

基于个体资产的资产定价模型的实证检验：解决风险溢价估计中的变量内误差偏差

为了减少固有的变量误差偏差，投资组合被广泛用于风险溢价估计；但投资组合可能会diversify away(分散)，从而掩盖个别资产的相关风险或回报相关特征。我们提出了一个解决方案，允许使用单个资产(individual assets)，同时避免偏差。它取决于特定的工具变量，根据交替观测结果计算出的因素敏感性(β 's)。给出了大截面和时间序列的闭型渐近性。仿真表明，IV方法提供了无偏风险溢价估计和明确的足够功率的测试。实证研究发现市值和市账率这两个因子有显著的风险溢价(这是Fama-French三因子)。然而，在控制非 β 特征时，CAPM、规模、市场账面价值、投资、盈利能力和流动性调整CAPM的估计风险溢价不显著。

1.Introduction

金融经济学的一个基本原则是，投资者通过承担系统性风险来获得更高的平均回报。虽然这一观点已被广泛接受，但对于系统风险的identity或假定回报的大小，人们几乎没有达成共识。这并不是因为两种探索方式都缺乏努力。首先，许多candidates被认为是潜在的风险因素。其次，估计风险溢价的实证努力也有悠久而多样的历史。

从单因素CAPM(Sharpe, 1964; Lintner, 1965)和多因素APT, Ross (1976) 开始，第一支线已经提出了大量的风险因素候选者。其中，这些包括Fama-French的市值和市账率因素、人力资本风险(Jagannathan and Wang, 1996)、生产力和资本投资风险((Cochrane, 1996; Chen, Novy-Marx and Zhang, 2011; Eisfeldt and Papanikolaou, 2013)、消费风险的不同组成部分(Lettau and Ludvigson, 2001; Ait-Sahalia, Parker, and Yogo, 2004; Li, Vassalou, and Xing, 2006)、现金流和贴现率风险(Campbell and Vuolteenaho, 2004)和非流动性风险((Pastor and Stambaugh, 2003; Acharya and Pedersen, 2005)。Harvey, Liu and Zhu (2015)调查了文献，并报告说已经提出了300多个因素。

第二支线探索已经对许多人的风险溢价进行了经验估计，即Cochrane (2011)所说的因素称为“zoo”。大多数估计方法都遵循了Black, Jensen and Scholes (1972) (BJS)，并由Fama和Macbeth (1973) (FM)改进。它们最显著的特点是在测试资产定价模型时使用了投资组合，而不是单个资产。长期以来，这一直被认为是必不可少的，因为在估计风险溢价时存在一个固有的变量误差(EIV)问题。

通过通过BJS和FM方法进行跟踪，最能解决EIV问题。它涉及到两阶段回归：第一阶段是对单个资产收益的时间序列回归。这一阶段提供了因素负荷的估计，在金融文献中被广泛地称为“beta”(此后，我们将采用速写命名法“Beta”来表示“因子敏感性”或“因子加载”)。第二次通过根据第一次通过回归得到的beta值对资产回报进行截面回归。由于第二次通过的解释变量是估计值，而不是真实的beta，由此产生的风险溢价估计是有偏的和不一致的；当两阶段回归涉及多个因素时，偏差的方向是未知的。

由于有大量的单个资产($=N$)，EIV偏差可以通过使用投资组合而不是单个资产来减少。这个过程首先形成由某些个别资产特征分类的多元化投资组合，例如在初步样本期内估计的beta值。然后，第二阶段利用数据来估计投资组合的beta系数。最后，第三阶段使用数据对估计的投资组合beta进行了截面回归。BJS, Blume和Friend (1973) 和FM注意到投资组合的特异性风险较小；因此，变量中的误差偏差减少了(并且可以随着N的增长而完全消除)。

但使用投资组合，而不是单个资产，也有其自身的缺陷。由于降低了维度，test power是一个直接的问题；也就是说，不可避免的投资组合的解释变量比单个资产少(？特异性风险被对冲了？)。

对投资组合的多样化可以掩盖与投资组合分组程序无关的个别资产的横断面现象。例如，基本指数化的倡导者(Arnott, Hsu和Moore (2005))认为，高市场价值资产定价过高，反之亦然，但任何由市场价值本身以外的属性分组的投资组合都可能分散这种错误定价，使其无法察觉。

投资组合掩盖的另一个麻烦的结果涉及到平均回报和因素风险敞口(“beta”)之间的横断面关系。以单因子CAPM为例(尽管对任何线性因子模型都有同样的作用)。当且仅当用于计算贝塔的市场指数处于单个资产宇宙的均值/方差边界时,预期回报与贝塔之间的截面关系完全成立。beta/回报线的错误,无论是正的还是负,都意味着指数不在边界上。但是,如果个别资产按投资组合beta分类到投资组合中,并且个体误差与beta无关,那么符合投资组合收益和beta的类似线将显示出小得多的误差。这可能导致错误的推断,即指数处于有效边界。

