

大数据时代计量经济学的新发展与新应用^{*}

——第二届中国计量经济学者论坛(2018)综述

胡毅 陈海强 齐鹰飞

第二届中国计量经济学者论坛于2018年12月14—16日在中国科学院数学与系统科学研究院召开。本届论坛由东北财经大学、《经济研究》编辑部、厦门大学王亚南经济研究院与邹至庄经济研究中心、中国科学院预测科学中心联合主办,由中国科学院大学经济与管理学院承办。论坛旨在促进中国计量经济学理论研究与方法应用的发展。本届论坛邀请了美国麻省理工学院(MIT) Whitney Newey 教授、美国南加州大学萧政教授、《经济研究》编辑部刘霞辉研究员、美国堪萨斯大学蔡宗武教授、美国康奈尔大学洪永淼教授、中国社会科学院汪同三研究员、中国科学院汪寿阳研究员等海内外著名专家学者莅会并发表主题演讲。

Newey 教授研究了大数据在需求分析中的应用。传统研究分析商品价格变化对消费者剩余的影响仅能控制极少价格,往往导致遗漏变量偏误,而逐笔交易数据能解决遗漏变量问题,但因此会带来来自变量过多的问题。机器学习可实现降维并估计出平均需求与消费者剩余边界。此外,他提出利用控制方程处理总支出的内生性,从而得到一组消费者剩余的正交矩函数,最终估计出消费者剩余。Newey 教授将该方法应用于 Nielsen 交易数据来研究消费需求函数中的交叉价格效应。

萧政教授从计量经济学家视角审视了大数据分析。萧政教授指出,计量经济学通过模型来分析经济现象的关键特征,模型是现实世界的简化,建模时将不关键因素设定为误差项。与之相反,大数据分析使用“显微镜”寻找真实世界的“镜像”。如何调和两类方法,既充满了挑战,也给计量经济学发展带来了新契机。萧政教授认为,计量经济建模与大数据分析都能促进对方发展。最后,他总结提出大数据应用中有待研究的十大统计分析问题。

刘霞辉研究员总结了中国经济发展新阶段的特征。刘霞辉研究员认为,当前中国经济呈现五大特征:一是经济结构由产品生产经济转化为服务型经济;二是人力资本、技术创新、高新技术成为经济的主导力量;三是交通和通信设施逐步完善不断改变社会的生产组织行为和消费者的消费行为;四是以人的全面发展为中心的经济格局正在形成;五是网络型经济和社会形态逐步形成。刘霞辉研究员指出,这些新特征给经济学研究提出了新命题。

蔡宗武教授探究政策评估中非混淆(unconfounderness)假定的检验问题。鲁宾因果框架广泛应用于政策评估,其关键假定是非混淆条件,当假定满足时,一般使用回归方法、逆概率加权(IPW)、双重稳健(DR)、匹配(Matching)或分位数处理效应(QTE)方法,假定不满足时主要使用双重差分(DID)、工具变量(IV)、赫克曼选择(Heckit)、局部平均处理效应(LATE)等方法。对非混淆条件假定的检验,直接影响到政策评估方法的选择。他提出,可利用辅助变量,采用非参数方法对特定矩条件进行检验。

^{*} 胡毅,中国科学院大学经济与管理学院,邮政编码:100190,电子信箱:huyi@ucas.ac.cn;陈海强(通讯作者),厦门大学经济学院、王亚南经济研究院、教育部计量经济学重点实验室(厦门大学),邮政编码:361005,电子信箱:hc335@xmu.edu.cn;齐鹰飞,东北财经大学经济学院,邮政编码:116025,电子信箱:qiyingfei@sina.com。作者感谢方颖、洪永淼、刘霞辉、汪寿阳、王维国、张永山等论坛学术委员会和组委会为论坛做出的辛勤工作,以及为本文提供材料的博士后及研究生文强、万千、倪博、田露、闫玉等。

