数量经济研究 The Journal of Quantitative Economics
第 4 卷 第 1 辑 VoL. 4. No.1
2013 年 3 月 March 2013

近似因子结构设定下的稳健 GLS 估计

余壮雄[1] 王美今[2]
（1. 暨南大学产业经济研究院，510632；2. 中山大学岭南学院，510275）

摘要：共同因子和截面弱相关并存的近似因子结构，是目前计量经济建模中对客观经济现象中不同个体相关性的最好刻画。针对平稳 Panel 的误差近似因子结构模型的参数估计和检验，本文提出一种稳健的 GLS 估计（RGLS）来改进 GLS 估计的小样本性质。我们证明了，RGLS 估计是一致估计，而且比 ROLS 渐近有效，其 t 统计量收敛于标准正态；数值模拟的结果表明 RGLS 估计（特别是迭代的 RGLS）有非常良好的小样本表现。

关键词：近似因子结构 稳健 GLS 共同因子 截面弱相关

Robust GLS Estimation for Approximate Factor Models

Abstract: This paper focuses on estimation and inference for panel stationary models with approximate factor structure, which is the most reasonable setting of the correlation among different individual in economic reality. We propose a robust GLS estimator, RGLS, to improve the small sample performance of GLS. We prove that the RGLS estimator is consistent and asymptotic efficient than the ROLS, and its t-statistics sequential converge to standard normal distribution. Results of simulation also show that the small sample performance of the (iterated) RGLS estimator is very good.

Keywords: Approximate Factor Structure; Robust GLS; Common Factor; Weak Cross-Sectional Dependence

引 言

截面相关是面板数据分析中的一个重要设定，在经济与社会科学的实证分析中有着广泛的应用（Sarafidis and Wansbeek, 2012）。基于相关性的强度，截面相关可分为截面弱相关和强相关两种形式（Chudik et. al., 2009）：截面弱相关通常是系统内某个变量的冲击通过扩散传递给其它变量从而使得变量之间呈现某种相关性，此时变量的协方差阵对应的特征根阶数为 $O_p(1)$；截面强相关则是由系统外的冲击引起系统内各变量的同步波动，这些冲击也称为共同因子，此时协方差阵对应的特征根阶数为 $O_p(n)$。

对于解释变量满足外生性的情形，如果误差结构存在截面相关，普通的 LS 估计仍然可以获得参数的一致估计，但不再有效。为了改进估计量的有效性，可以利用 LS 估计的残差来估计截面相关矩阵，再使用 SUR 的方法（Zellner, 1962）①计算参数的可行 GLS 估计。遗憾的是，此时 FGLS 估计的可行性依赖于样本结构 $T \geq n$；当 $T < n$ 时，由第一步 LS 估计得到的残差计算的协方差阵不可逆，

① 一个较新的回顾可见 Moon 和 Perron（2006）。

FGLS 不再可行[②]；即便是$T \geq n$的情形，如果 T 比较接近 n，由于自由度损失太大，FGLS 估计的小样本表现也非常不好。

为了克服 FGLS 估计受制于样本结构的问题，必须对截面相关矩阵施加某些设定以减少待估参数的维度。其中一种思路是借鉴空间计量经济学（Anselin, 2001）的分析方法引入空间相关的结构；另一种思路则是将截面相关表示为共同因子（Chamberlin and Rothschild, 1983）的结构（早期的学者们并没有发现共同因子结构与普通的截面弱相关在强度上具有本质的区别)。由于共同因子似乎对应着经济系统背后的某种驱动力量或因素（Stock and Waston, 2002）[③]，误差因子结构设定就成为面板数据分析的一种基本研究模式。

Coakley 等（2002）基于误差因子结构的设定提出了残差主成分估计量（RPC），他们建议利用 LS 回归的残差计算协方差阵，再从协方差阵中提取最大的若干个特征根对应的主成分作为新增变量放入回归方程。但是，Pesaran（2003）的后续研究表明，当回归方程的解释变量与共同因子之间存在相关，即解释变量具有内生性时，RPC 估计量通常不是一致估计（Pesaran 给出了约束因子载荷系数为固定常数下的证明[④]）；由于解释变量与共同因子之间的相关性，第一步回归得到并用于计算协方差阵的残差不再是误差的一致估计，这种偏差通过影响提取的主成分最终导致增广回归系数的不一致。

Pesaran（2003）的研究质疑了 RPC 估计量这种两步回归方法的合理性，同时，也凸显解释变量内生性问题对误差因子结构设定的重要性。在截面相关矩阵一般化的设定下，解释变量与误差项的相关性无法很好刻画，此时假定解释变量外生是常用的设定形式；而共同因子的引入，使得解释变量内生性的描述在技术上成为一种方便形式；特别是，由于解释变量本身的数据生成过程往往包含共同因子的影响，解释变量与误差共同因子之间的这种相关设定可能更符合客观实际。

Kapetanios 和 Pesaran（2007）借鉴 Coakley 等（2002）的研究思路，提出了另一种在回归方程中加入主成分来改进估计量效率的方法（PC 估计量），不同的是他们提取的主成分来自所有变量（解释变量和被解释变量）的相关矩阵，不再依赖于第一步回归残差的一致性。Bai（2009）则在 Coakley 等（2002）的基础上提出了迭代的主成分估计（IPC），他证明了通过不断迭代计算主成分和增广方程的系数，Coakley 等（2002）的 RPC 估计会收敛到回归方程包含因子时的非线性回归系数，即得到参数的一致估计。

与主成分估计的思路不同，Pesaran（2006）提出了一种共同相关效应估计（Common Correlated Effects：CCE[⑤]）的方法，将所有变量的个体均值作为新增变量放入回归方程来改进回归系数的估计效率。与主成分方法相比，CCE 估计是一种比较稳健的估计，它不依赖于具体的共同因子及其个数，因此也不存在因子个数识别出错所带来的偏差。Pesaran 和 Tosetti（2007）进一步将 CCE 估计量扩展到同时存在空间弱相关和共同因子的情形，他们证明了 CCE 估计在这种设定下仍然是一致估计。

[②] 使用广义逆计算的 FGLS 估计的性质不具有良好性质的保证。

[③] 部分学者也在尝试识别共同因子对应的具体指标（Bai and Ng, 2006；Chen, 2012）。

[④] 后续的研究（Sarafidis and Wansbeek, 2012）提出了通过使得因子载荷系数的均值为 0 的处理来避开 Pesaran（2003）所提出的这一问题，处理的方法是第一步回归使用双向固定效应回归。

[⑤] CCE 根据设定有两个对应的估计量：CCEP 和 CCEMG，前者对应回归系数对不同个体同质，后者对应回归系数对不同个体异质。根据本文研究的模型设定，这里只使用 CCEP 作为比较。

误差因子结构的设定为截面相关的来源作出一种具有充实内容的解释，但是，后续的研究（Chudik et. al., 2009）表明因子结构与截面弱相关之间并不存在替换关系，两者之间具有明显的强度与表现形式差异。也就是说，现实世界更可能是共同因子与截面弱相关并存的情形，即近似因子结构。然而，一旦在因子结构的基础上引入截面弱相关，原本基于因子结构设定的估计量也将会失去有效性，因为这相当于在仅有截面弱相关时进行 LS 估计，唯一的区别是克服了解释变量与共同因子相关引起的内生性偏差；同时，其 t 统计量会因为弱相关的存在变得不再正确。

如何在近似因子结构下基于已有的各种估计方法构建具有较小偏差而且存在合适 t 统计量的估计方法是本文研究的目的。本文余下部分的结构安排如下：第二部分提出本研究的基本模型与假定；第三部分阐述各种已有的估计方法并提出本文的稳健 GLS（RGLS）估计；第四部分给出 RGLS 估计量及其 t 统计量的具体计算步骤；第五部分将模型扩展到具有个体效应以及解释变量与共同因子相关的情形；第六部分比较了几种估计量及其 t 统计量的小样本表现；最后为本文的结论。

1　基本模型与假定

考虑如下 Pool 模型：

$$y_{it} = x_{it}'\beta + \varepsilon_{it} \quad , \quad i = 1, \cdots, n; t = 1, \cdots, T \tag{1}$$

