

本期导读

研 究 前 沿

援助质量和援助机构排序

新 书 推 荐

全球监测报告 2010：危机之后的千年发展目标

专 家 视 点

转移支付：促进家庭发展

机 构 介 绍

欧洲复兴开发银行（EBRD）

援助质量和援助机构排序

Stephen Knack, F. Halsey Rogers和Nicholas Eubank

引言

如何对各援助机构的援助质量进行评估和比较？一些近期的文献认为援助机构的绩效可以通过一种有意义的、可信的方式，在全球维度进行测量和排序，从而就双边和多边援助机构的排序提出了各自的建议（Mosley, 1985; Roodman, 2006、2009; Center for Global Development, 2007; Easterly和Pfutze, 2008）。公开排序的目的旨在激励双边和多边援助机构通过名次的上升，来改善其实践、效率和效力。援助机构显然会关注这样的排名，在意公众的看法（Ramankutty等, 2009）。发展援助委员会（DAC）内部对“同僚压力”的持续关注也导致了援助比例的明显下降，该比例受制于援助国内部对商品和服务的购买（OECD 2009a）。

然而排序成功与否，极大地取决于其自身的效度和信度。由于对援助机构就援助的各个独立组成部分进行排序显然存在困难，本文对综合排名的信度进行了考察，加入了一些新的指标，并对各个组成部分赋予不同的权重。如果援助机构绩效的不同指标之间缺乏高度的相关性，那么最终的排名将会受到权重分配的巨大影响。来自Easterly和Pfutze（2009）的研究就证实，对指标各个组成部分（即该案例中的政策选择性和贫困选择性）的赋权，能够导致不同的机构排名。

腐败国家和非自由国家在（全球环境基金）的官方发展援助（ODA）中所占份额相对较低，最贫困国家也如是。正如本文所不惮承认的，某一方面的选择性，必然会以另一方面的选择性为代价……

在评估援助机构的援助质量时，本文在许多方面参考并扩展了Roodman（2006, 2009）、Easterly和Pfutze（2008）的成果。首先，采取一套更加全面的指标来评估援助机构的绩效。其中不仅包括了现有的指标，也加入了反映《巴

黎宣言》内容的新指标，所以更全面、更具代表性。特别是极大地借鉴了经合组织发展援助委员会（OECD-DAC）的巴黎宣言监测调查（SMPD），该成果包含了国家制度运用、援助可预测性，与国内其他援助机构的协调等方面的大量的新指标（OECD, 2008）。此外，我们还加入了关于通过多边援助机构所获得的双边援助份额的新指标。采取更多指标的优点之一在于，援助机构的综合排名不致对某一特定指标（或赋予其的特定权重）过于敏感。

其次，由于排序的目标之一是评估援助机构运营管理有效与否，所以我们依据不由援助机构直接控制的重要因素，对援助质量的大部分指标都进行了校正。例如，从管理的角度看，指责欧洲复兴开发银行（EBRD）没有关注世界上最贫困的国家是毫无意义的，因为该机构只允许在欧洲和中亚地区的中等收入国家开展工作。因此在排名时仅仅依据EBRD能够开展工作的国家，来测量其贫困选择性。我们事实上是在评估EBRD是否在其有权开展工作的地区瞄准了更为贫困的国家。后文的方法论部分将进一步阐述我们如何就援助机构的限制因素进行了校正，以及为何这样做。

第三，有关部门设置和项目推广的计算指标来自新的援助数据库（AidData）。该数据库由威廉玛丽学院以及杨百翰大学的研究人员共同开发，以OECD-DAC的信用报告系统（CRS）数据为基础，该数据在Easterly和Pfutze（2008）及Roodman（2006; 2009）的研究中已经采用；并加入了新的援助机构、受援者以及援助机构直接资助的项目等方面的数据。

最后，我们为2007年的工作提供了一个援助质量指标数据库，其中包含了校正后的指标及我们构建的各种指数及分类指数。在获得第三轮巴黎宣言监测调查的数据后，这一数据库

将能够进行更新并在2009年继续使用。需要注意的是，我们的援助机构样本与Easterly和Pfütze所采用的样本相似，他们的研究包含了DAC所有的双边援助机构及最重要的多边援助机构和全球基金，而我们则排除了缺失半数以上指标数据、或超过一个分类指数数据的援助机构，其中大多是新兴的、小型的、非DAC成员的援助机构。为了与巴黎宣言监测调查（我们的18个指标中，有8个来源于此）保持一致性，我们将联合国作为一个援助机构对待，取其各个机构数据的加和值。

收集和整合援助质量指标工作的成果既激动人心，又富有启示。首先，我们证明了能够将这些指标操作化为有逻辑的分类指数，以反映援助质量的主要因素：选择性、与国家制度的一致性、国内的和谐性、专业性。每个因素都是改善援助质量的立足点。我们还建立了这些指标的分类指数，它们彼此高度相关，能够测量援助质量的各个重要因素。援助机构在这些分类指数方面的排序，能够为援助机构所用，审视自身，明确相应的优势和劣势所在。

第二，我们基于这些分类指数，提供了一个新的援助机构质量综合排序。与其他的排序相比，这一排序覆盖了援助质量的更多方面，可能是对援助机构整体质量更公平的反映。

第三，我们证明了权重确实会影响排序：由于各个援助机构在不同的分类指数上排序结果不同，那么最终的综合排名就必然受到这些分类指数权重的影响。我们认为排序应该考虑对权重的敏感性，公开赋权的原因，使读者自行决定是否赞同。为了提高透明度，我们采取了不同的方式对分类指数进行汇总，表明排序结果是如何随权重而变化的。

援助机构排序：挑战、激励和方法

任何试图对援助机构的援助质量进行排序的努力都将举步维艰。首先，大多数关于援助机构绩效的指标都是基于对援助管理的最佳实践的信仰，这种信仰貌似可靠，实则未经验证。政策选择性在文献中获得了最高的关注度（World Bank, 1998; Burnside和Dollar, 2000），然而政策质量与援助效果之间的经验联系始终莫衷一是（Hansen和Tarp, 2000、2001; Easterly, Levine和Roodman, 2003、2004），评估选择性的标准也遭到质疑（Amprou, Guillaumont

和Guillaumont Jeanneney, 2007）。就其它的援助质量指标而言，例如援助机构对国家制度的利用等，虽然在援助机构内部对其重要性有着更多的共识，却更加缺乏经验证据的辅佐。虽然进来情况有所好转，Klein和Harford（2005）几年前的结论仍然基本正确：

显然，在影响援助效果的因素方面缺乏更好的理论支撑和经验证据的情况下，试图对援助质量进行令人信服的测量极其困难……这种警告的重要性不亚于得出一个简单的结论，因为得出结论所用的指标，要么仅仅进行了有限的实证研究，要么就是闭门造车，导致我们期待这些指标能够直观地通过相似的事实传递截然不同的信息。本文关于援助质量的初步探索的真正价值在于，强调了援助质量的重要性及我们的无知程度。但愿将来对援助质量和援助机构绩效的测量能够基于对援助效果的严格评估。

在选择关于援助质量的可靠指标时，我们遵循了Easterly和Pfütze（2008）的方法：

一个理想的援助机构是什么样的？学术界关于援助政策的文献及援助机构自身，就“最佳实践”的诸多方面达成了一致……但是，如果将这些共识作为标准，那么援助机构在实践中是否能以其自称正确的方式运作？

其次，排序暗含的假设是：这些指标所测量的援助特征，与更好的产出密切相关；并假设一切努力都单一地指向更好的产出。也就是说，要么假设某些特征更明显些会更好（例如和谐性或选择性），要么至少认为所有的援助机构都有所欠缺，且在短期内改善无望。我们也沿袭了这种现存排序所用的假设。

