:每一个被抽中的省际迁移者都有 28 个潜在迁入省份:1 个实际迁往的省份,加上 27 个有可能迁往而实际未迁往的省份。

其次,我们给这一迁移分析矩阵中的每一条记录(行)设定另一个新的变量:“实际省际迁移”。我们用变量“实际省际迁移”来揭示省际人口迁移的迁入地选择过程,即在众多的潜在迁入省份中,迁移者“实际”选择了哪一个省份,而放弃了哪些省份?当一个被调查者的“迁入省份”与某一条记录中的“潜在迁入省份”的值相同时,他在该条记录中的“实际省际迁移”变量上就取值为 1;而他在其它记录中的“实际省际迁移”变量上的取值均为 0。仍以前面由北京到上海的迁移为例,这个省际迁移者的可能迁移去向有 28 个:天津、河北、……、上海、……、新疆等。实际生活中他迁移到上海了。因此,对这个迁移者来说,在表示“潜在迁入省份”为天津、河北、……、新疆等 27 个省份的记录中的“实际省际迁移”

变量上的取值均为 0，而在表示“潜在迁入省份”为上海的记录中的“实际省际迁移”变量上的取值为 1（见表 1）。变量“实际省际迁移”将是本文以下分析的目标变量。

这样，用于迁入地选择分析的迁移分析矩阵之基础部分就建立起来了。

第三步工作主要是选择用于反映迁出地与迁入地之间差异的合适指标，并对这些指标进行赋值。

那么，应该选择哪些指标来反映迁出省份与（潜在）迁入省份之间的差异呢？人口迁移研究文献表明：迁出地与迁入地之间在经济发展、社会发展等方面的差异以及两地之间的空间距离等是影响人们作出是否迁移决策的主要因素（Poston 和 Mao, 1998; Liang 和 White, 1997; Oberai 和 Singh, 1983; 王桂新, 1997 等）。

表 1 省际人口迁移分析矩阵的构建（例）

记录 编号	迁出 省份	迁入 省份	潜在迁 入省份	实际省 际迁移	人口 密度 1	人口 密度 2	密度差
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8) = (7) - (6)
1	北京	上海	北京	*			
2	北京	上海	天津	0	北京密度	天津密度	天津密度 - 北京密度
3	北京	上海	河北	0	北京密度	河北密度	河北密度 - 北京密度
.....							
8	北京	上海	上海	1	北京密度	上海密度	上海密度 - 北京密度
.....							
.....							
27	北京	上海	青海	0	北京密度	青海密度	青海密度 - 北京密度
28	北京	上海	宁夏	0	北京密度	宁夏密度	宁夏密度 - 北京密度
29	北京	上海	新疆	0	北京密度	新疆密度	新疆密度 - 北京密度

* 此条记录由于“迁出省份”和“潜在迁入省份”相同，在省际人口迁移中是不可能发生的。因此，此记录将被删除。

为此，我们选择了以下指标来反映迁出省份与迁入省份及潜在迁入省份之间的差异：每两个省份之间在农民人均纯收入、农村劳动力人均耕地面积、城镇职工人均工资收入、人均固定资产投资额、产业分化程度指数、人口密度、15 岁以上人口识字率等指标上的差异。产业分化程度指标 MI 反映的是一个地区产业发展的发达程度。一般来讲，一个地区的 MI 越高，其产业发达程度越高，相应的就业机会就越多，越容易成为吸引人口迁入的地区（Poston 和 Mao, 1998）。鉴于空间距离在人口迁移决策过程中具有特殊重要的影响作用，在此我们将迁出省份与（潜在）迁入省份之间的空间距离也纳入我们的分析模型之中。

这些指标的赋值办法是，考虑到这些指标对人口迁移的影响应该有一定的“时滞”，即 1990 年的人口迁移应该受 1990 年以前社会经济因素的影响而不是 1990 年当年社会经济因素的影响，我们选取 1988 年各省的相应指标值。当然，1988 年的选取本身带有一定程度的主观随意性。在选择自变量的取值时，到底应该选择离调查时点多远的指标值为宜，尚待进一步的研究。但有一点是清楚的：只要条件允许，选取调查时点之前的指标总是比选取调查时点上的指标更好，因为这样可以尽可能避免“同时性偏差”（参见段成荣 b, 2000）。与此同时，有些指标如识字率，只有人口普查时才能得到，这时，我们明知存在“同时性偏差”，也不得不利用调查时点上的信息。

