arXiv:2603.02509v1 [stat.ME] 3 Mar 2026

含滞后与反馈协变量的纵向数据半分区广义矩方法

Niloofar Ramezani[1] 和JeffreyR.Wilson[2]

[1]生物统计学系，弗吉尼亚联邦大学[2]经济学系，亚利桑那州立大学

[1]ramezanin2@vcu.edu [2]jeffrey.wilson@asu.edu

摘要

我们提出一种半分区广义矩方法（GMM）框架，用于在边际建模范式下分析含时变协变量的纵向数据。该方法旨在解决聚合方法与完全分区GMM模型的局限性。聚合方法因假定效应恒定而掩盖时间动态，而完全分区方法虽能提供时间特异性，却以模型复杂度增加和稳定性下降为代价——尤其是在中等样本量或较深滞后结构的情形下。本方法通过分别估计同期系数，并将滞后矩条件分组为结构化集合，从而区分即时效应与滞后效应，同时保留滞后特异效应的灵活性。由此得到的模型在统计上既有效率又可解释，能够捕捉关键的时间变化，同时减轻与完全分区相关的方差膨胀和收敛难题。该框架能够容纳反馈，支持连续结局与二元结局，并利用Broyden–Fletcher–Goldfarb–Shanno（BFGS）算法进行可靠优化。通过模拟，我们表明：当分组滞后结构近似于潜在滞后模式时，半分区广义矩估计相对于完全分区模型可实现相应的覆盖率，并具有具有竞争力的效率。在膝骨关节炎与青少年肥胖临床数据集上的应用证实，该方法能够恢复一致且可解释的效应，并避免与细粒度分区相关的不稳定性。

关键词：广义矩方法；纵向数据分析；时变协变量；半分区模型；反馈效应；边际模型。

1 引言

纵向研究通过随时间反复测量个体，对于理解医学、行为和流行病学研究中的动态过程至关重要。当研究对象接受多次测量，且研究重点在于随时间变化的模式时，纵向数据便由此形成，其主要特征是同一研究对象的重复测量之间存在依赖关系 [1]。然而，此类数据的分析复杂性源于同一研究对象的重复测量之间的相关性，以及协变量和结局随时间变化的特性。传统的横断面模型无法考虑这些依赖关系，可能导致有偏推断和低效估计 [2]。

边际模型——例如基于广义估计方程（GEE）或广义矩方法（GMM）的模型——已成为在纵向设定中估计总体平均效应的基础工具。当总体平均推断是主要关注点 [3, 4] ，或基于当前协变量的期望响应值是关键关注量时[5]，这些方法都适用。这类模型侧重于对响应的边际均值建模，而不要求显式建模个体内变异；因此，当研究目标聚焦于总体趋势而非个体轨迹时，这类模型尤具吸引力 [6, 7]。本文明确采用边际均值建模框架：将时间 t 的期望结局建模为当前和滞后协变量的函数，而不以既往结局或个体特异随机效应为条件。这不同于条件混合模型或自回归模型，后者会根据结局的历史信息来描述其演变。

边际建模中一个持续存在且研究不足的挑战，是如何恰当处理时变协变量。这些协变量可能表现出结局的反馈，其关联强度可能随时间变化，或同时产生即时效应与延迟效应。若处理不当，时变协变量

会使标准估计方程失效，并使边际效应估计产生偏倚。Lai and Small [8] 基于此类协变量与既往及未来响应的关系，将其开创性地分为三类（I、II和III），为利用矩条件进行更为原则性的估计奠定了基础。尽管某些条件时间序列模型（例如混合数据抽样（MIDAS））也会对滞后协变量进行聚合，但这些方法针对的是条件期望与预测，而我们的关注点仍是重复测量数据的边际均值结构。

基于这一分类，广义矩估计框架中已形成两种主要的建模策略。聚合GMM方法将所有有效矩条件结合起来，为每个协变量估计一个系数。这种方法在计算上效率较高，但可能掩盖有意义的时间变化。与之相反，完全分区GMM模型将协变量的每个滞后期分别处理，从而提供最灵活的时间分辨率。然而，这类模型会引入多个额外参数，尤其是在随访时长较长时，并且常会出现标准误膨胀、奇异性问题或收敛失败——样本量不足时尤为如此。

针对这些局限，我们提出一种半分区GMM模型，在灵活性与效率之间进行策略性权衡。我们的方法将同期（滞后期为0）效应的估计与滞后效应的估计分开，并允许将滞后矩条件归入具有可解释性的区块，例如短期与长期滞后，或将其汇总为一个平均效应。这种结构既保留了矩条件框架的丰富性，又减轻了估计负担，并避免不必要的模型复杂度。

与完全分区方法 [9], 为每个滞后期估计一个系数不同，我们的模型通过限制待估参数的数量，缓解了方差膨胀，并能在完全分区变得不稳定的情形下提高计算稳定性。同时，它超越了聚合方法 [8, 10] ，通过明确建模时间依赖性的结构，而非将其折叠处理。此外，半分区模型能够容纳反馈过程（III型协变量），并支持连续型和二元结局。估计采用 Broyden–Fletcher–Goldfarb–Shanno（BFGS）算法实现；与以往 GMM 应用中使用的无导数方法相比，这种拟牛顿方法具有更好的数值表现。

半分区设定并非旨在作为建模的终点，而是灵活性与稳定性之间的一种务实折中。完全分区GMM仍可作为探索性诊断工具，用于识别滞后特异性模式，并据此构建更简约的半分区结构。

本文其余部分组织如下。第2节回顾GMM框架以及现有的聚合方法与完全分区方法。第3节详细介绍半分区GMM模型，包括其数学表述与估计策略。第4节提供模拟研究，以评估模型在不同滞后结构与反馈动态下的表现。第5节给出两项实际应用：一项来自膝骨关节炎研究的连续型结局，以及一项来自青少年肥胖纵向研究的二元结局。第6节讨论优势、局限性及未来研究方向，第7节作出结论。

2 广义矩估计法

广义矩估计最初由Hansen [11] 在计量经济学背景下提出，提供了一种强大而灵活的估计框架，在纵向数据情形中尤为有用 [12]。广义矩估计依赖于矩条件——数据与参数的函数的期望值——这些条件在真实数据生成过程下成立，而无需完整设定底层分布。这种半参数方法在纵向分析中具有优势，因为分布假设往往难以验证，而时变协变量会使基于似然的方法复杂化。令 $\boldsymbol{\beta}_0$ 表示真实参数向量，并令$\mathbf{x}_{it}$ 表示研究对象 i 在时刻 t的观测协变量，其中 $i = 1,\ \ldots\ldots,\ n$ 以及 $t = 1,\ \ldots\ldots,\ T$。在本文所考虑的边际建模情境中，这些矩条件刻画结局与协变量之间的总体平均关系，从而将我们的设定与规定结局完整动态演变及分布假设的条件时间序列模型区分开来。

