Asset mispricing in peer-to-peer loan secondary markets

Asset Mispricing in Peer-to-Peer Loan Secondary Markets

Mustafa Caglayan[a], Tho Pham[b], Oleksandr Talavera[c], Xiong Xiong[d]

Abstract

This study examines the presence of mispricing on Bondora, a leading European peer-to-peer lending platform, over the 2016-2019 period. By implementing machine learning methods, we measure the likelihood of success for loan resale on Bondora's secondary market and compare our predictions with the ex-post market outcomes. The differences observed uncover two phenomena which are related to the diverging perceptions of market participants on asset prices and associated fundamentals: some non-saleable assets are sold, while the resale of highly saleable assets is not successful. Sellers' pricing behaviour changes once they observe buyers' actions revealing the buyers' beliefs about the value of the asset. Our results are robust to various statistical and machine learning methods.

Keywords: mispricing, online secondary market, peer-to-peer lending, belief dispersion

JEL classification: G12, G20

[a] School of Social Sciences, Business School, Heriot-Watt University, Edinburgh EH14 4AS, United Kingdom.

[b] Department of Economics, University of Reading, Reading RG6 6UR, United Kingdom.

[c] Corresponding author. Department of Economics, Birmingham Business School, University of Birmingham, Birmingham B15 2TT, UK. Email: o.talavera@bham.ac.uk.

[d] College of Management and Economics, Tianjin University, China.

1. Introduction

One of the central principles of finance theory is that the price of an asset equals the present value of all future cash flows that the asset generates. However, this principle does not always hold in reality; the perceived fair price often diverges from that dictated by the fundamentals, and the difference can be substantial during periods of uncertainty. This observation has been extensively investigated in the literature, with a particular focus on market frictions and financial constraints as sources of mispricing (see, e.g., Brunnermeier and Pedersen, 2008; Duffie, 2010; Chien et al., 2012). Yet, little is known about the existence of mispricing in a market where transaction costs are marginal and capital constraints are not strictly binding. Questions remain as to whether assets could still be mispriced in such markets and what factors could affect (mis) pricing.

In this study, we seek answers to these questions by examining a detailed dataset obtained from *Bondora.com*, a prominent European peer-to-peer (P2P) lending online platform in Continental Europe (AltFi, 2019). In Bondora's secondary market, sellers can freely list their assets, which are part of the loans that originated from the very same platform's primary market, at a premium (or a discount). The platform provides a wide range of information on these assets for the consumption of buyers, including borrower characteristics, initial loan characteristics, as well as up-to-date loan performance. Of particular note is that, in this market, one can observe a substantial variation in the premium (discount) rates across the secondary market listings, even though they are, de facto, part of the same P2P loan being repaid. Further, on any day, buyers are able to choose from thousands of assets listed in the secondary market. In this context, one would expect an asset to be valued correctly by sellers, while buyers can appraise the same asset at a lower value, thereby making the listing unsuccessful in the secondary market. We define such a divergence in an asset's valuation between the buyers and the sellers as a *type 2 mispricing error*—a listing that is not sold, inspite of the high expectation that it would. It is also possible that a listing with a low sale probability could, in fact, find a buyer in the secondary market. In this case, investors would be committing a *type 1 mispricing error*.

It is a challenge to provide an incontestable definition of mispricing. Typically, finance literature discusses mis-valuation as a divergence of an asset's price from its fundamental value. Most known examples of this approach are stock/bond mispricing, initial public offering (IPO) under-pricing, acquisition overvaluation, tender offer mispricing (e.g., D'mello and Shroff, 2000; Dong et al., 2006; Sadka and Scherbina, 2007; Fleckenstein et al., 2014). The methods that researchers have used to determine mispricing vary: some researchers measure the fundamental value of the asset/firm, based on post-event (e.g., post-takeover) market performance, future cash flows, book value of equity, or residual income value; others use different asset pricing models that measure the "fundamental" price. In other words, the methods implemented to detect mispricing vary dependent upon the study's context and data availability.

In search of evidence of mispricing, in this study, we first estimate the probability of success of an asset's sale, given each transaction. A high predicted sale probability might indicate that the seller's price is close to the fundamental value of the asset (Walkling, 1985). In contrast, a low predicted asset sale probability might suggest over-valuation of an asset's worth by the seller. Next, we compare the predicted sale likelihood of each asset with the actual outcome of the transaction. If an asset, that has a low predicted likelihood of being sold, *is sold*, we interpret this as evidence that the buyer's asset valuation exceeds the seller's valuation. In contrast, when we consider an asset with a high predicted probability of sale, if the buyer's pricing is *less* than that of the seller, the asset would not be sold. In summary, we capture mispricing of an asset by comparing the estimated sale probability of that asset with the actual outcome. These observed differences between the theoretical predictions and actual outcomes would therefore provide evidence for the occurrence of type 1 or type 2 mispricing errors in the market.

For several reasons, P2P online platforms provide an ideal setting for exploring our research questions. First, online platforms are trading environments with fewer market frictions, e.g. transaction costs are small. Second, investors can invest (i.e., buy an asset) at any time without restrictions, due to the fact that online platforms have no opening/closing times. Third, investors on P2P lending platforms benefit from informational transparency, as historical transaction data are made publicly available on a daily basis. Consequently, on P2P platforms, the impact of a delayed search should be negligible.

To pursue our investigation, we extract data from Bondora, a leading P2P online lending platform in Europe. The estimation sample covers the 2016-2019 period and includes 126,147 loans originating from three countries (Estonia, Spain, and Finland). By applying a machine learning method—namely, least absolute shrinkage and selection operator (LASSO)—to this large data set, we first provide evidence of asset mispricing in the secondary market on *Bondora.com*. More specifically, we observe that some assets, despite being considered highly likely to be sold in the secondary market, are overlooked by potential buyers, leading to the failure of resale. In contrast, we find that some other assets, which are deemed to have a lower likelihood of being sold, can be highly valued by buyers, thereby making the resale successful.

We next examine whether secondary market dealers adjust their assets' prices when, upon scrutinizing past transactions, they discover a divergence between their asset valuation and that of the buyers. In particular, some assets may be valued highly by buyers and may be in high demand. We expect to find that such discoveries would then be reflected in sellers' subsequent asset pricing behaviours. Our results confim the following expectation: secondary market dealers tend to re-value their asset's price when they discover a divergence of valuation. After examining the most recent transaction data (new information), we have found that a seller can raise the listing price of an asset, when another asset with similar characteristics is sold. Alternatively, some sellers may revise their perception of an asset's value, due to high demand, and hence, stop listing a highly valued asset in the secondary market to gain from future loan repayments.

In connection with the phenomenon of asset mispricing, we further study the impact of investor agreement on assets' values and limited inattention on mispricing. We find that a higher degree of agreement on asset valuation is positively related to the number of type 1 mispriced notes (i.e., notes with low probabilities of sale but are sold). In contrast, a higher degree of agreement on asset valuation is negatively related to the number of notes which have high sale probabilities, but are not sold (i.e., type 2 mispriced notes). We argue that due to their cognitive constraints, investors, and especially buyers, may be unable to process all available information efficiently to form their trading decisions. As a result, in the presence of a high level of inattention, investors can inadvertently make mistakes in judging the value of assets, leading to an increase in the number of mispricing incidents.

Our examination builds on the empirical understanding of asset mispricing. Indeed, there exists a wide range of literature on asset mispricing and the sources of mispricing (e.g., Grossman and Vila, 1992; Caballero, 1995; Chowdhry and Nanda, 1998; Acharya and Pedersen, 2005; Duffie, 2010;). There is also empirical evidence supporting these theories (e.g., Brennan and Wang, 2010; Fleckenstein et al., 2014). Our study contributes to the literature on two major fronts. First, while the available empirical studies mainly focus on the existence of mispricing in stock markets, bond markets, or in takeover deals, we turn the attention to mispricing in the secondary markets on a P2P lending platform. Given the growing popularity of P2P lending and its secondary market activities, the results of this study will be beneficial to investors (mostly personal investors) in facilitating their trading decisions and to provide a basis for platforms to improving their functioning efficiency. Second, we propose a novel approach using machine learning methods to examine the existence of mispricing, which could be utilized by researchers to examine data in similar settings.

Our investigation also relates to the behavioural economics/finance literature that has examined the link between investors' behavioural biases and asset pricing. Among others, this paper contributes to the empirical studies which explore the impact of investors' dispersed beliefs on asset valuations (e.g., Diether et al, 2002; Avramov et al., 2009; Chatterjee et al., 2012). We also re-examine the influence of inattention on financial decisions in considering our findings (e.g., Reis, 2006; Cumming and Dai, 2011; Hirshleifer et al., 2011; Hébert and Woodford, 2017). However, we do not rule out other potential sources of asset mispricing. For example, even though a significant amount of information on the assets are available to investors, asymmetric information could still exist and amplify the degree of mispricing (Duffie and Rahi, 1995). Although we are not able to test this channel in the study directly, our results concerning the impact of inattention could also be applied to the case of asymmetric information.[1] Due to inattention and asymmetric information, investors might rely on the assigned credit ratings to make investment decisions (Brennan et al., 2009). However, credit ratings tend to neglect systematic risks and do not accurately reflect the asset quality (Marquesa and Pinto, 2020). Hence, credit rating – based on valuations, might also lead to mispricing of the assets' true value.

Overall, our findings reveal new phenomena in relation to pricing and implications with respect to the valuation of assets in secondary P2P online loan markets. First, we present evidence that some assets are not successfully sold, even though they merit a high chance of being sold, as justified by the prices and other important characteristics of the assets. In contrast, some other assets, despite the low predicted likelihood of sale, are, in fact, sold. Second, we show the speculative trading tendency exhibited by both sellers and buyers, which is induced by information conveyed from past transactions. That is, sellers tend to exploit the buyers' relatively high valuation of an asset by increasing (re-valuing) the asset's price, although this price does not necessarily reflect the asset's fundamentals. At the same time, buyers are willing to buy assets in high demand, even if the assets are overpriced, due to the perception/beliefs that these notes can be easily resold in the market at a later point in time. Third, we provide evidence that investors' inattention, caused by the limited information processing capability, has a positive impact on asset misvaluation.

[1] The impact of inattention is arguably more profound for buyers than sellers, since the number of assets for which buyers have to process information, is significantly larger than the number of assets for which sellers have to process information. In the context of asymmetric information, sellers are likely to have more information on the assets than buyers. Hence, in both cases, buyers are more likely to make mistakes in asset valuations.

The rest of this paper is organised as follows. In Section 2, we review the related literature and present our main hypotheses. Section 3 provides an overview of the Bondora lending platform. Our empirical strategy and data are shown in Section 4. Section 5 discusses the results. Finally, Section 6 concludes and provides several implications.

2. Literature review and hypothesis development

2.1. Theoretical models of asset mispricing

In efficient markets, rational investors are expected to react instantaneously when asset prices deviate from their fundamentals to push the prices back to their levels of equilibrium. However, a significant number of theoretical studies have shown that asset prices can diverge from their fundamental values in the presence of market frictions and financial constraints (Lewis et al., 2017). Some studies suggest that limited debt capacity plays an important role in driving asset prices away from their fundamentals (e.g., Grossman and Vila, 1992; Chowdhry and Nanda, 1998; Basak and Croitoru, 2000; Gromb and Vayanos, 2002; Brunnermeier and Pedersen, 2008). Slow-moving capital caused by delayed search (Wolinsky, 1990; Duffie, 2010), and investor inattention (Caballero, 1995; Lynch, 1996; Gabaix and Laibson, 2002; Chien et al., 2012) are proposed and examined as other important sources of asset mispricing.[2]

Additionally, the difficulty involved in raising equity capital could contribute to the divergence of asset prices from their fundamentals. For instance, Kondor and Vayanos (2019) show that the ability of arbitrageurs to provide liquidity to other traders, who seek to hedge their portfolio risk, is limited by their own capital. Thus, arbitrageurs' capital is an important state variable in determining asset prices. Similar results are found for the equity capital ratio of financial intermediaries (He et al., 2017) or the intermediary leverage ratios (Adrian et al., 2014).[3] Furthermore, transaction costs can lead to a difference between an asset's fundamental cash flow and its actual cash flow, which, in turn, impacts the asset price (Demsetz, 1968; Amihud and Mendelson, 1986; Boudoukh and Whitelaw, 1993; Acharya and Pedersen, 2005).[4]

Several empirical studies have found evidence to support the above theories (e.g., MacKinlay and Ramaswamy, 1988; Cornell and Shapiro, 1989; Longstaff, 2004; Brennan and Wang, 2010; Fleckenstein et al., 2014). Researchers conclude that (1) bonds (stocks) are often mispriced,

---

[2] See also Vayanos and Wang (2007), Weill (2007), Vayanos and Weill (2008), and Duffie and Strulovici (2012).

[3] For example, Xiong (2001), Kyle and Xiong (2001), and Basak and Pavlova (2013).

[4] See also Vayanos (1998), Vayanos and Vila (1999), and Huang and Wang (2009, 2010).

particularly during periods of distress and that (2) the drivers of mispricing are market liquidity, funding risk, and mispricing in other markets.

In addition to the rational-expectation models, the psychology-based asset pricing literature suggests that investors' behaviours/irrationalities also play an important role in explaining mispricing. Thus, a growing number of heterogeneous agent pricing models have been developed, in which investors have dispersed preferences or beliefs/expectations (e.g., Benartzi and Thaler, 1995; Barberis et al., 1998; Daniel et al., 1998; Hong and Stein, 1999; Barberis and Huang, 2001, 2008). The heterogeneous preferences and beliefs/expectations are generated by various behavioural biases, such as overconfidence, overextrapolation, heuristic simplification, and inattention, among others (Baker and Stein, 2004; Alti and Tetlock, 2014).[5]

Motivated by the existing literature, we hypothesize that mispricing can still exist in an environment with non-binding market frictions. More specifically, we aim to investigate the existence of asset mis-valuation in the secondary market of an online P2P lending platform. In this marketplace, where transaction costs are low, sellers can freely list assets that are part of a loan which originated from the platform's primary market. Yet, there is a significant variation in the listing prices across assets, even though they are, de facto, parts of the same P2P loans being repaid. In general, an asset under-priced by sellers should exhibit a higher likelihood of being sold, while the opposite should be observed for an overpriced asset.[6] However, as long as a buyer's valuation of an asset exceeds that of a seller's, a deal can be made. Thus, even an overvalued asset can be sold if buyers perceive the "net" value as positive i.e. buyers further overprice the asset. This is defined as type 1 mispricing in the study. In contrast, an undervalued asset might not find a buyer if the asset is further undervalued by the buyers, which is considered as type 2 mispricing error.

H1: Type 1 and Type 2 mispricing errors exist in P2P loan secondary market

[5] See Hirshleifer (2001) and Barberis and Thaler (2003) for comprehensive reviews of the behavioural asset pricing literature.

[6] To some extent, this is similar to the impact of bid premium on the success of a tender offer. As suggested by Walkling (1985), bidders need to offer a premium that exceeds the market price to ensure the deal's success and an under-priced offer will fail.

2.2. Investors' behaviours and asset (mis)pricing

2.2.1. Dispersion of beliefs

Among various heterogeneous agent pricing models, the Miller's model (Miller, 1977) documenting the role of dispersed opinions in the presence of short selling restrictions, is among the most appealing. It proposes that, in a market where short selling is limited/prohibited, investor disagreement can result in overvaluation, since the demand will stem, in the main, from the most optimistic investors. In this setup, investors could also exhibit speculative behaviour i.e. they will pay a price exceeding the assets' fundamental values in anticipation of opportunities to then resell to other investors, who are willing to pay an even higher price (Harrison and Kreps, 1978; Morris, 1996; Scheinkman and Xiong, 2003; Hong et al., 2006). In other words, in the absence of short selling, opinion divergence across market participants leads to the overpricing of assets and that mispricing is further amplified by the speculative phenomenon. Moreover, the dispersion of opinion can be persistent if investors agree to disagree and hence, infer information from prices incorrectly, or do not rely on prices to revise their beliefs (Banerjee, 2011).

Several empirical studies have attempted to test the prediction using different settings e.g. the trading environment with limited short-selling opportunities, versus the one where short selling does not exist and the results are inconclusive. For example, some studies find that the increased dispersion in investors'/analysts' forecasts/expectations is positively related to higher prices and lower returns (e.g., Diether et al., 2002; Park, 2005; Chatterjee et al., 2012). However, the negative disagreement – stock return link – disappears when credit rating is taken into account (Avramov et al., 2009). Moreover, Carlin et al. (2014) also acknowledges a positive link between disagreement and expected returns, return volatility, and trading volume in the context of the mortgage-backed asset market.

