第 10 卷第 2 期
2011 年 1 月

经 济 学
China Economic Quarterly

Vol. 10, No. 2
January, 2011

工资差异分解方法述评

郭继强　姜　俪　陆利丽*

摘　要　工资差异分解方法发展到今天，可谓均值分解和分布分解共分其天下。均值分解方法呈 Oaxaca-Blinder 分解及其改进、Brown 分解及其改进和 JMP1991 分解“三足鼎立”的局面，而 Oaxaca-Blinder 分解则是最基础也是最经典的一种方法。分布分解方法主要由 JMP1993 分解、DFL 分解、MM2005 分解、FL1998 分解、Lemieux 分解、Q-JMP 分解和 FFL 分解“七剑合璧”，其中前三种方法尤具奠基性。本文系统地梳理了这些分解方法之间的传承联结，以期形成一幅较为完整的工资差异分解方法演进的向导图。

关键词　工资差异，工资均值分解，工资分布分解，反事实工资

一、引　　论

在工资差异分解方法的演进史上，分解工资差异源自对劳动力市场中歧视的度量。Oaxaca 就是从界定歧视系数以度量歧视开始了他的工资差异均值分解。鉴于 Oaxaca (1973) 和 Blinder (1973) 差不多同时对两个组群（诸如男性与女性两个组群，抑或白人与黑人两个组群等。本文为叙述的前后一致性和一般性起见，将所考察的两个组群依据工资水平高低分别记为组群 H 和 L）之间的工资均值差异提出了几乎相同的分解方法，学术界通常将这种分解方法称为 Oaxaca-Blinder 分解（间或简称 Oaxaca 分解）。该分解将组群之间工资均值差异分解为由个体特征差异造成的可解释部分和由特征回报差异带来的不可解释部分，并把不可解释部分归因于歧视。因而，工资差异的均值分解方法常用于测度歧视的大小程度。

---

* 郭继强，浙江大学公共管理学院，浙江大学劳动保障与社会政策研究中心(LEPP)；姜俪、陆利丽，浙江大学公共管理学院。通信作者及地址：郭继强，杭州市浙大路 38 号浙江大学公共管理学院，310027；E-mail：jiqiangguo@163.com。作者感谢两位匿名审稿人中肯到位的评论和建议，但文责自负。作者还感谢国家社科基金重点项目“‘十二五’时期调整国民收入分配结构研究”(10AZD003)、国家社科基金项目“城乡一体化进程中扩大就业的理论和对策研究”(09BJL019)、国家自科基金重点项目“城乡劳动力市场整合机理与实现机制研究”(70933001)、教育部哲学社会科学研究重大课题攻关项目“建立城乡统一的劳动力市场，实现城乡劳动者平等就业研究”(06JZD0014)以及 LEPP 重大项目“我国收入不平等的度量及其矫正机制研究”的资助。

组群工资差异均值分解方法基本上是围绕着“在三个层面上应对两类问题”这一主线展开。“三个层面”是指 Oaxaca-Blinder 分解层面、Brown 分解层面和 JMP1991 分解层面。如果说 Oaxaca-Blinder 分解层面只是关注于“同工不同酬”问题；那么，Brown 分解通过吸纳和整合组群职业分布对工资差异的影响，形成了将职业分隔（occupational segregation）和同工不同酬纳入统一分析框架的新研究层面。前两个层面的均值分解都把工资均值差异的不可解释部分归因于歧视，但 JMP1991 分解却不然，而是把不可解释部分重点归因于不可观测技能差异（更一般地说是组群特异性），使得对不可解释部分成因的考察重心从歧视转变为不可观测技能差异[1]；另一方面，JMP1991 分解将不可解释部分即残差额表达成残差标准差与标准化残差分布差异之乘积，尽管还没有完全利用分布特性进行工资分布分解，但与以往只关注经典线性回归残差均值为零的特性从而忽略残差分布的均值分解方法相比，它转向了关注残差分布，进而推动着工资差异分解从均值视角向分布视角深化。

“两类问题”分别是指数基准问题以及诸如选择性偏差等计量技术问题。所谓指数基准问题（index number problem）是指这样的难题：在工资差异均值分解过程中因所选取的分解基准不同得到相异的分解结果，既无法客观准确地衡量各种因素或特性对工资差异的影响程度，也无法唯一地推断出真实的歧视程度。另一类问题是诸如样本选择问题和虚拟变量系数识别问题等一些计量技术问题。均值分解的样本选择问题是指由于对样本选择偏差纠正项差异的归因存在某种程度的主观性所导致的分解结果的多种可能性。虚拟变量系数识别问题则是指在将工资差异分解到各个单一协变量的效应时，各个虚拟变量的特征效应和系数效应会因基准组选择的不同而变化，从而导致分解结果的模糊性。这两类问题，在上述三个层面中均无法绕开。

客观地说，工资均值只能描述工资分布的集中趋势，反映分布的一个特征，而一个组群的工资状态毕竟是一种分布，特别是在组群的工资分布趋于离散的情况下，更需要对工资分布的不同区域进行针对性的逼近和刻画。此时，能够针对工资分布的不同区域进行解析的分布分解就拥有独特的优势。

在工资差异的分布分解方面，Fortin *et al*.（2010）建立了按照反事实工

[1] 其实，歧视和不可观测技能都会影响不可解释部分，但 JMP1991 分解方法仍无法识别歧视。当然，如何将不可观测技能与歧视进一步在理论和实证上区别开来，仍有待继续探索。

资分布的构造路径归纳工资差异分布分解方法的综述框架。[2] 笔者以为，Fortin 等的综述框架无疑具有高度概括性，但在此框架下却难以展现分解方法之间的传承联结。而本文则采取了不同的梳理思路，以分解方法的演进脉络为主线，同时考虑基本模型设定和反事实工资分布构造的差别，以递进和拓展的思路为读者提供一幅工资分解方法演进的向导图。因此，本文或许更适合那些对分布分解方法不甚熟悉却又感兴趣且期望能够较快把握工资差异分解方法的读者。

在笔者看来，迄今为止的工资差异分布分解方法的特点可大体上归结为"四类基本模型三种反事实分布"。具体说来，JMP1993 分解是工资差异分布分解的一个里程碑。JMP1993 分解基于经典线性回归模型（OLS 估计），通过将残差分布中的分位与个体在不可观测技能分布的分位相联系，经由残差分布变换来构建反事实工资分布，将同一组群工资分布变动分解为可观测特征变动、可观测特征回报率变动和不可观测技能变动。MM2005 分解利用条件分位回归模型，通过概率积分转换（probability integral transformation）构建反事实工资分布，解决了 JMP1993 分解中的异方差问题。不管是基于经典线性回归模型还是条件分位回归模型的分解方法，都无法避免模型参数线性设定的假定，而 DFL 分解基于半参模型（核密度估计），通过重置权重函数的思想构建反事实工资分布，突破了模型参数线性设定的局限。不过，前三个模型都无法将工资分布变动自然地分解至各个单一协变量特征效应，而 FFL 分解则基于 RIF 回归模型，利用 DFL 分解中重置权重函数构建反事实工资分布的方法，将工资分布变动成功地分解至单一协变量效应。此外，FL1998 分解、Q-JMP 分解以及 Lemieux 分解还从不同角度对工资分布分解进行了变异和拓展。

概言之，工资差异分布分解的方法基本上可以归纳成以下七种：(1) 基于经典线性回归的分布分解，包括 JMP1993 分解和 FL1998 分解；(2) 基于半参模型的分布分解，包括 DFL 分解和 Lemieux 分解；(3) 基于条件分位回归的分布分解，包括 MM2005 分解和 Q-JMP 分解；(4) 基于 RIF 回归的分布分解，即 FFL 分解。

---

[2] Fortin *et al*. (2010) 归纳了构造反事实工资分布的三条路径：

$$F(\ln w; t_w = 1, t_X = 0) = \int F(\ln w \mid X, t_w = 1) \mathrm{d}F(X \mid t_X = 0).$$

第一条路径是直接从变换工资值着手，将 $t=0$ 时期个体的工资值替换为反事实工资值，从而形成反事实工资分布。JMP1993、MM2005 和 Q-JMP 分解都是遵循这一路径构造了反事实工资分布。

第二条路径是从变换特征分布着手，将反事实工资分布作为 $t=1$ 时期工资分布的重置加权形式，即：

$$F(\ln w; t_w = 1, t_X = 0) = \int F(\ln w \mid X, t_w = 1) \psi_X(X) \mathrm{d}F(X \mid t_X = 1),$$

其中，$\psi_X(X) = \mathrm{d}F(X \mid t_X = 0) / \mathrm{d}F(X \mid t_X = 1)$。通过该路径构造反事实工资分布的方法有 DFL 分解、Lemieux 分解和 FFL 分解。

第三条路径是从变换工资条件分布着手，将 $t=0$ 时期的条件工资分布 $F(\ln w \mid X, t_w = 0)$ 替换为 $F(\ln w \mid X, t_w = 1)$。FL1998 分解、Donald *et al*. (2000) 和 Chernozhukov *et al*. (2009) 都通过该路径构造了反事实分布。

鉴于工资差异分解方法中均值分解和分布分解的区块特征非常明显，本文述评的结构安排自然要充分考虑这种状况。据此，本文的第二部分将介绍“工资差异均值分解”；第三部分考察“工资差异分布分解”；最后部分则是“评论性小结”。通过系统地梳理各种分解方法的理论基础、假设条件和适用范围，阐释各种分解方法之间的传承联结，扼要评价它们的学术贡献，本文希冀能够形成一幅较为完整准确的工资分解方法向导图。

二、工资差异均值分解

工资差异均值分解部分介绍四项内容：Oaxaca-Blinder 分解的形成与改进、Brown 分解及其修正、JMP1991 分解以及对均值分解中若干计量问题的矫正。Oaxaca-Blinder 分解是在承接 Becker（1957）市场歧视系数之意蕴基础上的计量展开。由于 Oaxaca-Blinder 分解存在指数基准问题，Cotton（1988）、Neumark（1988）以及郭继强和陆利丽（2009）就是通过优化无歧视工资结构来改进工资差异的均值分解。在 Brown 分解层面上，本文主要介绍 Brown 分解和 Appleton 分解，后者通过矫正双重指数基准问题来深化改进 Brown 分解。JMP1991 分解则是指 Juhn *et al*.（1991）在均值分解中引入分布工具的一种分解方法。在计量问题的矫正方面，本文着重讨论样本选择和虚拟变量系数识别问题。

（一）Oaxaca-Blinder 分解的形成与改进

1. Oaxaca-Blinder 分解的形成

Oaxaca（1973）在其经典论文“Male-Female Wage Differentials in Urban Labor Markets”中指出，以往对男女工资差异的研究多是描述性的，而该篇论文的目的则是要从数量上估计美国对女性工人歧视的平均程度以及各因素对男女工资差异的影响程度。如果记组群 H 和 L 在劳动力市场上的均衡工资分别为 w_H 和 w_L，这两个组群分别作为子样本的个体特征（个体禀赋）矩阵各为 X_H 和 X_L，相应的回归系数向量（或称工资结构）分别为 β_H 和 β_L，这两个组群的半对数形式的工资估计方程（通常以 Mincer 工资决定方程为基础）分别是 $\ln w_H = X_H\beta_H + u_H$，$\ln w_L = X_L\beta_L + u_L$；又记这两个组群的子样本个体特征向量的平均值分别为 $\overline{X}_H$ 和 $\overline{X}_L$，那么，根据最小二乘法（OLS）残差均值为零的性质，这两个组群的工资均值之差可表述成[3]：

$$\ln\overline{w}_H - \ln\overline{w}_L = \overline{X}_H\beta_H - \overline{X}_L\beta_L. \tag{2-1}$$

Oaxaca 承接 Becker（1957）所定义的市场歧视系数的基本内核并进行一

[3] 此处的 X 是 $n\times k$ 的矩阵，$\overline{X}$ 则是行向量；β 实际上是估计值 $\hat{\beta}$，本文为了表述的一致和简洁起见，将相关回归方程的估计参数中的“ˆ”（hat 符）去掉。同时，本文还对相关文献中同一概念或指标用不同符号形式的表达作了统一处理。

定的变形后提出了如下形式的歧视系数（discrimination coefficient，D）：$D=\frac{w_H/w_L-(w_H/w_L)^0}{(w_H/w_L)^0}$，其中，$(w_H/w_L)^0$表示无歧视时的均衡工资比。被观察到的现实工资往往是存在歧视时的工资，而无歧视时的工资则需要我们根据既有实际样本的信息加以估计和逼近。[4]无歧视工资也是我们进行工资差异均值分解和推断歧视程度的逻辑基点。若将 Oaxaca 对歧视的界定写成对数形式则有[5]：$\ln w_H-\ln w_L=\ln(w_H/w_L)^0+\ln(D+1)$。Oaxaca 还分别用组群 H 或 L 的实际观察到的工资充作无歧视时的劳动力市场工资，由此产生以下两种工资均值差异分解情形：

情形 1　将组群 H 的实际工资结构当做无歧视的劳动力市场工资结构时，无歧视状态下组群 H 和 L 均衡工资比的对数为 $\ln(w_H/w_L)^0=(\overline{X}_H-\overline{X}_L)\beta_H$。从而，(2-1) 式可分解为

$$\ln\overline{w}_H-\ln\overline{w}_L=(\overline{X}_H-\overline{X}_L)\beta_H+\overline{X}_L(\beta_H-\beta_L). \tag{2-2}$$

情形 2　将组群 L 的实际工资结构当做无歧视的劳动力市场工资结构时，组群 H 和 L 无歧视状态下的均衡工资比的对数为 $\ln(w_H/w_L)^0=(\overline{X}_H-\overline{X}_L)\beta_L$。从而，(2-1) 式可分解为

$$\ln\overline{w}_H-\ln\overline{w}_L=(\overline{X}_H-\overline{X}_L)\beta_L+\overline{X}_H(\beta_H-\beta_L). \tag{2-3}$$

不论是等式 (2-2) 还是等式 (2-3)，等式右边的第一项均表示即便不存在歧视组群 H 和 L 之间也存在的工资差异，亦即由组群 H 和 L 之间的个体特征差别引起的工资差异；第二项则是由组群 H 和 L 之间的工资结构差别引起的工资差异，或者说是存在歧视和不存在歧视两种状态下组群 H 和 L 之间工资差异之差额，Oaxaca 称之为歧视（歧视效应或歧视引起的工资差异）。显然，Oaxaca 将不能解释的工资差异部分都归因于歧视。[6]

比较等式 (2-2) 和 (2-3)，我们可以发现，选取组群 H 的实际工资结构 (β_H) 或选取组群 L 的实际工资结构 (β_L) 作为无歧视的劳动力市场工资结构，会得到不一致的分解结果，从而产生指数基准问题。[7]Oaxaca 已意识到

---

[4] 因此之故，无歧视状态也时常被称为反事实状态，该状态下的工资则是反事实工资。鉴于工资差异均值分解源自对歧视度量的关注，因此本文第二部分将会一直沿用无歧视状态的表达。

[5] $\ln(D+1)=\ln(w_H/w_L)-\ln(w_H/w_L)^0=\ln w_H-\ln w_L-\ln(w_H/w_L)^0$。

[6] 正是从这种意义上，Oaxaca 有时也将自己提出的分解方法称为残差法(the residual approach)。

[7] 例如，Ferber and Green(1982)在分解大学教授工资差异时发现，在一种基准下歧视只占工资差异的2%，而在另一种基准下却高达70%。又譬如，Cotton(1988)在研究白人与黑人工资差异时则表明，用两种不同的基准分解时歧视分别占工资差异的48.5%和95.4%。Démurger *et al*.(2007)曾提出，根据不同指数基准(男性和女性)得到的不同分解结果可以作为一个稳健性检验。笔者在一定程度和范围内赞同这种观点，但更在意指数基准选择的客观性，因为在不同的数据样本中，用不同的指数基准对组群工资差异分解的结果，既可能相当接近，也可能大相径庭；在结果出现大相径庭的情况下，所谓的稳健性就不能帮助我们判断了，我们只能转而信赖以更逼近“真正的”无歧视工资结构的那个指数基准所分解的结果。

了这一点，但囿于无法合理地认定到底应以哪个组群为基准，只能提出将分别按照这两个不同基准分解得到的估计结果作为度量歧视的一个可能的取值区间。

Oaxaca（1973）还指出了他所提出的工资分解方法仍存在一些难点：一是在无歧视状态下两个组群的工资结构是否会相同的问题；二是这种分解方法并没有考虑到两个组群的可选择性变动的个体特征对劳动力市场歧视的反馈效应。对于第一个难点，Butler（1982）认为即使在无歧视状态下两个组群的工资结构系数也会存在差别；而 Cotton（1988）则批评道，没有歧视时我们不能指望两个组群的工资结构系数差别会长期持续。至于第二个难点，则在某种程度上涉及变量的内生性问题。

在 Oaxaca 提出工资差异的上述分解的差不多同时，Blinder（1973）也对工资差异进行了类似形式的分解，故学术界通常将工资差异的上述经典分解方法合称为 Oaxaca-Blinder 分解。不过，Blinder 还着重考察了使用简约型和结构工资方程（reduced form and structural wage equations）对工资差异分解的影响，并在推导分解方法时突出了工资回归方程的常数项应包含在歧视之中。与 Oaxaca 的分析相比，Blinder 更加重视工资决定因素即变量的内生性问题。为此，Blinder 提出一个估计工资的结构方程组。但因这个结构方程组是不可识别的（underidentified），他转而采用两种次优方法（two second-best procedures）进行近似估计[8]：一种方法是使用简约型工资方程，二是在一定的假设条件下对结构方程组进行分块递归（block recursive）。显然，Blinder 敏感地触及了变量的内生性问题并在一定程度上力图加以处理。[9]

对于 Oaxaca-Blinder 分解中出现的指数基准问题，笔者认为，倘若歧视确实存在，两个组群中任一组群的实际工资结构都不能作为无歧视状态下的工资结构，原因在于任何一个组群的现实子样本工资回归方程的估计系数都已包含着歧视所造成的影响。如果能够寻找到“真正”的无歧视工资结构（记为 β^*），那么，组群间工资均值差异一般地可表述成如下分解式[10]：

$$\ln\overline{w}_H - \ln\overline{w}_L = (\overline{X}_H - \overline{X}_L)\beta^* + \overline{X}_H(\beta_H - \beta^*) + \overline{X}_L(\beta^* - \beta_L), \quad (2\text{-}4)$$

其中，$(\overline{X}_H-\overline{X}_L)\beta^*$ 是由组群间个体特征差异引起的工资差异，组群间工资差

[8] 大致地说，简约型工资方程估计的是在给定个人出生环境下的工资（或工资对数）条件期望，而结构工资方程估计的则是给定个人当前社会经济条件下的工资（或工资对数）条件期望。

[9] 后人普遍使用的仍是 Mincer 工资方程，至于内生性问题则多用其他方法譬如工具变量法等加以解决。况且，在笔者看来，强行划分简约型和结构工资方程，可能出现模型误设或遗漏变量等问题。

[10] Cotton（1988）和 Neumark（1988）都表达过相同或类似的分解形式。不难看出，Oaxaca-Blinder 分解是一般分解式（2-4）式的特例。当 $\beta^* = \beta_H$ 时，（2-4）式退化为（2-2）式；而当 $\beta^* = \beta_L$ 时，（2-4）式则简化为（2-3）式。

异的其余部分即 $\overline{X}_H(\beta_H-\beta^*)+\overline{X}_L(\beta^*-\beta_L)$ 一般归因于歧视。可以说，Reimers（1983）、Cotton（1988）、Neumark（1988）以及郭继强和陆利丽（2009）都是致力于寻找更具解释力的单一的无歧视工资结构（为表述简便，本文将其分别记为 β_R^*、β_C^*、β_N^* 和 β_G^*）。

2. Cotton 分解

Reimers（1983）率先将无歧视工资结构看做两个组群各自工资结构的加权平均，构建了一个宽泛的表达式：$\beta_R^*=D\beta_H+(I-D)\beta_L$。其中，$\beta_R^*$ 是 Reimers 意义上的无歧视工资结构，I 是一个单位矩阵，D 为一个任意对角矩阵，而 β_H 和 β_L 则是组群 H 和 L 各自的回归系数。由于对角矩阵的多种可能性，这一表达式也就蕴涵着无歧视工资结构选择的多重性和任意性。实际上，Reimers 在经验分析中选取了 $D=\mathrm{diag}(0.5,0.5,\cdots,0.5)=(0.5)I$ 作为对角矩阵。不过，Reimers 并没有细述为何将取值 0.5 的对角矩阵作为计算无歧视工资结构的加权矩阵。

Cotton（1988）是从分析歧视对两个组群的可能影响出发来寻求无歧视工资结构的。Cotton 指出，Oaxaca（1973）不恰当地描绘了 Becker 的最重要初始条件即无歧视状态下的工资结构，因为无论是以组群 H 还是以组群 L 为分解基准，劳动力市场歧视都被视为单向行为，要么单纯地歧视，要么单纯地偏爱另一组群。在 Cotton 看来，在无歧视状态下组群 H 和 L 的工资结构都将收敛于同一个无歧视工资结构（记为 β_C^*），从而，歧视将同时呈现双向行为：既歧视一个组群（本文为行文一致起见通常谓之直接歧视），而又偏爱另一个组群（称为反向歧视、偏爱性歧视、偏袒或逆歧视）。[11] 相应的，Cotton 分解可表述成

$$\ln\overline{w}_H-\ln\overline{w}_L=(\overline{X}_H-\overline{X}_L)\beta_C^*+\overline{X}_H(\beta_H-\beta_C^*)+\overline{X}_L(\beta_C^*-\beta_L). \quad (2\text{-}5)$$

公式（2-5）右边第一项是指由生产率特征或者说技能差异引起的工资差异。右边第二和第三项之和是总歧视，其中第二项表示组群 H 受偏爱或者说被优待（treatment advantage），其劳动力价值被高估（overvalued）所形成的反向歧视；第三项则表示组群 L 遭歧视或者说被不利对待（treatment disadvantage），其劳动力价值被低估（undervalued）所形成的直接歧视。

至于无歧视工资结构 β_C^*，Cotton 是在一些相当强的假定下获得的。这些假定包括：(1) 在无歧视状态下组群 H 得到的平均工资要低于现实中所得到的，而组群 L 则要高于现实中所得到的。(2) 无歧视工资结构将是那些影响组群 H 和 L 现行工资结构因素的函数，更简单的，假定 β_C^* 是 β_H 和 β_L 的线性函数。(3) 无歧视工资结构偏向于在劳动力市场中人数所占比重较高组群的实

[11] 在 Cotton 分析中具体是指白人与黑人两个组群。

际工资结构。因此，Cotton 的无歧视工资结构 β_C^* 可定义为

$$\beta_C^* = f_H\beta_H + f_L\beta_L, \tag{2-6}$$

f_H 和 f_L 分别代表组群 H 和 L 人数占就业总人数的比重。由于 $f_H + f_L = 1$，从而公式（2-6）也可以写成 $\beta_C^* = f_H\beta_H + (1 - f_H)\beta_L$。若令对角矩阵 $D = \mathrm{diag}(f_H, f_H, \cdots, f_H) = f_H(I)$，则公式（2-6）还可以进一步表述成

$$\beta_C^* = D\beta_H + (I - D)\beta_L. \tag{2-7}$$

显然，较之于 Reimers（1983）实际使用的 $D = 0.5(I)$（这其实可以看成 $f_H = 0.5$ 时的情形），Cotton 的无歧视工资结构 β_C^* 更具逻辑合理性，故 Cotton 分解也就为更多人所采用。[12]