第三，排序中所使用的一些指标忽略了援助机构之间的互动，而这一点至关重要。例如，如果大多数援助机构都向最贫困的国家聚集，就会改善它们的贫困选择性，然而边际产出却不一定能够得到提高。即使是援助效果的坚定拥趸，也会认识到援助对增长的边际效应会随着援助规模的增长而递减，最终可能降为负值。在这种情况下，如果“边际”援助机构能够避免扎堆，将其援助资源用于贫困程度略轻的国家，就会对贫困产生更为巨大的影响。

本文并不关注这些问题，转而思索以下问

题：即使能够假设所测量的援助机构的行为对于发展非常重要，就能够将它们简单地整合成具备效度和信度的单一指数、进行援助机构质量测量吗？或者对援助质量各个指标的测量都不具有相关性，所以任何综合排序都会受到赋予权重和指标分类时主观判断的严重影响？

在对援助机构进行排序时，方法与目标的契合非常重要。截至目前，对援助机构的排序简直成为“点名批评”：其目的是通过为援助机构设置基准，激励落后的机构提高援助额度，改善援助质量。Roodman/CGD以及Easterly和Pfutze的排序正是如此，虽然程度不同。Easterly和Pfutze自称在描述“援助机构的行为”，对透明度、部门设置、选择性、无效援助渠道的使用，以及间接成本等方面进行了考察。他们的目标明确地定位于激励援助机构在这些方面做出改进。以提高透明度为例：

如果本文能够促使援助机构变得更加透明，并为今后的研究提供可获得的质量数据，我们也就得偿所愿。也许某天我们能够基于产出对机构进行排序（Easterly和Pfutze，2009）。

Roodman/CGD的工作也试图推动援助机构在诸如援助附加条件、项目复制/部门设置及援助规模等方面做出改变。援助排序结果会作为年度发展职责指数（CDI）的一部分公开发布。CDI网站所述如下：

我们希望引发一场巅峰对决，获胜者应该为它们的成就而感到自豪。当然，所有的富裕国家都仍有进步的空间。几乎所有国家都至少在某一方面的得分低于平均水平……（Center for Global Development，2010）

正如Easterly和Pfutze所愿，援助机构能够通过改变自身的行为，提高在全球发展中心（CGD）“点名批评”名单上的排名。从而，排序就实现了改善援助机构行为的初衷。

要实现这样的目标，对方法论有何要求？为了评估援助机构如何完成其任务，我们认为最重要的是采用计量经济学的方法，对援助机构力不能及的因素进行控制。上文已经提到EBRD有限的地理活动范围的例子，而另一个例子则是非洲开发银行，如果进行全球比较，而非仅仅针对其职责所在的非洲国家进行考虑的话，它不可避免地会在政策选择性方面排名

更加靠后。通过针对使命的局限性进行校正之后，我们能够对援助机构的管理提出更多的问题：在其使命前提下，是否尽可能有效地减少贫困？例如根据指定的地理区域的政策条件和贫困状况，通过瞄准手段实现目标？就其他指标而言，我们也针对援助机构的总体援助预算进行了校正。例如，大型的援助机构会将其援助“复制”至更多的国家和地区，而在其他方面表现平平，也不应对此多加苛责。在测量对国家制度的利用时，我们针对各个援助机构所在国家整体援助方案的风险差异进行了校正，以免使那些针对公共财务管理系统薄弱的受援国提供了大量援助的援助机构在排序中沦于不利地位。下文对核心排序的计算中，包含了所有的校正。

要注意，排序的其他目的可能会需要不同的方法论。例如排序可以帮助援助国政府决定如何最好地分配其援助资源。援助基金应该通过本国的双边机构，还是多边机构和发展银行进行分配？如果通过多边机构，那么要选择哪个？就这一目标而言，未必所有的校正都合理。而如果目标在于最有效地配置边际援助金额，那么援助国的决定就应考虑所有影响这一目标的限制因素。例如如果援助国希望其边际资源直接进入贫困国家，就有可能不希望选择一个只能在中等收入地区开展活动的机构，即使该机构的贫困选择性得分很高也不行。下文中我们提出了一个忽略了所有校正的“配置指数”，可能会对这样的分配决策更有效。

数据及方法：构建分类指数

如上文所述，本文的主要目标之一在于采用一套更为综合的援助质量指标，对援助机构进行排序。本节将介绍我们是如何分解这些指标、对它们进行校正以提高效率，并汇总成为几个分类指数的。

在构建指数时，我们首先确定了援助质量的四个维度：选择性、一致性、和谐性和专业性。测量理论指出，测量质量的相似部分的指标应该归为一组，这也是我们构建分类指数的原因之一。我们确定的四个维度是来自于有关援助效果的文献，及针对援助的国际协议，当然我们也发现这种划分经得起实证数据的验证。当我们借助常识将不同的指标归入这四个分类指数时，发现正与这些分类指数内部及相

互之间的相关性不谋而合。在后文的结论部分，我们将讨论为什么这些分类指数本身就具备相当的重要性，甚至超越了在构建援助质量综合指数中的作用。

援助选择性

通常认为，在最需要的地方，即贫困人口为数众多的地方，以及有利于政策和制度环境增长和发展的地方，援助对发展产生的影响最大。这种直觉的信仰很难推翻，连Burnside和Dollar（2000）的实证结果也无济于事。因此，CGD及Easterly和Pfutze的排序都综合了政策和贫困选择性的测量，而Dollar和Levin（2006）则仅关注评估援助效果的可能趋势的变量。

在援助选择性变量方面，我们以Dollar和Levin（2006）的方法计算了对政策和贫困选择性的测量。特别是，我们对来自援助国*i*到受援国*j*的援助基于受援国*j*的3个变量进行了对数回归：受捐国*j*的人口对数，人均GDP（按照购买力评价校正之后），及其在世界银行的年度国别政策和制度评估（CPIA）中的总分。CPIA由“代表有效的减贫和增长策略的不同的政策和制度维度”的16个指标组成。这16个指标是由世界银行的工作人员会同外部著名发展专家设计，分为4类：宏观政策和财政政策，结构政策、社会政策及公共部门管理（包括腐败），用于分配世界银行的优惠贷款（IDA）。

就每个援助国都分别进行了这一个选择性回归。政策选择性简单地由CPIA值的部分回归系数进行测量，而贫困选择性则是由人均收入的部分回归系数进行测量（乘以-1，所以数值越大，表明低收入受援国得到了更好的瞄准）。

回归中所用的援助支出是2007年的数据，而3个自变量则滞后一年，更准确地反映援助国在进行分配时所掌握的信息。对大部分援助国而言，回归所用的样本包括125个受援国。而对于在特定区域开展活动的多边机构，则仅限符合条件的国家。

并非所有的受援国都有2007年的人口、GDP或CPIA值的数据。最明显的是伊拉克和阿富汗接受的援助占美国2007年度官方发展援助（ODA）总额的27%多，却缺少GDP和CPIA的数据。美国在政策选择度方面的排名已经靠后，如果加入这两个国家，分值会更低。

我们采用Dollar和Levin（2006）的方法来计算援助选择性这一分类指标，即通过对每个国家的政策和贫困弹性系数标准化为以0为平均值、1为标准差，随后赋予相同的权重。政策和贫困选择性之间的相关系数是.53，表明援助国在对二者之一加以选择时，（其他条件不变）也会考量另一个维度。

如果考虑所有的援助，即所有援助国援助额度的总和，那么长期来看，援助选择性较为稳定。我们的GDP和CPIA弹性系数在2007年度分别为-.52和+1.91，仅比Dollar和Levin（2006）报告中2002年的预测值-.49和+1.76略有边际改善。由于在选择性和受援国的增长或减贫之间构建切实的经验联系中存在困难，也就无从确切得知总体中的援助是否在贫困维度或政策维度具备“充分的”选择性，从而使减贫的影响最大化。隐含在各种援助排序实践中的假设证否了这一点，因为所有的援助国都仍然在从提高选择性中获取边际回报。

