大数据一定促进普惠金融实现吗？

——基于契约设计理论视角

洪洁瑛 武玙璠 李三希

目录

附录Ⅰ 关于情形 1 的证明

激励相容条件(1)可以写成下面的形式:

$$U_H^0 \geq U_L^0 + (\theta_H - \theta_L)(R - R_L), \tag{Ⅰ1}$$

$$U_L^0 \geq U_H^0 - (\theta_H - \theta_L)(R - R_H), \tag{Ⅰ2}$$

参与条件(2)可以写成下面的形式:

$$U_H^0 \geq W_H, \tag{Ⅰ3}$$

$$U_L^0 \geq W_L, \tag{Ⅰ4}$$

同时,有限责任条件为:

$$0 \leq R_H \leq R, \tag{Ⅰ5}$$

$$0 \leq R_L \leq R, \tag{Ⅰ6}$$

投资条件$L_H \geq I$和$L_L \geq I$可以写为:

$$U_H^0 \geq \theta_H (R - R_H), \tag{Ⅰ7}$$

$$U_L^0 \geq \theta_L (R - R_L), \tag{Ⅰ8}$$

信贷机构在这些限制条件下设计契约最大化他的效用(3)。该公式也可以写成

$V^0 = \frac{1}{2}\left(\theta_H R - I - U_H^0\right) + \frac{1}{2}\left(\theta_L R - I - U_L^0\right)$。

先求解没有约束条件下的最优解为$U_H^0 = W_H$和$U_L^0 = W_L$。通过考察约束条件,发现在$W_H \theta_L > W_L \theta_H$条件下,其他约束条件都成立,然而条件(Ⅰ2)和(Ⅰ7)无法同时成立,因而无法得到社会最优结果。

因此,猜测条件(Ⅰ2)和(Ⅰ7)恰好满足,后面会在求解最优化的结果后再回头判断其他约束条件是否成立。在(Ⅰ2)和(Ⅰ7)恰好满足时,得到$U_L^0 = \theta_L (R - R_H) = \frac{\theta_L}{\theta_H} U_H^0$,并带入信贷机构的效用函数,得到$V^0 = \frac{1}{2}\left(\theta_H R - I - U_H^0\right) + \frac{1}{2}\left(\theta_L R - I - \frac{\theta_L}{\theta_H} U_H^0\right)$。$V^0$对$U_H^0$求一阶导,得到下面的方程:$\frac{\partial V^0}{\partial U_H^0} = -\frac{1}{2}\left(1 + \frac{\theta_L}{\theta_H}\right) < 0$。

从上面的一阶导的结果,得到$U_H^0 = W_H$。根据$U_H^0 = \theta_H (R - R_H)$,可以求解得到$R_H = R - \frac{W_H}{\theta_H}$, $L_H = I$。在(Ⅰ2)式恰好满足时,可以解得 $U_L^0 = \frac{\theta_L}{\theta_H} W_H$、 $R_L = R - \frac{W_H}{\theta_H}$以及$L_L = I$。在这一解下,发现其他约束条件都成立。信贷机构的期望效用为$V^0 = \frac{1}{2}(\theta_H R - I - W_H) + \frac{1}{2}\left(\theta_L R - I - \frac{\theta_L}{\theta_H} W_H\right)$。

附录Ⅱ 未引进大数据时其他三种情形下的信贷机构最优契约设计求解

情形 2： $L_H \geq I$ 和 $L_L < I$。信贷机构只给高质量借款方提供足够的贷款额度，而给低质量借款方的额度小于其投资需求。低质量的借款方由于资金不足导致项目无法开展，信贷机构在这种情况下无法获得任何回报。此时，他的最优选择是给予低质量借款方的贷款额度设为 0，即 $L_L=0$。因此，第二种情况我们考虑 $L_H \geq I$ 和 $L_L=0$，分析也不会丧失一般性。

在第二种情况下，借方的参与约束条件为：

$$U_H^0=\theta_H(R-R_H)+(L_H-I)\geq W_H, \tag{Ⅱ1}$$

以及激励相容条件为：

$$W_L \geq \theta_L(R-R_H)+(L_H-I)=U_H^0-(\theta_H-\theta_L)(R-R_H)。 \tag{Ⅱ2}$$

公式（Ⅱ 1）确保了高质量借款方选择符合其真实类型的信贷契约时，所获得的效用至少不低于其外部效用。公式（Ⅱ 2）则表明，对于低质量借款方来说，他们选择不参与信贷更为有利，因为其外部效用超过模仿高质量借款方获得信贷所带来的效用。简而言之，高质量借款方会选择获得融资开展项目，而低质量借款方将选择放弃贷款。此外，还需要满足其他条件：$0 \leq R_H \leq R$ 代表借款方受到有限责任保护，以及 $L_H \geq I$ 代表高质量借款方获得足够融资。

信贷机构在满足（Ⅱ 1），（Ⅱ 2），（Ⅰ 5），（Ⅰ 7）条件下设计契约最大化自己的效用，其效用函数如下：

$$V^0=\frac{1}{2}\left(\theta_H R-I-U_H^0\right), \tag{Ⅱ3}$$

根据条件（Ⅱ 2）和（Ⅰ 7），得到下面的不等式：

$$\theta_H(R-R_H)\leq W_L+(\theta_H-\theta_L)(R-R_H), \tag{Ⅱ4}$$

通过求解不等式（Ⅱ 4），得到 $R_H \geq R-\frac{W_L}{\theta_L}$。同时，根据条件（Ⅱ 2）可以得到 $U_H^0 \leq \frac{\theta_H W_L}{\theta_L}$。因为 $W_H\theta_L > W_L\theta_H$，因此，$U_H^0 < W_H$。这与约束条件（Ⅱ 1）相悖。因此，不存在解使得所有的约束条件成立。这意味着，信贷机构无法设计出一个契约，使得只有高质量借款方获得贷款，而低质量贷款方被排除在外。

情形 3： $L_L \geq I$ 和 $L_H < I$。信贷机构向低质量借款方提供足够的贷款额度，而给高质量借款方的额度小于其投资需求。同样的，第三种情况我们只考虑 $L_L \geq I$ 和 $L_H=0$ 这特定情形，分析不会丧失一般性。

在第三种情况下，借方的参与约束条件和激励相容条件分别为：

$$U_L^0=\theta_L(R-R_L)+(L_L-I)\geq W_L, \tag{Ⅱ5}$$

$$W_H \geq \theta_H(R-R_L)+(L_L-I)=U_L^0+(\theta_H-\theta_L)(R-R_L), \tag{Ⅱ6}$$

公式（Ⅱ 5）代表低质量借款方选择自己真实类型相对应的契约所获得的效用大于等于外部效用。公式（Ⅱ 6）代表高质量借款方的外部效用超过模仿低质量借款方获得贷款所带来的效用。简而言之，低质量借款方会选择获得融资开展项目，而高质量借款方将选择放弃贷款。此外，依然需要满足其他条件：$0 \leq R_L \leq R$ 代表借款方受到有限责任保护，以及 $L_L \geq I$ 代表低质量借款方获得足够融资。

信贷机构在满足（Ⅱ 5），（Ⅱ 6），（Ⅰ 6），（Ⅰ 8）条件下设计契约最大化自己的效用。其效用函数如下：

$$V^0=\frac{1}{2}\left(\theta_L R-I-U_L^0\right), \tag{Ⅱ7}$$

通过最优化求解，我们得到最优契约为$U_L^0=W_L$，$R_L=R-\frac{W_L}{\theta_L}$， $I_L=I$，以及$U_H^0=W_H$，$R_H=0$， $I_H=0$。信贷机构的期望效用为$V^0=\frac{1}{2}(\theta_L R-I-W_L)$。

情形 4： $L_i<\mathrm{I}$， 对于任意的 $i=H, L$。信贷机构提供给两种类型借款方的贷款额度都小于其投资需求，借款方无法开展项目。同样的，我们只考虑$L_i=0$ 这特定情形，分析不会丧失一般性。在第四种情况下，借贷不会发生，信贷机构效用为 0。

经过对上述四种情况的详细分析，信贷机构将对比不同情况下所能获得的期望效用，选择能够最大化其效用的信贷契约设计。由于情形 2 无解，即不存在一个契约能使高质量借款方获得贷款而将低质量借款方排除在外，同时情形 4 情况下信贷机构获得的期望效用为 0。因此，信贷机构将主要在情形 1 和情形 3 之间进行比较。通过计算，可以得到：当 $X\geq 1$，信贷机构选择第一种情况契约；当 $X<1$，信贷机构选择第三种情况契约。

附录Ⅲ 引进大数据时的信贷机构最优契约设计求解及证明

情形 1：$L_i^j \geq I$，其中 $i,j=H,\ L$，即信贷契约给两种类型的借款方都提供足够的贷款额度。

基于数据信息 η_j，信贷机构的效用函数如下：

$$V^j=V\left(L_i^j,\ R_i^j \middle| \eta_j\right)=p\left(\theta_j R_j^j - L_j^j\right)+(1\text{-}p)\left(\theta_{-j} R_{-j}^j - L_{-j}^j\right), \tag{Ⅲ 1}$$

