

聪明的投资者:非完全市场化利率与风险识别^{*}

——来自 P2P 网络借贷的证据

廖理 李梦然 王正位

内容提要:利率市场化是我国金融改革的重要举措,但是目前还未有从微观数据角度探索利率市场化市场基础的研究。本文借助人人贷网络借贷平台的数据,试图做出尝试。实证结果表明:非完全市场化的利率部分反映了借款人的违约风险,但仍有相当高比例的违约风险未被反映在利率当中。进一步地,本文发现市场中的投资者是比较聪明的,他们能够借助借款人的公开信息识别相同利率背后所包含的不同违约风险。研究还发现,投资者的这种风险识别能力能够直接反映在一个成功订单的参与人数上,间接反映在订单募资成功所需的竞标时间上。这些结果说明,我国信贷市场中的投资者具有良好的风险判断能力,这为利率市场化提供了良好的基础。

关键词:非完全市场化利率 违约风险 风险识别 P2P 借贷

一、引言

2013年7月20日,随着中国人民银行宣布全面放开金融机构贷款利率管制,标志着利率市场化改革作为我国金融改革的重要突破口,开启了探索中国经济未来十年增长红利的新纪元。目前,除对金融机构的人民币存款利率上限还存在直接管制,在其他领域已经基本实现了通过市场机制进行定价、配置资源,由政府进行间接调控的局面。那么,利率市场化这一举措会对市场中投资参与者的投资决策带来怎样的影响?对市场中的融资主体、未来的经济增长会带来怎样的促进作用?

Mckinnon(1973)和Shaw(1973)最早在关于金融抑制和经济增长的研究中提出了“利率市场化”的概念,Mckinnon提出了金融抑制论,具体而言就是,政府通过国有银行垄断资金供给,设定一个很低甚至为负的利率,从而压制了金融中介机构实际债权存量,使得金融业被抑制而不能促进经济的有效增长。Shaw提出了金融深化论,他指出政府放弃对金融经济发展的过度干预,从而促进金融自由化,使得金融业能够优化社会资源配置,利率能够有效反映资金供求,提高到均衡水平,从而促进经济增长。这两大理论有力地证明了利率市场化对于金融发展的重要意义。但是利率市场化能够促进市场发展的前提是,市场中的投资者需要是足够聪明的,如此利率才能准确反映出市场的需求。目前为止,很少有实证研究从微观个体数据证明这一前提。

在中国金融改革的大环境下,2007年互联网金融行业伴随全球金融创新的浪潮涌入中国,其中P2P网络借贷平台是一种典型的互联网金融模式,所谓P2P网络信贷是指贷款人与投资者之间通过网络借贷平台而不是金融机构产生的无抵押贷款(Lin et al., 2009; Bachmann et al., 2011)。

^{*} 廖理,清华大学五道口金融学院,邮政编码:100084,电子信箱:liaol@pbcfs.tsinghua.edu.cn;李梦然,清华大学经济与管理学院博士生,邮政编码:100084,电子信箱:limr@pbcfs.tsinghua.edu.cn;王正位,清华大学五道口金融学院,邮政编码:100084,电子信箱:wangzhw@pbcfs.tsinghua.edu.cn。本文得到国家自然科学基金重点项目《中国城市居民家庭/消费者金融研究》(项目号:71232003)的资助。作者感谢匿名评审人的良好意见和建议,文责自负。

这种模式源自英国,是一种融合了互联网技术与小额信贷的新型民间借贷模式,凭借第三方网络平台,以信用贷款的形式将资金贷给有借款需求的人,它为个人与个人之间提供了公开透明的小额信用交易的可能。P2P网络信贷平台这种模式在带来无限商机的同时,也积累了大量个人与个人之间的交易数据,这些数据为研究信贷市场中投资者的参与行为以及利率市场化对融资效率的影响提供了一个理想“窗口”。

本文收集了中国人人贷 P2P 网络贷款平台的交易数据,这家 P2P 平台的利率设定方式是给定一个利率设定范围,由借款者自行设定利率的机制。这种机制有别于国外许多 P2P 网络贷款平台中所使用的竞价机制,即根据借款人信用分数及借贷记录等由平台直接设定利率的机制。这种由借款者自行设定利率的机制具有“市场化”成分,借款者在设定利率时也会适当考虑市场中的供求关系;但是,由于平台根据国家政策对借款利率进行了范围约束,尤其是对上限加以约束,使得这种机制下的利率可能不能完全反映出其应该对应的风险水平,因此这种利率机制被认为是“非完全市场化”的。本文将对这种非完全市场化下的利率如何对风险识别产生影响进行着重考察。

利用人人贷的交易数据,我们检验了:(1)这种非完全市场化的利率是否能够完全反映订单背后的违约风险?(2)投资者是否足够聪明,以至于能够识别相同利率背后不同的违约风险?(3)如果投资者能够识别,这种识别能力从哪些指标中体现?我们的结果证实,由于人人贷订单的利率是由借款人设定的,这种定价方式不能完全反映背后的违约风险。但投资者能够通过识别借款人的其他公开信息,来甄别相同利率背后不同的违约风险。这说明,投资者是足够聪明的,这为我国利率市场化有效发挥资源配置作用提供了市场基础。当利率确定的时候,对于违约风险更高的订单,投资者会采取更为保守的投资方式,从而导致完成订单所需要的人数更多,竞标时间更长。

本文的结构如下:第二部分对相关研究进行评述;第三部分是研究假设与实证方案设计;第四部分是样本数据与变量定义;第五部分是实证研究的结果;第六部分是结论。

二、相关研究评述

现代资本资产定价理论在于揭示风险和收益之间的关系,在无摩擦的市场环境之下,Sharp et al. (1964)指出,当市场达到均衡时,风险与收益之间呈现出线性关系。也就是说,在市场均衡的情况下,我们通过观测利率就能揭示其背后的风险。但是 CAPM 的一个前提假设是,市场中的参与者拥有完全对称的信息。然而在现实生活中,信息不对称是永远存在的,在不完全信息市场,由于市场交易双方对信息的掌握程度不同,对于双方的交易行为会造成影响 (Bagehot, 1971)。Akerlof (1970)最初根据旧车市场模型提出了逆向选择理论,是指交易双方中占有信息优势者会由于其具有更多信息而使另一方利益受损,信息劣势方难以做出正确决策,导致价格扭曲,失去了供需平衡,从而降低了市场效率。解决这种信息不对称的重要方式,就是要借助信息披露和信号理论。