广义矩估计程序从一组总体矩条件开始：

$$E[f(\mathbf{x}_{it}, \boldsymbol{\beta}_0)] = 0,$$

其中 $f(.)$是对应于矩条件的函数向量。样本对应量通过对受试者取平均而形成：

$$\frac{1}{n}\sum_{i=1}^{n} f(\mathbf{x}_{it}, \boldsymbol{\beta}).$$

GMM估计量通过最小化样本矩向量的二次型得到：

$$Q(\boldsymbol{\beta}) = \left(\frac{1}{n}\sum_{i=1}^{n} f(\mathbf{x}_{it}, \boldsymbol{\beta})\right)^{\top} \boldsymbol{W} \left(\frac{1}{n}\sum_{i=1}^{n} f(\mathbf{x}_{it}, \boldsymbol{\beta})\right)$$

其中 $\boldsymbol{W}$ 是一个半正定加权矩阵。在最优两步GMM中， $\boldsymbol{W}$ 取为 $\boldsymbol{S}^{-1}$的一致估计值，其中 $\boldsymbol{S}$ 表示样本矩条件˜[11]的长期协方差矩阵。 $Q(\boldsymbol{\beta})$的极小值点记为 $\boldsymbol{\beta}$ ，它是 $\boldsymbol{\beta}_0$的GMM估计值：

$$\hat{\boldsymbol{\beta}} = \arg\min Q(\boldsymbol{\beta}), \tag{1}$$

ˆ其中，arg min表示使函数达到最小值的 $\boldsymbol{\beta}$ 。在线性模型中， $\boldsymbol{\beta}$ 可能有闭式解；在非线性模型中，则需要进行数值优化。

2.1 聚合GMM

Lai and Small [8] 将GMM框架扩展到纵向边际回归模型，以处理时变协变量。他们的方法核心是根据协变量与结局过程在时间上的关联方向和性质，将协变量分为三类——I型、II型和III型。

I型协变量与响应的过去值和未来值均不相关。这类协变量通常是外生的，不受结局历史影响，因此在边际模型中最容易处理。例如，年龄、日历时间或交叉试验中的随机治疗分配。涉及I型协变量的矩条件在所有时间点均有效，因此能提供最丰富的估计方程集合。

II型协变量可能与既往结局相关，但假设其与未来响应条件独立。这种结构常见于自回归情境，或协变量反映受先前健康状态影响、但不受未来结局影响的受试者层面特征的纵向设计中。对于II型协变量，矩条件对当前和未来响应有效，但对过去的响应无效，因此得到的估计方程组有所缩减，但仍具有信息价值。

III型协变量会引入响应对协变量过程本身的反馈。在这种情况下，时间 t 的协变量可能是先前结局的函数，从而违反外生性假设。这种动态关系在行为或药物依从性研究中很常见，因为响应会影响未来的暴露。对于III型协变量，只有同期矩条件有效，这大大限制了可用估计方程的数量。

聚合GMM方法将给定协变量的所有有效矩条件合并为一个估计方程。这样，每个协变量只对应一个系数，实际上假设其效应在不同时点及所有有效滞后项上均保持恒定。这种合并策略简化了估计，并且通常能提高效率。然而，它可能掩盖重要的滞后特异性动态，并降低可解释性。当协变量效应存在异质性，或表现出延迟或反馈时，聚合处理可能不太理想。

2.2 完全分区GMM

为解决聚合方法的局限性，Irimata等人 [9] 提出了完全分区GMM框架，该框架针对每个时间依赖协变量的每个滞后期分别估计一个回归系数。该方法不假设效应随时间保持恒定，而是通过将给定滞后期观测到的每个协变量值与一个独立的回归参数相关联，显式地对滞后结构建模。这样便能详细而灵活地刻画协变量效应如何随时间变化，并深入了解其与响应之间的即时关联和延迟关联。

完全分区GMM模型通过将协变量矩阵重构为下三角形式来运行，其中每一行对应某个时间点的响应，每一列对应协变量的一个滞后值。对于在 $s = 1$、.....、t 测量的协变量，该设计矩阵可估计时间特异性效应，记为 $\beta^{[0]}$、$\beta^{[1]}$、$\beta^{[2]}$、.....、$\beta^{[T-1]}$，分别对应协变量的同期值和滞后值。在这一框架中，模型的系统性成分变为这些滞后特异性贡献的加权和，并相应地对矩条件进行分区，以反映所假设的时间结构。由于不对不同时期的数据进行合并，完全分区GMM能够揭示协变量效应中可能被聚合模型掩盖的滞后特异性变异。这在临床或行为研究中尤其有用，因为干预效应、暴露反应或依从行为可能随时间变化，也可能取决于近期历史。

然而，灵活性的提升也带来了显著的计算和统计成本。分区会使参数数量随时间点和协变量数量成比例增加，从而大幅提高估计问题的维度。在实践中，这可能导致数值不稳定、无法收敛或估计值不可靠，尤其是在研究对象数量并未显著多于特定滞后参数数量时。在二元结局模型中，这些问题会进一步加剧，因为稀疏事件计数会进一步降低每个滞后期的有效样本量。因此，尽管完全分区GMM有助于探索时间异质性，但它最适用于时间点相对较少或样本量非常大的研究。在典型的应用场景中，随访时间适中且检验效能有限，完全分区方法可能无法实施，或导致过拟合，因此有必要开发更简约的替代方法，例如半分区模型。在实践中，完全分区GMM也可作为探索性诊断工具，用于识别滞后特异性模式，并据此进一步制定简约的半分区设定。接下来介绍的半分区方法旨在保留滞后特异性建模的可解释性，同时避免完全分区设定固有的不稳定性和维度挑战。

3 半分区广义矩估计

我们提出了一种半分区GMM框架，融合了聚合GMM和完全分区GMM模型各自的优势。聚合GMM假设协变量效应在时间上恒定，而完全分区GMM模型则对每个滞后期分别建模；半分区方法在两者之间取得平衡，分别估计协变量的即时（同期）效应和滞后效应，并将后者归入一个或多个预定义结构中。这些预定义结构能够以受控方式表示滞后效应，例如将其合并为平均效应，或按相对时间邻近性分类（如近滞后项与远滞后项）。该方法尤其适用于随访期数适中的纵向数据集，因为过度参数化可能妨碍完全分区模型的收敛和稳定性。此外，该方法通过允许结局影响未来的协变量结构来适应反馈动态。对滞后协变量进行分组，能够以简约的方式概括时间结构，在降低维度的同时，无需假设底层的滞后特异效应彼此相等。

我们采用分组滞后组织结构来降低矩条件的维度。该结构在保留滞后特异效应灵活性的同时，整合多个滞后期的信息。当分组与潜在时间模式相吻合时，效率提升最为显著；模拟结果表明，即使存在中度不匹配，模型表现仍然稳定。