其中，下标 i 表示不同的个体，下标 t 表示不同的时期，参数 β 不随个体和时期变动；$X = (x_1, x_2, \cdots, x_T)$，$Y = (y_1, y_2, \cdots, y_T)$，$\varepsilon = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_T)$，$x_t' = (x_{1t}, x_{2t}, \cdots, x_{nt})$，$y_t' = (y_{1t}, y_{2t}, \cdots, y_{nt})$，$\varepsilon_t' = (\varepsilon_{1t}, \varepsilon_{2t}, \cdots, \varepsilon_{nt})$，$x_{it}' = (x_{it,1}, x_{it,2}, \cdots, x_{it,K})$。

假定 1（共同因子设定）：$\varepsilon_t = \Gamma v_t + \eta_t$，$v_t \sim i.i.d.(0, \Pi)$，$\eta_t \sim i.i.d.(0, \Sigma)$；$v_t$ 和 η_t 相互独立且都存在有限四阶矩。

其中，v_t 为 r 维的共同因子列向量，η_t 为 n 维的个体特有误差向量；Γ 为 $n \times r$ 的因子载荷系数矩阵，Π 为 $r \times r$ 的有限正定对角阵，Σ 为 n 维的有限正定矩阵[⑥]；r 为常数

同样的，不妨记 $v = (v_1, v_2, \cdots, v_T)$，$\eta = (\eta_1, \eta_2, \cdots, \eta_T)$，$\Pi = \mathrm{diag}(\pi_1, \pi_2, \cdots, \pi_r)$，$\Gamma = (\Gamma_1, \Gamma_2, \cdots, \Gamma_r)$；令 Σ 矩阵的谱分解形式为 $\Sigma = P\Lambda P'$，其中 $P = (P_1, P_2, \cdots, P_n)$为特征向量矩阵，$\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ 为对应的特征根矩阵。

假定 2（载荷参数设定）：$\dfrac{1}{n}\Gamma'\Gamma \xrightarrow{n\to\infty} \Phi$。

其中，Φ 为 $r \times r$ 的有限正定矩阵。

由假定 1 和 2 有，$\varepsilon_t \sim i.i.d.(0, \Omega)$ 且存在有限 4 阶矩，其中 $\Omega = \Gamma\Pi\Gamma' + \Sigma$ 为 $n \times n$ 的截面相关矩阵。易知，Ω 有 r 个特征根为 $O_p(n)$，n-r 个特征根为有限常数。

假定 3（外生性设定）：$\mathrm{E}(\varepsilon|X) = 0$，$\mathrm{E}(\varepsilon_t\varepsilon_s|X) = \begin{cases} \Omega & , \quad t = s \\ 0 & , \quad t \neq s \end{cases}$。

为了便于表示，不妨记 $x_t = (x_{t,1}, x_{t,2}, \cdots, x_{t,K})$，$x_{t,k}' = (x_{1t,k}, x_{2t,k}, \cdots, x_{nt,k})$。

假定 4（平稳性设定）：$\frac{1}{T}\sum_{t=1}^{T} x_{t,k}x_{t,l}' \xrightarrow{T\to\infty} \mathrm{M}_{kl}$，$\frac{1}{n}tr(\mathrm{M}_{kl}) \xrightarrow{n\to\infty} \varphi_{kl}$，对任意的

[⑥] 事实上，在下文的 OLS、ROLS 和 RGLS 估计中 Σ 允许为半正定矩阵的情况，只需要假定 Σ 中大于 0 的特征根的比例（渐近）大于 0 即可。

$k,l\in\{1,2,\cdots,K\}$。

其中，M_{kl} 为 $n\times n$ 的有限半正定矩阵，φ_{kl} 为大于 0 的常数。

则由假定 4 有：

$$\frac{1}{nT}\sum_{t=1}^{T}x_t'x_t\xrightarrow{n,T\to\infty}\Psi_X$$

其中，$\Psi_X=(\varphi_{kl})$ 为 $K\times K$ 的有限正定矩阵。

上式中的极限"$n,T\to\infty$"使用的是 Phillips 和 Moon（1999）定义的序贯极限的概念，即先固定 n 让 T 趋于无穷，再让 n 趋于无穷。

在假定 4 的基础上，假定有如下极限结果成立：

$$\frac{1}{n}tr(\Sigma)\xrightarrow{n\to\infty}\bar{\lambda}\ \ ,\ \ \frac{1}{n}tr(\Sigma\,\mathrm{M}_{kl})\xrightarrow{n\to\infty}\varphi_{\Sigma,kl}\ \ ,\ \ \frac{1}{n}tr(\Sigma^{-1}\mathrm{M}_{kl})\xrightarrow{n\to\infty}\varphi_{\Sigma1,kl}$$

$$\frac{1}{n}\Gamma'\mathrm{M}_{kl}\Gamma\xrightarrow{n\to\infty}\Phi_{kl}\ \ ,\ \ \frac{1}{n}\Gamma'\Sigma\,\Gamma\xrightarrow{n\to\infty}\Phi_{\Sigma}\ \ ,\ \ \frac{1}{n}\Gamma'\Sigma^{-1}\Gamma\xrightarrow{n\to\infty}\Phi_{\Sigma1}$$

$$\frac{1}{n}\Gamma'\Sigma\,\mathrm{M}_{kl}\Gamma\xrightarrow{n\to\infty}\Phi_{\Sigma,kl}\ \ ,\ \ \frac{1}{n}\Gamma'\Sigma^{-1}\mathrm{M}_{kl}\Sigma^{-1}\Gamma\xrightarrow{n\to\infty}\Phi_{\Sigma12,kl}$$

其中，$\bar{\lambda}$、$\varphi_{\Sigma,kl}$ 和 $\varphi_{\Sigma1,kl}$ 为有限常数，且 $\bar{\lambda}>0$；Φ_{kl}、Φ_{Σ}、$\Phi_{\Sigma1}$、$\Phi_{\Sigma,kl}$ 和 $\Phi_{\Sigma12,kl}$ 为 $r\times r$ 的有限矩阵，且有 $\Phi_{\Sigma}>0$ 和 $\Phi_{\Sigma1}>0$。

记 $\Psi_{\Sigma,X}=(\varphi_{\Sigma,kl}),\Psi_{\Sigma1,X}=(\varphi_{\Sigma1,kl}),\varphi_{\Pi,kl}=tr(\Pi\Phi_{kl}),\Psi_{\Pi,X}=(\varphi_{\Pi,kl})$；同样的，可知 $\Psi_{\Sigma,X}$、$\Phi_{\Sigma1,X}$ 和 $\Psi_{\Pi,X}$ 为 $K\times K$ 的有限正定矩阵。

2　基本估计量

2.1　OLS 估计和 GLS 估计

OLS 估计：

$$\hat{\beta}_{OLS}=\left(\sum_{t=1}^{T}x_t'x_t\right)^{-1}\sum_{t=1}^{T}x_t'y_t \tag{2}$$

$$var(\hat{\beta}_{OLS})=\left(\sum_{t=1}^{T}x_t'x_t\right)^{-1}\left(\sum_{t=1}^{T}x_t'\Omega x_t\right)\left(\sum_{t=1}^{T}x_t'x_t\right)^{-1} \tag{3}$$

$$t_{OLS}=\left(\sigma^2\sum_{t=1}^{T}x_t'x_t\right)^{-1/2}\sum_{t=1}^{T}x_t'y_t \tag{4}$$

实际计算时，标准误 σ^2 使用 $\hat{\sigma}^2=\frac{1}{nT}\sum_{t=1}^{T}\hat{\varepsilon}_t'\hat{\varepsilon}_t$ 代替，$\hat{\varepsilon}_t$ 为 OLS 估计的残差；协方差阵 Ω 使用一致估计量 $\hat{\Omega}=\frac{1}{T}\sum_{t=1}^{T}\hat{\varepsilon}_t\hat{\varepsilon}_t'$ 代替。