具体到 P2P 网络借贷市场,由于 P2P 网络借贷市场中交易双方互不相识,投资者的投资决策主要取决于借款人在网络上披露的信息内容及数量多少,因此其信息不对称情况更加严重,借款人通常具有更强的信息优势,更多的违约信息掌握在借款人一方。Stein (2002)指出信息是可以根据其可传递性分为:硬信息 (hard information) 和软信息 (soft information),硬信息是指能够被客观证实的内容,软信息是指不能被直接证实的内容。在 P2P 网络信贷市场中,Stein 的这种划分方式体现的更为具体,如借贷过程中借款人提供的住房抵押证明、结婚证等就可以归结为硬信息范畴,而在借贷过程中如借款人在其个人陈述中提到他是一个努力、勤奋、诚实的人等就可以被认为是软信息。

目前,大多数关于 P2P 借贷市场中的研究都证实了借款人所披露的硬信息是有助于降低信息不对称的情况,一般的硬信息包括信用等级、信用评分、种族、性别、婚姻状况、子女个数等因素,从

具体研究结果来看：

(1) 硬信息的披露对于交易成功与否具有一定的影响,使用 Prosper 网站的数据,Herzenstein et al. (2008)指出贷款人的信用、个人信息等都对借款成功率具有正显著影响;Lin (2009)指出借款人的信用等级越低,借款的成功率就越低;Puro et al. (2010)指出信用评级、总负债偿还比例以及当前逾期记录对于新的借款成功率具有重要影响;Pope & Sydnor (2011)指出年龄在 35—60 岁的人比其他年龄段的人更容易贷款成功;Ravina (2008)指出黑人的借款成功率较低,主要因为他们可以接受的贷款最高利率较低,所以不容易借到款;Ding et al. (2010)利用拍拍贷的数据,发现收入越高的人越容易借款成功。

(2) 硬信息的披露对于利率的高低具有一定的影响,Pope & Sydnor (2011)以及 Ravina (2008)指出种族信息的披露对于借款人订单成交的利率具有显著性影响,以上这些实证研究证明了硬信息的披露是有助于降低信息不对称的情况。有意思的是,借款人这些硬信息的披露也确实能够反映出其背后的违约风险。例如,Kumar (2007)指出信用等级越高和借款人披露其对资金账户认证的信息都与借款人的违约率呈负相关关系,但是借款额度却与违约率呈现出正相关关系。Klaft (2008)以及 Lin (2009)都指出借款人信用等级越低,其违约率的确更高。

另一些研究发现软信息方面的披露也会对借贷双方的交易行为造成影响。一般来说,照片、个人贷款的陈述等因素可作为衡量软信息的指标,Klaft (2008)指出通过照片判断长相越好看的人越容易贷款成功,但对利率没有影响。Duarte et al. (2012)指出,长相看上去越值得信任的人越容易贷到款,且利率更低。Herzenstein et al. (2011)以及 Larrimore et al. (2011)都指出个人陈述中提到自己过去的成功经历以及是值得信任的,将会提高其贷款成功率,陈述的内容越多越容易获得贷款。Michal (2012)指出对于信用等级较差的人披露越多的信息越有可能获得贷款,即使这些信息是未被认证过的也会增加其贷款成功率。有些情况下,软信息也会降低借贷双方的信息不对称程度。Duarte et al. (2012)发现那些长相看上去值得信任的人,其真实的违约率的确更低。但是,与硬信息不同的是,这些软信息的披露不一定能够降低信息不对称的情况,相反有可能会加剧信息不对称的程度。例如 Herzenstein et al. (2011)研究发现尽管文字性的描述有助于提高借款成功率,但是陈述内容越多的人往往具有更高的违约风险。

尽管以上的研究对于理解 P2P 借贷市场中的信息不对称以及信息披露的重要作用具有一定的意义,但是由于国外的研究数据主要集中于 Prosper 和 Lending Club 两家 P2P 网络借贷平台,他们都采用的是市场化的利率定价机制,在这种机制下,信息的不对称情况与违约风险可以通过利率定价得以体现;但是,在非完全市场化的利率机制下,利率未必能够完全反映违约风险,因此更多的信息不对称被利率定价机制所隐藏起来。因此过往的文献尚未回答我们在引言中所提出的问题,投资者在非市场化利率下的决策行为还有待进一步研究。

三、研究假设与实证设计

1. 研究假设

如前所述,在市场化利率定价机制下,每个投资者都可以通过各种途径收集信息,各自做出对借款人违约概率的判断,进而表达自己愿意参与借贷的利率;市场汇集每个投资者的利率定价,从而达到均衡利率。因此在这种机制下,利率就可以作为投资者对借款人违约风险的最终识别结果。

但是对于非完全市场利率机制,以人人贷网络借贷平台为例,其利率由借款人直接设定,投资者并不直接参与利率的定价。那么,在这种机制下,利率很可能无法完全反映出借款人的违约风险。我们想了解:如果两个订单具有相同的利率,但是其背后的违约风险不同的话,投资者会如何

对待这两个订单?

