

2017 年中国经济学奖获奖感言

陈晓红

耶鲁大学 Cowles 基金会

2017 年中国经济学奖颁奖仪式

2017 年 11 月 18 日，北京

感谢北京当代经济学基金会

- 我诚挚地感谢夏斌理事长和北京当代经济学基金会在 2016 年设立了中国经济学奖。
- 这项为了“鼓励理论创新，繁荣经济科学”的公益事业值得敬佩和赞赏。
- 我非常感谢今年评奖委员会对邹至庄教授和我在计量经济学领域研究工作的认可。

邹教授与我这一代华人经济学家

- 能与邹教授分享今年的中国经济学奖，我愧不敢当。
- 邹教授是所有中外经济学人共同尊敬的前辈和楷模。
- 邹教授在中国经济改革和经济学教育现代化方面也发挥了重大作用。
- 许多华人经济学家都是“邹氏检验派”；更多的经济学者（包括我）来自于邹教授从1985年开始创办的中美经济学培训班。

- 基金会给我这个殊荣是对我这一代华人计量经济学家和华人女性经济学家的肯定。
- 我希望此奖能够鼓励更多的年轻学者投身到计量经济学前沿的研究中去。

在教育经历方面，我很感谢

- 1976 年文化大革命结束；1977 年中国恢复高考；1978 年 12 月开始的改革开放。
- 教过我的所有中外老师，感谢他们对教学工作的热爱和无私奉献：
 - 1982 至 86 年，武汉大学数学系。
 - 1986 至 87 年，邹教授和黄达校长在人大创办的中美经济学培训班。
 - 1987 至 88 年，西安大略大学经济学硕士项目。
 - 1988 至 93 年，加州大学圣地亚哥分校经济学博士项目。

在学术科研方面，我很感谢

● 我的良师

- 在博士期间：Roger Gordon, Max Stinchcombe, the late Halbert White (我的博士导师);
- 在芝加哥大学任教期间：Lars Peter Hansen, Thomas Sargent (之后又在纽约大学共事);
- 在伦敦政经学院任教期间：Peter Robinson;
- 在普林斯顿大学期间（作为访问学者）：邹志庄, Bo Honore.

● 我的长期合作者和朋友：艾春荣, 范延琴, 洪翰, 李彤, 孙一啸, 肖志杰, Oliver Linton, Elie Tamer 等。

● 我的同事、合作者和朋友：萧政, 许成钢, Don Andrews, Richard Blundell, Javier Hidalgo, Sydney Ludvigson, Rosa Matzkin, Whitney Newey, Jim Powell, Ed Vytlačil 等。

● 我的学生和合作者：廖志鹏, 易艳萍, Tim Christensen, Demian Pouzo, Alex Torgovitsky 等。

● 我的父母，先生，姐姐、姐夫和哥哥、嫂子。

半非参模型和筛分方法

- 半非参模型是由有限维和无限维未知参数共同确定的概率分布族。
- 筛分方法通过在一系列渐近参数空间上（即筛空间）求解经验目标函数最优化问题来估计半非参模型。
- 两大类准则：所有用来估计非线性参数模型的现有准则（例如 Newey and McFadden, 94）都是合适的选择。
 - M-准则：最大似然估计 (ML), 准 ML, 分位数回归, 非线性最小二乘等；
 - 最小距离 (MD), 广义矩估计法 (GMM), 广义经验似然估计法 (GEL) 等。
- 两大类筛空间（例如 Chen, 07）
 - 有限维的线性或非线性筛：多项式, Hermite 多项式, 傅里叶级数, 样条函数, 小波, 人工神经网络 (ANN), 径向基, 脊波等；
 - 具有各种惩罚约束的无限维的线性或非线性筛。

筛分方法（续）

- 得益于可以选择不同的准则和筛空间，筛分方法在估计复杂的半非参的非高斯分布的不可分结构化模型时非常灵活；这类模型可能还具有内生性，潜在异质性或潜在状态相关性。
 - 准则的选择：取决于模型和研究者关注的参数。例如，没有内生性的非参数回归适用 M 准则；具有内生性的模型则适用 MD（或 GMM）准则；
 - 筛空间的选择：取决于我们所掌握的关于未知函数的先验信息，这些信息可能直接或间接地从经济或计量模型中得到。
- 筛分方法可以很容易地在未知函数上添加形状（单调性、凹性）、可加性、非负性和其他约束。
- 易于计算。一旦我们用有限维筛空间近似未知函数，则应用筛分方法就与参数非线性极值估计一样简单。