GLS 估计：

$$\hat{\beta}_{GLS}=\left(\sum_{t=1}^{T}x_t'\Omega^{-1}x_t\right)^{-1}\sum_{t=1}^{T}x_t'\Omega^{-1}y_t \tag{5}$$

$$var(\hat{\beta}_{GLS})=\left(\sum_{t=1}^{T}x_t'\Omega^{-1}x_t\right)^{-1} \tag{6}$$

$$t_{GLS}=\left(\sum_{t=1}^{T}x_t'\Omega^{-1}x_t\right)^{-1/2}\sum_{t=1}^{T}x_t'\Omega^{-1}y_t \tag{7}$$

其中，$\Omega^{-1}=\Sigma^{-1}-\Sigma^{-1}\Gamma(\Pi^{-1}+\Gamma'\Sigma^{-1}\Gamma)^{-1}\Gamma'\Sigma^{-1}$。

由于求逆的过程基本消除了 Ω 矩阵中 $O_p(n)$特征根的影响，GLS 估计事实上等价于先消除共同因子再对弱相关矩阵求逆的估计过程。Moon 和 Perron（2004）在研究 Panel 单位根检验时建议使用正交化来消除共同因子的影响，他们的这种方法可以看作是 GLS 估计的一种简化算法，即对协方差阵的逆矩阵使用简化计算。

MP 简化：

$$\Omega_m^{-1}=\Sigma^{-1}-\Sigma^{-1}\Gamma(\Gamma'\Sigma^{-1}\Gamma)^{-1}\Gamma'\Sigma^{-1}$$

由于 Π^{-1} 是有限常数矩阵，而 $\Gamma'\Sigma^{-1}\Gamma$ 则是$O_p(n)$的矩阵，因此在GLS估计中使用Ω^{-1} 和 Ω_m^{-1} 是渐近等价的。

2.2 ROLS 估计（Robust OLS）

Beck 和 Katz（1995）建议在计算 OLS-t 统计量时使用真实的协方差阵来校正截面相关的影响，他们把这种计算方法称为 PCSE（Panel-corrected Standard Errors）方法。Breitung 和 Das（2005）在 Panel 单位根检验中也使用了相同的方法来计算 OLS-t 统计量，但他们把这种方法称为 ROLS（Robust OLS）估计。由于 ROLS 估计更能体现在计算 OLS-t 统计量时使用的是稳健协方差阵，本文以下使用 ROLS 估计的叫法。

ROLS-t：

$$t_{ROLS}=\left(\sum_{t=1}^{T}x_t'x_t\right)^{1/2}\left(\sum_{t=1}^{T}x_t'\Omega x_t\right)^{-1/2}\left(\sum_{t=1}^{T}x_t'x_t\right)^{-1/2}\sum_{t=1}^{T}x_t'y_t$$

当 K=1 时，上式可简化为：

$$t_{ROLS}=\left(\sum_{t=1}^{T}x_t'\Omega x_t\right)^{-1/2}\sum_{t=1}^{T}x_t'y_t \tag{8}$$

数值模拟的结果（Beck 和 Katz，1995；Breitung 和 Das，2005）表明，ROLS-t 统计量的小样本表现要远远优于 OLS-t 和 GLS-t 统计量[⑦]的表现，甚至在 n>T 或协方差阵奇异的情形下，ROLS-t 统计量仍然具有相当良好的小样本表现[⑧]。

2.3 RGLS 估计（Robust GLS）

由于 GLS 估计在有限样本使用时可能达不到该有的效果，本文借鉴 ROLS 估计的思想对 GLS 估计进行适当的调整以改进其小样本表现。具体处理如下：将截面弱相关矩阵处理成单位阵进行 GLS 估计，在计算 t 统计量时使用 PCSE 方法。我们将这种方法称为稳健 GLS 估计（RGLS）。

RGLS 估计：

$$\hat{\beta}_{RGLS}=\left(\sum_{t=1}^{T}x_t'\Omega_*^{-1}x_t\right)^{-1}\sum_{t=1}^{T}x_t'\Omega_*^{-1}y_t \tag{9}$$

$$var(\hat{\beta}_{RGLS})=\left(\sum_{t=1}^{T}x_t'\Omega_*^{-1}x_t\right)^{-1}\left(\sum_{t=1}^{T}x_t'\Omega_*^{-1}\Omega\Omega_*^{-1}x_t\right)\left(\sum_{t=1}^{T}x_t'\Omega_*^{-1}x_t\right)^{-1} \tag{10}$$

---

⑦ 本文在提到某种估计方法的估计量或 t 统计量的小样本表现时都特指这种估计方法的可行估计。

⑧ Breitung 和 Das（2008）最近的研究表明在含有非平稳共同因子的单位根检验中 ROLS-t 统计量不再具有良好的性质与表现。

$$t_{RGLS}=\left(\sum_{t=1}^{T}x_t'\Omega_*^{-1}x_t\right)^{1/2}\left(\sum_{t=1}^{T}x_t'\Omega_*^{-1}\Omega\Omega_*^{-1}x_t\right)^{-1/2}$$
$$\left(\sum_{t=1}^{T}x_t'\Omega_*^{-1}x_t\right)^{-1/2}\sum_{t=1}^{T}x_t'y_t$$

当 K=1 时，上式可简化为：

$$t_{RGLS}=\left(\sum_{t=1}^{T}x_t'\Omega_*^{-1}\Omega\Omega_*^{-1}x_t\right)^{-1/2}\sum_{t=1}^{T}x_t'\Omega_*^{-1}y_t \tag{11}$$

其中，$\Omega_*=\bar{\lambda}^{-1}\cdot\Gamma\Pi\Gamma'+I$，$\bar{\lambda}=\frac{1}{n}tr(\Sigma)$；$\Omega_*^{-1}=I-\Gamma(\bar{\lambda}\cdot\Pi^{-1}+\Gamma'\Gamma)^{-1}\Gamma'$。

同样的，Ω_*^{-1} 的 MP 简化可表示为 $\Omega_{m*}^{-1}=I-\Gamma(\Gamma'\Gamma)^{-1}\Gamma'\equiv M_\Gamma$。

由式（9）可知，MP 简化的 RGLS 估计等价于在回归方程中加入共同因子载荷系数后的 LS 估计，其思想与 Coakley 等（2002）的 RPC 估计类似，只不过后者在回归方程中加入的共同因子本身。

如果 r=0，则有 $\Omega_*=I$，此时 RGLS 估计和 ROLS 估计等价。

对比 GLS 估计与 RGLS 估计，可以知道，式（1）的 GLS 估计等价于对消除共同因子影响后的数据进行 GLS 估计，而式（1）的 RGLS 估计则等价于对消除共同因子影响后的数据进行 ROLS 估计。相比 GLS 估计而言，RGLS 估计拥有了 ROLS 估计所有的优点，它同样不需要计算 $n\times n$ 的截面弱相关矩阵 Σ 的逆，而且允许 n>T 或 Σ 奇异的情形。

2.4 （渐近）有效性比较

注意到，$\Omega^{-1}\Omega=I$，$\Omega^{-1}\Omega\Omega_*^{-1}=\Omega_*^{-1}$，容易证明 GLS 估计比 OLS 估计和 RGLS 估计都要更有效。

又因为，

$$\frac{1}{nT}\sum_{t=1}^{T}x_t'\Omega_*^{-1}x_t=\frac{1}{nT}\sum_{t=1}^{T}x_t'x_t+o_p(1) \tag{12}$$

$$\frac{1}{nT}\sum_{t=1}^{T}x_t'\Omega_*^{-1}\Sigma\,\Omega_*^{-1}x_t=\frac{1}{nT}\sum_{t=1}^{T}x_t'\Sigma\,x_t+o_p(1) \tag{13}$$

