基于倾向得分匹配法的会计师事务所审计溢价分析

王 旭(副教授)

(镇江高等专科学校，江苏镇江 212003)

【摘要】为了控制样本选择偏误,本文采用倾向得分匹配法来分析国内会计师事务所审计费用溢价情况。与线性回归结论相比,在国际四大会计师事务所、国内六大会计师事务所和国内一般会计师事务所之间仍然存在审计费用溢价;而在控制选择偏误影响后,总体溢价程度相对降低。通过Rosenbaum边界分析表明,国际四大会计师事务所相对国内六大会计师事务所或一般会计师事务所的审计溢价结果对未观测到变量不敏感,结论是稳健的。

【关键词】审计溢价；倾向得分匹配法；Rosenbaum边界分析

一、引言

审计溢价是会计师事务所为弥补建立品牌所付出的成本而收取的较高的审计费用。国外多数研究认为,国际四大会计师事务所(简称国际“四大”所)在欧美成熟经济体能提供可靠的审计质量而得到市场的认可,取得了垄断性的市场份额和较高的审计费用(Simunic,1980;Peel和Roberts,2003)。在国内审计市场,国际“四大”所凭借其品牌相对于国内其他会计师事务所是否取得审计溢价并没有得到一致结论(耿建新和房巧玲,2006)。

Ireland和Lennox(2001)认为不一致的原因是研究方法存在问题,因为客户对会计师事务所的选择通常不是随机的,高质量客户倾向于选择大所审计,以往研究没有控制客户与审计师之间的自选择偏误。为了控制审计客户的自选择问题,本文借鉴Peel(2009)使用倾向得分匹配法(Propensity Score Matching),以2007~2012年间国内上市公司为研究对象,分析国际“四大”所(或国内“六大”所)与非国际“四大”所(或国内小所)相比审计费用溢价程度。结果显示,国际“四大”所、国内“六大”所、小所之间存在审计费用溢价,与线性回归结论相比,总体溢价程度相对降低,比如国际“四大”所相对非“四大”所只取得88%的审计溢价。通过Rosenbaum边界分析来衡量未观测到变量对处理效应误差的影响程度,结论显示国际“四大”所相对国内“六大”所或小所的审计溢价是稳建的。

二、文献综述与研究思路

Simunic(1980)首先对审计师是否收取溢价进行经验研究,结果显示,由于大所在审计时具有规模经济效应,所以“六大”所的审计收费略低于非“六大”所。而Francis(1984)、Palmrose(1986)研究则认为,无论在大客户市场还是在小客户市场,“六大”所审计收费都显著高于非“六大”所。Pong和Whittington(1994)研究发现,“六大”所的审计收费平均值要高于非“六大”所,但是在处理复杂的审计问题时,审计收费溢价低于非“六大”所而更有效率。Moizer(1997)研究认为,在一个竞争性日益激烈的市场中,不同会计师事务所之间的收费差异,代表了对会计师事务所高声誉资本投资的回报,但在不同的国家结论有所不同。Hay(2006)统计表明,在63份调查审计收费的研究中,支持在不同的国际“四大”所收取了溢价有42份,其余21份研究没有发现存在溢价。

国内伍莉娜(2003)、漆江娜(2004)等研究表明,国际“四大”所在中国审计市场审计费用存在显著溢价。李连军(2005)研究进一步发现,审计师声誉与审计费用的“质量溢价”呈正相关关系,不仅国际“四大”所的审计费用显著高于国内“六大”所，而且国内“六大”所的审计费用也显著高于国内其他所。吴应宇等(2008)研究认为,国内“六大”所针对大型客户收取了高额审计费用,而小客户群则不存在溢价情况。但是李爽和吴溪(2004)用2000年~2001沪深两市A股上市公司为样本进行实证研究,并未发现国内前“六大”所(包括国际“四大”所)具有收费溢价。耿建新等(2006)在对会计师事务所审计质量和客户特征控制后,发现国际“四大”所收取的审计费用并不比本土会计师事务所高。因此国际“四大”所在国内外是否收取了溢价仍没有得到一致结论。

Ireland和Lennox(2001)认为,由于客户对会计师事务所的选择通常不是随机的,高质量的客户倾向于选择大型会计师事务所审计,审计客户并没有被随机分配到国际“四大”所或者非国际“四大”所,将所有审计客户纳入一起分析无法克服样本选择偏误问题而导致结果有偏。因此研究者开始使用Heckman两步法控制自选择问