与 Oaxaca-Blinder 分解相比，Cotton 分解消解了 Oaxaca-Blinder 分解中存在的指数基准问题，并将总歧视进一步细分为反向歧视和直接歧视，因而是对 Oaxaca-Blinder 分解的一种改进。不过，需要指出的是，这仍是一种一定程度上的改进，因为倘若我们把 Oaxaca-Blinder 分解中分别用组群 H 或 L 作为基准时所得到的歧视量视为歧视的上界和下界，则 Cotton 分解所得到的歧视将介于这两个上下界范围之间，即

$$f_L\bar{X}_H(\beta_H - \beta_L) + f_H\bar{X}_L(\beta_H - \beta_L) \in (\bar{X}_L(\beta_H - \beta_L), \bar{X}_H(\beta_H - \beta_L)). \tag{2-8}$$

但实际上，无论是 Oaxaca、Reimers、Cotton 抑或其他研究者[13]，都未能从理论或逻辑上阐明为何所估计的歧视量必定落在这个上下界范围之内。而这一点也为 Neumark（1988）所诟病。

3. Neumark 分解

就在 Cotton（1988）对 Oaxaca-Blinder 分解改进的几乎同时，Neumark（1988）也提出了不同的改进方法。Neumark 将两个组群间工资歧视的实证估计与雇主歧视行为的理论模型相联系，认为运用于组群工资差异分解的无歧视工资结构（no-discrimination wage structure）应从一个歧视行为理论模型中推导出来。这其实是在构建无歧视工资结构的同时力图夯实理论基础。

Neumark 的雇主歧视行为理论模型是对 Arrows（1972）和 Becker（1957）雇主歧视模型的拓展，并且沟通了雇主歧视偏好与无歧视工资结构之

[12] Cotton 认为，他提出的无歧视工资结构隐含着在无歧视状态下不但实际总产出不变而且总工资也不变，受到影响只不过是收入和工作的分布而已。也就是说，歧视只会影响工资在两个组群之间的分配。因此，不能指望消除歧视的同时也就能解决劳动力使用效率不高的问题。

[13] Oaxaca(1973)认为，以组群 H 或 L 作为基准时分别所得到的歧视估计量构成了一个可能取值的区间。Oaxaca and Ransom(1994)也说过，真正的无歧视工资结构会在依照两个指数基准估算出来的歧视大小所构成范围内的某个点上。

间的连接。与 Cotton（1988）相似，Neumark 重点考察了雇主对两个组群既有偏袒（nepotism）又有歧视（discrimination）的一般情形。Neumark 对雇主歧视行为提出了两个基本假定[14]：（1）在表征雇主歧视偏好的效用函数中，每类劳动技能内部的各个组群的劳动投入都是零次齐次的；（2）在同一劳动技能类型中的劳动者都有相同的个体特征，同一劳动类型中同一组群的劳动者有相同的估计工资。Neumark 再借助一些近似处理推导出，全样本工资回归方程中的系数矩阵就是无歧视工资结构即[15]：

$$\beta_N^* = (X'\Omega X)^{-1}(X'\Omega\Lambda) = (X'X)^{-1}X'Y, \tag{2-9}$$

这里 Y 表示工资的对数即 $\ln w$；X 为包含组群 H 和 L 全部样本（pooled sample）的可观测特征矩阵。[16]全样本回归系数向量经过适当变形，还可以表述成两个组群子样本回归系数向量的矩阵加权模式（Oaxaca and Ransom，1994）：

$$\beta_N^* = \Omega_N\beta_H + [I - \Omega_N]\beta_L, \tag{2-10}$$

其中，$\Omega_N = (X'X)^{-1}X_H'X_H$。相应的，Neumark 分解可用下式表示：

$$\ln\overline{w}_H - \ln\overline{w}_L = (\overline{X}_H - \overline{X}_L)\beta_N^* + \overline{X}_H(\beta_H - \beta_N^*) + \overline{X}_L(\beta_N^* - \beta_L). \tag{2-11}$$

从形式上看，β_N^* 和 β_C^* 都是两个组群各自工资结构（回归系数向量 β_H 和 β_L）的加权平均，但实质上两者仍存在很大的差别。首先，两者的构建基础不同。β_C^* 更多来自经验和直觉，而 β_N^* 则源于雇主歧视行为理论模型。在 Neumark 看来，Oaxaca（1973）虽然把对无歧视工资结构的选择描述成著名的“指数基准问题”，但这种选择实际上取决于歧视行为的性质。特别是 Neumark 通过若干假定和利用最小二乘法原理还发现，由全样本计算所得的回归系数向量最接近无歧视工资结构，由此建立了雇主歧视偏好的形式与工资歧视估计结果之间的关系。其次，两者的权数不同。β_C^* 的权数是 $D = f_H(I)$，并且 $f_H \in (0,1)$；而 β_N^* 的权数则是 $\Omega_N = (X'X)^{-1}X_H'X_H$。显然，$\beta_N^*$ 并不是简单线性加权，而是矩阵加权，并且 Ω_N 的正负和大小取决于所探讨的样本具体特性。再次，两者的信息利用程度不同。$D = f_H(I)$ 利用的只是组群 H 和 L 两个子样本的劳动者数量的信息，而 Ω_N 则涵盖了更多的样本信息，包括由组群

[14] 更一般地说，Neumark(1988)的假定还包括：劳动供给和个人特性是固定的，并不对因消除歧视而产生的工资变动作出反应。

[15] 公式(2-9)中第一个等号表示使用分组数据时，用加权最小二乘法得到的全样本回归系数的表示方法；第二个等号表示使用个体的非分组数据时，用最小二乘法得到的全样本回归系数表示方法。

[16] 其中，$\Omega = \mathrm{diag}(M_1+F_1, M_2+F_2, \cdots, M_J+F_J)$，$M_j$ 和 F_j 是第 j（$j=1,2,\cdots,J$）类技能中男性劳动者和女性劳动者的数量；Λ 则是各类劳动技能的无歧视工资所组成的一个向量。第 j 类劳动技能的无歧视工资 $\Lambda_j = \dfrac{\ln w_{Mj}M_j + \ln w_{Fj}F_j}{M_j + F_j}$。$\ln w_{Mj}$ 和 $\ln w_{Fj}$ 分别表示第 j 类技能中男性劳动者和女性劳动者的工资。从这个无歧视工资公式可以看出，歧视对拥有某类技能（此处为第 j 类技能）的劳动者的工资总额并无影响，或者说，歧视只影响工资分布而不影响总额。

H 和 L 两个子样本所组成的总样本的信息、组群 H 和 L 的结构信息和个体特征分布信息等，而提高样本信息的利用率通常有助于提高理论和实证分析的可靠性。最后，两者所估计的歧视大小和范围不同。依据 β_C^* 的 Cotton 分解所得到的歧视量总是介于 Oaxaca-Blinder 分解所构成的歧视上下界范围之内。但是，根据 β_N^* 的 Neumark 分解的歧视量则为$[\overline{X}_H(X'X)^{-1}X_L'X_L+\overline{X}_L(X'X)^{-1}X_H'X_H](\beta_H-\beta_L)$，它却并不一定落在 Oaxaca-Blinder 分解所构成的歧视上下界范围之内。换言之，撇开所考察的具体样本的特性，我们就无法判定由 Oaxaca-Blinder 分解所构成的歧视上下界范围能否覆盖到所估计的歧视量。[17] 总之，正如 Oaxaca and Ransom（1994）已表达过的那样，Neumark 分解要比 Cotton 分解略胜一筹。

尽管 Neumark 分解具有较为坚实的理论基础，而且作为一种对 Oaxaca-Blinder 分解较好改进的方法而屡屡被人论及和使用（Oaxaca and Ransom，1994；Neuman and Oaxaca，2004；Appleton *et al*.，1999；葛玉好，2007），但郭继强和陆利丽（2009）认为，Neumark 推导出无歧视工资结构的假定相对严格，从另一视角反而给 Neumark 分解的普适性打了折扣，造成了 Neumark 分解的局限性。Neumark 关于雇主歧视行为的基本假定一，意味着在某一类劳动技能内部，雇主关心的只是组群 H 和 L 的劳动者相对比例而非绝对数，影响雇主效用的是组群 H 和 L 的分布。而现实劳动力市场能否满足假定就很值得怀疑，Oaxaca and Ransom（1994）以及 Appleton *et al*.（1999）也都对这个假定提出过质疑。而 Neumark 的另一个假定即“劳动供给和个人特性是固定的，并不对因消除歧视而产生的工资变动作出反应”假定，则表明了歧视并不会影响劳动者的技能选择。对此，Butler（1982）和 Reimers（1983）早已作过较为深入的结论相左的阐述，认为歧视会影响劳动者的技能选择。这就是说，歧视会影响不同技能类型中组群 H 和 L 的劳动者的构成，并导致 Neumark 全样本回归方程存在内生性问题，进而影响无歧视工资结构。[18] 总之，如果我们承认现实劳动力市场可以划分为 H 和 L 两个组群，它们之间的工资结构存在着系统性差别，而在全样本的工资回归方程中却又对组群因素不加区别和控制，那么，全样本估计系数就会与“真正的”无歧视状态下的工资结构产生系统性偏差。这也是郭继强和陆利丽（2009）分解的锲入点。

[17] Neumark(1988)的实证结果表明，用全样本估计的歧视造成的工资差异解释了工资总差异的57%，而 Oaxaca-Blinder 分解分别采用 H 和 L 组群的工资结构作为无歧视工资结构，得到的结果分别为69%和70%。显而易见，此时 Neumark 分解所得到的歧视量落在了由 Oaxaca-Blinder 分解所构成的歧视上下界范围之外。

[18] 歧视不仅来自于雇主歧视，雇员歧视和顾客歧视也可能影响到雇员的技能选择以及雇主对雇员支付的工资报酬大小。此外，单一地从雇主行为出发，在某种程度上还可以说只考虑到劳动力的需求方，而忽视了劳动力供给方与歧视的关联。

4. 郭继强和陆利丽（2009）分解

尽管 Neumark 的分析假定仍有待商榷，但客观地说，他的理论推导结果毕竟揭示了无歧视工资结构与全样本工资方程之间的某种关联。记全样本工资回归方程的表述式为

$$\ln w = \beta_{0N} + X_1\beta_{1N} + \varepsilon, \tag{2-12}$$

式中 β_{0N} 为常数项，β_{1N} 表示除常数项以外的所有控制变量的回归系数，X_1 是包含组群 H 和 L 的全部样本的个体可观测特征矩阵（不包括常数项）；$\ln w$ 是工资对数。（2-12）式就是 Neumark 的无歧视工资结构估计方程，相应的无歧视工资结构可记为

$$\beta_N^* = \begin{bmatrix} \beta_{0N} \\ \beta_{1N} \end{bmatrix}. \tag{2-13}$$

郭继强和陆利丽（2009）也力图从全样本视角挖掘和逼近"真正的"无歧视工资结构。改进之处是在工资方程（2-12）式中添加一个反映组群因素的虚拟变量（dummy variable）G，从而将工资回归方程修正为

$$\ln w = \beta_0 + \delta G + X_1\beta_1 + u, \tag{2-14}$$

其中，当样本属于组群 H 时，令 $G=1$；而当样本属于组群 L 时，则记 $G=0$。由于控制了组群因素对工资差异的影响，从而缓解了回归系数估算误差，提高了无歧视工资结构估算结果的一致性。较之于用方程（2-12）式估计出来的工资结构（β_{0N} 和 β_{1N}），用方程（2-14）式估计出来的工资结构（β_0 和 β_1）进一步向"真正的"无歧视工资结构逼近。[19]

一般说来，某一组群劳动者在劳动力市场中所占的比例愈高，无歧视工资就会愈偏向于该组群。这也符合 Becker 有效歧视理论中数量上和经济上的多数原则。因此，郭继强将新的无歧视工资结构推断为

$$\beta_G^* = \begin{bmatrix} \beta_0 + f_H\delta \\ \beta_1 \end{bmatrix}, \tag{2-15}$$

f_H 的含义同前亦即代表组群 H 的劳动者人数占总就业人数的比例。相应的，郭继强和陆利丽（2009）的分解表达式为

$$\ln\overline{w}_H - \ln\overline{w}_L = (\overline{X}_H - \overline{X}_L)\beta_G^* + \overline{X}_H(\beta_H - \beta_G^*) + \overline{X}_L(\beta_G^* - \beta_L). \tag{2-16}$$

---

[19] Oaxaca and Ransom(1994)曾从 Neumark 的雇主歧视理论出发，提出当雇主效用函数只与 H 组群和 L 组群的人数有关，而与技能水平无关时，可以通过在全样本工资回归方程中加入组群虚拟变量来度量歧视(具体地说，是用组群虚拟变量的系数与组群工资总差异之比来表示歧视带来的工资差异百分比)。但在郭继强(2009)看来，这样度量的工资差异只不过是该虚拟变量的系数即方程(2-14)式中的 δ，并没有考虑到由于个体特征系数差异所造成的歧视。实际上，郭继强和陆利丽(2009)加入 δ 的目的在于推断无歧视工资结构，而非 δ 本身。

郭继强和陆利丽提出的无歧视工资结构具有以下三个特性或命题：

命题1 无歧视工资结构 β_G^* 中常数项大小并不影响工资差异分解中所得到的总歧视大小的客观性。因而，构造无歧视工资结构 β_G^* 时可能隐含的主观性并不影响歧视估算的客观性。

命题2 无歧视工资结构 β_G^* 是两个组群实际观察到的工资结构的函数，但无歧视工资结构是否处于两个组群实际观察到的工资结构之间，以及估算的歧视是否介于由 Oaxaca-Blinder 分解分别根据两个组群实际观察到的工资结构计算的歧视所构成范围之内，则取决于样本自身的特性。

命题3 尽管郭继强和陆利丽（2009）分解与 Neumark 分解的结果都可以表达出来，但两种方法的分解结果的偏差方向却没有稳定的规定性，而是取决于具体样本的数据特性。

总而言之，鉴于 Neumark 对无歧视状态下工资结构的估计存在偏误，郭继强和陆利丽（2009）分解的改进是在估计无歧视工资结构时控制住了组群差别。当然，这只是对工资差异的均值分解方法进行了某种程度的改进，而且这种改进还限于一定的范围之内尤其是囿于参数模型或者线性模型范围之内。

（二）Brown 分解及其修正

1. Brown 分解

歧视可能通过职业分隔途径影响工资差异，但在1980年之前对工资差异的研究要么着眼于同工不同酬，要么聚焦于职业分隔（occupational segregation），而几乎没有考虑把这两方面有机结合的研究。从同工不同酬视角考察的 Oaxaca（1973）和 Blinder（1973），尽管考虑到了“职业”因素对组群工资差异的影响，但由于是把“职业”当做外生的虚拟变量加以控制，从而无法考察某一组群进入某个职业本身就可能存在的歧视。也就是说，将“职业”作为外生控制变量的做法，实际上就会把进入某种职业所可能出现的歧视也归结于可解释部分，造成所估计的歧视程度小于实际的歧视程度，低估了组群间工资的实际歧视程度。[20] 另一方面，既有的从职业分隔视角对歧视影响工资差异的机理研究大多是局部分析，即便估算职业分布大多也只是基于职业选择（occupational choice）视角（Sanborn，1964；Zellner，1972；Stevenson，1975）。而 Brown *et al*.（1980）则强调了组群之职业分隔对工资差异的

[20] Brown *et al*.(1980)已指出，有些研究通过在收入回归方程中加入表征职业的虚拟变量来处理，隐含着将男女间职业分布差异视为理所当然的，忽略了这些差异的潜在的歧视性质。

影响，从职业获得（occupational attainment）角度估计了职业分布，解析出了因职业进入的组别（比如性别）障碍而产生的歧视；同时，Brown 还将个人禀赋、职业分布、歧视与工资差异整合在一个统一的研究框架内，所形成的工资差异分解方法通常称为 Brown 分解。

Brown 首先将组群 H 和 L 的平均工资变形为[21]

$$\ln\overline{w}_H = \sum_j P_{jH}\ln\overline{w}_{jH}, \tag{2-17}$$

$$\ln\overline{w}_L = \sum_j P_{jL}\ln\overline{w}_{jL}, \tag{2-18}$$

其中，P_{jH}、P_{jL} 分别是组群 H 和 L 在第 j 种职业中的实际就业概率或比率，$\ln\overline{w}_{jH}$、$\ln\overline{w}_{jL}$ 分别是组群 H 和 L 在第 j 种职业中的平均工资，$\ln\overline{w}_H$、$\ln\overline{w}_L$ 是组群 H 和 L 的平均工资，$j=1,2,\cdots,J$ 表示社会经济中的 J 种职业。

其次，Brown 以 β_{jH} 作为指数基准，表示第 j 种职业的无歧视工资结构，并利用 $\ln\overline{w}_{jH}=\overline{x}_{jH}\beta_{jH}$、$\ln\overline{w}_{jL}=\overline{x}_{jL}\beta_{jL}$ 对工资差异进行分解[22]：

$$\begin{aligned}\ln\overline{w}_H - \ln\overline{w}_L &= \sum_j P_{jH}\overline{x}_{jH}\beta_{jH} - \sum_j P_{jL}\overline{x}_{jL}\beta_{jL} \\ &= \sum_j P_{jH}\overline{x}_{jH}\beta_{jH} - \sum_j P_{jL}\overline{x}_{jL}\beta_{jH} + \sum_j P_{jL}\overline{x}_{jL}\beta_{jH} - \sum_j P_{jL}\overline{x}_{jL}\beta_{jL} \\ &= \sum_j (P_{jH}\overline{x}_{jH} - P_{jL}\overline{x}_{jL})\beta_{jH} + \sum_j P_{jL}\overline{x}_{jL}(\beta_{jH} - \beta_{jL}).\end{aligned} \tag{2-19}$$

接着，Brown 强调了组群之职业分隔对工资差异的影响，并从职业获得（occupational attainment）角度估计了职业分布。

最后，Brown 将 $\hat{P}_{jL}$ 定义为假如样本 L 组群面临与 H 组群相同职业结构时处于 j 职业的比率，并以 $\hat{P}_{jL}$ 为职业（获得）结构基准进一步分解得

$$\begin{aligned}\ln\overline{w}_H - \ln\overline{w}_L &= \sum_j (P_{jH}\overline{x}_{jH} - \hat{P}_{jL}\overline{x}_{jH} + \hat{P}_{jL}\overline{x}_{jH} - P_{jL}\overline{x}_{jH} + P_{jL}\overline{x}_{jH} - P_{jL}\overline{x}_{jL})\beta_{jH} \\ &\quad + \sum_j P_{jL}\overline{x}_{jL}(\beta_{jH} - \beta_{jL}) \\ &= \sum_j P_{jL}(\overline{x}_{jH} - \overline{x}_{jL})\beta_{jH} + \sum_j P_{jL}\overline{x}_{jL}(\beta_{jH} - \beta_{jL}) \\ &\quad + \sum_j (\hat{P}_{jL} - P_{jL})\overline{x}_{jH}\beta_{jH} + \sum_j (P_{jH} - \hat{P}_{jL})\overline{x}_{jH}\beta_{jH}.\end{aligned} \tag{2-20}$$

---

21 Brown *et al*.(1980)原文中是男性和女性两个组群，本文为一般性起见统一用组群 H 和 L 表示。

22 显然，在这里，当 $\beta_{jH}=\beta_{kH}=\beta_H$，$\beta_{jL}=\beta_{kL}=\beta_L$ 时，j，k 表示 J 类职业中任何一类，$\sum(P_{jH}\overline{x}_{jH}-P_{jL}\overline{x}_{jL})=\overline{x}_H-\overline{x}_L$，$\sum P_{jL}\overline{x}_{jL}=\overline{x}_L$，此时，Oaxaca-Blinder 分解可以视为 Brown 分解中的一个特例。

(2-20) 式就是 Brown 分解公式。[23] (2-20) 式右边第一项 $\sum_j P_{jL}(\overline{x}_{jH}-\overline{x}_{jL})\beta_{jH}$ 表示职业内两个组群因个体特征差异所造成的工资差距,第二项 $\sum_j P_{jL}\overline{x}_{jL}(\beta_{jH}-\beta_{jL})$ 则是指职业内两个组群因受到不平等对待所造成的工资差距,第三项 $\sum_j(\hat{P}_{jL}-P_{jL})\overline{x}_{jH}\beta_{jH}$ 表示两个组群之间由于职业(获得)结构差别所造成的工资差距,第四项 $\sum_j(P_{jH}-\hat{P}_{jL})\overline{x}_{jH}\beta_{jH}$ 则是指由于职业资质差别所造成的工资差距[24]。第一项和第二项共同构成了职业内工资差异,第三项和第四项共同表征了职业间工资差异。可见,Brown 分解能够将工资差异按来源划分为职业内工资差异和职业间工资差异。

Brown 分解在贯通职业分隔和同工不同酬的视角下解析出了因职业进入的组别(比如性别)障碍而产生的歧视,在改进 Oaxaca-Blinder 分解对歧视测度的同时,优化了工资差异的均值分解。它与 Oaxaca-Blinder 分解相比,可以说是在分析层面上的一种提升。

然而,Brown 分解仍然遗留着指数基准问题,更确切地说是双重指数基准问题(dual index number problems):一是原 Oaxaca-Blinder 分解留存的工资结构指数基准问题(以下简称指数基准问题Ⅰ),二是 Brown 分解中出现的职业(获得)结构指数基准问题(以下简称指数基准问题Ⅱ)。[25] 客观地说,Brown 分解毕竟早在 1980 年就面世了,不可能预见到发表于 1988 年的 Cotton 分解和 Neumark 分解,故仍沿用了原 Oaxaca-Blinder 分解中的将组群 H 的现行工资当成无歧视工资结构的做法,从而存在着指数基准问题Ⅰ。另一方面,Brown 在估计无歧视职业(获得)结构时仅使用了假如 L 组群样本面临与 H 组群相同职业结构时 L 组群处于 j 职业的比率 P_j^L,这在某种程度上暗含着将 L 组群职业(获得)结构当做无歧视职业(获得)结构,并没有考虑到无歧视状态下职业(获得)结构的指数基准问题(即指数基准问题Ⅱ)。双重指数基准问题无疑加剧了在工资差异均值分解中对歧视程度估计的扭曲程度。

2. Appleton 分解

Appleton *et al*. (1999) 率先通过矫正双重指数基准问题来改进 Brown

---

[23] 需要指出的是,Brown *et al*. (1980) 的公式(3)值得商榷,因为 Brown 在无歧视工资结构中以男性为基准(即 $\beta_j^*=\beta_{jH}$),而在无歧视职业获得结构中的基准则是含糊的,即他的 $P_j^*=P_{jL}$ 和 $P_{jH}=P_{jH}^*$ 既不能说是以女性为基准(若以女性为基准,则有 $P_j^*=P_{jL}=P_{jL}^*$),也不能说是以男性为基准(若以男性为基准,则有 $P_j^*=P_{jH}=P_{jH}^*$)。笔者以为,在工资差异的均值分解中,即便存在指数基准问题,分解的基准也应该统一,要么以女性为基准,要么以男性为基准,否则,基准就不成其为基准了。

[24] Brown *et al*. (1980) 原文中的"occupation qualification"宜译为职业资质,意指从事某一职业所具备的技能。第四项实际上表示由个体特征差异带来的职业间工资差距。

