

附表 1

样本年度分布

年份	样本量	占比 (%)
2013	88827	22.151
2014	93757	23.380
2015	82163	20.489
2016	64027	15.967
2017	72234	18.013
总计	401008	100

附表 2

变量选择说明

分类	变量	变量类型	分类	变量	变量类型
个人特征	年份	连续	家庭信息	家庭包吃住人数	连续
	性别	分类		家庭月住房支出	连续
	民族	分类		家庭月食品支出	连续
	教育	分类		家庭月总支出	连续
	户口	分类		家庭月总收入	连续
	党员	分类		家庭规模	连续
	婚姻	分类		孩子数量	连续
	健康状况	分类		老人数量	连续
工作信息	周工作小时	连续	流动信息	老家位置	分类
	职业	分类		流动范围	分类
	行业	分类		流动城市数量	连续
	单位性质	分类	地区信息	迁入意愿	分类
	劳动合同	分类		流动省份	分类
	工作年限	连续		户口省份	分类
	月收入(结果变量)	连续			
	月收入变化	分类			
	医保参与	分类			
	社保参与	分类			

附表 3

不同人口属性群体的劳动收入风险

人口特征	分组	工资收入	SD	SK	KU	样本量
性别	女性	460.983	106.523	0.863	13.836	172698
	男性	597.613	127.738	0.163	11.238	228310
教育	高中及其以下	514.643	113.022	0.493	12.365	318840
	大专和本科	629.704	140.645	0.384	12.314	76734
	研究生	670.442	134.707	-0.067	12.488	5434
城乡	城镇	607.699	137.860	0.462	11.279	69349
	农村	524.359	114.575	0.465	12.582	331659

附表 4

不同工作特征群体的劳动收入风险

工作属性	分组	工资收入	SD	SK	KU	样本量
行业 ^①	第一产业（农、林、牧、渔）	487.946	91.855	0.002	15.078	35454
	第二产业（高技术制造业）	680.008	160.562	0.727	9.531	10609
	第二产业（其他制造业）	558.756	118.091	0.537	10.331	85336
	第二产业（非制造业）	636.458	138.161	0.075	9.508	45891
	第三产业（金融、信息传输、软件和信息技术服务）	661.129	147.017	0.252	10.435	19174
	第三产业（其他）	498.532	114.222	0.607	13.696	204544
单位性质	机关事业单位	529.929	134.327	0.863	10.463	10181
	国有、国有控股和集体企业	578.186	131.777	0.488	10.561	39942
	私营企业	530.794	116.441	0.456	12.636	313112
	外资和港澳台企业（独资和合资）	586.860	118.866	0.353	12.773	24993
	其他（社团民办组织等）	524.051	117.311	0.505	11.829	12780

①高技术制造业包括化学制品加工、医药制造、专用设备制造、交通运输设备制造、电器机械及制造、计算机及通讯电子设备制造和仪器仪表制造业。

附表 5

平均风险补偿估计（2017 年样本）

变量	系数	标准误	T 值	P 值	2.5%CI	97.5%CI	2.5%CI_adj	97.5%CI_adj	P 值_adj	机器学习模型	得分函数
<i>SD</i>	2.209	0.022	100.143	0.000	2.166	2.252	2.157	2.261	0.000	LassoCV	IV-Type
<i>SK</i>	-118.954	0.838	-141.901	0.000	-120.597	-117.311	-120.935	-116.973	0.000	LassoCV	IV-Type
<i>KU</i>	7.666	0.212	36.198	0.000	7.251	8.081	7.165	8.167	0.000	LassoCV	IV-Type
<i>SD</i>	2.140	0.024	88.784	0.000	2.092	2.187	2.084	2.196	0.000	LassoCV	Partialling Out
<i>SK</i>	-122.192	0.975	-125.371	0.000	-124.102	-120.282	-124.452	-119.932	0.000	LassoCV	Partialling Out
<i>KU</i>	7.808	0.231	33.849	0.000	7.355	8.260	7.273	8.343	0.000	LassoCV	Partialling Out
<i>SD</i>	0.873	0.022	40.068	0.000	0.830	0.915	0.821	0.924	0.000	RF	IV-Type
<i>SK</i>	-138.041	0.994	-138.913	0.000	-139.988	-136.093	-140.404	-135.677	0.000	RF	IV-Type
<i>KU</i>	7.726	0.165	46.826	0.000	7.403	8.049	7.334	8.118	0.000	RF	IV-Type
<i>SD</i>	0.787	0.031	25.229	0.000	0.726	0.848	0.717	0.858	0.000	RF	Partialling Out
<i>SK</i>	-126.562	1.079	-117.289	0.000	-128.677	-124.447	-128.998	-124.126	0.000	RF	Partialling Out
<i>KU</i>	6.750	0.128	52.587	0.000	6.498	7.002	6.460	7.040	0.000	RF	Partialling Out
<i>SD</i>	0.831	0.024	35.238	0.000	0.785	0.878	0.776	0.887	0.000	GBDT	IV-Type
<i>SK</i>	-135.351	0.824	-164.340	0.000	-136.966	-133.737	-137.295	-133.408	0.000	GBDT	IV-Type
<i>KU</i>	9.488	0.131	72.626	0.000	9.232	9.745	9.180	9.797	0.000	GBDT	IV-Type
<i>SD</i>	0.774	0.030	26.222	0.000	0.716	0.832	0.705	0.843	0.000	GBDT	Partialling Out
<i>SK</i>	-130.428	0.902	-144.545	0.000	-132.196	-128.659	-132.530	-128.325	0.000	GBDT	Partialling Out
<i>KU</i>	9.110	0.129	70.577	0.000	8.857	9.363	8.809	9.411	0.000	GBDT	Partialling Out
<i>SD</i>	1.081	0.035	30.737	0.000	1.012	1.150	0.997	1.165	0.000	MLPR	IV-Type
<i>SK</i>	-118.494	0.829	-142.968	0.000	-120.119	-116.870	-120.470	-116.518	0.000	MLPR	IV-Type
<i>KU</i>	8.008	0.117	68.541	0.000	7.779	8.237	7.730	8.287	0.000	MLPR	IV-Type
<i>SD</i>	0.895	0.029	31.338	0.000	0.839	0.951	0.829	0.960	0.000	MLPR	Partialling Out
<i>SK</i>	-75.271	0.636	-118.326	0.000	-76.517	-74.024	-76.729	-73.812	0.000	MLPR	Partialling Out
<i>KU</i>	6.810	0.119	57.173	0.000	6.576	7.043	6.537	7.083	0.000	MLPR	Partialling Out