题,例如Ireland和Lennox(2002)、McMeeking et al.(2006)研究发现,大所审计溢价显著为正;而Chaney et al.(2004)则研究发现,大所显著为负,大所的审计收费并不高于小所,相反客户如果选择与实际选择相反的会计师事务所,会支付更高的费用。Peel(2009)认为,Heckman估计在使用时可能由于不能满足其严格的假设条件,使得结论缺乏稳健性,因而会得出相反的结论。

针对以上问题,本文采用倾向得分匹配方法(Rosenbaum,1983),控制审计客户对会计师事务所的自选择问题,以中注协对全国前百家会计师事务所的评分排名为依据,将不同会计师事务所划分为国际"四大"所、国内"六大"所、国内小所三类,分析不同类型会计师事务所对客户收取的审计溢价情况,最后通过Rosenbaum边界分析未观察变量对匹配结论的影响程度。

三、研究方法

1. 倾向得分匹配法。Rosenbaum和Rubin于1983年提出利用倾向得分匹配法(PSM匹配法)来消除混淆因素(Confounding Factor)所产生的偏误,该方法已被广泛应用于药物治疗、政策实施效果评价等研究中。倾向得分匹配法通过均衡处理组(Treatment group)与控制组(Control group)之间的协变量分布,对非随机化研究中的混淆因素进行类似随机化的均衡制造一个"准随机"试验;其利用降维的思想,将多个特征变量浓缩成一个指标——倾向得分值,再通过它将处理组对象和控制组中预处理特征类似的对象相匹配,比较处理组与控制组结果差异。得到该因素的影响效果,可有效降低样本选择的偏误问题。

PSM匹配法是一种非参数统计技术,具有估计处理效应时不要求具有一定函数形式或服从特定假设的优点,比较适合于会计和财务领域。对于审计溢价问题,考虑到选择"四大"所审计的客户和选择非"四大"所的客户在某些关键变量上存在着系统差异,若将所有审计客户纳入一起分析,则无法克服样本选择偏误而导致结果不一致。这时可以借助PSM匹配法找到一组和选择"四大"所审计的企业特征相似的非"四大"所审计企业进行比较分析:"四大"所审计的企业同时被非"四大"所审计情况下两者的审计费用差异,即可认为是审计溢价。

PSM匹配方法运用步骤如下:

(1)选择变量后,采用Probit(或者Logit)模型估算倾向得分值。处理组为选择"四大"所(或国内"六大"所)进行审计的企业,控制组为不选择"四大"所(或国内"六大"所)进行审计的企业。

倾向得分被定义为在给定样本特征X的情况下,若企业选择"四大"所(或国内"六大"所)进行审计,则D=1,否则D=0,其概率表示如下:

$$P(x)=Pr[D=1|X]=E[D|X] \quad (1)$$

(2)根据估算的倾向得分值选择匹配方法对处理组和控制组样本进行匹配。常用的匹配方法有最近邻匹配法(Nearest neighbor matching)、核匹配法(Kernel matching)、半径匹配法(Caliper or Radius matching)、分层匹配法(Stratification matching)等方法。

(3)检查匹配样本是否满足共同支撑假设(Common Support Assumption)和条件独立分布假设(Conditional Independence Assumption)。共同支撑假设要求处理组和控制组样本特征分布有一定相同的量,匹配质量才高,否则会影响结论。条件独立分布假设是假设在一组不受事务所选择影响的可观测协变量情况下,审计费用大小独立于事务所的选择,这时企业选择会计师事务所审计是随机的,就可以比较相同特征的企业在选择和不选择"四大"所审计情况下审计费用的差异。如果在匹配后发现两个假设不满足,应调整匹配变量重新计算倾向得分或者改变匹配方法,重复多次直到假设得到满足为止。

(4)比较匹配后平均处理效应,分析处理效果。匹配后总体层次的平均处理效应ATE为:

$$\begin{aligned}ATE&=E[Y_1-Y_0]=E[Y_1]-E[Y_0]\\&=\{\pi E[Y_1|D=1]+(1-\pi)E[Y_1|D=0]\}-\{\pi E[Y_0|D=1]+(1-\pi)E[Y_0|D=0]\}\\&=\pi\{E[Y_1|D=1]-E[Y_0|D=1]\}+(1-\pi)\{E[Y_1|D=0]-E[Y_0|D=0]\}\end{aligned} \quad (2)$$