很显然,风险厌恶的投资者会在两个订单中更偏好风险较小者,这种偏好方式可以表现为两种形式:第一,个体投资者愿意在风险较小的订单上投资更多金额;第二,会有更多的投资者愿意投资于风险较小的订单。在第一种力量的作用下,只需要更少的投资者即可完成风险较小的订单,而对于风险较大的订单,则需要更多的投资者,因此我们可以提出如下研究假设:

假设 1:在非完全市场化利率机制下,对于相同的借款利率,借款人违约概率越大,订单募集资金成功所需要的投资者人数越多。

在第二种力量作用下,会有更多的投资者涌向风险较小的订单,这会导致风险较小的订单很快完成竞标;而对于风险更大的订单,则需要更长的竞标时间,因此我们可以得到如下假设:

假设 2:在非完全市场化利率机制下,对于相同的借款利率,借款人违约概率越大,项目募集资金所需的竞标时间越长。

2. 模型设定

如上文所述,当市场均衡的时候,市场化的利率就反映了违约风险,换句话说,对于市场化的利率(例如通过集合竞价得到的利率),通过观测借款人的借款利率就可以推测其违约概率,即:

$$\Pr(\text{default}_i | \text{rate}_i) = f(\text{rate}_i) + \varepsilon_i \quad (1)$$

如果利率是完全市场化的,那么相同的利率就应该对应着相同的违约风险。否则市场的力量会使更多的资金参与竞争违约风险较低的项目,从而导致新的均衡。但是如果利率是非市场化的,或者非完全市场化的,那么同一个利率可能对应不同的违约风险。仍以人人贷为例,人人贷网站是由借款人自行设定利率及借款额度,相同利率的两笔借款,其背后的风险也有可能不尽一致。

对于人人贷网站的利率设定方式,除了利率之外,我们可以用借款人的其它公开信息来辅助预测其违约概率,如公式(2)所示:

$$\Pr(\text{default}_i | \text{rate}_i, \text{other_info}_i) = f(\text{rate}_i, \text{other_info}_i) + \xi_i \quad (2)$$

如果公式(2)对违约率的预测显著高于公式(1),则我们可以推断 *other_info* 中包含了 *rate* 所不包含的信息——这也就意味着在非市场化的利率之下,利率并不能完全反映项目背后的违约概率。

在这种情况下, $\Pr(\text{default}_i | \text{rate}_i, \text{other_info}_i) - \Pr(\text{default}_i | \text{rate}_i)$ 则衡量了 *rate* 所不包含的违约风险。因此,我们定义:

$$\text{default_risk}_{\text{new}} \equiv \Pr(\text{default}_i | \text{rate}_i, \text{other_info}_i) - \Pr(\text{default}_i | \text{rate}_i) \quad (3)$$

根据前文的推理过程,如果投资者足够聪明,能够发现相同利率所对应的不同的违约风险,那么我们预期:(1)项目募集资金成功所需要的投资者人数(*num*)与 *default_risk_{new}* 成正比;(2)项目募集资金成功所需要的时间(*time*)与 *default_risk_{new}* 成正比。

四、数据与变量定义

1. 数据选取与网站特征描述

人人贷网站成立于 2010 年 5 月,由于该网站在成立初期有一些测试数据,会对研究具有一定干扰作用,因此本文选取人人贷网站 2011 年 1 月 1 日至 2013 年 4 月 27 日期间发布的所有有效借款订单数据作为研究分析的全样本,共包含 64539 个有效订单,其中涉及 2011 年的订单数为 15409 个,2012 年的订单数为 29546 个,2013 年的订单数为 19584 个。

人人贷的用户可以自愿将其个人信息披露于人人贷网站,包括工作城市、性别、年龄、单位规模(人数)、职业、工作时间、收入、毕业的学校、学历、入学时间、房产情况、车产情况及房贷、车贷等信息,特别的可以对“硬信息”提供客观证明由人人贷网站进行审核后得以认证,成为被认证过的信

息,包括:央行开具的个人征信报告、身份证复印件、工作证复印件或劳动合同、盖有单位公章的收入证明或工资卡银行流水、结婚证、户口本、个人视频、手机号、微博账号、汽车产权证明、房屋产权证明等,这些信息的认证情况参见表1。人人贷网站会根据借款人披露的信息及能够认证的信息情况为每位借款人进行信用评级,具体分为:AA、A、B、C、D、E、HR七个级别。

表1 认证信息及其描述性统计 (观测值:5681个)

| 变量 | 定义 | 释义 | 平均值 |
|--------------------------------|---------------------------|----------------|------|
| <i>idcard_cert</i> | 1 = 认证成功, 0 = 认证失败或没认证 | 提交身份证复印件 | 0.92 |
| <i>credit_cert</i> | 1 = 认证成功, 0 = 认证失败或没认证 | 提交征信报告 | 0.15 |
| <i>workpermit_cert</i> | 1 = 认证成功, 0 = 认证失败或没认证 | 提交工作证复印件 | 0.15 |
| <i>income_cert</i> | 1 = 认证成功, 0 = 认证失败或没认证 | 提交收入证明或银行流水 | 0.14 |
| <i>merriage_cert</i> | 1 = 认证成功, 0 = 认证失败或没认证 | 提交结婚证复印件 | 0.75 |
| <i>household register_cert</i> | 1 = 认证成功, 0 = 认证失败或没认证 | 提交户口本复印件或居住地证明 | 0.21 |
| <i>video_cert</i> | 1 = 认证成功, 0 = 认证失败或没认证 | 提交视频 | 0.53 |
| <i>mobile_cert</i> | 1 = 认证成功, 0 = 认证失败或没认证 | 提交手机实名认证 | 0.31 |
| <i>weibo_cert</i> | 1 = 认证成功, 0 = 认证失败或没认证 | 提交微博账号 | 0.35 |
| <i>car_cert</i> | 1 = 认证成功, 0 = 认证失败或没认证 | 提交汽车产权证明 | 0.46 |
| <i>house_cert</i> | 1 = 认证成功, 0 = 认证失败或没认证 | 提交房屋产权证明 | 0.41 |

人人贷网站的投标种类分为四种,信用认证标、实地认证标、机构担保标以及智能理财标。从表2可以看到,本文的样本中,包括信用认证标52097个,实地认证标4121个,机构担保标2054个,智能理财标92个,其中信用认证标是最主要的订单种类,占到全样本的90.28%。由于实地认证标、机构担保标以及智能理财标的担保方式和信息披露方式与我们要考察的目标存在较大的差异,并且在样本期间违约率均为零,因此本文主要对信用认证标进行深入研究,剔除了其他三类订单数据的样本。

表2 样本期间人人贷订单分布状况

| 订单状态 | | 信用认证标 | 实地认证标 | 机构担保标 | 智能理财标 | 总计 |
|------|-----|-------|-------|-------|-------|-------|
| 成功 | 还款中 | 500 | 3871 | 1119 | 1 | 5491 |
| | 违约 | 390 | 0 | 0 | 0 | 390 |
| | 已还清 | 5598 | 249 | 930 | 89 | 6866 |
| 失败 | | 51784 | 1 | 5 | 2 | 51792 |
| 总计 | | 58272 | 4121 | 2054 | 92 | 64539 |