如Dollar和Levin（2006）所述，使用世界银行构建的CPIA指标作为政策质量的测量手段，会在选择性排序中偏袒世界银行的IDA。因此他们以国际国别风险指南（ICRG）的“法治”指标和自由之家（FH）的民主指标来替代CPIA进行分析。采取这些变量时，一些援助国的排名略有变动，表明对政策变量选择存在一定的敏感性。虽然如此，使用ICRG/FH政策变量得到的政策选择性和整体的援助选择性的排序，与原本的排序结果的相关系数分别为.65和.77。此外，这种新的测量方式有着严重的不足：大多（如ICRG）仅覆盖了CPIA测量的发展相关政策的有限领域，且其政策覆盖面（如世界经济自由度指标）与受援国的援助覆盖面并不一致。

我们能够以非洲开发银行和亚洲开发银行的CPIA排序结果，来测试世界银行的偏差。这些排序采用相同的量表和调查问卷来获得世界银行的CPIA，但是更多地体现了非洲开发银行和亚洲开发银行工作人员的判断，而非世界银行工作人员的意见。我们就这两个地区开发银行对世界银行在73个国家援助分配的评价进行了选择性回归，首先使用世界银行自己的CPIA评估，然后使用非洲开发银行和亚洲开发银行的评估。在政策和贫困选择性的得分上，世界银行用其他开发银行排序的得分要比自身排序

评估的得分高，表明对世界银行CPIA评估的使用并没有导致有利于其自身的选择性排序偏差。

援助一致性

在《巴黎宣言》中，各援助国承诺“以伙伴国的国家发展战略、制度和程序为基础，提供全面支持”。随后的《阿克拉议程》中，援助国又承诺“通过尊重各国的发展优先序列，投资其人力资源和制度，充分利用其各个系统分配援助，提高援助的可预测性”，支持各国对发展项目的所有权（OECD，2009a）。这些声明表明了这样的共识：援助国提供援助的方式常常会破坏受援国的所有权和制度。

《巴黎宣言》将援助国的援助符合受援国的政策和制度的程度细化为几个指标，包括对该国公共财政管理制度的利用，政府采购制度的利用，避免设立重复的项目实施单位（PIU），援助的可预测性，有附加条件的援助，以及协调国家发展战略与技术援助等。OECD-DAC援助国多年来持续公布其援助是否带有附加条件。然而其他指标只能在近几年通过DAC的巴黎宣言监测调查（SMPD）进行测量，而SMPD旨在评估该宣言中一些定量目标的实现程度。

我们使用2008SMPD的数据，即2007年度的数据。该研究为我们提供了大量的援助质量指标，因此将在下文就其方法和内容进行详细介绍。为调查指标的可比性起见，来自其他渠道的指标也是基于2007年的援助数据，在2008年的数据可获得时亦然。

SMPD中排除了人道主义援助、债务减免和行政成本。其主要包括ODA，也包含来自多边发展银行的非优惠贷款。DAC的调查仅涉及54个受援国，然而援助国的覆盖面则相当全面，包含了所有的DAC援助国、双边和多边援助机构，以及垂直基金（例如全球基金和GAVI联盟）和一些非DAC双边援助国（如波兰和匈牙利）。调查中将所有的联合国机构视为一体，采用UNDP收集和发布的数据。DAC调查所示的针对54个受援国的援助规模，与DAC官方援助数据库（来自截然不同的上报系统）中相应的援助总额非常接近。

受援国政府决定是否自愿参与该调查，这使调查的代表性受到质疑。事实上，54个自愿

参与的受援国与其他受援国在许多方面有显著区别：它们更大（指人口规模）、更依赖援助（以人均援助计算）、也更贫困（以人均GDP测量）。此外，完成重债穷国（HIPC）确认过程以寻求债务减免的国家数量也过多。控制上述四个变量，则调查的参与性与政治开放程度、经济政策和治理的质量、殖民遗产及所处地区等都没有相关性。

虽然受援国的（自我）选择并非随机，却不能断言样本会导致对某一援助机构绩效测量的偏差。调查中一些援助机构的援助所占的比例与其他机构（如在样本中为数不多的、向小国或不那么贫困的受援国提供援助的国家）相比可能会相对较大，所以后者的绩效指标的随机测量误差也可能更大。在上报数据的国家中，有时并不含来自多边发展银行的非优惠贷款（World Bank，2010）。我们也不能完全否认某些援助机构对SMPD定义的阐释更加随心所欲的可能性。2008年的调查通过各种手段对此进行了规避，如以2006年第一轮调查的经验为基础，完善定义、引导和调查过程；并将援助机构在实践中的自主权控制到最小。这些举措切实提高了援助机构和受援国上报数据的可比性（OECD，2008；2007），然而援助国在某些情况下可以自主判断（如定义其在特定国家的援助是否包括“对以项目为基础的路径的支持”），这种自主性可能会影响它们在某些指标上的得分（World Bank，2010）。

援助机构在SMPD的一致性和和谐性指标中的绩效，会受到受援国自身特征的影响（Knack和Eubank，2009）。所有的援助机构都会尽可能地不去依赖陷于腐败和其他弊端的受援国的国家系统，如SMPD记录在民主刚果的援助机构通过该国的公共财政管理（PFM）系统拨付的援助比例为0%，而在摩洛哥这一比例则为79%。与其在任一受援国如何运作这一因素相比，援助国在未经校正的“利用国家制度”这一指标中的绩效更多地受到自身国家总体政策的风险差异的影响。因此，如果指责援助国越过而非校正受援国的国家制度，无异于指责它们在风险环境中所做的努力——其实是对政策选择性指标的重复计算。

因此，在进行援助机构绩效测量时，我们针对其开展工作的国家的特征对SMPD指标进行了校正。“利用国家制度”及其他指标都在

援助机构—受援国层次上进行测量，于是能够通过受援国的固定效应，控制国家政策条件的差异。详细地说，就是采用PFM系统（和其他SMPD指标）对全部的受援国和援助国进行回归。对这些变量而言，我们的援助国绩效指标就成为援助国的相关系数，反映了援助国利用受援国国家制度的潜在倾向。

SMPD指标的偏差是由援助机构和受援国双方的特征共同决定的。每个指标都选择：1）所有援助机构虚拟变量，2）所有受援国虚拟变量或3）上述二者，进行回归。在一些案例中，受援国的特征比援助机构的特征影响更大。受援国虚拟变量能够解释PFM系统中偏差的28%，而援助机构虚拟变量的解释力则为21%。如果受援国和援助国虚拟变量不相关的话，即如果援助机构在援助调查中“好的”或“坏的”受援国方面没有显著差异的话，进行如上文所述的那样针对受援国效应进行校正，将会略微改变援助机构绩效的测量结果。然而对于利用PFM制度等指标而言，仅对援助机构进行回归与仅对受援国进行回归时出现了相同的偏差，即得到的 R^2 系数比针对二者进行回归得到的 R^2 系数要小，这些指标表明校正程序可能会对最终排序产生显著影响。就利用PFM制度而言，校正对许多援助机构的排序影响甚微，甚至毫无影响：例如爱尔兰和泛美开发银行仍然占据第一和第二的位置，世界银行也仅仅从第7位上升到第6位。然而国际货币基金组织（IMF）则从第16位攀升到第5位，希腊从第26位跃到第14位。而排名下降的例子包括丹麦从第3位降至第8位，挪威从第6位降至第12位，芬兰从第9位降至第17位，以及加拿大从第10位降至第20位。

通过校正，我们建立起“与国家制度一致”的援助质量分类指数，包括5个来自SMPD指标的维度：1）对国家公共财政管理（PFM）制度的利用；2）对国家政府采购制度的利用；避免设立重复的项目实施单位（PIU）；4）援助的可预测性；以及5）协调技术援助与国家发展战略。同时我们加入了6）无附加条件援助的《巴黎宣言》指标，该指标（依照标准程序）根据DAC的信用报告系统计算（OECD，2009b）。

对国家的PFM制度的利用，是通过3个分类指标的平均值进行测量的，即援助机构的援助占受援国政府预算执行额度的比例、财务报告 and 审计制度。而对采购制度的利用，也通过援