当信贷机构获取的数据信息为 η_j 时，他将以概率 p 遇到类型为 j 的借款方，并获得期望收入 $\theta_j R_j^j\text{-}L_j^j$；以概率 1-$p$ 碰到类型-j 的借款方，并获得期望收入 $\theta_{-j} R_{-j}^j\text{-}L_{-j}^j$。因此，信贷机构的期望效用如等式（Ⅲ 1）所示。

对于任意 $i,\ j=H,\ L$，借款方选择信贷契约需要满足下面的激励相容条件：

$$\theta_i\left(R\text{-}R_i^j\right)+(L_i^j\text{-}I)\geq\theta_i\left(R\text{-}R_{-i}^j\right)+(L_{-i}^j\text{-}I), \tag{Ⅲ 2}$$

其中，$\theta_i\left(R\text{-}R_i^j\right)+\left(L_i^j\text{-}I\right)$ 是类型 i 的借款方选择与自己真实类型相对应的信贷契约 $\left(R_i^j, L_i^j\right)$ 所获得的期望效用。$\theta_i\left(R\text{-}R_{-i}^j\right)+\left(L_{-i}^j\text{-}I\right)$ 是类型 i 的借款方选择另一种类型对应的信贷契约 $\left(R_{-i}^j, L_{-i}^j\right)$ 所获得的期望效用。因此，激励相容条件（Ⅲ 2）保证了在数据信息 η_j 下的契约中，借款方选择与自己真实类型相对应的信贷契约总能获得更高的期望效用，从而没有动机去伪装成其他类型。

同样的，对于任意 $i,\ j=H,\ L$，借款方选择信贷契约需满足下面的参与条件：

$$\theta_i\left(R\text{-}R_i^j\right)+\left(L_i^j\text{-}I\right)\geq W_i, \tag{Ⅲ 3}$$

参与条件保证借款方选择自己真实类型对应的契约所获得的期望效用大于等于其外部效用。此外，还需要满足两个其他条件：$0\leq R_i^j\leq R$ 代表借款方受到有限责任保护，$L_i^j\geq I$ 代表项目可以获得足够融资。当信贷机构的数据信号为 η_j 时，类型 i 的借款方的期望效用为 $U_i^j=\theta_i\left(R\text{-}R_i^j\right)+(L_i^j\text{-}I)$。

当数据信息为 η_j 时，信贷机构对借款方类型分布进行贝叶斯更新，在满足借款方的激励相容条件、参与条件和其他约束条件的情况下，设计信贷契约菜单 $\left(L_i^j,\ R_i^j\right)$ 最大化自己的效用（Ⅲ 1）。

情形 2 和情形 3：$L_i^j\geq I$ 和 $L_{-i}^j=0$，其中 $i,j=H,\ L$（$i=H$ 对应第二种情况，$i=L$ 对应第三种情况）。借方的激励相容和参与条件为：

$$\theta_i\left(R\text{-}R_i^j\right)+\left(L_i^j\text{-}I\right)\geq W_i, \tag{Ⅲ 4}$$

$$W_{-i}\geq\theta_{-i}\left(R\text{-}R_i^j\right)+\left(L_i^j\text{-}I\right), \tag{Ⅲ 5}$$

公式（Ⅲ 4）代表类型 i 的借款方选择自己真实类型相对应的契约所获得的效用大于等于外部效用。公式（Ⅲ 5）代表类型-i 的借款方外部效用超过模仿其他类型借款方获得贷款所带来的效用。这表明，类型 i 借款方获得融资开展项目，而类型-i 借款方选择不要贷款。同时，依然需要满足条件 $0\leq R_i^j\leq R$ 以及 $L_i^j\geq I$，$L_{-i}^j=0$。

当数据信息 η_j 时，信贷机构对借款方类型分布进行贝叶斯更新，在满足（Ⅲ 4）、（Ⅲ 5）和其他约束条件下，设计信贷契约菜单 $\left(L_i^j,\ R_i^j\right)$ 最大化自己的效用，其效用函数如下：

$$V^j=V\left(L_i^j,\ R_i^j \middle| \eta_j\right)=p\left(\theta_i R_i^j\text{-}L_i^j\right)1_{j=i}+(1\text{-}p)\left(\theta_i R_i^j\text{-}L_i^j\right)1_{j\neq i}, \tag{Ⅲ 6}$$

其中，$1_{j=i}=1$ 如果 $j=i$，否则为 0；$1_{j\neq i}=1$ 如果 $j\neq i$，否则为 0。

情形 4：$L_i^j=0$，对于任意的 $i,j=H,\ L$。信贷机构不给借款方提供贷款。借款方将获得外部效用，而信贷机构的期望效用为 0。

证明：首先求解当信贷机构的数据信息是 η_H 时的信贷契约。

先证明情形 1 的结论，即 $L_H\geq I$ 和 $L_L\geq I$，最优化过程如下：

激励相容条件（Ⅲ 2）可以写成下面的形式：

$$U_H^H\geq U_L^H+(\theta_H\text{-}\theta_L)(R\text{-}R_L^H), \tag{Ⅲ 7}$$

$$U_L^H\geq U_H^H\text{-}(\theta_H\text{-}\theta_L)(R\text{-}R_H^H), \tag{Ⅲ 8}$$

参与条件（Ⅲ 3）可以写成下面的形式：

$$U_H^H \geq W_H, \tag{Ⅲ 9}$$

$$U_L^H \geq W_L, \tag{Ⅲ 10}$$

同时，有限责任条件为：

$$0 \leq R_H^H \leq R, \tag{Ⅲ 11}$$

$$0 \leq R_L^H \leq R, \tag{Ⅲ 12}$$

投资条件$L_H \geq I$和$L_L \geq I$可以写为：

$$U_H^H \geq \theta_H (R - R_H^H), \tag{Ⅲ 13}$$

$$U_L^H \geq \theta_L (R - R_L^H), \tag{Ⅲ 14}$$

信贷机构在这些限制条件下设计契约最大化他的效用，即公式（Ⅲ 1），该公式也可以写成 $V^H = p(\theta_H R - I - U_H^H) + (1-p)(\theta_L R - I - U_L^H)$。

先求解没有约束条件下的最优解为$U_H^H = W_H$和$U_L^H = W_L$。通过考察约束条件，发现当$W_H \theta_L > W_L \theta_H$时，其他约束条件都成立，然而条件（Ⅲ 8）和（Ⅲ 13）无法同时成立，因而无法得到社会最优结果。

当$W_H \theta_L > W_L \theta_H$时，猜测条件（Ⅲ 8）和（Ⅲ 13）恰好满足的，后面会在求解优化结果后再回头判断是否成立。（Ⅲ 8）和（Ⅲ 13）恰好满足，得到$U_L^H = \theta_L (R - R_H^H) = \frac{\theta_L}{\theta_H} U_H^H$，并代入信贷机构的效用函数，得到 $V^H = p(\theta_H R - I - U_H^H) + (1-p)\left(\theta_L R - I - \frac{\theta_L}{\theta_H} U_H^H\right)$。$V^H$对$U_H^H$求一阶导，得到：$\frac{\partial V^H}{\partial U_H^H} = -\left(p + (1-p)\frac{\theta_L}{\theta_H}\right) < 0$。

从上面的一阶导的结果，得到$U_H^H = W_H$。根据$U_H^H = \theta_H (R - R_H^H)$，可以求解得到$R_H^H = R - \frac{W_H}{\theta_H}$，$L_H^H = I$。根据$U_L^H = U_H^H - (\theta_H - \theta_L)(R - R_H^H)$，可以解得 $U_L^H = \frac{\theta_L}{\theta_H} W_H$、$R_L^H = R - \frac{W_H}{\theta_H}$以及$L_L^H = I$。在这一解下，发现其他约束条件都成立。信贷机构的期望效用为$V^H = p(\theta_H R - I - W_H) + (1-p)\left(\theta_L R - I - \frac{\theta_L}{\theta_H} W_H\right)$。

其次，证明情形 2 中的结论，条件（Ⅲ 4）和（Ⅲ 5）可以写为：

$$U_H^H \geq W_H, \tag{Ⅲ 15}$$

$$W_L \geq U_H^H - (\theta_H - \theta_L)(R - R_H^H), \tag{Ⅲ 16}$$

其他条件如（Ⅲ 11），（Ⅲ 13）。在这些约束条件下，信贷机构设计契约最大化效用（Ⅲ 6）。同时，可以将公式（Ⅲ 6）写成$V^H = p(\theta_H R - I - U_H^H)$。

根据条件（Ⅲ 16）和（Ⅲ 14），得到下面的不等式：

$$\theta_H (R - R_H^H) \leq W_L + (\theta_H - \theta_L)(R - R_H^H), \tag{Ⅲ 17}$$

通过求解不等式（Ⅲ 17），得到$R_H^H \geq R - \frac{W_L}{\theta_L}$。同时，根据条件（Ⅲ 11）得到$U_H^H \leq \frac{\theta_H W_L}{\theta_L}$。因为$W_H \theta_L > W_L \theta_H$，因此，$U_H^H < W_H$。这与约束条件（Ⅲ 15）相悖。因此，不存在解能使得所有的约束条件成立。也就是说，第二种情况不存在解。