附表 6

平均风险补偿估计 (2016 年样本)

变量	系数	标准误	T 值	P 值	2.5%CI	97.5%CI	2.5%CI_adj	97.5%CI_adj	P 值_adj	机器学习模型	得分函数
<i>SD</i>	2.012	0.027	75.423	0.000	1.959	2.064	1.949	2.075	0.000	LassoCV	IV-Type
<i>SK</i>	-108.089	1.288	-83.896	0.000	-110.614	-105.564	-111.121	-105.057	0.000	LassoCV	IV-Type
<i>KU</i>	6.183	0.260	23.795	0.000	5.674	6.692	5.572	6.795	0.000	LassoCV	IV-Type
<i>SD</i>	2.005	0.028	71.187	0.000	1.949	2.060	1.937	2.072	0.000	LassoCV	Partialling Out
<i>SK</i>	-111.005	1.665	-66.683	0.000	-114.268	-107.742	-114.996	-107.014	0.000	LassoCV	Partialling Out
<i>KU</i>	6.272	0.278	22.598	0.000	5.728	6.816	5.607	6.938	0.000	LassoCV	Partialling Out
<i>SD</i>	0.725	0.028	26.316	0.000	0.671	0.778	0.663	0.786	0.000	RF	IV-Type
<i>SK</i>	-140.251	1.601	-87.615	0.000	-143.388	-137.114	-143.853	-136.649	0.000	RF	IV-Type
<i>KU</i>	8.124	0.403	20.147	0.000	7.334	8.915	7.217	9.032	0.000	RF	IV-Type
<i>SD</i>	0.679	0.042	16.073	0.000	0.596	0.762	0.579	0.779	0.000	RF	Partialling Out
<i>SK</i>	-128.588	1.800	-71.442	0.000	-132.115	-125.060	-132.859	-124.316	0.000	RF	Partialling Out
<i>KU</i>	7.051	0.470	14.988	0.000	6.129	7.972	5.934	8.167	0.000	RF	Partialling Out
<i>SD</i>	0.788	0.028	27.882	0.000	0.732	0.843	0.720	0.856	0.000	GBDT	IV-Type
<i>SK</i>	-138.858	1.034	-134.327	0.000	-140.884	-136.832	-141.343	-136.373	0.000	GBDT	IV-Type
<i>KU</i>	9.881	0.330	29.987	0.000	9.236	10.527	9.089	10.674	0.000	GBDT	IV-Type
<i>SD</i>	0.726	0.033	21.782	0.000	0.661	0.792	0.647	0.806	0.000	GBDT	Partialling Out
<i>SK</i>	-134.765	1.083	-124.451	0.000	-136.887	-132.643	-137.355	-132.174	0.000	GBDT	Partialling Out
<i>KU</i>	9.727	0.337	28.869	0.000	9.067	10.387	8.921	10.533	0.000	GBDT	Partialling Out
<i>SD</i>	1.214	0.097	12.498	0.000	1.024	1.404	0.984	1.444	0.000	MLPR	IV-Type
<i>SK</i>	-214.475	4.670	-45.928	0.000	-223.628	-205.322	-225.524	-203.426	0.000	MLPR	IV-Type
<i>KU</i>	8.491	0.966	8.787	0.000	6.597	10.385	6.205	10.777	0.000	MLPR	IV-Type
<i>SD</i>	0.940	0.033	28.285	0.000	0.875	1.005	0.862	1.019	0.000	MLPR	Partialling Out
<i>SK</i>	-49.233	0.478	-102.906	0.000	-50.171	-48.296	-50.365	-48.102	0.000	MLPR	Partialling Out
<i>KU</i>	5.789	0.737	7.852	0.000	4.344	7.234	4.046	7.532	0.000	MLPR	Partialling Out

附表 7

平均风险补偿估计 (2015 年样本)