π是样本总体中选择"四大"所(D=1)进行审计的客户所占比例,Y_1和Y_0分别表示同一客户在选择和不选择"四大"所审计两种情况下的审计费用。

处理组的平均处理效应ATT为:

$$ATT=E[Y_1|D=1]-E[Y_0|D=1] \quad (3)$$

控制组的平均处理效应ATU为:

$$ATU=E[Y_1|D=0]-E[Y_0|D=0] \quad (4)$$

ATT表示被国际"四大"所审计公司在不选择国际"四大"所审计的情况下审计费用的差异;ATU表示被非国际"四大"所审计的公司在选择国际"四大"所审计的情况下审计费用的差异;ATE表示样本中选择国际"四大"所和未选择国际"四大"所审计的企业审计费用差异的加权平均,三者中ATT是分析处理效应的重要结果。

2. Rosenbaum边界敏感性分析。Heckman(2008)研究认为,当存在未观测到的混淆因素时,倾向得分匹配法不仅不能消除系统偏差,反而会带来新的偏差。Becker和Caliendo(2007)也认为,在研究中检查估计的处理效应对假设偏离的敏感程度越来越重要。因此我们参考Duvendack(2011)使用Rosenbaum边界分析法来分析匹配结果受未观测到的变量影响的敏感程度。

假设公司i选择国际"四大"所进行审计的概率是P_i,则此概率不仅由观察到的配对变量X_i所决定,而且受到未观察到的变量u_i的影响,由下式表示:

$$P_i=P(X_i,u_i)=P(D_i=1|X_i,u_i)=F(\beta X_i+\gamma u_i) \quad (5)$$

式中：γ表示变量u_i对企业是否选择“四大”所审计的影响程度；F表示logistic分布。

针对匹配公司i和j，其接受“四大”所审计的可能性分别由$P_i/(1-P_i)$和$P_j/(1-P_j)$表示，因而同时具有未观察到特征变量的处理组公司i和控制组公司j的概率比（odds ratio）由下式给出：

$$\frac{P_i/(1-P_i)}{P_j/(1-P_j)}=\frac{\exp(\beta X_i+\gamma u_i)}{\exp(\beta X_j+\gamma u_j)} \quad (6)$$

匹配过程要求i和j两家公司有相同或相近的可观察变量，即$X_i=X_j$，因此得到：

$$\frac{P_i/(1-P_i)}{P_j/(1-P_j)}=\exp[\gamma(u_i-u_j)] \quad (7)$$

这时两家公司在选择“四大”所审计的概率通过参数γ和未观察到的协变量u的差异共同决定。如果未观察到的变量对进入处理组概率没有影响（$\gamma=0$），或者未观察到的变量对处理组和控制组个体影响相同（$u_i=u_j$），则概率比等于1，表明不存在与未观察到变量相关的隐藏偏误。随着γ和u_i-u_j的大小变化，可以估计处理效果受其影响的敏感度。假设$\Gamma=e^{\gamma}$，确定的概率比范围如下：

$$\frac{1}{\Gamma}<=\frac{P_i/(1-P_i)}{P_j/(1-P_j)}<=\Gamma \quad (8)$$

式中：$\Gamma=1$（$\gamma=0$）表示不存在隐藏偏误，随着Γ值（未观察到变量u的系数γ的对数值）的增加，意味着未观察到的特征变量对处理组选择的影响也越大。Rosenbaum边界分析即针对不同Γ值计算的处理效应的置信区间来判断结论多大程度会被改变，如果包括零的置信区间的Γ值小于2，则意味着这时所估计的处理效应对未观测到的变量是敏感的（Duvendack，2011）。

四、样本选取与变量定义

1. 样本选取。本文选取2007～2012年沪深两市A股上市公司为研究样本，样本公司财务数据和治理数据来自CSMAR和CCER数据库，会计师事务所排名资料来自中国注册会计师协会每年公布的全国前百家会计师事务所综合排名，并进行了如下技术处理：①剔除银行、保险等金融类上市公司；②剔除财务信息、审计费用等数据缺失的公司；③为了控制异常值对模型结果的影响，对样本中连续变量进行上下1%分位缩尾处理。