其中，$\Psi_{\Sigma,X}=(\varphi_{\Sigma,kl})$，

$$\begin{aligned}\frac{1}{nT}\sum_{t=1}^{T}x_{t,k}'\Omega_{m*}^{-1}x_{t,l}&=\frac{1}{nT}\sum_{t=1}^{T}x_{t,k}'x_{t,l}-\frac{1}{nT}\sum_{t=1}^{T}tr\left[(\Gamma'\Gamma)^{-1}\Gamma'x_{t,l}x_{t,k}'\Gamma\right]\\&\xrightarrow{T\to\infty}\frac{1}{n}tr(\mathrm{M}_{kl})-\frac{1}{n}tr\left[(\Gamma'\Gamma)^{-1}\Gamma'\mathrm{M}_{kl}\Gamma\right]\\&\xrightarrow{n\to\infty}\varphi_{kl}\end{aligned}$$

$$\begin{aligned}\frac{1}{nT}\sum_{t=1}^{T}x_{t,k}'\Omega_{m*}^{-1}\Sigma\,\Omega_{m*}^{-1}x_{t,l}&=\frac{1}{nT}\sum_{t=1}^{T}x_{t,k}'\Sigma\,x_{t,l}-\frac{1}{nT}\sum_{t=1}^{T}tr\left[(\Gamma'\Gamma)^{-1}\Gamma'\Sigma\,x_{t,l}x_{t,k}'\Gamma\right]\\&\quad-\frac{1}{nT}\sum_{t=1}^{T}tr\left[(\Gamma'\Gamma)^{-1}\Gamma'x_{t,l}x_{t,k}'\Sigma\,\Gamma\right]\\&\quad+\frac{1}{nT}\sum_{t=1}^{T}tr\left[(\Gamma'\Gamma)^{-1}\Gamma'\Sigma\,\Gamma(\Gamma'\Gamma)^{-1}\Gamma'x_{t,l}x_{t,k}'\Gamma\right]\\&\xrightarrow{T\to\infty}\frac{1}{n}tr(\Sigma\,\mathrm{M}_{kl})-\frac{1}{n}tr\left[(\Gamma'\Gamma)^{-1}\Gamma'\Sigma\,\mathrm{M}_{kl}\Gamma\right]\\&\quad-\frac{1}{n}tr\left[(\Gamma'\Gamma)^{-1}\Gamma'\Sigma\,\mathrm{M}_{kl}\Gamma\right]\\&\quad+\frac{1}{n}tr\left[(\Gamma'\Gamma)^{-1}\Gamma'\Sigma\,\Gamma(\Gamma'\Gamma)^{-1}\Gamma'\mathrm{M}_{kl}\Gamma\right]\\&\xrightarrow{n\to\infty}\varphi_{\Sigma,kl}\end{aligned}$$

则 OLS 估计量和 RGLS 估计量的渐近协方差可计算如下：

$$
\begin{aligned}
V_0 =& asy.var(\hat{\beta}_{OLS}) \\
=& \lim_{n,T\to\infty}\left\{\left[\frac{1}{nT}\sum_{t=1}^{T}x_t'x_t\right]^{-1}\left[\frac{1}{nT}\sum_{t=1}^{T}x_t'(\Gamma\Pi\Gamma'+\Sigma)x_t\right]\left[\frac{1}{nT}\sum_{t=1}^{T}x_t'x_t\right]^{-1}\right\} \\
=& \Psi_X^{-1}(\Psi_{\Pi,X}+\Psi_{\Sigma,X})\Psi_X^{-1} \\
V_1 =& asy.var(\hat{\beta}_{RGLS}) \\
=& \lim_{n,T\to\infty}\left\{\left[\frac{1}{nT}\sum_{t=1}^{T}x_t'\Omega_*^{-1}x_t\right]^{-1}\left[\frac{1}{nT}\sum_{t=1}^{T}x_t'\Omega_*^{-1}\Omega\Omega_*^{-1}x_t\right]\left[\frac{1}{nT}\sum_{t=1}^{T}x_t'\Omega_*^{-1}x_t\right]^{-1}\right\} \\
=& \lim_{n,T\to\infty}\left\{\left[\frac{1}{nT}\sum_{t=1}^{T}x_t'x_t\right]^{-1}\left[\frac{1}{nT}\sum_{t=1}^{T}x_t'\Sigma\, x_t\right]\left[\frac{1}{nT}\sum_{t=1}^{T}x_t'x_t\right]^{-1}\right\} \\
=& \Psi_X^{-1}\Psi_{\Sigma,X}\Psi_X^{-1}
\end{aligned}
$$

所以有，$asy.var(\hat{\beta}_{OLS}) \geq asy.var(\hat{\beta}_{RGLS})$，即 RGLS 估计比 OLS 估计渐近有效。

2.5 渐近正态性

定理 1. 如果假定 1 至 4 成立，则有

$$\sqrt{nT}(\hat{\beta}_{OLS}-\beta)\xrightarrow{n,T\to\infty}N\left\{0\,,\,\Psi_X^{-1}(\Psi_{\Pi,X}+\Psi_{\Sigma,X})\Psi_X^{-1}\right\} \tag{14}$$

$$\sqrt{nT}(\hat{\beta}_{GLS}-\beta)\xrightarrow{n,T\to\infty}N\left\{0\,,\,\Psi_{\Sigma1,X}^{-1}\right\} \tag{15}$$

$$\sqrt{nT}(\hat{\beta}_{RGLS}-\beta)\xrightarrow{n,T\to\infty}N\left\{0\,,\,\Psi_X^{-1}\Psi_{\Sigma,X}\Psi_X^{-1}\right\} \tag{16}$$

证明：

由于各估计量的证明类似，此处只证明如下的核心结论：

$$\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}x_t'\Omega_{m*}^{-1}\varepsilon_t\xrightarrow{,n,T\to\infty}N(0,\Psi_{\Sigma,X})$$

注意到

$$\mathrm{E}\left(\frac{1}{\sqrt{n}}x_t'\Omega_{m*}^{-1}\eta_t\right)=0$$

$$\mathrm{E}\left(\frac{1}{n}x_{t,k}'\Omega_{m*}^{-1}\eta_t\eta_t'\Omega_{m*}^{-1}x_{t,l}\right)=\mathrm{E}\left(\frac{1}{n}x_{t,k}'\Omega_{m*}^{-1}\Sigma\,\Omega_{m*}^{-1}x_{t,l}\right)=\frac{1}{n}tr\left(\Omega_{m*}^{-1}\Sigma\,\Omega_{m*}^{-1}\mathrm{M}_{lk}\right)=\theta_{lk}$$

则有

$$\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}x_t'\Omega_{m*}^{-1}\varepsilon_t=\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\frac{1}{\sqrt{n}}x_t'\Omega_{m*}^{-1}\eta_t\xrightarrow{T\to\infty}N\left\{0\,,\,\Theta\right\}$$

其中，$\Theta=(\theta_{lk})$。

又由式（13）可知，$\Theta\xrightarrow{n\to\infty}\Psi_{\Sigma,X}$，所以有

$$\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}x_t'\Omega_{m*}^{-1}\varepsilon_t \xrightarrow{,n,T\to\infty} N(0,\Psi_{\Sigma,X})$$

证明完毕。

定理 2. 如果假定 1 至 4 成立，则在 $\beta=0$ 的零假设下 ROLS-t 统计量、GLS-t 统计量和 RGLS-t 统计量序贯收敛于标准正态分布。

（证明显然，省略）

3　RGLS 估计量的计算

3.1　可行的 RGLS 估计

在实证分析中一般不知道真实的协方差阵 Ω，在计算估计量或 t 统计量时只能使用 Ω 的一致估计，因此，上述的 ROLS、GLS 和 RGLS 估计就转化为相应的可行（Feasible）估计。

矩阵 $\hat{\Omega}_*^{-1}$ 实际计算如下：

$$\hat{\Omega}_*^{-1}=I-\hat{\Gamma}\left(\hat{\bar{\lambda}}\cdot\hat{\Pi}^{-1}+\hat{\Gamma}'\hat{\Gamma}\right)^{-1}\hat{\Gamma}'$$

其中，依照 Bai 和 Ng（2004）的估计方法，$\{\hat{\bar{\lambda}},\hat{\Gamma},\hat{\Pi}\}$ 的具体计算见下面的阐述。