再次，考察第三种情况。条件（Ⅲ 4）和（Ⅲ 5）可以写为：

$$U_L^H \geq W_L, \tag{Ⅲ 18}$$

$$W_H \geq U_L^H + (\theta_H - \theta_L)(R - R_L^H), \tag{Ⅲ 19}$$

其他条件如（Ⅲ 12）和（Ⅲ 14）。在这些约束条件下，信贷机构设计契约最大化效用（Ⅲ 6）。同时，可以将公式（Ⅲ 6）写成$V^H=(1-p)(\theta_L R-I-U_L^H)$。先求解没有约束条件下的最优解为$U_L^H=W_L$。后续考察约束条件。

约束条件(Ⅲ 19)得到$R_L^{\mathrm{H}}\geq R-\frac{W_H-W_L}{\theta_H-\theta_L}$,约束条件(Ⅲ 14)得到$R_L^{\mathrm{H}}\geq R-\frac{W_L}{\theta_L}$。当$W_H\theta_L>W_L\theta_H$时，$R-\frac{W_H-W_L}{\theta_H-\theta_L}\leq R-\frac{W_L}{\theta_L}$。

因此，$R_L^{\mathrm{H}}\geq R-\frac{W_L}{\theta_L}$。同时，发现其他约束条件也都成立。因此，得到最优契约为$U_L^H=W_L$，$R_L^{\mathrm{H}}=R-\frac{W_L}{\theta_L}$，$I_L^H=I$，以及$U_H^H=W_H$，$R_H^{\mathrm{H}}=0$，$I_H^H=0$。信贷机构的期望效用为$V^H=(1-p)(\theta_L R-I-W_L)$。

最后，第四种情况代表信贷机构将不会放款，借款方无法开展项目，获得外部选择效用。信贷机构效用为 0。

因为第四种情况信贷机构获得的期望效用为 0。第三种情况的效用大于等于 0。因此，信贷机构不会选择完全不放贷。而是在第一和第三种情况之间进行比较。通过计算，可以得到：当$X\geq\frac{1-p}{p}$，信贷机构选择第一种情况契约；当$X<\frac{1-p}{p}$，信贷机构选择第三种情况契约。

类似的，可以推导得到当信贷机构的数据信息是η_L时的信贷契。Q.E.D.

附录Ⅳ 对命题 1 的证明

① 当 $X<\frac{1-p}{p}$或 $X>\frac{p}{1-p}$时，也就是说 $X<1$ 且$\frac{1}{2}<p<\frac{1}{1+X}$，或 $X\geq1$ 且$\frac{1}{2}<p<\frac{X}{1+X}$两种情况时，无论信贷机构是否有数据，将提供同样的信贷契约。

② 当$\frac{1-p}{p}\leq X<1$ 时，也就是说 $X<1$ 且 $p\geq\frac{1}{1+X}$时，若信贷机构没有借款方数据，高质量借款方无法获得借款，低质量借款方获得信贷额度为 I，还款额度为 $R-\frac{W_L}{\theta_L}$，因此对应贷款利率为$r_L=\frac{R-\frac{W_L}{\theta_L}}{I}-1$。如果信贷机构获得数据信息为$\eta_H$，信贷契约为$R_H^H=R_L^H=R-\frac{W_H}{\theta_H}$，$L_H^H=L_L^H=I$。两种类型的借款方都获得信贷额度 I。两种类型的借款方获得利率$r_H^H=r_L^H=r_L=\frac{R-\frac{W_H}{\theta_H}}{I}-1$。同时，$r_L^H<r_L$。如果信贷机构获得数据信息为$\eta_L$，借款方获得的信贷契约与没有数据信息一样。

③ 当 $1\leq X\leq\frac{p}{1-p}$ 时，也就是说 $X\geq1$ 且 $p\geq\frac{X}{1+X}$时，若信贷机构没有借款方数据，两类借款方均获得信贷额度为 I 和利率$r_H=r_L=\frac{R-\frac{W_H}{\theta_H}}{I}-1$。如果信贷机构获得数据信息$\eta_L$，信贷契约为$R_H^L=0$，$L_H^L=0$；$R_L^L=R-\frac{W_L}{\theta_L}$，$L_L^L=I$。高质量借款方无法获得贷款，低质量借款方的信贷额度$L_L^L=I$和利率$r_L^L=\frac{R-\frac{W_L}{\theta_L}}{I}-1$。因此低质量借款方获得的利率更高。如果信贷机构获得数据信息η_H，借款方获得的信贷契约与没有数据时一样。Q.E.D.

附录Ⅴ 对命题 2 的证明

当 $X>\frac{p}{1-p}$或 $X<\frac{1-p}{p}$时，信贷机构所提供的契约与数据信息无关。因此，信贷机构在有无数据的两种不同情况下都获得相同的期望效用。

当 $1\leq X\leq\frac{p}{1-p}$时，如果信贷机构没有数据信息，其效用为$V^0=\frac{1}{2}\left(\theta_H R-I-W_H\right)+\frac{1}{2}\left(\theta_L R-I-\frac{\theta_L}{\theta_H}W_H\right)$。如果信贷机构有数据信息时，$\frac{1}{2}$概率会获得数据信息$\eta_H$，其期望效用为$V_H=p(\theta_H R-I-W_H)+(1-p)\left(\theta_L R-I-\frac{\theta_L}{\theta_H}W_H\right)$；其余$\frac{1}{2}$概率会获得信息$\eta_L$，其期望效用为$V_L=p(\theta_L R-I-W_L)$。因此，在有数据信息的情况下，信贷机构的期望效用为$\frac{1}{2}V_H+\frac{1}{2}V_L$。对比两种情况下信贷机构的期望效用可以发现，信贷机构在有数据情况下获得的期望效用更高，即$\frac{1}{2}V_H+\frac{1}{2}V_L-V^0>0$。①

当$\frac{1-p}{p}\leq X<1$ 时，证明思路类似。容易证明信贷机构在有数据情况下可以获得更高的期望效用。因此，通过比较信贷机构拥有数据信息和没有数据信息两种情况下的期望效用，可以得出大数据能够提高信贷机构的期望效用的结论。Q.E.D.

① $\frac{1}{2}V_H+\frac{1}{2}V_L-V^0=\frac{1}{2}\left[p\left(\theta_H R-I-W_H\right)+(1-p)\left(\theta_L R-I-\frac{\theta_L}{\theta_H}W_H\right)\right]+\frac{1}{2}p\left(\theta_L R-I-W_L\right)-\frac{1}{2}\left(\theta_H R-I-W_H\right)-\frac{1}{2}\left(\theta_L R-I-\frac{\theta_L}{\theta_H}W_H\right)$

$=\frac{1}{2}p\left(\theta_L R-I-W_L\right)-\frac{1}{2}(1-p)\left(\theta_H R-I-W_H\right)-\frac{1}{2}\left(\theta_L R-I-\frac{\theta_L}{\theta_H}W_H\right)=\frac{1}{2}p\left(\frac{\theta_L}{\theta_H}W_H-W_L\right)-\frac{1}{2}(1-p)\left(\theta_H R-I-W_H\right)$

$=\frac{1}{2}\left(\frac{\theta_L}{\theta_H}W_H-W_L\right)[p-(1-p)X]>0$

附录Ⅵ 对命题 3 的证明

证明：关于高质量借款方，当 $X\geq 1$ 时，在有数据信息时，如果$\frac{1}{2}<p<\frac{X}{1+X}$，高质量借款方获得融资的概率为 1。如果 $p\geq\frac{X}{1+X}$，高质量借款方只有在信贷机构数据信息为η_H时获得融资，其融资概率为 p。因此，数据信息准确度 p 对其融资概率的影响是非单调的，融资概率在 $p=\frac{X}{1+X}$时取值最小。当 $X<1$ 时，如果$\frac{1}{2}<p<\frac{1}{1+X}$时，高质量借款方获得贷款的概率为 0。当 $p\geq\frac{1}{1+X}$时，高质量借款方只有在信贷机构数据信息为η_H时获得融资，其融资概率为 p。因此，数据信息准确度 p 的增加会提高融资概率。

关于低质量借款方，当 $X\geq 1$ 时，在有数据信息时，如果$\frac{1}{2}<p<\frac{X}{1+X}$，低质量借款方的还款额度为 $R\text{-}\frac{W_H}{\theta_H}$。如果 $p\geq\frac{X}{1+X}$，低质量借款方在信贷机构数据信息为η_H时还款额度为 $R\text{-}\frac{W_H}{\theta_H}$；在信贷机构数据信息为$\eta_L$时还款额度为 $R\text{-}\frac{W_L}{\theta_L}$。因而，低质量借款热的期望还款额为$(1\text{-}p)\left(R\text{-}\frac{W_H}{\theta_H}\right)+p\left(R\text{-}\frac{W_L}{\theta_L}\right)=R\text{-}\frac{W_H}{\theta_H}+p\left(\frac{W_H}{\theta_H}\text{-}\frac{W_L}{\theta_L}\right)$。这表明，数据信息准确度 p 的增加会提高其期望融资成本（还款额）。当 $X<1$ 时，如果$\frac{1}{2}<p<\frac{1}{1+X}$时，低质量借款方的还款额度为 $R\text{-}\frac{W_L}{\theta_L}$。当 $p\geq\frac{1}{1+X}$时，低质量借款方在信贷机构数据信息为η_H时还款额度为 $R\text{-}\frac{W_H}{\theta_H}$；在信贷机构数据信息为 η_L 时还款额度为 $R\text{-}\frac{W_L}{\theta_L}$。因而，低质量借款方的期望还款额依然为 $(1\text{-}p)\left(R\text{-}\frac{W_H}{\theta_H}\right)+p\left(R\text{-}\frac{W_L}{\theta_L}\right)=R\text{-}\frac{W_H}{\theta_H}+p\left(\frac{W_H}{\theta_H}\text{-}\frac{W_L}{\theta_L}\right)$。这表明，数据信息准确度 p 对其期望融资成本的影响是非单调的，该成本在 $p=\frac{1}{1+X}$时最小。Q.E.D.