例 1a: 内生恩格尔曲线的形状不变系统

- Blundell et al. (03): 在某给定年份, 满足家户效用最优化的恩格尔曲线系统:

$$Y_{1\ell i} = h_{1\ell}(Y_{2i} - h_0(X_{1i})) + h_{2\ell}(X_{1i}) + \varepsilon_{\ell i}, \quad \ell = 1, \dots, N,$$

其中, $Y_{1\ell i}$ 是家户 i 在商品 ℓ 上的预算份额, Y_{2i} 是 i 在所有非耐用品上的支出的对数, 这里 Y_{2i} 是内生变量, 即 $E[\varepsilon_{\ell i} | Y_{2i}] \neq 0$ 。

- Blundell et al. (07): 半非参均值工具变量 (IV)

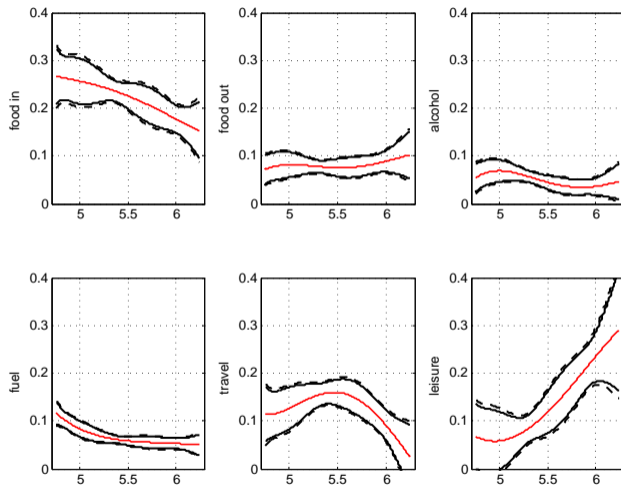
$$E[Y_{1\ell i} - \{h_{1\ell}(Y_{2i} - g(X'_{1i}\beta_1)) + X'_{1i}\beta_{2\ell}\} | X_{1i}, X_{2i}] = 0,$$

- Chen-Pouzo (09): 半非参分位数 IV

$$E[1(Y_{1\ell i} \leq h_{1\ell}(Y_{2i} - g(X'_{1i}\beta_1)) + X'_{1i}\beta_{2\ell}) | X_{1i}, X_{2i}] = \gamma \in (0, 1).$$

- 上述两种矩条件都使用筛最小距离 (MD) 方法估计。

图：估计恩格尔曲线



上图为恩格尔曲线的估计（红线），以及使用自举法得到的一致置信区间（UCB）（实黑线为 90%，虚黑线为 95%）。x 轴是家庭总支出的对数，y 轴是家庭预算份额。来源：Chen-Christensen（即将发表）

例 1b: 内生需求曲线的福利泛函

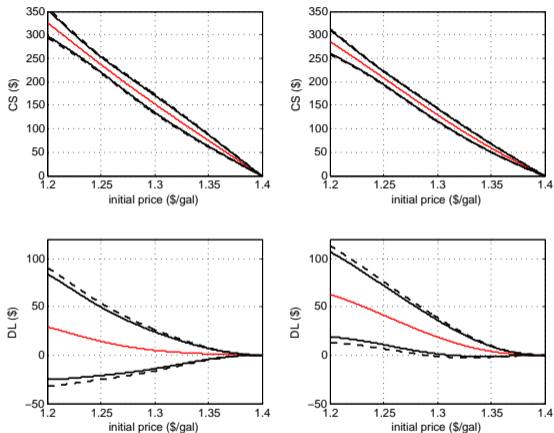
- 第 i 个家户对汽油的需求量: $Q_i = h_0(P_i, Y_i) + \epsilon_i$, 其中 P_i 代表价格, Y_i 代表 i 的收入。 P_i 是内生的。消费者剩余 (CS) 和无谓损失 (DL) 都是对价格 (即税) 变化如何影响消费者福利的度量。Hausman (81) 证明了当价格从 p^0 变为 p^1 时, 收入水平为 y 的消费者的 CS (用 $S_y(p^0)$ 表示) 满足:

$$\frac{\partial S_y(p(u))}{\partial u} = -h_0(p(u), y - S_y(p(u))) \frac{dp(u)}{du}, \quad S_y(p(1)) = 0,$$