变量	系数	标准误	T 值	P 值	2.5%CI	97.5%CI	2.5%CI_adj	97.5%CI_adj	P 值_adj	机器学习模型	得分函数
<i>SD</i>	2.080	0.020	103.414	0.000	2.041	2.119	2.032	2.128	0.000	LassoCV	IV-Type
<i>SK</i>	-80.113	0.581	-137.894	0.000	-81.252	-78.974	-81.499	-78.727	0.000	LassoCV	IV-Type
<i>KU</i>	5.348	0.092	58.105	0.000	5.168	5.528	5.128	5.568	0.000	LassoCV	IV-Type
<i>SD</i>	2.112	0.021	100.195	0.000	2.070	2.153	2.060	2.163	0.000	LassoCV	Partialling Out
<i>SK</i>	-81.808	0.612	-133.749	0.000	-83.007	-80.609	-83.301	-80.316	0.000	LassoCV	Partialling Out
<i>KU</i>	5.425	0.099	54.721	0.000	5.230	5.619	5.183	5.667	0.000	LassoCV	Partialling Out
<i>SD</i>	0.870	0.024	36.932	0.000	0.824	0.916	0.814	0.926	0.000	RF	IV-Type
<i>SK</i>	-75.815	0.859	-88.261	0.000	-77.498	-74.131	-77.862	-73.768	0.000	RF	IV-Type
<i>KU</i>	2.990	0.063	47.562	0.000	2.867	3.113	2.840	3.140	0.000	RF	IV-Type
<i>SD</i>	0.783	0.033	24.058	0.000	0.719	0.847	0.710	0.856	0.000	RF	Partialling Out
<i>SK</i>	-69.710	0.920	-75.799	0.000	-71.512	-67.907	-71.776	-67.644	0.000	RF	Partialling Out
<i>KU</i>	2.681	0.056	47.485	0.000	2.570	2.792	2.554	2.808	0.000	RF	Partialling Out
<i>SD</i>	0.803	0.023	35.185	0.000	0.758	0.848	0.745	0.860	0.000	GBDT	IV-Type
<i>SK</i>	-82.470	0.626	-131.640	0.000	-83.698	-81.242	-84.050	-80.889	0.000	GBDT	IV-Type
<i>KU</i>	4.159	0.064	64.649	0.000	4.033	4.285	3.996	4.321	0.000	GBDT	IV-Type
<i>SD</i>	0.710	0.028	25.573	0.000	0.656	0.765	0.647	0.774	0.000	GBDT	Partialling Out
<i>SK</i>	-79.562	0.690	-115.322	0.000	-80.915	-78.210	-81.143	-77.982	0.000	GBDT	Partialling Out
<i>KU</i>	4.039	0.069	58.462	0.000	3.903	4.174	3.880	4.197	0.000	GBDT	Partialling Out
<i>SD</i>	1.081	0.039	27.884	0.000	1.005	1.157	0.990	1.172	0.000	MLPR	IV-Type
<i>SK</i>	-75.144	0.741	-101.421	0.000	-76.596	-73.692	-76.882	-73.406	0.000	MLPR	IV-Type
<i>KU</i>	3.203	0.063	50.562	0.000	3.079	3.327	3.054	3.352	0.000	MLPR	IV-Type
<i>SD</i>	0.824	0.029	28.865	0.000	0.769	0.880	0.755	0.894	0.000	MLPR	Partialling Out
<i>SK</i>	-56.772	0.495	-114.740	0.000	-57.742	-55.802	-57.971	-55.573	0.000	MLPR	Partialling Out
<i>KU</i>	2.882	0.062	46.388	0.000	2.760	3.004	2.731	3.033	0.000	MLPR	Partialling Out

附表 8

平均风险补偿估计(2014 年样本)

变量	系数	标准误	T 值	P 值	2.5%CI	97.5%CI	2.5%CI_adj	97.5%CI_adj	P 值_adj	机器学习模型	得分函数
<i>SD</i>	1.955	0.021	92.895	0.000	1.914	1.996	1.907	2.002	0.000	LassoCV	IV-Type
<i>SK</i>	-50.627	0.390	-129.752	0.000	-51.392	-49.863	-51.509	-49.745	0.000	LassoCV	IV-Type
<i>KU</i>	2.922	0.042	69.837	0.000	2.840	3.004	2.827	3.016	0.000	LassoCV	IV-Type
<i>SD</i>	1.873	0.022	84.488	0.000	1.829	1.916	1.821	1.925	0.000	LassoCV	Partialling Out
<i>SK</i>	-51.747	0.415	-124.699	0.000	-52.560	-50.934	-52.719	-50.775	0.000	LassoCV	Partialling Out
<i>KU</i>	2.955	0.045	66.412	0.000	2.868	3.043	2.851	3.060	0.000	LassoCV	Partialling Out
<i>SD</i>	0.705	0.026	27.162	0.000	0.654	0.756	0.646	0.764	0.000	RF	IV-Type
<i>SK</i>	-43.943	0.473	-92.854	0.000	-44.870	-43.015	-45.018	-42.868	0.000	RF	IV-Type
<i>KU</i>	1.496	0.027	56.058	0.000	1.443	1.548	1.435	1.556	0.000	RF	IV-Type
<i>SD</i>	0.640	0.038	16.984	0.000	0.566	0.714	0.551	0.729	0.000	RF	Partialling Out
<i>SK</i>	-39.865	0.513	-77.710	0.000	-40.870	-38.859	-41.080	-38.650	0.000	RF	Partialling Out
<i>KU</i>	1.361	0.026	52.183	0.000	1.310	1.412	1.299	1.423	0.000	RF	Partialling Out
<i>SD</i>	0.778	0.023	34.087	0.000	0.733	0.822	0.721	0.835	0.000	GBDT	IV-Type
<i>SK</i>	-50.737	0.374	-135.682	0.000	-51.470	-50.004	-51.671	-49.803	0.000	GBDT	IV-Type
<i>KU</i>	2.491	0.036	69.496	0.000	2.421	2.562	2.402	2.581	0.000	GBDT	IV-Type
<i>SD</i>	0.703	0.029	24.246	0.000	0.646	0.760	0.634	0.772	0.000	GBDT	Partialling Out
<i>SK</i>	-49.035	0.413	-118.757	0.000	-49.844	-48.225	-50.019	-48.050	0.000	GBDT	Partialling Out
<i>KU</i>	2.442	0.038	63.615	0.000	2.367	2.518	2.351	2.534	0.000	GBDT	Partialling Out
<i>SD</i>	2.817	0.973	2.896	0.004	4.723	0.911	5.141	0.492	0.011	MLPR	IV-Type
<i>SK</i>	-52.841	0.598	-88.378	0.000	-54.012	-51.669	-54.270	-51.412	0.000	MLPR	IV-Type
<i>KU</i>	1.424	0.030	47.859	0.000	1.366	1.482	1.353	1.495	0.000	MLPR	IV-Type
<i>SD</i>	0.445	0.029	15.565	0.000	0.389	0.501	0.378	0.511	0.000	MLPR	Partialling Out
<i>SK</i>	-28.540	0.297	-96.120	0.000	-29.121	-27.958	-29.232	-27.848	0.000	MLPR	Partialling Out
<i>KU</i>	1.073	0.028	38.737	0.000	1.019	1.128	1.009	1.138	0.000	MLPR	Partialling Out