对每一条记录，我们设置两类新的变量：一类是表示迁出省份指标值的“变量 1”；另一类是表示（潜在）迁入省份指标值的“变量 2”。比如，在上面由北京迁到上海的例子中，在记录 8 中，变量“人口密度 1”（迁出省的人口密度）取值为北京市的人口密度，而变量“人口密度 2”（迁入省的人口密度）则取值为上海市的人口密度。而在记录 2 中，潜在的迁移方向是由北京市迁往天津市，因而变量“人口密度 1”同样取值为北京市的人口密度，变量“人口密度 2”（“潜在迁入省份”的人口密度）则取天津市的人口密度（参见表 1）。其余记录的人口密度取值同样处理。

有关农民人均纯收入、城镇职工人均工资收入、识字率、产业分化程度等指标的取值，完全按与“人口密度”相同的方法进行。

按照上述办法,我们可以将反映迁出省和(潜在)迁入省之间差异的变量赋值给各记录。然后,用“变量2”(潜在)迁入省份的变量值)减去“变量1”(迁出省的变量值),就可以得到(潜在)迁入省份与迁出省份之间在各项指标上的差异。

各省之间的空间距离按照省会城市之间的铁路运营里程计算(中国交通图册,1985,第133~134页)。

这样,用于分析迁入地选择影响因素的人口迁移分析矩阵就建立起来了。

3 结果

3.1 迁入地选择影响因素的简单相关分析

我们首先逐项考察各项省际差异指标与省际人口迁移迁入省份选择之间的关系,为此我们进行了各变量之间的相关分析。由表2可知,作为我们研究目标变量的“实际省际迁移”变量与我们选取的多数省际差异指标之间存在统计上显著的相关关系。

“城镇职工人均工资收入差异”、“农民人均纯收入差异”、“人口密度差异”、“15岁以上人口识字率差异”、MI差异及“人均投资差”与“实际省际迁移”变量之间存在统计上高度显著的正相关关系。这表明,一个省作为迁入省份较之另一个作为迁出省份的省在上述指标上高出越多,人口由该迁出省份迁往该迁入省份的可能性越高。

表2 各项省际差异指标、省际空间距离和“实际省际迁移”变量之间的相关系数

变量	实际省际迁移	人均工资差	人均耕地差	农民收入差	人口密度差	识字率差	MI差	空间距离	人均投资差
实际省际迁移	1								
人均工资差	.043 **	1							
人均耕地差	-0.003	-.070 **	1						
农民收入差	.069 **	.614 **	-.042 **	1					
人口密度差	.074 **	.421 **	-.224 **	.756 **	1				
识字率差	.050 **	.200 **	.303 **	.523 **	.289 **	1			
MI差	.047 **	.468 **	.499 **	.731 **	.294 **	.667 **	1		
空间距离	-.129 **	-.001	-.059 **	-.069 **	-.010	-.109 **	-.103 **	1	
人均投资差	.071 **	.700 **	.089 **	.886 **	.804 **	.515 **	.707 **	-.076 **	1

* $p < 0.05$ ** $p < 0.01$

省际“空间距离”与“实际省际迁移”变量之间存在统计上高度显著的负相关关系。这表明,两个省之间在地理位置上越是邻近,人口由其中一个省迁往另一个省的可能性越高;相反,两个省之间的空间距离越远,人口由其中一个省迁往另一个省的可能性越低。

“农村劳动力人均耕地差”与“实际省际迁移”变量之间的相关关系为负。但是,这一关系在统计上并不显著。这表明,我国实际省际迁移的发生概率与“农村劳动力人均耕地占有量的省际差异之间没有必然的联系。