$p: [0, 1] \rightarrow R$ 是一个连续可导函数, 满足 $p(0) = p^0$ 和 $p(1) = p^1$ 。

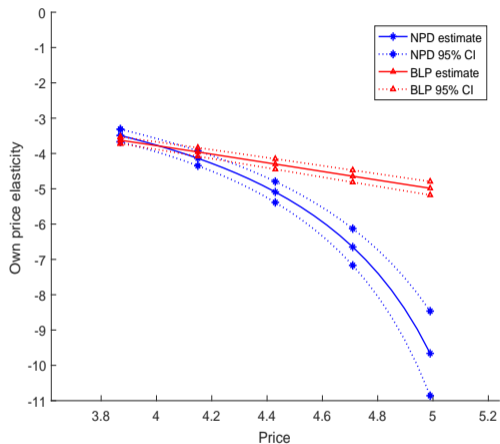
- Chen-Christensen (即将发表) 使用筛分非参数工具变量 (NPIV) 方法估计了内生汽油价格上涨对消费者 CS 和 DL 等福利影响; 文章使用了 Blundell et al. (12) 的数据。
- 产业组织中结构性需求: 今年耶鲁毕业生 Compiani 的求职论文 (Compiani, 17) 使用筛 NPIV 来估计内生需求, 目标是需求价格弹性。

图：汽油价格上涨导致的福利影响



使用筛 NPIV 方法估计（内生）汽油需求时，汽油价格升至 \$1.40/加仑所导致的消费者剩余 (CS) 和无谓损失 (DL) 变化 (红色实线) 以及通过自举法得到的 UCB (黑色实线表示 90%，黑色虚线代表 95%)。左侧代表家户收入为 \$72,500；右侧为 \$42,500。来源：Chen-Christensen (即将发表)

图：有机草莓的价格弹性函数



BLP 法对比筛 NPIV 法估计的自价格弹性函数。BLP 法依赖于参数假设，而筛 NPIV 使用 Chen-Christensen 中的结论进行统计推断。BLP 法低估了价格弹性。来源：Compiani (17, Yale JMP)

例 2a: 半参数资产定价模型

- 基于消费理论的资产定价模型:

$$E(M_{t+1}R_{j,t+1} - 1|\mathcal{I}_t) = 0, \quad j = 1, \dots, N,$$

$M_{t+1} = \frac{\partial U/\partial C_{t+1}}{\partial U/\partial C_t}$ 是消费的跨期边际替代率 (IMRS), 也被称作定价核或随机折现因子 (SDF)。

- Hansen-Singleton (82): $U = \sum_{t=0}^{\infty} \delta^t \left[(C_t^{1-\gamma} - 1)/(1 - \gamma) \right]$, $M_{t+1} = \delta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma}$ 。使用 GMM 估计**非条件矩约束模型**

$$E \left(\left[\delta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} R_{j,t+1} - 1 \right] \mathbf{z}_t \right) = 0, \quad j = 1, \dots, N,$$

- 很多研究对效用是消费的时间可分函数这一假设提出疑虑。

例 2b: 基于非线性习惯的资产定价模型

- Chen-Ludvigson (09): $U = \sum_{t=0}^{\infty} \delta^t \left[((C_t - H_t)^{1-\gamma} - 1) / (1 - \gamma) \right]$, 这里 $H_t = C_t g(c_t^*)$ 表示未知习惯的水平; $0 \leq g < 1$, g 是 $c_t^* = \left(\frac{C_{t-1}}{C_t}, \dots, \frac{C_{t-L}}{C_t} \right)$ 的第一个自变量的非减函数。 $M_{t+1} = \frac{\partial U / \partial C_{t+1}}{\partial U / \partial C_t}$ 。对外部习惯有 $\partial U / \partial C_t = C_t^{-\gamma} (1 - g(c_t^*))^{-\gamma}$; 对内部习惯有

$$\partial U / \partial C_t = C_t^{-\gamma} \left[(1 - g(c_t^*))^{-\gamma} - E_t \left\{ \sum_{j=0}^L \delta^j \left(\frac{C_{t+j}}{C_t} \right)^{-\gamma} (1 - g(c_{t+j}^*))^{-\gamma} \frac{\partial H_{t+j}}{\partial C_t} \right\} \right]$$