附表 9

平均风险补偿估计（2013 年样本）

变量	系数	标准误	T 值	P 值	2.5%CI	97.5%CI	2.5%CI_adj	97.5%CI_adj	P 值_adj	机器学习模型	得分函数
<i>SD</i>	1.821	0.022	81.247	0.000	1.777	1.865	1.768	1.873	0.000	LassoCV	IV-Type
<i>SK</i>	-29.257	0.248	-117.858	0.000	-29.744	-28.771	-29.840	-28.674	0.000	LassoCV	IV-Type
<i>KU</i>	1.126	0.029	38.321	0.000	1.068	1.184	1.057	1.195	0.000	LassoCV	IV-Type
<i>SD</i>	1.866	0.023	80.941	0.000	1.821	1.911	1.810	1.922	0.000	LassoCV	Partialling Out
<i>SK</i>	-29.657	0.259	-114.396	0.000	-30.165	-29.149	-30.287	-29.027	0.000	LassoCV	Partialling Out
<i>KU</i>	1.136	0.030	37.578	0.000	1.077	1.195	1.062	1.209	0.000	LassoCV	Partialling Out
<i>SD</i>	0.396	0.022	17.740	0.000	0.352	0.439	0.340	0.451	0.000	RF	IV-Type
<i>SK</i>	-23.905	0.384	-62.308	0.000	-24.657	-23.153	-24.852	-22.957	0.000	RF	IV-Type
<i>KU</i>	0.570	0.013	44.226	0.000	0.545	0.596	0.538	0.602	0.000	RF	IV-Type
<i>SD</i>	0.365	0.034	10.630	0.000	0.298	0.432	0.283	0.447	0.000	RF	Partialling Out
<i>SK</i>	-22.077	0.423	-52.162	0.000	-22.907	-21.248	-23.086	-21.069	0.000	RF	Partialling Out
<i>KU</i>	0.512	0.012	41.172	0.000	0.488	0.536	0.482	0.542	0.000	RF	Partialling Out
<i>SD</i>	0.786	0.023	34.514	0.000	0.742	0.831	0.734	0.839	0.000	GBDT	IV-Type
<i>SK</i>	-30.293	0.300	-100.947	0.000	-30.881	-29.704	-30.980	-29.605	0.000	GBDT	IV-Type
<i>KU</i>	1.182	0.023	51.760	0.000	1.137	1.227	1.129	1.234	0.000	GBDT	IV-Type
<i>SD</i>	0.702	0.029	24.580	0.000	0.646	0.758	0.635	0.768	0.000	GBDT	Partialling Out
<i>SK</i>	-29.616	0.336	-88.079	0.000	-30.275	-28.957	-30.403	-28.829	0.000	GBDT	Partialling Out
<i>KU</i>	1.150	0.024	47.517	0.000	1.103	1.198	1.093	1.207	0.000	GBDT	Partialling Out
<i>SD</i>	0.733	0.015	48.945	0.000	0.704	0.762	0.696	0.770	0.000	MLPR	IV-Type
<i>SK</i>	-29.103	0.372	-78.333	0.000	-29.832	-28.375	-30.017	-28.190	0.000	MLPR	IV-Type
<i>KU</i>	0.756	0.024	31.941	0.000	0.710	0.803	0.698	0.814	0.000	MLPR	IV-Type
<i>SD</i>	0.751	0.030	24.761	0.000	0.692	0.811	0.677	0.826	0.000	MLPR	Partialling Out
<i>SK</i>	-24.169	0.328	-73.692	0.000	-24.812	-23.527	-24.973	-23.366	0.000	MLPR	Partialling Out
<i>KU</i>	0.541	0.018	30.505	0.000	0.506	0.576	0.497	0.584	0.000	MLPR	Partialling Out

附表 10

异质性风险补偿估计（2017 年样本）

风险	系数	分位数	置信区间							
			normal_ci_l	normal_ci_u	percentile_ci_l	percentile_ci_u	pivotal_ci_l	pivotal_ci_u	normal_ci2_l	normal_ci2_u
SD	-0.869	0.1	-1.111	-0.628	-0.928	-0.805	-0.933	-0.811	-0.934	-0.805
SK	-58.318	0.1	-75.464	-41.171	-60.156	-56.258	-60.378	-56.479	-60.274	-56.361
KU	3.539	0.1	2.295	4.783	3.307	3.736	3.342	3.771	3.311	3.767
SD	0.249	0.2	0.074	0.424	0.172	0.338	0.160	0.326	0.163	0.335
SK	-85.349	0.2	-89.996	-80.702	-87.641	-83.153	-87.546	-83.057	-87.529	-83.169
KU	5.871	0.2	4.182	7.560	5.635	6.164	5.579	6.108	5.586	6.157
SD	2.067	0.4	2.015	2.119	2.000	2.119	2.015	2.134	2.009	2.125
SK	-112.142	0.4	-114.267	-110.016	-113.545	-110.800	-113.483	-110.739	-113.537	-110.746
KU	9.059	0.4	8.714	9.404	8.821	9.290	8.828	9.298	8.817	9.301
SD	2.380	0.5	2.354	2.406	2.348	2.408	2.352	2.412	2.346	2.414
SK	-111.991	0.5	-113.177	-110.805	-113.738	-110.830	-113.152	-110.244	-113.330	-110.653
KU	9.496	0.5	9.146	9.847	9.206	9.770	9.223	9.787	9.229	9.764
SD	2.511	0.6	2.473	2.549	2.486	2.543	2.480	2.536	2.483	2.540
SK	-111.118	0.6	-112.349	-109.888	-112.195	-109.793	-112.443	-110.041	-112.338	-109.899
KU	9.809	0.6	9.602	10.015	9.567	10.020	9.597	10.050	9.575	10.042
SD	2.729	0.8	2.702	2.756	2.702	2.757	2.701	2.756	2.700	2.757
SK	-107.597	0.8	-109.524	-105.669	-109.080	-106.194	-109.000	-106.113	-109.083	-106.110
KU	10.339	0.8	10.080	10.598	10.082	10.632	10.046	10.595	10.068	10.610
SD	2.839	0.9	2.735	2.942	2.807	2.871	2.806	2.870	2.806	2.872
SK	-104.754	0.9	-107.586	-101.921	-106.415	-103.556	-105.951	-103.092	-106.273	-103.234
KU	10.508	0.9	10.050	10.965	10.245	10.723	10.292	10.770	10.260	10.755