3.2 迁入地选择的多元回归分析

以上是从单个变量之间关系的角度考察了各省际差异指标与“实际省际迁移”变量之间的关系。但是,这些差异指标之间也是相互联系的。比如,“人均固定资产投资省际差异”指标与“农民人均纯收入差异”、“人口密度差异”及“MI省际差异”等指标之间均高度相关,且相关系数分别高达0.886、0.804和0.707。如此,仅仅通过相关分析难以准确把握各省际差异指标与“实际省际迁移”之间的关系。

因此,为了更准确地反映各项省际差异指标对迁入地选择的影响,有必要对迁入地选择进行多元回归分析。在分析中,变量“实际省际迁移”是因变量,而(潜在)迁入省份与迁出省份之间的各项差异指标和空间距离则是自变量。经过试验,在多元回归分析中,我们选取的自变量包括:城镇职工人均工资收入差、农民人均纯收入差异、识字率差异、人均固定资产投资差异、人口密度差异等差异指标和省际空间距

离。

由于因变量“实际省际迁移”是一个虚拟变量，故我们选用 logistic 回归模型来分析因变量与自变量之间的关系。

分析模型如下：

$$\ln [p / (1 - p)] = b_0 + b_1 \times x_i$$

其中， p 代表实际省际迁移发生概率； x_i 代表各省际差异指标和省际空间距离； b_0 和 b_1 是回归方程的系数。

回归结果如表 3 所示。由表中可以看到：

省际“农民人均纯收入差异”对“实际省际迁移”变量具有统计上显著的正影响作用。这表明，一个省作为迁入省份较之另一个作为迁出省份的省的农民人均纯收入高出越多，人口由该迁出省份迁往该迁入省份的可能性越高。

“人口密度差异”对“实际省际迁移”变量具有统计上高度显著的正影响作用。这表明，一个省作为迁入省份较之另一个作为迁出省份的省的人口密度高出越多，人口由该迁出省份迁往该迁入省份的可能性越高。

“15 岁以上人口识字率差异”对“实际省际迁移”变量具有统计上显著的正影响作用。这表明，一个省作为迁入省份较之另一个作为迁出省份的省的 15 岁以上人口识字率高出越多，人口由该迁出省份迁往该迁入省份的可能性越高。

省际“空间距离”对“实际省际迁移”变量具有统计上高度显著的负影响作用。这表明，两个省之间在地理位置上越是邻近，人口由其中一个省迁往另一个省的可能性越高；相反，两个省之间的空间距离越远，人口由其中一个省迁往另一个省的可能性越低。

上述关系与前面“相关分析”中所得结论基本相同。这说明，在控制了其它变量的作用以后，省际“农村人均纯收入差异”、“人口密度差异”、“15 岁以上人口识字率差异”和“空间距离”等因素对省际人口迁移迁入地选择的影响作用仍然保持不变。

表 3 省际人口迁移迁入地的选择分析结果

变量	B	S. E.	Wald	Exp (B)
城镇工资差 (元)	.0005	.0003	3.4716	1.0005
农村纯收入差 (元)	.0008 *	.0003	5.4903	1.0008
识字率差异	.0105 *	.0047	4.9880	1.0105
空间距离 (百公里)	-.0947 **	.0058	266.1147	.9096
人均固定资产投资差 (元)	-.0008 *	.0003	9.6246	.9992
人口密度差	.0008 **	.0002	21.5593	1.0008
截距	-1.8659 **	.0873	457.1963	
案例数:	17192			
卡方:	416.802			

* $p < 0.05$ ** $p < 0.01$

与此不同的是，在控制了其它变量的影响之后，省际“人均固定资产投资差异”对省际人口迁移迁入地选择的影响作用发生了变化，与前面相关分析中所得关系发生了方向上的根本变化。在控制其它因素的影响之后，省际“人均固定资产投资差异”对省际人口迁移迁入地选择成为统计上显著的负影响作用，也就是说，一个省作为迁入省份较之另一个作为迁出省份的省的人均固定资产投资额高出越多，人口由该迁出省份迁往该迁入省份的可能性越低。

