Marketplace Lending, Information Aggregation, and Liquidity*

Julian Franks[1], Nicolas Serrano-Velarde[2], and Oren Sussman[3]

March 16, 2020

[1]London Business School

[2]Bocconi University

[3]Saïd Business School, University of Oxford

Abstract

We address a puzzle whereby lending marketplaces, aimed at directly connecting retail lenders and borrowers, retreat from auctions and take on the role of price setting and credit allocation, despite evidence that retail investors possess valuable soft and nonstandard information. Our analysis uses a unique data set on 7,455 auctions and 34 million bids, from the leading British peer-to-business platform. We find that the main problem of the platform was its vulnerability to liquidity shocks, resulting in sizable deviations from information efficiency. These increased over time due to a growing role played by non-crowd players, particularly large investors and algorithms.

*We would like to thank Franklin Allen, Christophe Bisière, Steve Bond, Itai Goldstein, Alexander Guembel, Ulricke Malmendier, Ailsa Roel, Alminas Zaldokas and two anonymous referees for very helpful comments on an earlier draft. We would also like to thank Funding Circle Limited (London) for providing data for this study and, in particular, some of its officers who were extremely generous with their time in answering our many questions.

1 Introduction

The remarkable success of online lending platforms, which directly connect retail lenders with corporate or consumer borrowers, has led many to believe that a radically new model of credit finance is on the rise: decentralized, transparent, disintermediated and more information-efficient. That is, there is the potential for a major disruption of traditional business models, particularly banks, and a redrawing of the boundary between markets and institutions.[1]

Supporting this view is a growing body of evidence showing that much information is dispersed across retail investors and that, if aggregated properly, it can significantly improve the pricing of consumer and corporate loans. For example, Iyer, Khwaja, Luttmer, and Shue (2015) study auction prices from Prosper, a peer-to-peer (P2P) platform, and demonstrate that they predict default better than credit scores do. Hence, they conclude that "our results highlight how aggregating over the views of peers and leveraging non-standard information can enhance lending efficiency" (page 1). It is, therefore, puzzling that many lending platforms, including Prosper, have abandoned their original auction design in favor of posted prices (see Wei and Lin (2016)).[2] If auctions enhance information efficiency, why were they abandoned?

We investigate this puzzle using hand-collected data from the UK's leading peer-to-business (P2B) platform, Funding Circle (FC). The data contain detailed information on 7,455 multi-unit auctions of small and medium-sized enterprise (SME) loans executed online between 2010 and 2015. We observe every order that was submitted to the platform (34 million in total), whether ultimately accepted or rejected. Investor ID numbers allow us to study individual bidding patterns and how they are incorporated into the price-discovery process via an open order book. Like Prosper, in 2015, FC abandoned its auction design in favor of posted prices, and, in September 2017, it further limited investors' choice to platform-selected loan portfolios. FC is a highly successful operator, and, by 2017, its market share exceeded 50% of a P2B industry that was funding around 29.0% of new UK SME loans (see Zhang et al. (2018)).

[1] C.f. Allen and Gale (1995) or Levine and Zevros (1998). For FinTech's disruptive potential, see Philippon (2016), Morse (2015), Yermack (2015), and Goldstein, Jiang, and Karolyi (2019).

[2] A similar trend from unit auctions to posted prices is documented in eBay; see Einav, Farronato, Levin, and Sundaresan (2018).

We suggest a resolution of this puzzle using three interrelated findings. First, in spite of enhanced predictive power, auction prices often deviate, temporarily but significantly, from information efficiency. We provide evidence that the source of these random deviations are liquidity shocks arising from mismatches between flows of funds into and out of the platform, in line with Duffie's (2010) slow-moving capital. Second, to assist a substantial number of investors who lacked the time and expertise to actively participate in the auctions, FC operated an algorithmic "autobid" to which "passive investors" could delegate the submission of their orders. However, our results indicate that FC was not successful in optimally calibrating the autobid, leading to a decrease in information efficiency. Third, the quality of pricing deteriorated over time. We associate this development with a "vanishing crowd" phenomenon, whereby large investors and the autobid came to dominate the auction process.

With respect to the first finding, we estimate an efficient market hypothesis (EMH) equation in which the dependent variable is a credit default dummy, and the explanatory variables include the borrower's credit scores and the loan interest rates, as determined by the auction. We adjust the specification to deal with the truncation of our performance data so that some loans mature out of sample. We find that the interest rate coefficient is strongly significant, indicating that prices enhance default predictability, *over and above* credit scores, similar to Iyer et al.'s (2015) findings. At the same time, we reject key EMH predictions. First, we reject the hypothesis that the interest rate, adjusted for the loss given default, reaches the information-efficient benchmark. Second, we find that the credit scores retain predictive power, contradicting the EMH prediction that *all* of their information should have been absorbed into the interest rate.

In order to identify the source of deviations from information efficiency, we augment the EMH regression with various proxies for liquidity shocks. Under the null EMH hypothesis, these proxies should have no power to predict default because all relevant information should be absorbed into the price. Hence, once the EMH regression controls for the price, any other variable, whether originally containing information or not, should prove insignificant. In fact, we find that liquidity shocks still have predictive power, which can be interpreted as deviations from information efficiency.