洪永森教授介绍了时间序列模型平均问题。他从“卢卡斯批判”等案例出发,指出时间序列模型存在结构不稳定、结构转换以及模型不确定性等问题。模型平均能够降低模型不确定性。模型平均的权重可能随时间变化,因而需要发展时变模型平均方法。他提出基于局部交叉验证准则最小化的时变刀切法模型平均(TJMA)。蒙特卡洛模拟以及实证研究显示,该方法优于其他方法,特别是在非线性时间序列数据中表现更好。

汪同三研究员阐述了经济计量分析如何研究全面深化开放背景下的经济问题。他总结了美国发动贸易战的动机,认为中美发展差距快速缩小是重要因素。贸易摩擦对实体经济的影响微弱,心理影响大于实际影响。对比双方有利条件,他认为中国可采取“不对抗”、“按步伐开放”与“国家核心利益不退让”原则予以应对,而且“按步伐开放”要引入“竞争中性”原则。最后,他提出了经济计量分析评估政策应注重的三大问题。

汪寿阳研究员关注了大数据时代的经济预测问题。他指出,非结构化数据给经济预测带来挑战,而经济预测至关重要,提出具有广泛适用性与较高预测性能的预测方法非常急迫。他提出,可利用分解方法提取非结构化数据的关键特征,然后利用计量建模与深度学习方法对其进行预测,最后利用综合集成方法将所有预测集成。他以其团队关于原油价格预测等工作为例,展示上述方法的优越性。他认为,上述方法应用价值巨大,但是如何上升到理论层面、如何识别因果效应值得探究。

除主题演讲外,本届论坛共收到百余位计量经济学者的踊跃投稿,最终选取了来自北京大学、东北财经大学、对外经贸大学、华南理工大学、美国路易斯安那州立大学、南开大学、上海财经大学、上海交通大学、西北大学、西南财经大学、厦门大学、中国科学院、中国人民大学、中南财经政法大学、中山大学、中央财经大学、中银国际证券等海内外高校、研究机构的50位计量经济学者参会并报告研究成果。下面,围绕计量经济学理论、宏观计量经济学、微观计量经济学和金融计量经济学四个领域对本届论坛的研讨成果进行介绍。

一、计量经济学理论

针对非线性动态面板模型的估计与检验方面,周乾坤基于未知函数系数部分(unknown slope functions)的筛分近似(sieve approximation)提出了筛分二阶段估计(Sieve 2SLS),并研究了估计量的渐近性质,同时提出非参数的斜率恒定性参数检验。王满关注了面板数据模型中方差变异的诊断问题,并基于移动总和(MOSUM)统计量提出一步估计法,该方法可以同时识别出变异的次数与具体发生时间,还可以克服基于累计总和技术的检验所存在的效力损失问题。李思麒基于五组蒙特卡罗仿真实验,通过直接比较潜在结果与反事实结果之间的分布,来比较合成控制法(SCM)、面板数据模型(PDA)以及广义合成控制法(GSCM)三种方法进行政策评论的有效性。

在模型平均方法方面,王苗苗关注高维分位数回归的变量选择问题,提出两步估计法,第一步基于边际分位数对所有解释变量排序,第二步是采用留一法交叉验证(delete-one cross-validation method)对第一步选出来的变量进行模型平均,最终选出最优模型。作者还放松了模型权重之和等于1的传统假定,并基于经验过程理论证明了渐近性质。孙玉莹探究区间型数值时间序列(ITS)的变量选择与滞后阶数确定问题,以Mallows型准则选择ITS中间点,并提出一种新的、以权重准则选择ITS广度的方法,同时证明了相对应的中间点与广度的模型平均估计量的渐近最优性质。

因子模型方面,孙宇澄使用伴随微观结构噪音的异步高频金融数据,研究了连续时间下具有稀疏特征协方差矩阵的高维因子模型,提出了一种共同因子数量、即积分的协方差矩阵及其逆矩阵的一致估计量,该估计量基于平定实核估计(flat-top realized kernels)方法,具备对序列依赖噪声稳健性质和收敛率最优性质。韩猛提出并开发了一个平滑转换因子模型,该模型旨在通过允许加载平

滑转换因子来检测和模拟因子模型中的状态切换行为,推导出平滑转换因子模型的因子和参数的两阶段估计及其渐近性质,并提供了规范的检验过程。王霞基于离散傅里叶变换提出了识别因子模型是否存在结构突变的检验,该检验适用于存在任意形式的序列相关和横截面相关,对于各种类型的结构变换和单个或多个结构变换点均是有效的。