为了分析不同规模会计师事务所的审计溢价情况，将其划分为国际“四大”所、国内“六大”所、国内小所三类。国际“四大”所是指普华永道、毕马威、德勤、安永在国内的合作所，国内“六大”所是指按照中注协每年发布的综合评价前百名会计师事所中除“四大”所之外的排名在前十名的所，国内小所是指在排名前百名以后的一般会计师事务所。选取的最终样本为8 860家公司，其中“四大”所共审计了487家，国内“六大”所审计了3 320家，其他小所审计了5 053家。

2. 变量定义及描述性统计。参考伍利娜（2003）、陈冬华（2006）、吴应宇（2008）、郭梦岚（2009）的研究成果，选择影响审计费用的变量及其描述性统计如表1所示：

表1　变量定义及描述性统计

变量名	变量定义	样本数	平均值	标准差	最小值	最大值
lnFee	审计费用自然对数	8 860	13.284	0.639	10.714	17.517
lnTA	年末总资产对数	8 860	21.646	1.248	15.376	27.301
large1	股权集中度（第一大股东持股比例）	8 860	36.562	15.591	0.140	99.000
lev	财务杠杆（资产负债率）	8 860	0.456	0.214	0.007	0.999
soe	最终控制人有国有性质为1，否则为0	8 860	0.510	0.500	0	1
loss	净利润小于0为1，否则为0	8 860	0.086	0.281	0	1
turnover	营运状况（资产周转率）	8 860	0.700	0.621	0.000	15.896
opinion	审计意见为非标准意见为1，否则为0	8 860	0.035	0.184	0.000	1.000
InvnAssets	存货占总资产比	8 860	0.173	0.160	0.000	0.940
quick	速动比率	8 860	2.135	5.144	0.019	179.578
ROE	净资产收益率	8 860	0.163	7.805	-72.146	713.204
st	st公司为1，否则为0	8 860	0.057	0.231	0	1
eastregion	公司注册地位于东部地区则为1，其余为0	8 860	0.671	0.470	0	1
cenregion	公司注册地位于中部地区则为1，其余为0	8 860	0.876	0.330	0	1
big4	四大审计为1，非四大审计为0	8 860	0.055	0.228	0	1
big4mid6	四大审计为1，国内六大审计为0	3 807	0.128	0.334	0	1
mid6small	国内六大审计为1，小事务所审计为0	8 373	0.397	0.489	0	1
big4small	四大审计为1，小事务所审计为0	5 540	0.088	0.283	0	1

五、结果分析

1. OLS估计结果分析。本文先使用常规的审计费用模型来比较三类所审计费用溢价情况：

$$lnfee=\alpha_0+\beta_1 lnTA+\beta_2 large1+\beta_3 lev+\beta_4 soe+\beta_{15} loss+\beta_6 turnover+\beta_7 opinion+\beta_8 InvnAssets+\beta_9 quick+\beta_{10} ROE+\beta_{11} st+\beta_{12} eastregion+\beta_{13} cenregion+\tau DUM+\varepsilon \quad (9)$$

式（9）中变量含义见表1，DUM分别表示变量big4、big4mid6、mid6small、big4small的结果，如表2所示，模型控制了年度和行业，并考虑异方差稳健估计，同时还以Logit回归结果表示不同企业在选择国际“四大”所、国内“六大”所、国内小所三类所审计时的差异。