注：He 等(2023)提供了四种可用置信区间：1、normal confidence intervals(表中 normal_ci_l 和 normal_ci_u)；2、percentile confidence intervals(表中 percentile_ci_l 和 percentile_ci_u)；3、pivotal confidence intervals(表中 pivotal_ci_l 和 pivotal_ci_u)；4、bootstrap confidence intervals(表中 normal_ci2_l 和 normal_ci2_u)。下同。

附表 11

异质性风险补偿估计（2016 年样本）

风险	系数	分位数	置信区间							
			normal_ci_l	normal_ci_u	percentile_ci_l	percentile_ci_u	pivotal_ci_l	pivotal_ci_u	normal_ci2_l	normal_ci2_u
SD	-1.109	0.1	-1.392	-0.825	-1.172	-1.038	-1.180	-1.046	-1.176	-1.042
SK	-53.107	0.1	-62.152	-44.062	-55.023	-51.376	-54.838	-51.191	-54.950	-51.264
KU	3.246	0.1	2.088	4.405	2.987	3.524	2.969	3.506	2.958	3.535
SD	-0.091	0.2	-0.207	0.025	-0.182	0.016	-0.198	0.000	-0.194	0.012
SK	-75.568	0.2	-80.139	-70.997	-78.033	-73.607	-77.529	-73.103	-77.726	-73.410
KU	5.114	0.2	4.345	5.884	4.770	5.396	4.832	5.459	4.820	5.409
SD	2.034	0.4	1.893	2.174	1.958	2.115	1.952	2.109	1.949	2.118
SK	-102.780	0.4	-104.634	-100.927	-104.304	-101.286	-104.274	-101.257	-104.177	-101.384
KU	8.310	0.4	8.043	8.577	7.990	8.578	8.042	8.630	8.010	8.610
SD	2.444	0.5	2.392	2.497	2.408	2.487	2.402	2.481	2.403	2.486
SK	-102.849	0.5	-105.958	-99.739	-104.060	-101.759	-103.938	-101.637	-104.046	-101.651
KU	8.793	0.5	8.345	9.240	8.545	9.092	8.494	9.040	8.512	9.074
SD	2.619	0.6	2.587	2.651	2.591	2.650	2.588	2.647	2.588	2.650
SK	-101.196	0.6	-102.542	-99.850	-102.410	-99.998	-102.394	-99.982	-102.452	-99.940
KU	9.067	0.6	8.907	9.227	8.787	9.326	8.807	9.346	8.800	9.334
SD	2.814	0.8	2.738	2.890	2.787	2.843	2.784	2.840	2.785	2.842
SK	-96.375	0.8	-99.930	-92.820	-97.476	-95.563	-97.187	-95.274	-97.409	-95.342
KU	9.461	0.8	8.937	9.984	9.241	9.751	9.170	9.680	9.199	9.722
SD	2.936	0.9	2.902	2.971	2.914	2.967	2.906	2.959	2.910	2.962
SK	-93.553	0.9	-95.807	-91.299	-94.825	-91.918	-95.188	-92.281	-95.129	-91.977
KU	9.765	0.9	7.860	11.670	9.386	10.038	9.492	10.144	9.435	10.094

附表 12

异质性风险补偿估计（2015 年样本）

风险	系数	分位数	置信区间							
			normal_ci_l	normal_ci_u	percentile_ci_l	percentile_ci_u	pivotal_ci_l	pivotal_ci_u	normal_ci2_l	normal_ci2_u
SD	-0.757	0.1	-0.926	-0.589	-0.811	-0.705	-0.810	-0.704	-0.815	-0.700
SK	-35.196	0.1	-38.657	-31.735	-36.337	-34.266	-36.126	-34.055	-36.211	-34.181
KU	1.636	0.1	1.263	2.010	1.339	1.813	1.459	1.934	1.400	1.872
SD	0.116	0.2	0.021	0.211	0.059	0.178	0.053	0.173	0.055	0.177
SK	-52.842	0.2	-55.058	-50.625	-54.136	-51.713	-53.971	-51.547	-54.147	-51.537
KU	3.199	0.2	3.042	3.355	3.096	3.322	3.076	3.301	3.089	3.309
SD	2.103	0.4	2.056	2.150	2.029	2.164	2.042	2.177	2.036	2.169
SK	-79.551	0.4	-80.308	-78.794	-80.412	-78.326	-80.776	-78.690	-80.630	-78.472
KU	6.025	0.4	5.957	6.093	5.808	6.178	5.872	6.242	5.844	6.206
SD	2.447	0.5	2.434	2.460	2.413	2.473	2.421	2.482	2.417	2.478
SK	-80.891	0.5	-81.699	-80.082	-81.911	-79.980	-81.801	-79.870	-81.856	-79.925
KU	6.601	0.5	6.482	6.720	6.440	6.759	6.443	6.763	6.436	6.766
SD	2.567	0.6	2.538	2.596	2.545	2.588	2.547	2.590	2.546	2.588
SK	-80.498	0.6	-81.916	-79.081	-81.423	-79.517	-81.479	-79.574	-81.453	-79.544
KU	6.952	0.6	6.675	7.229	6.785	7.123	6.781	7.119	6.792	7.112
SD	2.761	0.8	2.730	2.792	2.735	2.790	2.732	2.787	2.733	2.789
SK	-77.901	0.8	-80.061	-75.741	-78.695	-77.015	-78.786	-77.107	-78.749	-77.053
KU	7.504	0.8	7.101	7.907	7.319	7.648	7.360	7.688	7.337	7.670
SD	2.893	0.9	2.835	2.952	2.870	2.927	2.860	2.916	2.867	2.920
SK	-75.854	0.9	-77.362	-74.346	-77.081	-74.791	-76.917	-74.628	-76.986	-74.722
KU	7.746	0.9	7.411	8.081	7.535	7.934	7.558	7.957	7.561	7.932