To deepen our understanding of the way in which liquidity shocks can drive prices away from fundamental values, we analyze two liquidity events that, although known in advance, still result in considerable mispricing of loans. The first event involves the March 2014 termination of a £20 million UK government program, which contributed a fixed 20% stake in each auction. Consistent with slow-moving capital, the price effect was felt ten days after the event, once existing market liquidity had dried up, and persisted for an extra 20 days, until new capital started flowing into the platform. A second event relates to the predetermined closing hour of auctions. We show that auctions closing outside of the peak hours, 4pm to 7pm, closed at interest rates above the information-efficient benchmark. This is in spite of the fact that an auction's closing time is perfectly anticipated, conditional on the randomly allocated opening time.

The second finding is that FC was not able to optimally calibrate the autobid. Over the sample period, on average, 48% of the funding was allocated via the autobid. This, in itself, should not have been a problem: if investors have no information, it is better to delegate bidding to an algorithm so that it can play a role similar to that of a market maker or an IPO arranger (c.f. Kyle (1985) or Biais and Faugeron-Crouzet (2002)). To perform this function effectively, the autobid has to be calibrated so that it prices in information contained in some funding flows while smoothing out the liquidity shocks. A detailed analysis of the autobid's activity reveals that this was not the case. On the one hand, the autobid insufficiently smoothed time series fluctuations in aggregate supply and demand for funding, thereby allowing liquidity shocks to drive the price away from information efficiency. On the other hand, the autobid overly smoothed cross-sectional variations by channeling excessive funding when low levels of active bidding should have signaled a higher default probability, thereby preventing relevant market information from being adequately incorporated into the interest rate. This result highlights the difficulties that FC designers faced: an information-efficient autobid required detailed knowledge of the joint distribution of funding flows and default probabilities, so as to "signal extract" the information from the funding flows. It is hard to see what the source of such knowledge would be, given the relatively small sample of loans, the small magnitude of the event to be estimated (the annualized default probability of an A-scored loan is just 2.9%) and, the considerable lag with which the default and recovery takes place.

Deviations from price efficiency explain at least 80% of the price variance. Interestingly, the nature and magnitude of the problem are not dissimilar to that found in mature corporate bond markets. A study by Collin-Dufresne, Goldstein and Martin (2001) reports that "regression analysis can only explain about 25 percent of the observed credit spread [monthly] changes." In addition, they find that "the dominant component of changes in monthly credit spreads in the corporate bond market is driven by local supply/demand shocks that are independent of both changes in credit risk and typical measures of liquidity."[3]

Regarding the third finding, we document a sharp drop in the predictive power of the interest rate over the sample period, which we relate to the changing composition of the investor population. We document a "vanishing crowd" phenomenon; that is, through the combined effect of declining active investment and increasing activity of large investors, only 25% of the funding was crowd-allocated by the end of the sample period. To interpret the result, we draw on recent literature on household finance, which suggests that large investors do not increase information efficiency. Using Swedish data, Bach, Calvet and Sodini (2016) find that large investors earn higher returns that can be fully explained by greater risk taking, rather than by "informational advantages or exceptional investment skill, [which] contribute only marginally to the high returns of the wealthy." Similarly, we document that large investors earn an extra 1% return on their loan portfolios relative to small investors, but that a substantial part of that extra return is generated simply by higher risk taking. Integrating the effect into our EMH framework provides evidence that the decline in the quality of the price is correlated with the changes in investor composition.

While sophisticated investors can diversify away deviations from information efficiency and might even profit from the arbitrage opportunities that they create, a non-diversified SME borrower would have to bear the cost of overpricing for the entire duration of the

[3] See, however, Feldhütter and Schaefer (2018) for a more benign view of price efficiency in corporate bond markets.

loan, which, in our sample, has a median length of three years. This observation is consistent with the reasons given by FC for its decision to replace auctions with posted prices: "(i) businesses are put off by a lack of certainty around the cost of their loan, which is important to them; (ii) the price of each loan will now be based on the risk (and term) of the loan, rather than the availability of investor funds; and, (iii) borrowers will know how much their loan will cost before the funding process, attracting more businesses to Funding Circle, which will create more lending opportunities for you."[4]

We believe that our results can provide insights into the development of online marketplace lending. They shed light on common concerns of regulators about liquidity provision and the changing nature of the investor population across lending platforms in the US and Europe. In 2016, a US Treasury white paper on online marketplace lending specifically addresses their potential liquidity risks and notes that "ongoing research will be necessary to monitor the liquidity of online marketplace lending and its impact on credit markets." Across lending marketplaces, the issue of liquidity is intertwined with the phenomenon of "vanishing crowds" documented in our paper. Warren Mead, head of financial technology at KPMG stated that "this is not just finance 'by the people, to the people' any more. Increasingly it's about big business now."[5] Our paper, therefore, connects these global trends with the design of online marketplaces.