2.2.2. Rational inattention

It has been widely documented in the behavioural economics literature that limited attention can affect decision making (Kahneman, 1973). Building on the information-processing constraint framework, proposed by Sims (1998; 2003), some studies model the costs of acquiring and processing information or model the decision makers' actions in response to the limited information processing capability (e.g., Reis, 2006; Matějka and McKay, 2015; Hébert and Woodford, 2017).

In general, the idea that investors are inattentive, which then affects their financial decision making, is empirically supported by several studies (e.g., Cumming and Dai, 2011; Hirshleifer et al., 2011; Gilbert et al., 2012). For instance, there is evidence of investors' inattention to stock market announcements released on Friday (DellaVigna and Pollet, 2009; Louis and Sun, 2010). Further, inattention, caused by traders' inability to effectively complete two or more tasks simultaneously i.e. dual-task interference, has also been acknowledged as a source of mispricing (Brown, 2014).

Building on these strands of literature, we argue that in our setting, in which information on thousands of listings is publicly available, investors will face information-processing constraints, leading to limited attention. At the same time, different investors might have different inferences of the same set of information, resulting in disagreement in asset valuations. In our analysis, we use the success rates of sales as a proxy of secondary market investor agreement.[7] Similar to the effect of disagreement on price and trading volume of debt contracts, documented by Hong and Sraer (2013), investor agreement could have two different impacts on listing price and volume of secondary market listings. On the one hand, the agreement can be inferred as a positive signal of prospective sales. That is, holders of similar listings can find buyers whose valuations are similar (if not higher) to asset holders' valuations. The high sale rates can also indicate that the listings of a similar type are in high demand by other investors. These inferences can strengthen the speculative incentives, encouraging sellers to increase the prices of subsequent listings. On the other hand, holders of similar listings might revise their beliefs about these assets' values, based on the buyers' beliefs and, therefore, increase their valuations. This could provide them with an incentive to retain these assets, in order to gain from loan repayments, instead of selling in the secondary market.

H2: Investor agreement is positively related to the listing price, but negatively related to sellers' propensity to sell

Moreover, the higher the level of distraction and/or inattention, the more likely investors are to make mistakes when evaluating asset prices. We further posit that the effect of inattention is more profound for buyers than sellers. The reason is intuitive: sellers presumably only need to process the information on the assets they intended to sell, as well as information on other assets which are part of the same initial P2P loan. In contrast, buyers have to choose from

[7] In our case, a high level of agreement can be understood as either (1) most investors correctly price the assets or (2) most investors agree to disagree. Intuitively, the latter is more plausible.

thousands of secondary market listings. As a result, one would expect an increase in mis-valuation, particularly that of the buyers, if the degree of inattention is high.

H3: Limited attention is positively related to mispricing

3. Overview of the Bondora P2P platform

Founded in 2008, Bondora has become one of the best-known peer-to-peer online lending platforms in Continental Europe (Altfi, 2019). According to P2P market data, as of April 2020, Bondora is in top 10 Euro P2P lending platforms by funding amount, with a market share comparable to other leading platforms e.g. Twino (based in Latvia), Fellow Finance (based in Finland), or October (based in the Netherlands). To draw comparisons, there are both similarities and differences between Bondora's and its competitors' business models. For instance, while some platforms mainly lend to (small) businesses (e.g., Finexcap, Folk2Folk, or Funding Circle), or property investors (e.g., Landbay, AssetzCapital), Bondora and some others (e.g., Twino or Zopa) mainly focus on consumer loans.[8] In contrast from the platforms which only admit investors from the country of the headquarters and/or other European countries, Bondora is among the first allowing investors worldwide to invest in its loans. Bondora's loan products were first offered in Estonia; then the marketplace was extended to borrowers in three more countries namely Finland, Spain, and Slovakia.

Bondora's requirements for investing and borrowing are relatively straightforward. Any individual who is over the age of 18 and is a citizen of a European Union country, Switzerland, Norway, or a country approved by Bondora, can register and invest. To borrow money, an individual must create a loan application and provide personal information, contact information, socio-demographic information, information about income and outstanding liabilities, and other supporting data. After an application is submitted, Bondora checks the provided data, including the debt level of the applicant. Once the information is validated, a credit score is assigned to the applicant and a conditional loan offer is made.[9]

Originally, Bondora focused only on retail loans targeting middle-income borrowers who sought mid-sized loans (typically ranging from €500 to €10,000), with a term to maturity ranging from three to 60 months. Since 2016, investors have had three options for investing in the primary

---

[8] That said, loans issued in Bondora consist of both consumer and business loans.

[9] The credit score is determined using Bondora's risk scoring model. Each loan application is assigned one credit score at loan origination and the score does not change over the loan's lifetime. More detail about credit scoring can be found at https://support.bondora.com/hc/en-us/articles/212798989-Risk-scoring.

market. The first option (Go & Grow) is when investors choose to invest in Bondora, instead of individual loans. That is, investors simply deposit money in their Go & Grow accounts and the funds will be allocated by Bondora across primary market loans, based on the investors' preferences e.g. investing purposes (e.g., for retirement) and investing plan (e.g., duration of investment, capital amount, and planned monthly investment). The second option is Portfolio Manager, in which investors can specify their risk preferences e.g. ultra-conservative, conservative, balanced, progressive, or opportunistic. The final option is Portfolio Pro, in which investors have more control over their investment preferences e.g. bid size, choosing loans by countries, credit ratings, or interest rate ranges. The first two options are tools which automatically match investors with borrowers based on their stated preferences.

In March 2013, Bondora launched its secondary market, allowing investors to buy or sell the holdings of the P2P loans that originated from its primary market, with a discounted (or a mark-up/premium) rate, so that an investor can sell the remaining outstanding principal of the P2P loan at a discount or for a premium. A P2P loan listing cannot be placed in the secondary market for longer than 30 days.[10] That is, a note that is not sold within 30 days after its first being listed in the secondary market is automatically removed from the market and is assigned the *Failed* status. If the note is sold, its status is recorded as *Successful*. If the investor decides to cancel the sale, the status will appear as *Cancelled*. Appendix Figure 1 provides an overview of the Bondora secondary market's section for user interface. Within the interface, all important criteria e.g. credit rating, interest rate charged, loan status, scheduled future payments, internal rate of return, and price are clearly displayed. More detailed information about a particular P2P loan and the borrower of that loan can be accessed by clicking on the listing, by expanding the overview, and through the public reports.[11]

There are several features of the platform and its secondary market which are worth noting. First, like other P2P lending platforms, the transaction costs are marginal: investors are not charged for their investments in both primary and secondary markets.[12] Second, investment in the Bondora primary market is 100% automatic, while investment in the secondary market is mostly manual.

---

[10] Henceforth, we will interchangeably use loan part, note, asset, and listing to refer to an asset that an investor posts on the secondary market. In this context, a listing on a secondary market is an asset that an investor posts and is different from a listing that a *borrower* posts to borrow funds from a P2P platform from the *primary market*.

[11] These features can also be added to the secondary market interface for convenient access by changing the setting.

[12] In Bondora, investors are charged a fee of EUR 1 when they withdraw money from their Go and Grow accounts.

One exception is when investors choose the Portfolio Manager option. If investors opt for secondary market investment in pre-setting, secondary market investments will be automatically made, based on the investors' preferences. However, the Portfolio Manager only acquires the notes at par or a discounted value, as well as the notes of loans that do not have overdue or default status at the time of transaction. Third, Bondora provides investors with all relevant loan and borrower information to facilitate investors' decision-making. The information includes (1) borrower profile e.g. job history, age, marital status, education, income, and homeownership status; (2) loan characteristics e.g. credit rating, amount, duration, and purpose; and (3) loan performance e.g. information about loan repayment, collection process or loan schedule. The information is included in Bondora's public reports that are updated daily.

As stated by Bondora (2016), investing in the secondary market is generally riskier than investing in the primary market, for several reasons. More specifically, Bondora allows notes of an overdue, or a defaulted loan, to be sold in the secondary market.[13] Some investors might be interested in buying these loan parts at a discount in exchange for a later payment through the recovery process. However, the actual returns on this kind of investment depend on the collection and recovery efforts. Furthermore, an investor may be willing to pay a premium to buy notes of a loan issued to a low-risk borrower. However, if the paid premium is too high, the investor's actual returns might be far lower than expected. Having said this, investments in the secondary market might not necessarily be riskier than those in the primary market, if the markets are efficient. That is to say, conditional upon the efficiency of the market, the prices paid in the secondary market should compensate for the higher risks relating to the investments. Thus, after adjusting for risks, returns in the primary and secondary markets should be consistent.

4. Data and empirical strategy

4.1. Data and sample

We collect three sets of publicly available data from Bondora.com, covering the period of January 2016 – December 2019. These include (1) loan dataset, which contains information about all public loans, (2) historic payment information, which includes all received payments of granted loans, and (3) secondary market transactions, which include all investments in the secondary

[13] Bondora's take on secondary market loans differs from how secondary markets are managed in other platforms. For instance, investors of FundingCircle can only sell loan parts in the secondary market at par. For Twino's secondary market, there is no clearly defined section of this market in the user interface and defaulted loans cannot be sold in the secondary market.

market. After merging these datasets using the unique loan identification numbers, we take several additional steps to clean the data. First, we drop all loans that are listed in the secondary market and whose primary market information is missing. All loans with less than 36 months of maturity are also excluded from the data. Second, given that loans originating in Slovakia account for a small proportion of the data, we exclude them from the sample.[14] Third, data on the number of listings and the discount rate (or mark-up) are trimmed at the top 1% level of distribution, in order to alleviate the influence of extreme observations. Finally, any listing that has either (1) negative outstanding principal, (2) negative total unpaid principal, or (3) outstanding principal larger than its unpaid total principal is also excluded.

Table 1 presents the descriptive statistics for 65,183 loans in our sample, of which 39,701 originate from Estonia, 9,993 originate from Spain, and 15,489 originate from Finland. Our sample contains only longer-term loans (36-60 months), yielding an average maturity of around 4.1 years. The average loan size is around €2,530, though the size can vary from only €115 to slightly over €10,600. In comparison, the average size of loans originating from Finland is the largest (€3,381), followed by loans originating from Estonia (€2,437) and loans originating from Spain (€1,580). It can also be seen that the interest rate is relatively high; on average, borrowers must pay an interest rate of over 23% on their loans. Borrowers from Spain are charged the highest interest rates, with an average of nearly 80%, while in Finland and Estonia these rates are 42% and 23%, respectively.[15] Reflecting the average interest rates, the average default rate in Estonia is lower than the sample default rate (37.0% versus 48.6%), while the default rates in Spain and Finland are significantly higher at around 67%. More specifically, for every ten loans granted in Spain and Finland, approximately 6 of them default. This suggests a relatively lower level of risk of P2P investment in Estonia, when compared with P2P investment in the other two countries.[16]

(Table 1 about here)

It should be noted that Bandora's secondary market is reasonably active. Panel A of Table 1 shows that the monthly average number of secondary-market loan listings is 158, of which 62 (about 39%) are successful. Further, most notes are listed in the secondary market at a discount,

[14] As of April 2020, the loan amount originating in Slovakia accounts for less than 1% of the total amount of all loans. Note that our findings are quantitatively similar if we include Slovakia in the analysis.

[15] In Bondora, interest rates are fixed interest rates.

[16] Summary statistics by loan types (i.e., consumer loans vs. business loans) are presented in Appendix Table 3.

with an average discount rate of 2.4%. This is to be expected, when one accounts for the term structure of interest rates: Fixed-income assets with shorter effective maturities are expected to be sold at a discount. A comparison of the listing price of successful and failed listings (Table 2) shows that, on average, both successful and failed listings are marketed at a discount in the secondary market. Further, the average discount rate of the successful listings is, indeed, higher than that of the latter (7.5% versus 2.7%), which is expected. In addition, the successful sale of notes in the secondary market happens, on average, during the first eight months following the initial loan offer.[17] It seems that selling notes of a "mature" loan in the secondary market is a challenging endeavour. Moreover, the number of days on which these notes are listed is significantly higher for failed listings than it is for the successful ones. This could be because some sellers must try and fail several times before they recognise the "true" sale price of their notes in the secondary market.

(Table 2 about here)

Figure 1 plots the evolution of the key variables describing Bondora's secondary market on a monthly basis. The number of secondary market listings shows a generally upward trend during the 2016 – 2019 period, with the exception of a spell of decline in the first half of 2018. This pattern is also observed when we track the number of listings by country. Correspondingly, the success rate of secondary market listings increased from 2016 till early 2018, declined between January 2018 – July 2018, then recovered following this. The figure also shows that the secondary market in Estonia is most active, with an average monthly success rate of around 45-50%, while the success rates in Spain and Finland are lower—around 25-30%.

(Figure 1 about here)

Figures 2-3 show the success rate over the lifetime of the loans (top panels), the number of listings (middle panels), and the number of distinct loans in the secondary market (bottom panels). Overall, we observe that investors tend to sell their notes at the beginning of a loan's lifetime. For instance, the number of 36-month loans listed in the secondary market during the first six months of their maturity is around twice as many as the number of loans listed during their final 12 months of maturity. Likewise, most of the 60-month loans are listed in the secondary market during the first 12 months of maturity. A similar pattern is reflected in the evolution of the number of listings. Regarding the success rate of sales over a loan's lifetime, we again acknowledge the

[17] Around 250 days for listings which were successfully sold and 485 days for those which failed to sell.

declining rate when the loans approach their maturity. Specifically, the success rates of the 36-month loans' listings and the 60-month loans' listings remain at around 30% and 20% during the final 12 months of their maturity, respectively.

(Figures 2 and 3 about here)

Figure 4 presents further detail in relation to the pricing behaviour on Bondora's secondary loan market. In essence, one would expect sellers to set a discount for a secondary market listing following the first scheduled loan repayment, i.e., after the borrower begins to pay against the original loan. Moreover, the price of a listing should decrease over time. However, this is not observed in the data. For instance, prior to the first scheduled payment, the number of secondary listings with a premium is around 4-5 times larger than those with a discount (Panels A and B). This observation could also be related to the design of Bondora's secondary markets. As the notes are listed for 30 days, it might be rational for sellers to first try to obtain a high price for the listings and to subsequently lower the price, until a buyer is found.[18]

The dominance of the premium listings is observed for eight months from the first scheduled payment. From the ninth month onwards, we observe that the listings with a discount dominate the platform, i.e., investors have stronger incentives to sell their notes at a discount when the loan approaches maturity and when the remaining outstanding principal is small (Panel D). Moreover, during the period between the 4^{th} and 8^{th} months following the first scheduled payment, the average price of the listings, on which more than half of the principal is repaid, is relatively high with an average premium of around 5% (Panel C). Overall, these observations suggest the existence of irrationality in sellers' pricing behaviour: Sellers are inclined to ask for a premium on their listings when they should be selling their notes at a discount.

(Figure 4 about here)

4.2. Empirical specifications

4.2.1. Identifying asset mispricing in P2P secondary loan markets

Before discussing our approach to identify mispricing in the P2P secondary loan markets, it is useful to consider the misvaluation hypothesis of takeovers in the P2P market context, even though there are differences in the setting of the two markets. More specifically, in the merger and acquisition (M&A)/takeover bids, bidders tend to use their stocks to acquire the overvalued targets, if the target is, in relative terms, less overvalued than the bidder (Shleifer and Vishny,

[18] We thank an anonymous referee for making this suggestion.

2003; Rhodes-Kropf and Viswanathan, 2004). From the perspectives of the bidders, they gain a favourable real exchange ratio by trading their assets for less overvalued target assets. From the targets' perspectives, they accept the offer, due to the overestimated potential synergies of the combination. In other words, a deal is made, even though both bidders and targets are overvalued, as long as the (perceived) gain (for both sides) is positive. Applying this notion in the context of our study, if buyers perceive that the sellers' valuation of the notes is lower than their own valuation, a sale will occur, i.e., buyers will gain from such trades. Since we do not directly observe the investors' asset valuation process, we will take advantage of the outcomes of these valuations in order to classify mispricing.[19]

Inspired by literature investigating the determinants of successful M&A/takeover deals (e.g., Baker et al., 2012; Edmans et al., 2012), we first use all relevant information available to all investors and the offered asset (sale) prices to predict the likelihood of an asset being sold for each listing (Model (1) below). In the spirit of Walkling (1985), a high predicted asset sale probability implies that sellers either correctly price, or under-price the listings relative to the market, while a low predicted success suggests the sellers' relative overvaluation. Variation in the predicted probability of listings which are parts of the same P2P loans will provide us with the preliminary evidence of mispricing behaviour experienced by sellers in the P2P secondary markets.

$$Sold_{i,t} = \alpha + Listing\ characteristics_{i,t}\beta + Loan\ characteristics_{i}\gamma + Borrower\ characteristics_{i}\delta + FEs + \varepsilon_{i,t} \quad (1)$$

where *i* and *t* refer to the note issued on the secondary P2P loan market and time, respectively. The dependent variable, *Sold*, is the binary outcome in the secondary market, which equals one if the note is sold and zero otherwise. Listing characteristics is a vector of listing-specific variables, including (1) the price (premium/discount/flat rate) that the seller requests (*Listing price*), (2) a dummy variable that equals one if the loan had a late interest payment at the time that its note was listed in the secondary market and zero otherwise (*Late payment*), and (3) the natural log of one, plus the note's outstanding principal on the day that it is listed on the secondary market (*Principal*).