[25] 虽然 Brown *et al*. (1980) 在附录中给出了当选取不同组群作为指数基准的分解式,说明他们已意识到了双重指数基准问题。

分解。双重指数基准问题的矫正过程一般说来可以分为两步：

第一步　矫正指数基准问题Ⅱ。引入无歧视职业（获得）结构 P_j^*，即引入无歧视状态下全体劳动者（包括组群 H 和 L 的全体劳动者）在第 j 种职业中就业的概率或比率，从而，组群 H 和 L 的平均工资差异一般地可写成如下形式：

$$
\begin{aligned}
\ln\overline{w}_H-\ln\overline{w}_L &= \sum_j P_{jH}\ln\overline{w}_{jH}-\sum_j P_{jL}\ln\overline{w}_{jL}\\
&= \sum_j P_{jH}\ln\overline{w}_{jH}-\sum_j P_j^*\ln\overline{w}_{jH}+\sum_j P_j^*\ln\overline{w}_{jH}\\
&\quad -\sum_j P_j^*\ln\overline{w}_{jL}+\sum_j P_j^*\ln\overline{w}_{jL}-\sum_j P_{jL}\ln\overline{w}_{jL}\\
&= \sum_j P_j^*(\ln\overline{w}_{jH}-\ln\overline{w}_{jL})+\sum_j(P_{jH}-P_j^*)\ln\overline{w}_{jH}\\
&\quad +\sum_j(P_j^*-P_{jL})\ln\overline{w}_{jL}\\
&= A+B,
\end{aligned}
\tag{2-21}
$$

其中，$A=\sum_j P_j^*(\ln\overline{w}_{jH}-\ln\overline{w}_{jL})$ 是职业内的工资差异，表示由于组群 H 和 L 在各职业内的平均工资不同所造成的期望工资差异；$B=\sum_j(P_{jH}-P_j^*)\ln\overline{w}_{jH}+\sum_j(P_j^*-P_{jL})\ln\overline{w}_{jL}$ 则是职业间的工资差异，意指因组群 H 和 L 对各种不同职业的实际获得概率与无歧视时获得概率之间的差异所引起的期望工资差异之和。

第二步　矫正指数基准问题Ⅰ。此时，需要在工资差异分解中进一步引入无歧视工资结构。记组群 H 和 L 在第 j 种职业中的工资结构系数分别为 β_{jH} 和 β_{jL}，记第 j 种职业中的无歧视工资结构为 β_j^*，则

$$
\begin{aligned}
A &= \sum_j P_j^*(\ln\overline{w}_{jH}-\ln\overline{w}_{jL})=\sum_j P_j^*(\overline{x}_{jH}\beta_{jH}-\overline{x}_{jL}\beta_{jL})\\
&= \sum_j P_j^*(\overline{x}_{jH}\beta_{jH}-\overline{x}_{jH}\beta_j^*+\overline{x}_{jH}\beta_j^*-\overline{x}_{jL}\beta_j^*+\overline{x}_{jL}\beta_j^*-\overline{x}_{jL}\beta_{jL})\\
&= \sum_j P_j^*(\overline{x}_{jH}-\overline{x}_{jL})\beta_j^*+\sum_j P_j^*\overline{x}_{jH}(\beta_{jH}-\beta_j^*)+\sum_j P_j^*\overline{x}_{jL}(\beta_j^*-\beta_{jL}),
\end{aligned}
\tag{2-22}
$$

(2-22) 式中的第一项是职业内工资差异的可解释部分，第二项和第三项则构成了职业内歧视。

再用 P_{jH}^* 和 P_{jL}^* 分别表示无歧视状态下 H 组群和 L 组群在第 j 种职业中就业的概率或比率，则

$$
\begin{aligned}
B= & \sum_j (P_{jH}-P_j^*)\ln\overline{w}_{jH}+\sum_j (P_j^*-P_{jL})\ln\overline{w}_{jL} \\
= & \sum_j (P_{jH}-P_{jH}^*+P_{jH}^*-P_j^*)\ln\overline{w}_{jH}+\sum_j (P_j^*-P_{jL}^*+P_{jL}^*-P_{jL})\ln\overline{w}_{jL} \\
= & \sum_j (P_{jH}^*-P_j^*)\ln\overline{w}_{jH}+\sum_j (P_j^*-P_{jL}^*)\ln\overline{w}_{jL} \\
& +\sum_j (P_{jH}-P_{jH}^*)\ln\overline{w}_{jH}+\sum_j (P_{jL}^*-P_{jL})\ln\overline{w}_{jL}, \qquad (2\text{-}23)
\end{aligned}
$$

(2-23) 式中的第一项和第二项表示职业间组群工资差异的可解释部分，这是不同组群（包括 H 组群和 L 组群）由于个体可观测特征不同引起的职业获得差别所造成的工资差异；第三项和第四项则表示职业分隔中不可解释部分，被称为职业间歧视。

将（2-22）式和（2-23）式代入（2-21）式，就可以得到经过双重指数基准矫正后的 Brown 分解形式：

$$
\begin{aligned}
\ln\overline{w}_H-\ln\overline{w}_L= & \sum_j P_j^*(\overline{x}_{jH}-\overline{x}_{jL})\beta_j^*+\sum_j P_j^*\overline{x}_{jH}(\beta_{jH}-\beta_j^*) \\
& +\sum_j P_j^*\overline{x}_{jL}(\beta_j^*-\beta_{jL})+\sum_j (P_{jH}^*-P_j^*)\ln\overline{w}_{jH} \\
& +\sum_j (P_j^*-P_{jL}^*)\ln\overline{w}_{jL}+\sum_j (P_{jH}-P_{jH}^*)\ln\overline{w}_{jH} \\
& +\sum_j (P_{jL}^*-P_{jL})\ln\overline{w}_{jL}, \qquad (2\text{-}24)
\end{aligned}
$$

(2-24) 式中的第一项是职业内工资差异的可解释部分；第二项和第三项分别表示职业内的反向歧视和直接歧视，这两项构成了职业内（总）歧视，表现为同工不同酬。第四项和第五项是职业间工资差异的可解释部分；第六项和第七项则是指职业间反向歧视和直接歧视，这两者共同构成职业间歧视，代表因组群歧视引起职业进入障碍所造成的工资差异。与原 Brown 分解即(2-20)式相比，矫正后的 Brown 分解即（2-24）式是对原 Brown 分解的深化和一般化，而原 Brown 分解则只是矫正后的 Brown 分解形式的一个特例。[26]

显而易见，矫正 Brown 分解的关键在于能够得到无歧视职业（获得）结构和无歧视工资结构。对于无歧视工资结构，Appleton *et al*.（1999）直接借

---

[26] 若以 L 组群为基准，则 $P_j^*=P_{jL}=P_{jL}^*$，$\beta_j^*=\beta_{jL}$，从而，(2-24)式就可以退化为原 Brown 分解：

$$
\begin{aligned}
\ln\overline{w}_H-\ln\overline{w}_L= & \sum_j P_{jL}[(\overline{x}_{jH}-\overline{x}_{jL})\beta_{jL}+\overline{x}_{jH}(\beta_{jH}-\beta_{jL})] \\
& +\sum_j [(P_{jH}^*-P_{jL})\ln\overline{w}_{jH}+(P_{jH}-P_{jH}^*)\ln\overline{w}_{jH}],
\end{aligned}
$$

这里的 P_{jH}^* 相当于 Brown *et al*.（1980）中所说的 $\hat{P}_{jH}$，即假如 H 组群面临与 L 组群相同职业结构时 H 组群处于 j 职业的比率。笔者认为，$\hat{P}_{jH}$ 在某种程度上可以说是对 P_{jH}^* 的一种近似估计。同理，我们还可以写出原 Brown 分解中以 H 组群为基准的工资均值分解形式。

用了 Neumark（1988）用全样本工资回归方程获得的工资结构的加权形式作为无歧视工资结构的刻画。关于无歧视职业（获得）结构，Appleton *et al*.（1999）仍借助 Neumark（1988）无歧视工资结构加权形式中的权重，采用加权思路构造无歧视职业（获得）结构（参见（2-10）式）。尽管这种加权构造的思想和形式早已出现在 Reimers（1983）和 Cotton（1988）对无歧视工资结构的考察之中，但 Appleton 的"别样"之处则是进一步将其直接套用到对职业获得的分析上，提出了如下形式的无歧视职业（获得）结构：

$$\gamma_{Aj}^{*} = \Omega\gamma_{jH} + (I - \Omega)\gamma_{jL}, \tag{2-25}$$

其中，Ω 的含义同上，即 Neumark 无歧视工资结构的权重矩阵，而 γ_{jH} 和 γ_{jL} 则是第 j 种职业中分别用 H 组群和 L 组群的子样本进行多元 logit 模型估算所得到的系数向量，γ_{Aj}^{*} 表示 Appleton 意义上的无歧视职业（获得）结构[27]（non-discriminatory occupational attainment structure；在不至于引起歧义的情况下也可简称为"无歧视职业结构"），它可以说是逼近"真正"无歧视职业（获得）结构的一种有益尝试。

Appleton 之所以没有直接采用全样本数据而是用 H 组群和 L 组群的子样本各自的系数向量加权的形式估算无歧视职业（获得）结构，笔者以为，很可能是他已经认识到：在 H 和 L 两个组群的总体特征回报存在显著差异的情况下（否则，就没有必要对 H 组群和 L 组群的工资差异进行分别考察），如果直接在 Neumark 视角下用全样本回归的职业获得方程即多元 logit 模型 $\text{logit}P_j = \gamma_{0j} + \gamma_{1j}X + u$ 中进行回归[28]，并以相应的回归系数构造成无歧视职业（获得）结构，那么，回归估计的结果无疑将产生偏误；另一方面，由于所得到的无歧视职业分布将与现实劳动力市场中的职业分布相同，这就意味着歧视并不会改变整个劳动力市场的职业分布，而只是引起同一职业内 H 和 L 两个组群所占比重发生变化。这种推断结果或者说内在局限势必与劳动力市场的实际存在较大出入。

此外，Appleton"将 Neumark 无歧视工资结构的权重矩阵 Ω 径直充作无歧视职业（获得）结构权重矩阵"的做法值得商榷。首先，Appleton 没有也无法在理论上论证出"为什么能够借用 Neumark 无歧视工资结构的权重矩阵作为无歧视职业（获得）结构的权重矩阵"，因为职业获得机制不同于工资决定机制，两者的影响因素也各有不同。其次，就估计方法而言，工资决定估计所使用的是 Mincer 工资方程，全样本数据的 Mincer 工资方程采用经典线性回归，其估计结果亦即 Neumark 意义上的无歧视工资结构恰好可以写成加

[27] 无歧视第 j 种职业（获得）结构通常用 γ_j^{*} 表示，但本文为了将 Appleton 意义上的无歧视职业（获得）结构与其他的无歧视职业（获得）结构相区别，特将其记成 γ_{Aj}^{*}。

[28] P_j 为总样本中劳动者（包括 H 组群和 L 组群）在第 j 种职业中的职业获得概率。

权形式；而职业获得概率估计则采用多元 logit 模型，全样本多元 logit 模型回归得到的系数向量却难以简单地写成形如（2-25）式的加权形式。概言之，无歧视职业（获得）结构与无歧视工资结构的理论基础和估计方法存在着质的规定性差异。事实上，Appleton *et al*.（1999）自己也已表示，没有明确的证据可以证明零阶齐次假定是否满足现实劳动力市场，因而，全样本回归是不是无歧视工资结构的一个好的估计值并不明确，同时，无歧视职业结构中的权重是不是一个好的权重也不能确定。可见，Appleton 对直接承袭 Neumark 的分析思路以及无歧视职业（获得）结构套用权重矩阵 Ω，都是没有把握的。

（三）JMP1991 分解

JMP1991 分解是指 Juhn *et al*.（1991）使用的主要方法，它是在均值分解中引入分布工具的一种分解方法。就分解方法而言，JMP1991 分解通过将组群均值工资差异的不可解释部分亦即残差额（residual gap）表达成标准差与标准化残差分布差异之乘积，并与不可观测特征[29]（更一般地说是组群特异性）的价格和数量差相对应，以此进一步挖掘了工资差异不可解释部分的成因。

JMP1991 分解和 Oaxaca-Blinder 分解尽管都采取基于 OLS 的线性回归模型，但在 Oaxaca-Blinder 分解中，不同组群使用各自的工资方程，每个工资方程均满足零条件均值假定，两个组群之间的工资均值差异可分解成可解释部分和不可解释部分。而 JMP1991 分解则选定某个组群作为基准，假定基准组群工资方程满足零条件均值；在另一组群与基准组群具有相同工资结构系数的情况下，两个组群之间工资均值差异就可分解为可解释部分和残差额（不可解释部分）；再假定另一组群与基准组群具有相同的方差，从而残差额可进一步表述成标准差与标准化的残差分布差异之乘积。如果说 Oaxaca-Blinder 分解关注的是某一时期两个组群之间的工资差异，那么，JMP1991 分解的重点则在于不同时期组群之间工资差异的变化。

具体地说，Juhn *et al*.（1991）是在两个组群 H 和 L（例如白人与黑人组群）具有同方差假设下，以组群 H 为基准将某一时期（譬如 t 时期）组群 H 和 L 间工资差异的不可解释部分表述成为标准差与标准化残差分布差异之乘积。对组群 H 的工资回归方程而言，有 $\ln\overline{\varpi}_{Ht}=\overline{X}_{Ht}\beta_t$；对组群 L 而言，若以组群 H 的工资结构系数 β_t 为基准，则有 $\ln\overline{\varpi}_{Lt}=\overline{X}_{Lt}\beta_t+u_{Lt}$，$u_{Lt}$ 是 L 组群实际工资与采用 H 组群工资结构得到的工资估计值之差。于是，t 时期组群 H 和 L 之间的均值工资差异 $D_t=\ln\overline{\varpi}_{Ht}-\ln\overline{\varpi}_{Lt}=(\overline{X}_{Ht}-\overline{X}_{Lt})\beta_t-u_{Lt}$。在同方差

[29] 鉴于后面的分解方法，开始关注残差，尤其是不可观测特征（技能）对工资差异的影响，本文从此开始将个体特征分为可观测特征和不可观测特征。

假设下，组群 H 和 L 具有相同的残差分布离散程度。残差额（$-u_{Lt}$）可进一步看成是组群 H 的标准差 σ_t 与组群 L 的标准化残差差异 $\Delta\bar{\theta}_t$ 的乘积[30]，其中 $\Delta\bar{\theta}_t=0-\bar{\theta}_t=-\bar{\theta}_t$，表示组群 H 和 L 的不可观测特征分布的相对位置亦即组群 L 相对于组群 H 在不可观测特征上的差异。从而，

$$D_t=\ln\bar{w}_{Ht}-\ln\bar{w}_{Lt}=(\bar{X}_{Ht}-\bar{X}_{Lt})\beta_t-u_{Lt}=\Delta\bar{X}_t\beta_t+\sigma_t\Delta\bar{\theta}_t.$$

从第 $t=0$ 期到第 $t=1$ 期，组群 H 和 L 之间均值工资差异的变化为

$$\begin{aligned}D_1-D_0&=(\Delta\bar{X}_1\beta_1+\sigma_1\Delta\bar{\theta}_1)-(\Delta\bar{X}_0\beta_0+\sigma_0\Delta\bar{\theta}_0)\\&=(\Delta\bar{X}_1\beta_1-\Delta\bar{X}_0\beta_0)+(\sigma_1\Delta\bar{\theta}_1-\sigma_0\Delta\bar{\theta}_0).\end{aligned}$$

Juhn *et al*.（1991）以第 $t=0$ 期为基期进一步进行了如下分解：

$$D_1-D_0=(\Delta\bar{X}_1-\Delta\bar{X}_0)\beta_0+\Delta\bar{X}_1(\beta_1-\beta_0)+(\Delta\bar{\theta}_1-\Delta\bar{\theta}_0)\sigma_0+\Delta\bar{\theta}_1(\sigma_1-\sigma_0),\tag{2-26}$$

等式（2-26）右边的第一项 $(\Delta\bar{X}_1-\Delta\bar{X}_0)\beta_0$ 表示在固定价格 β_0 下，组群 H 和 L 的个人可观测特征差异变动对组群工资差异变动的贡献，也就是在固定价格下可观测数量变化的效应[31]；第二项 $\Delta\bar{X}_1(\beta_1-\beta_0)$ 表示组群 H 工资结构变动对组群工资差异变动的贡献，称之为价格效应；第三项 $(\Delta\bar{\theta}_1-\Delta\bar{\theta}_0)\sigma_0$ 抓住了组群 L 与组群 H 的相对工资分布的变动，也就是组群 L 相对于组群 H 的工资分布是上移还是下移，这可以看做是由不可观测特征差异变动所造成的工资差异或者说组群 L 的特异性效应[32]；第四项 $\Delta\bar{\theta}_1(\sigma_1-\sigma_0)$ 则捕捉了组群 H 工资分布的离散化程度变化或者说工资不平等（wage inequality）程度变化对组群工资差异变动的贡献，这是一种广义相对价格效应。第一项和第二项共同反映了由可解释部分变动造成的组群工资差异变动，而第三项和第四项则共同反映了不可解释部分即残差额变动造成的组群工资差异变动。

在实证上，针对自 1979 年以来美国黑人-白人工资趋同减缓的现象，Juhn *et al*.（1991）用 JMP1991 分解的结果表明，不可观测特征的广义相对价格上升是导致“黑人-白人工资趋同减缓”的重要原因。因而，黑人-白人工资趋同的减缓反映了一个更为广泛的工资不平等的增长趋势。

Juhn *et al*.（1991）中的 JMP1991 分解过程有以下一些特别需要注意的问题：(1) 两个组群具有相同方差假设的现实性未及验证。现实经济中高收入

---

[30] 标准化残差 $\theta_t=(\ln w_{Lt}-\bar{X}_{Lt}\beta_t)/\sigma_t$。标准化残差的分布是均值为 0、方差为 1 的正态分布。此外，Neuman and Oaxaca(2004)还曾将 θ_t 称为 JMP 均值标准工资残差(the JMP mean standardized wage residuals)。

[31] 这里所说的个人可观测特征或禀赋就是以前所说的个人特征或个人禀赋，这是为了与不可观测特征相对应。

[32] Juhn *et al*.(1991)研究的是白人与黑人组群的工资差异，他们曾将黑人特异性所造成的工资差异简称为纯黑人特异性效应(the true black-specific effect)。

组群 H 的方差通常要比低收入组群 L 的方差大，故而同方差性须经过检验，否则就不宜用 JMP1991 分解方法。(2) 组群特异性之内涵的模糊性。以 Juhn *et al*. (1991) 的研究为例，Juhn 研究的是黑人组群与白人组群（white-black）之间工资差异的变化，组群的特异性就表达为黑人特异性（black-specific）。若市场上没有歧视，黑人特异性自然可以对应于不可观测特征数量的差异；但若既有技能差异又存在歧视，则 Juhn 自己也认为，黑人特异性就不宜再只归因于不可观测特征数量的差异。Juhn 还尝试把不可观测特征与歧视分开进行定量考察，不过，这种区分的可操作性仍有待进一步探索。(3) 组群分解的基准问题。JMP1991 分解中有两个基准，一个是在同一时期以某一组群作为基准，另一个则是选取某一时期作为两个时期工资差异变化的基期，譬如 Juhn *et al*. (1991) 就是分别以白人组群和第 $t=0$ 期为基准，但如果考虑可以用黑人组群和第 $t=1$ 期为基准，则还会出现另外三种组合的基准，造成所得到的分解结果的差异性和不确定性。[33] 其实，在存在歧视的情况下，无论是以白人为基准还是以黑人为基准，都仍然遗留着指数基准问题。

总的说来，与以往只关注经典线性回归残差均值为零的特性而忽略残差分布的均值分解方法不同，JMP1991 分解将目光投向了残差分布及其解析，并推动着工资差异均值分解进一步向分布分解深化。从这个角度看，JMP1991 分解也是从均值分解向分布分解的过渡。同时，值得一提的是，在 JMP1991 分解中歧视关注度被降低。特别是当工资差异从均值分解转向分布分解，研究对象主要是同一组群不同时期工资分布变动时，更是如此。原因在于，劳动力市场歧视是与不同组群特异性相联系的，当研究同一组群工资分布变动时，不再涉及组群差异，也就没有再论及歧视问题。当然，工资分布分解方法也可以应用于不同组群工资分布差异分解，一旦涉及不同组群就必然无法回避歧视问题。不过，即使研究对象是不同组群工资分布差异，限于分析主旨，对歧视问题的处理通常也只是提及但很少作专门分析。

（四）对均值分解中若干计量问题的矫正

工资差异分解过程中除了指数基准问题以外，还会遇到其他若干计量问题，其中较为普遍且影响较大的是样本选择偏差纠正项归因和虚拟变量系数识别问题。Neuman and Oaxaca (2004) 对样本选择纠正项的归因问题分不同情况进行了细致的讨论并提供了可选的矫正方案。Yun (2005) 利用标准化

---

[33] 例如，Blau and Kahn (1997) 曾以第 $t=1$ 期作为基期，得到另一种分解形式。显然，选取不同的基期，将获得有差别的分解结果。Juhn *et al*. (1991) 显然已经意识了到这个问题，在实证中事实上使用了两个时期（若用多个时期则是所有期）的平均值作为基准（benchmark），从而有

$$D_1-\bar{D}=(\Delta X_1-\Delta\bar{X})\bar{\beta}+\Delta\bar{X}_1(\beta_1-\bar{\beta})+(\Delta\bar{\theta}_1-\Delta\bar{\theta})\bar{\sigma}+\Delta\bar{\theta}_1(\sigma_1-\bar{\sigma}).$$

但如此一来，分解的就不是两期之间的工资差异变化了。

系数方程解决了虚拟变量系数识别问题。[34]

1. 选择性偏差问题

Neuman and Oaxaca（2004）讨论了在不同组群工资均值差异分解中的 Heckman（Heckit）样本选择偏差纠正项应该如何归因的问题。其中，样本选择偏差纠正项来源于 Heckit 模型（Heckman，1979），它解决了实际样本不满足随机抽样的假定时，导致 OLS 估计量有偏误和不一致的问题。[35]因为实际工资数据只有在就业概率 $P_{ig}^{*}>0$ 时才能被观测到，所以被雇用个体的期望工资为

$$E(\ln w_{ig} \mid P_{ig}^{*}>0)=X_{ig}\beta_{g}+E(u_{ig} \mid \varepsilon_{ig}>-H_{ig}\gamma_{g})=X_{ig}\beta_{g}+\theta_{g}\lambda_{ig} \quad g=H,L, \tag{2-27}$$

其中，$\ln w_{ig}$ 是小时工资对数，P_{ig}^{*} 是就业的潜变量，H_{ig} 是决定个体劳动参与的特征向量，X_{ig} 是决定个体工资的个体可观测特征向量，γ_{g} 和 β_{g} 是对应的参数向量，同时扰动项 ε_{ig} 和 u_{ig} 满足 i. i. d.（独立同分布）并服从二元正态分布 $(0,0,\sigma_{\varepsilon g},\sigma_{ug},\rho)$。样本选择偏差纠正项 $\theta_{g}\lambda_{ig}$ 由参数 θ_{g} 和逆米尔斯比率（IMR）λ_{ig} 构成，其中参数 θ_{g} 被识别为就业决定方程和工资方程误差项的相关系数 ρ_{g} 与工资方程误差项标准差 σ_{ug} 的乘积，即 $\theta_{g}=\rho_{g}\sigma_{ug}$，而逆米尔斯比率 $\lambda_{ig}=\phi(H_{ig}\gamma_{g})/\Phi(H_{ig}\gamma_{g})$。[36]