完全分区模型。令 Y_{it} 表示研究对象 i 在时间 t 的结局，其中 $i = 1, \ldots, n$ and $t = 1, \ldots, T$，并令 $x_{ij(s)}$ 表示协变量 j 在时间 s 的取值，其中 $s = t, t-1, \ldots, 1$。对于每个协变量 j，完全分区模型同时纳入即时效应和滞后效应，如下所示：

$$g(\mu_{it}) = \beta_0 + \sum_{j=1}^{J} \left(\beta_j^{|0|} x_{ij(t)} + \sum_{k=1}^{t-1} \beta_j^{|k|} x_{ij(t-k)} \right), \tag{2}$$

其中，$\mu_{it} = E[Y_{it} \mid x_{ij(s)}]$, $g(\cdot)$ 是适当的连接函数，$\beta_j^{|0|}$ 是即时效应，$\beta_j^{|k|}$ 表示协变量 j 的滞后-k 效应。

一般半分区模型。为提高简约性并使估计更加稳定，可以对滞后效应进行分组，使同一区块内的所有滞后项共享一个系数。令 $\{G_{j1},\ldots,G_{jm_j}\}$ 表示协变量 j 的滞后项分区，其中 m_j 是第 j 个协变量分区中的元素个数。一般的分组滞后设定可写为：

$$g(\mu_{it}) = \beta_0 + \sum_{j=1}^{J}\left(\beta_j^{|0|}x_{ij(t)} + \sum_{g=1}^{m_j}\beta_{jg}\sum_{k\in G_{jg}}x_{ij(t-k)}\right), \tag{3}$$

其中，G_{jg} 是滞后索引的第 g组，β_{jg} 是所有滞后项 $k \in G_{jg}$的共同系数。

半分区模型的特殊情况。也可以采用更加灵活的半分区结构。例如，如果将"近"滞后项（例如，$k = 1$、.....、$t-r$）与"远"滞后项（例如，$k = t-r+1$、.....、$t-1$）分开分组，则模型变为：

$$g(\mu_{it}) = \beta_0 + \sum_{j=1}^{J}\left(\beta_j^{|0|}x_{ij(t)} + \beta_{j,\text{near}}\sum_{k=1}^{t-r}x_{ij(t-k)} + \beta_{j,\text{far}}\sum_{k=t-r+1}^{t-1}x_{ij(t-k)}\right).$$

作为一个具体情况，如果认为第一阶滞后具有不同的效应，而所有其余滞后项共享一个共同效应，则将第一阶滞后单独建模，并将其余滞后项归为一组，从而得到模型：

$$g(\mu_{it}) = \beta_0 + \sum_{j=1}^{J}\left(\beta_j^{|0|}x_{ij(t)} + \beta_j^{|1|}x_{ij(t-1)} + \beta_{j,\text{post-1}}\sum_{k=2}^{t-1}x_{ij(t-k)}\right). \tag{4}$$

最后一个特殊情况用于模拟研究，假设协变量 j 的所有滞后效应均由单个区块汇总：

$$g(\mu_{it}) = \beta_0 + \sum_{j=1}^{J}\left(\beta_j^{|0|}x_{ij(t)} + \beta_{j,\text{lag}}\sum_{k=1}^{t-1}x_{ij(t-k)}\right). \tag{5}$$

这种结构保留了检测即时效应的能力，同时使用较少的参数对滞后贡献进行建模。这些分组设定通过在每个滞后区块内施加等式约束来降低维度，在保持灵活的边际均值结构的同时，得到具有可解释性的滞后组效应。

在实际应用中，分组选择通常依据领域专业知识，或完全分区模型中观察到的探索性模式。这些模式有助于构建简约的半分区结构，而无须识别每个滞后特异性系数。例如，临床医生可能对治疗或暴露的时间变化特征有充分的先验认识，例如预期某种药物的作用主要是同期效应，或仅持续到接下来一至两次随访访视。在这类情况下，将滞后项分组成具有临床意义的区块，是在边际均值结构中体现这些知识的合理方法。如果缺乏这类信息，则可使用探索性的完全分区模型来识别滞后特异性系数中的模式（例如，幅度或符号相似的簇），并据此构建更简约的半分区设定。这种两步策略既保留了分组模型的可解释性和稳定性，又避免了估计大量参数，以及完全分区方法固有的方差膨胀。在大多数应用中，主要关注的参数是同期效应 $\beta^{|0|}$；即使滞后效应的分组结构与潜在时间模式存在一定偏差，其解释仍然稳健。

在此处考虑的边际设定中，分组及其相关滞后结构的设定没有条件模型（例如自回归模型）中那么关键。半分区方法旨在将边际协变量效应——在标准广义矩估计 [8] 中通常汇总所有滞后项——分解为可解释的组成部分，这些部分分别归因于特定的滞后项组，而无需估计每个单独的滞后特异性系数。

为确保矩条件有效，我们为每个协变量 j定义一个矩阵 M_j ，其中每个元素表示根据 Lai and Small [8]的分类框架，某个矩条件是否有效。该矩阵分为 T个部分： M_{j0}对应即时效应， M_{jk}对应滞后项 $k = 1$、.....、$T-1$,

每个部分对应一个特定的时间差 $t-s=k$。例如，在分组滞后设定下，如公式（5），与 M_{j1}、……、$M_{j(T-1)}$ 相关的矩条件会根据所选的分组结构进行组合。

连续结局的估计方程构造如下：

$$\frac{\partial \mu_{is}(\tilde{\boldsymbol{\beta}})}{\partial \beta_j^{[k]}}\left[Y_{it}-\mu_{it}(\tilde{\boldsymbol{\beta}})\right]=x_{ij(s-k)}\left[Y_{it}-\mu_{it}(\tilde{\boldsymbol{\beta}})\right],$$

其中，$\tilde{\boldsymbol{\beta}}$ 是 $\boldsymbol{\beta}$的起点。

对于二元结局，我们使用逻辑斯蒂均值函数

$$\mu_{it}(\boldsymbol{\beta})=\frac{\exp(\eta_{it}(\boldsymbol{\beta}))}{1+\exp(\eta_{it}(\boldsymbol{\beta}))},$$

其中

$$\eta_{it}(\boldsymbol{\beta})=\beta_0+\sum_{j=1}^{J}\left(\beta_j^{|0|}x_{ij(t)}+\beta_{j,\mathrm{lag}}\sum_{k=1}^{t-1}x_{ij(t-k)}\right)$$

矩条件变为：

$$\frac{\partial \mu_{is}(\tilde{\boldsymbol{\beta}})}{\partial \beta_j^{[k]}}\left[Y_{it}-\mu_{it}(\tilde{\boldsymbol{\beta}})\right]=x_{ij(s-k)}\mu_{is}(\tilde{\boldsymbol{\beta}})\left[1-\mu_{is}(\tilde{\boldsymbol{\beta}})\right]\left[Y_{it}-\mu_{it}(\tilde{\boldsymbol{\beta}})\right].$$