定义 $\hat{\Omega}=\frac{1}{T}\sum_{t=1}^{T}\hat{\varepsilon}_t'\hat{\varepsilon}_t$ 的谱分解为：

$$\hat{\Omega}=\mathrm{Q}\Upsilon\mathrm{Q}'=\begin{pmatrix}\mathrm{Q}_r & \mathrm{Q}_{n-r}\end{pmatrix}\begin{pmatrix}\Upsilon_r & 0\\ 0 & \Upsilon_{n-r}\end{pmatrix}\begin{pmatrix}\mathrm{Q}_r'\\ \mathrm{Q}_{n-r}'\end{pmatrix}$$

其中，Υ 和 Q 分别为 $\hat{\Omega}$ 的特征根矩阵与特征向量矩阵，Υ 中的特征根从大到小排列，Υ_r 和 Υ_{n-r} 分别表示由最大的 r 个特征根组成的对角阵和由最小的 n-r 个特征根组成的对角阵，Q_r 为 Υ_r 对应的特征向量矩阵。

给定共同因子的个数 r，则 $\{\hat{\bar{\lambda}},\hat{\Gamma},\hat{\Pi}\}$ 计算如下：

$$\hat{\Gamma}=\sqrt{n}\mathrm{Q}_r\quad,\quad\hat{\Pi}=\frac{1}{n}\Upsilon_r\quad,\quad\hat{\bar{\lambda}}=\frac{1}{n}tr(\Upsilon_{n-r})\tag{17}$$

通常，共同因子的个数 r 也是未知的。本文使用 Bai 和 Ng（2002）建议的共同因子选取准则 IC_1 来确定共同因子的个数。

定义 $\hat{v}=\frac{1}{n}\hat{\Gamma}'\hat{\varepsilon}$，$\hat{\xi}=\hat{\varepsilon}-\hat{\Gamma}\hat{v}$，Bai 和 Ng（2002）的选择准则 IC_1 如下：

$$\hat{r}=\arg\min_{r=0,1,\cdots,r_m}\left\{\log\left[\hat{\sigma}^2(r)\right]+r\cdot\log\left[\frac{nT}{n+T}\right]\frac{n+T}{nT}\right\}$$

其中，$\hat{\sigma}^2(r)=\frac{1}{nT}tr(\hat{\xi}'\hat{\xi})$，$r_m$ 为某个有限的常数。

3.2　迭代 RGLS 估计

可行的 RGLS 估计（FRGLS）使用 OLS 估计的残差来计算截面相关矩阵，其一致性依赖于 OLS 估计的一致性；但是，上文的证明表明，RGLS 估计渐近有效于 OLS 估计，同样的，也必定有 FRGLS 估计渐近有效于 OLS 估计，因此，使用 RGLS 估计的残差重新计算截面相关矩阵再计算 FRGLS 估

计，可以提高估计量的准确性，改进小样本表现。特别是对于解释变量与共同因子相关的情形，LS 估计的残差不再一致，但通过迭代可以使结果收敛于非线性估计的结果，而后者是一致的。

本文建议使用迭代的 RGLS 估计，具体计算如下：

STEP-1. 根据 OLS 估计的残差识别共同因子的个数，计算 $\{\hat{\hat{\lambda}}, \hat{\Gamma}, \hat{\Pi}\}$，从而计算可行的 RGLS 估计量及其 t 统计量；

STEP-2. 根据上一步可行 RGLS 估计的残差重新计算 $\{\hat{\hat{\lambda}}, \hat{\Gamma}, \hat{\Pi}\}$，从而计算可行的 RGLS 估计量及其 t 统计量。

STEP-3. 重复计算 STEP-2 直至可行的 RGLS 估计收敛。

如果只计算 STEP-1 和 STEP-2，得到的 FRGLS 估计称为两步迭代的 RGLS 估计。后面的模拟发现，RGLS 估计在本文设定的条件下收敛速度非常快，两步迭代的 RGLS 估计已经非常接近迭代至收敛的 RGLS 估计。

需要强调的是，基于 MP 简化的迭代 RGLS 估计在本质上也对应着 Bai（2009）的 IPC 估计，两者的区别在于前者在回归方程中增加因子载荷估计，而后者则是增加因子的估计。

4　模型扩展

4.1　包含参数随个体变动的共同解释变量

不妨记$\tilde{y} = (y_1', y_2', \cdots, y_T')'$，$\tilde{x} = (x_1', x_2', \cdots, x_T')'$，$\tilde{\varepsilon} = (\varepsilon_1', \varepsilon_2', \cdots, \varepsilon_T')'$，此时基本模型可表示如下：

$$\tilde{y} = \tilde{x}\beta + (z \otimes I_n)\alpha + \tilde{\varepsilon} \tag{18}$$

其中，z 为$T \times l$的外生解释变量矩阵，对不同的个体取值相同，α为$n \cdot l \times 1$的待估参数向量，l 为常数。

如果$z = i_T$，α对应个体固定效应或随机效应。

令$M_T = I_T - z(z'z)^{-1}z'$，式（18）可变换如下：

$$y_* = x_*\beta + \varepsilon_* \tag{19}$$

其中，$y_* = (M_T \otimes I_n)\tilde{y}$，$x_* = (M_T \otimes I_n)\tilde{x}$，$\varepsilon_* = (M_T \otimes I_n)\tilde{\varepsilon}$；则$\varepsilon_*$的协方差矩阵为$E[\varepsilon_*\varepsilon_*'] = M_T \otimes \Omega$。

此时，可行的 RGLS 估计可计算如下：

$$\hat{\beta}_{RGLS} = \left[\tilde{x}'(M_T \otimes \hat{\Omega}_*^{-1})\tilde{x}\right]^{-1} \tilde{x}'(M_T \otimes \hat{\Omega}_*^{-1})\tilde{y} \tag{20}$$

其中，$\hat{\Omega} = \frac{1}{T}\sum_{t=1}^{T} \hat{\varepsilon}_{*t}\hat{\varepsilon}_{*t}'$，$\hat{\varepsilon}_{*t}$为式（19）的 LS 估计的残差，$\hat{\Omega}_*^{-1}$可由$\hat{\Omega}$算出。

容易证明，此时$\hat{\Omega}$仍然是Ω的一致估计，因此，根据式（20）计算的可行 RGLS 估计量仍然是β的一致估计。

4.2　解释变量与共同因子相关

基本模型可表示如下：

$$y_t = x_t\beta + \varepsilon_t = x_t\beta + \Gamma v_t + \eta_t \tag{21}$$

其中，x_t和Γ与η_t无关。

由上文迭代 RGLS 估计的计算过程可知，基于 MP 简化的迭代 RGLS 估计对应如下方程组的解：

$$\hat{\beta} = \left(\sum_{t=1}^{T} x_t' M_{\hat{\Gamma}} x_t\right)^{-1} \sum_{t=1}^{T} x_t' M_{\hat{\Gamma}} y_t$$

$$\left[\frac{1}{T}\sum_{t=1}^{T}(y_t - x_t\hat{\beta})(y_t - x_t\hat{\beta})'\right]\hat{\Gamma} = \hat{\Gamma} V$$

其中，$\hat{\Gamma}$和 V 分别对应协方差阵$\hat{\Omega} = \frac{1}{T}\sum_{t=1}^{T}(y_t - x_t\hat{\beta})(y_t - x_t\hat{\beta})'$的特征向量和特征根矩阵。

Bai（2009）的分析表明[⑨]，上述的方程组会收敛到如下目标函数的解：

$$\min\left[\frac{1}{nT}\sum_{t=1}^{T}(y_t - x_t\hat{\beta} - \hat{\Gamma}\hat{v}_t)'(y_t - x_t\hat{\beta} - \hat{\Gamma}\hat{v}_t)\right] = \min\left[\frac{1}{nT}\sum_{t=1}^{T}\hat{\eta}_t'\hat{\eta}_t\right] \qquad (22)$$

由x_t和Γ的外生性可知上述的非线性最优化的解必定为真实参数的一致估计（详细的证明可见 Bai（2009））；可见，即使解释变量与共同因子存在相关，此时迭代的 RGLS 估计量仍然可以获得参数的一致估计。