附录Ⅶ 对命题 4 的证明

证明：当 $X\geq 1$ 时，在没有数据信息情况下，高质量和低质量借款方均获得合约（I，$R-\frac{W_H}{\theta_H}$），高质量的借款方获得效用 W_H，低质量的借款方获得效用 $\frac{\theta_L}{\theta_H}W_H$。因此，借款方的平均效用为 $\frac{1}{2}\left(W_H+\frac{\theta_L}{\theta_H}W_H\right)$。

当 $X\geq 1$ 时，如果 $\frac{1}{2}<p<\frac{X}{1+X}$，借款方的效用如没有数据信息时一致为 $\frac{1}{2}\left(W_H+\frac{\theta_L}{\theta_H}W_H\right)$。如果 $p\geq\frac{X}{1+X}$ 以及信贷机构获得信号 η_L 时，低质量的借款方的效用为下降为 W_L，而高质量的贷款人的效用一直为 W_H。如果 $p\geq\frac{X}{1+X}$ 以及信贷获得信号 η_H 时，低质量的借款方的效用依然为 $\frac{\theta_L}{\theta_H}W_H$，而高质量的贷款人的效用一直为 W_H。因此，如果借款方是低质量的，信贷机构获得信号 η_L 的概率为 p，获得信号 η_H 的概率为 1-p。因此，低质量的借款方的期望效用为 $pW_L+(1-p)\frac{\theta_L}{\theta_H}W_H$。此时，借款方的平均效用为 $\frac{1}{2}W_H+\frac{1}{2}(pW_L+(1-p)\frac{\theta_L}{\theta_H}W_H)=\frac{1}{2}W_H+\frac{1}{2}\frac{\theta_L}{\theta_H}W_H-\frac{1}{2}p\left(\frac{\theta_L}{\theta_H}W_H-W_L\right)$。因为 $W_L<\frac{\theta_L}{\theta_H}W_H$，所以随着数据精度 p 越高，低质量借款方的期望效用越小，借款方的平均效用也越小。我们得到，当 $X\geq 1$，数据信息准确度 p 的增加会损害借款方的利益。

当 $X<1$ 时，在没有数据信息的情况下，高质量借款方的期望效用为 W_H，而低质量借款方的期望效用为 W_L，因此，借款方的平均效用为 $\frac{1}{2}(W_H+W_L)$。在有数据信息的情况下，如果 $\frac{1}{2}<p<\frac{1}{1+X}$，借款方的效用如没有数据信息时一致为 $\frac{1}{2}(W_H+W_L)$。如果 $p\geq\frac{1}{1+x}$ 以及信贷机构得到信号为 η_L 时，高质量借款方的期望效用为 W_H，而低质量借款方的期望效用为 W_L。如果 $p\geq\frac{1}{1+x}$ 以及信贷机构得到信号为 η_H 时，低质量的借款方会得到更高的期望效用 $\frac{\theta_L}{\theta_H}W_H$，而高质量的贷款人的效用一直为 W_L。如果借款方是低质量的，信贷机构获得信号 η_L 的概率为 p，获得信号 η_H 的概率为 1-p。因此，当 $p\geq\frac{1}{1+x}$ 时，低质量的借款方的期望效用为 $pW_L+(1-p)\frac{\theta_L}{\theta_H}W_H$，借款方的平均效用为 $=\frac{1}{2}W_H+\frac{1}{2}\frac{\theta_L}{\theta_H}W_H-\frac{1}{2}p\left(\frac{\theta_L}{\theta_H}W_H-W_L\right)$。因此，随着数据精度 p 增加，借款方的期望效用呈现非单调变化。当数据的精度增加到 $p=\frac{1}{1+X}$ 时，借款方的期望效用达到最大。如果 p 小于这一值，借款方获得效用效用更低，如果 p 大于这一值，信号的精确会减少其期望效用。Q.E.D。

附录Ⅷ 对命题 5 的证明

当 $X\geq 1$ 时，如果$\frac{1}{2}<p<\frac{X}{1+X}$，信贷机构提供的契约不受数据信息的影响。其中，$\frac{1}{2}$概率借给高质量借款方，获得效用 $\theta_H R\text{-}W_H\text{-}I$；$\frac{1}{2}$概率借给低质量借款方，获得效用 $\theta_L R\text{-}\frac{\theta_L}{\theta_H}W_H\text{-}I$。因此，期望效用为 $\frac{1}{2}\left(\theta_L R\text{-}\frac{\theta_L}{\theta_H}W_H\text{-}I\right)+\frac{1}{2}(\theta_H R\text{-}W_H\text{-}I)$。如果 $p\geq\frac{X}{1+X}$，信贷机构面临四种可能情景。

情景一：信贷机构得到信号η_H，且借款方为高质量，该情景概率为$\frac{1}{2}p$，信贷机构获得$\theta_H R\text{-}W_H\text{-}I$。

情景二：信贷机构得到信号η_H，且借款方为低质量，该情景概率为$\frac{1}{2}(1\text{-}p)$，信贷机构获得$\theta_L R\text{-}\frac{\theta_L}{\theta_H}W_H\text{-}I$。

情景三：信贷机构得到信号η_L，且借款方为高质量，该情景概率为$\frac{1}{2}(1\text{-}p)$，信贷机构获得 0。

情景四：信贷机构得到信号η_L，且借款方为低质量，该情景概率为$\frac{1}{2}p$，信贷机构获得$\theta_L R\text{-}W_L\text{-}I$。

因此，信贷机构的期望效用为

$$\frac{1}{2}p(\theta_H R\text{-}W_H\text{-}I)+\frac{1}{2}(1\text{-}p)\left(\theta_L R\text{-}\frac{\theta_L}{\theta_H}W_H\text{-}I\right)+\frac{1}{2}p(\theta_L R\text{-}W_L\text{-}I)=\frac{1}{2}\left(\theta_L R\text{-}\frac{\theta_L}{\theta_H}W_H\text{-}I\right)+\frac{1}{2}p\left(\theta_H R\text{-}W_H\text{-}I+\frac{\theta_L}{\theta_H}W_H\text{-}W_L\right)$$

。当 $p=\frac{X}{1+X}$，$\frac{1}{2}\left(\theta_L R\text{-}\frac{\theta_L}{\theta_H}W_H\text{-}I\right)+\frac{1}{2}p\left(\theta_H R\text{-}W_H\text{-}I+\frac{\theta_L}{\theta_H}W_H\text{-}W_L\right)=\frac{1}{2}\left(\theta_L R\text{-}\frac{\theta_L}{\theta_H}W_H\text{-}I\right)+\frac{1}{2}(\theta_H R\text{-}W_H\text{-}I)$。此外，因为$\theta_H R\text{-}W_H\text{-}I+\frac{\theta_L}{\theta_H}W_H\text{-}W_L>0$，所以 p 的增加会提高信贷机构的期望效用。

当 $X<1$ 时，类似的，可以证明 p 的增加会提高信贷机构的期望效用。

附录Ⅸ 对命题 6 的证明

① （d_H=0，d_L=0)是均衡策略。

假设θ_H和θ_L的借款方选择不同的数据披露策略，即(d_H=1，d_L=0)和(d_H=0，d_L=1)是均衡策略，信贷机构在第二期决策中通过借款方的数据披露策略会准确得知借款方的真实类型。信贷机构的最优契约设计将如帕累托最优结果，即$R_H=R-\frac{W_H}{\theta_H}$、$R_L=R-\frac{W_L}{\theta_L}$和$L_H=L_L=I$。在这种情况下，低质量借款方有激励将信息披露策略从d_L改为d_H，以此获得更高的效用。因此， $d_H \neq d_L$不是均衡策略。

下面证明（d_H=0，d_L=0)是均衡策略。对于高质量借款方，他没有激励改变这一政策。如果他选择d_H=0，信贷机构将会按照原始预期提供信贷契约（定理 1），可以求解得到高质量借款方的期望效用为W_H。如果他改变策略选择d_H=1，这种情况下，信贷机构基于均衡外的贝叶斯预期，即原始的预期，作出最优契约设计。高质量类型的借款方在第一期的期望效用依然为W_H。因此，可以得到高质量类型的借款方并没有激励去改变他的选择。