助比例来测算。协调技术援助，则是以在援助机构全部技术援助中所占比例进行测量。援助机构在利用PFM制度、采购制度和协调技术援助几个方面的绩效，可以使用每一组援助机构—受援国的援助正值的SMPD数据，通过固定效应回归，从而就受援国特征的差异做出校正。PIU的数量也同样可以通过受援国的固定效应进行校正，同时需要就援助机构在每个国家的援助规模进行校正，因为更大的援助规模无疑需要更多的PIU。在构建一致性这一分类指数时，所有变量都标准化为以0为平均值，以1为标准差。而PIU的标准值要乘以-1，从而与其他变量保持一致，即越大的数值代表越高的绩效。

援助可预测性以援助机构2007年度的预计援助额占该年度实际援助额的比例进行测量。在许多情况下，实际援助额都高于而非低于预计援助额。我们采用世界银行（2009）的方法测算绝对偏差，即二者差额占其中较大值的比例。与世界银行报告不同的是，我们根据受援国的固定效应对该值进行了校正，来代表受援国任何导致援助机构预算超支或预算使用不足的特征。需要注意，这种测量可预测性的指标只能用于短期条件。Kharas（2008）强调了中期内援助的波动性这一重要问题，认为来自大型援助机构的援助会导致受援者的变化。当然，我们的研究样本中的多数援助机构及其他援助机构排序都并未出现在其研究中。

无附加条件的援助是以上报到CRS的2007年度援助承诺中，没有要求、也没有部分要求从援助国购买的比例进行测量的（OECD，2009b）。据估测，有附加条件的援助可能会使采购成本增加15%到30%（Jepma，1991）。许多国家不会上报其援助活动的附加条件。我们选择采用Easterly（2002）的方法，没有忽略那些未上报是否带有附加条件的援助，而是将其视为有附加条件。这类援助与技术援助不成比例，技术援助中有附加条件的比例更高。没有上报附加条件的动机是掩饰带有附加条件的援助比例过高的事实（OECD，2009b）；因此如果忽略所有未上报附加条件的援助，无论从分子还是分母来看，都会夸大无附加条件的援助所占的比例。在面临过高还是过低地估计透明度不高的援助机构的无附加条件的援助的比例时，我们宁可选择后者。二者之间，过低的估计与我们其他的援助一致性指标具有更高的相

关性，能够提高这一分类指数的内在信度。而非双边援助机构的所有援助都不带任何附加条件这一现象，是否值得表扬仍莫衷一是。它们没有能力提出任何附加条件，而它们建立的初衷恰恰是避免带有附加条件的援助遭遇的那些问题。例如，英国的援助取消了全部的附加条件，这是一个长足的进步，表明了其援助绩效的改善。因此在构建一致性分类指数和援助机构绩效的综合指数时，我们将多边援助机构的无附加条件的援助编码为缺失数据。

在援助一致性这一分类指数下的得分及援助机构排序，由6个指标计算得出，其相关系数的平均值为.27，而内在信度的 α 系数则为.68。

根据受援效应校正指标，旨在提高测量援助机构的援助质量的指标效度。然而援助一致性指数下的几个指标之间的相关性增强，并不一定会随之改善信度。如果校正能够消除没有相关性的指标带来的随机测量误差，那么该指数的内在信度应该提高。但如果校正主要影响到各个指标都具有的共同偏差，信度就会降低。那么，即使牺牲信度也难以改善效度：当使用校正过的指标构建4个分类指数时，内部相关性的平均值及 α 系数会增加或保持不变。

国家内部的和谐性

援助机构在《巴黎宣言》中承诺，在降低由援助机构和受援国政府导致的交易成本的目标指引下，实施统一的安排并简化程序。SMPD的测量中，用到了3个围绕国家主导型项目的援助机构活动的和谐性指标。第一个指标是通过以项目为基础的路径（PBA）提供的援助的比例，其中不仅包括直接的预算援助，也包括以政府主导的部门倡导路径（SWAp）形式实施的项目援助。从概念上看，这一和谐性指标显然与协调国家所有型发展项目下的援助及采用PBA并借助国家PFM制度和采购制度提供的所有直接预算支持这个问题有所重叠。

《巴黎宣言》中关于和谐性的其他两个指标是通过SMPD测量的，即与其他援助机构协调开展的任务和国别分析研究的比例。在和谐性这一分类指数下，上述3个指标之间相关系数的平均值为.30，得出的内在信度的 α 系数为.56。

专业性

通常认为，大多数援助机构的援助分散在

太多国家、部门和项目（World Bank, 2003）。援助机构和援助项目的急剧增加，过度扩张并损害了受援国政府管理援助的能力及其与援助机构的关系（Knack和Rahman, 2007）。援助机构在专业化方面畏首畏尾，又削弱了各国、各部门的专业能力，从而降低了援助效果，连并未卷入援助机构和援助项目急剧扩张的受援国也未能幸免。因此，我们的专业性分类指数从概念上有别于在受援国的和谐性实践指数。

《巴黎宣言》呼吁在援助机构之间进行更有效的劳动分工，声称“在全球、国家和部门层次上，援助的过度分散有损于援助效果”。援助机构承诺成立一个权威组织，负责以比较优势为前提，指导各个援助机构的实践。该宣言并未提出任何指标，在SMPD中也不曾提及。我们采用Acharya, Fuzzo de Lima和Moore（2006）、Knack和Rahman（2007）、Easterly和Pfütze（2008）以及Roodman（2006, 2009）的方法，借助DAC的数据库构建了几个2007年度的指标。

首先，我们计算了每个援助机构在2007年度除了债务减免、人道主义援助和行政成本之外的ODA总额的地理集中程度。这一指标以Herfindahl指数呈现，等于一个援助机构对其各受援国提供援助占援助总额比例的平方之和。大型的援助机构通常活动的地理范围更大，在更多的国家（部门）拥有更多的专业知识，其援助的集中程度就相对较小。而地理集中程度与ODA（的对数）的相关系数为.49，所以如果使用未加校正的指标进行绩效测量，那么小型援助机构就会因其规模较小而在排序中占得先机。因此我们在援助机构绩效指标中，计算了对2个变量进行回归的残差，即1）援助机构总ODA的对数和2）援助机构活动的地理范围中所有援助国所占的比例。大多数援助机构的这一比例为1，然而对于地区发展银行（AFDB, ASDB, IDB和EBRD）而言，则小于1。

我们同样采用Herfindahl的方式，以来自援助数据库（AidData）中各个援助机构2007年度的承诺值，测量了援助的部门集中程度。在这个测量中，即使一个援助机构在某个受援国的集中程度低，也不能说明该受援国的援助较为分散，因为这个援助机构可能对众多缺乏其他受援渠道的效果提供了援助。然而，Acharya, Fuzzo de Lima和Moore（2006）的研究表明，

这种现象极为罕见，可以排除。援助分散程度最高的援助机构，其受援国往往拥有众多的援助渠道（因此从受援国的角度看，援助更为分散）。

Easterly和Pfitze采用援助数据库和DAC的CRS数据库，加上来自其他援助机构的信息，对部门分散性进行了计算。在我们的援助机构绩效指标中，我们对2个变量的部门集中程度进行回归的残差来测算：1) 援助机构在援助数据库记录中的2007年度承诺援助总额的对数；和2) 在有限部门（即全球还是“垂直”基金）活动的援助机构的虚拟变量。最终指标计算的结果是，与在地理集中程度的绩效指标正相关，虽然相关程度并不显著（.23）。