由于本文需要研究订单的违约风险,因而对样本要求必须是已经筹标成功的订单。此外,正在还款的订单无法判断借款人最终是否会对订单全额违约。因此,本文最终选取订单的状态为“已还清”及“违约”两种类型的订单。

另一方面,值得说明的是:我国对非法集资有着严格的标准,个人吸收公众存款数额在 20 万元^①以上则有可能触犯这一标准。为避免 20 万元以上背后的法律风险对本文研究的干扰,我们把样本数据限定在 20 万元以下订单。共计得到 5681 个样本,其中 352 个违约订单。

根据人人贷网站规定,借款人在申请不同订单种类进行借款时,分别有相应的认证信息需要提交人人贷网站进行审核。对于信用认证标(现又叫工薪贷),需要提交审核的材料包括身份证复印件、个人征信报告、带公司公章的劳动合同或工作证明以及近 3 个月工资卡银行流水;人人贷网站目前的运营方式是不参与利率定价,2012 年前,对订单利率不设定下限,但需要根据国家有关条例对民间借贷的利率设定上限,2012 年以后,人人贷网站出台了新的利率规定,利率范围定为 10%—24%,信用认证标的利率是借款人自主设定的;人人贷网站规定借款额度必须在 3000—500000 之间,且为 50 的倍数,借款期限可以为 3、6、9、12、18、24 个月,还款方式均为等额本息,按月还款。当一个借款人发起一个订单时,投资者可以进行投标,投标的最低金额是 50 元,因此对于同一个订单,可能会有多个投资者共同参与,每个投资者投资的额度也不尽相同。一个订单自挂出后,网站规定其筹标期限为 7 天,当所有投资者的出资额度相加等于借款人需要借款的数额时,表示该订单筹款成功,即使期限未到,订单也不能再继续接受投资者投标;如果在规定期限内未能筹到款或是所有投资者愿意投资的额度相加不能达到订单需要借款的数额时,则该订单筹款失败。

2. 变量设置

根据研究需要,本文设定核心研究变量如下:

违约(*default*):当一个订单完成筹款后,如果未能按时还款超过约定时间 30 天后,人人贷网站将会将这种行为认定为严重逾期,网站对于该订单的状态会显示为“逾期”;另有一些订单逾期时间超过 90 天,人人贷网站将按照对投资者“本金保障”的约定,垫付本金,这时网站对于该订单的状态会显示为“已垫付”。本文将这两种状态下的订单都视作违约订单,并处理为虚拟变量。1 = 有违约现象,0 = 已还清。

利率(*rate*):每一个订单中,借款人都会自行设定一个可以接受的借款利率,2011 年对利率的设定只规定上限不得超过银行同类贷款的 4 倍,2012 年以后网站规定利率范围为 10%—24%。

人人贷利率的政策变化(*pre2012*):2012 年 1 月 1 日以后,人人贷借贷平台加强了对借款人风险的控制,不仅对借款的利率范围进行了限定,而且引入了订单预审机制。因此本文将 2012 年以前的数据设定为 1,即 $pre2012 = 1$,否则为 0。

借款人信用的风险水平(*risk level*):根据人人贷网站对于借款人会根据其提交的信息及交易记录给出信用等级,信用等级为 HR 时, $risk\ level = 1$ (HR 为信用最差的,即借款人信用的风险水平最高),其他等级的为 0。

订单投资者个数(*num*):一个成功订单中参与投资的投资者人数。

竞标时间(*time*):一个成功订单募集完成所需要的时间。人人贷网站规定,一个订单自发布起需在七天内完成募资,但大多数成功的订单都会在几个小时内完成,因此募资完成的速度对于本文的研究具有重要的意义。本文数据样本中,时间的单位为分钟。

借款期限(*month*):每一个订单中,借款人都会自行设定一个借款期限,借款期限可以为 3、6、9、12、18、24 个月。

^① 参见《关于审理非法集资刑事案件具体应用法律若干问题的解释》。

借款人需要借款的数额 (*money*):网站规定必须是 3000—500000 之间的数,并且是 50 的倍数,但不同信用等级的借款人的最高借款数额要根据网站提供给他的信贷额度所决定。

信贷额度 (*credit line*):是人人贷网站根据借款人的信用等级、信用评分及其他其所提供的认证信息给出的贷款额度,信贷额度限制了借款人需要借款的数额上限。

认证信息的个数 (*cert_num*):将每个借款人进行认证的信息个数进行统计,显示其个人信息认证披露的程度。

表 3 给出了变量的描述性统计。

表 3 变量描述性统计

| 变量 | 变量个数 | 平均值 | 标准差 | 最小值 | 最大值 |
|--------------------|------|----------|----------|------|---------|
| <i>default</i> | 5681 | 0.062 | 0.241 | 0 | 1 |
| <i>rate</i> | 5681 | 14.007 | 3.413 | 5.6 | 24.4 |
| <i>pre2012</i> | 5681 | 0.453 | 0.498 | 0 | 1 |
| <i>risk level</i> | 5681 | 0.323 | 0.468 | 0 | 1 |
| <i>num</i> | 5681 | 27.249 | 42.572 | 1 | 382 |
| <i>time</i> | 5681 | 372.118 | 1125.535 | 0 | 9996 |
| <i>month</i> | 5681 | 6.736 | 4.556 | 3 | 36 |
| <i>money</i> | 5681 | 22614.09 | 34870.42 | 3000 | 190000 |
| <i>credit line</i> | 5681 | 50314.25 | 75771.52 | 0 | 1000000 |
| <i>cert_num</i> | 5681 | 4.366 | 2.307 | 0 | 11 |

表 4 给出了变量的相关系数以及显著性水平,从表 4 可以看出一些有意思的现象:首先,借款人的风险等级和利率之间呈现出较强的正相关关系,说明人人贷利率的设定一定程度上反映了借款人的信用等级风险,而借款人的信用等级和其信息披露水平 (*cert_num*) 是有很大关系的;其次,利率和违约率之间有显著的正相关关系,说明人人贷利率的设定部分反映了借款人的违约概率。