- 在外部和内部习惯模型下对福利的影响效应并不相同。研究者需要使用非线性的习惯函数 $g(\cdot)$ 来识别是外部还是内部习惯。

- Chen-Ludvigson (09): **筛最小距离 (SMD) 法估计条件矩约束模型:**

$$E(M_{t+1}R_{j,t+1} - 1 | \mathbf{w}_t) = 0, \quad j = 1, \dots, N, \quad \mathbf{w}_t \subset \mathcal{I}_t,$$

在实证研究中 $\mathbf{w}_t = [\widehat{cay}_t, RREL_t, SPEX_t, \frac{C_t}{C_{t-1}}]'$ 。

- 用 **ANN** 筛近似未知习惯函数 $g()$, 其中 $L = 3, 4$ (季度)。
- 我们使用季度数据估计得到的部分实证结果: (1) 习惯函数是非线性的; (2) 相比外部习惯, 使用内部习惯可以显著更好的拟合数据; (3) δ, γ 的估计值都符合直觉; (4) 基于习惯生成的 SDF 的估计在解释股票回报的横截面差异上表现很好; (5) 在定价错误方面有更多的发现, 并且根据 HJ 定价错误比较不同的模型。

我过去在筛 MD 估计上的研究小结

对半非参条件矩模型 $E[\rho(Y, \beta_0, h_0(\cdot))|X] = 0$, 其中未知函数 $h(\cdot)$ 可以取决于内生变量,

- Ai-Chen (03): 对条件矩模型提出使用筛 MD 方法估计 $(\beta, h(\cdot))$, \sqrt{n} 渐近正态性和有效估计, 以及 β 的筛 Wald 统计量。
- Ai-Chen (07), Ai-Chen (12): 为得到具有不同条件集的一般误判模型里的 β 的 \sqrt{n} 渐近正态性 (07), 或在顺序矩约束里为得到 β 的有效估计 (12) 分别调整了筛 MD。
- Blundell-Chen-Kristensen (07): 得到对内生恩格尔曲线系统 (例 1a) 的未知函数 $h(\cdot)$ 的筛 NPIV 估计的非参数均方误差收敛速度。
- Chen-Pouzo (09, 12, 15): 得到对可能非平滑的广义残差 $\rho(\cdot)$ (如分位数 IV) 的筛 MD, 筛 Wald 和 QLR 检验的大样本性质 (一致性, 收敛速度, 极限分布)
- Chen-Christensen (15): 得到未知函数 h 的筛 NPIV 估计的非线性福利泛函的最优极大值范数收敛速度和一致置信区间, 应用于内生需求估计 (例 1b)。

例 3: 具有不可观察的异质性的久期模型

- Heckman and Singer (84): 根据下式生成 i.i.d. 数据 $\{T_i, X_i\}_{i=1}^n$:

$$p(T|X, \beta_0, h_0) = \int_{\mathcal{U}} g(T|X, u, \beta_0) f_U(u) du,$$

- $g(T|X, u, \beta_0)$: 久期 T 在给定不可观测的异质性 U 和可观测的 X 条件下的密度分布。例如, g 可以是 Weibull 密度:

$$g(T|X, u, \beta_0) = \beta_{0,1} T^{\beta_{0,1}-1} \exp \left[\beta'_{0,2} X + u - T^{\beta_{0,1}} \exp(\beta'_{0,2} X + u) \right].$$

- U 与 X 独立。如果使用的密度函数 $f_U(u) \equiv h^2(u)$ 有误, 那么 β_0 的估计值将会不一致。
- 半参数混合模型被广泛应用在这类问题中。

- 令 $\alpha_0 = (\beta_0, h_0) \in B \times \mathcal{H}$, 它可以通过筛 MLE 来估计:

$$\hat{\alpha}_n = \arg \max_{\beta \in B, h \in \mathcal{H}_n} \sum_{i=1}^n \log \left\{ \int_{\mathcal{U}} g(T_i | X_i, u, \beta) h^2(u) du \right\}$$

其中 \mathcal{H}_n 是筛空间; 随着 $n \rightarrow \infty$, \mathcal{H}_n 在 \mathcal{H} 中变得稠密。