在 Oaxaca-Blinder 分解的基础上，考虑样本选择纠正后的组群工资均值差异，可以分解为[37]

$$\ln\overline{w}_{H}-\ln\overline{w}_{L}=\overline{X}_{L}(\beta_{H}-\beta_{L})+(\overline{X}_{H}-\overline{X}_{L})\beta_{H}+(\theta_{H}\lambda_{H}-\theta_{L}\lambda_{L}). \tag{2-28}$$

不同组群间 Heckit 样本选择纠正项的差异 $(\theta_{H}\lambda_{H}-\theta_{L}\lambda_{L})$ 如何归因成为矫正选择性偏差后的遗留问题。Duncan and Leigh（1980）的处理方法是回避对样本选择偏差是否应归因于歧视还是个人特征差异的讨论，即将（2-28）等式右边的 $(\theta_{H}\lambda_{H}-\theta_{L}\lambda_{L})$ 进行移项，利用去除样本选择影响后的工资差异净值进行分解，将分解式写为

$$(\ln\overline{w}_{H}-\ln\overline{w}_{L})-(\theta_{H}\lambda_{H}-\theta_{L}\lambda_{L})=\overline{X}_{L}(\beta_{H}-\beta_{L})+(\overline{X}_{H}-\overline{X}_{L})\beta_{L}.^{38} \tag{2-29}$$

[34] 虽然样本选择纠正项的归因和虚拟变量系数识别问题在工资差异均值分解中已得到某种程度的矫正，但如何矫正工资差异分布分解中类似的计量问题仍值得探寻。

[35] 当进入劳动力市场的概率与某个观察不到而又影响工资高低的因素系统相关时，产生 OLS 估计量有偏误和不一致问题，从而需要纠正样本选择来解决偏误问题。

[36] 劳动参与选择方程为 $L_{ig}^{*}=H_{ig}\gamma_{g}+\varepsilon_{ig}$；工资决定方程为 $\ln w_{ig}=X_{ig}\beta_{g}+u_{ig}$。就业概率 $\mathrm{Prob}(L_{ig}^{*}>0)=\mathrm{Prob}(\varepsilon_{ig}>-H_{ig}\gamma_{g})=\Phi(H_{ig}\gamma_{g})$，其中 $\Phi(\cdot)$ 表示标准正态分布的累积分布函数，$\phi(\cdot)$ 表示标准正态分布的密度函数。

[37] Neuman and Oaxaca(2004)指出，暂时不探讨指数基准问题，而是关注于样本选择项的归因问题，为了方便表述，都以 H 组群作为分解基准。

[38] Reimers(1983)也利用了这一处理方法，同时采用 $\beta_{R}^{*}=D\beta_{H}+(I-D)\beta_{L}$，$D=\mathrm{diag}(0.5,0.5,\cdots,0.5)=0.5(I)$，对指数基准问题进行了某种程度的矫正。

Neuman and Oaxaca（2004）认为，（2-29）式并没有分解实际观测到的组群间工资差异（$\ln\overline{w}_{\mathrm{H}}-\ln\overline{w}_{L}$），所以应将组群间样本选择纠正项差异（$\theta_H\lambda_H-\theta_L\lambda_L$）纳入工资均值差异分解框架。Neuman and Oaxaca 将可观测的组群工资均值差异分解为歧视、个人特征和个人选择三个部分的贡献，并给出了四种带有不同主观判断的分解形式。

第一种是最直接的分解形式，就是将劳动参与选择所导致的工资差异全部认为是个人选择因素造成的。

$$\ln\overline{w}_{\mathrm{H}}-\ln\overline{w}_{L}=\overline{X}_L(\beta_H-\beta_L)+(\overline{X}_H-\overline{X}_L)\beta_H+(\theta_H\lambda_H-\theta_L\lambda_L), \tag{2-30}$$

（2-30）式第一项归因于歧视，第二项归因于个人特征差异，第三项是个人选择造成的工资差异。

第二种形式是将 H 组群和 L 组群的劳动参与选择方程中的系数差异归因为受劳动力市场歧视的影响，而劳动参与选择方程中的个体特征的差异和工资方程中的个体特征差异合并后，分解式为

$$\begin{aligned}\ln\overline{w}_H-\ln\overline{w}_L=&[\overline{X}_L(\beta_H-\beta_L)+\theta_H(\lambda_L^0-\lambda_L)]+[(\overline{X}_H-\overline{X}_L)\beta_H+\theta_H(\lambda_H-\lambda_L^0)]\\&+[(\theta_H-\theta_L)\lambda_L],\end{aligned} \tag{2-31}$$

其中 $\lambda_L^0=\sum_{i=1}^{N_L}\lambda_{iL}^0/N_L$，$\lambda_{iL}^0=\phi(H_{iL}\gamma_H)/\Phi(H_{iL}\gamma_H)$，$N_L$ 是组群 L 的样本个数，λ_L^0是当 L 组群面对与 H 组群相同劳动参与选择方程时的逆米尔斯均值。(2-31)式第一项归因于歧视，第二项归因于个人可观测特征差异，第三项归因于个人选择。

第三种形式是将劳动参与选择方程和工资方程的误差相关系数 ρ_s 的差异可以作为反映个体特征差异的部分，而 H 和 L 组群工资离散程度的差异仍无法推断是否由歧视导致，于是，组群工资差异均值分解形式为

$$\begin{aligned}\ln\overline{w}_H-\ln\overline{w}_L=&[\overline{X}_L(\beta_H-\beta_L)+\theta_H(\lambda_L^0-\lambda_L)]\\&+[(\overline{X}_H-\overline{X}_L)\beta_H+\theta_H(\lambda_H-\lambda_L^0)+(\rho_H-\rho_L)\sigma_{uL}]\\&+\rho_H(\sigma_{uH}-\sigma_{uL}),\end{aligned} \tag{2-32}$$

（2-32）式第一项归因于歧视，第二项归因于个人特征差异，第三项归因于个人选择。

第四种形式是将劳动参与选择方程的组群系数差异和劳动力市场对不同组群不可观测的选择因素反应不同所导致的工资差异都归结为歧视，那么分解式为

$$\begin{aligned}\ln\overline{w}_H-\ln\overline{w}_L=&[\overline{X}_L(\beta_H-\beta_L)+\theta_H(\lambda_L^0-\lambda_L)+(\theta_H-\theta_L)\lambda_L]\\&+[(\overline{X}_H-\overline{X}_L)\beta_H+\theta_H(\lambda_H-\lambda_L^0)]\end{aligned}$$

$$= [\overline{X}_L(\beta_H - \beta_L) + \theta_H\lambda_L^0 - \theta_L\lambda_L] + [(\overline{X}_H - \overline{X}_L)\beta_H + \theta_H(\lambda_H - \lambda_L^0)], \tag{2-33}$$

(2-33) 式第一项归因于歧视带来的工资差异，第二项归因于个人特征差异导致的工资差异。

由于 (2-30) 式、(2-31) 式、(2-32) 式和 (2-33) 式涉及四种不同的假设与价值判断，因而上述四种纠正样本选择偏差后的组群工资差异均值分解式难以理性地区分优劣。

2. 虚拟变量系数识别问题

工资差异均值分解通常基于 Mincer 工资决定模型，而在模型中经常会加入一些类别变量（例如是否结婚、是否参加过培训、所在地区等），这些类别变量在回归模型中一般用虚拟变量来表示，这时就需要选择虚拟变量的基准组。根据计量理论可知，基准组选择的不同并不影响计量模型的正确估计；进一步的，在工资差异均值分解的研究中如果只需要得到总的特征效应（可解释部分）和总的回归系数效应（不可解释部分），那么虚拟变量选择不同的基准组也不影响分解结果。然而，当工资均值差异需要分解到各个单一协变量的效应时，每个虚拟变量的特征效应和系数效应就会随着基准组选择的不同而变化，也就产生了虚拟变量系数识别问题。[39]

Yun (2005) 通过利用标准化的回归方程 (normalized regression) 提出了一种只需要做一次回归[40]就能解决虚拟变量系数和常数项[41]识别问题的矫正方法。为简便起见，假设没有其他解释变量[42]，只有虚拟变量，并且虚拟变量只有三组值。取第一类为基准组，那么估计后的对数工资均值方程是

$$\ln\overline{w}_g = \alpha_g + \beta_{2g}\overline{D}_{2g} + \beta_{3g}\overline{D}_{3g}, \tag{2-34}$$

其中，$g = H, L$；D_{2g} 表示属于第二类时为 1，其他为 0；D_{3g} 表示当属于第三类时为 1，其他为 0。那么，$\overline{D}_{2g}$ 和 $\overline{D}_{3g}$ 分别表示属于第二类和第三类的比率。为了获得平均效应，把 (2-34) 式写成下面的一个标准方程：

$$\ln\overline{w}_g = \alpha_g + \bar{\beta}_g + (\beta_{2g} - \bar{\beta}_g)\overline{D}_{2g} + (\beta_{3g} - \bar{\beta}_g)\overline{D}_{3g}, \tag{2-35}$$

其中 $\bar{\beta}_g = (\beta_{1g} + \beta_{2g} + \beta_{3g})/3$，因为第一类为基准组，所以 β_{1g} 等于 0，方程

---

[39] Oaxaca and Ransom(1999)认为，系数效应的具体分解（分解至各个单一变量的特征效应和系数效应）必定无法避免识别问题。这种不稳定性或者识别问题对劳动经济学家来说并不陌生，且这个问题已经困扰分解或歧视分析很长一段时期。

[40] 解决这个识别问题的一种直觉的想法就是，既然选择不同的基准组会得到不同的结果，那么应该通过改变次序让每一个虚拟变量中的每一类都当一次基准组，然后分别作回归方程，获得所有的分解结果，然后将这些结果平均，这个平均效应就可以避免识别问题。但这种方法比较麻烦，需要做很多次回归。

[41] 作为基准组的效应反映在常数项中，因此，在分析虚拟变量系数识别问题时，就会关联常数项。

[42] Yun(2005)从一个简化的具体例子出发来阐述他的方法，这个简化的例子可以直接拓展到一般化的方法。

(2-35)即为标准化的回归方程。[43]然后根据这个标准化的回归方程进行工资差异 Oaxaca-Blinder 分解如下：

$$\begin{aligned}\ln\overline{w}_H-\ln\overline{w}_L&=[(\alpha_H+\bar{\beta}_H)-(\alpha_L+\bar{\beta}_L)]\\&\quad+\{[(\beta_{2H}-\bar{\beta}_H)-(\beta_{2L}-\bar{\beta}_L)]\overline{D}_{2H}+[(\beta_{3H}-\bar{\beta}_H)-(\beta_{3L}-\bar{\beta}_L)]\overline{D}_{3H}\}\\&\quad+[(\overline{D}_{2H}-\overline{D}_{2L})(\beta_{2L}-\bar{\beta}_L)+(\overline{D}_{3H}-\overline{D}_{3L})(\beta_{3L}-\bar{\beta}_L)],\end{aligned}\tag{2-36}$$

其中第一项为常数项效应；第二项为虚拟变量的系数效应；第三项为虚拟变量的特征效应。采用标准化方程得到的均值分解结果就不会存在各个单一虚拟变量的系数效应和特征效应随着基准组选择不同而变化的问题，从而消减了工资差异均值分解中的虚拟变量系数识别问题。

三、工资差异分布分解

均值分解为考察工资差异提供了简洁明快的分析路径，但工资均值描绘的仅是工资分布的集中趋势，无法对劳动力市场中日趋离散的工资分布提供更为细致的考察。本部分关注从分布分解的视角研究工资差异，将已有的工资差异分布分解方法归为四类情形加以述评：一是基于经典线性回归的分布分解；二是基于半参模型的分布分解；三是基于条件分位回归方程的分布分解；四是基于 RIF 回归方程的分布分解。

（一）基于经典线性回归的分布分解

JMP1993 分解研究的是同一组群工资分布的变动成因，这标志着工资差异分解突破仅从均值讨论的瓶颈，真正进入分布分解的研究新阶段。[44]在 JMP1993 分解考察残差排位的基础上，FL1998 分解引入（综合）技能指数(skill index)，利用技能排位和工资排位相对应的思路，对同一组群工资分布变动和不同组群工资分布差异变动在排位回归（rank regression，亦称秩回归）的框架下进行分解，避开了对工资与技能之间关系具体形式的假定，改进并深化了 JMP1991 分解和 JMP1993 分解。[45]

---

[43] 如果分别选择第一类、第二类和第三类作为基准组，分别可以写出三个标准化回归方程。

$$\ln\overline{w}_g=\alpha_g^1+\bar{\beta}_g^1+(\beta_{2g}^1-\bar{\beta}_g^1)\overline{D}_{2g}+(\beta_{3g}^1-\bar{\beta}_g^1)\overline{D}_{3g}\quad（以第一类为基准组），$$
$$\ln\overline{w}_g=\alpha_g^2+\bar{\beta}_g^2+(\beta_{1g}^2-\bar{\beta}_g^2)\overline{D}_{1g}+(\beta_{3g}^2-\bar{\beta}_g^2)\overline{D}_{3g}\quad（以第二类为基准组），$$
$$\ln\overline{w}_g=\alpha_g^3+\bar{\beta}_g^3+(\beta_{1g}^3-\bar{\beta}_g^3)\overline{D}_{1g}+(\beta_{2g}^3-\bar{\beta}_g^3)\overline{D}_{2g}\quad（以第三类为基准组）.$$

用 k 表示用来做基准组的类别，$k=1,2,3$。尽管当 k 取不同的值时，$\bar{\beta}_{g1}^k$、$\bar{\beta}_{g2}^k$ 和 $\bar{\beta}_{g3}^k$ 的值会发生变化，但 $(\bar{\beta}_{g1}^k-\bar{\beta}_g^k)$、$(\bar{\beta}_{g2}^k-\bar{\beta}_g^k)$ 和 $(\bar{\beta}_{g3}^k-\bar{\beta}_g^k)$ 都将不变。

[44] JMP1991 分解可以认为是从分布视角探讨工资差异变动的萌芽，但因其讨论的对象仍是不同组群的工资均值差异，故它属于工资差异均值分解方法。

[45] FL1998 分解虽然不是基于经典线性回归的分解方法，但由于该方法的排位思想源于 JMP1993 分解，因此本文将该方法放在此处述评。实际上，我们在对各种分解方法综述时往往按第一种基础性方法归类，而后面的方法则是属于递进和拓展的情形。

1. JMP1993 分解

JMP1993 分解是指 Juhn *et al*. (1993) 使用的主要方法，用以找寻同一组群工资分布变动的成因，或者说探究同一组群内工资不平等变化的原因。它的主要思路是，通过把工资（回归）方程 $\ln w_{it}=X_{it}\beta_t+u_{it}$ 中的残差 u_{it} 设定成 $F_t^{-1}(\tau_{it}|X_{it})$，即个体 i 不可观测技能在组群中的排位（用分位 τ_{it} 表示）所对应累积分布的反函数之值，并使 τ_{it} 与工资方程残差分布中的分位相对应[46]，从而将同一组群工资分布的变动分解为由个人可观测特征的数量变化、特征价格（特征回报）变化以及由不可观测技能造成的变动。

在同一时期 t（时期），工资方程可解析为[47]

$$\begin{aligned}\ln w_{it} &= X_{it}\beta_t+u_{it} = X_{it}\beta_t+F_t^{-1}(\tau_{it}|X_{it})\\ &=[X_{it}\bar{\beta}+\bar{F}^{-1}(\tau_{it}|X_{it})]+X_t(\beta_t-\bar{\beta})+[F_t^{-1}(\tau_{it}|X_{it})-\bar{F}^{-1}(\tau_{it}|X_{it})],\end{aligned} \tag{3-1}$$

其中，$\bar{\beta}$ 是所有时期可观测特征回报的平均值，$\bar{F}^{-1}(\cdot)$ 为平均累积分布的反函数。方程右边第一项和第二项分别反映了个人特征数量变动效应和特征回报变动效应；第三项则捕捉了不可观测技能变动效应。[48]

同一组群在不同时期（譬如从 $t=0$ 期到 $t=1$ 期）工资不平等的变动，可以通过某些工资分布统计量的变动来反映。Juhn *et al*. (1993) 采取了用不同分位（比如，90 分位与 10 分位）工资差异的变动来表征不同时期工资不平等的变动。在分位视角下，工资方程可写成

$$\ln w_{t,\tau}=[X_{t,\tau}\bar{\beta}+\bar{F}^{-1}(\tau|X_{t,\tau})]+X_{t,\tau}(\beta_t-\bar{\beta})+[F_t^{-1}(\tau|X_{t,\tau})-\bar{F}^{-1}(\tau|X_{t,\tau})], \tag{3-2}$$

其中，$\ln w_{t,\tau}$ 表示 t 时期 τ 分位的工资[49]，$X_{t,\tau}$ 表示 t 时期不可观测技能处于 τ

---

[46] 在同一组群内 $\tau_{it}|X_{it}$ 不随时间变化的隐含假定下，个体在各时期的不可观测技能分布的分位与其在工资方程的残差分布中的分位相对应。Juhn *et al*. (1993) 采用的表述是百分位(Percentile)，而百分位只是分位(Quantile)的一种形式，故本文采用分位的说法。此外，在 Juhn *et al*. (1993) 中分位用 θ_{it} 表示，本文为行文一致起见，分位统一用 τ_{it} 表达，同时还可以避免与 Juhn *et al*. (1991) 中的标准化残差 θ_t 相混淆。

[47] Fortin *et al*. (2010) 认为，Juhn *et al*. (1993) 文章中没有给出在条件分布下的残差补缺(residual impute)过程，因此 JMP1993 分解方法在经验分析中通常假定残差 u_{it} 与 X_{it} 是独立的，也就是残差值仅与残差排名与残差分布函数有关，即 $F_t^{-1}(\tau_{it}|X_{it})=F_t^{-1}(\tau_{it})$。关于这个问题，仍值得进一步深究。不过，本文的 JMP1993 分解公式仍按照条件分布的形式给出。

[48] 为阐释这个问题，Juhn *et al*. (1993) 提出了 $\ln w_{it}^1$、$\ln w_{it}^2$ 和 $\ln w_{it}^3$ 三种状态：记 $\ln w_{it}^1=X_{it}\bar{\beta}+\bar{F}^{-1}(\tau_{it}|X_{it})$，它反映了 $F_t^{-1}(\cdot)$ 和 β 固定(分别为 $\bar{F}^{-1}(\cdot)$ 和 $\bar{\beta}$)，而 X_{it} 为 t 期时的工资；再记 $\ln w_{it}^2=X_{it}\beta_t+\bar{F}^{-1}(\tau_{it}|X_{it})$，它反映了既定 $\bar{F}^{-1}(\cdot)$，而 X_{it} 的数量和回报为 t 期时的工资；又记 $\ln w_{it}^3=X_{it}\beta_t+F_t^{-1}(\tau_{it}|X_{it})$，它其实就是 t 时期的工资方程，反映了 X_{it}、β 和 $F_t^{-1}(\cdot)$ 都为 t 期时的工资。显然，工资方程(3-1)式右边的第一项表征了 $\ln w_{it}^1$ 的状态，可以看做 X_{it} 数量产生的效应；第二项 $X_{it}(\beta_t-\bar{\beta})$ 可表述成 $(\ln w_{it}^2-\ln w_{it}^1)$，它是特征回报变动的效应；而第三项 $[F_t^{-1}(\tau_{it}|X_{it})-\bar{F}^{-1}(\tau_{it}|X_{it})]$ 则可表述为 $(\ln w_{it}^3-\ln w_{it}^2)$，它反映了残差变动效应，Juhn *et al*. (1993) 将其归结为不可观测技能造成的效应。

[49] 例如，$\ln w_{1,90}$ 表示 $t=1$ 时期的工资分布 90 分位点的工资值，其余 $\ln w_{1,10}$，$\ln w_{0,90}$ 和 $\ln w_{0,10}$ 也是按类似方式定义。

分位的那些人拥有的 X。从 $t=0$ 期到 $t=1$ 期 90 分位与 10 分位之间的工资差异的变动可写作$(\ln w_{1,90}-\ln w_{1,10})-(\ln w_{0,90}-\ln w_{0,10})$，不同分位之间工资差异的变动可以进一步分解为个人特征数量变化的效应、个人特征价格变化的效应以及不可观测技能变化（包括不可观测技能数量变化和价格变化）的效应。

值得一提的是，在上述不同分位之间工资差异变动的分解中，我们仍无法在不可观测技能变化的效应中分离出数量变动效应和价格变动效应。鉴于此，Juhn *et al*.（1993）在经验分析中，使用了 cohort（同一批人）方法来控制不可观测技能分布的稳定性，从而解析出了不可观测技能的价格效应。

Juhn *et al*.（1993）实证发现，1963—1989 年间美国男性工资不平等程度在增长。通过对不同分位之间工资差异变动进一步分解表明，男性工资不平等的增长主要源自可观测技能（如教育、经验等）和不可观测技能（残差）回报的上升，而非教育和经验的数量，并且不可观测技能的价格变动比可观测技能价格变动所起的作用更大。

JMP1993 分解与 JMP1991 分解相比，虽然两者的出发点都是利用工资回归方程中的残差信息，但在 JMP1991 分解中，残差额被设定成标准差与标准化残差分布差异之乘积，并与不可观测技能的广义相对价格和数量差相对应；而在 JMP1993 分解中，残差则被设定成与个体在不可观测技能分布所处的分位相联系，进而展开在分位上（更一般地说是工资分布上）的工资差异分解。对残差的不同设定反映了不同的研究目的，JMP1991 分解关注的是组群间均值工资差异变动的成因，而 JMP1993 分解则是寻找同一组群的工资分布变动抑或工资不平等变化的成因。

JMP1993 分解将工资差异从均值分解真正扩展到了分布分解，是研究工资差异分布分解的一次重要突破。同时，JMP1993 分解存在的缺陷也是明显的：由于 JMP1993 分解的工资方程仍然使用基于 OLS 估计的经典线性回归模型，还不能自然拓展到分位上，而且各个反事实工资分布变动加总后也不是总的工资分布变动（Autor *et al*.，2005；Lemieux，2002）[50]。对这些缺陷改进的一种思路是采用条件分位回归模型，会在下文的 Q-JMP 分解中作详细介绍。

2. FL1998 分解

FL1998 分解是指 Fortin and Lemieux（1998）使用排位回归技术解析组群之间工资分布差异变动的方法。这是通过构造包括可观测技能和不可观测技能的（综合）技能指数，将个体在（综合）技能分布中的排位与其在工资

[50] 基于 OLS 的经典线性回归模型要求满足零条件均值假设，但在实际经验分析中，50 分位所对应的残差可能与残差均值不一致，即在 X_t 条件下 50 分位上的残差不一定为零，故 JMP1993 分解还无法实现精确的工资分布分解。

分布中的排位相对应，并利用有序 Probit 模型（ordered Probit model），从而把组群间工资分布差异的变动分解为技能分布（skill distribution，或称技能构成，skill composition）变动、工资结构（wage structure）变动和各组群与参照组相对位置变动三个部分的效应。