While our analysis explores the comparative advantages and disadvantages of the P2B innovation relative to traditional models of credit intermediation, we believe that it is too early to conclude that P2B lenders will displace traditional commercial banks. In their favor, banks bridge the liquidity shocks that affect P2B pricing, but, in doing so, they have to engage in liquidity transformation. However, liquidity transformation is accompanied by credit risk and costly regulation, including capital requirements. In contrast, the instant clearing of investors' funds against accepted bids avoids FC's exposure to credit risk – i.e., "securitization at origination," which virtually eliminates any burden of regulation.[6] Substantial improvements and refinements of the new model are required before

[4]See https://www.fundingcircle.com/uk/fixedrate/, upon abandoning the auction design.

[5]Similarly, a report by Morgan Stanley highlights that the term peer-to-peer "is a misnomer" and that "conventional wisdom" underestimates the importance of institutional funding.

[6]It is noteworthy, however, that some of P2Bs' light regulation is based on the assumption that they carry no systemic risk, an assumption that has yet to be tested under the condition of an economic downturn.

its disruptive potential can be fully assessed.

Additional literature

The first strand of the literature that is relevant to our work highlights the diversity in current platform designs, typical of an innovative industry in the early stage of development. Like FC's platform, these designs experiment with alternative ways to combine both algorithms and human investors. Vallee and Zeng (2018) explore, theoretically and empirically, the connection between the information provision on P2P platforms and rents extracted by sophisticated investors. Their study is motivated by the decision of Lending Club, a P2P platform, to remove half of the 100 variables on borrower characteristics that it previously provided to its investors. D'Acunto, Prabhala, and Rossi (2018) study the implications of robo-advising for the portfolio choices and performance of investors in the Indian stock exchange. They document that the adoption of the delegated investment mechanism has heterogeneous effects across investors, with benefits decreasing in the amount of portfolio diversification. Grennan and Michaely (2019) study the operations of FinTechs that aggregate and synthesize public data. They find a reduction in the quality of information produced by online financial analysis and, as a result, a deterioration in information efficiency. Finally, several studies show how the design of peer-to-peer marketplaces affects the matching between borrowing households and contract terms (Hertzberg, Liberman, and Paravisini, 2017; Cespedes, 2017; Liskovich and Shaton, 2017).

The second strand of literature documents the significant amount of "soft and non-standard" information dispersed among investors about borrower default prediction. Lin, Prabhala and Viswanathan (2013) demonstrate that, due to their lower default probabilities, borrowers with friends are more likely to get their credit application approved and on better terms. Lin, and Viswanathan (2016) explore the effect of physical proximity – home bias – on the propensity to lend. Butler, Cornaggia and Gurun (2017) explore the relationship between the availability of local bank credit and the propensity to borrow from a P2P network. Ravina (2019) finds that "beautiful borrowers are 11.7% more likely to get a loan, pay similar interest rates as average looking borrowers with the same credentials, but default more often."

The rest of the paper is organized as follows: Section 2 describes the data, the auction design and the price-discovery process. Section 3 resolves econometric issues related to sample truncation problems and estimates the default probabilities and loss given default (LGD) rates. Section 4 executes the EMH analysis, and Section 5 analyzes the vanishing crowd phenomenon among investors. Section 6 provides some robustness tests, and Section 7 concludes.

2 The data, the platform and the auction

In this section, we provide a description of our data, the design of the FC auction and the price-discovery process. We also place the auction design in the context of the theoretical literature.

2.1 The data

Our data include *all* loans, 7,455 in total, auctioned between 2010Q4 and 2015Q1. During that period, the aggregate value of FC's loan book grew at a *weekly* rate of 2.4%, albeit erratically, with a standard deviation of 1.2%. Loan maturity in our sample is between six months and five years, with a median of three years (see Table 1). Loans are amortized in equal monthly payments. We track these repayments to the end of 2016. While we lack full performance data for 42% of the loans in our sample, we do have at least one-and-a-half years of performance data on each loan. Section 3 describes the method that we use in order to resolve the sample truncation problems that arise in the estimation of default probabilities and LGD.

[Table 1 About Here]

Loan size varies from £5,000 to £0.52 million, with a median of £50,000. According to the borrower's own report, the main purpose of the loan is to fund working capital, growth or new investments. The vast majority of borrowers are organized as limited companies. They come from all regions of the UK and from all sectors of the economy. The dataset contains no information about borrowers' bank relationships at the time of the auction. However, since the median SME borrower in our data is nine-years-old, it is safe to conclude that he did have substantial banking relationships at the time. Baeck, Collins and Zhang (2014) report survey results regarding borrowers' perception of the advantages offered by P2B funding: more willingness to take risk; higher speed of approval; and greater access to funds, with a marginally lower cost of borrowing. It is widely believed, however, that banks have failed to provide adequate funding to the SME sector (see the Breedon Review (2012)). In response, the UK government decided to support the development of alternative models of financial intermediation. In March 2013, the Department of Business Innovation and Skills (BIS) created a £20 million fund to be allocated via the FC platform.[7] This experiment provides an event study opportunity, which is described in Section 4.3 below.

A borrower's loan application is first evaluated by FC's own credit department. Some borrowers are rejected immediately due to suspected fraud or an unacceptably high default probability. Those that are approved are assigned a credit score, A+ for the lowest risk and D for the highest. The analysis is based on hard information, including the borrower's Experian credit rating, its credit history and its financial statements. The credit-scoring process allows for a certain level of discretion by FC's analysts. Borrowers are asked to provide a "prospectus," which is made public on the platform's website. While the auction is running, the platform usually opens a ledger where investors can post questions or ask for additional information. Borrowers are encouraged to respond fully to questions.