- Heckman and Singer (84): \mathcal{H}_n 为一阶样条筛; 一致性。
- Chen and Liao (14): \mathcal{H}_n 为 Hermite 多项式筛; 基于筛似然率 (LR) 的推断, 并应用在中国数据上。
- Chen, Tamer and Torgovitsky (11): 自举筛 LR 推断对点识别的潜在损失稳健。
- 这一理论可以直接应用于随机参数、测量误差等问题。

表格：中国二胎的久期分析¹

AIC	筛 MLE	标准误	高斯置信区间 ⁵	卡方置信区间 ⁵
log(久期)	3.3621	0.0975	(3.1710 3.5532)	(3.2720 3.4496)
常数	-5.3175	0.2624	(-5.8317 -4.8032)	(-5.4516 -5.1844)
第一胎性别 ²	0.0584	0.1256	(-0.1877 0.3045)	(-0.1182 0.2322)
教育年数	0.0088	0.0181	(-0.0267 0.0442)	(-0.0102 0.0275)
独生子女奖励 ³	-0.2270	0.1331	(-0.4880 0.0339)	(-0.4164 -0.0418)
户口类别 ⁴	0.7900	0.1769	(0.4433 1.1366)	(0.6427 0.9365)
BIC	筛 MLE	标准误	高斯置信区间	卡方置信区间
log(久期)	3.2404	0.0718	(3.0090 3.2905)	(3.0642 3.2321)
常数	-4.9435	0.2284	(-5.3232 -4.4281)	(-5.0064 -4.7476)
第一胎性别	0.0397	0.0840	(-0.0993 0.2298)	(-0.1100 0.2309)
教育年数	0.0075	0.0157	(-0.0253 0.0364)	(-0.0129 0.0235)
独生子女奖励	-0.2836	0.1118	(-0.4816 -0.0434)	(-0.4459 -0.0818)
户口类别	0.7326	0.1591	(0.4291 1.0526)	(0.5974 0.8824)

1. 样本量为 n=694; 2. 如果第一胎为女孩, 则性别哑变量值为 1, 否则为 0; 3. 如果家庭因为服从独生子女政策而得到奖励, 则独生子女奖励哑变量为 1, 否则为 0; 4. 如果家庭户口类别为农村户口, 则户口类别哑变量为 1, 否则为 0; 5. 95% 置信区间; 来源: Chen-Liao (14)

例 4: 半非参数 GARCH + 残差 Copula 模型

- Geanakoplos (10): 坏消息通常伴随着不确定性 (波动性) 上升。“新闻影响曲线”。
- Engle (10): “风险评估” 对于理解金融危机也很重要。
- 为了同时反映“新闻影响曲线”和尾部联合相关性, 我们考虑半非参 GARCH + 残差 Copula 模型。
- 我们基于三大资产类别: 抵押担保债券 (MBS)、股票和债券市场回报, 使用过去四年的每日交易数据来处理上述“新闻影响曲线”和风险评估问题。

Chen-Fan (06) SCOMDY 模型: 巴克莱 MBS 指数的超额回报 (S_t^e), 股票市场 (每日 Fama-French 因子) 超额回报 (M_t^e), 以及巴克莱债券指数的超额回报 (B_t^e):

$$\text{MBS 市场} : S_t^e = c_S + \rho_S S_{t-1}^e + \beta_S M_{t-1}^e + \sigma_{S,t} \varepsilon_{S,t}$$

$$\text{股票市场} : M_t^e = c_M + \rho_M M_{t-1}^e + \sigma_{M,t} \varepsilon_{M,t}$$

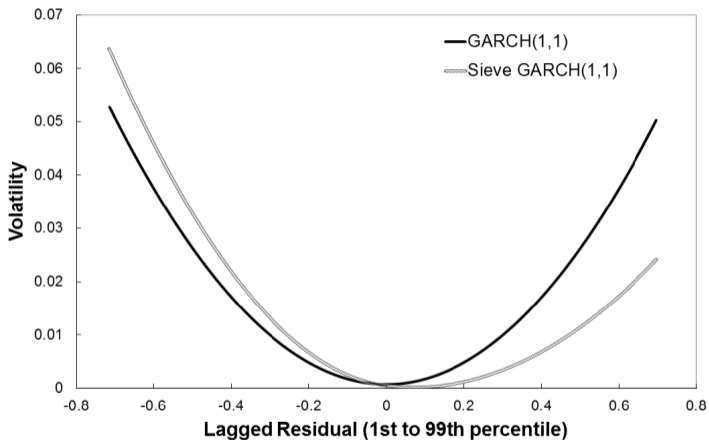
$$\text{债券市场} : B_t^e = c_B + \rho_B B_{t-1}^e + \beta_B M_{t-1}^e + \sigma_{B,t} \varepsilon_{B,t}$$