[19] Our approach is different from the traditional approach, where researchers attempt to derive asset prices from observed fundamental values.

The vector *Borrower characteristics* includes borrower-specific information that can be used in investment decision making in the secondary market. The information includes (1) the natural log of a borrowers' age (*Age*), (2) gender (*Gender*), (3) whether the borrower has other debts (*Existing debt*), (4) marital status (*Marital status*), (5) the level of education (*Education*), (6) employment status (*Employment*), (7) occupation (*Occupation*), (8) the natural log of total income (*Income*), and (9) homeownership type (*Home ownership*).[20]

The vector *Loan characteristics* contains loan information which can affect the probability of its notes being sold in the secondary market. These characteristics are (1) interest rate charged on loan (*Interest rate*), (2) loan purposes (*Use of loan*), (3) Bondora's credit rating (*Credit rating*). The vector *FEs* incorporates a set of dummies to account for the country effect, seasonality, and time effects, including (1) country ID, (2) day of month, (3) week of year, (4) day of week, (5) hour of day, and (6) month effects. These dummy variables are included to control for all calendar variations, as well as for country-specific characteristics/macroeconomic variations.

The key feature of our mispricing classification is that we need to have an effective prediction model. To achieve this, we implement a machine learning method, namely least absolute shrinkage and selection operator (LASSO), to estimate Model (1). In recent years, this method has gained popularity within asset pricing and forecasting literature, due to its high performance as a model selector and its prediction power (see, e.g., Tian et al., 2015; Sermpinis et al., 2018; Coad and Srhoj, 2019; Zhang et al., 2019). LASSO minimizes the residual sum of squares subject to a penalty (λ) on the absolute size of coefficient estimates. As λ increases, more coefficients are set to zero and dropped. Thus, the variance will decrease at the expense of increasing bias. The variance bias trade-off helps to improve the model's degree of prediction accuracy. To choose the optimal penalty level, we use the Extended Bayesian Information Criterion (EBIC). Due to the fact that our dependent variable is binary, we apply the logit-based LASSO method, which maximizes the log-likelihood subject to λ.[21]

Using the point estimates from the LASSO methodology, we construct the in-sample prediction of the probability of a listing in the secondary market being sold and compare the predicted probability with the realised outcome in order to determine mispricing types. We argue that buyers commit mispricing if (1) a listing with a low predicted likelihood of sale is sold or (2) a

[20] Detailed description of variables is presented in Appendix Table 1.

[21] For a detailed discussion of LASSO and variants of LASSO, see Tibshirani (1996; 2011). For a detailed discussion of applications of LASSO in Stata, see Ahrens et al. (2019).

listing with a high predicted likelihood of sale is not sold. In the case of the former, sellers are overpricing the value of the notes and this should discourage buyers from completing the sale (Hirshleifer et al., 2006). However, the fact that these notes are sold indicates buyers' overvaluation, relative to sellers. In the case of the latter, the notes should be sold, given the fact that the sellers' pricing behaviour is fair (or even potentially under-priced). Yet, the sale was not successful, suggesting that buyers under-price the notes lower than sellers' valuation.

In our investigation, investors commit type 1 mispricing (type 1 error) if the predicted sale probability is ≤ 25% (listings are overvalued by sellers) and the realised outcome is a success. By contrast, investors are considered committing type 2 mispricing (type 2 error) if the predicted probability of a successful outcome is ≥ 75% (listings are undervalued or correctly valued by sellers) but the realised outcome is a failure (i.e., no sale).[22]

4.2.2. <u>*Sellers' repricing behaviour*</u>

In this section, we investigate sellers' (re)pricing behaviour in a secondary loan market. In this context, the analysis that follows will be at the *loan level*, whereas the analysis in Section 4.2.1. is at the *listing level*. Our loan-level specification is as follows:

$$Y_{j,m} = \alpha + \beta_1 Last\ Success_{j,m} + \beta_2 Last\ Sold\ Price_{j,m} + X_{j,m}\gamma + \varepsilon_{j,m} \quad (2)$$

where *j* and *m* refer to loan and month indices, respectively. In this model, we employ two listing-specific variables as the dependent variable. The first one is *No Listings*, which is the natural log of one plus the number of current listings normalized by the number of listing days in month *m* for loan *j*. Examining this variable, we can predict the number of new listings posted for a P2P loan in the secondary loan market. The second dependent variable is *Price*, which is the average listing price for which sellers ask. It helps us understand how sellers value their assets.

Our main variable of interest is *Last Success*, defined as the share of the number of successful listings of the same P2P loan over the last three months. This variable is the proxy of the degree of investor agreement on the values of notes of the same P2P loans. Since pricing behaviour in the current month can also depend on the information derived from the past prices, *Last Sold Price*, which captures the average price (discount/mark-up/flat rate) of the successful listings of the given P2P loan over the last three months, is included. Furthermore, the model incorporates

[22] As a robustness check, we redefine mispricing types by changing the thresholds to 90% for type 1 errors and to 10% for type 2 errors. The results from this exercise are similar to our main findings and are available upon request.

several additional control variables, captured by vector ***X***. To control for the effect of loan maturity, we include *Loan Age*, which is the natural log of one plus the age of the loan (in months). *Outstanding Principal* is the ratio of the outstanding principal in the given month to the original principal. Month fixed effects and loan fixed effects are also added into the model to control for loan-specific characteristics, platform-specific characteristics, e.g. the platform popularity, and macroeconomic conditions. Model (2) is estimated using the fixed effect estimator.

4.2.3. Agreement, inattention, and mispricing

To quantify the effect of investor agreement and limited attention on mispricing, we employ the following loan-level model:

$$Y_{j,m} = \alpha + \beta_1 Last\ Success_{j,m} + \beta_2 Last\ Sold\ Price_{j,m} + \beta_3 Inattention_m + X_{j,m}\gamma + \varepsilon_{j,m} \quad (3)$$

where *j* and *m* refer to loan and month indices, respectively. We employ two indicators of mispricing as the dependent variable, *Y*. The first one is *Type 1*, which is the natural log of one plus the number of type 1 error incidents for loan *j* in month *m*. The second dependent variable is *Type 2*, which is the natural log of one plus the number of type 2 error incidents for loan *j* in month *m*.

Two proxies of inattention, $Listing^{active\ hours}$ and $Listing^{other\ loans}$, are used. $Listing^{other\ loans}$ is the natural log of the number of notes of other P2P loans (rather than loan *j*) listed in the secondary market in month *m*. Our argument is that, when investors, particularly buyers, have increased opportunities to buy notes of other loans, they pay less attention to listings of loan *j*. The second proxy, $Listing^{active\ hours}$, is the natural log of the number of notes listed during the arguably more active trading hours (days) i.e. between 8am – midnight on Monday – Friday. A higher measure indicates a greater flow of information (on the assets) that one needs to process. Hence, there is a high chance that buyers might be unable to efficiently use relevant information in making trading decisions. The remaining variables, depicted by vector ***X*,** are similar to those described in Section 4.2.2. We estimate Model (3) using a fixed-effect estimator with various sets of fixed effects.

5. Results

5.1. Examining loan mispricing

Figure 5 presents the distribution of the in-sample predicted probability of listings from estimating Model (1) using LASSO.[23] The figure presents the distribution for both *realized* outcomes, the dashed line for successful sales and the solid line for failures. Overall, our prediction model performs relatively well, as we observe a right-skewed distribution of the successful listings and a left-skewed distribution of the failed listings. That is, most assets that have high predicted sale probability are, indeed, sold in the secondary market. Similarly, most listings with a low predicted likelihood of sale are not sold. As such, the figure provides evidence of mispricing committed by sellers and buyers. More specifically, some assets which have very low predicted sale probability are sold (type 1 mispricing), while some others with very high predicted likelihood of sale are, in fact, not sold (type 2 mispricing).[24] While the peak for failed listings is around zero, the peak for successful listings is close to 70-80% (not 100% as one would expect in an ideal world).

(Figure 5 about here)

While this finding contributes to the literature as the first reported evidence on the existence of mispricing on P2P secondary markets, considering the uniqueness of the P2P online loan secondary markets, it is useful to recall several distinct and significant observations for this environment. First, we can rule out transaction costs as potential sources of mispricing, as they are marginal in P2P lending platforms. Further, up-to-date data on transactions in both secondary and primary markets are publicly available and accessible to all investors. At the same time, an asset is typically listed for 30 days[25], which provides enough time for potential buyers to consider the viability of an asset, before deciding whether to buy it. Thus, limited access to information and/or limited information itself is also unlikely to be a source of mispricing in this marketplace. Third, potential buyers' capital constraints might lead to type 2 mispricing, as financially constrained investors might be unable to buy high-quality assets. However, the possibility of this

[23] Recall that the listings are parts of loans that were funded in the primary market by several investors, including the current seller, who holds a fraction of the initial total amount raised from the platform.

[24] The ex-post default rate of high-quality loans (those with at least 75% predicted probability of being sold) is about 48.5%, while the ex-post default rate of low-quality loans (those with less than 25% predicted probability of being sold) is about 72.6%.

[25] Unless the sellers withdraw their listings, which is not observed in our data.

is low, given that each listing accounts for only part of a mid-sized loan. Thus, the size of each note constituting part of the initial loan is relatively small.

Given these unique features, we argue that the dispersion between sellers' and buyers' beliefs about an asset's value and investors' inattention would be the primary sources of mispricing in P2P secondary loan markets. Due to cognitive constraints, investors are unable to efficiently process all of the available (and relevant) information on all assets that are on the market. This can then lead to varying interpretations of the information attained by market participants, resulting in disagreement across sellers and buyers about the value of assets (Harris and Raviv, 1993; Banerjee and Kremer, 2010). For example, sellers could potentially list the prices of the notes significantly lower than the actual values in order to liquidate their holdings, which might distort investors' perceptions of the quality of the listings. This, coupled with the buyers' inability to correctly value the assets, makes them reluctant to buy highly discounted notes. Moreover, there may be a degree of irrationality among sellers, making them more likely to sell their notes at a premium, especially during the initial stages of an asset's lifetime. In this case, over-priced notes may be more attractive, due to the buyers' erroneous perception on the future cash flows associated with these assets. Hence, even though one would not attach a high probability of sale to such notes, potential buyers may be willing to pay a premium to purchase them.

Figures 6 and 7 provide further insights on the evolution of mispricing over the lifetime of loans. In particular, these two figures provide the changes in the share of mispricing for notes of loans, whose terms to maturity range from 36 to 60 months, respectively. When we examine 36-month loans, we see that the share of mispriced notes tends to increase over time and only drops during the last six months of a loan's maturity. In the case of 60-month loans, we observe a different pattern: the share of mispriced notes in Estonia is relatively stable over the lifetime of the loans. In Finland, the share of mispriced listings of 60-month loans gradually declines. Similarly, this figure for Spain, despite fluctuations, also shows a decrease when the loans approach their date of maturity. We also notice that, in general, the number of type 1 mispricing accounts for at least 60% of the total quantity of mispricing and the change in the number of this mispricing type appears to be in line with the overall pattern.

(Figures 6 and 7 about here)

5.2. Asset repricing

Given the evidence of mispricing in P2P loan secondary markets, we estimate Model (2) to examine the extent to which sellers re-value (re-price) their assets with the arrival of new information. The expectation is that sellers should be able to learn about the demand of their assets by observing the past transactions of the previous listings that belong to the same P2P loans and price their listings accordingly.[26]

The results presented in Table 3 suggest that *Last Success*, our indicator of investor agreement, is negatively associated with the number of listings. Holding all other variables constant, a one percentage point increase in the share of successful listings in the past three months leads to a decrease of 0.14% in the normalized number of listings of the same P2P loans in the current month. Additionally, the higher degree of investor agreement increases sellers' incentives to increase the price of their current listings. Specifically, a one percentage point increase in *Last Success* is associated with a 0.02 percentage point increase in the average price. Since there is a significant variation in the past success rates (one standard deviation in the estimation sample is 30.17 percentage points), the economic importance of the effect is considerable, despite the small point estimate. That is, a one standard deviation increase in *Last Success* can effectively increase the average listing price by 0.5 percentage points.

Regarding the impact of the sold prices, we observe that sellers are likely to list their assets in the secondary market more often if they see similar listings sold at a higher price in the past. However, the economic significance of this effect is rather weak: a one standard deviation (i.e., 30 percentage points) increase in the average transaction prices in the past is related to an increase of only 0.015% in the number of current month's listings. The higher sold prices in the last three months also have a positive link with the listing prices in the current month. We calculate that a one standard deviation increase in the *Last Sold Price* yields an increase of approximately 2.5 percentage points in the current listing price.

(Table 3 about here)

These results are consistent with the hypotheses presented in Section 2.2 and supports the predictions of the impact of investor agreement on debt overpricing in Hong and Sraer (2013). More specifically, the high degree of agreement on the value of listings of a P2P loan signals the positive prospects of sales of the same loan's notes in the present i.e. the high degree of investor optimism. It also indicates that the listings of this loan are in high demand i.e. the listings are highly valued by buyers. Furthermore, one could possibly interpret the high success rate of sales in the past as a signal that the asset valuation of buyers equals or exceeds the seller's valuation.

[26] Our results are similar when we estimate all models on the sub-samples of each country. These results are available upon request.

These signals, subsequently, could lead to two changes in the behaviours of owners of similar assets. First, some owners might revise their beliefs about the assets' value in order to "match" that of the buyers i.e. increasing perceived values. These investors might then keep the assets in order to gain from loan repayments, instead of selling in the secondary market. Second, some other sellers might want to exploit the buyers' optimism to liquidate their notes at a higher price. That is, these sellers could raise the prices of their listings in the subsequent periods to take advantage of buyers' beliefs about the asset values., The results of this is that, as optimism increases, a debt bubble can quietly emerge with a high price but a low (listing) volume.

Our findings are in line with the argument of Easley et al. (2002), that when investors hold different beliefs, some of those who are equipped with the "true" asset values could take advantage of the disparity in the perceived value of the assets. The results also complement the literature, which argues that the arrival of new information might induce speculative trading behaviour (e.g., Harrison and Kreps, 1978; Morris, 1996). More specifically, based on past transactions, sellers observe the difference between their beliefs and buyers' beliefs about an asset's value. As a result, they can realize future gains by setting higher prices on subsequent listings, taking advantage of the presence of optimistic buyers in the market. It should be noted that this behaviour is considered speculative, as the new (higher) prices might not necessarily reflect the true value of the notes.

The findings for the remaining explanatory variables are as expected. For instance, when the loan reaches the later stages of its term (higher *Loan Age)*, investors are not inclined to sell the asset, or even to sell it at a higher price to earn a "premium." This is meaningful, as loans that are due to mature carry less risk, due to the fact that the borrowers have a proven record that they pay on time. Given that the average interest rate is in the region of 30%, the investor who holds such an asset would not want to sell it at a price lower than the asset's perceived valuation. We also find that assets with a higher share of outstanding principal, *Outstanding Principal*, are less likely to be listed and demand a lower premium in the secondary market. This result is rational, as holders of such assets might want to retain them for a period of time in order to benefit from loan repayment. However, if the original investor becomes sceptical in relation to the future cash flow associated with such an asset, the investor would be more likely to sell it at a discount in order to avoid bearing a loss in the event of a future default.

5.3. Effects of agreement and inattention on mispricing

Thus far, we have provided evidence of seller's pricing behaviour, according to the degree of agreement inferred from a listing's past success. We follow this by examining the influence of the level of investor agreement and inattention on the subsequent mispricing errors. The results in Panel A of Table 4 show that the indicator of agreement, *Last Success*, is positively associated with the number of type 1 mispricings i.e. the number of notes that have low probabilities of being sold, but are in fact sold in the secondary market. Holding other variables constant, if the share of successful listings increases by one standard deviation, the number of type 1 mispriced listings of the same P2P loans will increase by 1.2%.