Fortin 和 Lemieux 首先提出了关于工资决定模型的两个基本假设[51]：工资结构变化对所有具有相同工资的工人的影响是相同的，工人在工资分布中的位置（或排位）是其自身技能的函数。在这两个假设下，个体的技能排位与其在工资分布中的位置（排位）具有一致性，即

$$\mathrm{rank}(\ln w_i) = \mathrm{rank}(r_i^*), \tag{3-3}$$

（3-3）式也可表示成 $F(\ln w_i) = F(r_j^*)$。其中，w_i 表示第 i 个工人的工资，r_j^* 表示第 i 个工人的（综合）技能指数；$F(\ln w_i)$ 和 $F(r_j^*)$ 分别表示工资和技能的累积分布函数。Fortin 和 Lemieux 之所以称自己的模型为排位回归模型，是因为在（3-3）式中并没有假定工资与技能指数之间存在显性关联，而仅假定一个工人在工资分布中的排位与其在技能分布中的排位相同。对于个体的技能指数，Fortin 和 Lemieux 进一步将其假定为可观测技能向量（个人可观测特征向量）的线性函数与不可观测技能之和，从而有

$$r_i^* = X_i b + \varepsilon_i, \tag{3-4}$$

其中，X_i 是可观测技能（教育、经验等）向量，b 是因子负荷向量（factor loading vector），或者技能“价格”，表征各类可观测技能在市场中的相对重要性[52]；ε_i 则表示不可观测技能，并假定 ε_i 满足独立同分布并服从标准正态分布[53]。不过，技能指数 r_i^* 是一个潜变量（latent variable），还需要通过转换模型（transformation model）[54]

$$\Lambda(\ln w_i) = X_i b + \varepsilon_i, \tag{3-5}$$

将技能指数与工资联系起来。其中，$\Lambda(\cdot)$是转换函数，而反函数 $\Lambda^{-1}(\cdot)$则是技能回报函数，$\ln w_i = \Lambda^{-1}(r_i^*)$表示技能 r_i^* 在劳动力市场中的价值。

Fortin 和 Lemieux 通过确定阈值（thresholds）将工资分布分割成 J 个有

[51] 假设一：工资结构变化对所有工资相同的工人的影响是相同的，不论性别如何。换言之，工资结构效应仅仅是工人在工资分布中位置或“排位”的函数。假设二：工人在工资分布中的位置是其自身技能的函数。这个假设秉承了工资决定的标准人力资本-竞争模型中的精神，在该模型中工资等于技能数量乘以技能价格。

[52] Fortin and Lemieux(1998)在文中采用 β 表示各种可观测技能的因子负荷向量，但由于本文已用 β 表示工资结构系数，为了避免混淆，这里用 b 表示因子负荷向量。

[53] 将综合技能函数残差标准化(均值为 0，方差为 1)是为了能够将工资分布变动中技能回报函数 $\Lambda^{-1}(\cdot)$ 的变化和综合技能指数自身的变化区分开来。

[54] 利用排位函数的单调性质，Fortin 和 Lemieux 对(3-3)式进行变换得 $r_i^* = \widetilde{F}^{-1}(F(\ln w_i)) \equiv \Lambda(\ln w_i)$。其中，$F^{-1}(\cdot)$ 和 $\widetilde{F}^{-1}(\cdot)$ 分别表示工资累积分布函数的反函数和综合技能累积分布函数的反函数。

序区间[55]，采用有序 Probit 模型估计转换模型的因子负荷向量 b，再借助阶跃函数（step function）得到转换函数 $\Lambda(\cdot)$和技能回报函数 $\Lambda^{-1}(\cdot)$的近似。[56]同时，他们还估计出个体进入某个工资区间的概率[57]，并把这个概率作为核密度估计中带宽的权重构造工资密度函数。[58]估算出了工资密度函数，就可获得工资分布函数。由于 $t(t=0,1)$时期某一组群 $g(g=H,L)$的工资分布可以看做 Λ_t^{-1}、X_t 和 b_t 的函数[59]，从而，t 时期 g 组群的工资分布函数可表述成

$$F_{gt}(\ln w) = F_{gt}(\Lambda_t^{-1};X_t;b_t). \tag{3-6}$$

$F_{gt}(\Lambda_t^{-1};X_t;b_t)$表示 g 组群带有协变量的无条件工资分布（unconditional distribution in the presence of covariates）。

为了分解从 $t=0$ 期到 $t=1$ 期的组群 g 工资分布的变动，定义组群 g 的三种反事实工资分布状态。

状态 1 $F_{g1}^{C1}(\ln w)\equiv F_{g1}(\Lambda_1^{-1};X_0;b_1)$，即 $F_{g1}^{C1}(\ln w)\equiv F_{g1}[F_{R1}^{-1}[F_{R1}(\Lambda_1^{-1};X_0;b_1)]]$。这是指组群 g 在技能回报函数和技能"价格"处于 $t=1$ 期，而技能数量却保持在 $t=0$ 期情形下的工资分布。$F_{R1}(\cdot)$表示参照组在 $t=1$ 期的工资分布；$F_{g1}[F_{Rt}^{-1}(\cdot)]$指 t 时期组群 g 与参照分布 $F_{Rt}(\ln w)$的相对位置函数，表示将参照分布的某个分位对应到 g 组群的分位点。[60]

---

[55] 合理选择区间数量 J 是非常关键的。J 太小，由阶跃函数（step function）估计得到的 $\Lambda^{-1}(\cdot)$ 不是精确近似；而 J 太大，又会错误估计 $\Lambda^{-1}(\cdot)$。因此，Fortin 和 Lemieux 选择可以使工资密度估计值与工资密度真实值相近的最小区间数量。

[56] 由于假定转换模型中的 ε_i服从 i. i. d 的标准正态分布，从而个体 i 进入某一区间$[a_{j-1}, a_j]$的概率为 $\Pr(a_{j-1}\leqslant \ln w_i\leqslant a_j)=\Pr(\Lambda(a_{j-1})\leqslant r_i^* < \Lambda(a_j))=\Pr(\Lambda(a_{j-1})\leqslant X_ib+\varepsilon_i<\Lambda(a_j))=\Phi(-X_ib+\Lambda(a_j))-\Phi(-X_ib+\Lambda(a_{j-1}))=\Phi(-X_ib+\lambda_j)-\Phi(-X_ib+\lambda_{j-1})$，其中，$j=1,2,\cdots,J$。记 $\Lambda(a_j)$为 λ_j，记 $\Lambda(a_{j-1})$为 λ_{j-1}；$a_0=\Lambda_0=-\infty$，$a_J=\Lambda_J=+\infty$；Φ 表示正态累积分布函数。通过上述等式构建有序 Probit 模型，由最大似然法得到 b 和 λ_{j-1}的估计值。转换函数为 $\Lambda(\ln w_i) = \sum_{j=0}^{J} I_{\{a_{j-1}\leqslant \ln w_i < a_j\}}\lambda_j$，其中 $I_{\{a_{j-1}\leqslant \ln w_i < a_j\}}$为指示函数，满足 $a_{j-1}\leqslant \ln w_i < a_j$ 时取 1，否则取 0。技能回报函数可通过转换函数的反函数获得：$\ln w=\Lambda^{-1}(r^*)=\sum_{j=0}^{J} I_{\{\lambda_j\leqslant r^* < \lambda_{j+1}\}}a_{j+1}$。由于估计值 λ_j对应临界值 a_j的转换函数值，故利用阶跃函数就可得到转换函数的近似 $\Lambda(\cdot)$ 和对应的技能回报函数 $\Lambda^{-1}(\cdot)$。

[57] $\pi_j^g = \pi_j(\Lambda_g^{-1},b^g,X^g) = \frac{1}{N_g}\sum_{i\in g}\{\Phi(-X_ib^g+\lambda_j^g)-\Phi(-X_ib^g+\lambda_{j-1}^g)\}$，$N_g$表示组群 g 中工人的数量。

[58] 核密度估计公式为 $f^g(\ln w) = \sum_{j=0}^{J}\frac{\pi_j^g}{h^*}K\left(\frac{\ln w-\ln w_j}{h^*}\right)$，其中，$\ln w_j$表示每个区间的工资均值。

[59] 根据工资决定模型 $\ln w_i^g=\Lambda_g^{-1}(X_ib^g+\varepsilon_i)$，组群 g 在 t 期工资水平直观上与 Λ_t^{-1}，X_t，b_t，ε_t四个因素相关。但由于 ε_t无法观测，在实际估算中 Λ_t^{-1}表述的是可观测技能和工资的关系，ε_t对工资影响落入 Λ_t^{-1}中，换言之，Λ_t^{-1}效应包括工资结构效应和不可观测技能效应。因此，组群 g 在 t 期的工资水平实际上就只与 Λ_t^{-1}，X_t，b_t三个因素相关联。

[60] $\tau_{gt}=F_{gt}[F_{Rt}^{-1}(\tau_{Rt})]$ 指 t 时期将参照分布上的分位 τ_{Rt}对应到 g 组群的分位 τ_{gt}。

状态 2　$F_{g1}^{C2}(\ln w) \equiv F_{g1}(\Lambda_1^{-1}; X_0; b_0)$，即 $F_{g1}^{C2}(\ln w) \equiv F_{g1}[F_{R1}^{-1}[F_{R1}(\Lambda_1^{-1}; X_0; b_0)]]$。它表示组群 g 在技能回报函数处于 $t=1$ 期，而技能构成（数量和价格）却保持在 $t=0$ 期情形下的工资分布。

状态 3　$F_{g1}^{C3}(\ln w) \equiv F_{g0}[F_{R0}^{-1}[F_{R1}(\Lambda_1^{-1}; X_0; b_0)]]$。这是指将技能构成和组群 g 与参照组群的相对位置保持在 $t=0$ 期时的工资分布。

从而，组群 g 从 $t=0$ 期到 $t=1$ 期的工资分布变动为

$$
\begin{aligned}
F_{g1}(\ln w) - F_{g0}(\ln w) = & [F_{g1}(\ln w) - F_{g1}^{C1}(\ln w)] \\
& + [F_{g1}^{C1}(\ln w) - F_{g1}^{C2}(\ln w)] + [F_{g1}^{C2}(\ln w) - F_{g1}^{C3}(\ln w)] \\
& + [F_{g1}^{C3}(\ln w) - F_{g0}(\ln w)],
\end{aligned}
\tag{3-7}
$$

其中，第一项$[F_{g1}(\ln w)-F_{g1}^{C1}(\ln w)]$表示由于技能数量变化引起的工资分布变动；第二项$[F_{g1}^{C1}(\ln w)-F_{g1}^{C2}(\ln w)]$表示由于技能“价格”变化引起的工资分布变动；第三项$[F_{g1}^{C2}(\ln w)-F_{g1}^{C3}(\ln w)]$代表参照组和组群 g 在扣除技能构成（数量和价格）变化对工资分布变动的影响后，剩余部分的相对位置变动对工资分布的影响，即相对位置剩余变动（residual improvements in the relative position of g）；第四项$[F_{g1}^{C3}(\ln w)-F_{g0}(\ln w)]$则表征由于工资结构变化引起的工资分布变动。

于是，从 $t=0$ 期到 $t=1$ 期的组群 H 与组群 L 之间工资差异的变动可分解成[61]

$$
\begin{aligned}
& [F_{1H}(\ln w) - F_{1L}(\ln w)] - [F_{0H}(\ln w) - F_{0L}(\ln w)] \\
& = \{[F_{1H}(\ln w) - F_{1L}(\ln w)] - [F_{1H}^{C1}(\ln w) - F_{1L}^{C1}(\ln w)]\} \\
& \quad + \{[F_{1H}^{C1}(\ln w) - F_{1L}^{C1}(\ln w)] - [F_{1H}^{C2}(\ln w) - F_{1L}^{C2}(\ln w)]\} \\
& \quad + \{[F_{1H}^{C2}(\ln w) - F_{1L}^{C2}(\ln w)] - [F_{1H}^{C3}(\ln w) - F_{1L}^{C3}(\ln w)]\} \\
& \quad + \{[F_{1H}^{C3}(\ln w) - F_{1L}^{C3}(\ln w)] - [F_{0H}(\ln w) - F_{0L}(\ln w)]\}.
\end{aligned}
\tag{3-8}
$$

（3-8）式的第一项是技能构成中的数量变动效应，第二项是技能构成中的价格变动效应，两者之和组成技能分布或技能构成变动效应；第三项是组群 L 相对位置剩余变动效应；第四项则是工资结构变动效应。

Fortin 和 Lemieux 的实证研究发现，尽管分解结果对参照分布（男性或者是全样本[62]）的选择是极为敏感的，但无论采用何种参照分布，女性相对于参照分布位置的剩余变动是美国性别工资差异下降的最重要因素。当然，这

---

[61] 当以 H 组群作为参照组时，由于 $F_{1H}^{C3}(\ln w)=F_{1H}^{C2}(\ln w)$，因此组群 H 不存在相对位置的变动。当以全样本作为参照组时，$t=1$ 期的 $F_{R1}(\ln w)$ 采用全样本真实的工资分布，而 $t=0$ 期的参照分布则采用 H 和 L 组群的加权分布，即 $\widetilde{F}_{R0}(\ln w) \equiv s_{H1}F_{H0}(\ln w)+s_{L1}F_{L0}(\ln w)$。

[62] FL1998 分解中所指的全样本是由男性和女性样本进行加权后组成。

种相对位置剩余的上升既可能来源于劳动力市场歧视的下降，也可能来源于女性不可观测技能的提高。

与 JMP1991 分解相比，FL1998 分解将组群间工资差异的变动从均值分解扩展到分布分解；并且，JMP1991 分解基于的是工资残差分布，而 FL1998 分解基于的则是真实的工资分布。与 JMP1993 分解关注同一组群工资分布的变动相比，FL1998 分解则聚焦于不同组群间工资分布差异的变动，同时还凸显了组群间相对位置的变动对组群工资分布差异的影响。如果说 JMP1993 分解是把个体的不可观测技能在组群中的排位与工资方程残差分布中的分位相对应，那么，FL1998 分解则将个体在技能分布中的排位与其在工资分布中的排位相对应，而后者的对应方式却可以避开对工资与技能之间关系作具体形式的（诸如线性关系之类）先验假定，这就有可能更好地揭示出工资与技能之间的真实关联。

尽管 FL1998 分解不必对技能回报函数（转化函数的反函数）作出具体形式的假定，但在估计转化模型时却暗含着因子负荷向量不随工资区间的改变而改变的假定，从而在某种程度上降低了分解模型的灵活性。同时，FL1998 分解在技能指数构造中还需要对不可观测技能作满足独立同分布并服从标准正态分布的假定。[63]

对于 FL1998 分解中因子负荷向量不随工资区间变化而仅有截距项随工资区间变动的局限，Chernozhukov *et al*.（2009）提出的基于分布回归（distribution regression）的分解方法则能有效地加以解决。Chernozhukov *et al*.（2009）对每一个工资值都用分离回归模型（separate regression model）估计，模型中的斜率参数和截距项都会随工资值的不同而不同。[64] 显然，这是一种更为灵活的分解方法，并能将最低工资等制度性因素纳入分析框架。Chernozhukov *et al*.（2009）运用该方法考察美国 1979—1988 年工资不平等变动时发现，劳动力个体特征分布变动是导致工资分布高端不平等变动的主要原因，而工会分布变动对解释美国总体工资不平等变动的作用并不大。

（二）基于半参模型的分布分解

JMP1993 分解的核心技术仍是 OLS 估计，要求工资方程满足经典线性回归模型假定，至于这样的假定到底距离现实多远，仍是一个值得深究的问题。而 DFL 分解的优点则是不需要对工资方程进行参数形式的设定，它利用半参模型，通过重置权重函数来构造反事实工资分布。Lemieux 分解在 DFL 分解

---

[63] JMP1991 分解和 JMP1993 分解对不可观测技能无需作正态分布假定，只要满足独立同分布就可以了。

[64] FL1998 分解中的条件工资分布为 $F(\ln w_i \mid X)=\Phi(-X_i b^g+\Lambda_g(a_j))$。Chernozhukov *et al*.（2009）中的条件工资分布为 $F(\ln w_i \mid X)=L(X_i b(\ln w_i))$，其中 $L(\cdot)$ 是关联函数（link function），可以是 logit、probit 等函数。

的基础上，借鉴 JMP1993 分解的思路，改进了 DFL 分解中没有直接分解特征回报变动效应的不足。

1. DFL 分解

DFL 分解是指 DiNardo *et al*.（1996）提出的这样一种方法，即用加权核密度估计（weighted kernel density estimation）刻画工资分布，并借助重置权重函数（reweighting function）构造反事实工资分布，从而以工资密度图可视化地呈现各因素对工资分布不同位置的影响。

具体而言，DiNardo *et al*.（1996）首先在 Rosenblatt（1956）和 Parzen（1962）的核密度估计量中引入与每一观测值相关的样本权重，形成加权核密度估计：

$$f_h(\ln w) = \sum_{i=1}^{n} \frac{\pi_i}{h} \left(\frac{\ln w - \ln w_i}{h} \right), \tag{3-9}$$

其中，f_h 是单变量密度 f 基于随机样本 $\ln w_1, \cdots, \ln w_n$ 的核密度估计；π_i 是权重，且 $\sum_{i=1}^{n} \pi_i = 1$；h 是带宽（bandwidth）；$K(\cdot)$是核函数（kernel function）[65]。

然后，DiNardo 等用重置加权函数构造反事实工资分布，并进行两个时期工资分布变化的分解。定义

$$\begin{aligned} f_t(\ln w) &= \int_{X \in \Omega_X} f(\ln w \mid X, t_w = t) \mathrm{d}F(X \mid t_X = t) \\ &\equiv f(\ln w; t_w = t, t_X = t) \end{aligned}$$

为 t 时期($t=0,1$)的工资密度。其中 $t_w=t$，$t_X=t$ 分别表示工资结构（特征回报）和个体特征所处的时期，$F(X \mid t_X=t)$表示 t 时期个体特征 X 的分布。再定义反事实工资分布 $f(\ln w; t_w=1, t_X=0)$为“保持 $t=0$ 时期的个体特征，而工资结构为 $t=1$ 时期的工资密度”。假定 $t=1$ 时期工资结构不依赖于个体特征 X，也就是忽略在一般均衡中个体特征分布变化对工资结构影响，则该反事实工资分布可表述成

$$\begin{aligned} f(\ln w; t_w = 1, t_X = 0) &= \int (\ln w \mid X, t_w = 1) \mathrm{d}F(X \mid t_X = 0) \\ &= \int (\ln w \mid X, t_w = 1) \psi_X(X) \mathrm{d}F(X \mid t_X = 1), \end{aligned} \tag{3-10}$$

式中 $\psi_X(X)$ 是重置权重函数，$\psi_X(X) = \dfrac{\mathrm{d}F(X \mid t_X=0)}{\mathrm{d}F(X \mid t_X=1)} = \dfrac{\Pr(t_X=0 \mid X)}{\Pr(t_X=1 \mid X)} \cdot$

[65] 在 DiNardo *et al*.(1996)中，π_i是 CPS(Current Population Survey)采样权重乘以日常工作小时数并标准化成和为 1 的权数。核密度估计中的一个关键问题是带宽的选择。DiNardo 等用 Sheather and Jones (1991)的方法选择最优带宽，提出对于以工作小时数为权重的密度估计的最优宽度在 0.05—0.08 之间。核函数则采用高斯核函数。

$\frac{\Pr(t_X=1)}{\Pr(t_X=0)}$。其中，$\Pr(t_X=1|X)$表示给定个体特征 X 下 $t_X=1$ 的概率。从而，在 $t=0,1$ 两个不同时期工资密度变化的可分解为

$$\begin{aligned} f_1(\ln w)-f_0(\ln w) &\equiv f(\ln w;t_w=1,t_X=1)-f(\ln w;t_w=0,t_X=0) \\ &=[f(\ln w;t_w=1,t_X=1)-f(\ln w;t_w=1,t_X=0)] \\ &\quad+[f(\ln w;t_w=1,t_X=0)-f(\ln w;t_w=0,t_X=0)]. \end{aligned} \tag{3-11}$$

分解式中的第一项表示因个体特征 X 分布变化造成的工资分布变动，可称为构成效应（composition effect，构成效应）；第二项代表由于工资结构变化引起的工资分布变动，称为工资结构效应（wage structure effect）。

接着，DiNardo 等进一步考察了两个制度性因素——工会化水平（unionization）和最低工资制度对工资分布的影响。[66]

最后，DiNardo 等将各种因素对工资密度变化的影响整合在同一个式子中，形成 DFL 分解。这些因素包括最低工资、工会化水平、其他个体特征以及供求因素，从而，在 $t=0$ 和 $t=1$ 期之间的工资密度变化可写成

$$\begin{aligned} f_1(\ln w)-f_0(\ln w) =&[f(\ln w;t_w=1,t_z=1,m_1,d_1,s_1) \\ &-f(\ln w;t_w=1,t_z=0,m_0,d_0,s_0)] \\ &+[f(\ln w;t_w=1,t_z=0,m_0,d_0,s_0) \\ &-f(\ln w;t_w=0,t_z=0,m_0,d_0,s_0)]. \end{aligned} \tag{3-12}$$

（3-12）式实际上就是（3-11）式的扩展式，工资密度的变化可以像 Oaxaca-Blinder 分解那样，解析为构成效应和工资结构效应两部分。其中，z_t、m_t、d_t 和 s_t 分别是第 $t(t=0,1)$期的包括是否参加工会在内的个人特征向量、最低工资、劳动供给和需求特征。

DFL 分解还可以继续将构成效应分解到各个对应因素上，但此时的工资结构效应已被归入了残差因素（residual factors）的效应。DiNardo 等也没有再进一步区分劳动力市场不同要素的价格变动对于工资不平等的影响。

值得注意的是，DFL 分解中存在顺序分解（sequential decomposition，或译为序贯分解）问题，即给定因素对工资分布变化的影响大小还与分解的顺序（ordering）相关，放置在分解顺序前面的因素往往会增大其影响程度或效应[67]。正因为如此，DiNardo 等还将分解因素的顺序颠倒，以确保自己并没

[66] 分析工会化水平对工资分布的影响，是在假定条件密度 $f(\ln w|u,X,t_w)$ 与工会化率（unionization rate）不相关的情况下，考虑“劳动者个体是否参加工会”与个体特征 X 的相关性。在探讨最低工资制度对工资分布的影响时，DiNardo 等设定了以下三个假设：(1) 最低工资对高于最低工资的那部分工资分布没有外溢效应（spillover effect）；(2) 在最低工资及以下的实际工资条件密度的形状只取决于实际最低工资；(3) 最低工资不影响就业概率。显然，在这三个假设下，基于 $t=0$ 期最低工资的 $t=1$ 期反事实工资条件密度是一个“拼接”的分布。

[67] 这主要是由于不同协变量之间的交互作用所带来的。

有过分强调欲考察的因素（主要指劳动力市场制度因素）的影响程度。

DiNardo *et al*.（1996）实证表明，制度因素（包括工会化水平和实际最低工资）和供求冲击都是造成 1979—1988 年间美国工资不平等程度上升的重要原因。最低工资的下降压低了工资分布的低端（compress the lower tail），尤其是对女性而言。更一般地说，DFL 分解能够直观明晰地显示各个变量对工资分布的哪个区位上施加了最大的影响，而这用以往的研究方法是很难做到的。