表 4 变量相关系数

| | <i>default</i> | <i>rate</i> | <i>pre2012</i> | <i>risk level</i> | <i>num</i> | <i>time</i> | <i>month</i> | <i>money</i> | <i>credit line</i> |
|--------------------|----------------|-------------|----------------|-------------------|------------|-------------|--------------|--------------|--------------------|
| <i>rate</i> | 0.14*** | | | | | | | | |
| <i>pre2012</i> | -0.08*** | 0.07*** | | | | | | | |
| <i>risk level</i> | 0.36*** | 0.36*** | 0.010 | | | | | | |
| <i>num</i> | 0.04*** | -0.03** | -0.28*** | -0.09*** | | | | | |
| <i>time</i> | -0.03** | 0.01 | 0.15*** | 0.04** | 0.29*** | | | | |
| <i>month</i> | 0.18*** | 0.15** | -0.10*** | 0.03** | 0.22*** | 0.14*** | | | |
| <i>money</i> | 0.09*** | 0.00 | -0.31*** | -0.14*** | 0.81*** | 0.15*** | 0.24*** | | |
| <i>credit line</i> | -0.02 | -0.27*** | -0.01 | -0.25*** | 0.34*** | 0.06*** | 0.07*** | 0.47*** | |
| <i>cert_num</i> | -0.14*** | -0.28*** | -0.23*** | -0.37*** | 0.05*** | -0.11*** | -0.02 | 0.07*** | 0.24*** |

注:***代表在 1% 水平下显著;**代表在 5% 水平下显著;*代表在 10% 水平下显著。

另一方面,在表 4 中还比较关注 *num* 和 *time* 两个变量,从相关系数可以看到,*num* 和违约率正相关,说明违约率较高的订单需要更多的投资者才能完成,这一点和我们的假设比较相符;但 *time* 和违约率负相关——这一定程度上违反我们的直觉,但这并不能说明投资者会“疯抢”风险高的订单。实际上,从 *time* 和 *risk level* 之间的相关性来看,二者是正相关关系,说明风险等级为 HR 的订单会被投资者拖得时间更长。当然,这些都是简单的相关关系,为了给论文提供更严谨的证据,还

需要更为详尽的实证分析。

此外,表 4 还说明订单的金额会对 *num*、*time* 构成显著影响,说明我们在后文中需要对该变量进行控制。

五、实证分析结果

1. 利率、其他信息与违约风险

在利率市场化的情况下,利率是借款人使用资金的成本,应当可以直接反映出违约风险的水平,也就是说借款人设定的利率越高意味着借款人违约的可能性越高,同时,相同的利率应当反映出相同的违约风险。因此,基于在上文中提到的 P2P 网络借贷平台非完全市场化利率的现状,本文首先需要探究人人贷借贷利率是否包含了借款人的违约风险?如果答案是肯定的,那么本文接下来将考察该利率是否能够全部反映出违约风险?换言之,将考察相同利率的两笔借款其背后的违约风险是否一致。

为了考察人人贷借贷利率是否包含了借款人的违约风险,我们可以根据模型中的公式(1),以借款利率为核心解释变量,以借款人最终是否违约为被解释变量,看借款利率对违约概率有没有预测功能。此外,为了排除借款期限、人人贷关于利率的限制这些因素的干扰,我们对这两个因素进行了控制。表 5 的模型(1)描述了回归的结果。

表 5 利率与其他信息对违约风险的
回归检验(Probit 回归)

| | 模型(1) | | 模型(2) | |
|------------------------------|-----------------------|--------------|-----------------------|--------------|
| | 被解释变量: <i>default</i> | | 被解释变量: <i>default</i> | |
| | 系数 | z-statistics | 系数 | z-statistics |
| <i>constant</i> | -2.742*** | -22.978 | -3.368*** | -13.908 |
| <i>rate</i> | 0.082*** | 10.680 | 0.025** | 2.229 |
| <i>month</i> | | | 0.067*** | 10.361 |
| <i>credit line</i> | | | 4e-06*** | 8.166 |
| <i>cert_num</i> | | | -0.132*** | -5.536 |
| <i>risk level</i> | | | 2.112*** | 15.427 |
| <i>pre2012</i> | | | -0.593*** | -7.111 |
| <i>Observations</i> | 5,681 | | 5,677 | |
| <i>Persudo R²</i> | 0.044 | | 0.370 | |

注:***代表在 1% 水平下显著;**代表在 5% 水平下显著;*代表在 10% 水平下显著。

公开信息,包括订单的期限、信贷额度、认证信息的个数、借款人的风险水平来辅助预测其违约概率。此外,为了排除人人贷对利率设定制度变化的影响,我们也对 2012 年之前的样本进行了控制。

结果如表 5 的模型(2)所示。从模型(2)可以看到,利率仍在 5% 的统计水平下显著,说明在引入了订单借款人其它公开信息之后,利率对于借款人的违约概率仍有一定的预测功能。同时,所引入的三个借款人的公开信息变量都在 1% 的水平下显著,具体而言,我们发现借款人认证信息的个数(*cert_num*)越多,借款人的违约概率越低,由(*risk level*)的系数显著为正可知当借款人信用等级为 HR 时,其违约概率最高,这意味着人人贷的信息认证体系一定程度上反映了个人诚信状况;个人信贷额度(*credit line*)在 1% 的水平下显著为正,这可能意味着人人贷对于违约概率较高的借款人,赋予了过高的信贷额度;总体来说,个人其它公开信息的显著性说明这些信息对于借款人的违

从模型(1)的结果可以看到,利率的系数为 0.082,在 1% 的统计水平下显著性。同时模型的拟 R^2 达到 4.4%。该结果说明:尽管人人贷的利率设定并非完全市场化的,但借款利率还是部分反映了借款人的违约风险。其内在的机制可能在于:对于那些风险较高而利率设定较低的订单,投资者虽然并未直接参与利率的设定,但是却可以采用“以脚投票”的方式,不参与这些订单,从而使得“幸存订单”的利率在一定程度上反映借款人的违约概率。