将这些分量在所有有效矩条件和受试者之间进行汇总，以形成经验矩向量 $G(\boldsymbol{\beta})$。利用公式（1），随后通过最小化标准二次型得到GMM估计量：

$$Q(\boldsymbol{\beta})=G(\boldsymbol{\beta})'\boldsymbol{W}G(\boldsymbol{\beta}). \tag{6}$$

其中，$\boldsymbol{W}$ 是一个正定权重矩阵。不同于以往依赖无导数优化方法（如 Nelder–Mead）的实现，我们采用 Broyden–Fletcher–Goldfarb–Shanno（BFGS）算法，该算法能够高效近似海森矩阵，并改善非线性情形下的收敛特性 [13]。

4 模拟

我们开展了模拟研究，将半分区GMM估计量与完全分区GMM以及仅使用滞后0期和滞后1期协变量的基准GMM模型进行比较，以评估其有限样本性能。我们主要关注不同时间相关结构和协变量类型下95%置信区间的覆盖率和平均区间长度。在前两个设定中，分组滞后结构与潜在时间模式相吻合，而设定III则用于评估存在中度不匹配时的表现。通过这一设计，我们可以评估稳健性，并展示分组滞后参数如何概括多个滞后效应的综合影响。这些模拟设定仅用于说明，并非面面俱到；我们的目标是突出若干有代表性的情形下稳定性—方差权衡，而非声称完全分区模型在所有情况下都不如本方法。在这些模拟中，我们采用公式（4）中的分组滞后设定。

4.1 设定I：具有平稳协变量的滞后期特定效应

此情景改编自 Lai and Small [8], ，模拟了结局同时取决于协变量当前值和滞后值的过程。具体而言，我们从以下模型生成数据：

$$\begin{aligned} Y_{it}&=\gamma_1 x_{it}+\gamma_2 x_{i,t-1}+b_i+e_{it},\\ x_{it}&=\rho x_{i,t-1}+\epsilon_{it}, \end{aligned} \tag{7}$$

其中，$(b_i,\ e_{it},\ \epsilon_{it})$ 是相互独立的正态变量，均值为零，方差分别为 4、1 和 1。 x_{it} 服从平稳 AR(1) 过程，即 $x_{i0}\sim N(0,\ \sigma_\epsilon^2/(1-\rho^2))$。

我们考虑 $i = 1, \ldots\ldots, n = 500$ 个个体在 $t = 1, \ldots\ldots, T = 5$ 这些时间点的数据。参数 γ_1、 γ_2 和 ρ 分别设为 1、1 和。5。注意，根据式 (7)，协变量在滞后期 0 和 1 处的分区广义矩估计最准确地表示了协变量与结局之间的关系。此外，我们注意到，当$\mu_{it} = \gamma_1 x_{it} + \gamma_2 x_{i,\ t-1}$

$$E[x_{is}(Y_{it} - \mu_{it})] = E[x_{is}(b_i + e_{it})] = 0, \quad \text{for any } 1 \le s \le T,$$

也就是说，在这种情形下，当模型纳入同期效应和一阶滞后效应时，该协变量符合 I 型的定义。然而，我们假设分析者不知道滞后结构，并使用预测变量 $\mu_{it} = \sum_{s=0}^{t-1} \gamma_{s+1} x_{i,\ t-s}$ 拟合完全分区和半分区 GMM 模型；在半分区情形下，我们采用约束 $\gamma_3 = \ldots = \gamma_T$，该约束与此情形下真实的数据生成过程相符，并对应于将所有高阶滞后项归入同一个区块。这种分组反映了一种常见的应用场景：研究者预期主要信号来自同期效应和一阶滞后效应，而其余滞后项则产生较小的平均影响。

表1展示了基于1,000个模拟数据集计算的经验覆盖率和95%置信区间的平均长度。对于 γ_1，所有方法的覆盖率都相近，略低于名义水平。对于 γ_2，半分区 GMM 与神谕估计量非常接近，且效率优于完全分区方法。这些结果表明，当真实滞后结构较为简单时，将高阶滞后项归组可获得与神谕模型相当的表现，并且相较于完全分区模型能够提高稳定性。

	γ_1		γ_2	
	置信区间覆盖率	平均置信区间长度	置信区间覆盖率	平均置信区间长度
仅滞后1期广义矩估计	.914	.0770	.94	.0844
半分区广义矩估计	.914	.0820	.94	.0845
（完全）分区广义矩估计	.916	.0823	.944	.0872

表1：设定I（平稳协变量、已知一阶滞后结构）的覆盖率和平均置信区间（置信区间）长度

4.2 设定 II：具有 III型协变量的反馈

第二个模拟设定引入了结局与协变量过程之间的反馈回路。具体而言，数据生成模型如下。

$$Y_{it} = \beta x_{it} + \kappa Y_{i,t-1} + u_{it},$$
$$x_{it} = \gamma Y_{i,t-1} + v_{it}$$

其中， u_{it} 和 v_{it} 是相互独立的正态随机变量，均值为零，方差分别为 σ_u^2和 σ_v^2。这一配置产生了III型协变量，即当前协变量值受到既往结局的影响。我们像之前一样设定 $n = 500$ 和 $T = 5$ ，并令 $\beta = 1$、 $\kappa = .4$、 $\gamma = .2$、 $\sigma_u^2 = \sigma_v^2 = 1$。与设定I一样，我们考虑预测变量 $\mu_{it} = \theta_1 x_{it} + \theta_2 x_{i(t-1)}$ （零阶和一阶滞后）以及$\mu_{it} = \sum_{s=0}^{t-1} \theta_{s+1} x_{i,\ t-s}$，在半分区情形下满足约束 $\theta_3 = \ldots = \theta_T$。尽管单滞后模型无法完全捕捉反馈结构，但数据生成过程中的高阶滞后效应可以忽略不计，因此将其作为比较的基准是合理的。

表 2 汇总了模拟结果。所有模型的覆盖率均接近名义水平。与单滞后广义矩估计相比，半分区广义矩估计显著提高了效率；完全分区模型则带来了额外但幅度有限的提升，代价是参数化程度更高。这些发现表明，即使在反馈导致有效矩条件有限的情况下，对高阶滞后项进行分组也能产生稳定且易于解释的估计值，同时不影响覆盖率。

	θ_1		θ_2	
	置信区间覆盖率	平均置信区间长度	置信区间覆盖率	平均置信区间长度
仅滞后1期广义矩估计	.944	.177	.944	.768
半分区广义矩估计	.944	.182	.944	.852
（完全）分区广义矩估计	.948	.188	.951	.923

Tab表 2：设定 II（反馈回路；III型协变量）的覆盖率和平均置信区间长度)