在考察某些制度性因素对工资分布变动的影响时，DFL 分解具有较强的诠释能力。例如，DFL 分解在解析最低工资变动对不平等影响时，不仅考察了最低工资线及以下人数的变动对不平等的影响，而且还涵盖了最低工资值的变动对工资不平等的影响。[68]再者，DFL 分解还能够将工资分布分解一般化到各种工资不平等统计量的分解。这是因为它得到了工资密度函数，将其函数变换后可简便地得到一系列表征工资不平等的统计量，诸如基尼系数、泰尔指数等。不过，DFL 分解探讨单一解释变量（协变量）变动对工资分布变动影响时，要求该解释变量是离散变量。若是连续变量则还需要进行离散化处理，否则就难以将工资分布的变动分解到各个因素上，但在经验分析中将类似工作经验年数等协变量转化成离散变量会带来大量运算（参见 Firpo *et al*.，2007b）。

2. Lemieux 分解

Lemieux 分解是指 Lemieux（2002）在 OLS 回归的基础上，通过把样本按照个体特征划分为若干单元（cell）的形式经过重置权重构造反事实工资，进而解析不同组群工资分布差异或同一组群不同时期工资分布变动成因的方法。这种方法有时也称为基于个体特征单元（cell-based）的分解方法。

具体地说，Lemieux 首先通过对各个特征变量分组将样本个体特征向量 X_{it} 划分组合为 J 个单元，从而，$X_{it}=[X_{i1t},\cdots,X_{ijt},\cdots,X_{iJt}]$。其中，$X_{ijt}$ 表示无遗漏个体特征信息的虚拟变量，$j=1,2,\cdots,J$。显然，每一个个体 i 都有其归属的单元。[69]记 ω_{it} 为个体 i 在采样中的权重，表示个体 i 在样本中的代表性程度；再记 P_{jt} 为 j 单元中样本数占总样本数的比重，由于 $\overline{X}_{jt}=\sum_{i}\omega_{it}X_{ijt}=\sum_{x_{ijt}=1}\omega_{it}=P_{jt}$，说明样本比重 P_{jt} 与虚拟变量 X_{ijt} 的平均值 $\overline{X}_{jt}$ 相等，因而 P_{jt} 完整表达了特征

[68] 在 JMP1993 分解中工资分布是由工资均值和残差拟合的间接估计，最低工资的影响无法从工资均值的变动中直接分离出来，故 JMP1993 分解中只能通过放入虚拟变量（个体是否处于最低工资之下）来控制最低工资线及以下的人数变动对于工资不平等变动的影响。

[69] 按照每个特征变量的特性先各自分组，然后把所有组进行排列组合，每个组合就是一个单元。比如，现在有两个个体特征变量教育水平和性别，教育水平分为 7 类，性别分为 2 类，那么一共就有 14 个单元。个体特征的类别均用虚拟变量表示。若个体 i 属于第 k 类，则有 $X_{it}=[X_{i1t},\cdots X_{ijt},\cdots X_{ikt}.,\cdots X_{iJt}]=[0,\cdots 0,\cdots 1,\cdots 0]$。

分布构成，P_{jt} 的变动也反映了特征分布的变动。

其次，Lemieux 用 OLS 估计工资方程得到的回归系数和基于单元的重置权重，构造两种反事实工资 $\ln w_{it}^{a}$ 和 $\ln w_{it}^{b}$。$\ln w_{it}^{a}$ 表示个体特征回报（回归系数）发生变动时的工资，$\ln w_{it}^{b}$ 表示个体特征回报与特征分布发生变动时的工资。记 $t=0$ 时期工资方程为 $\ln w_{i0}=X_{i0}\beta_0+u_{i0}$。利用该工资方程中的系数向量，可构造个体特征回报为 $t=0$ 期的反事实工资

$$\ln w_{i1}^{a} = X_{i1}\beta_0 + u_{i1}. \tag{3-13}$$

基于 $\ln w_{i1}^{a}$ 构造特征回报和协变量（个体特征）分布为 $t=0$ 时期的反事实工资为

$$\ln w_{i1}^{b} = \psi_i \omega_{i1} \ln w_{i1}^{a} = \psi_i \omega_{i1}(X_{i1}\beta_0 + u_{i1}) \equiv \psi_{i1}^{a}(X_{i1}\beta_0 + u_{i1}), \tag{3-14}$$

其中，ω_{i1} 是 $t=1$ 时期个体 i 的样本权重，$\psi_i = \sum_j X_{ij1}P_{j0}/P_{j1}$ 是重置权重因子，表示通过借助各个特征单元的权重变化对个体重置权重；$\psi_{it}^{a}=\psi_i\omega_{i1}$ 是反事实权重因子，表示对 $\ln w_{i1}^{a}$ 赋予反事实权重因子，使得协变量（个体特征）分布变动到 $t=0$ 时期。

得到个体的反事实工资 $\ln w_{it}^{a}$ 和 $\ln w_{it}^{b}$ 后，就可通过核密度估计得到反事实工资密度函数。于是，从 $t=0$ 期到 $t=1$ 期的工资分布变动可以分解为

$$\begin{aligned} f(\ln w_1) - f(\ln w_0) = & [f(\ln w_1) - f(\ln w_1^{a})] \\ & + [f(\ln w_1^{a}) - f(\ln w_1^{b})] + [f(\ln w_1^{b}) - f(\ln w_0)]. \end{aligned} \tag{3-15}$$

分解式的第一项表示回归系数（特征回报）变动效应，第二项表示特征分布变动效应，第三项为残差变动效应。

与 JMP1993 分解相比，Lemieux 分解能够更完整地分离协变量分布变动的效应。[70] 在构造反事实工资 $\ln w_{it}^{a}$ 时，Lemieux 分解与 JMP1993 分解相同，通过变换 OLS 估计工资方程的特征回报 β_t 来构造。然而，在构造反事实工资 $\ln w_{it}^{b}$ 时，JMP1993 分解直接采用变换工资方程的 X_t 构造，也就是仅变换了 $\ln w_{i1}^{a}$ 中的 $X_{i1}\beta_0$，而没有对 u_{i1} 进行变换。[71] Lemieux 指出，个体特征分布变动将通过两个不同途径影响工资分布（可用工资方差表示）[72]：一是直接通过协变量的方差-协方差矩阵 $\Omega_{X,t}$ 变动的效应，二是特征分布构成变动以影响残差

[70] 当工资方程满足经典线性假定时，JMP1993 分解能够简便而准确地解析出工资分布变动的成因。但在现实中尤其是在工资分布分解中，经典线性假定通常难以满足，往往会出现诸如异方差等问题，导致 JMP1993 分解的残差分布的变动效应包含了协变量分布变动的影响，使协变量分布变动效应无法完整解析出来，同时还引发工资分布分解次序问题。

[71] 在 JMP1993 分解中，所构造特征回报和协变量分布为 $t=0$ 时的反事实工资为 $\ln w_{i1}^{b*}=X_{i0}\beta_0+u_{i1}$。

[72] 在 $\ln w_{it}=X_{it}\beta_t+u_{it}$ 中，不同时期工资方差变动源于三个部分：一是估计系数变动，二是协变量的协方差矩阵（包括协变量自身方差以及与其他变量的协方差）变动，三是残差方差变动。

分布间接影响工资分布。举例来说，如果更有经验的工人可能出现更大的残差方差时，工作经验年数的上升不但会直接影响工资方差，还会通过增加残差方差使得工资方差上升。鉴于此，Lemieux 通过基于单元的反事实重置权重 ψ_{i1}^{a} 对 $\ln w_{i1}^{a}$ 整体进行了变换，构造出更为贴合现实要求的 $\ln w_{i1}^{b}$，解决了 JMP1993 分解无法完全解析个体特征分布变动效应的问题。

在数据类型合适的情况下，比如调查数据本身以单元编码[73]，Lemieux 分解还可以弥补 DFL 分解对价格效应处理的不足。[74]值得注意的是，基于单元的 Lemieux 分解在研究不同时期工资方差变动时是具有优势的，而结合了反事实重置权重技术后，该方法还可以分解其他工资分布统计量的变动，比如 90—10 分位差，GINI 等。此外，Lemieux 分解还可以应用于残差分布变动的分解。

Lemieux（2002）开展了两个实证分析，其中一个是关于加拿大的两个省（Alberta 和 British Columbia）工资分布差异的分解，结果显示这两个省工资分布的形状差别很明显，Alberta 省是一个工资总体水平高，且工资离散度低的省份。与之相反，British Columbia 省是一个工资总体水平低，但工资离散度却较高的省份。这两个省份工资分布的高端部分差异主要源自两者残差差异，而工资分布的低端部分差异主要来自于回归系数差异。特征分布和残差差异缩小了两个省份的工资分布的低端部分差异。另一个实证研究是有关美国工资变动的分解。经验研究表明 1973—1999 年工资残差方差上升主要是由劳动力构成效应导致的。然而，在 1979—1989 年间工资残差方差上升则是归因于不可观测技能价格上升。同时，特征构成回报、不可观测技能价格和特征分布变动不能很好地解释工资变动的非线性部分。特别是，无法捕捉最低工资制度对工资分布低端的影响。这种情况在 20 世纪 80 年代美国女性工资分布变动中特别明显。

不过，Lemieux 基于特征单元的分解方法常常会面临数据不支持问题。Lemieux 指出，从估计的角度出发，将数据按照教育和经验年数主观地分为不同单元存在局限。虽然这种局限可以通过设定较小的单元来解决，比如按照教育和经验年数逐年构建若干个单元。然而，在经验分析中，实证数据往往无法达到该要求，可能出现某些单元没有样本信息的问题。这时，就仍然需要求助于 DFL 分解的重置权重技术，解决基于单元的工资分布分解中样本信息不够的问题。[75]

[73] 加拿大 LES 数据就是该种类型的数据，Lemieux（2002）利用该数据对 Alberta 省和 British Columbia 省工资分布差异进行了分析。

[74] DFL 分解将工资分布变动扣除协变量分布变动效应后作为剩余项，并没有明确分离出价格效应。同时，它仅对工资总体不平等变动进行了分析，而未利用重置权重的方法分析工资残差不平等的变动。

[75] 当数据不合适或不可行时，基于单元的重置权重计算就无法实施。Lemieux 提出借助 DFL 分解中的重置权重因子计算方法，再结合经典线性回归模型分解工资分布变动，也就是借助 logit 或者 probit 模型扩展基于经典线性回归模型的分解方法。

Lemieux（2006）采用 Lemieux（2002）提出的方法研究了 1973—2003 年美国工资残差不平等的变动。结果表明，该期间工资残差不平等上升主要来源于劳动力构成效应。这主要是基于更高教育水平和更有经验劳动力的工资会出现更大离散，而随着劳动力市场中此类工人人数增加就会导致工资残差不平等上升。也就是说，当不可观测因素存在异方差时，工资残差变动中的劳动力构成效应是不能忽视的。

（三）基于条件分位回归的分布分解

一般说来，随机变量 Y 的分布函数定义为 $F(Y)=\Pr(Y\leqslant y)$，Y 的 τ 分位数函数 $Q(\tau)$ 定义为 $Q(\tau)=\inf\{Y:F(Y)\geqslant\tau\}$，$\forall\tau\in(0,1)$。在参数线性假设下，Koenker and Bassett（1978）提出了如下估计参数的形式[76]：

$$\min_{\beta}\Big[\sum_{\{i:Y_i\geqslant X_i\beta\}}\tau\,|Y_i-X_i\beta|+\sum_{\{i:Y_i<X_i\beta\}}(1-\tau)\,|Y_i-X_i\beta|\Big], \tag{3-16}$$

其中，Y_i 是被解释变量，X_i 是解释变量（协变量，covariance）的向量，τ 是估计中所取的各分位值 $\tau\in(0,1)$，β［或写成 $\beta(\tau)$］是各分位点估计系数向量。对于回归线（面）上方的点，赋以 τ 的权重，而对位于回归线（面）下方的点，则赋以（$1-\tau$）的权重，然后将离差绝对值加权求和并使其最小化。可见，（3-16）式是加权最小绝对离差（weighted least absolute deviations，WLAD，间或称为 LAD）。最小化该目标函数就可得到分位数估计

$$Q_\tau(Y\,|\,X)=X\beta(\tau). \tag{3-17}$$

在不同分位上得到不同的回归参数，可以推断为被解释变量条件分布在不同分位上对解释变量变动的反应存在差异。经典线性回归模型利用的是样本的平均信息，描述的是解释变量对被解释变量均值的影响，或者说是被解释变量关于解释变量的条件期望，相应的回归系数表达的是解释变量对被解释变量的平均边际效果（局部效应，partial effect）。但在现实中，我们不仅关心样本的平均特征，而且也在意因变量分布的局部特征和信息。例如，最近媒体流行的我国职工工资“被平均”就是这种情况的典型反映。然而，分位数回归（这里特指条件分位数回归，conditional quantile regression）则可针对被解释变量条件分布的任何一个分位展开估计，从而更加细致全面地揭示被解释变量条件分布与解释变量之间的关系。[77]

---

[76] 虽然 Koenker and Bassett(1978)自己称这是分位数回归(quantile regression/regression quantiles)，这其实是有条件的分位回归(conditional quantile regression)。

[77] 与普通最小二乘法相比，分位数回归能够：(1) 刻画解释变量对被解释变量条件分布的位置和形状的影响，揭示更多的信息。(2) 能够估计具有异方差性的模型。(3) 对异常值有更强的包容性，从而估计更加稳健。当数据出现异常值时，最小绝对离差比最小离差平方和更能削弱异常值对回归模型的影响。(4) 在误差分布非正态的情形下，分位数回归比最小二乘法估计更加有效(Buchinsky and Hahn，1998)。总之，分位数回归的假设条件更宽松，估计结果更稳健，挖掘信息也更丰富。

不过，条件分位数回归仍有其局限，主要是条件分位的偏效应只能表示解释变量对被解释变量条件分布影响的程度，并不能展现解释变量对被解释变量分布的效应，不能反映解释变量分布变动对被解释变量分布变动的边际影响，从而无法将被解释变量的分布直接分解到各个解释变量上。因此，我们需要思考如何将条件分位回归模型得到的参数估计运用于构建带有解释变量（协变量）的无条件反事实分布，进而开展工资分布变动的分解。可以说，MM2005 分解和 Q-JMP 分解在相当大程度上就是在做这些工作。

1. MM2005 分解

MM2005 分解是指 Machado and Mata（2005）所使用的分解工资分布变化中各因素贡献的方法。这种方法是在用条件分位回归（conditional quantile regression）估计工资条件分布的基础上，借助概率积分转换得到工资边际密度函数的一致估计，再通过随机替换构造所需的反事实工资分布，以解析同一组群不同时期工资分布变动的成因。

Machado 和 Mata 首先循着 Koenker and Bassett（1978）提出的条件分位回归的思路，用最小化绝对离差估计（LAD）方法估计 τ 分位上的回归模型 $Q_\tau(\ln w|X)=X\beta(\tau)$，从而得到以 X 为条件的工资分布。

可是，条件分布并不能反映协变量 X 分布变化对工资分布的影响，从而条件分布还不能直接用于工资分布变动的分解，须将其转换为无条件工资分布或者工资边际密度函数（marginal density function，边缘密度函数）。[78] 为此，Machado 和 Mata 通过抽取服从 $U[0,1]$ 均匀分布的 τ，并利用概率积分变换定理和条件分位函数与总体分位函数的一致性[79]，形成带有协变量的工资边际密度函数[80]，用于构造所需的反事实工资分布。

这个变换过程大致可分为以下四个步骤：（1）从均匀分布 $U[0,1]$ 中生成样本量为 m 的随机样本 $u_1,\cdots,u_m$，即随机生成服从均匀分布的 τ 值，m 是指 τ 的个数。（2）利用 t 时期数据 X_t（$n_t\times k$ 的协变量矩阵）对每个 $\{u_i\}$ 估计 $Q_{u_i}(\ln w|X_t)$，做 m 次分位回归，从而产生 m 列分位回归估计系数 $\beta_t(u_i)$。（3）从特征矩阵 X_t 中随机抽取 m 行，表示为 $\{X_{it}^*\},i=1,\cdots,m$。[81]（4）最后根据 $\ln w_i^*(t)\equiv X_{it}^*\beta_t(u_i)$ 获得 $\{\ln w_{it}^*\}_{i=1}^m$。$\{\ln w_{it}^*\}_{i=1}^m$ 是满足所要求分布的样本量为 m 的随机样本。利用该样本就可以得到 t 时期（带有协变量的）工资边际

---

[78] 只有当总体中所有劳动力拥有相同的可观测特征时，工资的条件分布才同时又是工资的无条件分布。

[79] 概率积分变换定理：如果 U 是 $[0,1]$ 均匀分布的随机变量，则 $F^{-1}(U)$ 的分布是 F。

[80] 工资边际密度函数 $f(\ln\hat{w})=\iint\limits_{X,\tau}\hat{Q}_\tau(\ln w|X)g(X)f(\tau)\partial X\partial\tau$。因 τ 是从 $[0,1]$ 均匀分布中抽取的随机变量，$f(\tau)=1$，从而上式可简化为 $f(\ln\hat{w})=\iint\limits_{X,\tau}\hat{Q}_\tau(\ln w|X)g(X)\partial X\partial\tau$。（参见 Autor *et al.*，2005）。从上式可知，工资边际密度函数本身并不包含协变量，但正文中之所以用"生成带有协变量的工资边际密度函数"，是为了凸显协变量对工资边际密度函数的影响。

[81] 此处是一个有放回的重复随机抽样过程。

密度函数 $f^*(\ln w_t;X_t)$。[82]

记 $f^*(\ln w_1;X_1)$和 $f^*(\ln w_0;X_0)$分别是 $t=1$ 和 $t=0$ 期的工资边际密度函数。[83]若将上述过程的步骤（3）中的 X_1 替换为 X_0，就可以得到第一种反事实的工资边际密度 $f^*(\ln w_1;X_0)$。它表示特征分布保持在 $t=0$ 期，而特征回报为 $t=1$ 期的反事实工资边际密度。于是，从 $t=0$ 期到 $t=1$ 期工资分布的变动可分解为

$$\begin{aligned}
&\nu[f(\ln w_1)]-\nu[f(\ln w_0)] \\
&\quad = \nu[f^*(\ln w_1;X_1)]-\nu[f^*(\ln w_0;X_0)]+\text{residual} \\
&\quad = \{\nu[f^*(\ln w_1;X_1)]-\nu[f^*(\ln w_1;X_0)]\} \\
&\qquad +\{\nu[f^*(\ln w_1;X_0)]-\nu[f^*(\ln w_0;X_0)]\}+\text{redidual}, \quad (3\text{-}18)
\end{aligned}$$

其中 $\nu(\cdot)$表示分布的统计量（比如分位等）。分解式右边的第一项和第二项分别是协变量分布变化（个体特征分布变化）的效应和条件分位回归系数变化（特征回报变化）的效应[84]，而第三项则是剩余项（residual），它是由于工资边际密度估计误差所致。[85]

为了进一步从第一项构成效应中解析出某个协变量（或称特征，用 z 表示）的贡献，或者说进一步考察某个协变量分布变动对工资分布的影响，则还需要通过重置权重构建第二种反事实工资边际密度，即仅有协变量 z 的分布发生变化而其他协变量 X_{-z} 保持不变时的反事实工资边际密度。[86]具体步骤如下：(1) 根据上述的(1)—(4)过程，生成 $t=1$ 时期样本量为 m 的工资随机样本$\{\ln w_{it}^*\}_{i=1}^m$。(2) 假设某个协变量 z 可以分成 J 类，用 C_{jt} 表示其中的第 j 类。[87] (a) 从 $t=1$ 期的工资样本$\{\ln w_{i1}^*\}_{i=1}^m$中选取第 j 类子样（subsample）组成$\{\ln w_{i1}^*\}_{i\in I_j}$，其中 $I_j=\{i=1,\cdots,m \mid z_{i1}\in C_{j1}\}$。(b) 从 $t=1$ 期的第 j 类子样中生成随机样本，使其样本量大小与 $t=0$ 期第 j 类子样相同，即 $m\times f_{j0}$[88]，从而替换$\{\ln w_{i1}^*\}_{i\in I_j}$。(3) 对于协变量 z 的其余每一类，重复步骤（2）。由此获得协变量 z 分布保持在 $t=0$ 期，而其余协变量分布以及特征回报为 $t=1$ 时

---

82 笔者以为，考虑到要构造反事实工资边际密度函数，用 $f^*(\ln w;X_t,\beta_t)$ 的形式或许能够更清晰地表述 t 时期的工资边际密度函数。而正文中之所以仍采用 $f^*(\ln w_t;X_t)$，是为了与 Machado and Mata(2005) 的表达保持一致。

83 直接从工资数据出发也可以估计 t 期的无条件工资分布，但这个分布中由于不含变量 X 的信息，故而无法利用它来构造反事实分布。

84 这两项分别相当于本文前述分解中的构成效应和工资结构效应。

85 Machado 和 Mata 在分解中并未对这个剩余项作进一步处理。但其实证结果显示，剩余项对工资分布变动影响不可忽视。

86 式中 X_{-z}表示除了协变量 z 之外其他协变量组成的矩阵。

87 如果 z 是连续性变量，则需要将其转化为离散变量。

88 f_{jt}表示 t 期第 j 类子样占总体的比重。

期的工资样本[89]，用于构造反事实的工资边际密度 $f^*(\ln w_1;X_{-z,1},z_0)$。于是，当其他协变量 X_{-z} 保持不变，仅有协变量 z 的分布发生变化对工资分布变动的贡献可由下式得到：

$$\nu[f^*(\ln w_1;X_{-z,1},z_1)]-\nu[f^*(\ln w_1;X_{-z,1},z_0)]. \tag{3-19}$$

总之，MM2005 分解在条件分位回归模型的基础上，利用概率积分变换得到带有协变量的工资边际密度估计。然后，通过随机抽样和替换过程构造反事实工资分布，就能够对同一组群不同时期工资分布的变动进行分解。此外，如果再利用无条件的重置权重过程，就可构造仅有某个协变量分布发生变动的反事实分布，从而进一步解析出该协变量的构成效应。[90]在实证上，Machado and Mata（2005）据此能够识别出大多数国家的工资不平等增加的源泉。他们将这些方法运用于葡萄牙 1986—1995 年的数据发现，教育水平的上升是造成工资不平等加剧的重要原因。

MM2005 分解固然与 JMP1993 分解的研究对象相同，但基于条件分位回归的 MM2005 分解却能够解决 JMP1993 分解中存在的两大缺陷[91]。更为重要的是，MM2005 分解为工资差异分布分解提供了一个统一和一致的分析框架，并成为研究工资分布差异和工资不平等变动的常用工具。

同时，MM2005 分解还是一个可供进一步拓展和深化研究的基础平台。例如，Melly（2005）解决了 MM2005 分解中存在不同分位回归线交叉的问题，并将 MM2005 分解中的“特征回报变动的效应”进一步细分为中位数系数变动（changes in median coefficient）效应和残差变动（changes in residual）效应。而 Autor *et al*.（2005）就是循着 Melly 的这条思路形成了本文接着将要介绍的 Quantile-JMP 分解（简称 Q-JMP 分解）。又譬如，Melly（2006）指出，MM2005 分解所利用的概率积分变换并不能够确保条件分位函数估计与总体分位函数估计是一致的。为此，Melly 提出了一种更为有效的分位回归估计方法，并且证明，只有当 $m\to\infty$ 时，MM2005 分解的估计结果才

---

89 上述过程表明当某个协变量 z 分布保持 $t=0$ 期状态，特征回报为 $t=1$ 时期的反事实工资分布为 $\int f_1(\ln w\,|z)\mathrm{d}F_0(z)$。其中 $f_1(\ln w|z)$ 表示 $t=1$ 期给定 z 条件时的条件工资密度，$F_0(z)$ 表示 $t=0$ 期的协变量 z 分布。对于分类变量 z 而言，其频率可以刻画出自身的分布状况。通过对 $t=1$ 期分类变量 z 频率重置权重，就可以得到与 $t=0$ 时期一致的协变量分布。