对于低质量借款方，她也没有激励改变这一政策。如果他选择d_L=0，信贷机构会按照原始预期提供信贷契约（定理 1），可以得到：当 $X \geq 1$ 时，低质量借款方的期望效用为$\frac{\theta_L}{\theta_H}W_H$；当 $X<1$ 时，低质量借款方在第一期的期望效用为W_L。如果他选择d_L=1，信贷机构基于均衡外的贝叶斯预期，即原始的预期，作出最优契约设计，那么他的效用依然不会变化。因而也没有激励去改变。因此，（d_H=0，d_L=0)是均衡策略。

② 当 X<1 时，（d_H=1，d_L=1)是均衡策略。

对于高质量借款方，如果他选择d_H=1，信贷机构会基于数据信息更新其分布情况，并提供合适的信贷契约（如定理 2），高质量的借款方在第一期的期望效用为W_H。如果高质量借款方改为d_H=0，那么信贷机构会基于均衡外的贝叶斯预期，即原始的预期，作出最优契约设计。高质量借款方在第一期的期望效用依然为W_H。因此，高质量借款方没有激励改变他的选择。

对于低质量借款方，如果他选择d_L=1，信贷机构会基于数据信息更新其分布情况，并提供对应的契约（如定理 2）。当 $X<\frac{1-p}{p}$时， 借款方获得W_L。当 $X>\frac{p}{1-p}$，借款方获得$\frac{\theta_L}{\theta_H}W_H$。当$\frac{1-p}{p} \leq X \leq \frac{p}{1-p}$，借款方的期望效用为$(1-p)\frac{\theta_L}{\theta_H}W_H+pW_L$。如果他选择$d_L$=0， 那么信贷机构会基于均衡外的贝叶斯预期，即原始的预期，作出最优契约设计。那么，当 X<1 时， 借款方获得W_L。当 $X \geq 1$，借款方获得$\frac{\theta_L}{\theta_H}W_H$。因此，当 X<1 时，低质量借款方没有激励改变策略。但是，当 $X \geq 1$，低质量借款方会选择d_L=0。综上，当 X<1 时，（d_H=1，d_L=1)是均衡策略。Q.E.D.

附录X 对命题 7 的证明

①当$\frac{1-p}{p}\leq X<1$时，如果借款方选择(d_H=1，d_L=1)，信贷机构会按照借款方提供的数据信息更新借款方类型的概率分布预期并制定最优的契约，如命题 6 的证明，可以求解得到：高质量借款方获得期望效用W_H；低质量借款方的期望效用是$(1-p)\frac{\theta_L}{\theta_H}W_H+pW_L$，因为面对低质量借款方，信贷机构会以概率$(1-p)$观测到$\eta_H$和概率$p$观测到$\eta_L$。如果借款方选择($d_H$=0，$d_L$=0)，信贷机构会按照原始预期制定最优契约。如命题 6 的证明，可以求解得到：高质量借款方获得期望效用W_H；低质量借款方的期望效用是W_L。因此通过对比，得到借款方会选择(d_H=1，d_L=1)，因为这一均衡策略会给低质量借款方带来更高的效用，而不损害高质量借款方的效用。

②当$1\leq X\leq\frac{p}{1-p}$时，如果借款方选择(d_H=1，d_L=1)，信贷机构会按照数据信息更新借款方的类型分布的预期并制定最优的契约，如命题 2 的证明，可以求解得到：高质量借款方获得期望效用W_H；低质量借款方的期望效用是$(1-p)\frac{\theta_L}{\theta_H}W_H+pW_L$，因为此时信贷机构会以概率$(1-p)$观测到$\eta_H$和概率$p$观测到$\eta_L$。如果借款方选择($d_H$=0，$d_L$=0)，信贷机构会按照原始预期制定最优契约，如命题 1 的证明，可以求解得到：高质量借款方获得期望效用W_H；低质量借款方的期望效用是$\frac{\theta_L}{\theta_H}W_H$。因此通过对比，得到借款方会选择($d_H$=0，$d_L$=0)，因为这一均衡策略会给低质量借款方带来更高的效用，而不损害高质量借款方的效用。Q.E.D.

附录XI 不同数据所有权下的社会福利分析

当信贷机构拥有数据时，如果 $X>\frac{p}{1-p}$，信贷机构提供的信贷契约与数据信息无关。根据表 3，我们可以得到高质量借款方的期望效用为 W_H,低质量借款方的期望效用为$\frac{\theta_L}{\theta_H}W_H$，因为两种借款方的概率各为$\frac{1}{2}$。因此，借款方的平均期望效用为$\frac{1}{2}W_H+\frac{1}{2}\frac{\theta_L}{\theta_H}W_H$。此时，信贷机构面临两种情况：概率$\frac{1}{2}$会遇到高质量借款方获得效用 $\theta_H R\text{-}W_H\text{-}I$，概率 $\frac{1}{2}$ 会遇到低质量借款方获得效用 $\theta_L R\text{-}\frac{\theta_L}{\theta_H}W_H\text{-}I$。因此，其期望效用为 $\frac{1}{2}(\theta_H R\text{-}W_H\text{-}I)+\frac{1}{2}\left(\theta_L R\text{-}\frac{\theta_L}{\theta_H}W_H\text{-}I\right)$。社会福利为借款方和信贷机构期望效用之和，即 $\frac{1}{2}(\theta_H R\text{-}W_H\text{-}I)+\frac{1}{2}\left(\theta_L R\text{-}\frac{\theta_L}{\theta_H}W_H\text{-}I\right)+\frac{1}{2}W_H+\frac{1}{2}\frac{\theta_L}{\theta_H}W_H=\frac{1}{2}(\theta_H R\text{-}I)+\frac{1}{2}(\theta_L R\text{-}I)$。

如果$\frac{1-p}{p}\leq X\leq\frac{p}{1-p}$，根据表 3，我们可以得到高质量借款方的期望效用为 W_H。低质量借款方的期望效用为 $pW_L+(1\text{-}p)\frac{\theta_L}{\theta_H}W_H$，因为面临低质量借款方，信贷机构获得信号 η_L的概率为 p，获得信号 η_H的概率为 1-p。此时，借款方的平均期望效用为$\frac{1}{2}W_H+\frac{1}{2}\left(pW_L+(1\text{-}p)\frac{\theta_L}{\theta_H}W_H\right)=\frac{1}{2}W_H+\frac{1}{2}\frac{\theta_L}{\theta_H}W_H\text{-}\frac{1}{2}p\left(\frac{\theta_L}{\theta_H}W_H\text{-}W_L\right)$。此外，如命题 5 的证明所示，信贷机构的期望效用为$\frac{1}{2}\left(\theta_L R\text{-}\frac{\theta_L}{\theta_H}W_H\text{-}I\right)+\frac{1}{2}p\left(\theta_H R\text{-}W_H\text{-}I+\frac{\theta_L}{\theta_H}W_H\text{-}W_L\right)$。社会福利为借款方和信贷机构期望效用之和，求解得到$\frac{1}{2}(\theta_L R\text{-}I)+\frac{1}{2}W_H+\frac{1}{2}p(\theta_H R\text{-}W_H\text{-}I)$。

如果 $X<\frac{1-p}{p}$，根据表 3，我们可以得到高质量借款方的期望效用为 W_H,低质量借款方的期望效用为 W_L。因此，借款方的平均期望效用为$\frac{1}{2}W_H+\frac{1}{2}W_L$。此时，信贷机构只与低质量借款方有信贷契约，因此，其期望效用为$\frac{1}{2}(\theta_L R\text{-}W_L\text{-}I)$。社会福利为$\frac{1}{2}W_H+\frac{1}{2}(\theta_L R\text{-}I)$。因此，我们得到表 6 结果。

当借款方拥有数据所有权时，当 $X<1$ 时，如果$\frac{1-p}{p}\leq X<1$ 时，借款方选择($d_H=1$，$d_L=1$)均衡策略。因此，信贷机构会获得借款方的数据并据此提供信贷契约。这与信贷机构拥有数据信息情况一致。因此，借贷双方期望效用和社会福利在两种不同的数据所有权下一致。当 $X<\frac{1-p}{p}$时，信贷机构是否拥有数据都会提供一样的信贷契约。因此，借贷双方期望效用和社会福利在两种不同的数据所有权下也会一致。

当 $X\geq 1$ 时，如果$\frac{1}{2}<p<\frac{X}{1+X}$，数据信息不会影响信贷契约，借贷双方期望效用和社会福利在两种不同的数据所有权下一致。如果 $p>\frac{X}{1+X}$时，借款方拥有数据信息时会选择不披露信息，这时信贷机构提供的契约与没有数据信息时一致。因此，$X\geq 1$ 时，借款方的期望效用为$\frac{1}{2}W_H+\frac{1}{2}\frac{\theta_L}{\theta_H}W_H$，信贷机构的期望效用为$\frac{1}{2}(\theta_H R-W_H-I)+\frac{1}{2}(\theta_L R-\frac{\theta_L}{\theta_H}W_H-I)$，社会福利为$\frac{1}{2}(\theta_H R-I)+\frac{1}{2}(\theta_L R-I)$。我们得到表 7 结果。Q.E.D.