4.3 设定III：模型误设

我们重新考察设定I，并按公式（7）修改数据生成方式如下：

$$
\begin{aligned}
Y_{it} &= \gamma_1 x_{it} + \gamma_2 x_{i,t-1} + \gamma_3 x_{i,t-2} + \gamma_4 x_{i,t-3} + \gamma_5 x_{i,t-4} + b_i + e_{it}, \\
x_{it} &= \rho x_{i,t-1} + \epsilon_{it},
\end{aligned} \tag{8}
$$

其中 $\gamma_3 = 0.3$、$\gamma_4 = 0.2$ 和 $\gamma_5 = 0.1$；其他所有量均保持不变。此设定用于评估当真实的滞后特异效应在不同滞后期之间存在差异，但半分区模型将其分组时的稳健性。在这种情况下，分组参数概括的是不等滞后效应的综合影响，而非分别恢复每个系数，这体现了分组滞后设定的简约性。

我们考虑与之前相同的方法。请注意，只有完全分区广义矩估计能够捕捉修改后的滞后结构，而滞后 1 广义矩估计，以及将与第一期之后的滞后项相关的系数约束为相等的半分区广义矩估计，设定均不正确。表 3 总结了这一修改后设定的结果。半分区广义矩估计的覆盖率分别在第一个和第二个系数上略有下降和中度下降。半分区估计量仍比单滞后广义矩估计稳健得多，后者表现出严重的覆盖不足。在此设定中，完全分区估计量的表现最佳，正如预期，但代价是需要估计多个额外参数。当分组设定有误时，半分区估计量仍能对滞后效应提供稳定的概括，并在过度简化的单滞后模型与高维完全分区模型之间提供一种实用的折中方案。

	γ_1		γ_2	
	置信区间覆盖率	平均置信区间长度	置信区间覆盖率	平均置信区间长度
仅滞后1期广义矩估计	.342	.0834	.361	.0902
半分区广义矩估计	.906	.0822	.820	.0848
（完全）分区广义矩估计	.915	.0822	.944	.0872

表 3：设定III（模型误设）的覆盖率和平均置信区间长度。

4.4 总结

在前两个模拟设定中，半分区GMM模型始终保持适当的覆盖率，并且相较于完全分区模型具有更高的效率。当真实滞后结构较为简单或高阶滞后较弱时，分组所得估计值与神谕完全分区模型相当，但置信区间略短（即抽样变异性较小）。当协变量效应随时间变化，但大部分信号集中于即时效应或一阶滞后效应时，半分区方法提供了一种稳健且简约的替代方案。即使在模型设定错误的情况下，半分区估计量仍能对滞后效应提供稳定的概括，并且相较于单滞后广义矩估计，其可靠性显著更高，同时所需参数少于完全分区模型。这些结果支持在实际纵向设定中使用该方法，尤其是在同时重视可解释性和估计稳定性的情况下。

5 真实数据示例

我们将半分区GMM框架应用于两个真实世界数据集，以评估其相对于聚合式和完全分区GMM方法的表现。这些数据集涵盖连续结局和二元结局，包含多个随访波次，并纳入时变协变量，从而为所提出的方法提供了有力的实证验证。在这两个应用中，我们关注的是边际的总体平均效应，而非条件动态，这与前文提出的建模框架一致。

5.1 Osteoarthritis Initiative数据

我们的第一个应用基于 Osteoarthritis Initiative (OAI)，这是一项聚焦膝骨关节炎的大型多中心纵向研究 [14]。我们从该数据集中选取了 3,717 名参与者，他们在三个随访波次中均有完整测量数据。我们关注的结局是西安大略和麦克马斯特大学骨关节炎指数 (WOMAC) 残疾评分——这是一项广泛使用的连续指标，用于测量疼痛、僵硬和身体功能。我们对双膝评分取平均值，以得到统一的下肢残疾指标。关键预测变量包括性别（时不变）和两个时变协变量：年龄和体重指数 (BMI)，二者均为已知的骨关节炎进展风险因素。模型将年龄、BMI 和性别纳入 WOMAC 评分的预测变量。年龄和 BMI 被视为时变协变量，以反映它们可能随随访时间而变化并影响功能结局。

我们采用了全部三种 GMM策略——聚合式、完全分区式和半分区式——来估计边际效应，结果汇总于表4。所有模型均显示，较高的 BMI 和女性性别与 WOMAC 评分升高显著相关，而年龄未显示显著的同期效应。这些结果与既有文献一致，并进一步证实了边际GMM在捕捉已确立关联方面的有效性。

然而，不同模型之间也呈现出重要差异。完全分区方法虽然在理论上具有灵活性，但为年龄和体重指数的滞后效应引入了若干额外参数，而这些参数在统计上均不显著。这种扩展增大了标准误，并降低了估计效率。这些规律与随访时间较短的情形下估计大量特定滞后参数所面临的挑战一致，因为此时高阶滞后效应预计较弱。相比之下，半分区GMM模型对体重指数和性别即时效应的点估计几乎相同，同时通过对滞后效应进行分组，实现了更高的简约性。这种分组表明，年龄或体重指数的任何延迟影响都微乎其微，同时避免了估计大量特定滞后参数时标准误增大的问题。

这些结果表明，在随访时间较短的情况下，半分区GMM模型能够更高效、直观地表征纵向关系。它成功保留了完全分区的核心优势——即对滞后动态进行建模的能力，同时避免了其缺陷，尤其是在滞后效应较弱或临床意义不重要时。通过将边际效应分解为同期成分和分组滞后成分，半分区模型能够捕捉即时效应和延迟效应，同时避免过拟合。从实践角度看，分组滞后参数应被理解为对持续暴露的稳定概括，而非声称所有滞后期的效应都相同。

这些发现进一步证实了既往证据，即较高的体重指数和女性性别与骨关节炎相关疼痛及残疾增加有关，而年龄——尽管常与更差的结局相关——在与其他协变量一同建模时，其影响可能更为有限或间接 [15, 16, 17]。

5.2 Add Health数据

第二个应用使用了全国青少年至成年健康纵向研究（Add Health）的数据，该研究是一项具有全国代表性的队列，在四轮随访中追踪青少年进入成年期[18]。二元结局为肥胖状态，依据标准体重指数临界值定义。时变协变量包括身体活动、电视观看小时数、抑郁症状和饮酒。种族作为时间不变协变量纳入。