Orders submitted to the platform must be backed by funds that are transferred in advance to a special FC account. If a bid is accepted, that account is immediately debited, while the purchased loan parts are simultaneously transferred to the investor's custodian account. Such instantaneous settlement implies that at no stage in the process does FC have ownership of any part of the loan, and, hence, there was no need to allocate any (regulatory) capital against credit risk. The process might be described as securitization at the point of origination.

Once the loan is issued, FC collects the monthly repayments and distributes them among the investors who funded the loan. In the event of default, FC acts as a Diamond

[7]Buchack et. al. (2018) show that in the residential mortgage market, both technology and regulation explain the penetration of FinTech. Similarly, De Roure et al. (2019) and Tang (2019) focus on consumer credit markets to show that regulatory requirements can account for the retreat of banks, especially in the higher-risk segment of households.

(1984) delegated monitor to manage a resolution, including, if necessary, litigating on behalf of the investors; the mean number of investors per loan is 571, with a median of 459 (see Table 1). For these services, investors are charged a 1% annual fee on the outstanding principal of performing loans, while borrowers are charged a one-time issuance fee that is a function of loan maturity and the borrower's rating.

2.2 Auction design and function

In a multi-unit auction, each investor submits orders for a part of the total amount of the loan. The minimum loan part that the platform accepts is £20. The mean number of loan parts per loan is 805, and the median investor stake, as a proportion of the value of the loan, is 0.48% (see Table 1).

Only limit orders are accepted; each order must specify a quantity and a price. Auctions run continuously, day and night. Typically, auctions are scheduled for seven days (168 hours), but a few auctions are allocated more time. It follows that, conditional on the opening hour, there is perfect foresight of the closing hour, an important fact in the event study analysis of Section 4.3 below. However, borrowers may terminate the auction prematurely, in which case the terms of the loan are determined by the orders submitted at that point.

FC auctions price discriminate, so that each accepted order pays the submitted interest rate. At any time during the auction time, orders submitted up to that point can be sorted by price to form an increasing, stepwise, supply curve. Figure 1 presents such supply curves, for an arbitrary A-scored loan, 24, 96 and 168 hours from the open. Orders are normalized by the size of the loan, in this case £15,000. By construction, given that the size of the loan is fixed in advance, the normalized demand curve is just a vertical line at one unit. When the auction closes, the intersection point between supply and demand, in this case 6.6%, defines the marginal closing interest rate –namely, the highest interest rate of any accepted order. All orders submitted at an interest rate below the marginal rate are accepted, while orders submitted at an interest rate higher than the marginal rate are discarded. Orders tied at the marginal rate are prioritized by submission time. It follows that the "borrowing rate" – i.e., that which is charged to the borrower – is a weighted average of all the accepted orders, in this case 6.49%. By construction, the borrowing rate always lies below the marginal closing rate.

[Figure 1 About Here]

Another important property of the FC design is that orders are not retractable: once submitted, an order cannot be withdrawn. It follows that the downward shift of supply curves over auction time, as depicted in Figure 1, is a necessary consequence of the design of the FC auction. Nonretractablity plays an important role in limiting investors' ability to engage in manipulation (see Section 2.3 below for a detailed discussion).

We defer the detailed analysis of default probabilities to Section 3 below. To derive statistics on the interest rate r, both the borrowing and the marginal closing rates, we run an OLS regression:

$$r_i = \alpha + \theta \times FE\_Score_i + \gamma \times FE\_QIssue_i + \varepsilon_i \tag{1}$$

where FE_Score and FE_QIssue are fixed effects (FEs) for the credit scores and the quarter of issuing loan i. Table 3 shows that the borrowing rate and the marginal rate for an A-scored loan are 8.47% and 8.97%, respectively. Both rates increase by, roughly, 1% from one credit score to the next. In A-scored loans, the difference between the marginal and the borrowing rate is $50bp\,(8.97 - 8.47)$. The annualized default probability for an A-scored loan is 2.9%, while the LGD is 34.2%, as shown in Section 3 below. It follows that, given the Bank of England's base rate of just 0.5%, FC was offering investors what appeared to be a generous return for this particular time series of loans.

Many investors, particularly small ones, may not have the required time or expertise to actively participate in the bidding process. For such investors, FC operates an algorithmic bidding device, the "autobid." On average, about half of the platform's funding is provided by "passive investors" (see Table 1). The autobid selects a diversified portfolio to match the risk profile chosen by the investor. The number of "active" (non autobid) investors per loan remains very high, with a mean of 200. It is worth noting, however, that active investors do not participate equally. The mean share of the largest investor is 8%,while the mean share of the top five and the top 20 investors is 18% and 29%, respectively.

Section 5 provides a detailed analysis of the investors' composition across size categories.

[Table 2 About Here]

While the auction is running, the order book is open for the inspection of all interested parties, with the platform computing and reporting the marginal interest rate, continuously. The order book provides information on the identity of the investor[8] who submitted the order and his bid in terms of the interest rate, the amount, and the time. The platform does not report whether the order is placed directly or via the autobid, although sophisticated investors can infer that information with some degree of confidence.