附录 XII 模型拓展分析一：考虑大数据能够降低信贷机构信息收集成本的情况

在基本模型设定的博弈时序中加入信贷机构数据信息收集的环节，具体如下：

第 0 期，自然决定借款方的类型。第 1 期，信贷机构支付一定的成本收集数据，从而获得数据信息精度 p。假设数据信息收集成本函数 $c(p)=\frac{1}{2}A(p-\frac{1}{2})^2$，其中$\frac{1}{2}\leq p\leq 1$，$A>0$。$c(p)$是 p 的增函数。这意味着，信贷机构搜索更多数据信息以提高信息精确度需要付出更大的成本。大数据的作用体现为参数 A 的下降，即大数据可以降低数据信息收集成本。第 2 期，信贷机构通过大数据分析获得与借款方类型相关的信息，并基于该信息决定信贷契约。第 3 期，借款方根据自身类型从契约菜单中选择合适的契约；第 4 期，借款方按照契约约定，用项目收入进行还款。第 2 期及之后的信贷机构和借款方的最优决策，我们在第四节已经分析。因此，这里聚焦第 1 期的信贷机构的数据收集选择 p。

当 $X\geq 1$ 时，信贷机构选择数据信息精度 p 最大化其效用。根据表 6 中第 2 期信贷机构的期望效用，我们得到第 1 期时信贷机构效用函数如下：

$$V(p)=\begin{cases}\frac{1}{2}(\theta_H R-W_H-I)+\frac{1}{2}\left(\theta_L R-\frac{\theta_L}{\theta_H}W_H-I\right)-\frac{1}{2}A\left(p-\frac{1}{2}\right)^2, & p<\frac{X}{1+X}\\ \frac{1}{2}\left(\theta_L R-\frac{\theta_L}{\theta_H}W_H-I\right)+\frac{1}{2}p\left(\theta_H R-W_H-I+\frac{\theta_L}{\theta_H}W_H-W_L\right)-\frac{1}{2}A\left(p-\frac{1}{2}\right)^2, & p\geq\frac{X}{1+X}\end{cases}; \quad \text{(XII 1)}$$

类似的，当 $X<1$ 时，第 1 期信贷机构效用函数如下：

$$V(p)=\begin{cases}\frac{1}{2}(\theta_L R-W_L-I)-\frac{1}{2}A\left(p-\frac{1}{2}\right)^2, & p<\frac{1}{1+X}\\ \frac{1}{2}\left(\theta_L R-\frac{\theta_L}{\theta_H}W_H-I\right)+\frac{1}{2}p\left(\theta_H R-W_H-I+\frac{\theta_L}{\theta_H}W_H-W_L\right)-\frac{1}{2}A\left(p-\frac{1}{2}\right)^2, & p\geq\frac{1}{1+X}\end{cases}; \quad \text{(XII 2)}$$

信贷机构效用函数是 p 的函数。通过最优化求解，我们得到以下结论：

命题 9：伴随着数据收集成本参数 A 的下降，信贷机构有激励收集更多数据来提高数据信息精度。

证明：

为了后续方便计算，我们定义 $m=\theta_H R-W_H-I$，$n=\theta_L R-W_L-I$，$z=\theta_L R-\frac{\theta_L}{\theta_H}W_H-I$，那么 $X=\frac{m}{n-z}$，$\frac{X}{1+X}=\frac{m}{m+n-z}$，$\frac{1}{1+X}=\frac{n-z}{m+n-z}$。

① 当 $X\geq 1$ 时，根据公式（XII 1），如果 $p<\frac{m}{m+n-z}$，$V(p)$随着 p 的增加而下降，因此，$p=\frac{1}{2}$时，信贷机构的效用最大。如果 $p\geq\frac{m}{m+n-z}$，我们通过求导得到一阶条件$\frac{1}{2}(m+n-z)=A\left(p-\frac{1}{2}\right)=0$，因此 $p=\frac{1}{2}+\frac{m+n-z}{2A}$。下面我们分两种情况讨论：情况一 $m>3(n-z)$；情况二 $m\leq 3(n-z)$。

情况一：$m>3(n-z)$

当 $A\leq m+n-z$ 时，$\frac{1}{2}+\frac{m+n-z}{2A}\geq 1$，因此，信贷机构的效用在 $p=\frac{1}{2}$或 $p=1$ 时取得最大值。我们比较 $V\left(\frac{1}{2}\right)$与 $V(1)$的大小。$V\left(\frac{1}{2}\right)=\frac{m+z}{2}$，$V(1)=\frac{m+n}{2}-\frac{A}{8}$。我们可以得到当 $A\leq 4(n-z)$时，$V(1)\geq V\left(\frac{1}{2}\right)$，因此信贷机构选

择 p=1，对应的期望效用为 $V(1)=\frac{m+n}{2}-\frac{A}{8}$。当 $4(n-z)<A\leq m+n-z$ 时，$V(1)<V\left(\frac{1}{2}\right)$，因此，信贷机构选择 $p=\frac{1}{2}$，对应的期望效用为 $V\left(\frac{1}{2}\right)=\frac{m+z}{2}$。

当 $A>m+n-z$ 时，$\frac{1}{2}+\frac{m+n-z}{2A}<1$。因此，信贷机构效用在 $p=\frac{1}{2}$或 $p=\frac{1}{2}+\frac{m+n-z}{2A}$时取得最大值。我们比较 $V\left(\frac{1}{2}\right)$与 $V\left(\frac{1}{2}+\frac{m+n-z}{2A}\right)$的大小。$V\left(\frac{1}{2}\right)=\frac{m+z}{2}$，$V\left(\frac{1}{2}+\frac{m+n-z}{2A}\right)=\frac{(m+n-z)^2+2A(m+n+z)}{8A}$。我们可以计算得到，当 $m>3(n-z)$ 且 $A>m+n-z$ 时，我们可以得到 $V\left(\frac{1}{2}+\frac{m+n-z}{2A}\right)<V\left(\frac{1}{2}\right)$。因此，信贷机构选择 $p=\frac{1}{2}$，对应的期望效用为 $V\left(\frac{1}{2}\right)=\frac{m+z}{2}$。

因此，在 $m>3(n-z)$情况下，当 $A\leq4(n-z)$时，信贷机构选择 p=1；当 $A>4(n-z)$时，信贷机构选择 $p=\frac{1}{2}$。

情况二：$m\leq3(n-z)$

当 $A\leq m+n-z$ 时，$\frac{1}{2}+\frac{m+n-z}{2A}\geq1$，因此，信贷机构的效用在 $p=\frac{1}{2}$或 $p=1$ 时取得最大值。我们比较 $V\left(\frac{1}{2}\right)$与 $V(1)$的大小。$V\left(\frac{1}{2}\right)=\frac{m+z}{2}$，$V(1)=\frac{m+n}{2}-\frac{A}{8}$。我们可以得到当 $A\leq m+n-z$ 且 $m\leq3(n-z)$时，$V(1)\geq V\left(\frac{1}{2}\right)$，因此信贷机构选择 p=1，对应的期望效用为 $V(1)=\frac{m+n}{2}-\frac{A}{8}$。

当 $A>m+n-z$ 时，$\frac{1}{2}+\frac{m+n-z}{2A}<1$。因此，信贷机构效用在 $p=\frac{1}{2}$或 $p=\frac{1}{2}+\frac{m+n-z}{2A}$时取得最大值。我们比较 $V\left(\frac{1}{2}\right)$与 $V\left(\frac{1}{2}+\frac{m+n-z}{2A}\right)$的大小。$V\left(\frac{1}{2}\right)=\frac{m+z}{2}$，$V\left(\frac{1}{2}+\frac{m+n-z}{2A}\right)=\frac{(m+n-z)^2+2A(m+n+z)}{8A}$。通过计算得到，当 $m+n-z<A\leq\frac{(m+n-z)^2}{2(m-n+z)}$时，我们可以得到 $V\left(\frac{1}{2}+\frac{m+n-z}{2A}\right)\geq V\left(\frac{1}{2}\right)$。因此，信贷机构选择 $p=\frac{1}{2}+\frac{m+n-z}{2A}$，对应的期望效用为 $V\left(\frac{1}{2}+\frac{m+n-z}{2A}\right)=\frac{(m+n-z)^2+2A(m+n+z)}{8A}$。而当 $A>\frac{(m+n-z)^2}{2(m-n+z)}$时，我们可以得到 $V\left(\frac{1}{2}+\frac{m+n-z}{2A}\right)<V\left(\frac{1}{2}\right)$。因此，信贷机构选择 $p=\frac{1}{2}$，对应的期望效用为 $V\left(\frac{1}{2}\right)=\frac{m+z}{2}$。

因此，在 $m\leq3(n-z)$情况下，当 $A\leq m+n-z$ 时，信贷机构选择 p=1；当 $m+n-z<A\leq\frac{(m+n-z)^2}{2(m-n+z)}$时，信贷机构选择 $p=\frac{1}{2}+\frac{m+n-z}{2A}$；当 $A>\frac{(m+n-z)^2}{2(m-n+z)}$时，信贷机构选择 $p=\frac{1}{2}$。