由于二元结局、随访波次较多，以及协变量之间可能存在时间反馈和行为相互依赖，该数据集构成了更具挑战性的建模情境。

不分区				半分区				完全分区			
	系数	SE	p值		系数	SE	p值		系数	SE	p值
截距	-13.8	1.65	<**0.001**	截距	-14.44	1.67	<**0.001**	截距	-14.46	1.67	<**0.001**
Sex	2.75	0.33	<**0.001**	Sex	2.60	0.33	<**0.001**	Sex	2.60	0.33	<**0.001**
Age	-0.01	0.02	0.69	Age	0.03	0.02	0.15	Age	0.03	0.02	0.09
BMI	0.68	0.04	<**0.001**	BMI	0.66	0.04	<**0.001**	BMI	0.65	0.04	<**0.001**
-	-	-	-	年龄（滞后项 1-2）	-0.01	0.005*	0.23	年龄（滞后 1）	-0.02	0.01	0.17
-	-	-	-	体重指数（滞后项 1-2）	-0.01	0.01	0.61	年龄（滞后 2）	-0.002*	0.01	0.86
-	-	-	-	-	-	-	-	体重指数（滞后 1）	-0.01	0.02	0.71
-	-	-	-	-	-	-	-	体重指数（滞后 2）	0.002*	0.02	0.92

表4：不同分区策略下WOMAC残疾评分的GMM估计值注1：数值均四舍五入至小数点后两位，但 $\leq$ 0.005除外；为保留精度，这些数值保留三位小数，并以星号（*）标记。注2：显著变量的P值以粗体显示

在此类情境中，有效的矩条件更加有限——尤其是对于受先前结果影响的协变量而言——因此模型稳定性和简约性尤为重要。

我们采用了全部三种GMM策略，结果汇总于表5。完全分区GMM模型虽在理论上具有灵活性，但得出的结果不稳定，有时甚至违背直觉。例如，按滞后期进行分区后，电视观看时长与肥胖之间关联的方向发生逆转，且标准误增大。这些异常现象与高维二元模型中估计大量特定滞后参数所面临的挑战一致；在这类模型中，多重共线性和信息稀疏可能扭曲统计推断。总体而言，这些模式表明完全分区模型对数据过度拟合，并且对滞后协变量之间的相关性高度敏感。相比之下，聚合GMM模型得出的估计值较为稳定，但将所有有效矩条件合并为单个系数，从而掩盖了重要的时间细微差异。

半分区模型的表现优于另外两种方法。它成功还原了屏幕使用时间增加、身体活动减少与肥胖风险升高之间预期的关系。尤为重要的是，该模型还捕捉到了抑郁症随时间持续产生的影响，包括具有统计学显著性的滞后效应——这些结果与此前关于青少年心理健康和肥胖风险轨迹的研究相符。通过对滞后效应进行分组，该模型在保留时间影响的可解释性概述的同时，稳定了估计。在这一应用中，分组滞后效应应被视为对较长期暴露的连贯概述，而不应被理解为各个滞后特异性系数完全相同。

这些发现凸显了半分区GMM模型在现实应用中的价值。该模型能够处理纵向二元结局中动态协变量行为的复杂性，而无需加入不必要的参数，使模型负担过重。该模型稳定统计推断、还原滞后关联并得出符合临床实际的结果，这些能力彰显了它在应用纵向研究中的实用优势。

这些结果与既往研究相符；既往研究表明，抑郁症状 [19, 20], 延长电视观看时间 [21], 以及减少身体活动 [22, 23] 均与青春期和青年期肥胖风险升高有关。半分区GMM模型为全部三种协变量估计出了显著的即时效应和滞后效应，其方向与文献中关于抑郁症和屏幕使用时间的发现一致。分组滞后估计值为持续的行为影响提供了稳定的概述，无需估计每个单独的滞后系数。值得注意的是，该模型还检测到身体活动滞后效应的符号反转，这与此前的发现相呼应：身体活动的长期影响可能会因行为和环境因素 [23]而减弱或受到混杂。

5.3 总结

本节展示的应用为半分区GMM模型提供了有力证据，表明该模型不仅在理论上具有吸引力，在多种纵向研究情境中也具有实际优势。在骨关节炎研究中，结局为连续型，且随访波次相对较少。半分区模型所得估计值在统计上与完全分区所得结果无法区分，同时避免了后者所伴随的标准误膨胀和过度参数化。分组滞后结构以简约而富有信息的方式分解了时间影响。重要的是，半分区模型保留了滞后协变量效应的可解释性，

不分区				半分区				完全分区			
	系数	SE	p值		系数	SE	p值		系数	SE	p值
截距	-1.99	0.13	<**0.001**	截距	-3.23	0.15	<**0.001**	截距	-2.59	0.19	<**0.001**
Race	0.07	0.10	0.46	Race	0.18	0.09	0.05	Race	0.11	0.12	0.38
电视观看时长	0.01	**0.002***	<**0.001**	电视观看时长	0.01	**0.002***	<**0.001**	电视观看时长	-0.21	0.10	**0.04**
酒精	0.10	0.06	0.13	酒精	-0.05	0.06	0.45	酒精	0.01	**0.004***	<**0.001**
活动	-0.57	0.04	<**0.001**	活动	-0.17	0.03	<**0.001**	活动	-0.41	0.20	**0.04**
抑郁症	0.77	0.10	<**0.001**	抑郁症	0.26	0.10	**0.01**	抑郁症	-0.09	0.05	0.10
-	-	-	-	电视观看时长（滞后1至3期）	**0.006***	**0.001***	<**0.001**	电视观看时长（滞后1期）	**0.004***	0.01	0.60
-	-	-	-	酒精（滞后项1-3）	0.04	0.05	0.35	电视观看时长（滞后期 2）	0.01	0.01	0.16
-	-	-	-	活动（滞后项 1-3）	0.13	0.03	<**0.001**	电视观看时长（滞后期 3）	**-0.003***	0.01	0.79
-	-	-	-	抑郁症（滞后项 1-3）	0.40	0.09	<**0.001**	酒精（滞后 1）	0.09	0.20	0.66
-	-	-	-	-	-	-	-	酒精（滞后 2）	0.04	0.21	0.85
-	-	-	-	-	-	-	-	酒精（滞后 3）	-0.32	0.31	0.29
-	-	-	-	-	-	-	-	活动（滞后 1）	-0.07	0.10	0.45
-	-	-	-	-	-	-	-	活动（滞后2）	0.14	0.13	0.26
-	-	-	-	-	-	-	-	活动（滞后 3）	-0.06	0.18	0.73
-	-	-	-	-	-	-	-	抑郁症（滞后 1）	0.46	0.43	0.28
-	-	-	-	-	-	-	-	抑郁症（滞后2）	0.98	0.52	0.06
-	-	-	-	-	-	-	-	抑郁症（滞后 3）	1.03	0.71	0.15

表5：不同分区策略下肥胖状态的GMM估计值注1：除 ≤ 0.005之外，数值均四舍五入至小数点后两位；为保留精度，≤ 0.005报告至小数点后三位，并以星号（*）标记。注2：显著变量的P值以粗体显示