表 2　　OLS审计模型和Logit事务所选择模型

	四大所与非四大所		四大所与国内六大所		国内六大所与小所		四大所与小所	
	OLS(1)	logit(2)	OLS(3)	logit(4)	OLS(5)	logit(6)	OLS(7)	logit(8)
lnTA	0.338***	1.233***	0.376***	1.110***	0.317***	0.230***	0.330***	1.387***
	(56.10)	(23.41)	(41.00)	(19.14)	(52.70)	(9.36)	(43.33)	(22.88)
large1	-0.001**	0.003	-0.001***	0.000	-0.000	0.003**	-0.001***	0.008**
	(-2.26)	(0.79)	(-2.99)	(0.03)	(-1.00)	(2.14)	(-2.76)	(2.16)
lev	-0.025	-3.374***	-0.003	-3.636***	-0.005	-0.639***	-0.057*	-2.938***
	(-0.92)	(-7.98)	(-0.07)	(-7.58)	(-0.18)	(-4.41)	(-1.66)	(-6.73)
soe	-0.099***	0.002	-0.122***	0.197	-0.090***	-0.227***	-0.095***	-0.161
	(-10.66)	(0.02)	(-7.97)	(1.44)	(-9.72)	(-4.52)	(-8.15)	(-1.21)
loss	0.040**	0.270	0.041	0.072	0.026*	0.115	0.055***	0.369
	(2.56)	(1.21)	(1.52)	(0.29)	(1.67)	(1.33)	(2.94)	(1.55)
turnover	0.083***	0.140*	0.064***	0.088	0.080***	0.121***	0.096***	0.159*
	(8.38)	(1.84)	(5.30)	(1.07)	(8.08)	(3.29)	(5.94)	(1.79)
opinion	0.104***	0.975**	0.069*	1.124***	0.105***	0.173	0.114***	1.077***
	(4.35)	(2.53)	(1.67)	(2.70)	(4.39)	(1.25)	(3.91)	(2.73)
InvnAssets	-0.058*	-1.098***	-0.086	-1.302***	-0.074**	-0.043	-0.025	-1.388***
	(-1.78)	(-3.07)	(-1.56)	(-3.33)	(-2.29)	(-0.28)	(-0.64)	(-3.58)
quick	-0.004***	-0.309***	-0.003***	-0.408***	-0.004***	0.003	-0.004***	-0.220***
	(-4.70)	(-4.67)	(-2.86)	(-5.37)	(-4.72)	(0.61)	(-2.85)	(-3.70)
ROE	0.001***	-0.033	0.001***	-0.053	0.001***	0.005	0.002	-0.041
	(6.81)	(-0.25)	(8.76)	(-0.21)	(5.68)	(1.16)	(0.97)	(-0.44)
st	0.156***	-0.451	0.167***	-0.454	0.136***	-0.011	0.175***	-0.453
	(7.89)	(-0.93)	(4.48)	(-0.90)	(6.92)	(-0.10)	(7.53)	(-0.92)
eastregion	0.119***	0.757***	0.165***	0.791***	0.118***	0.326***	0.078***	0.762***
	(11.05)	(4.98)	(8.52)	(4.79)	(11.05)	(5.50)	(6.02)	(4.69)
cenregion	0.059***	0.045	0.026	0.231	0.077***	-0.560***	0.079***	-0.179
	(4.20)	(0.19)	(1.10)	(0.91)	(5.54)	(-6.92)	(4.39)	(-0.71)
big4	0.745***							
	(-26.46)							
big4mid6			0.629***					
			(-21.23)					
mid6smal					0.091***			
					(-10.24)			
big4small							0.806***	
							(-27.94)	
Year	控制		控制		控制		控制	
industry	控制		控制		控制		控制	
_cons	5.805***	-29.072***	5.248***	-25.337***	6.191***	-4.925***	5.921***	-32.142***
	(46.70)	(-25.89)	(27.77)	(-20.67)	(49.84)	(-9.87)	(37.36)	(-24.59)
r2	0.616		0.690		0.496		0.660	
chi2		1 096.949			845.049		186.017	1 134.539
N	8 860	8 860	3 807	3 807	8 373	8 373	5 540	5 540

注:*、**、***分别表示统计值在10%、5%、1%的水平上显著,括弧中数据为标准差。

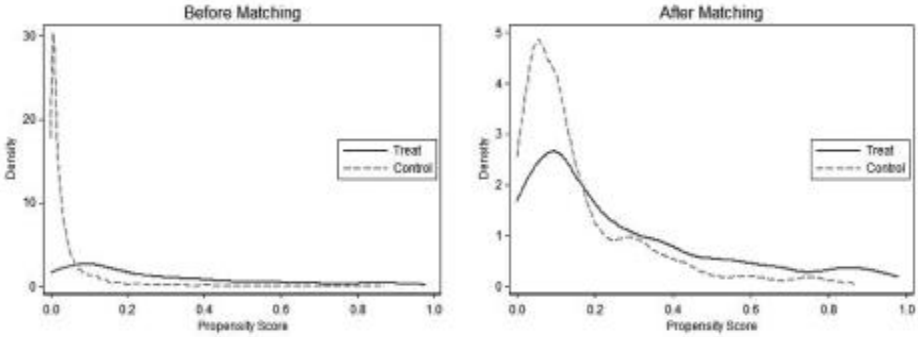
由表2可以看出,在控制其他变量时,会计师事务所变量DUM的系数均显著,big4的回归系数为0.745,表明国际“四大”所比非“四大”所有110.6%的溢价;big4mid6的系数为0.629,即国际“四大”所比国内“六大”所有87.6%的溢价;mid6small的系数为0.091,即国内“六大”所比国内小所有9.5%的溢价;而big4small的系数为0.806,即国际“四大”所比国内小所有123.9%的溢价。