Table 2 reports mean daily order flows for both the autobid and active investors within 24-hour intervals of auction time – i.e., auction "days." Only auctions that complete the full seven days allocated to them are included. Daily orders are normalized by loan size and, then, averaged across auctions. Active investors are sorted by the total amount of their daily orders into size groups, small, medium and large, $-£100$, $£100 - £1000$ and $£1000+$,respectively. The autobid is highly active in the opening hours, and it is quite common for it to submit orders in excess of the amount of the loan. However, $61\% \, (= 1 - 0.25/0.65)$ of the day one autobid orders (by value) are eventually rejected (see variables *Flow* and *Exec.* Table 2). It follows that the autobid is programmed to submit, on day one, a large number of small orders at different interest rates, whereby higher ones are ultimately rejected, as they end up above the marginal interest rate at the close of day seven. Sorting these day one orders by the submitted interest rate, from the lowest to the highest, we interpret the autobid policy as the submission of an upward-sloping supply curve, as shown in Figure 2. The slope, the intercept and the exact shape of this autobid supply schedule had to be calibrated by the designers of the FC platform. Much of the EMH analysis, in Section 4 below, addresses the question of whether that calibration was done in a manner that maximized information efficiency.

[Figure 2 About Here]

Nothing compels active investors to place orders at the early stage of the auction: they can wait and bid just prior to the close, free riding on information of other investors with-

[8]Typically, investors mask their real identity behind a code name.

out contributing any of their own. Clearly, if all investors were to adopt such a strategy, the process would degenerate into a sealed-bid auction. Table 2 demonstrates that this is not the case: by the end of day one, active investors have submitted orders totaling 34% $(= 1.11 + 0.18 + 0.05 - 1)$ in excess of the size of the loan. As for the autobid, most of these orders are ultimately rejected. The implication is that a substantial proportion of active investors voluntarily participate in the price-discovery process. Their reason for doing so is beyond the scope of this paper.[9] Suffice it to say that similar results are documented in the literature (see Biais, Hillion and Spatt, (1999) and Bellia, Pellizon, Subrahmanyam, Uno and Yuferova (2017) on participation in pre-open bidding on the Paris Bourse).

However, not all investors adopt such a cooperative bidding strategy. Variable *New* in Table 2 reports the new component of the order flow – namely, orders submitted by an investor who did not participate in the previous days of the same auction. It follows that 49% $(= 0.23/0.47)$ of the orders placed by large investors on the last day are new in that respect. The numbers for medium and small investors are 61% and 79%, respectively. Hence, the pattern of entry into the auction is U-shaped in auction time for all size groups: strong at the open, weakening in days two to six, gaining strength again towards the close.

Given the open order book, the investors who participate in the price-discovery process face a simple decision: whether or not to undercut the marginal interest rate, at that point in auction time, so as to guarantee that their bid is accepted. If many of them do, the marginal interest rate will fall. As a result, at the close, active investors' accepted orders are clustered at or just below the marginal interest rate; the mean difference between accepted active bids and the marginal closing rate is $37bp$ (see Table 1). In contrast, passive investors' orders are spread out below the marginal closing interest rate, with a mean difference of $101bp$. For a diagrammatic representation of the borrowing rate, see the shaded area below the autobid's supply schedule and the marginal rate in Figure 2.

2.3 Discussion: FC's auction design and the EMH

Unfortunately, except for the largest platforms, the academic literature does not provide

[9]See, however, Admati and Perry (1990) on contributions to a joint project when commitments and enforceable contracts are not available.

a comprehensive survey of the auction mechanisms used in online lending marketplaces. However, a leading FinTech blogger ("p2pmoney") constructed a classification of UK and European online marketplaces and found that out of 36 platforms, 14 employed non-uniform pricing in their auctions. In addition, price discrimination is a prominent mechanism in other financial auctions, notably the sale of Treasury Bills. Brener, Galai and Sade (2009) survey 48 OECD countries and find that 24 use price-discriminatory auctions; nine use uniform-price auctions; nine use both; and six use other designs. Interestingly, empirical research fails to find significant revenue differences between discriminatory- and uniform-price auctions. Using occasions in which US Treasury bills were sold simultaneously, via both uniform and discriminatory auctions, Nyborg and Sunderesen (1996) find only a modest difference in outcomes. Hortaçsu and McAdams (2010) take a structural approach, estimating bidding strategies in Turkish treasury auctions and then using the results to simulate the effect of a regime change. They conclude that "by our point estimate, the switch from a discriminatory to a uniform price auction would lead to a gain of expected revenue that is at most 0.12 percent of the realized revenue. ... However, taking sampling variation into account, we cannot reject the hypothesis that such a switch would lead to no difference in revenue."

Quantity discrimination is standard in initial offerings of shares for public listing (IPO), which Biais and Faugeron-Crouzet (2002) analyze as an auction derived from a mechanism design problem. Their solution sets the price at a "significant discount relative to the market clearing price" while rationing particular investors, so that the price does "not adjust to demand too strongly." Benveniste and Wilhelm (1990) model a closely related problem and derive a solution that price discriminates. One interpretation is that price and quantity discrimination are close substitutes, used in order to incentivize revelation of information but, at the same time, to guarantee a fair return to the uninformed so as to elicit their participation. In that respect, the FC design can be rationalized: while price discrimination delivers higher returns to active (and better informed) investors, the autobid caps these returns so as to protect the uninformed.