In contrast, Panel B of Table 4 presents a negative link between *Last Success* and the number of type 2 mispricings i.e. the number of notes that have high probabilities of being sold, but are in fact not sold in the secondary market. The magnitude of the effect is larger compared to the effect on type 1 mispricings: the number of type 2 mispriced listings increases by around 3%, with a one standard deviation increase in *Last Success*. When we examine the impact of the average price of the listings sold earlier in the market (*Last Sold Price*) on mispricing, we find that the coefficients associated with this variable are positively related to the number of type 1 mispricing incidents, but negatively related to the number of type 2 counterparts. The scale of the effects, however, is minor: a change of one standard deviation in the past transaction prices leads to a change of between 0.03-0.06% in the number of mispriced notes. This result confirms that the past transaction price plays a negligible role in determining investors' behaviour in the secondary market.

(Table 4 about here)

These results complement and help to rationalize our findings in Section 5.2. As we have shown earlier, after learning about the agreement on assets' values, owners of assets that are similar to those highly valued by buyers tend to keep these assets in their portfolios, instead of selling them in the secondary market. This reaction, eventually, leads to a reduction in the supply of such notes. At the same time, the high degree of agreement could also induce speculative trading behaviour among buyers (e.g., Hong et al., 2006). That is, buyers would have an incentive to buy assets in high demand with the belief that they could always resell these assets at a later point in time. Due to the differences in supply and demand, one would expect that these notes can be sold at ease, which effectively increases the likelihood of being sold. Rational sellers can also take advantage of the situation, or can, more precisely, exploit the buyers' high valuation of assets and/or speculative trading, in order to gain higher profits by selling notes at higher prices.

Our empirical model also provides evidence that investors' inattention has a positive effect on the mispricing incidents, regardless of the types of mispricing. Yet, the effect on type 2 errors is larger than the impact on type 1 mispricing. More specifically, a 10% increase in the number of notes of other loans listed during the active hours leads to an increase of only 0.1-0.4% in the number of notes with low sale probabilities but are sold, and an increase of 0.5-2.1% in the number of notes with high sale likelihoods but are not sold. Similar effects are observed when we use the number of other loans' listings as the proxy for investors' inattention.

These findings suggest the important role of investors' (in)attention in making trading decisions. That is, the abundance of available information on various loans' notes reduces buyers' capability to process information efficiently and hence, make them more likely to make errors when valuating the price of assets. When the error is an overvaluation of the assets that are already overpriced by sellers, the number of type 1 mispricing incidents increases. Conversely, the number of type 2 mispriced notes increases when buyers undervalue the assets that are priced correctly or under-priced by sellers. Since the effect of inattention is more significant in the latter case, one can infer this as evidence of the dominance of buyers' under-pricing tendency in the presence of limited attention.[27]

In investigating the effects of loan age, we observe that as the maturity date approaches, the number of type 1 mispricing incidents increases while the number of type 2 mispricing incidents decreases. Further, a higher remaining outstanding principal, *Outstanding Principal*, is positively associated with the two types of mispricing. These results are meaningful and in line with our previous findings: Sellers are reluctant to list notes, whose terms are near completion in the secondary market, but once they list these assets, they tend to list them at higher prices. This, consequently, limits the supply of assets with high sale probabilities, while increasing the supply of assets with low sale probabilities (i.e., the overvalued assets). Similarly, an asset with a higher outstanding principal is often listed in the secondary market at a discount, which attracts buyers

[27] In a recent study, Cumming and Hornuf (2020) find that the number of competing loan applications on Zencap, a German P2P lending platform, on a given day is negatively related to the number of investments and the bid amounts on that day. To some extent, this can be considered as an evidence of lenders' under-pricing behaviour in the presence of limited attention, which is consistent with our findings.

who overvalue the asset. At the same time, a big discount on an asset's price could make sceptical buyers, who already undervalue the asset, increasingly unwilling to buy the asset.

5.4. Robustness checks

5.4.1. Selection bias

In regressions that employ the average listing price and mispricing as the dependent variable, we observe the outcome, only if a loan is listed in the secondary marketplace. Thus, the results from these models might suffer from sample selection bias, if the decision to sell a note in the secondary market is not random. To address this concern, we employ the following system of equations:

$$Y_{j,m} = \beta_0 + X_{j,m}\beta_1 + \delta_j + \varepsilon_{j,m} \tag{4.1}$$

$$Listed_{j,m}^{secondary\ market} = \alpha_0 + Z_{j,m}\alpha_1 + \gamma_j + \epsilon_{j,m} \tag{4.2}$$

Equation (4.1) is the outcome equation which models the impact of investor agreement and inattention on (1) listing price, (2) the number of type 1 mispricing errors, and (3) the number of type 2 mispricing errors. These outcomes are only observed if at least one note of loan *j* is listed in the secondary market in month *m* ($Listed^{secondary\ market}$ = 1). Vector *X* contains all explanatory variables specified in Equation (3). Equation (4.2) is the selection equation which models the selection into being listed in the secondary market. In this equation, vector *Z* is a superset of *X*: it contains all variables in *X* and an instrument, namely *New auctions*, which is the natural log of one, plus the number of new loan auctions in the primary market for each maturity in month *m*. This instrument represents the investment opportunities/costs that might affect investors' decisions to sell their loans in the secondary market, while not necessarily affecting the listing price and the listing's outcome.

It should be noted that both equations include loan fixed effects, δ_j and γ_j, to control for the unobserved heterogeneity, which is potentially correlated with other covariates and the two error terms. Given the combination of the endogeneity issues, due to the correlated unobserved time-invariant effects, and sample selection bias, we cannot use a standard Heckman correction to estimate our model. Instead, we adopt the framework proposed by Wooldridge (1995) and later extended by Semykina and Wooldridge (2010), which allows for correcting for sample selection in fixed effects models. This approach can be summarised as follows. First, using Mundlak's (1978) modelling device, the unobserved loan effect is modelled as a linear function of $Z_{j,m}$, $\overline{Z}_j$, and an error term, a_j. The selection equation becomes:

$$Listed_{j,m}^{secondary\ market} = \alpha_0 + Z_{j,m}\alpha_1 + \overline{Z}_j\alpha_2 + v_{j,m} \tag{4.3}$$

where $v_{j,m} = a_j + \epsilon_{j,m}$. For each time period, Equation (4.3) is estimated using a probit model to calculate the inverse Mills ratios, $\widehat{\lambda_{j,m}}$.

Next, similar to the above step, the loan fixed effect in Equation (4.1) is replaced with a linear projection of the time average of all variables in X and the instrument. To correct for selection bias, the calculated $\widehat{\lambda_{j,m}}$ is plugged into Equation (4.1).

$$Y_{j,m} = \beta_0 + X_{j,m}\beta_1 + \overline{X}_j\beta_2 + \overline{New\ auctions_j}\beta_3 + \theta\widehat{\lambda_{j,m}} + u_{j,m} \tag{4.4}$$

Finally, one could add the interaction terms between $\widehat{\lambda_{j,m}}$ and time dummies, which allow the selection correction to be different in each time period, so that consistent estimates can be obtained. Equation (4.4) is estimated using the OLS estimator with the bootstrap standard errors.

The results in Table 5 are generally in line with the findings that we reported using fixed effect models. Specifically, the degree of investor agreement has a positive association with the listing price and the number of type 1 mispricing incidents, while it is negatively associated with the number of type 2 mispricing errors. Moreover, the estimates' magnitudes in Table 5 are relatively similar to those in Tables 3 and 4. One exception is the magnitude of *Last Success* in the estimations with the number of type 1 mispricing errors as the dependent variable: the estimated coefficients on this variable, after correcting for sample selection, are larger. Again, although statistically significant, the effect of the average sold price in the past, *Last Sold Price*, on mispricing, is still not economically important.

(Table 5 about here)

With regard to the impact of investors' inattention on the number of type 1 mispriced notes, we observe that the coefficients on both $Listing^{active\ hours}$ and $Listing^{other\ loans}$ are no longer significant, both statistically and economically. In contrast, the effect of inattention on the number of type 2 mispricing errors remains significant and the scale of its impact is even larger. These results support our previous argument that, when provided with a large volume of information, buyers are more likely to undervalue assets.

5.4.2. <u>*High vs. low credit rating loans*</u>

Since pricing behaviours and investment strategies may differ across assets with different credit ratings, we re-estimate models (2) and (3) on the sub-samples of high- and low-credit rating loans to verify this possibility. The high-credit rating sub-sample includes loans for which the maximum expected loss is 0-9% (i.e., loans that have AA, A, B, or C rating). The low-credit rating group consists of loans for which the maximum expected loss ranges from 9% to more than 25% (i.e., loans that have D, E, F, or HR rating).

As can be seen in Table 6, while the impact of *Last Success* on the sellers' propensity to list their assets in the secondary market is similar, the impact on listing prices is different. In particular, owners of the notes belonging to high-rating loans tend not to change their pricing behaviour based on the successful rates of similar listings in the past. Conversely, owners of the notes belonging to low-rating loans are likely to increase the prices after observing the high degree of agreement on the values of similar notes. This strengthens our argument that the sellers have a tendency to take advantage of the "mismatch" in investors' beliefs about assets' values i.e. the buyers' valuation is equal to, or exceeds that, of the sellers. Evidently, they are more likely to do so when trying to sell the lower rating (and thus, lower quality) assets, which is a rational behaviour from the sellers' perspectives.

(Table 6 about here)

More importantly, although the impact of investor agreement on the supply of highly demanded notes is similar for the notes of high-rating loans and that of low-rating loans, the impact on mispricing is stronger for the low-rating group (Table 7). There are several reasons that can explain this finding. First, as suggested by Hong and Sraer (2013), the upside payoff of a less risky debt claim is more limited than that of a riskier counterpart. Consequently, its valuation is less sensitive to investor disagreement and the resale options for such an asset are lower. In this sense, it is expected that buyers are less likely to display speculative behaviour when trading low-quality notes. Second, while buyers' tendency to buy the low-rating assets might be irrational, it might not be uncommon on Bondora's secondary market. As stated on the website, a number of secondary market investors "have a strategy of buying defaulted loans at a very low price, with the plan to reap the rewards later on when the loans are generating a cash flow from recoveries".[28]

There is also a significant difference in the effects of inattention on the number of notes that have high probabilities of sale but are not sold between the high- and low-rating groups. Specifically, the effect on the former group is nine times larger than the effect on the latter. This is sensible, given that credit rating is one of the main determinants of the probability of

[28] See https://www.bondora.com/en/secondary-market (accessed 9 September 2020).

sale. Thus, the high-credit group should contain more assets that have high sale probabilities.[29] As a result, due to limited attention, buyers are more likely to undervalue the high-quality notes, leading to more type 2 mispricing errors.

(Table 7 about here)

5.4.3. *Predicting probability of sale*

In this section, we will discuss the alternative machine learning methods, that are also gaining increasing popularity in economics and finance research as a prediction/classification tool (e.g., Lahmiri and Bekiros, 2019; Mai et al., 2019; Bianchi et al., 2020; Tobek and Hronec, 2020). The first alternative is the Random Forest (RF) classifier, which consists of multiple decision trees. An example of a decision tree, in the context of this study, is presented in Appendix Figure 2. One can interpret this conceptual decision tree as follows: starting with the root node (whether the listing price is greater than 50%), the tree is expanded into different branches/sub-trees, of which the variables and cut-off values used to split on, are selected to minimize the forecast error. Each decision tree in a random forest model is then trained on a random sample drawn from the training dataset and only a random subset of predictor variables is used in a single tree. At the end, for each data point i.e. each observation, there will be **N** predicted classes obtained from **N** decision trees and the final classification will be made by averaging these predictions. In doing so, the built forest is the combination of **N** weakly correlated trees, which helps reduce the variance of the entire model (at the cost of increasing bias).

However, it should be noted that, through its construction, the predicted class probability obtained from an RF model is not equivalent to the fitted probability obtained from a logit-based LASSO model. More specifically, in the former case, the predicted probability of the sale of observation A is, indeed, the fraction of trees in the forest that classify A as a successful listing. Thus, we do not directly compare this RF-based probability with the LASSO-based probability, but rather use it to classify mispricing errors in the same manner discussed in Section 4.2.1, and then we compare the similarity in mispricing classification across methods.

The second alternative is the Artificial Neural Network (NN) algorithm which is considered to be one of the most complex machine learning methods. The neural network model, applied in this study, is a traditional feedforward network, which is made up of (1) the input layer which

[29] The average predicted probability of sale of notes in the high-quality group is 57% and that in the low-quality group is 50%.

contains all predictors; (2) the hidden layers whose nodes (or neurons) are nonlinear activation functions of the input; and (3) the output layer which consists of the predicted outcomes. The output of one layer is the input of its proceeding layer. For the purpose of simplified illustration, we present an example of a feedforward network model with two hidden layers in Appendix Figure 3. In this example, the input layer has **N** nodes which are **N** predictors in Model (1). Each neuro H_{0K} (K=[1;7]) in the first hidden layer has a corresponding bias (b_K) and is connected with the input X_i (i=[1;N]) through weight ($w_{K,i}$) and bias: $\sum_{i=1}^{N} X_i \times w_{K,i} + b_K$. This weighted summation is passed on a nonlinear activation function, such as a logistic sigmoid function or a hyperbolic tangent function, to obtain the output O_{0K} of neuro H_{0K} (K=[1;7]). These outputs in hidden layer 1 are then used to calculate the output O_{1J} of neuro H_{1J} (J=[1;4]) in the second hidden layer using the same procedure. The final classification is made by applying the procedure to the outputs of hidden layer 2.

The third method is the k-Nearest-Neighbour (KNN) algorithm. As suggested by the name, in this method, the class (e.g., success or failure of an observation) is predicted based on the known outcomes of its nearest neighbours i.e. the observations that are very similar to the one for which we want to predict the outcome. When k>1, the classification is computed by the majority vote. Hence, similar to the RF algorithm, the KNN-based probability does not have the same meaning as the LASSO-based one: for a target observation, the KNN-based probability of sale is the proportion of successful listings among its k nearest neighbours.

To implement these methods using our data, we randomly select 100,000 failed listings and 100,000 successful listings for training. Each of the algorithms is then trained on the training dataset and the trained model is used to predict the likelihood of sale of the remaining listings i.e. the unseen dataset. For the RF classifier, we perform an exhaustive grid search to determine the optimal tuning parameters and find the optimal number of decision trees of 100. The NN classifier is built with one hidden layer with 100 neuros and the rectified linear unit function is employed as the activation function. For the KNN classifier, the number of nearest neighbours is 10. We find that the mispricing classifications obtained from these methods are significantly similar to the one obtained from our baseline LASSO (the similarity scores are around 92%).[30]

We would like to stress that, although the performance of all methods is relatively similar, we prefer LASSO in our analysis due to two advantages. First, the use of LASSO allows us to preserve the full sample for analysis. In contrast, when using other methods, only a part of the

[30] Detailed outputs from each of the three methodologies are available from the authors upon request.

data can be used for the analysis and the prediction results can be sensitive to the changes in the training dataset, which is randomly drawn from the full sample. Second, the model selection feature in LASSO has practical implications regarding the importance of each characteristic/category in predicting the sale probability, which are not easy to achieve using other methods. For instance, the KNN algorithm does not provide any estimated parameters. Similarly, the fact that the RF classification is aggregated over classifications from all trees makes it difficult to determine the important predictors, although the decision rule in each tree is clear (Suss and Treitel, 2019).

6. Conclusion

Building on the extensive literature on asset mispricing, this study investigates the existence and persistence of mispricing in a secondary P2P loan market platform. In our examination, we also explore the extent to which investors reprice their secondary market listings, when provided with access to information relating to past transactions.

Our study contributes to the literature in two ways. Existing studies have suggested that several factors lead to asset mispricing, including investor inattention, delayed search, capital constraints, and transaction costs. However, these factors are documented for mispricing using data extracted from a stock market or a bond market. In our case, we document the existence of mispricing from an online marketplace, where transaction costs are low, the level of informational transparency is high, investors can trade any time at their convenience, and capital constraints are not strictly binding. Second, we take our examination a step further by investigating the extent to which the dispersion of investors' asset valuations and investors' inattention can affect the degree of mispricing.

To carry out the investigation, we use data from Bondora.com, one of the leading online P2P lending platforms in Europe. The data contains information related to (1) transactions and characteristics of loans that originated in the primary market, (2) transactions in the secondary market, and (3) the repayment history of primary market loans. Applying LASSO, a machine learning technique, to a sample of 65,183 P2P loans, we identify the mispricing of listings on Bondora's online secondary market. In particular, we find that some listings with a low likelihood of success are, in fact, successfully sold (type 1 mispricing), while others that have a high likelihood of success have failed to sell (type 2 mispricing).