这证实了大部分信号是通过同期关联捕捉到的，因此证明了对滞后项进行分组所带来的简约性是合理的。在Add Health研究中，结局为二元型，且时间结构更丰富，半分区方法的优势则更加明显。完全分区模型产生的系数估计值不稳定，有时甚至有悖直觉，这很可能是因为估计大量特定滞后参数引入了多重共线性和过拟合。相比之下，半分区模型不仅稳定了统计推断，还正确地恢复了行为协变量（如看电视、抑郁症、身体活动）与肥胖之间已知的关系。该模型在保持可解释性的同时，避免了估计众多较弱滞后效应所引入的噪声。通过捕捉即时效应和分组滞后效应，而无需使用数量庞大的参数，该模型兼顾了灵活性与简约性。

在这两个示例中，半分区广义矩估计均展现出一致、易于解释且高效的性能。它捕捉了时变协变量的动态影响，同时没有牺牲统计效能或模型稳定性。这些实证结果进一步强化了分组的理论依据：当高阶滞后效应较小或较为分散时，半分区能够对时间依赖性提供稳定且具有科学意义的概括。这些结果凸显了半分区框架在现实纵向应用中的实际优势，尤其是在重复测量次数适中、滞后效应存在但区分度不明显的情况下。总之，实证证据验证了半分区的理论依据：它在保留协变量—响应关系基本时间动态的同时，能够避免完全滞后分解通常伴随的噪声、低效率和解释混乱。因此，在建模延迟协变量效应十分重要、但又必须兼顾有限样本和设计约束的许多应用场景中，半分区广义矩估计方法不应被视为一种折中方案，而应被视为首选的默认方法。

6 讨论与后果

当协变量随时间变化并与结局呈现反馈关系时，纵向数据会带来复杂的分析挑战。在这类情形下，要准确估计边际效应，需要采用能够灵活兼顾时间变化和受试者内相关性的模型。完全分区GMM模型能够提供最精细的时间分辨率——分别估计每个滞后期的效应——但也需要在参数膨胀、标准误增大和模型不稳定等方面付出相当大的代价。相比之下，聚合GMM模型将所有有效矩条件合并为每个协变量的单一估计，从而提高效率，却完全掩盖了时间动态。本文提出的半分区GMM模型提供了一种有原则且实用的解决方案，能够克服这两种极端方法的局限

。该模型通过明确估计同期效应，并将滞后效应归入有结构的集合，在保留关键信息的同时避免估计大量参数。这种平衡对于随访时间中等的应用尤其有价值，因为高阶滞后效应往往较弱，但仍具有科学意义。与完全分区模型不同，后者在滞后阶数中等或样本量有限时往往会出现标准误膨胀和收敛失败；半分区模型则降低了参数空间的维度。这既提高了统计功效和计算稳定性，也保留了使GMM区别于广义估计方程（GEE）等较简单替代方法的丰富矩结构。

此外，尽管聚合GMM模型在设计上通过取平均来提高效率，但这种做法也会平均掉可能存在异质性的效应。因此，在即时协变量效应与延迟协变量效应方向不同或强度不一的情况下，聚合GMM模型可能导致误导性的结论。半分区模型通过保留这些区别来降低这一风险，从而提供更准确、更易于解释的估计值，更好地反映潜在过程的动态特性。

从实证结果来看，我们的模拟研究表明，就覆盖率和效率而言，半分区方法的表现与完全分区模型相当甚至更优，尤其是在存在滞后异质性或反馈的情形下。当分组结构设定错误时，该方法同样表现出较强的稳健性：它能够对平均滞后效应给出稳定的估计值，同时避免完全分区可能带来的不稳定性。在真实数据应用中，该模型得到了稳定且易于解释的效应估计，并避免了完全分区下出现的不稳定性和反直觉估计值，尤其是在二元结局情形下。这些发现进一步表明，半分区并非仅仅是一种折中方案，而是许多实际场景下更优的建模策略。

从应用角度看，半分区模型为包含时变协变量的纵向研究提供了一种务实的折中方案。它不仅超越了会掩盖时间动态的单一汇总效应，也避免了为每个滞后项分别估计系数所带来的不稳定性和解释负担。在我们的临床实例中，将滞后项分成少数几个区块，所得效应估计值既具有统计稳定性，也便于科学解释，使研究者能够区分协变量的即时效应与更持久的效应，同时避免过拟合。

6.1 优势

半分区GMM框架将时间效应估计的可解释性与参数缩减带来的效率相结合，推动了纵向边际建模的发展。与完全分区相比，它提供了一种更稳定、可扩展的替代方案；与完全汇总相比，则显著提高了估计值的真实度和信息量。该框架能够纳入反馈结构，并适用于连续型和二元结局，进一步拓宽了其适用范围。它能够适应不同结局类型、纳入反馈效应，并依托稳定的优化方法，因此对现代纵向健康研究尤其具有吸引力。

6.2 局限性

与所有建模框架一样，半分区广义矩估计也有其局限性。当协变量及其滞后项之间存在共线性时，尤其是在纳入许多时变预测变量的情况下，其性能仍可能受到影响。此外，该模型目前假设数据完整且按固定间隔收集，未对缺失或不均匀随访进行校正。若能扩展模型以处理不规则观测时间或信息性脱落，将进一步提升其实用性。对滞后效应进行分组虽有助于提高简约性，但如果一两个滞后项的影响远大于其他项，这种做法可能会掩盖有意义的异质性。

6.3 未来方向

尽管这对本文提出的方法并非至关重要，但我们认为，通过设计有原则的策略来选择所考虑的 $T-1$ 滞后项的最优分组，可以取得进一步改进。所提出的方法允许任意固定地分组为连续块（例如，对于 $T=6$的 $\{1,2\}$、{3,4}、{5}）。共有 $1+\sum_{k=1}^{T-2}\binom{T-2}{k-1}$ 种此类分组方式，其数量随 T迅速增加。对于中等的 T，它

仍然可以通过合适的模型选择准则评估每一种分组方式；对于较大的 T取值，惩罚估计是一种可行方法。

从思想上看，这种对候选分组的搜索与时间序列计量经济学中的从一般到具体策略相类似（例如 Hendry 的从一般到具体（GETS）框架；见 [24, 25]），尽管我们的重点仍然是边际均值结构，而非条件动态。相关的条件模型——例如 MIDAS设定——也探讨类似的滞后结构问题，但其目标是条件预测关系 [26, 27]。相比之下，半分区GMM框架专为纵向研究而设计，适用于科学研究关注总体平均协变量效应的情形。

在边际GMM设定下，开发用于选择或惩罚滞后组结构的原则性数据驱动程序，是当前研究的重点，相关成果将发表于另一篇论文。

进一步的方法学改进可以包括扩展模型以处理缺失或不规则数据，以及开发用于诊断检验和矩有效性的工具。这些发展将进一步拓展半分区模型已有的广泛实用性，并巩固其作为处理带有时变协变量的纵向数据的首选方法的地位。