附表 13

异质性风险补偿估计（2014 年样本）

风险	系数	分位数	置信区间							
			normal_ci_l	normal_ci_u	percentile_ci_l	percentile_ci_u	pivotal_ci_l	pivotal_ci_u	normal_ci2_l	normal_ci2_u
SD	-0.882	0.1	-0.933	-0.832	-0.939	-0.832	-0.933	-0.826	-0.935	-0.829
SK	-19.441	0.1	-20.238	-18.644	-19.985	-18.838	-20.044	-18.896	-20.049	-18.833
KU	0.404	0.1	0.198	0.611	0.300	0.538	0.271	0.508	0.280	0.529
SD	-0.268	0.2	-0.349	-0.187	-0.328	-0.217	-0.320	-0.208	-0.325	-0.212
SK	-27.077	0.2	-27.805	-26.349	-27.692	-26.317	-27.836	-26.461	-27.775	-26.379
KU	1.126	0.2	0.901	1.351	1.022	1.245	1.006	1.229	1.014	1.238
SD	1.728	0.4	1.668	1.787	1.651	1.812	1.644	1.805	1.650	1.806
SK	-46.043	0.4	-46.865	-45.221	-46.859	-45.106	-46.980	-45.227	-46.902	-45.185
KU	2.703	0.4	2.659	2.747	2.617	2.784	2.622	2.789	2.618	2.788
SD	2.370	0.5	2.351	2.390	2.330	2.409	2.332	2.411	2.333	2.408
SK	-50.036	0.5	-51.638	-48.434	-50.732	-49.397	-50.675	-49.340	-50.744	-49.329
KU	3.194	0.5	2.975	3.414	3.113	3.260	3.129	3.276	3.116	3.273
SD	2.560	0.6	2.540	2.579	2.534	2.588	2.531	2.585	2.533	2.586
SK	-50.548	0.6	-51.076	-50.020	-51.128	-49.872	-51.224	-49.969	-51.219	-49.878
KU	3.462	0.6	3.409	3.515	3.361	3.539	3.385	3.563	3.375	3.549
SD	2.820	0.8	2.787	2.854	2.791	2.841	2.800	2.850	2.795	2.846
SK	-48.532	0.8	-49.796	-47.269	-49.358	-47.931	-49.134	-47.707	-49.272	-47.792
KU	3.847	0.8	3.743	3.950	3.744	3.964	3.729	3.949	3.740	3.953
SD	2.898	0.9	2.857	2.940	2.874	2.925	2.872	2.923	2.871	2.926
SK	-46.736	0.9	-48.107	-45.365	-47.512	-45.809	-47.663	-45.960	-47.586	-45.885
KU	4.076	0.9	3.815	4.338	3.965	4.213	3.940	4.188	3.949	4.204

附表 14

异质性风险补偿估计 (2013 年样本)

风险	系数	分位数	置信区间							
			normal_ci_l	normal_ci_u	percentile_ci_l	percentile_ci_u	pivotal_ci_l	pivotal_ci_u	normal_ci2_l	normal_ci2_u
SD	-0.918	0.1	-1.047	-0.790	-0.959	-0.874	-0.963	-0.878	-0.962	-0.875
SK	-13.059	0.1	-14.730	-11.388	-13.553	-12.514	-13.604	-12.565	-13.564	-12.553
KU	-0.488	0.1	-0.814	-0.162	-0.544	-0.426	-0.550	-0.432	-0.550	-0.426
SD	-0.528	0.2	-0.628	-0.427	-0.583	-0.476	-0.580	-0.473	-0.585	-0.471
SK	-15.259	0.2	-16.444	-14.075	-15.906	-14.794	-15.725	-14.613	-15.797	-14.722
KU	-0.455	0.2	-0.527	-0.383	-0.532	-0.391	-0.518	-0.377	-0.523	-0.386
SD	1.123	0.4	1.055	1.190	1.009	1.237	1.008	1.236	1.018	1.227
SK	-22.518	0.4	-22.899	-22.138	-23.121	-21.887	-23.150	-21.916	-23.166	-21.870
KU	0.544	0.4	0.407	0.681	0.416	0.684	0.404	0.672	0.411	0.677
SD	2.296	0.5	2.261	2.331	2.232	2.352	2.240	2.360	2.235	2.357
SK	-27.456	0.5	-27.901	-27.011	-28.042	-26.886	-28.025	-26.869	-28.017	-26.895
KU	1.335	0.5	1.307	1.363	1.261	1.398	1.272	1.409	1.271	1.399
SD	2.607	0.6	2.570	2.645	2.579	2.635	2.580	2.636	2.580	2.634
SK	-29.072	0.6	-29.739	-28.405	-29.632	-28.501	-29.643	-28.512	-29.649	-28.495
KU	1.567	0.6	1.519	1.616	1.522	1.626	1.509	1.613	1.515	1.619
SD	2.894	0.8	2.859	2.930	2.871	2.919	2.870	2.917	2.869	2.920
SK	-29.867	0.8	-30.740	-28.994	-30.402	-29.298	-30.436	-29.332	-30.458	-29.276
KU	1.930	0.8	1.843	2.017	1.862	2.000	1.860	1.998	1.860	2.000
SD	2.982	0.9	2.882	3.082	2.950	3.007	2.956	3.014	2.953	3.010
SK	-29.372	0.9	-30.602	-28.142	-30.085	-28.610	-30.134	-28.659	-30.124	-28.620
KU	2.173	0.9	1.919	2.428	2.066	2.244	2.103	2.281	2.081	2.266

附录 1 劳动收入风险与工资收入决定

1. 工资收入中的风险因素

假定个体在接受 S 年教育后面临两种就业选择：第一、寻求一份稳定工作，往后每年将获得一个无风险工资收入 Y_s ；第二、寻求一份非稳定工作，往后每年将获得一个风险工资收入 Y ，且 $E(Y) = \mu$ 。通常认为个体具有风险厌恶特征，因此非稳定工作的工资收入往往要高于稳定工作才会被选择。引入绝对风险溢价，选择 1 和选择 2 的工资收入关系如下：

$$\Theta = E(Y) - Y_s = \mu - Y_s \quad (1)$$

假定个体一旦做出选择，往后不存在更换工作的可能性。均衡状态要求两份工作的终身效用相等：

$$\int_s^{\infty} U(Y_s) e^{-\delta t} dt = E \left(\int_s^{\infty} U(Y) e^{-\delta t} dt \right) \quad (2)$$