Given the unique features of the marketplace, e.g., a high level of informational transparency and low transaction costs, we attribute the existence of mispricing to the dispersion of beliefs about asset values among sellers and buyers, as well as the rational inattention caused by cognitive constraints. Particularly, due to limited information processing ability and the fact that a large amount of information on assets is available, investors in the P2P secondary market are unable to process all relevant information to make their trading decisions. Thus, buyers and sellers might interpret the same set of information differently, leading to a difference in their beliefs/valuations of asset values. As a result, some listings, although they are priced correctly by sellers, still cannot be sold as they are undervalued by buyers. Conversely, others, despite being overpriced, are sold. Furthermore, investors, particularly buyers, are more likely to make mistakes in asset valuation when the volume of information to process is large i.e. inattention increases levels of mispricing. These arguments are confirmed in our examination, where we disentangle the effects of belief dispersion from the effects of inattention. Our results are robust to different statistical and machine learning methods.

Several implications can be drawn from our findings. First, the existence of asset mispricing in the P2P secondary market can be attributed to investors' limited ability to process information and to a mismatch in beliefs about asset values. However, these behavioural issues are potentially due to the presence of the large volume of information that investors are required to analyse, which is not resolved simply by a credit rating system. Hence, to facilitate trading in this marketplace and to improve the market efficiency, the platform could introduce tools to help both sellers and buyers, so that both parties can process available information more efficiently. For example, it would beneficial to the sellers if the predicted probability of a sale, based on the most important asset characteristics selected by machine learning algorithms and the proposed prices, could be made available to them. Equipped with additional information, sellers could adjust the price of their assets to maximize the chance of success before listing them in the market. Further, the platform could also provide buyers with a comparison tool, which allows them to compare statistics, such as expected loan repayments of several listings of interest. As a result, asset mispricing and speculative trading could be reduced rendering the market operations more efficient. Second, the growing popularity of P2P lending platforms, generally, and their secondary markets, particularly, requires an in-depth understanding of pricing/trading behaviours in these marketplaces. We believe that an investigation, using data from other platforms while implementing a similar methodology to ours, would be useful in this regard.

Acknowledgements

We are grateful to Douglas Cumming, the Managing Editor, and three anonymous reviewers for valuable comments and suggestions, which have helped to improve this paper. We also thank the participants at the 2019 Money, Macro and Finance Society Conference, 2019 Warsaw International Economic Meeting, and 2019 Econometric Research in Finance Workshop.

Funding: This research was supported by Economic and Social Research Council (ESRC) grant ES/P004741/1.

Declarations of interest: none

References

Acharya, V.V. and Pedersen, L.H., 2005. Asset pricing with liquidity risk. *Journal of Financial Economics*, 77(2), pp.375-410.

Adrian, T., Etula, E. and Muir, T., 2014. Financial intermediaries and the cross-section of asset returns. *The Journal of Finance*, 69(6), pp.2557-2596.

Ahrens, A., Hansen, C.B. and Schaffer, M.E., 2020. Lassopack: Model selection and prediction with regularized regression in Stata. *The Stata Journal*, 20(1), pp.176-235.

AltFi, 2019. Altfi - Peer-To-Peer Lending- State Of The Market Report 2019. [Online] Available at: <https://www.altfi.com/state-of-the-market> [Accessed 3 June 2020].

Alti, A. and Tetlock, P.C., 2014. Biased beliefs, asset prices, and investment: A structural approach. *The Journal of Finance*, 69(1), pp.325-361.

Amihud, Y. and Mendelson, H., 1986. Asset pricing and the bid-ask spread. *Journal of Financial Economics*, 17(2), pp.223-249.

Avramov, D., Chordia, T., Jostova, G. and Philipov, A., 2009. Dispersion in analysts' earnings forecasts and credit rating. *Journal of Financial Economics*, 91(1), pp.83-101.

Baker, M. and Stein, J.C., 2004. Market liquidity as a sentiment indicator. *Journal of Financial Markets*, 7(3), pp.271-299.

Baker, M., Pan, X. and Wurgler, J., 2012. The effect of reference point prices on mergers and acquisitions. *Journal of Financial Economics*, 106(1), pp.49-71.

Banerjee, S. and Kremer, I., 2010. Disagreement and learning: Dynamic patterns of trade. *The Journal of Finance*, 65(4), pp.1269-1302.

Banerjee, S., 2011. Learning from prices and the dispersion in beliefs. *The Review of Financial Studies*, 24(9), pp.3025-3068.

Barberis, N. and Huang, M. 2001. Mental accounting, loss aversion, and individual stock returns. *Journal of Finance*, 56(4), pp.1247-1292.

Barberis, N. and Huang, M. 2008. Stocks as Lotteries: The implications of probability weighting for security prices, *American Economic Review*, 98(5), pp.2066-2100.

Barberis, N. and Thaler, R. 2003. A survey of behavioral finance, ch. 18, p. 1053-1128 in Constantinides, G.M., Harris, M. and Stulz, R. M. eds., Handbook of the Economics of Finance, Vol. 1, Part 2, Elsevier.

Barberis, N., Shleifer, A. and Vishny, R. 1998. A model of investor sentiment, *Journal of Financial Economics*, 49, issue 3, pp.307-343.

Basak, S. and Croitoru, B., 2000. Equilibrium mispricing in a capital market with portfolio constraints. *The Review of Financial Studies*, 13(3), pp.715-748.

Basak, S. and Pavlova, A., 2013. Asset prices and institutional investors. *American Economic Review*, 103(5), pp.1728-58.

Benartzi, S. and Thaler, R. 1995. Myopic loss aversion and the equity premium puzzle. *The Quarterly Journal of Economics*, 110(1), pp.73-92.

Bianchi, D., Büchner, M. and Tamoni, A., 2020. Bond risk premiums with machine learning. *The Review of Financial Studies*, forthcoming.

Boudoukh, J. and Whitelaw, R.F., 1993. Liquidity as a choice variable: A lesson from the Japanese government bond market. *The Review of Financial Studies*, 6(2), pp.265-292.

Brennan, M.J. and Wang, A.W., 2010. The mispricing return premium. *The Review of Financial Studies*, 23(9), pp.3437-3468.

Brennan, M.J., Hein, J. and Poon, S.H., 2009. Tranching and rating. *European Financial Management*, 15(5), pp.891-922.

Brown, A., 2014. Information processing constraints and asset mispricing. *The Economic Journal*, 124(575), pp.245-268.

Brunnermeier, M.K. and Pedersen, L.H., 2008. Market liquidity and funding liquidity. *The Review of Financial Studies*, 22(6), pp.2201-2238.

Caballero, R.J., 1995. Near-rationality, heterogeneity, and aggregate consumption. *Journal of Money, Credit and Banking*, 27(1), pp.29-48.

Carlin, B.I., Longstaff, F.A. and Matoba, K., 2014. Disagreement and asset prices. *Journal of Financial Economics*, 114(2), pp.226-238.

Chatterjee, S., John, K. and Yan, A., 2012. Takeovers and divergence of investor opinion. *The Review of Financial Studies*, 25(1), pp.227-277.

Chien, Y., Cole, H. and Lustig, H., 2012. Is the volatility of the market price of risk due to intermittent portfolio rebalancing? *American Economic Review*, 102(6), pp.2859-96.

Chowdhry, B. and Nanda, V., 1998. Leverage and market stability: The role of margin rules and price limits. *The Journal of Business*, 71(2), pp.179-210.

Coad, A. and Srhoj, S., 2019. Catching Gazelles with a Lasso: Big data techniques for the prediction of high-growth firms. *Small Business Economics*, pp.1-25.

Cornell, B. and Shapiro, A.C., 1989. The mispricing of US Treasury bonds: A case study. *The Review of Financial Studies*, 2(3), pp.297-310.

Cumming, D. and Dai, N., 2011. Fund size, limited attention and valuation of venture capital backed firms. *Journal of Empirical Finance*, 18(1), pp.2-15.

Cumming, D. and Hornuf, L., 2020. Marketplace lending of SMEs. *CESifo Working Paper* 8100.

Daniel, K., Hirshleifer, D. and Subrahmanyam, A., 1998. Investor psychology and security market under-and overreactions. *The Journal of Finance*, 53(6), pp.1839-1885.

DellaVigna, S. Pollet, J.M. 2009. Investor inattention and Friday earnings announcements, *Journal of Finance*, 64(2), pp.709-749.

Demsetz, H., 1968. The cost of transacting. *The Quarterly Journal of Economics*, 82(1), pp.33-53.

Diether, K.B., Malloy, C.J. and Scherbina, A., 2002. Differences of opinion and the cross section of stock returns. *The Journal of Finance*, 57(5), pp.2113-2141.

D'mello, R. and Shroff, P.K., 2000. Equity undervaluation and decisions related to repurchase tender offers: An empirical investigation. *The Journal of Finance*, 55(5), pp.2399-2424.

Dong, M., Hirshleifer, D., Richardson, S. and Teoh, S.H., 2006. Does investor misvaluation drive the takeover market? *The Journal of Finance*, 61(2), pp.725-762.

Duffie, D. and Rahi, R., 1995. Financial market innovation and security design: An introduction. *Journal of Economic Theory*, 65(1), pp.1-42.

Duffie, D. and Strulovici, B., 2012. Capital mobility and asset pricing. *Econometrica*, 80(6), pp.2469-2509.

Duffie, D., 2010. Presidential address: Asset price dynamics with slow-moving capital. *The Journal of Finance*, 65(4), pp.1237-1267.

Easley, D., Hvidkjaer, S. and O'hara, M., 2002. Is information risk a determinant of asset returns? *The Journal of Finance*, 57(5), pp.2185-2221.

Edmans, A., Goldstein, I. and Jiang, W., 2012. The real effects of financial markets: The impact of prices on takeovers. *The Journal of Finance*, 67(3), pp.933-971.

Fleckenstein, M., Longstaff, F.A. and Lustig, H., 2014. The TIPS-treasury bond puzzle. *The Journal of Finance*, 69(5), pp.2151-2197.

Gabaix, X. and Laibson, D., 2002. The 6D bias and the Equity-Premium Puzzle. *NBER Macroeconomics Annual*, pp.257-330.

Gilbert, T., Kogan, S., Lochstoer, L. and Ozyildirim, A., 2012. Investor inattention and the market impact of summary statistics. *Management Science*, 58(2), pp.336-350.

Gromb, D. and Vayanos, D., 2002. Equilibrium and welfare in markets with financially constrained arbitrageurs. *Journal of Financial Economics*, 66(2-3), pp.361-407.

Grossman, S.J. and Vila, J.L., 1992. Optimal dynamic trading with leverage constraints. *Journal of Financial and Quantitative Analysis*, 27(2), pp.151-168.

Harris, M. and Raviv, A., 1993. Differences of opinion make a horse race. *The Review of Financial Studies*, 6(3), pp.473-506.

Harrison, J.M. and Kreps, D.M., 1978. Speculative investor behavior in a stock market with heterogeneous expectations. *The Quarterly Journal of Economics*, 92(2), pp.323-336.

He, Z., Kelly, B. and Manela, A., 2017. Intermediary asset pricing: New evidence from many asset classes. *Journal of Financial Economics*, 126(1), pp.1-35.

Hébert, B. and Woodford, M., 2017. Rational inattention and sequential information sampling. NBER Working Paper W23787.

Hirshleifer, D., 2001. Investor psychology and asset pricing. *Journal of Finance*, 56(4), pp.1533-1596.

Hirshleifer, D., Lim, S.S. and Teoh, S.H., 2011. Limited investor attention and stock market misreactions to accounting information. *The Review of Asset Pricing Studies*, 1(1), pp.35-73.

Hirshleifer, D., Subrahmanyam, A. and Titman, S. 2006. Feedback and the success of irrational investors. *Journal of Financial Economics*, 81(2), pp.311-338.

Hong, H. and Sraer, D., 2013. Quiet bubbles. *Journal of Financial Economics*, 110(3), pp.596-606.

Hong, H. and Stein, J.C., 1999. A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of Finance*, 54(6), pp.2143-2184.

Hong, H., Scheinkman, J. and Xiong, W., 2006. Asset float and speculative bubbles. *The Journal of Finance*, 61(3), pp.1073-1117.

Huang, J. and Wang, J., 2009. Liquidity and market crashes. *The Review of Financial Studies*, 22(7), pp.2407-2443.

Huang, J. and Wang, J., 2010. Market liquidity, asset prices, and welfare. *Journal of Financial Economics*, 95(1), pp.107-127.

Kahneman, D., 1973. Attention and effort (Vol. 1063). Englewood Cliffs, NJ: Prentice-Hall.

Kondor, P. and Vayanos, D., 2019. Liquidity risk and the dynamics of arbitrage capital. *The Journal of Finance*, 74(3), pp.1139-1173.

Kyle, A.S. and Xiong, W., 2001. Contagion as a wealth effect. *The Journal of Finance*, 56(4), pp.1401-1440.

Lahmiri, S. and Bekiros, S., 2019. Can machine learning approaches predict corporate bankruptcy? Evidence from a qualitative experimental design. *Quantitative Finance*, 19(9), pp.1569-1577.

Lewis, K.F., Longstaff, F.A. and Petrasek, L., 2017. Asset mispricing. NBER Working Paper W23231.

Longstaff, F.A., 2004. The flight-to-liquidity premium in US treasury bond prices. *Journal of Business*, 77(3), pp.511-526.

Louis, H. and Sun, A., 2010. Investor inattention and the market reaction to merger announcements. *Management Science*, 56(10), pp.1781-1793.

Lynch, A.W., 1996. Decision frequency and synchronization across agents: Implications for aggregate consumption and equity return. *The Journal of Finance*, 51(4), pp.1479-1497.

MacKinlay, A.C. and Ramaswamy, K., 1988. Index-futures arbitrage and the behavior of stock index futures prices. *The Review of Financial Studies*, 1(2), pp.137-158.

Mai, F., Tian, S., Lee, C. and Ma, L., 2019. Deep learning models for bankruptcy prediction using textual disclosures. *European Journal of Operational Research*, 274(2), pp.743-758.

Marques, M.O. and Pinto, J.M., 2020. A comparative analysis of ex-ante credit spreads: Structured finance versus straight debt finance. *Journal of Corporate Finance*, forthcoming.

Matějka, F. and McKay, A., 2015. Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105(1), pp.272-98.

Miller, E.M. 1977. Risk, uncertainty, and divergence of opinion, *The Journal of Finance*, 32(4), pp.1151-1168.

Morris, S., 1996. Speculative investor behavior and learning. *The Quarterly Journal of Economics*, 111(4), pp.1111-1133.

Mundlak, Y., 1978. On the pooling of time series and cross section data. *Econometrica*, 46, pp.69-85.

Park, C., 2005. Stock return predictability and the dispersion in earnings forecasts. The Journal of Business, 78(6), pp.2351-2376.

Reis, R. 2006. Inattentive consumers. *Journal of Monetary Economics*, 53(8) pp.1761-1800.

Rhodes-Kropf, M. and Viswanathan, S., 2004. Market valuation and merger waves. *The Journal of Finance*, 59(6), pp.2685-2718.

Rhodes-Kropf, M., Robinson, D.T. and Viswanathan, S., 2005. Valuation waves and merger activity: The empirical evidence. *Journal of Financial Economics*, 77(3), pp.561-603.

Sadka, R. and Scherbina, A., 2007. Analyst disagreement, mispricing, and liquidity. *The Journal of Finance*, 62(5), pp.2367-2403.

Scheinkman, J.A. and Xiong, W., 2003. Overconfidence and speculative bubbles. *Journal of Political Economy*, 111(6), pp.1183-1220.

Semykina, A. and Wooldridge, J.M., 2010. Estimating panel data models in the presence of endogeneity and selection. *Journal of Econometrics*, 157(2), pp.375-380.

Sermpinis, G., Tsoukas, S. and Zhang, P., 2018. Modelling market implied ratings using LASSO variable selection techniques. *Journal of Empirical Finance*, 48, pp.19-35.

Shleifer, A. and Vishny, R.W., 2003. Stock market driven acquisitions. *Journal of Financial Economics*, 70(3), pp.295-311.

Sims, C.A., 1998. Stickiness. In Carnegie-Rochester Conference Series on Public Policy (Vol. 49, pp. 317-356). North-Holland.

Sims, C.A., 2003. Implications of rational inattention. *Journal of Monetary Economics*, 50(3), pp.665-690.

Suss, J. and Treitel, H., 2019. Predicting bank distress in the UK with machine learning. Bank of England Staff Working Paper No. 831.

Tian, S., Yu, Y. and Guo, H., 2015. Variable selection and corporate bankruptcy forecasts. *Journal of Banking & Finance*, 52, pp.89-100.