90 在 Autor *et al*.(2005)看来，MM2005 分解能将由 JMP1993 分解引入分析不平等的“全变异解释”(full variance accounting)技术与由 DFL 分解(DiNardo *et al*.，1996)提出的核重置权重(kernel reweighting)相关联。

91 JMP1993 分解中存在的两大缺陷分别是：第一，JMP1993 分解是建立在 OLS 估计工资方程的基础上，它提供的是工资分布的条件均值估计。除非残差项满足同方差对称分布的假定，否则，该模型就不能自然拓展到工资分布的各分位上。第二，JMP1993 分解各个反事实分布变动加总后无法得到总的工资分布变动。具体地说，通过调整 $t=0$ 时期工资分布的三个部分($\Delta X,\Delta\beta,\Delta F^{-1}(\cdot|X)$)无法再现 $t=1$ 时期的工资分布。因为 $X_t\beta_t$ 和 $F^{-1}(\cdot|X_t)$ 是随机变量，而 $F(\ln w_{it})$ 是两者加和的分布，所以 $F(\ln w_{it})$ 不但取决于 $X_t\beta_t$ 和 $F^{-1}(\cdot|X_t)$ 的各自分布，还取决于它们的联合分布(协方差)，然而 JMP1993 分解中没有计算协方差对工资分布的影响(参见 Autor *et al*.，2005)。

与 Melly（2006）的估计结果相同。

MM2005 分解也存在着其他一些不足。首先，也是最重要的，MM2005 分解在将构成效应分解到单一协变量过程中，提出的无条件重置权重方法仍有待商榷。比如考察工会的影响，Machado 和 Mata 提出保持其他因素不变，重置工会的频率，就可得到仅有工会分布变动的反事实分布。可是，在经验分析中无条件重置权重的过程却可能会改变某些与工会变量相关的协变量分布，从而无法准确估计工会对工资分布变动的影响（参见 Firpo *et al.*, 2007b）。[92] 其次，MM2005 分解需要在许多分位上展开反事实的拟合计算，这会带来大量的运算过程。此外，条件分位回归的估计量只有在分位函数设定正确的前提下才满足一致性，而在经验分析中要判断模型设定正确与否却是困难的。不过，Machado and Mata（2005）已坦承自己的这一局限。

2. Q-JMP 分解

Quantile-JMP 分解（简称 Q-JMP 分解）是指 Autor *et al.*（2005）和 Melly（2005）提出的这样一种方法：通过在 MM2005 分解的分位分解技术（quantile decomposition technique）平台上展开 JMP1993 分解的思路，将同一组群的工资（总体）分布变动解析为个体特征分布变动、组间价格（between-group price）变动和残差价格（residual price，亦称组内价格，within-group price）变动三部分效应。

众所周知，MM2005 分解是以条件分位回归为基础将工资（总体）分布变动归因于劳动者个体特征分布变动、特征价格（特征回报，亦即分位回归系数向量）变动以及由工资边际密度估计误差所致的剩余项[93]。但在 Autor 等看来，这个剩余项实际上往往很小，于是，工资分布变动就可由个体特征分布变动和特征价格变动两项来刻画。[94] 更重要的是，Autor 等对工资分布变动的分解并没有仅仅停留于此，而是继续循着 Melly（2005）的做法，借鉴 JMP1993 分解的思路[95]，将个体 i 在 t 时期的工资分位回归方程变形为

$$\begin{aligned} Q_\tau(\ln w_{it} \mid X_{it}) &= X_{it}\beta_t(\tau) = X_{it}\beta_t(50) + [X_{it}\beta_t(\tau) - X_{it}\beta_t(50)] \\ &\equiv X_{it}\beta_t^b + X_{it}\beta_t^w, \end{aligned} \tag{3-20}$$

其中，$\beta_t(\tau)$是用条件分位回归估计 $t(t=0,1)$ 时期各个分位点工资方程的回归

92 要获得构成效应中单一协变量的影响，在重置权重时应该采用 DFL 分解所提出的条件重置权重函数。不过，条件重置权重的分解当面对协变量不是离散变量时也会遇到困难。

93 参见公式(3-18)中的“residual”项。

94 若条件分位回归模型能够完美刻画条件工资分布，则 MM2005 分解中的剩余项就变成零。此外，Autor *et al.*(2005)利用美国 CPS 数据的实证研究也表明，真实工资分布与基于分位回归模型拟合的工资分布在各种统计量上的差异都非常小。

95 JMP1993 分解的基本思路是，用工资集中趋势和残差分布来表征工资总体分布，并将工资总体分布变动分解为特征分布变动、特征回报(组间价格)变动和残差分布变动三部分。若假定不可观测技能分布不随时间改变，则残差分布的变动就可推断为是由不可观测技能价格造成的。

系数向量（特征价格向量），$\beta_t^b \equiv \beta_t(50)$ 是 50 分位的回归系数，$\beta_t^w \equiv \beta(\tau)-\beta(50)$ 是 τ 分位的回归系数与 50 分位回归系数向量之差，β_t^b 和 β_t^w 还可分别称为组间价格向量和组内价格向量。根据（3-20）式，在 MM2005 分解中的特征价格变动效应就可进一步细分为组间价格变动和残差价格变动两部分。[96]再记 $f(\ln w_t; g_t(X), \beta_t^b, \beta_t^w)$ 为由条件工资分布拟合成无条件的工资分布[97]，记 $\nu(\cdot)$ 为分布的统计量（比如分位等）。从而，工资总体分布从 $t=0$ 到 $t=1$ 时期的变动可由工资分布统计量 ν 的变动来度量，利用以下分解公式能够获得工资变动中各因素的贡献[98]：

$$\begin{aligned}\Delta\nu &= \nu[f(\ln w; g_1(X), \beta_1^b, \beta_1^w)] - \nu[f(\ln w; g_0(X), \beta_0^b, \beta_0^w)] \\ &= \{\nu[f(\ln w; g_1(X), \beta_0^b, \beta_0^w)] - \nu[f(\ln w; g_0(X), \beta_0^b, \beta_0^w)]\} \\ &\quad + \{\nu[f(\ln w; g_1(X), \beta_1^b, \beta_0^w)] - \nu[f(\ln w; g_1(X), \beta_0^b, \beta_0^w)]\} \\ &\quad + \{\nu[f(\ln w; g_1(X), \beta_1^b, \beta_1^w)] - \nu[f(\ln w; g_1(X), \beta_1^b, \beta_0^w)]\},\end{aligned} \tag{3-21}$$

其中，分解式第一项是个体特征分布 $g_t(X)$ 的变动效应，第二项是组间价格向量 β_t^b 的变动效应，最后一项是组内价格向量（残差价格向量）β_t^w 的变动效应。组间价格变动效应表征了组间工资不平等（between-group inequality）变化，描述工资分布的集中趋势[99]；组内价格变动（残差价格变动）效应则反映了组内工资不平等（within-group inequality）的变化。Autor *et al*.（2005）对美国工资总体分布变动的经验研究发现，20 世纪 70 年代开始高端部分（upper-tail）不平等上升主要源自组间和残差价格变动；90 年代工资分布低端部分（lower-tail）不平等下降也是由组间和残差价格变动造成的。

Autor 等还将上述分解框架进一步运用于残差分布变动分解。如果在工资总体分布 $f(\ln w; g_t(X), \beta_t^b, \beta_t^w)$ 中令 $\beta_t^b = \beta^b = 0$，$f(\ln w; g_t(X), \beta^b = 0, \beta_t^w)$ 则变成工资残差分布。[100]残差分布变动可以分解为特征分布变动效应和残差价格变动效应：

---

[96] 在 Melly(2005)原文中，特征价格变动效应细分为中位数系数变动效应和残差变动效应。与 Autor *et al*.(2005)细分的内容和大小是一致的，只不过称呼不同而已。

[97] Autor *et al*.(2005)在估计分位回归模型时，共取了 501 个分位点。以 0.2 为区间获得 500 个分位点[0.1，0.3，…，99.7，99.9]，再加上一个 50 分位点。此外，在 Autor *et al*.(2005)原文中工资边际密度函数表示为 $f(g_t(X), \hat{\beta}_t^b, \hat{\beta}_t^w)$，但本文为了更清晰和简洁起见将其表达为 $f(\ln w_t; g_t(X), \beta_t^b, \beta_t^w)$。

[98] Autor *et al*.(2005)原文是在分位 τ 上展开以 $t=0$ 为基期的分解。鉴于每个部分对工资总体分布变动的贡献依赖于分解顺序，故 Autor 等根据劳动力特征、组间价格和组内价格的顺序展开分解后，再逆序进行了分解。

[99] 尽管 Autor *et al*.(2005)表明中值和均值回归系数的差异在经验分析中并不显著，但仍需注意两者存在的差异。

[100] 这里的残差是指 $X_{it}\beta_t^w$。

$$\Delta\nu_R = \{\nu[f(\ln w; g_1(X), \beta^b = 0, \beta_0^w)] - \nu[f(\ln w; g_0(X), \beta^b = 0, \beta_0^w)]\} + \{\nu[f(\ln w; g_1(X), \beta^b = 0, \beta_1^w)] - \nu[f(\ln w; g_1(X), \beta^b = 0, \beta_0^w)]\}, \tag{3-22}$$

其中，第一项是特征分布变动效应，表示保持 $t=0$ 时期的残差价格，特征分布 X 变动对残差分布变动的影响[101]；第二项是残差价格效应。显然，(3-22) 式可以看做是 (3-21) 式在不同时期均设定 $\beta^b=0$ 的特例。

Lemieux (2006) 以为，劳动力构成变动是导致美国劳动力市场工资残差不平等上升的主要因素。[102]但是，Autor *et al*. (2005) 利用 Q-JMP 分解不但分析了劳动力构成变动对工资总体不平等与工资残差不平等变动的影响，还仔细考察了构成效应对工资分布高端和低端不平等变动的影响，并拒绝了 Lemieux 的构成效应假说 (Lemieux composition hypothesis)。Autor 等的经验研究拒绝该假说主要基于以下两个理由：一是实证结果表明劳动力构成变动主要是对低于中值工资的分布产生影响，而工资高端不平等扩大主要是由劳动力市场价格变动引起。二是 20 世纪 70 年代以后，劳动力构成效应对工资总体不平等和工资残差不平等变动的影响都是次要的，而特征价格变动才是主要因素。

与 JMP1993 分解相比，基于分位回归模型的 Q-JMP 分解既能够自然拓展到工资分布的各分位上，又可以将个体特征分布 $g_t(X)$ 的变动效应，组间价格向量 β_t^b 和残差价格向量 β_t^w 的变动效应加总为工资总体分布变动，因此，Q-JMP 分解相比 JMP1993 分解更为优越。

Q-JMP 分解的不足之处在于：一是没能将构成效应进一步分解到各个单一自变量；二是由于 Q-JMP 沿袭 MM2005 分解的分位分解框架，对条件分位回归模型作了参数线性的假定，该假定是否成立无法进行实证检验；三是 Autor *et al*. (2005) 自己也承认，工资分位回归方程包含哪些个体特征变量，将直接影响到 Q-JMP 分解中组间价格效应和残差价格效应的大小，而哪些个体特征变量进入工资分位回归方程，在某种程度上存在着固有的任意性。

（四）基于 RIF 回归的分布分解

Koenker and Bassett (1978) 提出的分位回归考察的是协变量 (解释变量) 对条件分位的影响，这其实是条件分位回归。在 Firpo *et al*. (2007b, 2009) 看来，当分布统计量为分位时，RIF 回归模型能够直接估计协变量对被解释变量的影响，更便于工资分布变动的分解。FFL 分解正是以 RIF 回归

[101] 此处特征分布变动效应是其对工资分布总体变动效应的一个部分。换言之，特征分布变动通过两条途径影响工资分布总体变动：一方面是直接效应，另一方面是通过影响残差分布间接影响工资分布总体变动。

[102] Autor *et al*. (2005) 引用的是 Lemieux(2005)，但后者又于 2006 年正式发表。

模型作为核心的工资差异分布分解方法。

FFL 分解是 Firpo *et al*.（2007b，2009）在借助 DiNardo *et al*.（1996）的重置权重函数（reweighting function）构造反事实工资分布从而把工资分布差异分解为构成效应和工资结构效应的基础上，对分布统计量的再集中影响函数（recentered influence function，RIF）进行回归，继而将上述两个效应进一步细分到每个协变量上的方法。为了区别以往的条件分位回归，Firpo *et al*.（2009）把这种分布统计量为分位时的 RIF 回归称为无条件分位回归（unconditional quantile regression）。[103]

FFL 分解方法的具体步骤分如下两步：

步骤一　利用 DiNardo *et al*.（1996）的重置权重函数将工资分布变动（或差异）分解为构成效应（composition effect）和工资结构效应（wage structure effect）。

$t(t=0,1)$时期的工资分布通常可以直接写成 $F_t=F(\ln w_t)$，但为了获得反事实工资分布，我们还需要将 t 时期的工资分布变换成带有协变量的工资边际分布（工资边际密度函数）。由于个体 i 的工资 $\ln w_{it}$ 取决于可观测的个体特征 X_{it} 和不可观测的因素 ε_{it}，则

$$\ln w_{ti} = s_t(X_i, \varepsilon_i), \tag{3-23}$$

其中，$s_t(\cdot,\cdot)$为劳动力市场中的工资结构函数（wage structure function）。从而，t 时期的工资分布还可以表达成 $F_t=F(\ln w_t; X_t, \varepsilon_t)$。

在构造工资结构为 $t=0$ 期而协变量分布为 $t=1$ 期的反事实工资边际分布 F_C 之前，Firpo *et al*.（2007b）明确提出了识别 F_C 的两个假设条件[104]：（1）可忽略性（ignorability）假定[105]，即给定 X 时，$F(\varepsilon|X,t=0)=F(\varepsilon|X,t=1)$；（2）支持重叠（overlapping support）假定，即对于所有 X，$P(X)=\Pr[t=1|X]<1$ 且 $\Pr[t=1]>0$，其含义是可观测的个人特征 X 在不同组群是重叠的，不会仅出现在某一期中。[106]在这两个假设条件下，若将 $t=0$ 和 $t=1$ 的随机样本整合为一个样本，则该混合截面样本对数工资的边际密度函数乘以重置权重因子 $\omega_C(t,X)$就能够得到 F_C。重置权重因子 $\omega_C(t,X)$的计算公式如下[107]：

---

[103] RIF 回归模型是指利用再集中影响函数(RIF)得到自变量变动对因变量分布统计量的影响。在 Firpo *et al*.(2009)看来，当分布统计量为分位时，RIF 回归模型能够直接估计协变量对被解释变量的影响，因此可称其为无条件分位回归模型。

[104] 值得注意的是，在构造反事实分布时都暗含着可忽略性假定和重叠支持假定，而 Firpo *et al*.(2007b)则是明确地指出了这两个假定并加以阐释。

[105] 在该假定下，构成效应才是单纯的协变量分布变动效应(即 $F(X|t=1)$ 和 $F(X|t=0)$ 差异效应)。否则，构成效应反映的是协变量和不可观测因素联合分布变动效应。

[106] 如果不支持重叠假定，那么，$P(X)=\Pr[t=1|X]$ 就非 0 即 1。

[107] FFL 分解中计算重置权重因子的基本思想与 DiNardo *et al*.(1996)相同。

$$\omega_C(t,X) \equiv \left(\frac{P(X)}{1-P(X)}\right)\left(\frac{1-t}{P}\right), \qquad (3\text{-}24)$$

其中，$P(X)=\Pr[t=1|X]$是指当给定可观测的个人特征 X 时，个体属于 $t=1$ 的概率，被称之为"倾向得分"（propensity-score），可通过 logit 或 Probit 模型得到。$P=\Pr[t=1]$表示个体来自 $t=1$ 的概率。同样，$t=1$ 和 $t=0$ 期的工资边际分布 F_1 和 F_0，也可以通过分别施加权重 $\omega_0(t)\equiv\frac{1-t}{1-P}$和 $\omega_1(t)\equiv\frac{t}{P}$得到。不言而喻，$F_1$ 和 F_0 还能够分别从 $t=0$ 和 $t=1$ 时期样本中直接得到。

另一方面，工资分布也可用分布的统计量来刻画，记为 $\nu_t=\nu(F_t)$。于是，在获得反事实工资边际分布 F_C 之后，不同时期工资分布变动 $\nu(F_1)-\nu(F_0)$就可分解为

$$\nu(F_1)-\nu(F_0)=[\nu(F_1)-\nu(F_C)]+[\nu(F_C)-\nu(F_0)]. \qquad (3\text{-}25)$$

公式右边的第一项表示工资结构效应，第二项则为构成效应。

步骤二 在由步骤一得到 $\nu(F_1)$、$\nu(F_0)$和 $\nu(F_C)$的基础上，通过分布统计量为分位时的再集中影响函数回归，获得类似 Oaxaca-Blinder 分解的形式，进而将工资分布变动（或差异）分解到各个单一协变量上。

Firpo *et al*.（2007b）利用 Firpo *et al*.（2007a，2009）提出的计算协变量分布变动对分布统计函数的偏效应方法[108]，将影响函数（influence function，IF）加回到分布统计中[109]，形成再集中影响函数（recentered influence function，RIF）：

$$\mathrm{RIF}(\ln w_t;\nu_t,F_t)=\nu(F_t)+\mathrm{IF}(\ln w_t;\nu_t,F_t). \qquad (3\text{-}26)$$

根据影响函数期望为零的性质，可以得到 RIF 的期望等于分布统计 $\nu(F_t)$。[110]继而，根据重律期望法则（the law of expectation），$\nu(F_t)$还可以写成 RIF 条件期望的期望：$\nu(F_t)=E[\mathrm{RIF}(\ln w_t;\nu_t,F_t)]=E[E[\mathrm{RIF}(\ln w_t;\nu_t,F_t)\,|X_t]]$，$t=0,1$。记再集中影响函数条件期望为 $m_t^{\nu}(X)=E[\mathrm{RIF}(\ln w_t;\nu_t,F_t|X)]$，Firpo 等称之为 RIF 回归模型，可用于计算在其他影响因素一定条

[108] Firpo *et al*.(2007b)原文引用的文章[即 Firpo *et al*.(2007a)]早在 2006 年和 2007 年就有不同版本的工作论文，后来于 2009 年正式发表[即 Firpo *et al*.(2009)]。

[109] 影响函数通常用来测量分布统计对特异值的稳定性。在某种意义上也可以说，影响函数类似于导数，反映的是因变量分布统计函数 $\nu(F)$ 的变化率。

[110] 影响函数期望值为零的性质，即$\int_{-\infty}^{\infty}\mathrm{IF}(y;\nu)\cdot \mathrm{d}F(y)=0$。从而，对于 $t=0,1$，有

$$\begin{aligned} E[\mathrm{RIF}(\ln w_t;\nu_t,F_t)] &= \int \mathrm{RIF}(\ln w_t;\nu_t,F_t)\cdot \mathrm{d}F(\ln w_t) \\ &= \int(\nu(F_t)+\mathrm{IF}(\ln w_t;\nu_t,F_t))\cdot \mathrm{d}F(\ln w_t)=\nu(F_t). \end{aligned}$$

件下协变量对无条件分布的偏效应。特别的，倘若 RIF 回归满足参数线性，则可进一步写成[111]

$$m_t^{\nu}(X) = E[\mathrm{RIF}(\ln w_t; \nu_t \mid X)] = X_t \beta_t^{\nu}. \tag{3-27}$$

根据重律期望法则，$\nu(F_t) = E[X_t \beta_t^{\nu}] = E[X_t]\beta_t^{\nu}$；个体特征分布保持在 $t=1$ 时期的反事实分布的分布统计量 $\nu(F_C) = E[X_1 \beta_C^{\nu}] = E[X_1]\beta_C^{\nu}$。显然，此时的 $\nu(F_1)$、$\nu(F_0)$ 和 $\nu(F_C)$ 都是带有协变量的无条件分布，说明 RIF 回归模型能够直接估计协变量对工资分布的影响。

综上，不同时期工资分布变动可表达成以各个协变量为基础的构成效应和工资结构效应，即

$$\begin{aligned}
\nu(F_1) - \nu(F_0) &= [E(X_1)\beta_1^{\nu} - E(X_1)\beta_C^{\nu}] + [E(X_1)\beta_C^{\nu} - E(X_0)\beta_0^{\nu}] \\
&= E(X_1)(\beta_1^{\nu} - \beta_C^{\nu}) + [E(X_1) - E(X_0)]\beta_0^{\nu} + R_0 \\
&= \sum_{k=1}^{K}[E(X_{1,k}) - E(X_{0,k})]\beta_{0,k}^{\nu} \\
&\quad + \sum_{k=1}^{K}[E(X_{1,k})(\beta_{1,k}^{\nu} - \beta_{C,k}^{\nu})] + R_0,
\end{aligned} \tag{3-28}$$

其中，第一项表示个体特征分布构成效应细分在各个协变量上的总和；第二项是工资结构效应细分在各个协变量上的总和；第三项 $R_0 = E[X_1](\beta_C^{\nu} - \beta_0^{\nu})$ 表示近似误差，在实证中包括由条件期望线性设定导致的误差和由一阶近似所产生的误差。[112]

Firpo *et al*.（2007b）的实证结果显示，构成效应可以解释美国 1988—2005 年间工资不平等上升的大部分原因，而从单一协变量效应来看，教育和工会是美国近期工资分布变动的主导力量。

FFL 分解通过 RIF 回归模型能将工资结构效应和构成效应分解到各个协变量，从而可以看做 Oaxaca-Blinder 分解拓展至工资分布分解上的一般形式。DFL 分解通常只能够用来将工资分布变动（或差异）分解出协变量总体的构

---

111 无条件分位回归模型一般说来可以采取灵活的函数设定形式。这里之所以假定模型为参数线性，主要基于三方面原因：一是已能够提供 X 分布变动效应的一阶近似，尤其是当分布统计是分位时，即 RIF 回归为无条件分位回归时，线性设定情况下的 RIF 回归得到的估计结果与更为灵活的非线性设定条件下回归结果非常相似。二是不影响从第一步利用重置权重函数得到的总的工资结构效应和构成效应。三是能够与 Oaxaca-Blinder 分解形式对应，为分解提供更为简洁的推断解释（参见 Firpo *et al*.，2007b）。

112 在估算 RIF 回归模型时，需借助 Von Mises(1947)展开式，用线性近似非线性分布统计的方法获得。当这种近似效果较好时，误差项 R_0 就较小。实际上，当 RIF 回归是线性回归时，我们还可以利用重置权重估计得到的构成效应 $\nu(F_C) - \nu(F_0)$ 和利用 RIF 回归方法得到的 $[E(X_1) - E(X_0)]\beta_0$ 之间的差异 R_0，作为 RIF 回归估计的一种特殊检验。譬如，Firpo *et al*.(2007b)对 1988—2005 年美国工资分布变动的经验分析表明，由步骤一得到的总的工资结构效应和构成效应与步骤二加总得到的两个效应之间的差异是微小的。

成效应和工资结构效应；倘若欲继续分解单一协变量的构成效应，则协变量必须是离散变量（虚拟变量）。而FFL分解则能放松DFL分解中单一协变量效应分解需要离散变量的局限。