用于会计师事务所选择的Logit模型包含的变量和OLS模型相同,结果显示,企业对选择“四大”所、国内“六大”所、小所审计时有一定倾向差异。总体来说公司规模、资产负债率、非标准审计意见、速动比率、公司注册地位于东部地区等变量影响了公司选择不同会计师事务所进行审计,规模较大的企业、注册地位于东部地区、资产负债率较低、资产流动性较好的的企业倾向于选择国际“四大”所进行审计,而具有第一大股东持股比例较高、资产周转率较快、大股东具有国有性质、公司所在地位于中部地区等特征的企业倾向于选择国内“六大”所进行审计。

2. 倾向得分匹配估计分析。下面以国际“四大”所和非国际“四大”所样本为例,说明利用最近邻匹配法进行倾向得分匹配的过程。

(1)计算客户选择国际“四大”所审计的倾向得分值。下图分别呈现了处理组和控制组在匹配前后的倾向得分值的核密度函数,从图中可以看出,二者分布存在明显差异,控制组非国际“四大”所审计的企业集中在倾向得分较低的部分(大多位于0.2以下),而国际“四大”

所审计的企业相对分布较平稳，因此直接比较这两组样本公司之间的审计费用差异，所得到的统计结果会有偏的。同时倾向得分显示，两组之间有很大的重叠区，符合共同支撑假设要求。相比之下，在完成匹配后，两组样本倾向得分值的概率分布已经比较接近，匹配效果较好。



处理组和控制组在匹配前（左图）、匹配后（右图）的核密度

（2）利用最近邻匹配法进行样本配对。进行样本匹配时设置强制共同支撑（**Common support**），即在处理组中倾向得分值高于控制组中最大值或者低于最小值的观察对象会被丢弃，最终匹配成功**8 660**个样本。

（3）检查**PSM**匹配结果是否满足平行假设。**Oakes**和**Kaufman**（**2006**）认为，匹配后的变量标准偏差低于**10%**是可以接受的。从表**3**中可以看出，匹配后各变量在处理组和控制组间不再存在显著差异，除了变量**eastregion**为**7%**外，其余变量的偏差都小于**5%**，匹配结果较好。

表 3　　匹配前后变量的差异对比

Variable	Unmatche	Mean		%reduct		t-test	
	Matched	Treated	Control	%bias	|bias|	t	p>|t|
lnTA	U	23.453	21.541	147.6		35.06	0.000
	M	23.286	23.333	-3.6	97.6	-0.57	0.572
quick	U	1.038 4	2.198 6	-30.3		-4.84	0.000
	M	1.054 2	1.108 8	-1.4	95.3	-0.69	0.489
opinion	U	0.0184 8	0.036 07	-10.8		-2.05	0.040
	M	0.0195 2	0.018 44	0.7	93.8	0.12	0.904
eastregion	U	0.8069 8	0.663 32	33.0		6.58	0.000
	M	0.796 1	0.765 73	7.0	78.9	1.11	0.265
turnover	U	0.7579 5	0.696 83	10.6		2.11	0.035
	M	0.7531 2	0.758 51	-0.9	91.2	-0.14	0.888
InvnAssets	U	0.1577 7	0.174 4	-10.5		-2.22	0.026
	M	0.1591 1	0.155 76	2.1	79.9	0.31	0.754
lev	U	0.5153 4	0.452 4	31.6		6.32	0.000
	M	0.512 4	0.522 66	-5.2	83.7	-0.85	0.395

（4）倾向得分匹配估计结果分析。表**4**列出了**PSM**匹配结果**ATT**值、模型对应的**Rosenbaum**边界分析的**Hodges-Lehmann**估计值和Γ临界值、**OLS**回归结果。从表**4**可以看出，**ATT**值和**HL**值非常相近，表明匹配的样本没有极端异常值。在不考虑未匹配样本的情况下，**ATT**值及**HL**值都显示，国际“四大”所和国内“六大”所存在显著的审计溢价。总体来说，国际“四大”所比非“四大”所收取了较高的审计溢价（约**88%**），但低于**OLS**回归结果的**110.64%**，说明在控制样本自选择后，审计溢价程度降低，**OLS**回归会高估审计溢价。另外，与**OLS**结果一致，国内“六大”所对小所溢价（**11%~12%**）比“四大”所对国内“六大”所存在的溢价（**73%~74%**）、“四大”所对小所溢价（约**103%**）程度低得多。