Tibshirani, R. 1996. Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society: Series B*, 58(1), pp.267–288.

Tibshirani, R. 2011. Regression shrinkage selection via the LASSO: a retrospective. *Journal of the Royal Statistical Society: Series B.*, 73(3), pp.273-282.

Tobek, O. and Hronec, M., 2020. Does it pay to follow anomalies research? Machine learning approach with international evidence. *Journal of Financial Markets*, forthcoming.

Vayanos, D. and Vila, J.L., 1999. Equilibrium interest rate and liquidity premium with transaction costs. *Economic Theory*, 13(3), pp.509-539.

Vayanos, D. and Wang, T., 2007. Search and endogenous concentration of liquidity in asset markets. *Journal of Economic Theory*, 136(1), pp.66-104.

Vayanos, D. and Weill, P.O., 2008. A search-based theory of the on-the-run phenomenon. *The Journal of Finance*, 63(3), pp.1361-1398.

Vayanos, D., 1998. Transaction costs and asset prices: A dynamic equilibrium model. *The Review of Financial Studies*, 11(1), pp.1-58.

Walkling, R.A., 1985. Predicting tender offer success: A logistic analysis. *Journal of Financial and Quantitative Analysis*, pp.461-478.

Weill, P.O. 2007. Leaning against the wind. *The Review of Economic Studies*, 74(4), pp.1329-1354.

Wolinsky, A., 1990. Information revelation in a market with pairwise meetings. *Econometrica*, 58, pp.1-23.

Wooldridge, J.M., 1995. Selection corrections for panel data models under conditional mean independence assumptions. *Journal of Econometrics*, 68(1), pp.115-132.

Xiong, W., 2001. Convergence trading with wealth effects: an amplification mechanism in financial markets. *Journal of Financial Economics*, 62(2), pp.247-292.

Zhang, Y., Ma, F. and Wang, Y., 2019. Forecasting crude oil prices with a large set of predictors: Can LASSO select powerful predictors? *Journal of Empirical Finance*, 54, pp.97-117.

Tables

Table 1. Summary statistics by loans for the full sample

	Mean	Min	Max	SD	N
	(1)	(2)	(3)	(4)	(5)
	Panel A. All				
Original principal	2,530.145	115.000	10,630.000	2,199.551	65,183
Default	0.486	0.000	1.000	0.500	65,183
Interest rate	36.647	7.430	264.310	31.215	65,183
Maturity	49.160	36.000	60.000	11.316	65,183
No. of listings	157.957	1.000	2,425.000	193.630	65,183
No. of success	62.005	0.000	1,109.000	70.423	65,183
Listing price	-2.360	-100.000	40.000	9.753	65,183
Share of principal	0.006	0.000	1.000	0.026	65,183
	Panel B. Estonia				
Original principal	2,437.134	200.000	10,630.000	2,191.028	39,701
Default	0.370	0.000	1.000	0.483	39,701
Interest rate	23.429	8.080	76.080	8.804	39,701
Maturity	49.115	36.000	60.000	11.47	39,701
No. of listings	158.303	1.000	2,425.000	196.875	39,701
No. of success	70.679	0.000	1,109.000	77.032	39,701
Listing price	-0.960	-100.000	33.104	7.843	39,701
Share of principal	0.006	0.000	1.000	0.028	39,701
	Panel C. Spain				
Original principal	1,580.415	115.000	10,630.000	1,592.950	9,993
Default	0.678	0.000	1.000	0.467	9,993
Interest rate	80.142	12.000	264.310	53.812	9,993
Maturity	46.834	36.000	60.000	10.977	9,993
No. of listings	149.685	1.000	2,107.000	161.081	9,993
No. of success	43.132	0.000	720.000	45.786	9,993
Listing price	-7.839	-100.000	40.000	13.924	9,993
Share of principal	0.009	0.000	1.000	0.034	9,993
	Panel D. Finland				
Original principal	3,381.282	375.000	10,630.000	2,256.128	15,489
Default	0.661	0.000	1.000	0.473	15,489
Interest rate	42.465	7.430	230.390	18.210	15,489
Maturity	50.777	36.000	60.000	10.851	15,489
No. of listings	162.406	1.000	2,081.000	204.043	15,489
No. of success	51.948	0.000	895.000	61.524	15,489
Listing price	-2.415	-90.000	39.000	9.654	15,489
Share of principal	0.003	0.000	0.943	0.010	15,489

This table reports the descriptive statistics for all loans as well as loans generated in Estonia, in Spain, and in Finland (Panels A-D, respectively). *Original principal* is the size of loan (in Euro). *Default* is a dummy variable, which equals to one if the loan's status is default and zero otherwise. *Interest rate* is the interest rate (%) charged on loans. *Maturity* is loan maturity (in months). *No. of listings* is a loan's total number of secondary market listings. *No. of success* is a loan's total number of successful listings in the secondary market. *Share of principal* is the ratio of the listing's principal to the original principal at the time of being listed in the secondary market. *Listing price* (%) is the mark-up/discount rate asked by loan sellers.

Table 2. Summary statistics for failed and successful listings

	Fail			Success			Diff
	Mean	SD	N	Mean	SD	N	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Listing price	-2.712	22.415	6,254,412	-7.495	21.544	4,041,692	4.783***
Share of principal	0.003	0.011	6,254,412	0.003	0.008	4,041,692	0.000***
No. of listed days	220.069	147.070	6,254,412	156.077	121.166	4,041,692	63.992***
Days since loan origination	485.314	381.517	6,254,412	250.471	299.132	4,041,692	234.843***
Days till maturity	1,250.547	777.817	2,630,347	1,362.849	841.257	2,438,908	-112.302***

This table reports the t-test for mean difference between failed and successful listings. *Listing price* (%) is the mark-up/discount rate asked by loan sellers. *Share of principal* is the ratio of the listing's unpaid principal to the total loan amount at the time of being listed in the secondary market. *No. of listed dates* is the number of distinct days on which a P2P loan has at least one listing in the secondary market. *Days since loan origination* is the gap between the day when a loan is originated and the day when its note is listed in the secondary market. *Days till maturity* is the gap between the day when a note of a loan is listed in the secondary market and the loan's maturity date.

Table 3. Effects on normalised number of listings and average listing price

	No Listings	Price
	(1)	(2)
Last Success	-0.1370***	1.6616***
	(0.0015)	(0.0773)
Last Sold Price	0.0005***	0.0849***
	(0.0000)	(0.0008)
Outstanding Principal	-0.1370***	-21.6450***
	(0.0034)	(0.1769)
Loan Age	-0.0401***	0.6336***
	(0.0009)	(0.0476)
Observations	748,638	748,820
R-squared	0.3976	0.5434
Month FE	YES	YES
Loan FE	YES	YES

This table presents the results for Model (2). In Column (1), the dependent variable is the natural log of one plus the number of current listings normalized by the number of listing days in the current month (*No Listings*). In Column (2), the dependent variable is the average price (discount, mark-up, or flat rate) asked by the sellers (*Price*). *Last Success* is the share of the number of successful listings over the last 3 months. *Last Sold Price* is the average price (discount, mark-up, or flat rate) of the successful listings for a given loan over the last 3 months. *Loan Age* is the logarithm of 1 plus the age of loans (in months). *Outstanding Principal* is the ratio of the outstanding principal in the given month to the original principal. In all estimations, a constant term is included but not reported. Robust standard errors are in parentheses. *, **, and *** indicate significant level at 10%, 5%, and 1%, respectively.

Table 4. Effects on types 1 and 2 mispricing

	(1)	(2)	(3)	(4)
	Panel A. Type 1 mispricing			
Last Success	0.0428***	0.0352***	0.0428***	0.0346***
	(0.0040)	(0.0031)	(0.0040)	(0.0031)
Last Sold Price	0.0016***	0.0019***	0.0016***	0.0019***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Outstanding Principal	0.1829***	0.1923***	0.1829***	0.1923***
	(0.0115)	(0.0072)	(0.0115)	(0.0072)
Loan Age	0.0778***	0.1020***	0.0778***	0.1020***
	(0.0025)	(0.0019)	(0.0025)	(0.0019)
$\text{Listing}^{\text{active hours}}$	0.0360***	0.0072**		
	(0.0039)	(0.0034)		
$\text{Listing}^{\text{other loans}}$			0.0339***	0.0136***
			(0.0037)	(0.0036)
Observations	753,099	748,820	753,099	748,820
R-squared	0.0429	0.2663	0.0429	0.2664
	Panel B. Type 2 mispricing			
Last Success	-0.0913***	-0.0978***	-0.0913***	-0.0954***
	(0.0022)	(0.0018)	(0.0022)	(0.0018)
Last Sold Price	-0.0012***	-0.0011***	-0.0012***	-0.0011***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Outstanding Principal	0.1187***	0.1199***	0.1187***	0.1197***
	(0.0062)	(0.0043)	(0.0062)	(0.0043)
Loan Age	-0.0351***	-0.0270***	-0.0351***	-0.0273***
	(0.0017)	(0.0011)	(0.0017)	(0.0011)
$\text{Listing}^{\text{active hours}}$	0.0500***	0.2051***		
	(0.0024)	(0.0020)		
$\text{Listing}^{\text{other loans}}$			0.0472***	0.2278***
			(0.0022)	(0.0021)
Observations	753,099	748,820	753,099	748,820
R-squared	0.0810	0.3025	0.0810	0.3038
Month FE	YES	NO	YES	NO
Year FE	NO	YES	NO	YES
Month of year FE	NO	YES	NO	YES
Loan FE	YES	YES	YES	YES

This table presents results for Model (3). The dependent variables are *Type1Mispricing* and *Type2Mispricing* (Panels A and B, respectively). These variables are defined as the ratios of the number of type 1 (or type 2) error to total number of listings. Type 1 error is the loan that has predicted success probability of ≤ 25% but has realised outcome of being sold. Type 2 error is the loan that has predicted success probability of ≥ 75% but has realised outcome of being failed. *Last Success* is the share of the number of successful listings over the last 3 months. *Last Sold Price* is the average price (discount, mark-up, or flat rate) of the successful listings for a given loan over the last 3 months. *Loan Age* is the logarithm of 1 plus the age of loans (in months). *Outstanding Principal* is the ratio of the outstanding principal in the given month to the original principal. *Other loans* is the natural log of the number of other P2P loans of which notes are listed in the secondary market. $\text{Listing}^{\text{active hours}}$ is the share of noted listed during the arguably more active trading hours (days) i.e. between 8am – midnight on Monday – Friday. In all estimations, a constant term is included but not reported. Robust standard errors are in parentheses. *, **, and *** indicate significant level at 10%, 5%, and 1%, respectively.

Table 5. Sample selection correction

	Price	Type 1	Type 1	Type 2	Type 2
	(1)	(2)	(3)	(4)	(5)
Last Success	1.6150***	0.0728***	0.0735***	-0.0889***	-0.0894***
	(0.1073)	(0.0045)	(0.0044)	(0.0047)	(0.0047)
Last Sold Price	0.0854***	0.0017***	0.0017***	-0.0012***	-0.0012***
	(0.0006)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Outstanding Principal	-19.8641***	0.1605***	0.1633***	0.1301***	0.1251***
	(0.2903)	(0.0031)	(0.0033)	(0.0027)	(0.0028)
Loan Age	-1.3810***	0.1118***	0.1104***	-0.0501***	-0.0477***
	(0.0475)	(0.0037)	(0.0037)	(0.0022)	(0.0022)
$Listing^{active\ hours}$		0.0039		0.0807***	
		(0.0042)		(0.0020)	
$Listing^{other\ loans}$			0.0060		0.0721***
			(0.0041)		(0.0020)
Observations	807,222	807,222	807,222	807,222	807,222
Month FE	YES	YES	YES	YES	YES
Loan FE	YES	YES	YES	YES	YES

This table presents results for models (2) and (3), correcting the sample selection bias. The results in Column (1) are based on Model (2) where the dependent variable is the average price (discount, mark-up, or flat rate) asked by the sellers (*Price*). The results in Columns (2)-(5) are based on Model (3) where the dependent variables are *Type1Mispricing* (Columns (2) and (4)) and *Type2Mispricing* (Columns (3) and (5)). These variables are defined as the ratios of the number of type 1 (or type 2) error to total number of listings. Type 1 error is the loan that has predicted success probability of ≤ 25% but has realised outcome of being sold. Type 2 error is the loan that has predicted success probability of ≥ 75% but has realised outcome of being failed. *Last Success* is the share of the number of successful listings over the last 3 months. *Last Sold Price* is the average price (discount, mark-up, or flat rate) of the successful listings for a given loan over the last 3 months. *Loan Age* is the logarithm of 1 plus the age of loans (in months). *Outstanding Principal* is the ratio of the outstanding principal in the given month to the original principal. *NewAuctions* is the natural log of one plus the number of new loan auctions in the primary market for each maturity – country – month. In all estimations, a constant term is included but not reported. Standard errors which are obtained using a panel bootstrap procedure with loan as the cluster are in parentheses. *, **, and *** indicate significant level at 10%, 5%, and 1%, respectively.

Table 6. Effects on normalised number of listings and average listing price - High vs. low credit rating loans

	No Listings		Price	
	High rating	Low rating	High rating	Low rating
	(1)	(2)	(3)	(4)
Last Success	-0.1565***	-0.1234***	-0.0189	2.6072***
	(0.0023)	(0.0020)	(0.1023)	(0.1093)
Last Sold Price	0.0004***	0.0005***	0.0809***	0.0816***
	(0.0000)	(0.0000)	(0.0012)	(0.0010)
Outstanding Principal	-0.0845***	-0.2010***	-16.6223***	-23.9514***
	(0.0056)	(0.0045)	(0.2510)	(0.2491)
Loan Age	-0.0512***	-0.0275***	2.0562***	-0.1242*
	(0.0014)	(0.0012)	(0.0618)	(0.0683)
Observations	302,948	445,173	303,001	445,302
R-squared	0.4313	0.3806	0.4919	0.5525
Month FE	YES	YES	YES	YES
Loan FE	YES	YES	YES	YES

This table presents the results for Model (2) with sub-samples of loans with high credit ratings (Columns (1)-(2)) and loans with low credit ratings (Columns (3)-(4)). In Columns (1) and (2), the dependent variable is the natural log of one plus the number of current listings normalized by the number of listing days in the current month (*No Listings*). In Columns (3) and (4), the dependent variable is the average price (discount, mark-up, or flat rate) asked by the sellers (*Price*). *Last Success* is the share of the number of successful listings over the last 3 months. *Last Sold Price* is the average price (discount, mark-up, or flat rate) of the successful listings for a given loan over the last 3 months. *Loan Age* is the logarithm of 1 plus the age of loans (in months). *Outstanding Principal* is the ratio of the outstanding principal in the given month to the original principal. In all estimations, a constant term is included but not reported.. Robust standard errors are in parentheses. *, **, and *** indicate significant level at 10%, 5%, and 1%, respectively.