5　小样本表现

为了简化分析，我们只考虑 $K = 1$、含有个体固定效应的 Panel 模型。

DGP 设定如下：

$$\begin{cases} y_{it} = \alpha_i + x_{it}\beta + \varepsilon_{it} \;,\; i = 1,2,\cdots,n; t = 1,2,\cdots,T \\ \varepsilon_{it} = \gamma_i' v_t + \eta_{it} \;,\; x_{it} = c_0 + c_1\gamma_0' v_t + \rho x_{i,t-1} + u_{it} - \theta u_{i,t-1} \\ \begin{bmatrix} v_t \\ \eta_t \\ u_t \end{bmatrix} \sim i.i.d.N\left\{0, \begin{bmatrix} \Pi & 0 & 0 \\ 0 & \Sigma_1 & 0 \\ 0 & 0 & \Sigma_2 \end{bmatrix}\right\} \end{cases}$$

其中，$\alpha_i \sim i.i.d.U[1,3]$，$\beta = \begin{cases} 0 & , H_0 \\ 0.1 & , H_1 \end{cases}$；$\gamma_{ij} \sim i.i.dU(0,2)$，$j = 1,\cdots,r$。

共同因子的结构我们考虑如下单一共同因子和多元（3 个）共同因子两种情况。

单一共同因子（r=1）：$\Pi = 1$。

多元共同因子（r=3）：$\Pi = diag(\pi_1,\pi_2,\pi_3)$，$\pi_j \sim i.i.d.U(0.5,1.5)$，$j = 1,\cdots,3$。

弱相关矩阵 Σ_1 和 Σ_2 的设定则参考 Chang（2002）的一般化设定形式：

$\Sigma_1 = P\Lambda_1 P'$ ，$\Sigma_2 = P\Lambda_2 P'$ ，$P = A(A'A)^{-1/2}$ ，$A = (a_{ij})$ ，$a_{ij} \sim i.i.d.U(0,1)$ ；$\Lambda_1 = diag(\lambda_{11},\lambda_{12},\cdots,\lambda_{1n})$ ，$\Lambda_2 = diag(\lambda_{21},\lambda_{22},\cdots,\lambda_{2n})$ ，$\lambda_{1i} \sim i.i.d.U(0,2)$ ，$\lambda_{2i} \sim i.i.d.U(0.5,1.5)$，$\Lambda_1$ 和 Λ_2 相互独立。

解释变量 X 的自相关结构设定为：$c_0 = 1$ ，$\rho = 0.5$ ，$\theta = -0.5$；$c_1 = \{0,1\}$分别对应 X 是否包含共同因子，即是否存在内生性，γ_0为所有元素为 1 的矩阵。

---

⑨ 本文的结果与 Bai（2009）对应的非线性最优化目标函数是等价的。其中，本文从协方差求逆的思路导出消除共同因子载荷系数的计算过程，而 Bai（2009）则是消除共同因子的计算过程。

样本设定为 $n,T=\{20,30,50,100\}$，在选取共同因子时最大可能的共同因子个数设为 $r_m=6$；计算迭代 RGLS 估计时设定迭代的最高次数为 6[⑩]；变量的初始值全部设定为 0，重复模拟次数 B=1000，外生参数一旦生成在重复模拟中不变。

为了比较各种估计方法的小样本性质，我们考察如下 3 个指标：

检验水平或功效（Size/Power）：$\frac{1}{B}\sum_{i=1}^{B}1(|t_i|>1.96)$；

平均绝对偏差（A-Bias）：$\frac{1}{B}\sum_{i=1}^{B}|\hat{\beta}_i-\beta|$；

均方误差（MSE）：$\frac{1}{B}\sum_{i=1}^{B}|\hat{\beta}_i-\beta|^2$；

其中，检验的 Size 或 Power 反映不同估计方法下 t 统计量的扭曲程度，平均绝对偏差和 MSE 反映不同估计方法下估计量的偏差及其集中程度，事实上 MSE 也是反映不同估计量相对样本有效性的等价指标。

表 1 至 3 分别给出了上述 3 个指标的数值模拟结果，由于各种情况下的模拟结果非常接近，为了节省篇幅，本文只报告了多元共同因子设定下的数值模拟结果，平均绝对偏差和 MSE 只报告$\beta=0$的情形。在各种情况下 Bai 和 Ng（2002）的 IC_1准则表现良好，正确识别出共同因子的概率基本处于 0.4 至 1 之间，并且随着样本 n、T 的增大迅速上升，正确识别出存在共同因子的概率基本为 1；在计算协方差阵逆矩阵时使用 MP 简化和未使用简化的模拟结果非常接近，因此对 RGLS 估计我们只报告使用 MP 简化计算的模拟结果；模拟的结果表明迭代的 RGLS 估计收敛的速度非常快[11]。

表 1 给出了各种估计方法下 t 统计量的检验 Size 和 Power。如表所示，对于 X 外生的情形，OLS-t 和 GLS-t 检验的 Size 存在严重的扭曲，ROLS-t 检验的 Size 非常接近 0.05 的名义水平，本文提出的 RGLS-t 检验的 Size 也很接近 0.05 的名义水平；GLS-t 和 RGLS-t 检验的 Power 非常高，即使不进行 Size 调整，RGLS-t 检验的 Power 也要稍微高于 GLS-t 检验的 Power，OLS-t 和 ROLS-t 检验的 Power 则相对要低很多。对于 X 内生的情形，普通 OLS-t、GLS-t 和 ROLS-t 都出现严重的扭曲，一步的 RGLS-t 同样也存在严重的扭曲，但迭代的 RGLS-t 则表现非常良好，检验的水平非常接近名义水平 0.05。总的来看，就检验的 Size 而言，ROLS-t 在 X 外生时要稍微优于 RGLS-t 检验，两者都明显优于 OLS-t 和 GLS-t，但在 X 内生时，迭代的 RGLS-t 要远远优于其它几种检验[12]；就检验的 Power 而言，RGLS-t 要稍微优于 GLS-t 检验，而且它们都明显优于 OLS-t 和 ROLS-t 检验。

表 1　Size/Power

			C1=0					C1=1				
	n	T	OLS	ROLS	GLS	RGLS	IRGLS	OLS	ROLS	GLS	RGLS	IRGLS
Size	20	20	69	37	832	102	102	1000	1000	1000	893	132
		30	70	33	610	74	80	1000	1000	1000	882	84

[⑩] 估计参数的变动的平方小于 0.00001 则认为是收敛。

[11] 对于 X 外生的设定，绝大多数情形下迭代的次数不超过 3 次就已经收敛；而对于 X 内生的设定，绝大多数情形下迭代的次数不超过 6 次就已经收敛

[12] 事实上，其它的几种估计量在此时都不是一致估计。

		50	74	57	360	70	78	1000	1000	1000	998	117
		100	46	39	134	46	50	1000	1000	1000	978	81
	30	20	65	31	312	94	88	1000	1000	1000	971	112
		30	74	47	900	76	74	1000	1000	1000	971	96
		50	75	52	537	63	65	1000	1000	1000	960	78
		100	53	39	198	60	61	1000	1000	1000	998	79
	50	20	122	39	44	90	79	1000	1000	596	986	84
		30	103	54	197	53	50	1000	1000	1000	914	61
		50	89	48	912	61	63	1000	1000	1000	997	76
		100	68	54	433	61	59	1000	1000	1000	999	64
	100	20	202	51	233	84	70	1000	1000	627	983	91
		30	134	42	121	70	64	1000	1000	170	998	85
		50	106	46	71	58	61	1000	1000	1000	1000	74
		100	86	46	945	47	49	1000	1000	1000	1000	73
	20	20	333	286	935	756	760	1000	1000	1000	999	804
		30	433	372	941	864	861	1000	1000	1000	1000	888
		50	634	598	1000	999	999	1000	1000	1000	1000	999
		100	975	970	1000	1000	1000	1000	1000	1000	1000	1000
	30	20	387	317	895	946	949	1000	1000	1000	1000	966
		30	534	442	989	991	993	1000	1000	1000	1000	995
		50	865	840	1000	1000	1000	1000	1000	1000	1000	1000
Power		100	996	994	1000	1000	1000	1000	1000	1000	1000	1000
	50	20	558	389	603	976	979	1000	1000	969	1000	987
		30	874	792	997	1000	1000	1000	1000	1000	1000	1000
		50	955	919	1000	1000	1000	1000	1000	1000	1000	1000
		100	998	997	1000	1000	1000	1000	1000	1000	1000	1000
	100	20	828	617	385	1000	1000	1000	1000	163	1000	1000
		30	921	808	819	1000	1000	1000	1000	829	1000	1000
		50	993	985	1000	1000	1000	1000	1000	1000	1000	1000
		100	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000