(分层测试)测试投资组合通常根据与平均回报相关的公司特征进行组织,例如,规模和对市场的账面价值。对已知的可以预测回报的特征进行排序,有助于产生跨测试资产的平均回报的合理变化。但Lewellen, Nagel, and Shanken (2010)指出,对特征的排序也为测试投资组合提供了很强的因素结构。Lewellen等人(2010)表明,即使是与排序特征弱相关的因素,也可以解释不同测试投资组合的平均回报的差异,而不管这些因素背后的理论的经济价值。

Test portfolios are typically organized by firm characteristics related to average returns, e.g., size and book-to-market. Sorting on characteristics that are known to predict returns helps generate a reasonable variation in average returns across test assets. But Lewellen, Nagel, and Shanken (2010) point out sorting on characteristics also imparts a strong factor structure across test portfolios. Lewellen et al. (2010) show that as a result even factors that are weakly correlated with the sorting characteristics would explain the differences in average returns across test portfolios regardless of the economic merits of the theories that underlie the factors.

最后,风险溢价的统计显著性和经济规模可能严重取决于测试投资组合的选择。例如,当根据相应的特征对测试投资组合进行排序时,Fama and French的规模以及账面到市场的风险因素的定价显著,但当仅根据势头对测试投资组合进行排序时,它们不会获得显著的风险溢价。

Finally, the statistical significance and economic magnitudes of risk premiums could depend critically on the choice of test portfolios. For example, the Fama and French size and book-to-market risk factors are significantly priced when test portfolios are sorted based on corresponding characteristics, but they do not command significant risk premiums when test portfolios are sorted only based on momentum.

为了克服投资组合分组的缺陷,同时避免了EIV偏差,我们开发了一种新的程序来估计风险溢价,并使用单个资产来检验其统计显著性。我们的方法采用了工具变量技术,一个标准的计量经济学解决EIV问题的方法。我们定义了一组特定的表现良好的工具,并随后将我们的方法称为IV方法

In an effort to overcome the deficiencies of portfolio grouping while avoiding the EIV bias, we develop a new procedure to estimate risk premiums and to test their statistical significance using individual assets. Our method adopts the instrumental variables technique, a standard econometric solution to the EIV problem. We define a particular set of well-behaved instruments and hereafter refer to our approach as the IV method.

具体来说,我们的IV方法首先从数据样本中可用的一部分观察结果中估计单个资产的测试值。这些成为第二阶段横截面回归的“独立”变量。然后,使用完全不同的样本观察,它重新估计相同的beta,这成为第二阶段横截面回归中的“工具”变量。

To be specific, our IV method first estimates betas for individual assets from a portion of the observations available in the data sample. These become the “independent” variables for the second-stage cross-sectional regressions. Then, using completely different sample observations, it re-estimates the same betas, which become the “instrumental” variables in the second-stage cross-sectional regressions.

我们将探索一下这个基本方案的几种变体。一种变体使用序列子样本的观测来估计贝塔和贝塔工具;第一次T观测的betas和观测T+1到2T的beta工具。然后在2T+1中进行横截面回归观察。整个过程向前滚动一个周期,并重复到最后一次可用的观察结果,从而生成一个用于统计检验的风险溢价估计数的时间序列。

We explore several variants of this basic scheme. One variant estimates betas and beta instruments using observations from sequential subsamples; betas from the first T observations and beta instruments from observations $T+1$ to $2T$. A cross-sectional regression then is run for observations in $2T+1$. The entire procedure is rolled forward by one period and repeated up until the last available observation, thereby generating a time series of risk premium estimates for statistical testing.

基本方案的另一种变体是从偶数周期的观测中估计贝塔，例如，2、4个月，和奇异时期的观测beta工具，如1、3个月，... $2T-1$ 。然后使用 $2T+1$ 中的返回来运行第二阶段的横截面回归。在第二阶段横截面回归中，betas及其工具的作用是可以互换的。

Another variant of the basic scheme is to estimate betas from observations in even-numbered periods, e.g., months 2, 4, ... $2T$, and beta instruments from observations in odd-numbered periods, e.g., months 1, 3, ... $2T-1$. The second-stage cross-sectional regressions are then run using returns in $2T+1$. The roles of betas and their instruments are interchangeable in the second-stage cross-sectional regressions.