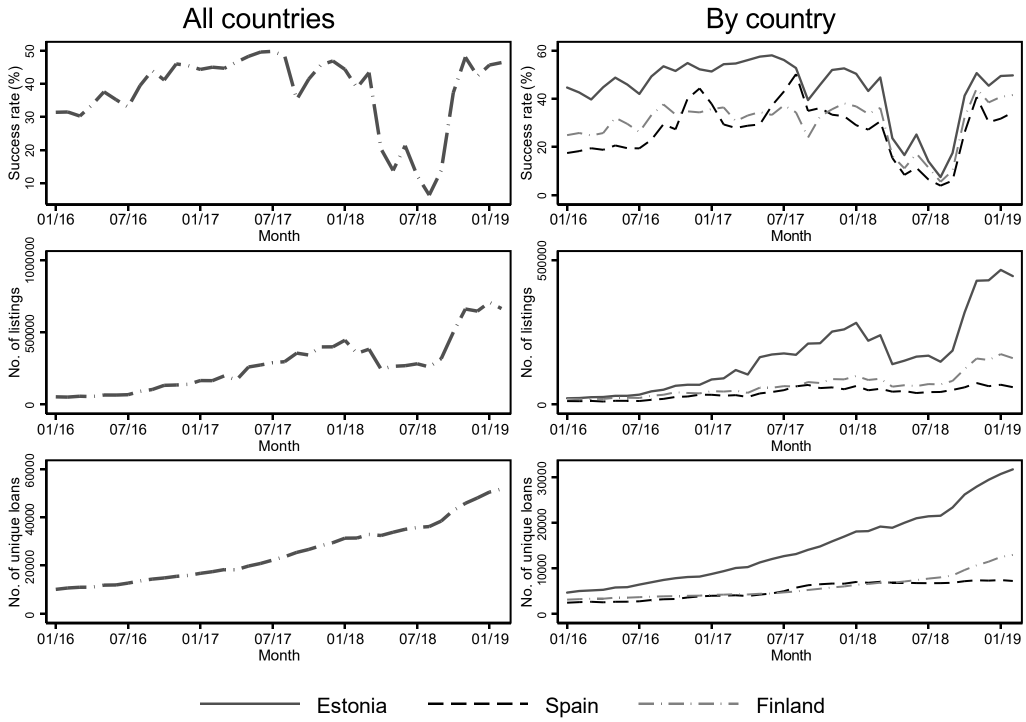
Table 7. Effects on mispricing – High vs. low credit rating loans

	High rating		Low rating	
	(1)	(2)	(3)	(4)
	Panel A. Type 1 mispricing			
Last Success	0.0074	0.0074	0.0794***	0.0794***
	(0.0064)	(0.0064)	(0.0051)	(0.0051)
Last Sold Price	0.0017***	0.0017***	0.0015***	0.0015***
	(0.0001)	(0.0001)	(0.0000)	(0.0000)
Outstanding Principal	0.2796***	0.2795***	0.0702***	0.0702***
	(0.0205)	(0.0205)	(0.0144)	(0.0144)
Loan Age	0.0669***	0.0670***	0.0945***	0.0945***
	(0.0041)	(0.0041)	(0.0033)	(0.0033)
$Listing^{active\ hours}$	0.0358***		0.0294***	
	(0.0072)		(0.0048)	
$Listing^{other\ loans}$		0.0337***		0.0277***
		(0.0068)		(0.0045)
Observations	304,305	304,305	448,276	448,276
R-squared	0.0282	0.0282	0.0589	0.0589
	Panel B. Type 2 mispricing			
Last Success	-0.0358***	-0.0358***	-0.1204***	-0.1204***
	(0.0032)	(0.0032)	(0.0029)	(0.0029)
Last Sold Price	-0.0010***	-0.0010***	-0.0012***	-0.0012***
	(0.0001)	(0.0001)	(0.0000)	(0.0000)
Outstanding Principal	0.1085***	0.1085***	0.0894***	0.0894***
	(0.0114)	(0.0114)	(0.0072)	(0.0072)
Loan Age	-0.0985***	-0.0985***	0.0092***	0.0092***
	(0.0027)	(0.0027)	(0.0021)	(0.0021)
$Listing^{active\ hours}$	0.1065***		0.0130***	
	(0.0039)		(0.0029)	
$Listing^{other\ loans}$		0.1005***		0.0122***
		(0.0036)		(0.0028)
Observations	304,305	304,305	448,276	448,276
R-squared	0.1046	0.1046	0.0800	0.0800
Month FE	YES	YES	YES	YES
Loan FE	YES	YES	YES	YES

This table presents results for Model (3) for sub-samples of loans with high credit ratings (Columns (1)-(2)) and loans with low credit ratings (Columns (3)-(4)). The dependent variables are *Type1Mispricing* and *Type2Mispricing* (Panels A and B, respectively). These variables are defined as the ratios of the number of type 1 (or type 2) error to total number of listings. Type 1 error is the loan that has predicted success probability of ≤ 25% but has realised outcome of being sold. Type 2 error is the loan that has predicted success probability of ≥ 75% but has realised outcome of being failed. *Last Success* is the share of the number of successful listings over the last 3 months. *Last Sold Price* is the average price (discount, mark-up, or flat rate) of the successful listings for a given loan over the last 3 months. *Loan Age* is the logarithm of 1 plus the age of loans (in months). *Outstanding Principal* is the ratio of the outstanding principal in the given month to the original principal. *Other loans* is the natural log of the number of other P2P loans of which notes are listed in the secondary market. $Listing^{active\ hours}$ is the share of noted listed during the arguably more active trading hours (days) i.e. between 8am – midnight on Monday – Friday. In all estimations, a constant term is included but not reported. Robust standard errors are in parentheses. *, **, and *** indicate significant level at 10%, 5%, and 1%, respectively.

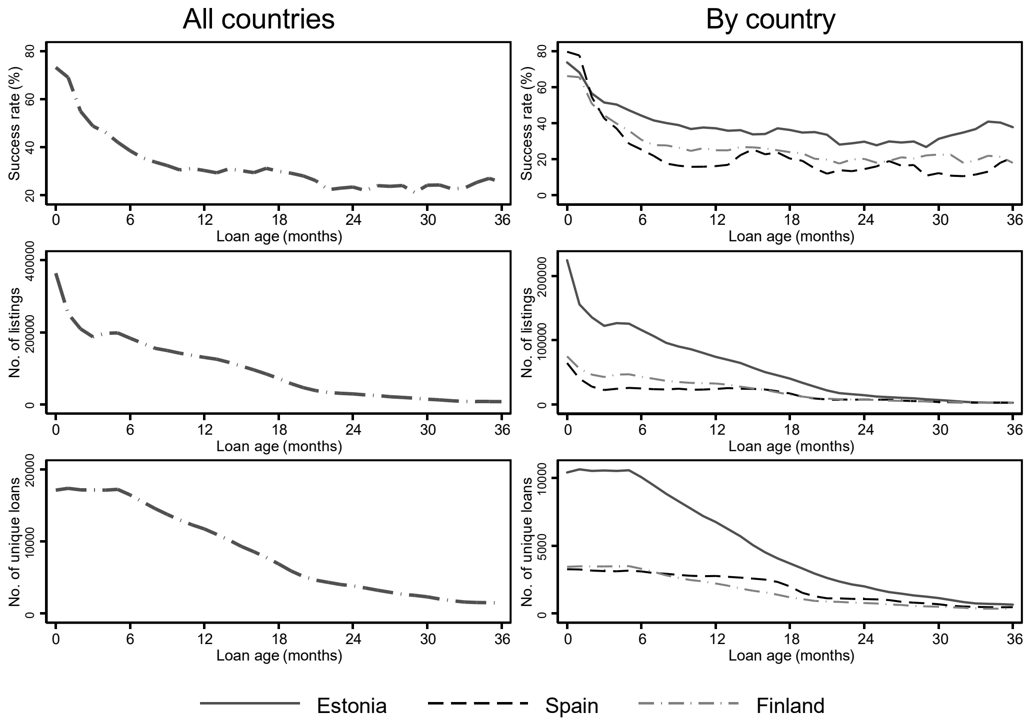
Figures

Figure 1. Statistics by month



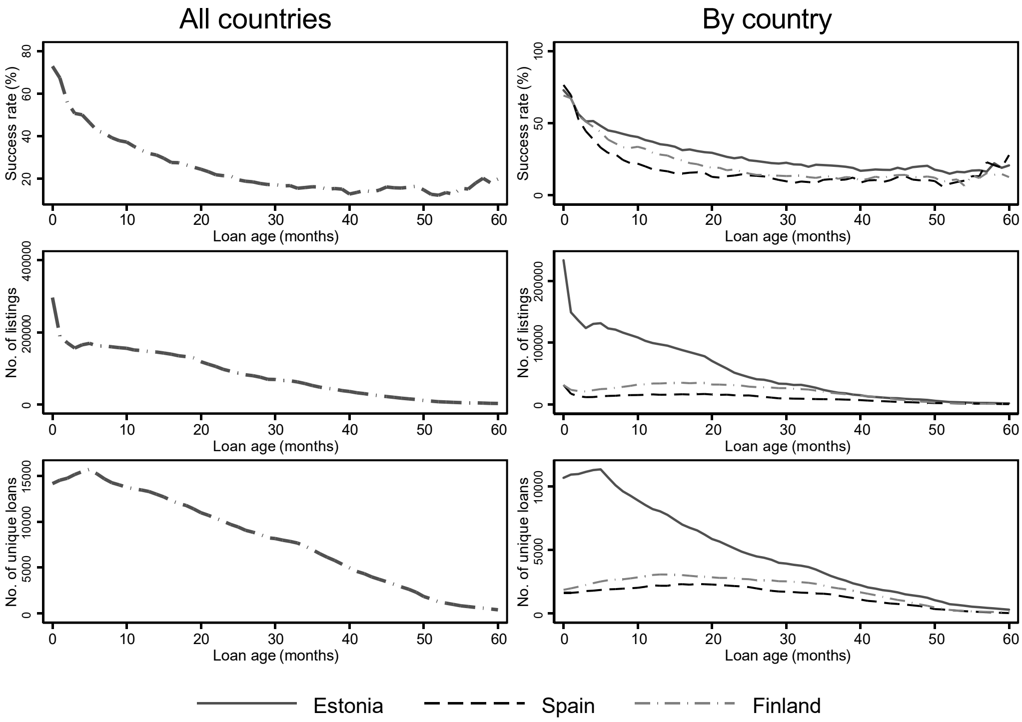
This figure shows the average success rate, number of secondary market listings, and number of unique loans listed in the secondary market over the examined period (the top, middle, and bottom panels, respectively). The left panels present statistics for the full sample. The right panels present statistics by country. The solid line, dashed line, and dotted-dashed line represent Estonia, Spain, and Finland, respectively.

Figure 2. Statistics by loan age, 36-month loans



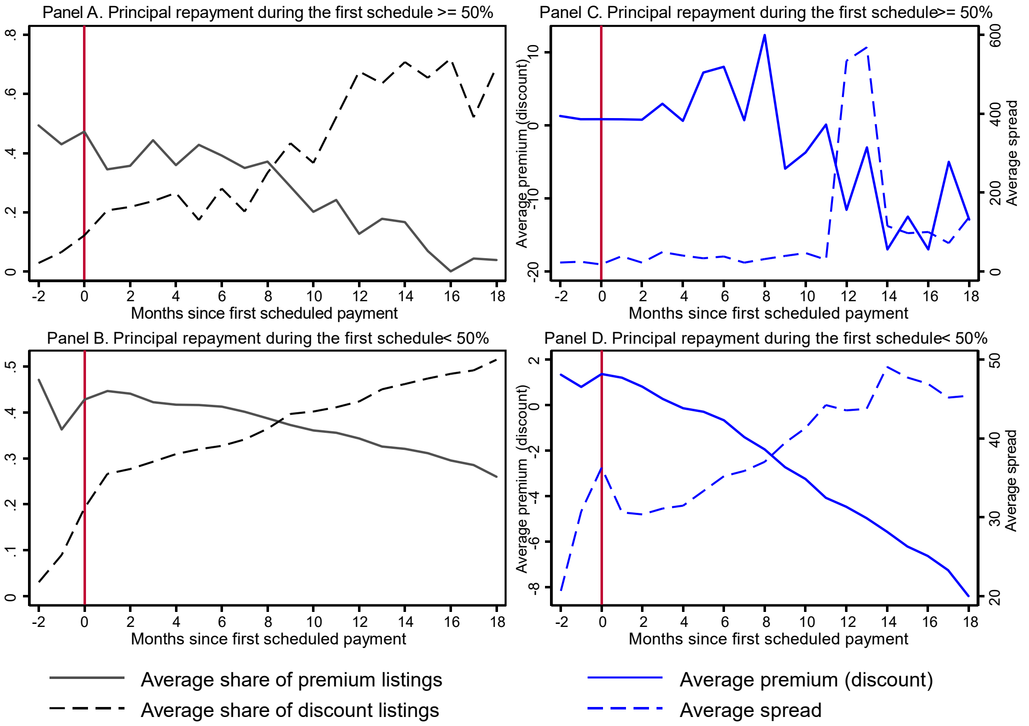
This figure shows the average success rate, number of secondary market listings, and number of unique loans listed in the secondary market over the lifetime of 36-month loans (the top, middle, and bottom panels, respectively), conditional on being listed in the secondary market in at least 6 months. The left panels present statistics for the full sample. The right panels present statistics by country. The solid line, dashed line, and dotted-dashed line represent Estonia, Spain, and Finland, respectively.

Figure 3. Statistics by loan age, 60-month loans



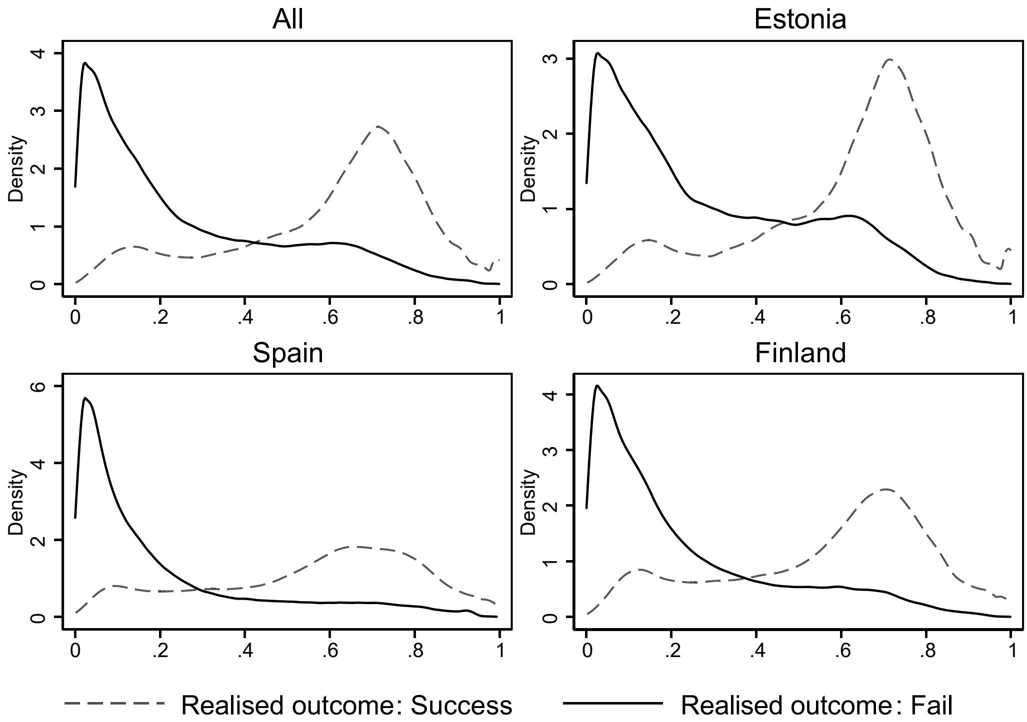
This figure shows the average success rate, number of secondary market listings, and number of unique loans listed in the secondary market over the lifetime of 60-month loans (the top, middle, and bottom panels, respectively), conditional on being listed in the secondary market in at least 6 months. The left panels present statistics for the full sample. The right panels present statistics by country. The solid line, dashed line, and dotted-dashed line represent Estonia, Spain, and Finland, respectively.

Figure 4. Statistics on pricing behaviour, 36 month loans



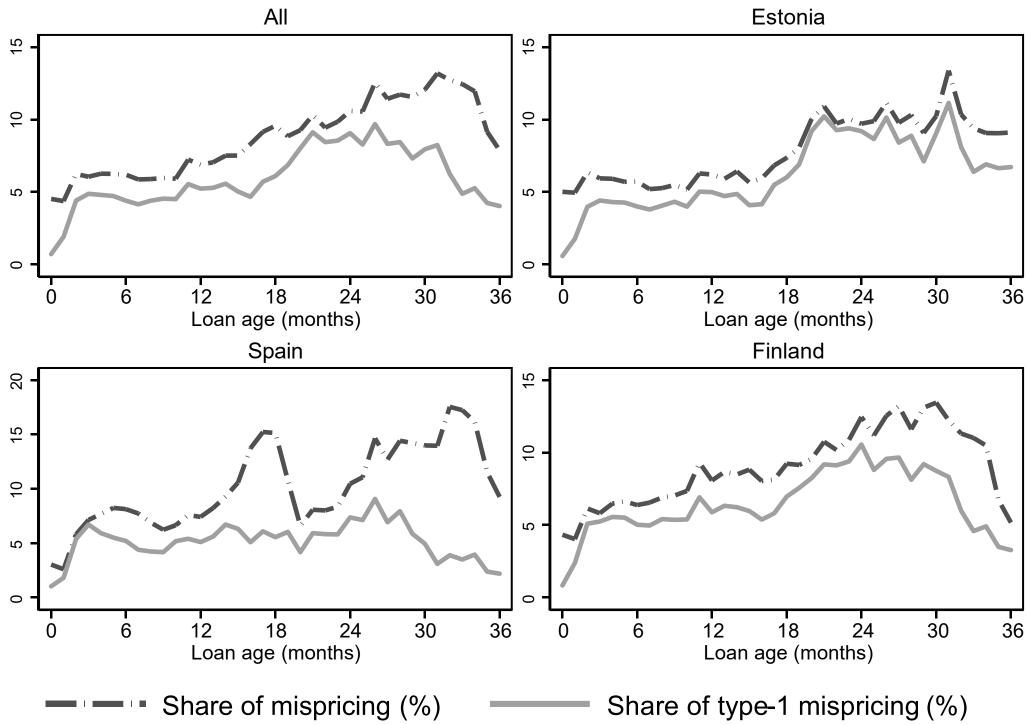
This figure shows the summary statistics for pricing behaviour in the secondary market for 36-month loans. The y-axis represents the gap (in months) relative to the month of the first scheduled repayment. The left panels show statistics for the average share of listings with premium and listings at discount. The right panels show statistics for the average of loan median premium (discount) rates and the average spread (the difference between the maximum and minimum rate).