注：RGLS 对应使用 MP 简化的 RGLS 估计量，IRGLS 对应迭代至收敛的 MP 简化的 RGLS 估计；GLS 估计在 T 小于 n 时使用广义逆求逆。表中报告的结果为 Size/Power 乘以 1000 得到。

表 2　A-Bias

		C1=0				C1=1			
n	T	OLS	GLS	RGLS	IRGLS	OLS	GLS	RGLS	IRGLS
20	20	5.92	5.65	3.44	3.42	42.76	40.19	14.64	3.36

	30	5.14	3.89	2.71	2.70	48.43	34.18	12.01	2.65
	50	3.66	1.96	1.69	1.66	44.08	20.80	12.66	1.59
	100	1.94	0.86	1.07	1.07	35.85	10.89	6.27	1.05
30	20	5.13	3.47	2.41	2.37	44.14	26.73	15.08	2.35
	30	4.30	4.17	1.94	1.94	50.05	47.69	11.96	1.94
	50	2.64	1.88	1.39	1.40	37.99	22.92	7.19	1.39
	100	1.74	0.94	0.92	0.93	38.50	13.88	7.20	0.91
50	20	4.92	3.47	2.16	2.09	46.91	10.48	14.23	2.04
	30	2.87	1.85	1.43	1.43	36.61	17.82	7.91	1.42
	50	2.31	2.27	1.11	1.11	40.78	39.73	7.66	1.09
	100	1.59	0.97	0.78	0.78	44.01	20.69	5.74	0.77
100	20	3.66	8.80	1.40	1.35	41.96	12.33	11.09	1.35
	30	2.75	3.27	1.08	1.05	44.67	3.29	11.95	1.04
	50	1.83	0.99	0.78	0.77	44.32	13.45	8.78	0.75
	100	1.18	1.17	0.55	0.55	42.93	42.29	7.55	0.53

注：表中报告的数值为平均绝对偏差乘以 100 得到。

表 3　MSE

		C1=0				C1=1			
n	T	OLS	GLS	RGLS	IRGLS	OLS	GLS	RGLS	IRGLS
20	20	5.48	4.98	1.88	1.88	188.26	166.61	24.57	1.83
	30	4.09	2.37	1.14	1.15	238.40	119.86	16.52	1.12
	50	2.17	0.61	0.45	0.43	196.18	44.07	17.01	0.39
	100	0.60	0.12	0.17	0.18	129.47	12.13	4.28	0.17
30	20	4.16	1.91	0.93	0.90	199.48	74.81	25.20	0.88
	30	2.92	2.74	0.59	0.58	254.32	230.99	15.75	0.58
	50	1.10	0.56	0.30	0.30	145.92	53.40	5.74	0.29
	100	0.47	0.14	0.14	0.14	148.99	19.48	5.49	0.13
50	20	3.81	1.90	0.75	0.70	224.95	14.13	22.08	0.67
	30	1.32	0.55	0.32	0.32	137.30	33.49	7.20	0.32
	50	0.85	0.82	0.19	0.19	167.87	159.34	6.28	0.19
	100	0.41	0.15	0.09	0.09	194.58	43.11	3.49	0.09
100	20	2.13	11.98	0.32	0.29	181.32	19.14	13.74	0.29
	30	1.22	1.65	0.18	0.18	202.34	1.71	15.22	0.17
	50	0.54	0.15	0.10	0.10	198.06	18.64	8.12	0.09
	100	0.23	0.22	0.05	0.05	184.98	179.56	5.90	0.04

注：表中报告的数值为 MSE 乘以 1000 得到。

表 2 和表 3 分别给出了各种估计方法下估计量的平均绝对偏差和均方误差。如表所示，对于 X 外生的情形，即使此时 OLS 和 GLS 估计量都是一致估计，与 RGLS 估计相比，两者的估计偏差都比较大，在绝大多数的情况下 OLS 的估计偏差最大，但 GLS 估计在 n 远大于 T 的情况下会出现严重的偏差；当 T 远大于 n 时，GLS 估计可以得到最小的偏差，此时 RGLS 的估计偏差要比它略大一点，但 OLS 的估计偏差仍然要大很多。而对于 X 内生的情形，OLS 和 GLS 估计量出现非常严重的偏差，这一偏差是由估计量的不一致所引起；一步的 RGLS 估计也存在严重的偏差，但仍远小于 OLS 和 GLS；只有迭代的 RGLS 具有非常小的偏差。总的来看，就估计量的平均绝对偏差和 MSE 而言，RGLS 估计，特别是迭代的 RGLS 估计，要明显优于 OLS 和 GLS 估计。

综合上述结果可知：检验水平或功效的结果表明，对于 X 外生的情形，ROLS 和 RGLS 估计量都非常接近对应的名义水平，其表现明显优于 OLS 和 FGLS 估计量，OLS 和 FGLS 估计量的 t 统计量存在严重的扭曲；而对于 X 内生的情形，只有迭代的 RGLS 估计仍然具有良好的表现，其它几个估计量都出现严重的扭曲。估计偏差和 MSE 的结果表明，RGLS 估计量的小样本偏差要远小于 OLS 和 FGLS 估计量，仅当 X 外生且 T 远大于 n 的时候，FGLS 估计量的偏差才略低于 RGLS 估计量。可见，RGLS 估计量（特别是迭代的 RGLS）的小样本表现要明显优于 FGLS 估计量与（R）OLS 估计量。

6 结论

误差近似因子结构设定下系数的估计与检验是计量经济学关注的一个重要问题。本文关注的是近似因子结构设定下 GLS 估计量小样本性质的改进方法，我们提出了一种稳健的 GLS 估计（RGLS），克服了普通 FGLS 估计量由于求逆过程带来的小样本偏差过大的问题。本文提出的 RGLS 估计可以简单地分解成两部分的工作，其一是对数据进行部分的 GLS 估计以消除共同因子的影响，其功能接近于对共同因子进行正交变换；其二是对变换后的数据进行稳健 OLS 估计。

我们证明了在相当弱的假定条件下 RGLS 估计是一致估计，而且比 ROLS 估计量渐近有效，其对应的 t 统计量序贯收敛于标准正态分布。数值模拟的结果表明，在绝大多数的情况下 FRGLS 估计量的估计偏差和 MSE 要小于 OLS 和 FGLS 估计量；RGLS-t 统计量的检验 Size 非常接近 0.05 的名义水平，远远优于 OLS-t 和 GLS-t，其检验的 Power 更是在 4 个估计量中居首；特别是对于 X 内生的情形，迭代的 RGLS 估计仍然具有非常良好的小样本表现，但其它几个估计量都存在严重的扭曲。总的来看，RGLS 估计（特别是迭代的 RGLS）具有非常良好的小样本表现，明显优于 OLS 和 FGLS 估计。

参考文献

Anselin L.2001.Spatial econometrics, in: B. Baltagi (ed.), A Companion to Theoretical Econometrics. Blackwell, Oxford,310-330.

Bai J, Ng S.2004.A panic attack on unit roots and cointegration.Econometrica,72(4): 1127-1177.

Bai J,Ng S.2002.Determining the number of factors in approximate factor models. Econometrica,70(1): 191-221.

Bai J,Ng S.2006.Evaluating latent and observed factors in macroeconomics and finance. Journal of Econometrics, 131(1): 507-537.

Bai J.2009. Panel data models with interactive fixed effects. Econometrica,77(4):1229-1279.

Beck N, Katz J N. 1995.What to do (and not to do) with time-series cross-section data.The American Political Science Review, 89(3): 634-647.