本文更想了解的是:人人贷交易中相同的借款利率是否具有相同的违约风险?根据公式(2)的设定,我们可以在控制利率的情况下,引入订单以及借款人的其它

约行为也有较强的预测作用,而这些预测作用并没有完全体现在利率当中。另一方面,观测模型的解释力可以发现,在引入了借款人其它公开信息之后,模型的解释力由原来的4.4%提升为37%,再次说明对于预测借款人的违约概率而言,借款人的这些公开信息包含了大量未被包含在利率之中的信息。事实上,全部样本的违约率为6.3%,因此如果以6.3%作为违约的阈值^①对表5中的模型进行判断准确度验证,那么模型(1)预测借款人违约、且借款人确实违约的预测准确率为69.32%,预测借款人违约、但借款人实际没违约的预测错误率为45.06%;模型(2)预测借款人违约、且借款人确实违约的预测准确率为96.02%,预测借款人违约、但借款人实际没违约的预测错误率为24.90%。这些结果也说明在Probit模型中引入这些信息可以使得违约概率的预测更为精准。

2. 投资者对于风险的识别

根据上文,我们已经发现利率不能完全反映出违约风险——换句话说,相同的借款利率,其背后的违约风险可能是不一样的。那么投资者能否识别其背后的差异呢?

根据公式(3)的设定,我们首先要识别出由借款人其它公开信息能反映出来的、而利率没有反映出来的那部分违约风险。为此,分别利用表5中模型(1)和模型(2)估算借款人的违约概率,后者与前者的差额即反映了新信息所反映出的违约概率。

如前文所述:如果投资者足够聪明,那么他们就应该能够发现相同利率的订单其背后对应的不同违约风险,因此我们对订单的投资者个数以及竞标时间进行检验,得到结果如表6所示:

表6 投资者风险识别能力的检验(OLS回归)

| | 模型(3) | | 模型(4) | |
|-----------------------------|-------------|---------|-------------|--------|
| | 被解释变量: num | | 被解释变量: time | |
| | 系数 | t 值 | 系数 | t 值 |
| constant | 10.138 *** | 7.181 | 226.973 *** | 3.598 |
| default_risk _{new} | 9.842 *** | 3.288 | 350.863 *** | 2.623 |
| rate | -0.3712 *** | -3.833 | 2.727 | 0.630 |
| money | 0.001 *** | 103.698 | 0.005 *** | 11.146 |
| Observations | 5677 | | 5677 | |
| R ² | 0.658 | | 0.024 | |

注:***代表在1%水平下显著;**代表在5%水平下显著;*代表在10%水平下显著。

被利率所包含的违约风险,因此表6的结果充分验证了我们此前提出的两个假设,即在控制了利率之后,借款人风险越高,需要更多的投资者才能满足其资金需求;在控制了利率之后,借款人风险越高,需要更长的竞标时间才能完成筹款。

实际上,表6还有更多的经济含义:default_risk_{new}显著的一个前提条件是投资者是足够聪明的,否则即使相同的利率背后隐藏着不同的风险,投资者也未必能够意识到这一点。因此,default_risk_{new}显著无疑证实了投资者的确是很聪明的——当他们无法直接决定利率设定时,选择了其它路径作为风险出口。在本文中这些路径就直接反映在投资参与的人数上,间接反映在了订单筹资完成所需要的竞标时间上。

3. 稳健性检验

作为对上文实证结果的稳健性检验,我们首先考虑了样本选取是否影响本文的结果。尤其是,

根据表6,我们发现:无论是模型(3)还是模型(4)中,default_risk_{new}都在1%的水平下显著为正,这意味着对于一个筹资成功的订单,其所需要的投资者人数(num)与default_risk_{new}成正比,从具体系数来看,default_risk_{new}每增大10个百分点,就额外需要1个人才能完成订单;一个订单筹资完成所需要的竞标时间(time)与default_risk_{new}成正比,从具体系数来看,default_risk_{new}每增大10个百分点,就额外需要35分钟才能完成订单。值得一提的是,这里default_risk_{new}的经济意义是没有

^① 如果表5中Probit模型预测借款人违约的概率超过6.3%就判定其违约,否则就判定其不会违约,该方法参考了Palepu(1986)、赵勇、朱武祥(2000)等人的研究。

之前我们把样本限定为订单金额在 20 万元以下的样本,那么对于不同借款额度,前面的结果是否依然稳健?为此,我们把订单的借款额度分别限定在 5 万元和 10 万元,对上述两个问题进行稳健性检验,表 7 和表 8 显示了稳健性检验的结果。结果表明,无论借款额度限定在 5 万元还是 10 万元,相较于之前的结果,变量系数和符号都比较稳定,其显著性水平也没有明显变化,说明上文结果具有较好的稳健性。

表 7 稳健性检验一:订单金额在 5 万元以下

Panel A:利率与其他信息对违约风险的回归检验(Probit 回归)

| | 模型(5) | | 模型(6) | |
|------------------------|---------------|--------------|---------------|--------------|
| | 被解释变量:default | | 被解释变量:default | |
| | 系数 | z-statistics | 系数 | z-statistics |
| constant | -2.873*** | -22.392 | -3.217*** | -12.363 |
| rate | 0.085*** | 10.603 | 0.029** | 2.522 |
| month | | | 0.059*** | 8.868 |
| risk level | | | 1.879*** | 11.753 |
| credit line | | | 3.49e-06*** | 4.163 |
| cert_num | | | -0.145*** | -5.710 |
| pre2012 | | | -0.486*** | -5.564 |
| Observations | 4980 | 4976 | | |
| Persudo R ² | 0.0543 | 0.326 | | |

Panel B:投资者风险识别能力的检验(OLS 回归)

| | 模型(7) | | 模型(8) | |
|-----------------------------|-----------|--------|------------|-------|
| | 被解释变量:num | | 被解释变量:time | |
| | 系数 | t 值 | 系数 | t 值 |
| Constant | 11.208*** | 15.267 | 245.771*** | 3.979 |
| default_risk _{new} | 13.749*** | 6.885 | 719.421*** | 4.282 |
| rate | -0.426*** | -8.658 | 3.468 | 0.837 |
| money | 0.001*** | 71.578 | 0.003** | 2.353 |
| Observations | 4976 | 4976 | | |
| R ² | 0.528 | 0.006 | | |