金融计量方面,田丁石构建了广义自回归期望分位数(GARE)模型,该模型可以包含用于测度尾部风险的期望分位数滞后项,通过构造一个“线性指数”新密度族,证明任何一致QML估计量(QMLE)的拟似然函数必定是“蜚线性指数”(tick-linear-exponential)族之一。谢海滨利用随机混合分布(MIDIS)来刻画资产价格的动态变化,该方法使用两个正向和负向独立冲击的随机组合对资产价格变化进行建模,发现负向冲击分布决定了金融风险的尾部特征。苏栝芳针对市场中可能存在的未知个数且未知发生时间的间断性结构性改变,首次提出了建立具有多个间断点的非参数局部常量估计,并通过一个序贯检验来确定间断点个数以及推导统计量的极限分布。杨利雄发现当时间序列分数频率傅里叶形式存在结构性变化时,dickey-fuller单位根检验往往存在伪拒绝的问题,从而提出傅里叶分数频率的区间估计,并证明了该条件区间的渐近性质。

二、宏观计量经济学及其应用

在经济增长、产业发展方面,陶长琪将产业结构低级形态地区的创新要素集聚、制度质量和创新要素集聚引致的技术溢出变量纳入内生增长模型,基于30个省市数据,使用空间门槛面板模型,发现不同严格程度的制度质量对创新要素集聚产生不同的影响效应,使创新要素出现不同的集聚形式从而影响产业结构高端化。师博基于1992—2016年省级面板数据探讨创新对高质量发展的影响,发现通过市场潜能传导机制,专利存量与高质量发展之间存在显著的U型关系,发明专利助力高质量发展的效应最为显著。欧阳志刚研究了省级工业行业的非充分发展与路径选择,发现各地区工业行业生产要素使用不充分并呈现出非充分发展状态,影响各行业发展的重要因素不尽相同,各地区应根据自身产业结构因地制宜制定优先发展行业和发展路径。

在经济波动方面,赵宣凯使用代理变量结构向量自回归(proxy SVAR)模型,利用原油价格高频数据,识别“黑天鹅事件”对中国宏观经济的影响,发现资本市场能较快较明显地反映“黑天鹅事件”带来的冲击。黄乃静探讨美国经济政策不确定性影响中国宏观经济活动的传导机制和溢出效应,使用结构向量自回归模型,发现美国经济政策不确定性升高将引起中国股指回报率下降、外商直接投资增加以及对美出口减少,与美国金融市场联系较密切、对外国投资者较具吸引力以及出口业务对美国市场依赖程度较高的行业,对美国经济政策不确定性冲击的反应更明显。

在财政、税收、货币政策方面,郭俊聪剖析地方竞争的动机,通过建立具有时变特征的固定效应空间向量自回归模型,刻画了中国2001—2016年县级政府财政支出的动态变化,研究发现2009—2014年区域竞争程度降低源于“四万亿”财政扩张政策,而2014—2016年竞争程度上升源于经济下行压力。刘定通过构建包含货币政策与财政政策的一般均衡模型,结合1993—2017年的数据,估计和分析了中国财政政策和货币政策的类型,发现政策类型属于被动型货币政策和积极型财政政策的组合,这样的政策组合容易引发通胀、带来债务问题。胡华研究了个人所得税平均税率变动的收入调节功能,发现收入调节功能表征指数与平均税率同向变动,但免征额、边际税率等要素决定着该指数与平均税率大小,而2018年的个税改革将极大削弱个税的收入调节功能。

在贸易和汇率方面,李鑫茹基于世界投入产出表,通过基于国民收入视角的贸易收益核算方法,重新估算了2014年中美、中欧双边贸易差额,发现在国民收入口径下,中美贸易差额约为1560亿美元,比贸易总值和增加值口径分别缩小34%和17%,中欧贸易差额约为424亿美元,比贸易总值和增加值口径分别缩小52%和34%。程欣基于2005—2017年的汇率日度数据,使用网络分析