IV方法对有限长度的时间序列产生 n -一致的风险溢价估计。在一个更一般的情况下，既允许横截面 N 的大小和时间序列 T 的长度无界增长，我们证明了IV方法提供了 NT -一致的风险溢价估计。我们还发展了在 N 和 T 的这两种不同情况下的IV估计量的渐近分布。

The IV method produces N -consistent risk premium estimates for a finite length of time-series. In a more general case that allows both the size of cross-section N and the length of time-series T to grow without bounds, we prove that the IV method provides NT -consistent risk premium estimates. We also develop the asymptotic distributions of the IV estimator for those two different scenarios of N and T .

虽然大样本特性可以提供一些指导，但在实际应用中，检查各种估计的小样本性能是很重要的。为此，我们进行了一些模拟实验。我们选择与实际数据相匹配的模拟参数。模拟结果表明，即使在用于估计因子敏感性的相对较短的时间序列内，IV方法也能产生无偏的风险溢价估计。相比之下，我们发现使用OLS适应第二阶段回归的标准方法(以下我们将这种标准方法称为OLS方法)存在严重的EIV偏差。模拟结果还表明，IV方法的均方根误差明显低于OLS方法。例如，在随时间变化保持常数因子敏感性下的单因子模型模拟中，我们发现OLS估计器，如果使用单个股票，在使用2640次时间序列观测估计beta值时会显著偏向于零。相比之下，当只有264个时间序列观测值可用时，IV估计器就可以产生几乎无偏不倚的风险溢价估计值（见图1）

While large sample properties can provide some guidance, it is important to examine the small sample performance of various estimators for practical applications. To do so, we conduct a number of simulation experiments. We choose simulation parameters matched to those in the actual data. Simulation results verify that the IV method produces unbiased risk premium estimates even for relatively short time-series used to estimate factor sensitivities. In contrast, we find that the standard approach that fits the the second stage regressions using OLS (hereafter we will refer to this standard approach as the OLS method) suffers from severe EIV biases. The simulations also show that the root-mean-squared errors of the IV method are substantially lower than those of the OLS method. For example, in simulations with a single factor model under time-constant factor sensitivities, we find that the OLS estimator, if used with individual stocks, is significantly biased toward zero even when betas are estimated with 2640 time series observations. In contrast, the IV estimator yields nearly unbiased risk premium estimates when only 264 time-series observations are available (see Figure 1).

就测试规模(即I型错误)和power(即II型错误)而言，我们发现基于IV估计的传统的t检验是很好的（在零假设真正风险溢价为零），并且他们是相当强大的（在替代假设下即真正风险溢价等于因子实现的样本平均值）在小样本估计测试。对于Fama-French三因子模型，也发现了类似的结果。

In terms of test size (i.e., type I error) and power (i.e., type II error), we find that the conventional t-tests based on the IV estimator are well specified (under the null hypothesis that true risk premiums are zero) and they are reasonably powerful (under the alternative hypothesis that true risk premiums equal the sample means of factor realizations) in small samples for estimating betas. For the Fama-French three-factor models, similar results are found.

根据实际数据，我们采用IV法估算文献中提出的几个风险因素的风险溢价，包括CAPM、Fama and French 的三因素和五因素模型（1993和2014）、Hou, Xue, and Zhang的（2014）的q因素资产定价模型，以及Acharya and Pedersen（2005）的流动性调整资本资产定价模型(LCAPM)。这些风险因素在用投资组合进行测试时，在经验上是成功的。与原始论文相比，当控制相应的非 β 特征，我们发现这些因素没有一个与个股回报横截面的显著风险溢价。

With actual data, we apply the IV method to estimate the risk premiums for several risk factors proposed in the literature, which include the CAPM, the three-factor and five-factor models of Fama and French (1993 and 2014), the q-factor asset pricing model of Hou, Xue, and Zhang (2014), and the liquidity-adjusted capital asset pricing model (LCAPM) of Acharya and Pedersen (2005). These risk factors have been empirically successful when they were tested with portfolios. In contrast to the original papers, when controlling for corresponding non- β characteristics, we find that none of these factors is associated with a significant risk premium in the cross-section of individual stock returns.

这种未能发现显著风险溢价的原因并不是由于IV方法缺乏test power。我们提出了模拟证据，表明基于IV方法的t检验在替代的假设下，即真实的风险溢价等于因素实现的样本均值，提供合理的高功率。例如，当真实的HML风险溢价为正时，在时间常数和时变因子敏感性下，零假设(即HML的零风险溢价)的拒绝率分别为84.2%和89.6%。在分析真实数据时，在缺乏非 β 特征的情况下，我们发现有一些证据表明，SMB和HML在个股回报的横截面中具有显著的风险溢价。然而，当横断面回归中包含相应的非 β 特征时，SMB和HML的beta的定价证据大大减弱。有和没有非 β 特征的定价证据的明显差异表明，不显著的风险溢价不是由于缺乏IV方法的测试能力。