鉴于**OLS**结果提供了审计溢价的全样本（存在选择偏误问题）平均估计值，而**PSM**估值时样本则限制在共同支撑区域，因而三种结果只在一定样本范围具有可比性。从表**4**还可以看出，通过**PSM**匹配控制观察到的偏误后，仍然提供了有力的证据，支持如下结论：国际“四大”所相对具有相似特征的国内“六大”所收取了较高溢价，后者相对于所匹配的小所收取了较低溢价。

表**4**还列出了**Rosenbaum**边界临界Γ值，当它导致处理效果在**5%**水平上不显着。汇总数据显示，模型**1**（“四大”所与非“四大”所）中只有未观测到的变量导致客户在两类事务所之间出现选择偏误的Γ值在约**6.2**时，超过**88%**的溢价估计才会出现统计上不显著（模型**2**和模型**4**也能得到类似的估计）。而国内“六大”所对国内小所具有较小的审计溢价（**11%~12%**），并且由于临界Γ值为**1.4**，表明结论相对未观察到的混淆因素影响更为敏感。

表 4　　PSM结果对比分析

参数估计	四大所与非四大所 模型1	四大所与国内六大所 模型2	国内六大所与小所 模型3	四大所与小所 模型4
处理组均值	14.515	14.523	13.316	14.458
控制组均值	13.882	13.966	13.205	13.746
ATT估计值	0.632	0.557	0.111	0.711
HL估计值	0.629	0.549	0.101	0.711
OLS估计值	0.745	0.629	0.091	0.806
审计溢价（ATT）	88.14%	74.54%	11.74%	103.60%
审计溢价（HL）	87.57%	73.15%	10.63%	103.60%
审计溢价（OLS）	110.64%	87.57%	9.53%	123.89%
RBs临界Γ值	6.2	5.0	1.4	8.8
HL临界Γ值	6.0	4.8	1.39	8.5

注：**HL**估计值为**Hodges-Lehmann**估计，审计溢价为对应值的指数值，**Rosenbaum**Γ临界值为在**5%**水平时的Γ值。

3. **Rosenbaum** 边界敏感性分析。表 **5** 以模型 **3** 为例，列示了 **PSM** 估计后的 **Rosenbaum** 边界分析值，其中 **sig+**、**sig-** 是 **Wilcoxon** 符号秩检验值，**t-hat+**、**t-hat-**、**CI+**、**CI-** 是 **Hodges-Lehmann** 点估计值。

表 5　Rosenbaum 边界分析值（国内"六大"所与小所）

Rosenbaum bounds for delta（N=3 318 matched pairs）						
Gamma	sig+	sig-	t-hat+	t-hat-	CI+	CI-
1.0	0	0	0.100 903	0.100 903	0.080 956	0.120 751
1.1	2.3e-14	0	0.076 836	0.125 048	0.056 765	0.145 005
1.2	3.4e-08	0	0.054 773	0.147 051	0.034 833	0.167 438
1.3	0.000 30	0	0.034 644	0.167 693	0.014 898	0.188 082
1.4	0.053 00	0	0.016 224	0.186 659	-0.003 609	0.207 295
1.5	0.528 10	0	-0.000 77	0.204 43	-0.020 838	0.225 114
1.6	0.950 39	0	-0.016 785	0.220 916	-0.036 774	0.241 734
1.7	0.999 13	0	-0.031 616	0.23 628	-0.051 884	0.257 466
1.8	0.999 99	0	-0.045 581	0.250 994	-0.065 978	0.272 296
1.9	1	0	-0.058 891	0.264 848	-0.079 213	0.286 342
2.0	1	0	-0.071 4	0.278 076	-0.091 811	0.299 822
* gamma - log odds of differential assignment due to unobserved factors						
sig+ - upper bound significance level						
sig- - lower bound significance level						
t-hat+ - upper bound Hodges-Lehmann point estimate						
t-hat- - lower bound Hodges-Lehmann point estimate						
CI+ - upper bound confidence interval（a=0.95）						
CI- - lower bound confidence interval（a=0.95）						