7 结论

对含时变协变量的纵向数据进行建模一直面临持续挑战，尤其是当协变量效应随时间变化或涉及来自先前结果的反馈时。现有基于广义矩估计的方法往往走向两个极端：聚合模型通过合并所有有效矩条件来提高效率，但代价是掩盖时间动态；完全分区模型则提供最大的时间特异性，但在实际应用中常常参数过多且不稳定。本文提出一种半分区GMM框架，有策略地平衡灵活性、效率与可解释性。通过分别估计即时协变量效应，并将滞后效应分组为结构化区块，本模型保留了捕捉延迟协变量影响的能力，同时减轻了削弱完全分区方法的标准误增大和收敛问题。该框架适用于多种协变量类型，包括反馈驱动结构（III型），并支持连续型结局与二元结果建模。通过模拟研究，我们证明半分区模型以显著更少的参数，持续获得与完全分区模型相当或更优的覆盖率与效率。在两项实际应用中——考察骨关节炎中的功能残疾以及青少年的肥胖风险——半分区广义矩估计得到了稳定、可解释的结果，成功恢复了同期与滞后关联，同时避免了完全分区估计值中观察到的不稳定性与噪声。综合来看，这些发现支持将半分区广义矩估计视为许多纵向应用中的首选建模策略，而非一种折中。它为希望在时间保真度与统计及计算可处理性之间取得平衡的分析者提供了务实的解决方案。通过在不过拟合的前提下保留时间动态的关键特征，该框架拓展了边际GMM估计的实际适用范围，并填补了纵向数据分析工具箱中的一项关键空白。

参考文献

[1] Zeger SL和Liang KY。离散与连续结局的纵向数据分析。*Biometrics*1986; : 121–130.[2] Bena J和McIntyre S。相关事件时间数据的生存方法。载于*MWSUG 2008*会议。第12–4页。[3] Diggle PJ和Taylor-Robinson D。纵向数据分析。载于流行病学手册。Springer，2024。第1–34页。[4] Hansen LP。广义矩方法估计。宏观计量经济学与时间序列分析2010; : 105–118.[5] Sullivan Pepe M和Anderson GL。关于含纵向数据及一般相关响应数据的边际回归模型统计推断的警示。*Communications in statistics-simulation andcomputation* 1994; 23(4): 939–951.[6] Lee Y和Nelder JA。条件模型与边际模型：另一种观点。*Statistical Science* 2004; URL https://doi.org/10.1214/088342304000000305.[7] Zorn CJ。相关数据的广义估计方程模型：综述及应用。*American Journal of Political Science* 2001; : 470–490.[8] Lai TL和Small D。含时变协变量的纵向数据边际回归分析：一种广义矩方法。*Journal of the Royal Statistical Society Series B: StatisticalMethodology* 2007; 69(1): 79–99.[9] Irimata KM、Broatch J和Wilson JR。纵向数据的分区广义矩估计逻辑回归模型。*Statistics in Medicine* 2019; 38(12): 2171–2183.[10] Lalonde T、Wilson JR和Yin J。广义矩方法逻辑回归模型。*Journalof Data Science* 2014; 11(4): 715–738.[11] Hansen LP。广义矩方法估计量的大样本性质。*Econometrica* 1982; 50(4): 1029–1054.[12] Hall AR。广义矩方法。牛津大学出版社2005; .[13] Broyden CG。一类双秩最小化算法的收敛。*Journal of the Instituteof Mathematics and Its Applications* 1970; 6(1): 76–90.[14] Osteoarthritis Initiative。Osteoarthritis Initiative：一项膝关节健康研究。https://oai.ucsf.edu。访问于2025年6月21日。[15] Weiss E、Elman I、Rosenthal DI等。膝骨关节炎患者中体重指数、年龄与疼痛的关系：来自OAI的结果。*Rheumatology* 2014; 53(11): 2095–2101. DOI:10.1093/ rheumatology/keu202.[16] Fang J、Allison M、Wright J等。膝骨关节炎患者疼痛与功能的性别差异：心理困扰的作用。*BMJ Open* 2015; 5(9): e008542. DOI:10.1136/ bmjopen-2015-008542.[17] Ojeda F、Tío L、Castro-Domínguez F等。性别、年龄和体重指数在膝骨关节炎治疗决策中的作用：保守治疗与全膝关节置换。*Journal of OrthopaedicSurgery and Research* 2025; 20: 152. DOI:10.1186/s13018-025-05552-2. URL https://doi.org/10. 1186/s13018-025-05552-2.[18] Harris KM、Halpern CT、Whitsel E等。全国青少年至成年健康纵向研究：研究设计。https://addhealth.cpc.unc.edu/documentation/study-design, 2009. 访问于2025年6月。

[19] Anderson SE, Cohen P, Naumova EN 等. 在一项前瞻性社区研究中，抑郁症和焦虑障碍与体重变化的关联：对儿童进行随访直至成年期. 《儿科与青少年医学档案》2006; 160(3): 285–291. DOI:10.1001/archpedi.160.3.285. URL https://jamanetwork.com/journals/jamapediatrics/fullarticle/204656.[20] Lindberg L, Hagman E, Danielsson P 等. 肥胖儿童和青少年的焦虑与抑郁：一项瑞典全国性研究. 《*BMC*医学》2020; 18(1): 30. DOI:10.1186/s12916-020-1498-z. URL https://bmcmedicine.biomedcentral.com/articles/10.1186/s12916-020-1498-z.[21] MacDonell N 和 Hancox RJ. 儿童期和青少年期的电视观看与成年中期的代谢综合征. 《儿科学》2023; 152(2): e2022060768. DOI:10.1542/peds.2022-060768. URL https: // publications.aap.org/pediatrics/article/152/2/e2022060768/192843.[22] Rauner A, Mess F 和 Woll A. 青少年的身体活动、身体素质与超重之间的关系：对2000年及以后发表的研究进行的系统综述. 《*BMC*儿科学》2013; 13(1): 19. DOI:10.1186/1471-2431-13-19. URL https://bmcpediatr.biomedcentral.com/articles/10.1186/ 1471-2431-13-19.[23] Lundh H, Arvidsson D, Greven C 等. 肥胖儿童的身体活动和久坐行为——探讨儿童与父母之间的横断面关联. 《活动、久坐与睡眠行为杂志》2025; 4(1): 2. DOI:10.1186/s44167-025-00072-0. URL https://jassb. biomedcentral.com/articles/10.1186/s44167-025-00072-0.[24] Hendry DF. 《动态计量经济学》. 牛津大学出版社, 1995.[25] Hendry DF. 《计量经济学：炼金术还是科学？》《经济学》2000; 67(267): 387–406.[26] Ghysels E, Santa-Clara P 和 Valkanov R. 预测波动率：充分利用不同频率采样的收益数据. 《计量经济学杂志》2006; 131(1–2): 59–95.[27] Andreou E, Ghysels E 和 Kourtellos A. 混合采样频率的回归模型. 《计量经济学杂志》2010; 158(2): 246–261.