

3.模型和方法

3.1 环境库兹涅茨曲线（EKC）

研究环境污染与经济增长之间关系的理论之一是 EKC。EKC 由美国学者格罗斯曼和克鲁格（1991）提出的。他们发现经济增长与环境污染之间存在倒 U 型关系。当经济发展水平较低时，污染程度相对较低，随着增长的进展，污染程度也会增加。但是，当经济发展到一定阶段后，也就是说，达到拐点，收入的进一步增加将逐步改善环境质量。考虑到中国城市空气污染物存在 EKC 关系，我们选择人均 GDP 的二次方程作为解释变量。这个模型表达为：

$$\ln(Pollution_i) = \alpha + \beta_1 \ln(PCGDP) + \beta_2 [\ln(PCGDP)]^2 + \varepsilon$$

其中， $Pollution_i$ 表示 SO₂、NO₂ 和 PM₁₀ 的年平均浓度。PCGDP 表示人均 GDP。当 $\beta_1 > 0$ 和 $\beta_2 < 0$ 时，我们可以确定空气污染物似乎具有 EKC 的特征。

3.2 贝叶斯模型平均（BMA）方法

模型不确定性是计量经济模型化过程中的常见现象。这个问题主要来自两个方面：变量选择和变量测量。根据经济学理论。我们通常可以确定模型包含特定变量，但大多数变量都是任意包含或排除的（Magnus 等，2010）。其次，对同一变量使用不同的度量通常会导致不同甚至相互矛盾的结论。以不同的空气污染物为例。空气污染物的影响因素很复杂，研究人员往往无法确定这个模型需要引入哪些变量，特别是本文考虑的三种不同的空气污染物。因此，在以前的文献中，不同学者在构建计量经济模型时经常使用不同的解释变量。BMA 方法可以有效率地解决上述问题。

以下由 Magnus 等（2010 年）提出的，考虑线性回归模型：

$$Y = \alpha + \beta X + \gamma Z + \varepsilon, \varepsilon \sim N(0, \delta^2)$$

其中 Y ($n \times 1$) 表示因变量， X ($n \times k_1$) 和 Z ($n \times k_2$) 表示解释变量， ε 是误差向量， α 是截距， β 和 γ 是未知参数向量。我们假设 $k_1 \geq 1, k_2 \geq 1$ ，和列矩 (X, Z) 是满秩列。进一步假设误差向量 $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$ 是独立同分布 $\text{IID} \sim N(0, \delta^2)$ 。

在上述线性回归模型中， X 描述了所有模型的解释变量，有时称为核心变量，因为 X 变量包含根据经济理论得出的影响 Y 变量的重要因素。对于不同的线性模型， Z 变量表示不同模型中的唯一解释变量，称为辅助变量。通常 Z 中的解释变量的类型和数量不相同。我们必须区分 X 变量和 Z 变量，因为我们确信模型必须包含 X 变量。但是，我们无法确定哪些辅助变量需要被包含到模型中。当模型中的辅助变量不同时， X 的回归系数不相同。因此， Z 变量的选择变得至关重要。

矩阵 Z 中有 k_2 个解释变量，替代线性模型达到 2^{k_2} 种。假设 $M = \{M_1, M_2, \dots, M_K\}$ 是包含所有模型的模型空间，其中 $K = 2^{k_2}$ 。任意模型中 $M_i \in M$ ：

$$Y = \alpha + \beta X + \gamma_i Z_i + \varepsilon$$

其中， Z_i 是 Z 的一个子集， γ_i 是 γ 的一个子集。

我们假设单个模型时正态线性模型，那么模型 M_i 的可能性由下式给出：

$$p(y | \alpha, \beta, \gamma_i, \sigma^2, M_i) \propto (\sigma^2)^{-n/2} \left\{ \exp \left[-\frac{1}{2\sigma^2} (Y - \alpha - \beta X - \gamma_i Z_i) \right] \right\}$$

在实施 BMA 方法之前，需要设置参数和模型的先验分布。第一个是参数的先验分布。根据 O'Hagan(1994)和 Magnus (2010) 等人的研究，系数分布假设如下：

$$p(\sigma^2 | M_i) \propto \sigma^{-2}, p(\beta | \sigma^2, M_i) \propto 1$$

$$\gamma_i | \beta, \sigma^2, M \sim N(0, \sigma^2 V_i)$$

根据 Zellner (1986) 的理论， V_i 是一个 G-先验，定义为正定矩阵 $V_i = g_i Z_i' W Z_i$ ，其中， $W = I_n - X(X'X)^{-1}X'$ 。