这些关于集中程度的指数并不能很好地反映援助的分散状况，原因至少有二：首先，当一个援助机构增加其援助国家或部门数量却仅给它们分配极少的援助资源时，这些数据值就会下降。例如，如果一个拥有10亿美元年度预算的援助机构，为20个新的受援国分别提供50万美元的援助，不会对该援助机构的集中程度指数造成明显影响，然而该援助机构及其受援国的行政成本却会大幅增加。其次，在全球层次上计算的部门集中程度，可能会与在受援国层次上的计算结果大相径庭，对援助质量的判断也就截然不同。假设援助机构A在教育领域向10个受援国分别投入100万美元，在健康领域向另外10个受援国也分别投入100万美元，而援助机构B则分别向20个受援国的上述两个领域分别投入50万美元，那么援助机构A和援助机构B的部门集中程度均为0.5，虽然援助机构A在每个受援国仅援助一个部门，而援助机构B则在每个受援国援助了两个部门。在这个例子中，援助机构B更有可能导致项目的扩张和受援国行政负担的增加。

要解决这种集中程度测量的不足，我们借助PLAID数据系列计算了一个以援助机构为单位的指标，即每个受援国接受援助的部门的平均数量。我们对援助机构的援助承诺总额的对数和有限部门领域活动的援助机构的虚拟变量进行回归，将得到的残差作为绩效指标。基于2007年度每个援助机构的受援国数量（采用DAC的援助分配数据）能够得到类似的指标，该指标以援助机构的援助额度和地理范围进行了校正。在排除援助额度的共同作用之后，受

援部门和国家数量的相关系数为.48；以援助额度和地理范围进行校正之后，相关系数为.37。

在援助专业性方面的第5个指标是基于援助数据库的项目平均规模，采用2007年度的援助承诺数据。Roodman（2006；2009）Knack和Rahman（2007）指出，不同的援助项目迅速复制会为受援国带来许多不必要的交易成本，削弱其行政能力。Roodman提出了一种关于项目生产能力的独到假设来评估将援助分散到过多的项目中去导致的效率损失。例如，他假设在一定的援助额度下，最优的项目规模将在治理状况较好的国家发挥更佳的效益，因为玩忽职守产生的边际效应更低。与Roodman不同，我们能够更灵活地使用这些数据，而无需就效率损失做出任何预测。我们就所有的援助机构、受援国和部门虚拟变量，及除项目*i*之外援助机构提供的援助总额（的对数），对项目*i*的规模（以当前美元价格的对数）进行了简单回归。如果出于治理不善或其他考虑，所有的援助机构都倾向于在特定的受援国分散其援助资源，那么就需要控制受援国变量。而同样地，也需要控制部门虚拟变量，以避免援助机构在最优规模较大的活动中的参与引发的系统差异。所以在援助机构变量相关系数中，特定援助机构的影响就不受其预期开展活动的受援国或援助部门特征的影响。

我们发现，项目规模与援助机构向受援国提供的其他援助的总额呈现显著的正相关。即使排除这种规模效应，全球层次上的援助机构虚拟变量相关系数仍然受到其他规模效应的影响。即使控制援助机构在项目所在受援国的援助规模，仍然呈现出预算规模更大的援助机构项目规模越大的规律。因此，需要对我们的项目规模指标做进一步的校正，以来自援助数据库的援助机构2007年度援助承诺数据总和的对数，对援助机构虚拟变量相关系数预测值进行回归，取其残差作为描述平均项目规模的绩效指标。我们计算出的项目规模指标，与Roodman的2007年“规模权重”指标的相关性为.73。

《巴黎宣言》呼吁援助机构将援助工作交予在某些国家或部门活动更为频繁的同僚。SMPD和其他任何调查都未就此收集任何数据，然而普遍的共识是，由多边援助机构担此重任能够事半功倍。通过多边机构分配预算，能够减少上文提到的援助分散问题。多边机构

通常有特定的地理和部门活动范围，通过它们进行预算分配会削弱现有援助关系中的援助机构—受援国—援助项目三位一体的模式。与大多数双边机构（特别是新兴的、小型的援助机构）相比，多边机构的特殊使命也赋予了它们在更多国家和地区开展工作的专业知识。因此，我们的第6个援助专业性指标即，每个双边援助机构的ODA总额中通过（除欧盟外的）多边援助机构拨付的比例。对于多边援助机构而言，这一指标编码为缺失数据。

援助专业性的最后一个指标为行政成本在援助总额中所占的比例，以各援助机构向DAC上报的数据为准。过多地在受援国、部门、项目和其他维度上分散援助的援助机构，其行政成本将会更高。

我们将上述7个指标标准化，取其绝对值（如行政成本、受援国数量及受援部门数量平均值等）并求平均数，构建了援助专业性这一分类指数。这7个标准化指标的内部相关系数为.26，于是此分类指标内部信度的 α 系数为.71。

其他指标

我们没有选择一些研究者建议的其他指标。例如Mosley（1985）、White和Woestman（1994）及Bulow和Rogoff（2005）出于透明度及其他考虑，向发展银行建议以贷款的形式代替赠款，这种援助归还期限的慷慨性（即赠款是援助的“最佳”形式），也是援助质量的指标之一。我们同意前人的观点，然而正如Easterly（2007）指出，赠款“可能会引发其他问题，如无法催生归还贷款的长期投资等等”。因此我们认为，选择赠款还是贷款会是一个相当模糊的援助质量指标。这个问题更多地涉及援助数量，而Roodman（2006；2009）在进行定性校正之前，就简单地排除了援助数额的利息和本金。

Easterly和Pfitze将粮食援助和技术援助也做为不良援助的指标之一，认为这两种援助倾向于从援助国购买。他们（和我们）已经对无附加条件的援助进行了直接测量，所以粮食援助和技术援助通常（但未必总是）带有附加条件的事实对援助机构的影响意义不大。他们指出，技术援助“常常反映援助国而非受援国的优先考虑”，然而这个问题已经在我们的排序中，以技术援助所占比例与受援国项目的一致

程度这一SMPD指标的形式给予了关注。

除附加条件和缺乏协调导致影响甚微之外，很难对更倾向于提供技术援助的援助机构提出更多指责。Annen和Kosempel（2009）在其援助增长效应模型中指出，技术援助比其他形式的援助更为优越，它更加能够促进知识转移。援助机构在《巴黎宣言》和《阿克拉行动议程》中承诺帮助受援国提高其统计能力，完善其公共财政管理、采购和信息发布制度。而援助机构在这些领域的援助多采取技术援助的形式。

结果：援助机构得分和排名

我们对每个援助机构就上述4个分类指数中的18个指标分别进行打分排序并计算总分。

分类指数排名

在进行综合排序之前，有必要首先进行分类指数排序。在判断援助机构表现优劣方面，与反映各个截然不同的领域的行为的综合排名相比，这些分类指数更加能够反映切实的绩效。

首先，在选择性分类指数方面，世界银行排名第一，亚洲开发银行（ADB）位居第二，排名第三的荷兰在所有双边援助机构中则占据首位。我们对2007年度援助分配的排序结果与Dollar和Levin（2006）在2002年进行的排序结果大体相同。在这两年都表现良好的援助机构包括世界银行、IMF、丹麦和荷兰，而美国、葡萄牙和希腊却两度垫底。一些地区发展银行在两次排名中的成绩差别很大。由于我们的排序经过了校正，亚洲开发银行和非洲开发银行在我们的选择性排序中比在Dollar和Levin（2006）的排序中明显靠前。

在一致性方面，爱尔兰排名第一，随后是亚洲开发银行、全球疫苗免疫联盟（GAVI）和丹麦。日本在这一指数中排名很高（第8名），然而在其他3个指数中都表现不佳。联合国、美国、奥地利、葡萄牙、希腊，以及几个新兴的双边援助机构排名最后。葡萄牙和希腊的援助中，未上报是否带有附加条件的比例最高，如何处理这些缺失数据就直接影响到了它们的排名。如果将未上报数据按照上报数据中的比例计算带有附加条件的援助比例，那么希腊的排名将从第32位上升至第28位，而葡萄牙仍在38个援助机构中位于第37位。在处理这些数据时，