省际“人均固定资产投资差异”与“实际省际迁移”之间关系的这种变化，正是进行多元回归分析的必要性所在。在相关分析中，两个变量之间的相关系数中除包含有二者之间的相互关系外，还包含有其他

因素的作用；而在多元回归分析中，其他因素的作用得到了有效的控制，此时，两个变量之间显现出的关系，是它们之间更为直接、更为真实的关系。那么，又如何解释省际“人均固定资产投资差异”对“实际省际迁移”的负影响呢？一般认为，人们的迁移受投资水平的影响。哪里投资多，人们就迁往哪里。但在实际的迁移活动中，由于文化、心理、习惯等的作用，更多的人更乐于就近实现迁移。这种“就近”的迁移，往往不是人均投资额差异“最大化”的选择。

同样，“城镇职工人均工资收入差异”对实际省际迁移迁入地选择的影响也发生了变化。虽然前者对后者具有正影响，但其影响作用在统计上不显著。这表明，省际城镇职工人均工资收入差异对省际人口迁移迁入地的选择没有影响作用。

4 结论

4.1 在理论上，我们可以将完整的人口迁移决策过程分为决定是否迁移和决定向何处迁移两个方面。以往的文獻对于决定是否迁移研究较多，而对于决定向何处迁移则很少涉及。应该说，这两个方面无疑是有联系的，影响这两个方面的因素也会有许多共同点。但它们毕竟是一个过程的两个方面，各自的影响因素及其影响机制也会有所不同。本文通过建立人口迁移分析矩阵，有效地将地区之间的各种差异指标和个人的迁入地选择过程结合在一起，分析了省际社会、经济差异指标和空间距离指标对于个人迁入地选择过程的影响作用。这样一种分析思路，将有助于更好地分析研究影响人口迁移的因素及其影响机制，特别是有关迁入地选择的影响因素及其作用机制。

4.2 1985年7月1日至1990年6月30日间，我们选取的多数表示省际社会、经济差异的指标对我国省际人口迁移的迁入省份选择过程具有显著的影响作用。总起来讲，一个作为潜在迁入省份的省在农民人均纯收入、人口密度、15岁以上人口识字率、MI指标等指标上越是高于一个作为迁出省份的省，则人口由该迁出省份迁往该迁入省份的可能性越高；一个作为潜在迁入省份的省在人均固定资产投资额上越是高于一个作为迁出省份的省，则人口由该迁出省份迁往该迁入省份的可能性越低；两个省份之间的空间距离对两个省之间的人口迁移发生概率起着“障碍”作用，空间距离越远，两省之间的省际迁移发生概率越小。

4.3 多元回归分析的结果是在控制其它因素的影响作用之后，单个自变量对因变量的影响作用，因而它能够更准确地反映变量之间的相互关系和作用机制。因此，凡在理论上能够确认某个目标变量受到多个变量影响的，在分析时应尽可能选择多元回归分析。

参考文献：

- 1 Liang, Zai and Michael J. White. 1997. Market Transition, Government Policies, and Interprovincial Migration in China: 1983~1988. *Economic Development and Cultural Change* (1997): 321~139.
- 2 Oberai, A. S. and Singh, H. K. Manmohan. 1983. *Causes and Consequences of Internal Migration*. Oxford University Press.
- 3 Poston, Dudley L., Jr., and Michael Xinxiang Mao. 1998. Interprovincial Migration in China, 1985 - 1990. *Research in Rural Sociology and Development* 7: 227~250.
- 4 段成荣 a. 中国省际人口迁移研究. 中国人民大学博士学位论文, 2000
- 5 段成荣 b. 影响省际人口迁移的个人特征分析. *人口研究*, 2000; 4
- 6 中国交通图册. 铁路出版社, 1985
- 7 王桂新. 中国人口分布与区域经济发展. 一项人口分布经济学的探索研究. 华东师范大学出版社, 1997
- 8 杨云彦. 中国人口迁移与发展的长期战略. 武汉大学出版社, 1994

(责任编辑：姚 远 收稿日期：2000 - 07)