在模型 **3** 中，当临界Γ值介于 **1.39** 到 **1.4** 时，处理效果结论在 **5%** 水平上不显著，这说明影响企业是否选择"六大"所审计的未观察到变量仅仅增加 **40%** 会导致审计溢价在 **5%** 的水平上不显著。这时未观察到的变量对事务所选择的影响很重要。**Hodges-Lehmann** 点估计的 **95%** 置信区间包含了零（**HL** 估计值为 **0.018**），表明审计费用溢价接近于零。因此与表 **2** 中配对变量是否选择"四大"所的 **logit** 回归系数相比较，此数值大约是资产负债率（**lev**）和公司所在地在中部地区（**cenregion**）系数的 **0.639** 和 **0.560**。也就是说，未观察到变量 **u** 对事务所的选择影响力至少要像资产负债率一样大，才会影响到 **ATT** 的结论。

表 **4** 中模型 **1**、模型 **2**、模型 **4** 的Γ临界值非常高，**95%** 置信区间分别为 **6.2**、**5.0**、**8.8**，**Hodges-Lehmann** 点估计分别为 **6.0**、**4.8**、**8.5**，意味着未观察到变量对事务所选择的影响力要达到这么高才会影响 **ATT** 值的变化。**Duvendack** 和 **Palmer-Jones**（**2011**）认为，临界值大于 **2** 时可以得出未观测到变量不敏感的结论。并且这样的敏感性分析描述了一个"最坏的情况"，只是说明处理效应的估计可能通过隐藏偏误而改变，但这并不表示这些偏误是一定存在的（**Becker** 和 **Caliendo**，**2007**）。因此我们可以合理地得出结论，这三个模型的 **ATT** 结果对未观察到变量是不敏感的，估计审计费用溢价的结果是稳健的。

为进一步证实结果的稳健性，本文还检验了不同匹配算法和不同倾向得分计算方式的可靠性。在匹配方法改为核匹配和半径匹配法时，结果显示审计溢价仍然存在，而且比最近邻匹配方法计算的结果更高（**ATT** 分别为 **0.661** 和 **0.683**）。并且不论用 **Probit** 还是用 **Logit** 模型计算倾向得分，结论也不没改变，国际"四大"所、国内"六大"所、小所之间显著存在着不同程度的审计溢价。

六、结论

为了控制样本选择偏误，本文采用倾向得分匹配法来分析国内会计师事务所审计溢价情况。与线性回归结论相比，在国际"四大"所、国内"六大"所、国内小所之间仍然存在审计费用溢价情况，在控制选择偏误影响后，总体溢价程度降低，比如国际"四大"所相对非"四大"所只取得 **88%** 的审计溢价，只有国内"六大"所相对国内小所获得 **11%** 溢价，略高于回归估计溢价水平。

我们还通过计算 **Rosenbaum** 边界分析来衡量未观测到变量对处理效应误差的影响程度，结果表明，国际"四大"所的审计溢价结果对未观测到变量不敏感，而国内"六大"所相对小所的溢价则容易受未观测到变量的影响。总体来说，模型的 **ATT** 结果是稳健的，为会计师事务所审计溢价的存在提供了新的证据。

主要参考文献

Craswell A. T.，Francis J. R.，Taylor S. L.. Auditor brand name reputations and industry specializations［J］**. Journal of accounting and economics，1995（20）.**

Peel M. J.，Roberts R.. Audit fee determinants and auditor premiums: Evidence from the micro-firm sub-market［J］**. Accounting and Business Research，2003（33）.**

耿建新，房巧玲.国际四大所与我国本土大所审计收费比较研究——来自我国证券审计市场的初步证据［J］. 当代财经，**2006（1）.**

Rosenbaum，P.R.，'Hodges-Lehmann Point Estimates of Treatment Effect in Observational Studies'，Journal of the American Statistical Association，1993（88）.

Francis J. R.. The effect of audit firm size on audit prices: A study of the Australian market［J］**. Journal of accounting and Economics，1984（6）.**

Palmrose Z. V.. Audit fees and auditor size: Further evidence［J］**. Journal of accounting research，1986.**

Pong C. M.，Whittington G.. The determinants of audit fees: Some empirical models［J］**. Journal of Business Finance & Accounting，1994（21）.**

伍丽娜.盈余管理对审计费用影响的分析——来自中国上市公司首次审计费用披露的证据［J］.会计研究，**2003（12）.**