Price discrimination can also be used to avoid collusion in Walrasian auctions. Adapting a well-known argument by Wilson (1979) and Back and Zender (1993) to our context, investors can collude by submitting steep supply curves that clear the market above the competitive interest rate. Collusion is sustained because any unilateral deviation from that strategy results in a sharp fall in the interest rate, to the detriment of all investors, including those who have deviated, placing such behavior off the equilibrium path. Price discrimination breaks such collusion by imposing a heavy cost on the submission of steep supply curves: since each order that forms the supply curve is executed at a different interest rate, those orders that are executed at a low interest rate impose a heavy cost on the strategy.[10]

Collusion is closely related to the problem of price manipulation. Chakraborty and Yilmaz (2004) analyze the problem using a market micro structure setting. They describe manipulation as follows: "An insider who knows that the prospects of a certain asset are not good, might actually start buying the asset in order to drive its price up and then sell it without its price falling too fast." Clearly, such "pump and dump" strategies require that investors can place both buy and sell orders. It seems that FC was aware of the problem and blocked such manipulation by making orders nonretractable. Thus, an investor cannot deter other investors from participation in an auction by lending heavily early on, thereby depressing the interest rate, but then hiking it up by retracting the early orders when the auction approaches closure. In other settings, similar strategies are implemented through short selling. Goldstein and Guembel (2008) provide a rigorous analysis of manipulation strategies and conclude that "manipulation is profitable only via sell orders ... [which] is implicitly understood to be relevant, for example, by regulatory bodies [and platform designers] ... who introduced restrictions on short sales."

Market micro structure, as in Kyle (1985) or Glosten and Milgrom (1985), differs from the mechanism design approach in several respects. Notwithstanding the differences, there are two important commonalities. First, to operate effectively, markets need intermediaries, such as IPO arrangers or market makers. Interestingly, Biais and Faugeron-Crouzet

[10]Rock's (1986) seminal results are worth mentioning here. Consider a posted price offering, determined according tothe arranger's own information –i.e., its credit scores. That price is fixed, unaffected by investors' demand for the offering, whether strong or weak. The allocation of the offering is executed either pro rata or on a first-come, first-served basis. The likely outcome is that informed investors will receive the securities with high expected returns, leaving the low return securities to the uninformed - a winner's curse. In this light, it is easy to understand FC's initial decision to try an auction design. However, it is also easy to see why the posted-price system was quickly abandoned, driving FC to the position of an asset manager.

(2002) recognize that the role of such intermediaries could be automated, although they admit that it may prove difficult to "translate into [the] explicit computer algorithms the rather implicit rules" that human specialists have developed over many years of experience. A second commonality between the two approaches is that prices either reach or get close to the EMH price. Since matching the FC design with a structural model goes well beyond the scope of this paper, the EMH provides us with a convenient reduced-form conceptual framework to analyze the data. We work out the precise specification in Section 4.1 below.

The main role of the (algorithmic) market maker is to provide liquidity, which requires that it has two attributes. First, the autobid must hold sufficiently large reserves, which it can inject when gaps arise between the flow of funds into and out of the platform. Second, the market maker must be able to separate random funding gaps from informative supply fluctuations. While the former requires the injection of more liquidity, the latter requires that market conditions are passed through into the price. That is, the autobid needs to be calibrated so as to smooth out liquidity shocks but allow flows that contain information to influence the price. However, to do so, autobid designers need to know the joint distribution of order flows and default probabilities in order to signal extract any information in those flows.

Failing to have sufficient liquid reserves results in deviations from EMH pricing until new capital flows into the platform, an effect that Duffie (2010) calls slow-moving capital.[11] Such temporary deviations from information efficiency are even more likely to occur in a rapidly growing platform such as FC's, which had to expand simultaneously into both the investor and the borrower populations. Perfect synchronization between these two processes is highly unlikely. However, we interpret slow-moving capital somewhat more broadly: not only is the platform growing rapidly, but the composition and the characteristics of the investor population were also changing fast. It is hard to see how prior knowledge of the joint distribution of order flows and default probabilities could be obtained in such circumstances.

Finally, large, wealthy and sophisticated investors could also provide liquidity to the system, side by side with the market maker. Indeed, in theory, it is difficult to make a

[11]See Gromb and Vayanos (2018), and Shleifer (1986).

clear distinction between the two, a point made in Kyle (1989). It follows, however, the these large investors are likely to face problems similar to those of the market maker–that is slow-moving capital and imperfect knowledge of the joint distribution of order flows and default probabilities.

3 Default probabilities and LGD

The fact that our loan-performance data stop at the end of 2016 raises a truncation problem: a loan that does not default within the sample might still default subsequently, when we no longer observe it. Since FC loans are amortized in equal monthly payments, the timing of default plays an important role in our analysis, as it determines the pre-default recovery rate. A similar problem arises for post-default recoveries: the shorter the time we observe a loan post-default, the less likely we are to observe substantial recoveries. After correcting for these problems, we combine pre-default and post-default recoveries into the LGD rates, conditional on the credit scores. These corrections play an important role in adjusting the interest rates in Section 4's EMH analysis.

