This appendix contains the proof of the following proposition stated in the paper.

**Proposition 1** There exist securitization thresholds \( \alpha, \bar{\alpha} \in (0, 1) \), with \( \alpha < \bar{\alpha} \), such that in equilibrium (i) a lender acquires soft information only if \( \alpha \leq \alpha \) and the hard information signal is \( x_L \), and (ii) a lender does not acquire soft information only if \( \alpha \geq \bar{\alpha} \).

**Proof:**

Observe that, in equilibrium, all loans are offered at the interest rate \( r_h \), and all offers are accepted by the borrower, regardless of type. Thus, soft information is valuable only if lender screens out (i.e. does not lend to) borrowers with signal \( y_L \).

Now, suppose the hard information signal is \( x_L \) and investors believe the lender screens such loans based on soft information. That is, investors believe the lender collects information and lends only to borrowers with soft information signal \( y_H \). Note that if it lends to a borrower with soft information signal \( y_L \), the expected payoff on the loan is \( v_L + \lambda_L (v_H - v_L) \), which is less than zero, given assumption (i). Given the postulated lending strategy, all loans with hard information signal \( x_L \) must have generated the soft information signal \( y_H \). Therefore, \( P(x_L) = 1 + v_L + \lambda_L (v_H - v_L) \). Hence, if the lender sells the loan, it obtains a payoff \( P(x_L) - 1 - c = v_L + \lambda_L (v_H - v_L) + c \). If it retains the loan, its payoff depends on its actual screening and lending strategy. Suppose that, as investors expect, the lender acquires soft information and does not make loans to borrowers with soft information signal \( y_L \). Then, its payoff from a retained loan is again \( v_L + \lambda_L (v_H - v_L) - c \).

The probability that a borrower with hard information signal \( x_L \) generates a soft information signal \( y_H \) is \( \mu_L (1 - \mu) \gamma_H \). Note that \( \lambda_L = \frac{\mu_L \gamma_H}{\mu_L (1 - \mu) \gamma_L} \), so the probability of obtaining soft information signal \( y_H \) given hard information signal \( x_L \) may also be written as \( \frac{\mu_L \gamma_H}{\lambda_L} \). Since borrowers with soft information signal \( y_L \) are screened out, the lender’s expected payoff when it screens loans on soft information is \( \frac{\mu_L \gamma_H}{\lambda_L} [v_L + \lambda_L (v_H - v_L)] - c \).

Now, suppose the lender deviates and offers a loan to all borrowers with signal \( x_L \). Since soft information is not being used in any way, it is optimal to not acquire it and save on the cost \( c \). If the lender sells the loan, it again obtains a payoff \( P(x_L) - 1 \). If it retains the loan, its payoff is \( v_L + \mu_L (v_H - v_L) \). The payoffs following the deviation are earned with probability 1. Hence, its expected payoff from the loan is

\[
(1 - \alpha) [v_L + \mu_L (v_H - v_L)] + \alpha [v_L + \lambda_L (v_H - v_L)] = v_L + \mu_L (v_H - v_L) + \alpha (\lambda_L - \mu_L) (v_H - v_L).
\]
It is a best response for the lender to screen on soft information if and only if
\[
\frac{\mu \ell \gamma_h}{\lambda_{th}} (v_\ell + \lambda_{th} (v_h - v_\ell)) - c \geq v_\ell + \mu_\ell (v_h - v_\ell) + \alpha (\lambda_{th} - \mu_\ell)(v_h - v_\ell),
\]
(2)

or
\[
\alpha \leq \frac{-(1 - \mu \ell \gamma_h/\lambda_{th})v_\ell - \mu_\ell (1 - \gamma_h)(v_h - v_\ell) - c}{(\lambda_{th} - \mu_\ell)(v_h - v_\ell)}.
\]
(3)

Define \( \alpha = \frac{-(1 - \mu \ell \gamma_h/\lambda_{th})v_\ell - \mu_\ell (1 - \gamma_h)(v_h - v_\ell) - c}{(\lambda_{th} - \mu_\ell)(v_h - v_\ell)} \). Then, in equilibrium the lender acquires soft information only if \( \alpha \leq \alpha \). Assumption (ii) ensures that \( \alpha > 0 \), and it is immediate from the payoff expressions that when \( \alpha = 1 \), the payoff from screening on soft information is strictly lower.