然而，FFL分解依然存在以下一些不足：一是协变量对单个工资结构效应的贡献，会因RIF回归模型中的虚拟变量基组选择的不同而不同，也就是存在虚拟变量基组选择的问题。[113]二是FFL分解需要采用线性近似非线性分布统计的方法来计算统计量的RIF函数值，这种线性逼近必然会造成误差（Chernozhukov *et al.*，2009）。三是当选择不同时期工资结构作为工资分布分解的基准时，得到的工资结构效应和构成效应的推断结果将不唯一，故存在指数基准的问题。[114]

四、评论性小结

反事实工资是进行工资分解的基准，但反事实工资并没有在现实中直接呈现，而需要我们通过理论抽象和实证分析加以领悟和推断。寻找合理可靠的反事实工资，始终是准确有效地分解工资差异的关键之一。另一方面，基本模型的设定也对工资差异的分解起着举足轻重的作用。因为进行工资差异分解的前提是建立起各影响因素与工资之间的关联，这就需要选择某一基本模型来刻画这种关联。于是，本文从基本模型设定和反事实工资分布构造两个维度梳理工资差异分解方法的演进脉络，形成一幅工资分解方法演进的向导图（见图1）。

在图1中，用点划线将该图分为上下两个区域，分别表征均值分解和分布分解两大部分。图中各种分解方法之间若用实线连接，表示箭头所指的这种方法是在已有分解方法的基础上，直接采用其分解技术并加以改进或深化；若用虚线连接，则表示仅借鉴已有分解方法的某个基本思想，但不直接采用已有方法的分解技术，而是构建另一种相对独立的分解方法。

在均值分解区域，Oaxaca-Blinder分解是最基础的经典分解。围绕着Oaxaca-Blinder分解，Cotton分解、Neumark分解和郭继强分解是从解决指数基准问题出发对均值分解进行的改进；另一方面，Brown分解统一整合了同工不同酬和职业分隔，Appleton分解则尝试着解决Brown分解中的双重指数基准问题。此外，JMP1991分解利用残差信息探究了工资差异不可解释部分的成因，并推动着工资差异从均值分解向分布分解过渡。

---

[113] 分类变量基组选择问题是指分类变量对个人特征回报效应的贡献大小会随基组选择的不同而变动。

[114] 例如，当以 $t=1$ 期的工资结构作为分解基准时，分解式为

$$\nu(F_1)-\nu(F_0)=[E(X_0)(\check{\beta}_C-\check{\beta}_1)]+[E(X_1)-E(X_0)]\check{\beta}_1+R_1.$$

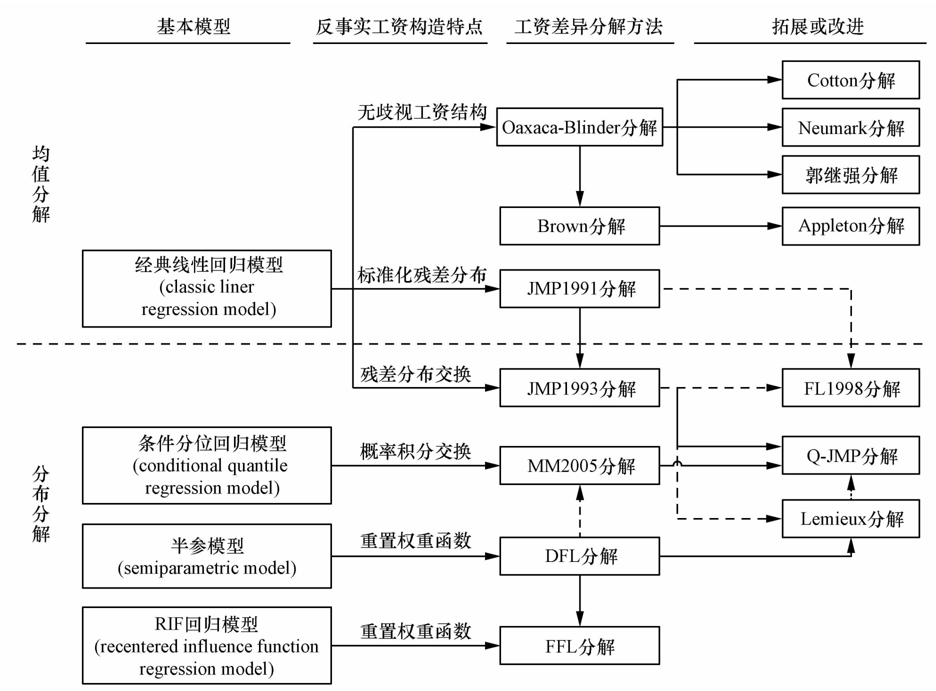

图 1　工资差异分解方法演进向导图

在分布分解区域，以 JMP1993 分解思想为基础，通过直接利用或借鉴 DFL 反事实分布构造技术，对分布分解方法进行拓展或改进。基本思路是逐步放松经典线性回归模型的假定条件，在不同的回归模型基础上构造相适应的反事实工资分布。当随机样本满足经典线性回归模型假定时，JMP1993 分解就成为相适宜的工资分布分解方法。然而，现实的样本数据很难同时满足参数线性和同方差这些相对苛刻的假定。DFL 分解方法就是直接与放松参数线性假定相联系的，该方法是在半参模型的框架下展开的，通过重置权重函数对样本重新赋权重，直接构造出无条件的反事实工资分布，避免了对工资回归方程的线性设定。在 JMP1993 分解基础上发展出了 FL1998 分解，该方法借鉴了 JMP1993 分解中排位的思想。而 Lemieux 分解则是以特征单元为分析对象，结合 JMP1993 分解特征回报效应的步骤，并利用 DFL 分解中重置权重的技术，解决了 DFL 分解中特征回报效应不能直接分离的缺陷。MM2005 分解通过利用条件分位回归模型，并借鉴 DFL 反事实分布构造中的重置权重思想，就可以放松 JMP1993 分解中的同方差假定，进而派生出 Q-JMP 分解。此外，为了进一步精进分布分解方法，FFL 分解结合了 DFL 分解技术，并在 RIF 回归模型的框架下展开分解，可将工资分布变动分解至各个单一的协变量效应，成为 Oaxaca-Blinder 分解在工资分布分解中的一般化，找到了贯通工资差异均值分解和分布分解的桥梁。特别需要强调的是，虽然工资差异分解方法是个逐步放松假定和精细化分解的演进过程，但并不存在新方法一定

比以往方法更占优势的递进关系，而是需要根据研究的目标、研究的对象、研究的问题和数据特点做出相适宜选择。

我国现有的关于工资差异分解的实证研究，考察组群间工资均值差异的研究相对丰富。其中，大多研究使用的是 Oaxaca-Blinder 分解（Gustafsson and Li，2000；王美艳，2005；姚先国和赖普清，2004；邓曲恒，2007；邢春冰，2008）；一部分文献结合 Oaxaca-Blinder 分解和 Cotton 分解开展比较研究（张丹丹，2004；谢嗣胜和姚先国，2006）。既有的研究职业分隔对工资差异影响的文献则多采用 Brown 分解（Meng，1998；Meng and Zhang，2001；王美艳，2005；李实和马欣欣，2006）。葛玉好（2007）使用过 Appleton 分解方法研究部门选择影响性别工资差异的程度。郭继强等（2010）针对 Appleton 分解存在的缺陷，提出了矫正 Brown 分解中双重指数基准的进一步改进方法。

近年来，关于工资分布差异或变动的研究越来越为学者们所关注。从分解技术上看，对同一组群不同时期工资分布变动的分解与对同一时期不同组群工资分布差异的分解步骤并无不同，只是研究对象的区别以及对分解结果解释和推断差别。考察我国工资分布差异的实证研究中，涉及 MM2005 分解方法的研究较多，不仅有研究性别工资分布差异的问题（陈建宝和段景辉，2009），还有将该分解方法运用于研究中国城乡消费不平等的来源（Qu and Zhao，2008）。然而，大多数运用 MM2005 分解的研究是以不同组群工资（收入）作为研究对象，将该方法直接用于工资不平等的研究较少，仅有 Xing（2010）同时利用 DFL 分解和 MM2005 分解研究了 1995—2007 年中国城镇职工工资残差不平等。相对于 MM2005 分解，Q-JMP 分解则更多地被用于研究同一组群工资不平等的变动及原因（姚先国和李晓华，2007）。FFL 分解既被用于中国城镇性别工资分布差异的研究，又被用于城镇工资不平等变动的研究（Chi and Li，2008；迟巍等，2008）。据目前笔者所掌握的文献来看，虽然未有研究较为深入地运用 DFL 分解针对我国工资不平等问题展开讨论，但由于 FFL 分解过程的第一步骤就是 DFL 分解，而 MM2005 分解在将个体特征效应进行细分时也借鉴了 DFL 分解的思想，因此在某种程度上 DFL 分解已被运用于中国问题的研究。Meng *et al*.（2010）采用 Lemieux 分解讨论了工资收入方差的变动及成因。邢春冰和罗楚亮（2009）综合 DFL 分解和 FL1998 分解两种方法讨论了农民工与城镇职工工资收入差距的形成原因。此外，虽然 JMP1991 分解原本针对的是不同组群工资差异的均值分解，但也有学者应用该方法研究我国性别工资分布差异的变动及其成因（Zhang *et al*.，2008）。当然，这样的工作还可以见仁见智地商榷。

对于工资差异分解方法的未来发展，笔者以为，工资差异均值分解中备受关注的两类问题——指数基准问题以及样本选择性偏差和虚拟变量识别等计量技术问题，在分布分解中也同样不可避免，值得关注和深入研究。FFL

分解也可以进一步拓展，以形成分解工资不平等和残差不平等的统一分析框架。再者，如何将反事实工资构造与政策评估中的处理效应（treatment effect）问题有机贯通，从而为制定公共政策提供更科学有效的实证研究支撑，很可能是工资差异分解方法发展的新空间。

参考文献

[1] Altonji, J., and R. Blank, "Race and Gender in the Labor Market", in Ashenfelter, O., and D. Card (eds.), *Handbook of Labor Economics*, volume 3C. Elsevier, 1999, 3143—3260.

[2] Appleton, S., J. Hoddinott and P. Krishnan, "The Gender Wage Gap in Three African Countries", *Economic Development and Cultural Change*, 1999, 47(2), 289—312.

[3] Arrow, K., "Models of Job Discrimination", in Pascal, A. (ed.), *Racial Discrimination in Economic Life*. Lexington Mass: D. C. Heath, 1972.

[4] Autor, D., F. Lawrence, and S. Melissa, "Rising Wage Inequality: The Role of Composition and Prices", Harvard Institute of Economic Research Working Papers No. 2096, 2005.

[5] Becker, G., *The Economics of Discrimination*. Chicago: The University of Chicago Press, 1957.

[6] Blau, F., and L. Kahn, "The Gender Earnings Gap: Learning from International Comparisons", *American Economic Review*, 1992, 82(2), 533—538.

[7] Blau, F., and L. Kahn, "Wage Structure and Gender Earning Differentials: An International Comparison", *Econometrica*, 1996, 63(250), 3—8.

[8] Blau, F., and L. Kahn, "Swimming Upstream: Trends in the Gender Wage Differential in the 1980s", *Journal of Labor Economics*, 1997, 15(1), 2—42.

[9] Blinder, A., "Wage Discrimination: Reduced Form and Structural Estimates", *Journal of Human Resources*, 1973, 8(4), 436—455.

[10] Brown, R., M. Moon, and B. Zoloth, "Incorporating Occupational Attainment in Studies of Male-Female Earnings Differentials", *Journal of Human Resources*, 1980, 15(1), 3—28.

[11] Buchinsky, M., and J. Hahn, "An Alternative Estimator for the Censored Quantile Regression Model", *Econometrica*, 1998, 66(3), 653—671.

[12] Butler, R., "Estimating Wage Discrimination in the Labor Market", *Journal of Human Resources*, 1982, 17(4), 606—21.

[13] 陈建宝、段景辉，"中国性别工资差异的分位数回归分析"，《数量经济技术经济研究》，2009 年第 10 期，第 87—97 页。

[14] Chernozhukov, V., I. Fernandez-Val, and B. Melly, "Inference on Counterfactual Distributions", forthcoming in *Econometrica*, 2009.

[15] Chi, W., and B. Li, "Glass Ceiling or Sticky Floor? Examining the Gender Earnings Differential across the Earnings Distribution in Urban China, 1987—2004", *Journal of Comparative Economics*, 2008, 36(2), 243—263.

[16] 迟巍、黎波、余秋梅，"基于收入分布的收入差距扩大成因的分解"，《数量经济技术经济研究》，2008 年第 9 期，第 52—64 页。

[17] Cotton, J., "On the Decomposition of the Wage Differentials", *Review of Economics and Statistics*, 1988, 70(2), 236—243.

[18] Démurger, S., M. Fournier, and Y. Chen, "The Evolution of Gender Earnings Gaps and Discrimination in Urban China, 1988—95", *Developing Economies*, 2007, 45(1), 97—121.

[19] 邓曲恒,"城镇居民与流动人口的收入差异——基于 Oaxaca-Blinder 和 Quantile 方法的分解",《中国人口科学》,2007 年第 2 期,第 8—16 页。

[20] DiNardo, J., N. Fortin, and T. Lemieux, "Labor Market Institutions and the Distribution of Wages, 1973—1992: A Semiparametric Approach", *Econometrica*, 1996, 64(5), 1002—1044.

[21] Donald, S., D. Green, and H. Paarsch, "Differences in Wage Distributions between Canada and the United States: An Application of a Flexible Estimator of Distribution Functions in the Presence of Covariates", *Review of Economic Studies*, 2000, 67(4), 609—633.

[22] Duncan, G., and D. Leigh, "Wage Determination in the Union and Nonunion Sectors: A Sample Selectivity Approach", *Industrial and Labor Relations Review*, 1980, 34, 24—34.

[23] Ferber, M., and C. Green, "Traditional or Reverse Sex Discrimination? A Case Study of a Large Public University", *Industrial and Labor Relations Review*, 1982, 35(4), 550—564.

[24] Firpo, S., N. Fortin, and T. Lemieux, "Unconditional Quantile Regressions", NBER Technical Working Paper, No. 339, 2007a.

[25] Firpo, S., N. Fortin, and T. Lemieux, "Decomposing Wage Distributions Using Recentered Influence Function Regressions", Mimeo, Department of Economics, University of PUC-RIO. 2007b.

[26] Firpo, S., N. Fortin, and T. Lemieux, "Unconditional Quantile Regressions", *Econometrica*, 2009, 77(3), 953—973.

[27] Fortin, N., and T. Lemieux, "Rank Regressions, Wage Distributions, and the Gender Gap", *Journal of Human Resources*, 1998, 33(3), 610—643.

[28] Fortin, N., T. Lemieux, and S. Firpo, "Decomposition Methods in Economics", Working Paper, 2010.

[29] 葛玉好,"部门选择对工资性别差距的影响:1988—2001 年",《经济学(季刊)》,2007 年第 6 卷第 2 期,第 607—628 页。

[30] 郭继强、陆利丽,"工资差异均值分解的一种新改进",《经济学(季刊)》,2009 第 8 卷第 4 期,第 1257—1280 页。

[31] 郭继强、姜俪、陆利丽,"双重指数基准矫正下 Brown 分解方法新改进",2010 年工作论文(该论文 2009 年入选第九届中国经济学年会)。

[32] Gustafsson, B. and S. Li, "Economic Transformation and the Earnings Gap in Urban China", *Journal of Population Economics*, 2000, 13(2), 305—329.

[33] Heckman, J., "Sample Selection Bias as a Specification Error", *Econometrica*, 1979, 47(1), 153—161.

[34] Juhn, C., K. Murphy, and B. Pierce, "Accounting for the Slowdown in Black-White Wage Convergence", in Kosters, M. (ed.), *Workers and Their Wages: Changing Patterns in the United States*. Washington: American Enterprise Institute Press, 1991.

[35] Juhn, C., K. Murphy, and B. Pierce, "Wage Inequality and the Rise in Returns to Skill", *Journal of Political Economy*, 1993, 101(3), 410—442.

[36] Katz, L., and K. Murphy, "Changes in Relative Wages, 1963—1987: Supply and Demand Factors", *Quarterly Journal of Economics*, 1992, 107(1), 35—78.

[37] Koenker, R., and G. Jr Bassett, "Regression Quantiles", *Econometrica*, 1978, 46(1), 33—50.

[38] Lee, L-F., "Generalized Econometric Models with Selectivity", *Econometrica*, 1983, 51 (2), 507—512.

[39] Lemieux, T., "Decomposing Changes in Wage Distributions: A Unified Approach", *Canadian Journal of Economics*, 2002, 35(4), 646—688.

[40] Lemieux, T., "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?" *American Economic Review*, 2006, 96(3), 462—498.

[41] 李春玲、李实，“市场竞争还是性别歧视——收入性别差异扩大趋势及其原因解释”，《社会学研究》，2008 年第 2 期，第 94—117 页。

[42] 李利英、董晓媛，“性别工资差异中的企业效应”，《经济研究》，2008 年第 9 期，第 122—135 页。

[43] 李实、马欣欣，“中国城镇职工的性别工资差异与职业分割的经验分析”，《中国人口科学》，2006 年第 5 期，第 3—13 页。

[44] Machado, J., and J. Mata. "Counterfactual Decompositions of Changes in Wage Distributions Using Quantile Regression", *Journal of Applied Econometrics*, 2005, 20(4), 445—465.

[45] Melly, B., "Decomposition of Differences in Distribution Using Quantile Regression", *Labour Economics*, 2005, 12(4), 577—590.

[46] Meng, X., "Gender Occupational Segregation and its Impact on the Gender Wage Differential among Rural-Urban Migrants: a Chinese Case Study", *Applied Economics*, 1998, 30 (6), 741—752.

[47] Meng, X., and J. Zhang, "The Two-Tier Labor Market in Urban China Occupational Segregation and Wage Differentials between Urban Residents and Rural Migrants in Shanghai", *Journal of Comparative Economics*, 2001, 29 (3), 485—504.

[48] Meng, X., K. Shen, and S. Xue, "Economic Reform, Education Expansion, and Earnings Inequality for Urban Males in China, 1988—2007", IZA Discussion Papers No. 4919, 2010.

[49] Neuman S., and R. Oaxaca, "Wage Decomposition with Selectivity-corrected Wage Equation: A Methodological Note", *Journal of Economic Inequality*, 2004, 2(1), 3—10.

[50] Neumark, D., "Employers' Discriminatory Behavior and the Estimation of Wage Discrimination", *Journal of Human Resources*, 1988, 23(3), 279—295.

[51] Oaxaca, R., "Male-Female Wage Differentials in Urban Labor Markets", *International Economic Review*, 1973, 14(3), 693—709.

[52] Oaxaca. R., and M. Ransom, "On Discrimination and the Decomposition of Wage Differentials", *Journal of Econometrics*, 1994, 61(1), 5—21.

[53] Parzen, E., "On Estimation of a Probability Density Function and Mode", *Annals of Mathematical Statistics*, 1962, 33(3), 1065—1076.

[54] Qu, Z., and Z. Zhao, "Urban-Rural Consumption Inequality in China from 1988 to 2002: Evidence from Quantile Regression Decomposition", IZA Discussion Papers No. 3659, 2008.

[55] Reimers, C., "Labor Market Discrimination against Hispanic and Black Men", *Review of Economics and Statistics*, 1983, 65(4), 570—579.

[56] Rosenblatt, M., "Remarks on Some Nonparametric Estimates of a Density Function", *Annals of Mathematical Statistics*, 1956, 27(3), 832—837.

[57] Sanborn, H., "Pay Differences between Men and Women", *Industrial and Labor Relations Review*, 1964, 17 (4), 534—550.

[58] Sheather, S., and M. Jones, "A Reliable Data-Based Bandwidth Selection Method for Kernel Density Estimation", *Journal of the Royal Statistical Society*, Series B (Methodological), 1991, 53 (3), 683—690.

[59] Stevenson, M., "Relative Wages and Sex Segregation by Occupation", in Lloyd, C. (ed.), *Sex, Discrimination, and the Division of Labor*. New York: Columbia University Press, 1975.

[60] Sung, Y., J. Zhang, and C. Chan, "Gender Wage Differentials and Occupational Segregation in Hong Kong, 1982—1996", *Pacific Economic Review*, 2001, 6 (3), 345—359.

[61] Teo, S., "Occupational Segregation and its Effect on Estimates of the Gender Wage Differential: Evidence from Brunei", *Asian Economic Journal*, 2003, 17(4), 342—360.

[62] Von Mises, R., "On the Asymptotic Distribution of Differentiable Statistical Functions", *Annals of Mathematical Statistics*, 1947, 18(3), 309—348.

[63] 王美艳,“城市劳动力市场上的就业机会与工资差异——外来劳动力就业与报酬研究”,《中国社会科学》,2005年第5期,第35—46页。

[64] 王美艳,“中国城市劳动力市场上的性别工资差异”,《经济研究》,2005年第12期,第35—43页。

[65] Wooldridge, J., *Introductory Econometrics: A Modern Approach*, Second Edition. South-Western College Publishing (a division of Thomson Learning), 2003.

[66] 谢嗣胜、姚先国,“农民工工资歧视的计量分析”,《中国农村经济》,2006年第4期,第49—55页。

[67] Xing, C., “Residual Wage Inequality in Urban China, 1995—2007”, IZA Discussion Papers No. 5003, 2010.

[68] 邢春冰,“农民工与城镇职工的收入差距”,《管理世界》,2008年第5期,第55—64页。

[69] 邢春冰、罗楚亮,“农民工与城镇职工的收入差距——基于半参数方法的分析”,《数量经济技术经济研究》,2009年第10期,第74—86页。

[70] 姚先国、赖普清,“中国劳资关系的城乡户籍差异”,《经济研究》,2004年第7期,第83—90页。

[71] 姚先国、李晓华,“市场化与工资不平等增长:变动程度及影响因素”,《浙江大学学报(人文社会科学版)》,2007年第1期,第161—167页。

[72] 姚先国、李晓华,“工资不平等的上升:结构效应与价格效应”,《中国人口科学》,2007年第1期,第36—43页。

[73] Yun, M., “A Simple Solution to the Identification Problem in Detailed Wage Decompositions”, *Economic Inquiry*, 2005, 43(4), 766—772.

[74] Zellner, H., “Discrimination against Women, Occupational Segregation, and the Relative Wage”, *American Economic Review*, 1972, 62 (1/2), 157—160.

[75] 张丹丹,“市场化与性别工资差异研究”,《中国人口科学》,2004年第1期,第33—41页。

[76] Zhang, J., J. Han, P. Liu, and Y. Zhao, “Trends in the Gender Earnings Differential in Urban China, 1988—2004”, *Industrial and Labor Relations Review*, 2008, 61(2), 224—243.

[77] Zveglich, J., and Y. Rodgers, “Occupational Segregation and the Gender Wage Gap in a Dynamic East Asian Economy”, *Southern Economic Journal*, 2004, 70 (4), 850—875.

Decomposition Methods for Wage Differentials: A Survey

JIQIANG GUO LI JIANG LILI LU

(*Zhejiang University*)

Abstract Wage decompositions are used to decompose the wage differentials in a distributional statistic into various explanatory factors. The methods can be divided into mean decomposition and distributional decomposition. The former is an improvement based on the canonical Oaxaca-Blinder decomposition. The latter includes seven main approaches, where JMP1993, DFL and MM2005 lay a foundation for them. This paper systematically exploits the differences and the connections among these methods and also proposes a comprehensive guide map to understand the development of the decomposition methods for wage differentials.

JEL Classification J31, J71, C14