To resolve the truncation problem in the estimation of the probability and the timing of default, we use a stacked regression methodology; see Cameron and Trivedi (2005) and Soyeshi (1995) for a more comprehensive discussion of the methodology and its advantages. The idea is simple: rather than estimating the default probability per loan, we estimate the *periodic* default probability, which we condition on the "life cycle" of the loan. Hence, we use $81,049$ loan-performance (three months) quarters for the $7,455$ loans in our sample. Loans drop out of the panel either when they mature or one quarter after defaulting; note that performance data for the entire sample end in late 2016. We estimate the transition probability from performance to non-performance using OLS, with the default indicator as the dependent variable. In Section 6, we report the robustness of our results relative to a Logit/Probit specification of the stacked regression, as well as to the Cox Proportional Hazard duration model.

More accurately, the regression specification is:

$$Default_{it} = \theta \times FE\_Score_i + \gamma \times FE\_QIssue_i + \lambda \times FE\_SLife_{it} + \varepsilon_{it}. \tag{2}$$

$Default$ is the performance indicator set equal to one if loan i defaults during performance quarter t, and zero otherwise. FE_Score_i and FE_QIssue_i are credit score and period of issue FEs as in equation (1) above. FE_SLife_{it} are FEs for the four stages in the loan's life cycle. That is, we partition the loan's "life" into four equal intervals, indicated by the Table 3 variables *Early*, *Mid-Early*, *Mid-Late*, and *Late*. For example, the *Early* dummy, receives a value of one if performance quarter t of loan i falls in the first quartile of the loan's life, and zero otherwise.[12] For ease of interpretation, we include all four life-cycle dummies in the regression and omit the constant (to avoid perfect multicollinearity).

Results are presented in Column (3) of Table 3. To derive the annualized unconditional default probability for an A-scored loan, we add up the quarterly default probabilities across the four stages of the loan's life cycle: $0.007 + 0.01 + 0.008 + 0.004$, equal to 2.9%. To find the annualized probability of default for an AA-scored loan, wesubtract $1.6\%(= 4 \times 0.04)$. The annualized unconditional default probabilities across credit scores are presented in Column (1) of Table 4.

[Table 4 About Here]

Applying Bayes' Law to the unconditional probabilities we derive the *conditional* probabilities of default given the stage in the loan's life cycle. That is, for an A-scored loan, we normalize the vector of unconditional life-cycle probabilities, $(0.007, 0.01, 0.008, 0.004)$, by the overall probability of default, 2.9%, which yields the life-cycle pattern of default probabilities, $(0.23, 0.35, 0.28, 0.14)$. That is, conditional on default, there is a probability of 23% that the loan will default during the early stage of its life cycle and 35% that it will default during the mid-early stage, etc.

[Figure 3 About Here]

The conditional default patterns for all credit scores are plotted in Figure 3. Given that FC loans are amortized in equal monthly payments, these life cycle patterns determine pre-default recoveries. That is, an A-scored loan that defaults early has already repaid between zero and 25% of the debt, capital and interest; the corresponding figure for a loan

[12] Consequently, a life cycle stage for a one-year loan is three months long, whereas it is nine months long for a three-year loan. The method is robust to alternative ways of controlling for the impact of time on default probabilities.

that defaulted mid-early is between 25% and 50%, etc. Note that, conditional on default, there is a $42\%(= 0.28 + 0.14)$ probability that the loan will not make it to the mid-point of its life, and so its pre-default recovery is below 50%. For the conditional recovery, prior to default, of an A-scored loan, we multiply the vector of conditional default probabilities (above) by a vector of conditional recovery rates $(0.125, 0.375, 0.625, 0.875)$. The results for all credit scores are reported in column (2) of Table 4. Evidently, in line with the argument above, conditional recovery rates to default are between 40% and 50%.

Post-default, FC manages to achieve significant recoveries. Again, the sample truncation problem arises: the earlier the default in calendar time, the longer is the observed recovery period. Hence, for a loan that defaulted in December 2014, we observe 24 months of recovery efforts by FC (up to the end of 2016), while for a loan defaulting in November 2016, we observe only one month of recovery efforts. The shorter the observed recovery period, the lower is the *reported* (not the actual) recovery rate. To address the problem, we estimate the post-default recovery equation:

$$RRecoveryPost_i = \alpha + \theta \times FE\_Score_i + \gamma \times FE\_QIssue_i + \lambda \times log\left(MRecovery_i\right) + \varepsilon_i, \quad (3)$$

where $RRecoveryPost$ is the post-default recovery rate of the debt outstanding at the point of default, and $MRecovery_i$ is the length of the recovery period, in months, from the point of default to the end of the sample, logged. Results are presented in column (4) of Table 3. As expected, the coefficient of the recovery period variable is positive and highly significant. Recovery rates do not differ significantly across credit scores. Column (3) of Table 4 reports the extrapolated post-default recovery rates by substituting 60 months into the estimated equation (3), capturing the five years during which, in FC's own assessment, recoveries are expected to be made.