方法构造了“一带一路”沿线主要货币汇率相互影响的动态网络,刻画人民币对沿线国家(地区)货币的影响力,发现这种影响力在提出“一带一路”倡议前后有明显上升。李欣珏通过多元自适应可变窗算法构建了具有优良预测能力的人民币汇率预测模型,以及时监测模型参数的时变特征,探测最大化同质区间。该方法仅给出了一种如何选择最佳窗宽(观测信息)的方法,也显著地提高了参数模型对人民币汇率预测的能力。

在环境规制方面,谌仁俊拓展了减排目标政策下行为分析模型,探讨总量减排目标政策对工业增长的潜在影响,并由此探寻保增长下的最优减排目标。周亚虹和忻恺评估了《京津冀协同发展规划纲要》对京津冀地区环境治理的影响,发现该政策在保持经济增长的前提下,在环境治理上呈现出显著的政策效应:在政策实施后第二年,工业二氧化硫的产生和排放量分别减少了25%和15%。张俊荣在传统非竞争型投入产出模型基础上建立区分企业规模的投入产出模型,发现大型企业在产出、增加值以及出口方面贡献了最主要的力量,在我国经济结构中占据主要地位,在增加值率和单位出口增加值上,中型企业最高,而小型企业的单位出口隐含碳排放最高。

三、微观计量经济学及其应用

在家庭金融方面,尹志超研究土地流入带来的土地经营规模扩大与家庭信贷可得性之间的关系,发现土地流入对家庭生产经营正规借贷、生产经营民间借入均有显著正向影响。土地流入增加了农业土地经营规模,通过提高生产效率、扩大生产要素投入进而提升农户信贷可得性。李程探讨房地产泡沫与居民杠杆之间的关系,在信贷模型中引入杠杆、财富因素,发现居民杠杆选择取决于其对财富和风险的权衡,房地产泡沫与居民杠杆率呈现倒“U”型关系,通过测算发现居民杠杆率的阈值为90.942%。

在劳动供给与住房需求方面,贺建风探讨中国居民创业选择的影响因素,利用贝叶斯模型平均(BMA)方法发现,个人层面的婚姻状况和风险偏好程度、家庭层面的房产状况和政治关系、社区层面的民族文化以及区域层面的FDI和行业管制是解释创业选择的核心因素。另外,社会保障水平和经济增速等因素对居民创业行为亦有较大影响。李秀婷利用空间计量模型研究了空气质量对中国房地产价格的影响,发现房价和PM2.5浓度在全国的分布存在空间聚集效应,PM2.5浓度会对房价产生负向作用。2013年以后,空气质量对房价的影响不断增加,并且经济越发达的城市空气质量对房价的影响越强。

在企业经营行为方面,朱平芳和纪园园评估了“加计扣除”政策对非国有高技术产业创新的影响,发现“加计扣除”政策确实能刺激非国有高技术产业研发经费投入、专利产出的规模及其相对增长率的提升,有助于增强企业自主创新的能力。戴美虹在多产品异质性企业理论框架下,揭示了互联网触发“创造性破坏”并影响出口企业创新的微观作用机理,探讨了互联网对于出口企业创新绩效的微观经济效应。李兵基于中国省级和市级报纸数据构造了城市媒体联系指数,从需求方和供给方识别媒体报道偏差来源,研究发现媒体报道受到省委书记和流动人口偏好的影响,城市间正向媒体报道市场偏差远高于负面媒体报道市场偏差,在市场化程度较高的地区,媒体报道更加倾向于读者偏好。

四、金融计量经济学及其应用

在宏观金融方面,陈良源基于结构向量自回归(SVAR)模型构造了货币政策冲击序列,借助随机波动率模型构建货币政策不确定性指数(MPU),进而构造了货币政策不确定性因子,实证发现该因子是中国股市一个重要的资产定价因子。徐承龙将随机利率以及与宏观经济风险相关的信用风险结构引入含有政府担保的可违约债券定价模型,研究发现中国债券市场中确实存在隐性担保问题,且不同类型的隐性担保概率差异较大。尚玉皇提出带区制转移的混频 Nelson-Siegel 利

