



吴裕彬

大数据如何改变经济咨询服务行业？

◎ 独立经济学家 吴裕彬 / 文

未来，经济学家将更多地使用大数据挖掘技术，经济学数据分析被时间序列垄断的局面将被打破，预测类算法将取代变量之间的因果分析，机器学习将取代统计模型。>>

大数据已经彻底改变了国家经济的方方面面，但其在经济学领域却没有太大反响，这是一件非常奇怪的事。经过多年经济学和大数据应用前景的研究，笔者深感到大数据技术可以对经济政策分析和经济学研究产生非常深远的影响。

大数据在经济学数据使用方面的潜力

首先，大数据将彻底改变经济学使用数据的实时性。大数据收集的都是实时数据，现在很多企业都在利用实时数据，奇怪的是经济学研究却主要使用汇总数据，很少使用实时数据。汇总数据一般最精准的就是当天的数据，比如汇率，而像通胀数据则是当月的。

其次，大数据将使经济学可调用的数据集规模极度扩大。经济学研究调用的一般是成千上万个时间点（一般最小的时间点是天）的数据，相对于大数据而言，这样的研究样本是很小的，所以经济学研究对研究者的统计学功力依赖很大。然而大数据可以调用数百万、数千万甚至是更多的数据，因此研究者的统计学功力就显得没那么重要。

第三，大数据可以同时观察多个变量之间的互动。经济学常用的数据分析方法是时间序列分析，一般只能研究两个变量之间的互动，比如狭义货币是如何影响通胀的。

第四，大数据结构更少，但层次更多。比如在零售领域，传统经济学的数据收集呈矩形，有 N 个观察时间点和 K 个变量，一般 K 远远小于 N 。而大数据记载的只是一系列消费行为，其数据并非矩形，也没有更复杂的结构，你可以用统计学方法把这些数据构造出无限多个矩形数据集。

最后，大数据将彻底改变经济学对数据记录之间关系的看法。传统经济学认为每一个记录的数据都是独立的，或者可以集结成面板数据，归根结底也就是时间序列的衍生物。但大数据却非如此，比如社交网络上人与人之间的互动数据是高度复杂的，传统的经济学模型无法揭示社交网络上人与人之间的互动关系，必须使用大数据的方法。

大数据在经济政策分析和经济学研究领域的应用场景

企业运用大数据的场景主要是记录运营过

程和结果，并构建涵盖范围广泛的预测类算法。比如 Amazon 和 Netflix 应用预测模型为客户推荐影视剧和书籍。预算类算法的可使用范围远远超越了电子商务。比如在医疗保险领域，保险公司通过将病人的付费情况和治疗效果数据导入预算类算法，可以计算其“风险系数”，然后通过风险系数来调整保费标准，而在大数据技术之前，“风险系数”是由病人的病史和相应的统计学分析方法来确定的。其实，大数据若应用到政府领域，也可以极大地改变经济政策分析和经济学研究。

首先，大数据会极大地释放出政府公共事务管理数据的潜在能力。通过管理税收系统、社保系统以及法规条例，政府积累了海量的颗粒状数据。公共管理数据在很大程度上没有被充分利用，主要原因是政府有关部门缺乏大数据硬件、软件和人才基础，另外这些数据也没有通过开放的端口给私营数据供应商使用，而且各地方政府的数据收集标准不一，难以统一维护和管理。在这方面，许多欧洲国家走在世界前列，其中央政府将各级地方政府的教育、医保等数据整合成了针对全国人口的大数据库。

公共管理数据的潜力非常巨大，这些数据涵盖不同个人、企业和机构相当长期的各类行为和状况，一般是面板结构的，数据质量也很高。而且由于这些数据集的涵盖面是普遍的，其可以和其它涵盖面更具选择性的数据集搭配使用。

如果政府向私营数据供应商有限开放这些公共管理数据的应用端口，对经济政策分析和经济学研究无疑将产生巨大推动。比如经济学家 Thomas Piketty 和 Emmanuel Saez 利用美国国

未来，政府公共事务管理数据的应用端口将更趋开放，其数据衍生产品前景远大，许多重要的经济指标将实时发布，公共部门的议题探讨和决策过程将更多地获取大数据的支持。>>

税局的数据构建了美国最富家庭占全国收入比例的历史数据集。他们的相关研究成果对奥巴马以来的决策者产生了极大的影响，美国最富家庭所占全国收入比例和所占全国纳税总额比例的严重不对称以及日益失衡成为了决策者和立法者探讨税收政策改革的一个话题焦点。

其次，大数据可以改变政府测算经济活动的方式。政府在经济活动的测算方面扮演着极其重要的角色，比如通货膨胀率、失业率和 GDP 等等的测算都是由政府主导的。一般而言，政府都是通过调研的方式来测算经济活动的。比如国家统计局会派出调研员去商店手动收集成千上万商品的价格，然后将这些数据汇总成不同的通胀指数——CPI 就是其中之一。然而大数据技术可以更大规模地收集物价数据，甚至可以做到实时收集。比如，由麻省理工学院斯隆商学院教授 Alberto Cavallo 和 Roberto Rigobon 发起的 10 亿物价项目（BPP），通过成百上千个电商网站和手机应用的端口可以收集数以 10 万计商品的实时价格数据，从而可以实时发布通胀数据，而不是像国家统计局那样每个月发布一次。

再次，大数据在企业部门的一个巨大作用就是使企业的议题探讨和决策过程更大程度上得到大数据分析的支持，并且有些公司还在用大数据进

行广泛的运营实验，相比之下，公共部门在这些方面进展十分迟缓。

政府收集了大量的公共事务管理数据，这些数据对经济政策分析和经济学研究助益极大，但利用效率却极低。比如，医保部门有过去几十年的每一笔医保索赔记录，稍加整理就可以得出无数个人的病史数据集，通过大数据挖掘和预测类算法，可以得出关于不同治疗方案和过程的非常详细的成本效益分析，从而使医保支出的效益大大提高。以此类推，政府部门完全可以用大数据技术对医保和税收等涉及财政收支的政策法规改革调整进行一定范围的精准实验，再根据效果去调整相关政策法规。

经济咨询与大数据结合前景远大

未来，经济学家将更多地使用上面所描述的大数据挖掘技术，经济学数据分析被时间序列垄断的局面也将被打破，预测类算法将取代变量之间的因果分析，机器学习将取代统计模型。

政府公共事务管理数据的应用端口将更趋开放，其数据衍生产品前景远大，许多重要的经济指标将可以做到实时发布，公共部门的议题探讨和决策过程将更多地获取大数据的支持。根据 Statista 的数据，经济咨询服务（经济咨询公司）在美国的市场规模已超过 300 亿美元（近 2000 亿元人民币），这并不包括银行、投行和公共机构内部的经济分析部门，一旦包括起来，其市场规模将更大。经济咨询服务未来在中国也应该成为一个千亿元级别的产业，大数据技术将使该行业发生根本性的技术革命，前景不可估量。■