我们坚持按照OECD-DAC的传统做法，将部分带有附加条件的援助视为带有附加条件而非不带附加条件（OECD-DAC，2009）。除了都要从发达国家而非发展中国家采购商品和服务之外，部分带有附加条件与完全带有附加条件截然不同。大多数援助机构的大多数带有附加条件的援助都是完全带有附加条件。所以如果将部分带有附加条件的援助视为不带附加条件的援助，不会显著影响一致性指数下的排名。欧盟上报到CRS的所有援助都部分带有附加条件，所以其不带附加条件的援助比例会从0%直接上升至100%，排名也从第31位升至第21位。

在和谐性分类指数下，爱尔兰继续占据首位，丹麦和GAVI紧随其后。英国、澳大利亚和（特别是）联合国在这一指数下的得分要高于其他3个分类指数。亚洲开发银行排名第14，而在其他3个指数下则排到前两位。这一指数排名中垫底的是8个小型双边机构和日本（第35位）。

在专业性分类指数下前七位中有5个是多边机构，另外2个是新兴的小援助国（波兰和匈牙利）。有3个双边援助国在这一指数下的排名不高，而在其他3个指数中的排名则较为靠前，即丹麦（第14位）、荷兰（第18位）和爱尔兰（第23位）。几个大型援助机构在这一指数下表现较好，即欧盟（第10位）、法国（第12位）和美国（第17位）。而一些新兴的小的双边援助国（希腊、土耳其、韩国和捷克）和EBRD在这一指数下排名垫底。

综合排名及权重的影响

如果对4个分类指数不做加权处理，直接计算平均值得到的“综合”得分，那么位于前5位的分别是亚洲开发银行、世界银行、丹麦、爱尔兰和GAVI，而最后5位则是葡萄牙、捷克、韩国、希腊和奥地利。总体来看，大型援助机构的综合得分相对较低，例如美国（第32位）和联合国机构（第30位）。

在进行排序的38个援助机构中，有22个是OECD-DAC成员（包括欧盟）。这22个援助机构的综合得分平均值为-.08，比6个非DAC成员的双边援助机构（包括韩国，其在2010年1月才加入DAC）的平均值-.63要高得多（ $p=.06$ ）。而对于6个多边发展银行（包括IMF）而言，平均值为.78，比属于DAC成员的双边援助机构的平均值也要高得多（ $p=.01$ ）。其余4个援助机

构（UN，IFAD，全球基金和GAVI）则与DAC双边援助机构没有显著差异。

在所有援助机构在4个分类指数下的排序结果中，我们取一个援助机构在某一分类指数下得分与其他3个分类指数下得分之差（取绝对值）的平均值，两两比较，计算6对排序差异的平均值。在4个分类指数的排序中，排名较高的（如世界银行和IFAD）和排名较低的（如希腊）援助机构都有排名变化不大的例子。排名变化显著的援助机构多是新兴的小型双边援助机构，例如匈牙利和波兰等，当然也不乏多边机构，如非洲开发银行、EBRD和联合国等。对这种排名变化最显而易见解释是这些援助机构在各个方面的作为的确存在差异，但同时也是测量援助质量这一宽泛概念时的不确定性和不精确性的反映。而对那些排名变动较大的援助机构而言，我们的测量手段能够反映其“真实”排序结果的不确定性更加凸显。

对排序结果的另一个质疑是，排名采用了单一来源即SMPD的数据，其提供了和谐性指数中所有3个指标，以及一致性指数中6个指标中的3个的数据来源。然而，在与一致性分类指数得分的相关性方面，无附加条件的援助指标与5个SMPD指标的相关性几乎同等显著（前者为.60，后者的均值为.63）。此外，这两个主要由SMPD数据支持的分类指数与综合指数的相关性（.77）不及选择性分类指数与综合指数的相关性（.83）。唯一与综合指数显著相关的指标并非这8个来自巴黎调查的指标，而是贫困选择性指标（.82）。如上文所述，我们不能完全排除一些援助机构随心所欲地理解调查要求，从而得到较高的绩效分值的可能性。但即便如此，由此导致的任何偏差都不足以对综合排序结果产生明显影响。

我们进一步分析了不同的权重选择对综合排名的影响。在这次综合排序中，我们对所有的指标而非分类指数求平均值。随后加入了Easterly和Pfitze（2008）的透明度指标和间接成本指标，以总计20个指标重新计算排序。显然，这两种计算方式中，都对于指标更多的分类指数赋予了更高的权重，即专业性指数和一致性指数。方法和内容的这些变化会对排序的结果造成影响，但总体情况变化不大：位于前5位的援助机构在先前的排序中均位列前八，而位于最后5位的援助机构在先前的排序中也都

位于最后8位。在双边机构中，仅有2个小型援助国（匈牙利和卢森堡）在两次排序中变动超过6个位置。EBRD是唯一变动较大的多边机构，由于其在专业性指数（该指数占20个指标中的7个）和Easterly和Pfitze（2009）的透明度指数中得分较差，两次排序的名词相差10位以上。

我们引入了分配指数进行进一步讨论。如上文所述，分类指数并非综合指数（即“点名批评”）的改头换面，而是指向截然不同的目标：帮助指导援助资源的分配。在计算分配指数时，考虑到援助机构无法控制的因素（特别是地理活动范围），我们没有进行如前几次排序中那样的校正。我们将多边援助机构的无附加条件的援助编码为100，因为它们的援助没有附加条件这一事实应该成为双边机构、欧盟和多边机构的榜样。尽管目的不同，分配指数的排序结果与综合指数的排序结果并未呈现出显著差异。仅有2个地区银行，即IDB和EBRD的排名上升超过5个位置。由于这两个地区银行多在中等收入地区活动，在未就地理范围进行校正的情况下其贫困选择性必然受到影响，这样的结果也就合情合理。在4个分类指数中，选择性指数受到的影响最大，但与先前排序结果的相关系数仍然达到.84；对其他3个分类指数而言，相关系数则从0.94到1.0不等。总而言之，采用分配指数的方法并没有显著影响排序结果，只是将几个双边机构的排名提升到2个地区发展银行前面而已。

有无数种方法对指标进行校正和综合，其中一些可能与我们的思路不谋而合。例如，可以基于组成结构、各组成部分产生的影响或因素分析的方法，对指标或分类指数赋予权重。我们尝试了这种方法，结果表明会降低专业性指数及其各个指标的权重，在该指数中表现上佳的援助机构就会受到影响。然而，以主要因子分析或因素分析的方式，打分排序的结果则不会在整体上受到多大影响。

与其他援助质量和数量指数的对比

我们将自己的综合指数排序与Easterly和Pfitze（2008）以及Roodman（2006；2009）的援助质量指数进行了对比。Roodman的工作是全球发展中心年度发展承诺指数（CDI）的一部分，他的假设坚定而合理，强调援助的选择

性、援助的附加条件，认为项目扩张会影响援助的效率，从而对以质量校正的援助额进行测量。虽然最终用于CDI的是关于援助数量的指数，但Roodman仍然构建起一个援助质量指数，即上述以质量校正的援助额除以“净援助额”或未经质量校正的援助额。

Easterly和Pfitze（2008）采用“最佳援助实践”指数对援助机构进行了排序。他们的选择性指数（基于贫困、腐败和民主）和专业性指数（地理、部门）与我们的这两个指数十分相似，虽然其专业性指数的设计更有利于小型援助机构。他们的第三个指数是“无效援助渠道”，提供有附加条件的援助、粮食援助和技术援助的援助机构在该项中得分较低。将粮食援助和技术援助划分为无效援助渠道的原因是它们大多要求从援助国购买商品和服务，所以可以将其简单地称之为援助附加条件指数。他们的其他两个分类指数时基于最新的间接成本和透明度数据，分别通过提供稳定的、可及的运营成本信息以及向OECD-DAC上报全部援助数据进行测量。虽然透明度和间接成本至关重要，特别是对援助研究的意义不言而喻，但毕竟不是援助效果辩论的核心议题，也并未体现在《巴黎宣言》中。此外，Easterly和Pfitze也承认要想精确测量这两个指标，的确是困难重重。