假定工资收入与时间 t 无关，因此积分和折现均可忽略：

$$U(Y_s) = E[U(Y)] \quad (3)$$

将式 (3) 等号左侧在 μ 处进行一阶泰勒展开：

$$U(Y_s) = U(\mu) + (Y_s - \mu)U'(\mu) = U(\mu) - \Theta U'(\mu) \quad (4)$$

(1) 方差风险与绝对风险溢价。接下来将收入风险（方差风险）引入工资收入决定方程。对 $U(Y)$ 在 μ 处进行二阶泰勒展开：

$$U(Y) = U(\mu) + (Y - \mu)U'(\mu) + \frac{1}{2}(Y - \mu)^2 U''(\mu) \quad (5)$$

其中：

$$U'(\mu) = \left. \frac{dU}{dY} \right|_{Y=\mu} ; \quad U''(\mu) = \left. \frac{d^2U}{dY^2} \right|_{Y=\mu}$$

对式 (5) 求期望：

$$E[U(Y)] = U(\mu) + \frac{1}{2}m_2 U''(\mu) \quad (6)$$

其中， m_2 为随机变量 Y 的二阶中心矩即方差： $m_2 = E[(Y - \mu)^2]$ 。将式 (3)、式 (6)

带入式 (1) 后整理得到绝对风险溢价：

$$\Theta = \frac{1}{2}V_a m_2 \quad (7)$$

其中，绝对风险厌恶指标 $V_a = -\frac{U''(\mu)}{U'(\mu)}$ ，度量了风险对工资收入的边际影响，由于通常认

为个体具有方差风险厌恶特征，因此 V_a 也被称为风险补偿。

(2) 高阶风险与绝对风险溢价。高阶风险的引入非常直接，对 $U(Y)$ 在 μ 处进行四阶泰勒展开：

$$U(Y) = U(\mu) + (Y - \mu)U'(\mu) + \frac{1}{2}(Y - \mu)^2 U''(\mu) + \frac{1}{6}(Y - \mu)^3 U'''(\mu) + \frac{1}{24}(Y - \mu)^4 U^{(4)}(\mu)$$

其中， $U'''(\mu) = \frac{d^3 U}{dY^3} \Big|_{Y=\mu}$ ； $U^{(4)}(\mu) = \frac{d^4 U}{dY^4} \Big|_{Y=\mu}$ 。

遵循同样的推导逻辑，可以得到绝对风险溢价：

$$\Theta = \frac{1}{2}V_a m_2 - \frac{1}{6}V_a F_a m_3 + \frac{1}{24}V_a F_a R_a m_4 \quad (8)$$

其中， $m_3 = E[(Y - \mu)^3]$ ， $m_4 = E[(Y - \mu)^4]$ ，分别为工资收入随机变量的三阶和四阶中心矩即偏度和峰度。关于绝对风险厌恶指标或者风险补偿如下：

$$F_a = -\frac{U'''(\mu)}{U''(\mu)}; R_a = -\frac{U^{(4)}(\mu)}{U'''(\mu)}$$

附录 2 平均风险补偿估计模型

1. 实证模型

通过构造满足内曼正交(Neyman Orthogonal)性质的矩条件，配合交叉拟合(Cross Fitting)

方法, Chernozhukov 等(2018)提出利用机器学习算法预测优势帮助计量模型有效识别目标参数的去偏差/双重机器学习方法(Debiased/Double Machine Learning, DML), 为大数据背景下如何在计量经济模型中引入机器学习算法提供了一个灵活的应用框架。经典的半参数回归模型如下:

$$Y = D\theta_0 + g_0(Z) + U, E[U|Z, D] = 0 \quad (9)$$

其中, Y 为结果变量(工资收入), D 为处置变量(收入风险), 处置效应 θ_0 (平均风险补偿) 为实证分析目标。 Z 为同时影响结果变量与处置变量的混杂因素, U 表示满足条件均值独立的随机干扰项。半参数模型与传统线性回归模型最大的差异在于只关注少量感兴趣变量的参数识别, 假定其为线性。其余混杂因素的具体函数形式并不进行假定, 完全通过数据驱动的方式进行估计, 极大放宽了传统线性回归模型的模型约束假设。既然 Z 为混杂因素, 那么其与处置变量之间必然存在某种显著关系:

$$D = m_0(Z) + V, E[V|Z] = 0 \quad (10)$$

其中, $m_0 \neq 0$, 属于观测性研究的常见特征即处置选择非随机。 V 为随机干扰项。

(1) 估计方法。方法 1: 回归调整估计。关于半参数模型即式(26)的一个自然估计方法是: ①初始阶段给定 θ_0 的一个猜测值 θ_0' 。②基于非参数方法(例如随机森林、核回归)使用混杂因素 Z 对 $Y - D\theta_0'$ 进行拟合, 得到冗余参数(Nuisance Parameter) $\hat{g}_0(Z)$ 。③利用最小二乘方法对 $Y - \hat{g}_0(Z)$ 和 D 进行回归, 得到目标参数的估计值 $\hat{\theta}_0$ 。④重复步骤 1-3 直至收敛。回归调整方法对应矩条件为:

$$E[(Y - D\theta_0 - g_0(Z))D] = 0 \quad (11)$$

方法 2: 倾向得分调整估计。观测性研究中处置变量和混杂因素之间通常存在相关性, 因此可以仿照回归调整的思路可以利用 D 与 Z 之间的关系 $m_0(Z)$ 进行倾向得分调整。首先通过非参数方法学习得到冗余参数 $\hat{m}_0(Z)$ 。利用剔除混杂因素后 $D - E[D|Z]$ 和 $g_0(Z) + U$ 之间的无关性构建矩条件如下:

$$E[(Y - D\theta_0)(D - E[D|Z])] = 0 \quad (12)$$

方法 3: 内曼正交估计。无论回归调整还是倾向得分调整, 基本思想都是希望从结果变量和处置变量中将混杂因素进行剔除。因此可以综合两种调整方式, 采用 Frisch-Waugh-Lovell 定理描述的参数估计方式。①基于非参数方法使用 Z 对 Y 和 D 进行预测, 得到冗余参数 $\hat{E}[Y|Z]$ 和 $\hat{E}[D|Z]$ 。②剔除混杂因素影响(Partialling Out)后得到残