Breitung J,Das S. Panel unit root tests under cross-sectional dependence. Statistica Neerlandica,59(4):414-433.

Breitung J, Das S.2008. Testing for unit roots in panels with a factor structure. Econometric Theory, 24(1): 88-108.

Chamberlain G, Rothschild M.1983.Arbitrage,factor structure, and mean-variance analysis on large asset markets. Econometrica, 51(5):1281-1304.

Chang Y, 2002.Nonlinear IV unit root tests in panels with cross-sectional dependency, Journal of Econometrics.110(2): 261-292.

Chen L.2012. Identifying observed factors in approximate factor models: estimation and hypothesis testing. working paper.

Chudik A, Pesaran M H, Tosetti E.2009. Weak and strong cross section dependence and estimation of large panels. working paper.

Coakley J, Fuertes A, Smith R.2002. A principal components approach to cross-section dependence in panels. Working paper.

Kapetanios G, Pesaran M H.2007.Small sample properties of cross section augmented estimators for panel data models with residual multi-factor structures, in the refinement of econometric estimation and test procedures: finite sample and asymptotic analysis, Garry Phillips and Elias Tzavalis (ed.), Cambridge University Press, Cambridge.

Moon H R, Perron B.2006.Seemingly unrelated regressions.Working paper.

Moon H R, Perron B.2004.Testing for a unit root in panels with dynamic factors.Journal of Econometrics, 122(1): 81-126.

Pesaran M H,2003. Estimation and inference in large heterogenous panels with cross section dependence. Working paper.

Peasaran M H.2006.Estimation and inference in large heterogeneous panels with a multifactor error structure. Econometrica, 74(4):967-1012.

Pesaran M H, Tosetti E.2007. Large panels with common factors and spatial correlations. Working paper.

Phillips P C B, Moon H R.1999. Linear regression limit theory for nonstationary panel data. Econometrica, 67(5): 1057-1111.

Sarafidis V,Wansbeek T.2012. Cross-sectional dependence in panel data analysis. Econometric Reviews, 31(5):483-531.

Stock J H, Watson M W.2002.Macroeconomic forecasting using diffusion indexes. Journal of Business & Economic Statistics, 20(2): 147-162.

Zellner A.1962.An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. Journal of the American Statistical Association, 57(298):348-368.

附　录

此处我们报告了各种参数设定下，RGLS 估计与 RPC、IPC 和 CCE 估计量的小样本比较，由于后面三个估计量的提出都主要针对因子结构的情形，对于近似因子结构下 t 统计量的计算并没有明确的说明，我们这里不报告它们的检验 Size 和 Power。

RPC 估计量的计算如下：

$$\hat{\beta}_{RPC} = \left(x_* M_{\hat{F}} x_*'\right)^{-1} x_* M_{\hat{F}} y_*'$$

其中，对于任意的 z，$z_* = z \cdot M_T$，$M_T = I_T - i_T i_T'/T$；$M_{\hat{F}} = I_T - \hat{F}(\hat{F}'\hat{F})^{-1}\hat{F}'$，$\hat{F}$为协方差阵$\hat{\Omega} = \frac{1}{n}(y_* - \hat{\beta}_{LS} x_*)'(y_* - \hat{\beta}_{LS} x_*)$的最大的 r 个特征根对应的特征向量。

IPC 估计量对应迭代的 RPC 估计量。

CCE 估计量的计算如下：

$$\hat{\beta}_{CCE} = \left(xM_{\bar{z}}x'\right)^{-1} xM_{\bar{z}}y'$$

其中，$\bar{z} = (\bar{z}_1', \cdots, \bar{z}_T')'$，$\bar{z}_t = \{\bar{y}_t, \bar{x}_t, 1\}$，$M_{\bar{z}} = I_T - \bar{z}(\bar{z}'\bar{z})^{-1}\bar{z}'$。

由附表 1 和附表 2 的结果可知，对于 X 外生的情形，RPC、IPC、CCE 和 RGLS 估计量都表现非常良好，其对应的偏差都很小，其中 CCE 略差。对于 X 内生的情形，RPC 和一步的 RGLS 不再一致，具有严重的偏差，其中，RPC 尤其严重；相对的，IPC 和 CCE 和迭代的 RGLS 的表现仍然非常良好，具有很小的偏差，其中的，CCE 略差。

附表 1　A-Bias

		C1=0					C1=1				
n	T	RPC	IPC	CCE	RGLS	IRGLS	RPC	IPC	CCE	RGLS	IRGLS
20	20	3.46	3.42	3.71	3.44	3.42	20.34	3.41	3.68	14.64	3.36
	30	2.70	2.67	2.94	2.71	2.70	25.23	2.69	2.90	12.01	2.65
	50	1.75	1.70	2.03	1.69	1.66	18.91	1.64	2.03	12.66	1.59
	100	1.09	1.09	1.20	1.07	1.07	11.59	1.05	1.20	6.27	1.05
30	20	2.40	2.40	3.05	2.41	2.37	20.26	2.40	3.19	15.08	2.35
	30	1.96	1.94	2.16	1.94	1.94	23.60	1.97	2.19	11.96	1.94
	50	1.39	1.39	1.65	1.39	1.40	15.71	1.39	1.65	7.19	1.39
	100	0.93	0.93	1.08	0.92	0.93	15.21	0.92	1.07	7.20	0.91
50	20	2.08	2.05	2.47	2.16	2.09	19.91	2.02	2.50	14.23	2.04
	30	1.43	1.44	1.54	1.43	1.43	11.47	1.42	1.55	7.91	1.42
	50	1.11	1.11	1.25	1.11	1.11	16.71	1.09	1.24	7.66	1.09
	100	0.78	0.78	0.95	0.78	0.78	17.98	0.77	0.95	5.74	0.77
100	20	1.34	1.34	1.59	1.40	1.35	14.82	1.34	1.61	11.09	1.35
	30	1.03	1.03	1.31	1.08	1.05	16.68	1.03	1.31	11.95	1.04
	50	0.77	0.77	0.94	0.78	0.77	17.24	0.76	0.94	8.78	0.75
	100	0.55	0.55	0.66	0.55	0.55	16.21	0.53	0.65	7.55	0.53

附表 2　MSE

		C1=0					C1=1				
n	T	RPC	IPC	CCE	RGLS	IRGLS	RPC	IPC	CCE	RGLS	IRGLS
20	20	1.90	1.88	2.19	1.88	1.88	46.23	1.88	2.12	24.57	1.83
	30	1.16	1.14	1.38	1.14	1.15	68.31	1.16	1.35	16.52	1.12
	50	0.48	0.45	0.65	0.45	0.43	37.46	0.41	0.65	17.01	0.39
	100	0.18	0.18	0.23	0.17	0.18	14.00	0.17	0.22	4.28	0.17
30	20	0.93	0.92	1.48	0.93	0.90	45.05	0.92	1.59	25.20	0.88
	30	0.59	0.58	0.76	0.59	0.58	60.06	0.61	0.77	15.75	0.58

	50	0.30	0.30	0.44	0.30	0.30	26.17	0.29	0.44	5.74	0.29
	100	0.14	0.14	0.18	0.14	0.14	23.86	0.14	0.18	5.49	0.13
50	20	0.70	0.69	0.96	0.75	0.70	43.85	0.67	0.98	22.08	0.67
	30	0.33	0.33	0.38	0.32	0.32	15.09	0.32	0.38	7.20	0.32
	50	0.19	0.19	0.25	0.19	0.19	29.51	0.19	0.24	6.28	0.19
	100	0.09	0.09	0.14	0.09	0.09	33.36	0.09	0.14	3.49	0.09
100	20	0.29	0.29	0.40	0.32	0.29	25.76	0.29	0.41	13.74	0.29
	30	0.17	0.17	0.27	0.18	0.18	30.08	0.17	0.27	15.22	0.17
	50	0.10	0.10	0.14	0.10	0.10	31.52	0.09	0.14	8.12	0.09
	100	0.05	0.05	0.07	0.05	0.05	27.02	0.04	0.07	5.90	0.04