虽然我们采取不同的综合指数得到了类似的排序结果，却与Easterly和Pfitze以及Roodman的排序结果有较大差异。我们的“综合”指数与Roodman排序结果的相关系数是.60，与Easterly和Pfitze排序结果的相关系数则是.25。在我们的综合指数中排在前三位的援助机构中，仅有世界银行在他们的两次排序中也能位列前五。亚洲开发银行的排序勉强能够持平，而另外3个援助机构在Easterly和Pfitze的排名中十分靠后。就排名较差的援助机构而言，CGD的排名与我们基本一致，却与Easterly和Pfitze的排名截然不同：在大型援助机构中，美国、日本和法国的排名都上升了至少15位。这3个援助国在透明度和间接成本中的得分较高，而我们则没有采用这两个指数。

我们的指数与Roodman指数相关性并不显著，而与Easterly和Pfitze指数的相关系数则非常低，并不意味着我们的指数在测量援助质量时具有更高的效度。值得注意的是，他们的指

数之间的相关系数更低（.19），可能是由于Easterly和Pfutze对援助机构绩效的测量超出了常规的援助质量的维度，他们对透明度和间接成本赋予了较高的权重，导致其排序与其他排序之间的相关性较弱。

与其他排序方式相比，我们的排序方法更好地契合了“内容效度”（Carmines和Zeller，1979）的标准。也就是说，我们的18个指标更全面、更典型地反映了《巴黎宣言》和关于援助效果的文献中的“最佳”援助实践。我们控制与绩效无关的变量（例如在测量援助的地理集中程度时，援助机构规模这一变量）所进行的校正，也提高了指标和指数的效度。

最后，与Easterly和Pfutze的工作相比，我们的综合指数和分类指数具有更高的内部信度。虽然指标和分类指数之间较高的内部信度并不意味着它们就一定能够更好地测量援助质量这个宽泛的概念，但内部信度较低无疑更令人心存疑虑。我们的4个分类指数之间的内部相关系数为.37，而Easterly和Pfutze的5个分类指数之间的内部相关系数则仅为.14。在指数信度的 α 系数方面，我们的指数为.70，而Easterly和Pfutze的指数则为.44。我们的4个分类指数中，内部相关系数均值最低的为.26，而Easterly和Pfutze的5个分类指数中，有2个的内部相关系数均值都是负值。

援助质量与相对于援助国的国民总收入（GNI）而言的援助数量（均采用2007年的数据）之间存在怎样的相关性？在双边援助国中，二者存在显著相关性，相关系数达.66。换句话说，致力于提供大规模援助的国家，也会采取与国际社会要求相应的方式来保证援助质量。

结论

本文介绍了测量双边和多边援助机构提供的援助质量的一个新指数。该指数包含了更广泛的指标，覆盖了援助领域最具共识的4个主要方面：选择性、一致性、和谐性和专业性。其中借鉴了前人的指标，采用最近对OECD-DAC的《巴黎宣言》的监测调查得到的新数据和援助数据库的项目数据进行测算。在有必要的情况下，我们对原始指标进行校正以排除外部因素的影响，以不牺牲透明度为前提切实提高效度。此外，我们还就援助质量的上述4个维度分别构建了分类指数，经验表明各个指数中所包

含的指标高度相关，能够用于测量同一现象。

可以借助这些指标和分类指数，以及最终的综合指数，对不同援助机构的援助质量加以比较。诚然，如上文所述，各个单独的指标可能更具说服力。将各个指标综合起来难免会涉及权重的问题，而不同的权重无疑会影响最终的排序。与综合指数相比，独立的指标和分类指数能够更加明确地进行操作。此外，独立指标也不能追踪历时变化，建立在各个标准化指标的基础上的综合指数也只能用于比较各个援助机构之间的相对绩效。

同时，我们相信综合指数必有用武之地。针对单一指数的排序结果能够对排名较低的援助机构施加压力，促使它们寻求提供援助的“最佳实践”。如果没有这样一个综合指数，这些援助机构就会仅仅关注得分较高的几个指标，裹足不前。综合指数恰好能够避免这个问题。

在设计指标时，我们已经尽可能全面地覆盖《巴黎宣言》中关于援助质量的定义，然而其中一些因素目前很难测量，被排除在最终的指数之外。例如将援助任务交予其他援助机构可能会影响某个援助机构的和谐性和专业性，但缺乏系统的测量方式。而我们对专业性的测量更多地关注数量而非质量：专业性应该在援助机构具备比较优势的方面有所体现，而我们缺乏数据对此进行评估。另一个难以定性测量的方面超出了《巴黎宣言》的内容，即援助的知识维度：援助机构对我们理解和解决发展问题的贡献。在数据上报和运营成本公开中的透明度只是援助机构信息提供功能的一小部分而已。一些援助机构对其项目进行了系统评估并公布结果。理想状况是对每个援助机构的工作、研究和国别分析进行质量、数量和透明度的评估。与援助质量相关的另一个维度更加难以定性分析，即援助者实践新方法的意愿：创新能够带来学习，一些援助机构更具创新精神，另一部分则仍在保守观望。如果加入所有这些维度，可能会改变我们的排序结果，这也更加提醒我们不要对排序的微小差异大惊小怪。

本文的目的是贡献于对援助质量和援助机构排序的测量，最终改善援助机构的绩效。我们无意贬低其他排序方法和定性研究的努力，这些成果仍在持续产生影响。CGD/Roodman指数贡献于CDG发展承诺指数的定期更新，而

Easterly和Pfutze的排序也即将更新。此外，Brookings/CGD的“官方发展援助质量”测量也已经启动。在下一轮巴黎宣言监测的数据可用时，我们也同样希望更新自己的指数和排序。从长期来看，我们认同由各个独立的智库和研究者而非援助机构工作人员开展排序工作的重

要价值，并希望本文能够在数据来源、方法运用、结果阐述等方面影响排序研究，并对排序结果的使用者予以启发。

来源：<http://www.worldbank.org>

全球监测报告 2010：危机之后的千年发展目标进程

世界银行和国际货币基金组织（IMF）发布的《全球监测报告 2010：危机之后的千年发展目标进程》这份年度报告系列中的第七期，从发展的角度审视全球经济危机，评估对发展中国家的影响，包括对其经济增长、减贫进程以及其他千年发展目标的影响，并提出了发展中国家和国际社会政策响应的优先重点领域。

据前几年的全球监测报告反映，2000 年设立的千年发展目标在 21 世纪的最初几年取得了重大进展，但自 2006 年起实现难度逐渐增大，而全球经济危机的影响使得这些目标的实现更不容乐观。2010 年的报告指出，发展中国家的减贫目标虽然可以实现，但预计危机给发展造成了长期严重危害，实现其他千年发展目标进程受损，包括饥饿、儿童与孕产妇健康、性别平等、清洁水源、疾病防控等，并将继续影响 2015 年之后的长期发展前景。

全球经济危机让人类付出了什么样的代价？《全球监测报告 2010》审视了大萧条以来最严重的经济衰退对发展中国家的贫困状况和人类发展带来的影响。虽然复苏已经开始，但危机的影响将是长期和难以度量的。危机前各国在减贫方面取得了显著进展，而受到危机的影响，未来的减贫速度将会放缓，非洲的低收入国家受到的影响会最为严重。在危机面前，所有发展中国家的居民都难以幸免。即使是收入较高的发展中国家，生活水平在日均 1.25 美元贫困线以上的居民也都在想方设法购买较便宜的食品，推迟其他物品的购买，减少上医院的次数，延长工作时间，或者通过做兼职贴补家用。