② 当 $X<1$ 时，根据公式（XII 2），如果 $p<\frac{n-z}{m+n-z}$，$V(p)$随着 p 的增加而下降，因此，$p=\frac{1}{2}$时，信贷机构的效用最大。如果 $p\geq\frac{n-z}{m+n-z}$，我们通过求导得到一阶条件$\frac{1}{2}(m+n-z)=A\left(p-\frac{1}{2}\right)=0$，因此 $p=\frac{1}{2}+\frac{m+n-z}{2A}$。下面我们分两种情况讨论：情况一 $m>\frac{1}{3}(n-z)$；情况二 $m\leq\frac{1}{3}(n-z)$。

情况一：$m>\frac{1}{3}(n\text{-}z)$

当 $A\leq m+n\text{-}z$ 时，$\frac{1}{2}+\frac{m+n\text{-}z}{2A}\geq 1$，因此，信贷机构的效用在 $p=\frac{1}{2}$或 $p=1$ 时取得最大值。我们比较 $V\left(\frac{1}{2}\right)$与 $V(1)$的大小。$V\left(\frac{1}{2}\right)=\frac{n}{2}$，$V(1)=\frac{m+n}{2}\text{-}\frac{A}{8}$。我们可以得到：当 $A\leq m+n\text{-}z$ 时，$V(1)\geq V\left(\frac{1}{2}\right)$，因此信贷机构选择 p=1，对应的期望效用为 $V(1)=\frac{m+n}{2}\text{-}\frac{A}{8}$。

当 $A>m+n\text{-}z$ 时，$\frac{1}{2}+\frac{m+n\text{-}z}{2A}<1$。因此，信贷机构效用在 $p=\frac{1}{2}$或 $p=\frac{1}{2}+\frac{m+n\text{-}z}{2A}$时取得最大值。我们比较 $V\left(\frac{1}{2}\right)$与 $V\left(\frac{1}{2}+\frac{m+n\text{-}z}{2A}\right)$的大小。$V\left(\frac{1}{2}\right)=\frac{n}{2}$，$V\left(\frac{1}{2}+\frac{m+n\text{-}z}{2A}\right)=\frac{(m+n\text{-}z)^2+2A(m+n+z)}{8A}$。我们可以计算得到：当 $m+n\text{-}z<m\leq\frac{(m+n\text{-}z)^2}{2(n\text{-}m\text{-}z)}$时，我们可以得到 $V\left(\frac{1}{2}+\frac{m+n\text{-}z}{2A}\right)\geq V\left(\frac{1}{2}\right)$。因此，信贷机构选择 $p=\frac{1}{2}+\frac{m+n\text{-}z}{2A}$，对应的期望效用为 $V\left(\frac{1}{2}+\frac{m+n\text{-}z}{2A}\right)=\frac{(m+n\text{-}z)^2+2A(m+n+z)}{8A}$。当 $m>\frac{(m+n\text{-}z)^2}{2(n\text{-}m\text{-}z)}$时，我们可以得到 $V\left(\frac{1}{2}+\frac{m+n\text{-}z}{2A}\right)<V\left(\frac{1}{2}\right)$。因此，信贷机构选择 $p=\frac{1}{2}$，对应的期望效用为 $V\left(\frac{1}{2}\right)=\frac{n}{2}$。

因此，在 $m>\frac{1}{3}(n\text{-}z)$情况下，当 $A\leq m+n\text{-}z$ 时，信贷机构选择 p=1；当 $m+n\text{-}z<A\leq\frac{(m+n\text{-}z)^2}{2(n\text{-}m\text{-}z)}$时，信贷机构选择 $p=\frac{1}{2}+\frac{m+n\text{-}z}{2A}$；当 $A>\frac{(m+n\text{-}z)^2}{2(n\text{-}m\text{-}z)}$时，信贷机构选择 $p=\frac{1}{2}$。

情况二：$m\leq\frac{1}{3}(n\text{-}z)$

当 $A\leq m+n\text{-}z$ 时，$\frac{1}{2}+\frac{m+n\text{-}z}{2A}\geq 1$，因此，信贷机构的效用在 $p=\frac{1}{2}$或 $p=1$ 时取得最大值。我们比较 $V\left(\frac{1}{2}\right)$与 $V(1)$的大小。$V\left(\frac{1}{2}\right)=\frac{n}{2}$，$V(1)=\frac{m+n}{2}\text{-}\frac{A}{8}$。我们可以得到：当 $A\leq 4m$ 时，$V(1)\geq V\left(\frac{1}{2}\right)$，因此信贷机构选择 p=1，对应的期望效用为 $V(1)=\frac{m+n}{2}\text{-}\frac{A}{8}$。当 $4m<A\leq m+n\text{-}z$ 时，$V(1)<V\left(\frac{1}{2}\right)$，因此，信贷机构选择 $p=\frac{1}{2}$，对应的期望效用为 $V\left(\frac{1}{2}\right)=\frac{n}{2}$。

当 $A>m+n\text{-}z$ 时，$\frac{1}{2}+\frac{m+n\text{-}z}{2A}<1$。因此，信贷机构效用在 $p=\frac{1}{2}$或 $p=\frac{1}{2}+\frac{m+n\text{-}z}{2A}$时取得最大值。我们比较 $V\left(\frac{1}{2}\right)$与 $V\left(\frac{1}{2}+\frac{m+n\text{-}z}{2A}\right)$的大小。$V\left(\frac{1}{2}\right)=\frac{m+z}{2}$，$V\left(\frac{1}{2}+\frac{m+n\text{-}z}{2A}\right)=\frac{(m+n\text{-}z)^2+2A(m+n+z)}{8A}$。通过计算得到：当 $A>m+n\text{-}z$ 时，我们可以得到 $V\left(\frac{1}{2}+\frac{m+n\text{-}z}{2A}\right)<V\left(\frac{1}{2}\right)$。因此，信贷机构选择 $p=\frac{1}{2}$，对应的期望效用为 $V\left(\frac{1}{2}\right)=\frac{n}{2}$。

因此，在 $m\leq\frac{1}{3}(n\text{-}z)$情况下，当 $A\leq4m$ 时，信贷机构选择 p=1；当 $A>4m$ 时，信贷机构选择 $p=\frac{1}{2}$。Q.E.D.

命题 9 告诉我们，当大数据极大地降低数据信息收集成本时，信贷机构有激励过度收集数据。第四节分析告诉我们，数据信息精度 p 的上升可能损害借款方利益，但是这对信贷机构却是有利的。因此，命题 9 也解释了现实中机构过度收集数据的现象。在中国，移动互联网金融业务曾经历一个无序攫取用户数据的“野蛮生长期”，众多网络借贷平台利用爬虫和 SDK 插件获取借款方通讯录、定位、消费等敏感信息，超出业务必要范围。低数据成本驱动下，这种过度收集十分普遍。不少贷款类或信用评分类 APP 在申请权限时“一揽子”索取通讯录、相册、定位等无关信息，不给权限就无法使用，严重背离必要性原则。比如，“你我贷”和“极融”曾被媒体曝光存在个人信息过度攫取的问题：用户在申请贷款时表面仅授权了 5 份主协议，但实际却触发了超过 700 份第三方授权文件，涉及近 30 家金融机构和 20 余个助贷平台，其中部分第三方协议还嵌套了大量子协议（仅榕树贷款一家就高达 171 份）。这也印证了命题 9 的结论，当数据收集技术变得廉价易得，信贷机构有动机过度收集数据，因此需要通过法律和行业规范加以约束。

附录 XIII 模型拓展分析二：考虑借款方质量为连续分布的情景

这里考虑借款方质量为连续分布的情景。借款方类型$(\theta, W(\theta))$是私人信息。假设$\theta\in[0,1]$，其分布密度函数为 $f(\bullet)$，分布函数为 $F(\bullet)$。信贷机构给借款方提供信贷契约菜单（$L(\theta), Y(\theta)$），其中 $L(\theta)$为贷款额度，$Y(\theta)$为借款方项目成功时的还款额。为了简化分析，设定 $L(\theta)=I$（经检验不影响模型的一般性，与离散时的结论一致）。该设定代表信贷机构如果贷款给借款方开展项目，那么其贷款额度刚好等于项目所需投资额。此外，设定$\frac{\partial(\frac{W(\theta)}{\theta})}{\partial\theta}>0$，代表没有信息不对称情况下，信贷机构从风险高的投资中会获得更高的回报（与离散情况的假设类似）。

对于任意两种借款方$\theta,\tilde{\theta}$，信贷机构提供的契约需满足借款方的激励相容条件：

$$\theta(R-Y(\theta))-W(\theta)\geq\theta(R-Y(\tilde{\theta}))-W(\theta), \tag{XIII 1}$$