注:*** 代表在 1% 水平下显著;** 代表在 5% 水平下显著;* 代表在 10% 水平下显著。

另一方面,表 5 中模型(1)和模型(2)在使用 probit 回归时,我们是将认证信息进行了数量统计,另一种方式是把每个认证信息都作为单独的变量放进回归模型。我们在表 9 的模型(14)中,加入了每一个认证信息的哑变量,结果表明,模型的 R² 明显提高,再次说明人人贷的利率设定方式不能包括所有的信息从而不能准确的反映违约风险的概率水平。表 7 中的 Panel B 表明:我们使用本文的方法识别出利率所未包含的违约风险之后,该违约风险与 num 成正比,与 time 成正比。这与上文的结论仍然一致。

此外,为控制订单提交时间的影响(比如,相比于高峰时间,深夜提交的订单可能需要更长的时间),我们引入 24 小时(23 个哑变量)的控制变量之后, default_risk_{new} 系数变化不大,且依然在 1% 的水平下显著。^① 另外,由于人人贷网站 2012 年之前对利率的设定只规定上限不得超过银行同类贷款的 4 倍,而在 2012 年以后网站规定利率范围为 10% - 24%。为了避免这种政策的影响,我们对 2012 年之后的样本单独进行了检验,结果依然稳健。^②

六、结论与启示

利率市场化是我国金融改革的重要举措,对实现资源有效配置的意义不言而喻。但是利率市场化提高资源配置的一个重要前提是市场的有效性,即:需要市场中有一批聪明的投资者,他们能够识别市场化利率背后的风险。

① 感谢匿名审稿人提出这一有益建议。

② 限于篇幅,此处没有列示上述两部分结果,有兴趣的读者可以向作者索取。

表 8 稳健性检验二:订单金额在 10 万元以下

Panel A: 利率与其他信息对违约风险的回归检验 (Probit 回归)

| | 模型 (9) | | 模型 (10) | |
|-----------------------|----------------|--------------|----------------|--------------|
| | 被解释变量: default | | 被解释变量: default | |
| | 系数 | z-statistics | 系数 | z-statistics |
| constant | -2.764 *** | -22.822 | -3.364 *** | -13.376 |
| rate | 0.082 *** | 10.612 | 0.025 ** | 2.258 |
| month | | | 0.065 *** | 10.006 |
| risk level | | | 2.100 *** | 13.972 |
| credit line | | | 4.26e-06 *** | 7.073 |
| cert_num | | | -0.139 *** | -5.726 |
| pre2012 | | | -0.563 *** | -6.676 |
| Observations | 5465 | 5461 | | |
| Pseudo R ² | 0.046 | 0.361 | | |

Panel B: 投资者风险识别能力的检验 (OLS 回归)

| | 模型 (11) | | 模型 (12) | |
|------------------|------------|--------|-------------|-------|
| | 被解释变量: num | | 被解释变量: time | |
| | 系数 | t 值 | 系数 | t 值 |
| constant | 11.249 *** | 10.015 | 270.401 *** | 4.425 |
| default_risk_num | 9.782 *** | 3.860 | 403.636 *** | 2.927 |
| rate | -0.415 *** | -5.418 | 1.901 | 0.456 |
| money | 0.001 *** | 89.519 | 0.002 *** | 3.991 |
| Observations | 5,461 | 5,461 | | |
| R ² | 0.601 | 0.005 | | |

注: *** 代表在 1% 水平下显著; ** 代表在 5% 水平下显著; * 代表在 10% 水平下显著。

表 9 稳健性检验三: 以哑变量替代认证个数

Panel A: 利率与其他信息对违约风险的回归检验 (Probit 回归)

| | 模型 (13) | | 模型 (14) | |
|-------------------------|----------------|---------|----------------|---------|
| | 被解释变量: default | | 被解释变量: default | |
| | 系数 | t 值 | 系数 | t 值 |
| constant | -2.742 *** | -22.978 | -2.980 *** | -11.334 |
| rate | 0.082 *** | 10.680 | 0.021 * | 1.823 |
| month | | | 0.066 *** | 9.964 |
| credit line | | | 0.000 *** | 7.397 |
| risk | | | 2.199 *** | 15.758 |
| idcard_cert | | | -0.793 *** | -6.106 |
| credit_cert | | | -0.969 ** | -1.987 |
| workpermit_cert | | | 0.175 | 0.322 |
| income_cert | | | 0.019 | 0.027 |
| marriage_cert | | | 0.099 | 1.212 |
| household_register_cert | | | 0.008 | 0.064 |
| video_cert | | | -0.167 ** | -1.983 |
| mobile_cert | | | 0.063 | 0.593 |
| weibo_cert | | | -0.156 | -1.511 |

续表 9

| Panel A:利率与其他信息对违约风险的回归检验(Probit 回归) | | | | |
|--------------------------------------|---------------|-------|---------------|--------|
| | 模型(13) | | 模型(14) | |
| | 被解释变量:default | | 被解释变量:default | |
| | 系数 | t 值 | 系数 | t 值 |
| <i>car_cert</i> | | | -0.045 | -0.531 |
| <i>house_cert</i> | | | 0.038 | 0.426 |
| <i>pre2012</i> | | | -0.772*** | -7.890 |
| <i>Observations</i> | 5681 | 5677 | | |
| <i>R</i> ² | 0.044 | 0.390 | | |

| Panel B:投资者风险识别能力的检验(OLS 回归) | | | | |
|-----------------------------------|-----------|---------|------------|--------|
| | 模型(15) | | 模型(16) | |
| | 被解释变量:num | | 被解释变量:time | |
| | 系数 | t 值 | 系数 | t 值 |
| <i>constant</i> | 10.156*** | 7.191 | 227.486*** | 3.606 |
| <i>default_risk_{new}</i> | 7.248*** | 2.536 | 355.821*** | 2.787 |
| <i>rate</i> | -0.373*** | -3.853 | 2.711 | 0.626 |
| <i>money</i> | 0.001*** | 103.646 | 0.005*** | 11.108 |
| <i>Observations</i> | 5,677 | 5,677 | | |
| <i>R</i> ² | 0.658 | 0.024 | | |