对于 g_i 的分布，根据 Fernandez (2001) 等人提出的信息标准， g_i 的分布选择为 $g_i = 1/\max(n, k_2^2)$ 。

其次，对于单个模型 M_i 的先验分布，在缺乏先验信息的情况下，我们假设它遵循均匀分布 $p(M_i) = 2^{-k_2}$ 。有关更多有关于 MDA 方法的详细信息，请参阅 Magnus (1997 年) 等人的研究。

BMA 方法以后验包含概率 (Posterior Inclusion Probability) 作为权重，并且同时使用后验包含概率作为解释变量选择的标准，这个方法可以有效地提高模型的解释能力。Raftery (1997) 等人指出，对于选择“真实”模型，MBA 线性模型优于其他传统的逐步回归模型。与线性模型相比，MAB 方法具有通过考虑模型的不确定性来处理变量选择问题的优点。Madigan 和 Raftery (1994) 指出，在对数评分标准下，BMA 模型的加权平均预测结果优于所有单一模型的结果。通过使用 BMA 方法，可以预先平均所有可能的单个模型，并且通过设置单个模型的先验概率来计算解释变量的 PIP (后验包含概率)。解释变量的重要性按后验包含概率排序。它可以避免人工选择解释变量引起的信息偏差。

在本文中，有三个因变量：SO2 年平均浓度，NO2 年平均浓度和 PM10 年平均浓度。核心变量是人均 GDP，人均 GDP 的平方和城市化率的平方，以及总共十个的辅助变量，这些将在第四章中讲述。

4. 数据描述

在 2012 年，中华人民共和国国务院颁布了环境空气质量新标准，并逐步在各个城市实施。2013 年和 2014 年，只有一部分实施新标准的中国城市公布了其空气污染数据。本文使用 2012 年中国 282 个城市的年度数据。本文使用 2012 年中国 282 个城市的年度数据。原因是：首先，我们考虑数据的可用性和一致性；其次，来自于表 2 和表 3，大多数变量的标准差较高，表明城市之间存在较大差异。此外，相对于本文考虑的 12 个解释变量，282 个城市大样本数据以确保可靠的估算结果。因此，在分析影响变量对空气污染的影响方面，2012 年 282 个城市的数据是一个不错的选择。截面数据也广泛用于最近的环境分析中 (Galeotti 等, 2016)。基于环境经济学的相关研究，我们构建了一个包含 12 个解释变量的计量经济模型。表格 1 描述了数据源。一些缺失的数据通过插值补充。

4.1 空气污染变量

空气污染数据包括 SO2 年平均浓度，NO2 年平均浓度和 PM10 年平均浓度。根据中国环境保护局的统计，颗粒物是中国城市的主要污染物，占 2012 年总天数的 90% 以上。空气

污染物的主要来源是人类生产过程的排放，各类机动车尾气排放（包含 SO₂，NO₂ 和其他污染物），和工业粉尘的排放（工业 SO₂ 和工业 NO₂）。表 2 描述了中国 282 个城市三种空气污染物的统计分析。

表格 1

变量	数据源
年平均 SO ₂ 浓度 (SO ₂)	2012 年中国环境质量
年平均 NO ₂ 浓度 (NO ₂)	
年平均 PM ₁₀ 浓度 (PM ₁₀)	
人口密度 (PD)	CEIC 中国数据库
建成区绿色覆盖率 (GCR)	2013 年中国城市统计年鉴
外国直接投资 (FDI)	
第二产业比例 (SI)	
第一产业比例 (PI)	
民用拥有汽车 (CMV)	
固定资产投资 (FAI)	
行政区域总面积 (UA)	
出口依赖 (ED)	
人均 GDP (PCGDP)	
城市化率 (Urb)	
温度 (T)	2013 年中国各省统计年鉴
	中国省市年度统计公报