通过上面激励相容条件，我们得到 $Y(\theta)=Y(\tilde{\theta})=Y$。

借款方的参与条件如下：

$$U(\theta)=\theta(R-Y)-W(\theta)\geq 0, \tag{XIII 2}$$

存在一个$\theta^{*}\in[0,1]$使得 $R-Y=\frac{W(\theta^{*})}{\theta^{*}}$。对于$\theta\in[0,\theta^{*}]$，$\theta(R-Y)-W(\theta)\geq 0$，借款方会选择接受贷款开展项目；而对于$\theta\in(\theta^{*},1]$，$\theta(R-Y)-W(\theta)<0$，借款方选择不参与。信贷机构在满足激励相容和参与条件下的设计最优契约以最大化自己的效用：

$$max\int_0^{\theta^{*}}(\theta R-I-W(\theta)-U(\theta))f(\theta)\,d\theta。 \tag{XIII 3}$$

下面在此连续类型分布的框架下考察大数据的影响。为了简化求解，假设θ在$[0,1]$上均匀分布。信贷机构拥有借款方数据，通过分析获得下面的信号：$Pr(S=H|\theta)=\theta$和$Pr(S=L|\theta)=1-\theta$。也就是说，对于类型为θ的借款方，信贷机构以θ的概率获得一个高质量信号，以 $1-\theta$的概率获得一个低质量信号。因此，借款方项目成功的概率越大，信贷机构获得高质量信号的概率也同样变大。信贷机构基于不同的信号对借款方类型分布进行贝叶斯更新，并设计信贷契约最大化自己的效用。效用函数如下所示：

$$max\int_0^{\theta^{*}}(\theta R-I-W(\theta)-U(\theta))f(\theta|S)\,d\theta, \tag{XIII 4}$$

通过最优化求解,我们发现在借款方类型是连续分布的情况下，大数据不一定能够提高借款方利益，促进普惠金融。相反，大数据甚至会导致更少的借款方能被纳入信贷体系获得贷款。

证明：通过借款方的效用函数，我们得到 $U(\theta^{*})=\theta^{*}(R-Y)-W(\theta^{*})=0$。我们得到 $Y=R-\frac{W(\theta^{*})}{\theta^{*}}$，因此

$U(\theta)=\theta(R-Y)-W(\theta)=\theta\left(R-R+\frac{W(\theta^{*})}{\theta^{*}}\right)-W(\theta)=\theta\frac{W(\theta^{*})}{\theta^{*}}-W(\theta)$，将其代入方程（9），我们得到：

$$max\int_0^{\theta^{*}}\left(\theta R-I-\theta\frac{W(\theta^{*})}{\theta^{*}}\right)f(\theta)\,d\theta, \tag{XIII 5}$$

对上面的方程求一阶条件，我们得到：

$$\left(\theta^{*}R-I-W(\theta^{*})\right)f(\theta^{*})+\left(\frac{W(\theta^{*})}{\theta^{*2}}-\frac{W'(\theta^{*})}{\theta^{*}}\right)\int_0^{\theta^{*}}\theta f(\theta)\,d\theta=0\,, \tag{XIII 6}$$

我们先分析信贷机构没有数据的情况，因为θ属于均匀分布，该一阶条件为

$$\left(\theta^{*}R-I-W(\theta^{*})\right)+\left(\frac{W(\theta^{*})}{\theta^{*2}}-\frac{W'(\theta^{*})}{\theta^{*}}\right)\frac{\theta^{*2}}{2}=0,$$

由上式求解出最优的θ^*。在信贷机构没有数据的情况，$\theta\leq\theta^*$的借款方获得贷款，而$\theta>\theta^*$借款方选择不参与。获得贷款的借款方的还款额为$R-\frac{W(\theta^*)}{\theta^*}$。

下面考察信贷机构有数据的情况。当信贷机构观察到信号为高质量时，后验密度函数为$f(\theta|s=H)=2\theta$。此时，其优化的一阶条件为：

$$\left(\theta_H^*R-I-W\left(\theta_H^*\right)\right)2\theta_H^*+\left(\frac{W(\theta_H^*)}{\theta_H^{*2}}-\frac{W'(\theta_H^*)}{\theta_H^*}\right)\frac{2}{3}\theta_H^{*3}=2\theta_H^*\left[\left(\theta_H^*R-I-W\left(\theta_H^*\right)\right)+\left(\frac{W(\theta_H^*)}{\theta_H^{*2}}-\frac{W'(\theta_H^*)}{\theta_H^*}\right)\frac{1}{3}\theta_H^{*2}\right]=0, \tag{XIII 7}$$

因此，$\left(\theta_H^*R-I-W\left(\theta_H^*\right)\right)+\left(\frac{W(\theta_H^*)}{\theta_H^{*2}}-\frac{W'(\theta_H^*)}{\theta_H^*}\right)\frac{1}{3}\theta_H^{*2}=0$。我们得到，$\theta\leq\theta_H^*$的借款方获得贷款，而$\theta>\theta_H^*$借款方没有获得贷款，获得贷款的借款方的还款额为$R-\frac{W(\theta_H^*)}{\theta_H^*}$。整理可得：

$$\left(\theta_H^*R-I-W\left(\theta_H^*\right)\right)+\left(\frac{W(\theta_H^*)}{\theta_H^{*2}}-\frac{W'(\theta_H^*)}{\theta_H^*}\right)\frac{1}{3}\theta_H^{*2}=\left(\theta_H^*R-I-W\left(\theta_H^*\right)\right)+\left(\frac{W(\theta_H^*)}{\theta_H^{*2}}-\frac{W'(\theta_H^*)}{\theta_H^*}\right)\frac{1}{2}\theta_H^{*2}-\left(\frac{W(\theta_H^*)}{\theta_H^{*2}}-\frac{W'(\theta_H^*)}{\theta_H^*}\right)\frac{1}{6}\theta_H^{*2}=0, \tag{XIII 8}$$

因为$\frac{\partial(\frac{W(\theta)}{\theta})}{\partial\theta}>0$，所以$-\left(\frac{W(\theta_H^*)}{\theta_H^{*2}}-\frac{W'(\theta_H^*)}{\theta_H^*}\right)\frac{1}{6}\theta_H^{*2}>0$。那么$\left(\theta_H^*R-I-W\left(\theta_H^*\right)\right)+\left(\frac{W(\theta_H^*)}{\theta_H^{*2}}-\frac{W'(\theta_H^*)}{\theta_H^*}\right)\frac{1}{2}\theta_H^{*2}<0$。而$\left(\theta^*R-I-W\left(\theta^*\right)\right)+\left(\frac{W(\theta^*)}{\theta^{*2}}-\frac{W'(\theta^*)}{\theta^*}\right)\frac{\theta^{*2}}{2}=0$。因为方程（XIII 5）的二阶导小于 0，我们得到$\theta_H^*>\theta^*$。也就是说，当信贷机构获得高质量信号时，更多借款方可以获得贷款。此外，因为$R-\frac{W(\theta_H^*)}{\theta_H^*}<R-\frac{W(\theta^*)}{\theta^*}$，借款方需要的还款额度也更低。

当信贷机构观察到信号为低质量时，后验密度函数为$f(\theta|s=L)=2(1-\theta)$。此时，其优化的一阶条件为：

$$\left(\theta_L^*R-I-W\left(\theta_L^*\right)\right)2(1-\theta_L^*)+\left(\frac{W(\theta_L^*)}{\theta_L^{*2}}-\frac{W'(\theta_L^*)}{\theta_L^*}\right)\left(\theta_L^{*2}-\frac{2}{3}\theta_L^{*3}\right)=2(1-\theta_L^*)\left[\left(\theta_L^*R-I-W\left(\theta_L^*\right)\right)+\left(\frac{W(\theta_L^*)}{\theta_L^{*2}}-\frac{W'(\theta_L^*)}{\theta_L^*}\right)\frac{\theta_L^{*2}(1-\frac{2}{3}\theta_L^*)}{2(1-\theta_L^*)}\right]=0, \tag{XIII 9}$$

因此，$\left(\theta_L^*R-I-W\left(\theta_L^*\right)\right)+\left(\frac{W(\theta_L^*)}{\theta_L^{*2}}-\frac{W'(\theta_L^*)}{\theta_L^*}\right)\frac{\theta_L^{*2}(1-\frac{2}{3}\theta_L^*)}{2(1-\theta_L^*)}=0$。我们得到，$\theta\leq\theta_L^*$的借款方获得贷款，而$\theta>\theta_L^*$借款方没有获得贷款，获得贷款的借款方的还款额为$R-\frac{W(\theta_L^*)}{\theta_L^*}$。类似的，我们可以证明$\theta_L^*<\theta^*$。因此，当信贷机构获得低质量信号时，更少的借款方可以获得贷款。此外，因为$R-\frac{W(\theta_L^*)}{\theta_L^*}>R-\frac{W(\theta^*)}{\theta^*}$，借款方需要的还款额度也更高。也就是说，当信贷机构拥有数据信息时，不一定会提高借款方获取贷款概率或降低借款方融资成本。Q.E.D.

上述分析表明本文的主要结论依然成立。这一拓展也进一步检验和强化了本文核心结论的稳健性，展示了数据的“双刃剑”作用，为现实中金融科技公司在小微企业融资业务中所出现的“普而不惠”及“既不普又不惠”等现象提供了理论支撑。

注：该附录是期刊所发表论文的组成部分，同样视为作者公开发表的内容。如研究中使用该附录中的内容，请务必在研究成果上注明附录下载出处。