注:***代表在1%水平下显著;**代表在5%水平下显著;*代表在10%水平下显著。

根据本文的分析,在非完全市场化的利率下,相同的利率可能对应不同的违约风险,如果投资者是聪明的,他们就应该能识别出不同的违约风险;同时由于他们无法控制利率来调节投资风险,必须寻求其他路径来规避风险。

实证结果表明:人人贷借贷平台非完全市场化的利率部分反映了借款人的违约风险;但仍有相当高比例的违约风险未被反映在利率当中——借款人的一些基本公开信息对于预测这部分风险具有一定帮助。进一步,我们发现市场中的投资者是比较聪明的,他们能够发现相同利率背后,没有被利率所包含的违约风险,在非完全市场化利率的环境下,投资者无法直接左右利率的设定,因此他们选择了其它风险规避方式,在本文中,我们发现这种规避方式能够直接反映在一个成功订单的参与人数上,间接反映在订单募资成功所需的竞标时间上。

以上研究有一定的政策启示。首先,在非市场化的利率环境下,尤其是当利率较低而不能完全反映投资风险的时候,需要更多的投资者以及更多的时间才能完成资金的募集——扩展来说,当投资者的数目不足以支撑项目资金需求的时候,非市场化的利率就会造成资金募集的失败。这是一种效率上的损失,也就是说,非市场化的利率有可能压制了市场中实际债权的存量,使得金融业被抑制而不能促进经济的有效增长。反过来说,利率的市场化改革则有助于优化社会资源配置,从而更好地让金融市场服务于经济增长。另一方面,从研究结果来看,我国网络信贷市场中的投资者具有一定的风险识别能力,这为我国的利率市场化提供了良好的市场基础。

参考文献

赵勇、朱武祥,2000:《上市公司兼并收购可预测性研究》,《经济研究》第4期。

Akerlof, G. A., 1970, "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism", *Quarterly Journal of Economics*, Vol. 84, 488—500.

- Bachmann, A., A. Becker, D. Buerckner, M. Hilker, M. Lehmann, and P. Tiburtius, 2011, "Online Peer-to-peer Lending-a Literature Review", *Journal of Internet Banking and Commerce*, Vol. 16, 1—18.
- Bagehot, W., and J. Treynor, 1971, "The Only Game in Town", *Financial Analysts Journal*, Vol. 27, 12—17.
- Duarte, J., S. Siegel, and L. Young, 2012, "Trust and Credit: the Role of Appearance in Peer-to-peer Lending", *Review of Financial Studies*, Vol. 25, 2455—2484.
- Herzenstein, M., S. Sonenshein, and U. M. Dholakia, 2011, "Tell Me a Good Story and I May Lend You My Money: The Role of Narratives in Peer-to-peer Lending Decisions", SSRN working paper.
- Herzenstein, M., R. Andrews, U. Dholakia, and E. Lyandres, 2008, "The Democratization of Personal Consumer Loans? Determinants of Success in Online Peer-to-peer Lending Communities", SSRN working paper.
- Klaftt, M., 2008, "Peer to Peer Lending: Auctioning Microcredits Over the Internet", SSRN working paper.
- Kumar, S., 2007, "Bank of One: Empirical Analysis of Peer-to-peer Financial Marketplaces", *Proceedings of the American Conference on Information Systems*, 1—8.
- Larrimoreet, L., L. Jiang, J. Larrimore, D. Markowitz, and S. Gorski, 2011, "Peer to Peer Lending: the Relationship Between Language Features, Trustworthiness, and Persuasion Success", *Journal of Applied Communication Research*, Vol. 39, 19—37.
- Lin, M., N. R. Prabhala, and S. Viswanathan, 2013, "Judging Borrowers by the Company They Keep: Social Networks and Adverse Selection in Online Peer-to-peer Lending", *Management Science*, Vol. 59, 17—35.
- Mckinnon, R., 1973, *Money and Capital in Economic Development*, Brookings Institution Press.
- Michal, J., 2012, "Do Unverifiable Disclosures Matter? Evidence from Peer-to-peer Lending", *Accounting Review*, Vol. 87, 1385—1413.
- Palepu, K. G., 1986, "Predicting Takeover Targets: A Methodological and Empirical Analysis", *Journal of Accounting and Economics*, Vol. 8, 3—35.
- Pope, D. G., and J. R. Sydnor, 2011, "What's in a Picture? Evidence of Discrimination from Prosper. Com", *Journal of Human Resources*, Vol. 46, 53—92.
- Puro, L., J. E. Teich, H. Wallenius, and J. Wallenius, 2010, "Borrower Decision Aid for People-to-people Lending", *Decision Support System*, Vol. 49, 52—60.
- Ravina, E., 2008, "The Effect of Beauty and Personal Characteristics in Credit Markets", Working Paper, Columbia University.
- Sharp, W. F., 1964, "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk", *Journal of Finance*, Vol. 19, 425—442.
- Shaw, E., 1973, *Financial Deepening in Economic Development*, Oxford University Press.

The Intelligent Investor: Not-Fully-Marketized Interest Rate and Risk Identify ——Evidence from P2P Lending

Liao Li, Li Mengran and Wang Zhengwei
(Tsinghua University)

Abstract: Interest rate marketization is a very important step in China's financial reform. However there is little research on exploring the market foundation of interest rate marketization from a micro-data perspective. Using data from a peer-to-peer lending website, this paper finds that the not-fully-marketized interest rate only can predict a small proportion of borrower's default risk. Furthermore, the paper finds that investors are intelligent enough to identify these different default risks indicated by the same interest rate, via evaluating borrowers' public information. This paper also finds that the risk identify ability is directly reflected in numbers of participants in completed biddings, and indirectly reflected in time required to raise these biddings. The results suggest investors in China's credit market have the basic ability in risk identifying, which can form the market basis of China's interest rate marketization.

Key Words: Not-fully-Marketized Interest Rate; Default Risk; Risk Identify; P2P Lending

JEL Classification: G11, G21, G23

(责任编辑:王利娜)(校对:张 涵)