$$\text{波动性} : \sigma_{i,t}^2 = \omega_i + \theta_i \sigma_{i,t-1}^2 + h_i (\sigma_{i,t-1} \varepsilon_{i,t-1}), \quad i \in \{S, M, B\},$$

对 $i \in \{S, M, B\}$, 有 $E(\varepsilon_{i,t}) = 0$, $E(\varepsilon_{i,t}^2) = 1$ 。 $(\varepsilon_{S,t}, \varepsilon_{M,t}, \varepsilon_{B,t})'$ 不同期是独立的, 但在同一期内服从联合分布 $F(\varepsilon) = C(F_1(\varepsilon_1), F_2(\varepsilon_2), F_3(\varepsilon_3); \Sigma, \nu)$, 其中 $C(\cdot) : [0, 1]^3 \rightarrow [0, 1]$ 是 Student's t-copula, 其未知参数为 (Σ, ν) , 未知边缘分布为 $F_i(\cdot)$, $i \in \{S, M, B\}$ 。

- 这三个估计的“新闻影响曲线”全都表现出相同的不对称性：由坏消息导致的波动性上升比好消息更多。对于 MBS 和股票，一些好消息实际上减少了波动性。这一结果同 Fostel and Geanakoplos (10) 一致。此外，在股市中绝大多数好消息对波动性没有太大的影响。Linton and Mammen (05) 也有类似的结论。
- 我们发现 (i) 对债券的冲击和对 MBS 的冲击高度正相关，(ii) 对 MBS 的冲击和对股票的冲击是中度负相关的，(iii) 对债券的冲击和对股票的冲击也是中度负相关的。
- 借助估计得到的半非参 GARCH 和残差 Copula 相关性参数，我们可以很容易地计算由 MBS、股票和债券组成的投资组合的 VaR。

图: MBS: 新闻影响曲线



当 MBS 受到负面新闻冲击时, 应用 Sieve-GARCH(1,1) 预测的波动性比标准 GARCH(1,1) 更大; 而当 MBS 受到正面新闻影响时, 应用 Sieve-GARCH(1,1) 预测的波动性比标准 GARCH(1,1) 更小。来源: Chen (2013)

我过去在筛 M 方法上的研究小结

在筛 M 估计和推断方面：

- Chen-Shen (98): 半非参时间序列模型的一般筛 M 估计的收敛速度和极限分布。ANN 非线性筛可以作为其中的一个例子。
- Chen-White (99): 使用 ANN 非线性筛估计条件密度和高维协变量的函数的条件分位数。
- Chen-Liao-Sun (14): 线性筛 M-估计量的收敛速度可能慢于 \sqrt{n} 的可估泛函的极限分布，弱相关时间序列模型的筛 Wald 检验。
- Chen-Tamer-Torgovitsky (11): 对潜在部分识别的半非参似然模型的自举法筛 LR 推断
- Chen-Liao (15): 对弱相关数据的筛半参两步法 GMM 估计，其中第一步可以是线性筛 M 或筛 MD 估计。

展望：发现更多未知

- 目前仍然缺乏关于收敛速度慢于 \sqrt{n} 的泛函的推断结果，例如其一致置信区间或基于非线性筛（如 ANN，或“深度学习”）的筛 Wald 和 QLR 检验等。
- 在为分析具有非线性非高斯潜在结构的半非参动态模型（如 DSGE）提出的基于数值模拟的筛分方法上需要更深入的研究。
- 大数据时代：更复杂的经济模型，更多政策反事实分析的可能性，更多由内生性、潜在异质性、潜在动态状态相关性、非平稳性、测度误差、变量（或模型）选择所带来的问题。
- 惩罚筛方法仍然适用，但我们需要更谨慎的研究随机优化算法和计量性质之间的相互作用，并且相应修改现有模型和推断理论。

谢谢大家!