Figure 5. Distribution of predicted probability of success



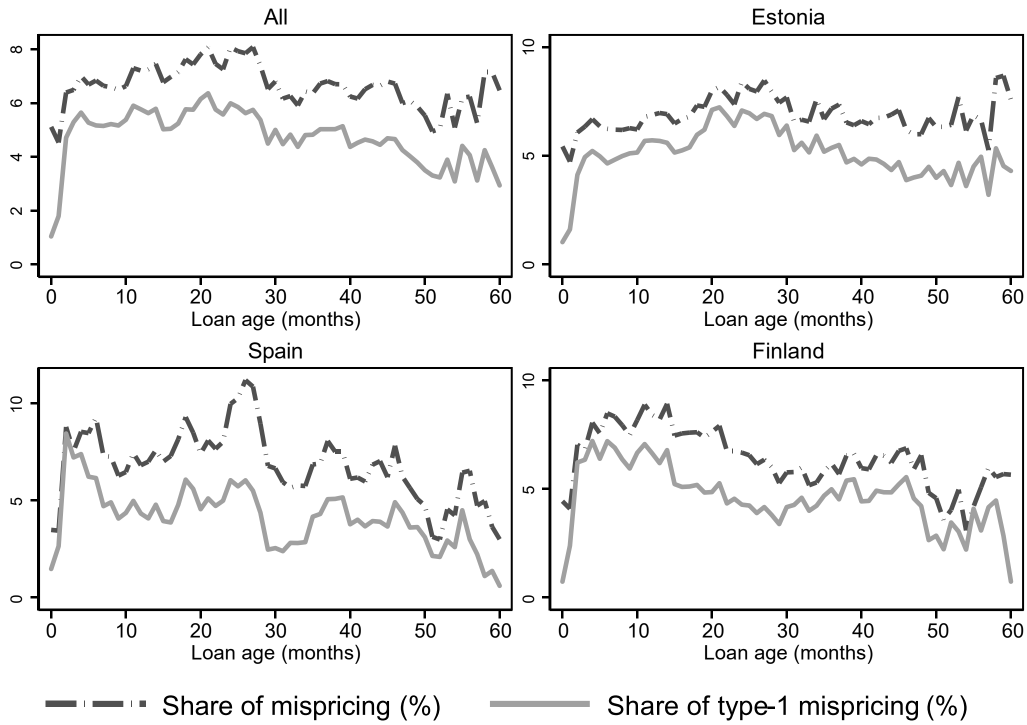
This figure shows the distribution of the within-sample predicted probability obtained from estimating Model (1). The solid line and dashed line represent listings with success and fail as the realised outcome, respectively.

Figure 6. Share of mispricing by loan age (36-month loans)



This figure shows the evolution of the share of mispricing and type 1 mispricing (dashed line and solid line, respectively) over the lifetime of 36-month loans, conditional on being listed in the secondary market in at least 6 months.

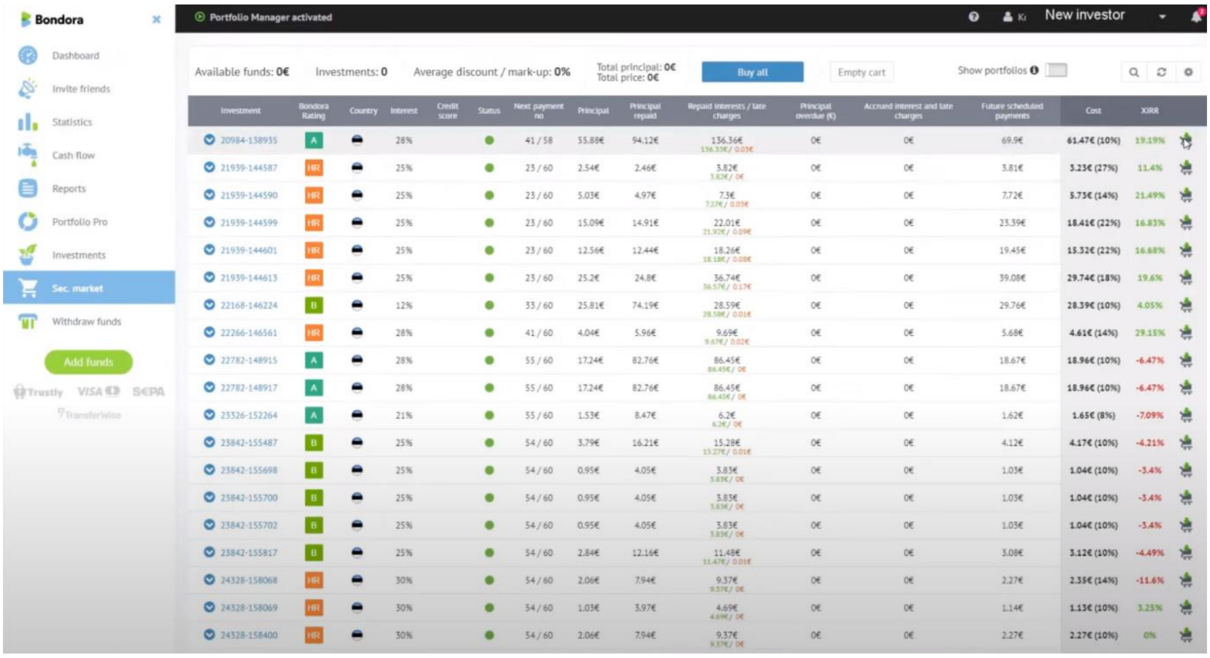
Figure 7. Share of mispricing by loan age (60-month loans)



This figure shows the evolution of the share of mispricing and type 1 mispricing (dashed line and solid line, respectively) over the lifetime of 60-month loans, conditional on being listed in the secondary market in at least 6 months.

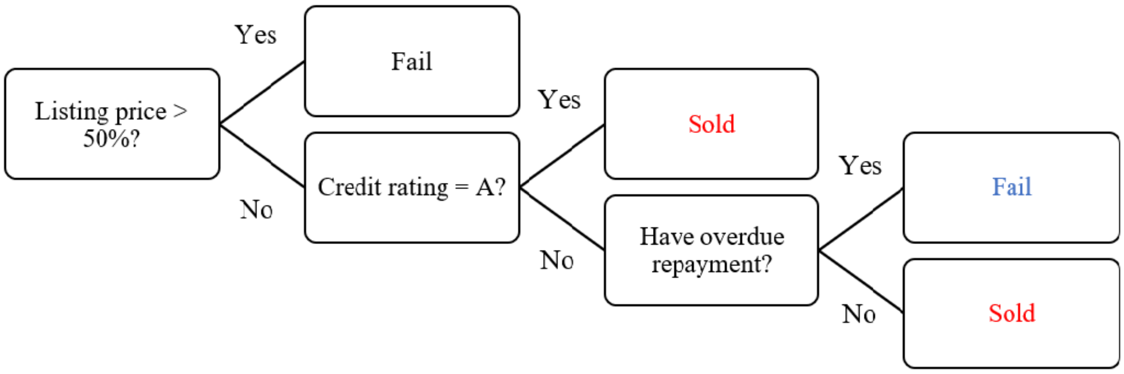
Appendix

Appendix Figure 1. Bondora secondary market interface



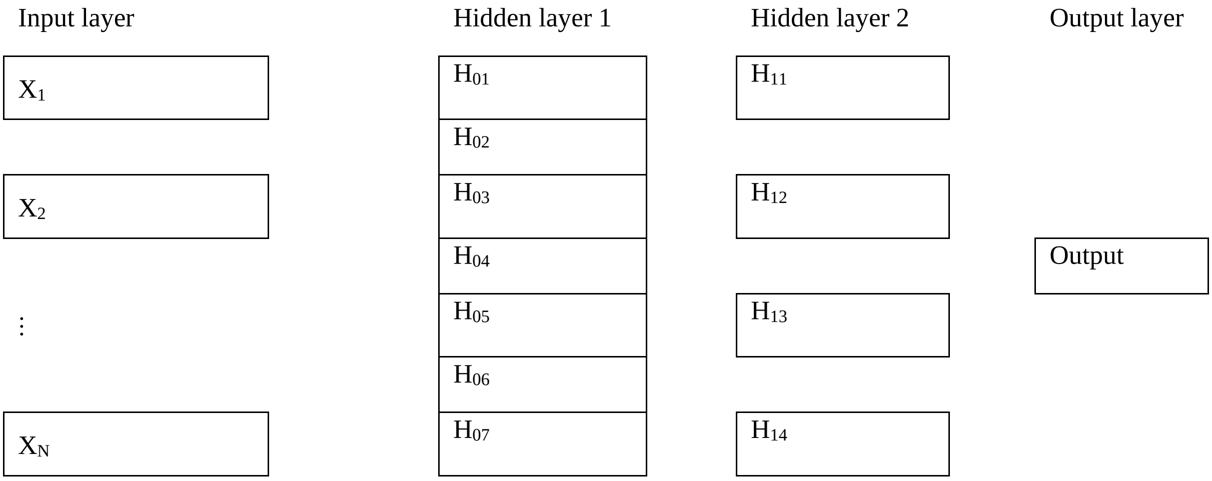
This figure shows the snapshot of Bondora secondary market interface (Source: https://support.bondora.com/en/what-is-the-secondary-market).

Appendix Figure 2. Illustration of a decision tree



This figure illustrates a conceptual example of a decision tree in the context of this study.

Appendix Figure 3. Illustration of a neural network with 2 hidden layers



This figure illustrates a conceptual neural network with 2 hidden layers.

Appendix Table 1. Variable description

Variable	**Description**	**Note**
Panel A. Listing level analysis		
Dependent variables		
Sold	1 if a listing in secondary market is sold, 0 otherwise	
Independent variables		
Listing characteristics		
Listing price	Price (premium/flat/discount) of a listing requested by the seller	
Late payment	1 if the loan has a late interest payment by the time its note is listed in the secondary market, 0 otherwise	
Principal	Natural log of one plus the listing's outstanding principal on the day it is listed on the secondary market	
Borrower characteristics		
Age	Natural log of the borrower's age	
Gender	0 male, 1 female, 2 undefined	
Existing debt	1 if the borrower has other debts, 0 otherwise	
Marital status	Categorical variable indicates the borrower's marital status	See Appendix Table 2 for detailed categories
Education	Categorical variable indicates the borrower's education level	
Employment	Categorical variable indicates the borrower's employment status	
Occupation	Categorical variable indicates the borrower's occupation	
Home ownership	Categorical variable indicates the borrower's home ownership status	
Income	Natural log of total income	
Loan characteristics		
Interest rate	Interest rate charged on loan on the primary market	
Use of loan	Categorical variable indicates the loan's purpose	See Appendix Table 2 for detailed categories
Credit rating	Categorical variable indicates the loan's credit rating assigned by Bondora	
Panel B. Loan level analysis		
Dependent variables		
No Listings	Natural log of one plus the number of secondary market listings of a loan in a given month normalized by its number of listing days in that month.	
Price	Average listing price of a loan's secondary market listings in a given month	
Type 1	Natural log of one plus the number of type 1 error incidents for a P2P loan in a given month	
Type 2	Natural log of one plus the number of type 2 error incidents for a P2P loan in a given month	
Independent variables		
Last Success	Share of the number of successful secondary market listings of the same P2P loan over the last 3 months	
Last Sold Price	Average price of the successful secondary market listings of the same P2P loan over the last 3 months	
Loan Age	Natural log of one plus the age of the loan (in months)	

Outstanding Principal	Ratio of the outstanding principal of the loan in the given month to the original principal	
$Listing^{active\ hours}$	Natural log of the number of notes of other P2P loans listed in the secondary market in a given month	
$Listing^{other\ loans}$	Natural log of the number of notes of all loans listed on the secondary market between 8am – midnight on Monday – Friday in a given month	

This table describes all variables used in the analysis at listing level (Panel A) and at loan level (Panel B). All variables in Panel A are raw data collected from Bondora.com while variables in Panel B are authors' calculations using data collected from Bondora.com.

Appendix Table 2. Loan composition

	Type	No. of loans	%
No. of existing debts	0	8,269	12.69
	1	10,778	16.53
	2	10,040	15.40
	3	8,203	12.58
	4	6,689	10.26
	≥ 5	21,204	32.53
Homeownership	Homeless	3	0.00
	Owner	24,169	37.08
	Living with parents	9,967	15.29
	Tenant, pre-furnished property	12,399	19.02
	Tenant, unfurnished property	3,370	5.17
	Council house	846	1.30
	Joint tenant	1,131	1.74
	Joint ownership	2,348	3.60
	Mortgage	7,928	12.16
	Owner with encumbrance	536	0.82
	Other	2,483	3.81
Occupation	Other	5,614	8.61
	Mining	90	0.14
	Processing	2,069	3.17
	Energy	395	0.61
	Utilities	204	0.31
	Construction	2,247	3.45
	Retail and wholesale	2,414	3.70
	Transport and warehousing	1,660	2.55
	Hospitality and catering	1,581	2.43
	Info and telecom	1,130	1.73
	Finance and insurance	678	1.04
	Real estate	359	0.55
	Research	389	0.60
	Administrative	573	0.88
	Civil service & military	1,039	1.59
	Education	885	1.36
	Healthcare and social help	1,727	2.65
	Art and entertainment	387	0.59
	Agriculture, forestry and fishing	686	1.05
	Not specified - Business loans	41,021	62.93
Employment status	Unemployed	18	0.03
	Partially employed	748	1.15
	Fully employed	19,858	30.46
	Self-employed	873	1.34
	Entrepreneur	1,349	2.07
	Retiree	1,308	2.01
	Not specified - Business loans	41,021	62.93
Marital status	Married	7,431	11.40
	Cohabitants	6,014	9.23
	Single	7,982	12.25
	Divorced	2,340	3.59
	Widow	395	0.61
	Not specified - Business loans	41,021	62.93
Education	Primary education	4,010	6.15
	Basic education	4,538	6.96

	Vocational education	13,788	21.15
	Secondary education	25,500	39.12
	Higher education	17,343	26.61
Gender	Male	46,969	72.06
	Female	15,708	24.10
Credit rating	A	3,438	5.27
	AA	1,956	3.00
	B	8,639	13.25
	C	11,143	17.09
	D	10,575	16.22
	E	9,913	15.21
	F	9,214	14.14
	HR	10,271	15.76
Use of loan	Loan consolidation	4,562	7.00
	Real estate	665	1.02
	Home improvement	6,468	9.92
	Business	1,352	2.07
	Education	893	1.37
	Travel	1,102	1.69
	Vehicle	1,860	2.85
	Other	6,307	9.68
	Health	953	1.46
	Business-related use	41,021	62.93

This table shows the number of loans for each characteristic/category.

Appendix Table 3. Summary statistics – by loan types

	(1)	(2)	(3)	(4)	(5)
	Mean	Min	Max	SD	N
	Panel A. Consumer loans				
Original principal	2,925.937	115.000	10,630.000	2,213.472	24,162
Default	0.511	0.000	1.000	0.500	24,162
Interest rate	34.249	8.080	262.900	24.913	24,162
Maturity	51.687	36.000	60.000	10.647	24,162
No. of listings	227.865	1.000	2,425.000	251.350	24,162
No. of success	69.814	0.000	1,109.000	85.767	24,162
Listing price	-6.002	-100.000	40.000	13.488	24,162
Share of principal	0.011	0.000	1.000	0.042	24,162
	Panel B. Business loans				
Original principal	2,297.017	149.000	10,630.000	2,157.622	41,021
Default	0.471	0.000	1.000	0.499	41,021
Interest rate	38.059	7.430	264.310	34.313	41,021
Maturity	47.671	36.000	60.000	11.433	41,021
No. of listings	116.780	1.000	1,807.000	133.382	41,021
No. of success	57.406	0.000	802.000	59.083	41,021
Listing price	-0.216	-90.000	39.000	5.621	41,021
Share of principal	0.003	0.000	0.231	0.003	41,021

This table presents summary statistics of consumer loans (Panel A) and business loans (Panel B).