差: $\hat{W} = Y - \hat{E}[Y|Z]$ 和 $\hat{V} = D - \hat{E}[D|Z]$ 。③对 $\hat{W}$ 和 $\hat{V}$ 进行回归得到目标参数估计值 $\hat{\theta}_0$ 。

对应的矩条件为:

$$E\left[\left((Y - E[Y|Z]) - (D - E[D|Z])\theta_0\right)(D - E[D|Z])\right] = 0 \quad (13)$$

(2) 偏差来源。以回归调整估计为例说明方法 1 和方法 2 偏差来源。求解式 (11) 可得

$$\begin{aligned} \sqrt{n}(\hat{\theta}_0 - \theta_0) &= \left(\frac{1}{n} \sum_{i=1}^n D_i^2\right)^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^n D_i U_i \equiv a \\ &+ \left(\frac{1}{n} \sum_{i=1}^n D_i^2\right)^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^n D_i (g_0(Z_i) - \hat{g}_0(Z_i)) \equiv b \end{aligned}$$

如果处置变量不存在内生性, 那么非极端条件 (Mild conditions) 下 $a \sim N(0, \Sigma)$ 。对于 b , 进一步可写为:

$$b \approx (ED^2)^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^n m_0(Z_i) (g(Z_i) - \hat{g}_0(Z_i))$$

$m_0(Z_i)$ 一般不等于零, 此时如果 $g_0(Z_i) - \hat{g}_0(Z_i)$ 无法收敛到零, 那么意味着参数估计存在偏误。遗憾的是, 常见的机器学习方法, 例如 Lasso、Ridge、Boosting 等, 为了防止过度拟合以及解决高维问题, 通常施加正则化项 (Regularization Terms), 导致 g_0 的估计 $\hat{g}_0$ 存在偏误, 称之为正则化偏误 (Regularization Biases)。与此同时, 机器学习算法的预测偏误收敛速度较慢, 均方根误差上通常为 $n^{-\varphi_g}$ ($\varphi_g < \frac{1}{2}$), 导致 $\sqrt{nn}^{-\varphi_g} \rightarrow \infty$, 从而 $b \rightarrow \infty$ 。

综上, 可以看到正则化偏误是回归调整估计不可靠的真正原因。倾向得分调整方法与之类似, 不再赘述。

(3) 正交化去偏差。现在考虑内曼正交估计, 求解式 (13) 可得:

$$\begin{aligned} \sqrt{n}(\hat{\theta}_0 - \theta_0) &= \left(\frac{1}{n} \sum_{i=1}^n V_i^2\right)^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^n V_i U_i \equiv a^* \\ &+ \left(\frac{1}{n} \sum_{i=1}^n V_i^2\right)^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^n (m_0(Z_i) - \hat{m}_0(Z_i)) (\ell_0(Z_i) - \ell_0(Z_i)) \equiv b^* \\ &+ \left(\frac{1}{n} \sum_{i=1}^n V_i^2\right)^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^n U_i (m_0(Z_i) - \hat{m}_0(Z_i)) \equiv c^* \end{aligned}$$

第一项通常仍会满足 $a^* \sim N(0, \Sigma)$ 。第二项 b^* 变成了两个正则化偏差的乘积:

$(m_0(Z_i) - \hat{m}_0(Z_i))(\ell_0(Z_i) - \ell_0(Z_i))$ ，此时均方误差意义上的收敛速度变成了 $n^{-(\varphi_g + \varphi_m)}$ 。

市面上常用的机器学习算法收敛速度虽然不及 $n^{\frac{1}{2}}$ ，但是 $\varphi > \frac{1}{4}$ 仍然可以满足，因此两项

偏差乘积可以实现 $o\left(n^{\frac{1}{4}}\right)$ 的收敛速度，意味着 $\sqrt{nn}^{-(\varphi_g + \varphi_m)} \rightarrow 0$ ，从而 $b^* \rightarrow 0$ ，此时

$\sqrt{n}(\hat{\theta}_0 - \theta_0)$ 才有可能实现渐进正态分布。由此可知，根据矩条件式(13)估计目标参数时似乎可以规避机器学习算法的正则化偏差。Chernozhukov 等(2018)针对这一性质给出理论证明，发现矩条件式(13)满足内曼正交，意味着局部区域内目标函数对于冗余参数的估计偏差并不敏感。

(4) 交叉拟合去除相关性。最后考虑第三项 c^* ：

$$\left(\frac{1}{n} \sum_{i=1}^n V_i^2\right)^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^n U_i (m_0(Z_i) - \hat{m}_0(Z_i))$$

Chernozhukov 等(2018)认为，如果使用相同样本估计冗余参数和目标参数，那么 U_i 和 $(m_0(Z_i) - \hat{m}_0(Z_i))$ 的相关性将难以消除，根本原因在于 $\hat{m}_0(Z_i)$ 存在过度拟合。由此提出通过交叉拟合方法控制上述相关性，实现 $c^* = o_p(1)$ 。基本思想是随机二分样本 $Sample_A$ 和 $Sample_B$ ，前者用于训练估计冗余参数的机器学习算法，后者则用于识别目标参数。为了充分利用样本信息， $Sample_B$ 反过来用于提供冗余参数模型，使用 $Sample_A$ 再次估计目标参数。最终平均两次结果即可得到满足渐进正态分布的目标参数估计值。