Next, suppose the hard information signal is \( x_\ell \) and investors believe the lender is not screening on soft information. Then, \( P(x_\ell) = 1 + v_\ell + \mu_\ell (v_h - v_\ell) \), so that the payoff to a lender from a sold loan is \( P(x_\ell - 1) = v_\ell + \mu_\ell (v_h - v_\ell) \). If the lender does not acquire soft information, it obtains a payoff \( v_\ell + \mu_\ell (v_h - v_\ell) \) regardless of whether the loan is sold or retained. Suppose the lender deviates, acquires soft information, and lends only to borrowers with soft information signal \( y_h \). Then, its payoff from retaining the loan is \( v_\ell + \mu_\ell (v_h - v_\ell) \). The payoff from selling the loan remains \( v_\ell + \mu_\ell (v_h - v_\ell) \). However, these payoffs are earned only with probability \( \frac{\mu \gamma_h}{\lambda_{th}} \). Also, the lender incurs the screening cost \( c \). Therefore, its overall payoff from the deviation is
\[
\frac{\mu \ell \gamma_h}{\lambda_{th}} [\alpha (v_\ell + \mu_\ell (v_h - v_\ell)) + (1 - \alpha) (v_\ell + \lambda_{th} (v_h - v_\ell))] - c =
\frac{\mu \ell \gamma_h}{\lambda_{th}} [v_\ell + \lambda_{th} (v_h - v_\ell) - \alpha (\lambda_{th} - \mu_\ell)(v_h - v_\ell)] - c.
\]
(4)

Thus, it is optimal for the lender to not screen on soft information only if
\[
v_\ell + \mu_\ell (v_h - v_\ell) \geq \frac{\mu \ell \gamma_h}{\lambda_{th}} [v_\ell + \lambda_{th} (v_h - v_\ell) - \alpha (\lambda_{th} - \mu_\ell)(v_h - v_\ell)] - c,
\]
(5)

or
\[
\alpha \geq \frac{-(1 - \mu \ell \gamma_h/\lambda_{th})v_\ell - \mu_\ell (1 - \gamma_h)(v_h - v_\ell) - c}{(\mu \ell \gamma_h/\lambda_{th})(\lambda_{th} - \mu_\ell)(v_h - v_\ell)}.
\]
(6)

Define \( \bar{\alpha} = \frac{-(1 - \mu \ell \gamma_h/\lambda_{th})v_\ell - \mu_\ell (1 - \gamma_h)(v_h - v_\ell) - c}{(\mu \ell \gamma_h/\lambda_{th})(\lambda_{th} - \mu_\ell)(v_h - v_\ell)} \). Then, in equilibrium, the lender does not acquire soft information only if \( \alpha \geq \bar{\alpha} \). Assumption (ii) ensures that \( \bar{\alpha} > 0 \), and it is immediate that, when \( \alpha = 1 \), the payoff from screening on soft information is strictly lower. Note that \( \bar{\alpha} > \alpha \), since \( \mu \ell \gamma_h < \lambda_{th} \).

To complete the proof, suppose the hard information signal is \( x_h \). Then, \( P(x_h) = 1 + v_\ell + \mu_h (v_h - v_\ell) \). Since \( v_\ell + \lambda_{th} (v_h - v_\ell) > 0 \) by assumption (i), even if the lender were to retain the loan, it would continue to lend to a borrower with soft information signal \( y_\ell \). Hence, it is immediate that the lender will not acquire soft information. ■