数据来源：aCEIE 经济数据公司是亚洲经济的权威提供者研究资料 (<https://insights.ceicdata.com/>)。

4.2 解释变量

影响空气质量的变量非常多。参考以往的文献和环境经济学理论，本文选择了以下十二个解释变量。

4.2.1 核心变量

- (1) 人均 GDP (PCGDP)。关于环境污染与收入水平之间关系的最有影响力的研究是 EKC 理论。EKC 理论广泛应用于现有文献中的环境分析，如空气污染物，水污染物和 CO₂ 排放的研究。常用的表达式是二次型，因此本文以 PCGDP 和 PCGDP 的平方为核心变量，探讨环境质量与收入水平之间的关系。
- (2) 城市化率 (Urb)。本文将城市人口占总人口的比例作为城市化率。城市化进程中最明星的变化是人口的聚集和消费模式的变化。这样人口众多的国家，城市化与环境污染之间存在着密切的关系。因此，本文选择城市化率作为研究城市化对空气质量影响的核心变量。

4.2.2 辅助变量

- (1) 经济发展变量。经济发展变量包括外国直接投资 (FDI)，固定资产投资 (FAI) 和出口依赖 (ED)。在本文中，我们使用外国直接投资所代表的外国资本的实际利用率。关于外国直接投资对环境质量的影响有两种不同的看法。一些学者认为，外国直接投资可以通过技术转让为东道国带来技术进步。他们还认为，外国直接投资可以通过技术分散带来经验和技术转让，改善环境晚污染，并最终改善东道国的环境质量 (Eastin 和 Zeng, 2009 年)。另一部分学者认为，外国直接投资从环境质量标准高的国家转移到标准较低的国家，将导致东道国环境质量恶化的趋势 (Chichilnisky, 1994)。

对于固定资产投资，政府面临经济增长和环境污染之间的双重压力，并在固定资产投资

和环境污染投资之间做出决策。当政府追求快速的经济增长和固定资产投资时，这将不可避免地导致对环境污染的投资，而大规模的固定资产投资可能会削弱环境保护。然而，环境污染的压力反而迫使政府加大对环境保护的投入。

出口依赖性代表总出口与总 GDP 的比率。它被认为是调查外贸对环境质的影响。关于外贸对空气质的影响尚无共识。一些人认为更自由的贸易可以改善环境质量（Antweiler 和 Taylor, 1998）。然而，污染假说理论认为，环境污染密集型产业将通过对外贸易转移到环境质里标准较低的国家。这将加剧这些国家的污染。

(2) 产业结构变量。本文采用的产业结构变量包括第一产业比例(PI)和第二产业比例(SI)。

中国已进入中后期工业化阶段：第二产业比里高，产业结构也不同。在中国 282 个城市中，2012 年第二产业比例最低，占 17.1%，最高占 87.96%。密集的工业生产过程导致污染物排放。产业结构的假设变化表明，人类社会将从以农业为基础的低污染社会转变为以工业为导向的高污染社会，最终转变为服务型低污染社会。

(3) 城市变量。城市变量包括城市人口密度 (PD)，建成区绿色覆盖率 (GCR)，年平均温度 (Tem)。行政区域 (UA)土地总面积和城市民用机动车辆 (CMV)。随着中国城市化水平的不断提高，城市常住人口和流动人口将以较高的速度增长，导致城市人口压力严重。2015 年，中国有 13 个城市，人口超过 1000 万。城市人口的增加对城市空气质有很大影响。首先，城市人口的增加带来了交通和运输的迫切需求。其次，它增加了对能源消耗，商品消费和公共基础设施的需求。

城市的年平均气温反映了城市气候条件。中国南部和北部地区的气温差异很大。中国北方冬季非常寒冷；因此，它需要燃烧大量的煤来加热，导致更多的废气排放到环境中。相比之下，南方气温相对较高，特别是中国东南部沿海地区，年均气温较高，环境污染程度较低。因此，我们将温度视为空气质量的重要因素，并将其作为辅助变量

表 3 描述了解释变量的汇总统计。各变量之间的相关系数见附录 B。正如预期的那样，城市化率与人均 GDP 呈正相关。

指导教师评语：所选文献与毕业论文内容相关性较高，该译文能准确表达作者原意及还原技术细节，字数在 3000 字以上。英语水平过关，译文文字通顺，表意清晰，符合翻译要求。