This failure to find significant risk premiums is not due to the lack of test power of the IV method. We present simulation evidence that the t-tests based on the IV method provide reasonably high power under the alternative hypotheses that the true risk premiums equal the sample means of factor realizations. For example, when the true HML risk premium is positive, the rejection rates of the null hypothesis (i.e., zero risk premium for HML) are 84.2% and 89.6% under time-constant and time-varying factor sensitivities, respectively. When analyzing real data, in the absence of non- β characteristics, we find some evidence that SMB and HML command significant risk premiums in the cross-section of individual stocks returns. However, this pricing evidence of SMB and HML betas is substantially weakened when corresponding non- β characteristics are included in the cross-sectional regressions. This stark difference in pricing evidence without and with non- β characteristics indicates that insignificant risk premiums are not due to the lack of test power of the IV method.

我们的论文也有助于一个关于测试资产定价模型的大量文献。随着时间序列的长度无限增长，Shanken（1992）表明，由于beta的估计精度的提高，EIV偏差可以忽略不计。他还推导了OLS方法的FM标准误差的渐近调整。Jagannathan and Wang（1998）将Shanken的渐近分析扩展到时间序列回归中的条件异质误差的情况。Shanken and Zhou（2007）和Kan, Robotti and Shanken（2013）将结果扩展到错误指定的模型misspecified models。这些论文的证据和分析主要集中在测试投资组合。我们的论文以个股作为测试资产，并提出了IV方法来减轻测试资产定价模型中的EIV偏差，这可能比投资组合更严重。据此，我们给出了IV估计量的一个“双”渐近理论，其中横截面的大小和时间序列的长度在无界同时增长。双渐近反映了计量经济学的最新发展，例如Bai（2003），它比单一渐近更适合于个股。

Our paper also contributes to a large literature on testing asset pricing models. As the length of time-series grows indefinitely, Shanken (1992) shows that the EIV bias becomes negligible because the estimation accuracy of betas improves. He also derives an asymptotic adjustment for the FM standard errors of the OLS method. Jagannathan and Wang (1998) extend the Shanken's asymptotic analysis to the case of conditionally heterogeneous errors in the time series regression. Shanken and Zhou (2007) and Kan, Robotti and Shanken (2013) extend the result to misspecified models. The evidence and analyses in those papers mainly focus on test portfolios. Our paper focuses on individual stocks as test assets and proposes the IV method to mitigate the EIV bias in testing asset pricing models, which is likely more severe with individual stocks than with portfolios. Accordingly, we provide a "double" asymptotic theory of the IV estimator, in which the size of cross-section and the length of time-series grow simultaneously without bounds. The double asymptotics reflects a recent development in the econometrics, e.g., Bai (2003) and it is more appropriate for individual stocks than single asymptotics

在测试资产定价模型中使用个股是文献中的一个最新进展。Kim (1995) 利用滞后beta修正了EIV偏差来推导市场风险溢价MLE估计量的封闭解。Kim提出的解决方案是基于Theil (1971) 的调整。Litzenberger and Ramaswamy (1979)、Kim and Skoulakis (2014) 以及Chordia等人 (2015) 提出的其他方法也类似, 产生了EIV修正项, 以获得n-一致的风险溢价估计量。相比之下, IV方法不需要任何对EIV偏倚的校正项。为了避免EIV偏倚, Brennan等人 (1998) 在第二阶段的回归分析中, 提倡将风险调整收益作为因变量。。然而, 他们的方法并没有估计各种因素的风险溢价。这些现有的论文都没有像我们的论文那样提供双渐近理论。

Using individual stocks in testing asset pricing models is a recent development in the literature. Kim (1995) corrects the EIV bias using lagged betas to derive a closed-form solution for the MLE estimator of market risk premium. The solution proposed by Kim is based on the adjustment by Theil (1971). Other methods proposed by Litzenberger and Ramaswamy (1979), Kim and Skoulakis (2014), and Chordia et al. (2015) are similar, producing the EIV correction terms to obtain N-consistent risk premium estimators. In contrast, the IV method does not require any correction term for the EIV bias. To avoid the EIV bias, Brennan et al. (1998) advocate risk-adjusted returns as dependent variable in the second-stage regressions. However, their method does not estimate the risk premiums of factors. None of these existing papers provides double asymptotic theories as in our paper.

2. Risk-Return Models and IV Estimation

2.1. Multifactor Asset Pricing Models

许多资产定价模型预测, 风险资产的预期回报与其与某些风险因素的协方差呈线性相关。K-factor资产定价模型的一般规范可以写为:

$$E(r^i) = \gamma_0 + \sum_{k=1}^K \beta_k^i * \gamma_k$$

其中 $E(r^i)$ 是股票i的预期超额回报, β_k^i 是股票i对因子k的敏感度, γ_k 是因子k的风险溢价。