率期限结构模型,该模型可以精确捕捉到期限结构因子的变化趋势,发现考虑区制转移特征后,利率期限结构的斜率因子对经济周期和通货膨胀的预测作用增强。

在资产定价方面,李明阳基于股利贴现框架模型并结合适应性预期,构造了动态结构性面板模型以考察经济基本面因素对上交所上市股票的定价能力。实证结果显示,基本面因素对股票价格存在解释力,但远不及非经济基本面因素。石磊磊利用相互一致性偏好来解释交易量动态均衡形成的机制,并使用价量概率波二阶微分方程,提供了一种用价-量分布来计量价格波动的不确定性和个人交易偏好的新方法。张传海基于2014年6月—2015年12月市场极端条件下上证50指数成分股的高频交易数据,发现投资者情绪对指令不平衡具有显著影响,并且指令不平衡对收益具有显著的预测能力,进而表明在极端市场条件下中国股市有效性较低。陈海强基于2013—2018年中国证监会公布的市場操纵案例及沪深两市股票交易高频数据,构造基于逻辑回归的操纵预警模型,并估计出个股市場操纵概率得分指标MM-score,发现MM-score能够很好识别出市場操作,基于MM-score的多空策略能够产生显著超额收益,同时发现MM-score是特征因子而非风险因子。

在金融风险与市場稳定性方面,徐秋华以上市A股为研究样本,采用混合copula函数模型对崩盘敏感性指标进行估计,并利用双重差分模型考察融资融券交易制度的市場效应,研究发现融资融券加剧了股票的崩盘敏感性,增加了个股伴随大盘暴涨和暴跌的概率。潘宁宁基于2004—2016年中国A股上市公司数据,发现机构投资者持股能降低股价崩盘敏感性,其中专业机构投资者会通过提升股票流动性、降低投资者负面情绪以降低股价崩盘敏感性,但是非专业机构投资者持股加剧了股价崩盘敏感性。龚玉婷将宏观层面低频解释变量纳入高频数据下的连接函数模型,构造了copula-MIDAS-X模型研究原油价格与美国股票市場价格的相依性,发现原油市場与美国股票市場共同下跌主要受原油总需求因素与股票市場特有冲击(如金融危机)影响。

在銀行与公司金融研究方面,焦文姝利用168家商業銀行2011—2017年的面板数据,考察2017年全面加强金融监管对銀行个体风险的影响,发现金融强监管显著降低了銀行的个体风险。周波利用2014—2016年A股上市公司数据,考察了高管政治联系对公司参与PPP项目的影响,发现公司高管存在中央级的政治联系时,企业参与PPP项目的概率越高,而省级政治联系或地市级政治联系对公司参与PPP项目不存在显著影响。李义超基于技术水平异质性视角,研究了研发投入及产品市場竞争对企业财务绩效的影响,发现当期和滞后一期的研发投入均与财务绩效显著负相关,而滞后三期的研发投入与财务绩效则呈显著正相关。吴武清使用参数和半参数模型分析了融资成本、融资规模和全要素生产率对杠杆调整度的影响,发现提高融资成本并不能明显提升去杠杆速度,肯定了限制融资规模对于去杠杆速度的高敏感性。

在互联网金融与基金投资方面,王博基于“人人贷”平台2010—2015年微观借贷数据,从金融发展理论和网络经济学双重视角发现传统金融通过扩展融资渠道、加大市場竞争、降低机会成本等路径,对互联网金融参与率存在替代效应。田正磊以中国2005—2017年的开放式基金数据为样本,通过基金重仓持股数据构建基金网络,发现网络中其他基金的业绩分布通过“互补效应”和“替代效应”对该基金的资金流产生交叉影响。

五、结 语

本届论坛聚集了国内外计量经济学理论和应用领域的资深专家和青年学者,为中国计量经济学领域的广大学者提供了互相学习、加深交流、共同提高的优质平台。论坛期间,与会学者积极展示最新研究成果,热烈探讨改进完善意见。这有助于推动中国计量经济学理论与应用水平的提高,为丰富和发展中国特色社会主义经济理论创造有利条件。

(责任编辑:陈小亮)(校对:张 涵)