危机对人类发展指标（HDI）也将产生严重影响：

- 到 2015 年，撒哈拉以南非洲的极贫人口将会增加 2000 万，全球的极端贫困人口将会增加 5300 万，这种情形将会持续到 2020 年。

- 2009 年至 2015 年间 5 岁以下儿童和婴儿死亡人数将分别增加 120 万和 26.5 万人。

- 到 2015 年无法完成初等教育的学龄儿童将增加 5 万人。

- 由于危机的影响，到 2015 年无法获得清洁饮用水的人数将增加 1 亿。

报告指出，如果听任贫困国家的经济复苏自由发展，任由危机造成广泛的国内政策失误和制度崩溃，对人类发展带来的负面影响，尤其是对妇女儿童负面影响将会是灾难性的。

国际金融机构和国际社会对危机做出了迅速和强有力的反应，但要维持复苏和恢复千年发展目标进程的势头，还需要付出更大的努力。发展中国家也需要推进重大的政策改革，加强体制机构，从而在财政紧张的情况下力争提升服务提供的效率。然而，与以往的危机不同，此次危机并非是由发展中国家的政策失误造成的。所以，发展成果的改善也取决于全球经济能否快速复苏，继而改善出口状况、贸易条件和负担得起的资本流动，同时满足对低收入国家的援助承诺。虽然对发展中国家的国际援助在 2008 年达到创纪录的水平，但援助量仍远低于预期，经济危机给捐助国本身造成的财政影响也将严重影响他们实施国际援助。因此，世界银行建议，国际金融机构之间需要加强协调，通过提升援助的有效性改善捐助活动，减少附加条件的援助份额等。

令人担忧的是经济危机牵制了发达国家政府的主要精力，其对发展中国家的援助出现了停滞的势头，而在全球金融市场动荡、进出口贸易变化剧烈的情况下，各个发展中国家的增长势头也遭到了遏制。欧洲国家在危机之后的中期阶段可能减少发展援助资金，而这个时期恰恰是低收入国家最需要援助资金的阶段。低收入国家如何扭转危机之后中期阶段的不利趋势，维持复苏和恢复千年发展目标进程？世界银行建议，发展中国家必须继续保持良好的政策，改善服务质量和效率，提高国内社会安全网络的效力。

来源：<http://www.worldbank.org>

转移支付：促进家庭发展

长期贫困研究中心

转移支付项目以减少贫困和脆弱性为目标，瞄准最穷的地区，对个人或家庭造成直接的影响。越来越多的研究表明，此类项目在促进家庭消费和提高家庭生产能力方面是成功的，并认为，大规模的转移支付项目能够支持贫困家庭和贫困社区的微观层次的增长。

阻碍贫困家庭发展的因素

贫困家庭在自身财富积累的投资方面的能力较弱，特别是面对市场的能力较低，而且难于获得贷款。同时，由于健康和教育方面的困难使他们无法提高自身能力。因此，穷人的生计是极度脆弱的。

转移支付具有刺激微观增长的潜力

如果进行合理的设计，转移支付项目能够促进家庭增长。

改善获得贷款的情况

如果贫困家庭能够获得定期的转移支付，金融机构将对其还款能力更有信心，减少贷款风险方面的考虑。这一情况曾在巴西得到证明。同时，通过促进家庭储蓄和投资，转移支付本身也作为提高家庭财产积累的有效方式。因此，有些时候，通过家庭生产的状况的改善，受益人的财产增长远超过了转移支付本身的额度。

提高生产能力

证据表明，很多家庭将转移支付的收入用于教育和健康方面的投资，特别是对于孩子。虽然低收入国家的社会服务利用情况受到供给不足的影响，但对于人力资源的投资预计会在总体上提高人们将来的生产能力和收入。

提供更高的确定性和安全感

脆弱性和不安全感导致了贫困家庭在资源分配方面的低效率，使他们更多地关注低风险、低回报的生产活动，并为应对风险而更多地持有现金，导致投资减少。不安全感同时也让他们更为注重当前消费而非长远投资。转移支付

能够通过提供补充性的收入来减少限制因素，从而减轻消费风险，并通过其综合保障的特点保护贫困家庭的消费、财产和投资安全。

促进家庭资源分配的改善

支出决策是影响家庭增长的关键性因素。通过对家庭关键成员的影响，转移支付能够使家庭支出更具生产性。特别是在转移支付通过妇女进行使用的情况下，她们更加倾向于对于孩子教育或农业生产方面的投资。

对于家庭增长的促进依赖于转移支付项目本身的设计

项目设计是转移支付项目影响家庭增长的关键因素，下面列出了一些主要观点：

- 如果转移支付项目的目标是增加贷款渠道和促进生产投资的多样化，那么转移支付必须是定期的、可依赖的，其期限必须足够长以能影响家庭的消费—投资决策。

- 项目的水平要能够反映项目目标。支付额度低的项目在影响家庭投资方面的作用将是极其有限的。

- 支付方式要能够改善家庭的资源分配，通过妇女进行支付通常能达到更好的效果。

- 转移支付影响家庭增长的效果很大程度上依赖于对人力资源的投资，因此，健康、教育等相关的社会服务的供给显得尤为重要。

- 补充性的财产积累干预同样能够更好地促进家庭增长。孟加拉国将转移支付项目与小额信贷和技术培训等相结合，取得了很好的效果。

- 外在的财产保护措施也能够有效促进转移支付的影响效果。转移支付具有较强的社会保障性质，将保险纳入相关项目之中对于项目效果的促进将是有效的和值得的。

来源：www.chronicpoverty.org

· 机构介绍 ·

欧洲复兴开发银行 (EBRD)

欧洲复兴开发银行 (EBRD) 成立于 1991 年, 宗旨是在加强民主、尊重人权、保护环境等前提下, 帮助和支持东欧、中欧国家向市场经济转化, 以调动上述国家中个人及企业的积极性, 促使他们向民主政体和市场经济过渡。目前欧洲复兴开发银行在资金来源于 61 个国家和 2 个政府间组织, 在中东欧地区的 29 个国家开展工作。

欧洲复兴开发银行的主要使命集中在三个领域: 项目投资、政策对话和咨询服务。欧洲复兴开发银行依据其受惠国在不同阶段之特殊需求, 协助受惠国进行结构及产业性经济改革, 提升竞争力、私有化及企业精神, 并借由投资增进受惠国私人产业活动、强化金融机构及法制、发展基础建设、振兴私人产业。为促进变革, 欧洲复兴开发银行鼓励公私部门共同融资及外国直接投资, 协助筹集国内资金, 并提供相关领域之技术合作。欧洲复兴开发银行与国际金融机构及其它国际或国家组织密切合作, 并提倡合乎环保的可持续发展。同时欧洲复兴开发银行也活跃在国际市场, 投资于信息产业和当地资本市场的发展。

欧洲复兴开发银行的主要业务包括:

- 提供必要的技术援助和人员培训;
- 帮助受援国政府制订政策及措施, 推动其经济改革, 帮助其实施非垄断化、非中央集权化及非国有化;
- 为基本建设项目筹集资金;
- 参加筹建金融机构及金融体系, 其中包括银行体系及资本市场体系;
- 帮助支持筹建工业体系, 尤其扶持中小企业的发展。其经营方针是“发展银行业务及商业投资银行的业务兼顾”。
- 欧洲复兴开发银行的主要研究工作集中在经济领域, 由首席经济学家主持, 主要针对项目移交、国别战略和部门政策及其他相关的策略、政策、操作性问题开展工作。其主要出版物包括其年度报告和《可持续报告》。

2010 年的欧洲复兴开发银行第 19 次年会有超过 2000 名代表参加, 共同探索了在经济复苏的情况下受援国的需求, 并对中东欧地区的商业机会进行了评估。

主 办：中国国际扶贫中心

编辑：《国际减贫动态》编辑部

协 办：中国农业大学人文与发展学院

地址：北京市朝阳区太阳宫北街 1 号

主 编： 吴 忠

邮编：100028

副主编： 黄承伟

电话：010-84419641

 李小云

传真：010-84419658

本期编译： 唐丽霞

电子信箱：zhangdeliang@iprcc.org.cn

 赵丽霞

网址：www.iprcc.org.cn

责任编辑： 张德亮