The last step is to combine recoveries pre- and post-default in columns (2) and (3) of Table 4, respectively, so as to construct an overall estimate of LGD. Column (2) of Table 4 shows that an A-score loan in default has already repaid 45.7% of the loan by the time of default, and is expected to pay off 37% of the remainder post default. It follows that total recovery, conditional on default, is 65.8%, so that LGD is 34.2% .

FC's recovery rates are remarkably high compared with those of unsecured loans and trade credit for UK SMEs (see Franks and Sussman (2005)). It seems that these can be explained by FC's innovative use of certain properties of English bankruptcy law. Although FC's loans are unsecured, they are usually guaranteed, personally, by the owners of the SME. It follows, that corporate default can lead to personal bankruptcy, which has more-severe consequences in the UK than in the US. First, protection for personal assets, including homes, is nonexistent. Second, many restrictions apply to bankrupt individuals. For example, while in bankruptcy a person cannot "borrow more than £500 without informing the lender, ... act as a director of a company without the court's permission, ... create, manage or promote a company without the court's permission."[13] It is common practice for British banks to freeze the bank accounts of bankrupt individuals or to refuse to open new accounts for them. The UK, unlike the US, has a centralized corporate register that facilitates the tracking of individuals' credit histories.

According to A. Jackson (2016), head of recovery at FC, the platform used the above properties of English law in order to innovate a "survival for revival" recovery strategy. While FC does not agree to haircuts on defaulted loans, it is prepared to show great flexibility in rescheduling loans in default, with a tough position on triggering personal bankruptcy if the borrower fails to meet the deferred payments. As a consequence, FC placed 25% of the owners of defaulting SMEs into personal bankruptcy. Jackson argues that "for Funding Circle, 90-95% of recoveries come through the personal guarantor. ... A conservative estimate of recovery is 40p in the £ over a five-year period post default," remarkably close to our estimates.

4 Market efficiency, mispricing and liquidity

Our EMH analysis provides three main results. First, because FC auctions aggregate some private, possibly nonstandard, information dispersed across the investor population, the interest rate better predicts the loan's probability of default, over and above the credit scores. Second, although such information is priced in, the price still fails the test of information efficiency, as there are sizable deviations of the price from its information-

[13] See www.gov.uk/bankruptcy/restrictions.

efficient benchmark. The problem is exacerbated over time. Third, these deviations from efficient pricing are due to funding shortages and a poorly calibrated autobid.

4.1 The EMH specification

To derive the EMH specification, we augment equation (2) above with an LGD-adjusted interest rate, r^*, a proxy for systematic risk, Sys, and a vector for various proxies for liquidity shocks, Liq:

$$Default_{it} = \beta \times r_i^* + \theta \times FE\_Score_i + \gamma \times FE\_QIssue_i + \lambda \times FE\_SLife_{it} + \delta \times Sys_i + \eta \times Liq_i + \varepsilon_{it}. \tag{4}$$

Information efficiency implies that the augmented interest rate *enhances* default predictability such that a 1% increase in r^* predicts a 1% higher default probability, and that since r^* absorbs *all* the information contained in $Dscore$ and Liq, these variables lose their predictive power.

Consider a risk-neutral competitive environment in which the interest rate, r, is determined by the lender's participation constraint:

$$1 + \rho = (1 - \pi_i^e)(1 + r_i) + \pi_i^e (1 - LGD_i^e)(1 + r_i), \tag{5}$$

where ρ is the risk-free rate; π is a loan's probability of default; and the e superscript is an expectations indicator. Linearizing and rearranging equation (5), we obtain:

$$\pi_i^e \approx \alpha^* + r_i^*, \tag{6}$$

where $\alpha^* = -\frac{\rho}{LGD_i^e}$ and $r_i^* = \frac{r}{LGD_i^e}$. Hence, there is a one-to-one relationship between the adjusted interest rate and the expected default probability. Intuitively, when LGD is 100%, a 1% increase in a loan's expected probability of default implies a 1% increase in the interest rate. However, if LGD is only 50%, a 1% increase in default probability implies only a 0.5% increase in the unadjusted interest rate, r, so that the LGD-adjusted interest rate, r^*, increases by 0.5/0.5, with an expected unit β coefficient in equation (4).

Under the null EMH hypothesis, since the adjusted interest rate, r^*, incorporates *all* relevant information, public as well as private (to the extent that it is revealed in the bidding process), it should not be correlated with equation (4)'s error term. Otherwise, r^* predicts ε, which, in turn, predicts default. Rational investors, including algorithmic ones, should use that information in order to revise their bidding strategies. By doing so, they drive the interest rate towards the information-efficient price, up to the point at which *all* of that information is priced in and the correlation between r^* and ε vanishes. An important implication of this argument is that r^* can be included on the right-hand side of equation (4), though it is determined by the actions of the investors.

By a similar argument, the credit score coefficient, θ, as well as the liquidity shock coefficient, η, should lose their economic and statistical significance.[14] The argument for θ is straightforward: investors should extract *all* the information contained in the publicly observed credit scores, discard any "noise" that the credit scores might have, and price it into the interest rate, so that the credit score is left with no predictive power. Likewise, the regression should reject any predictive power in the liquidity shocks. This is obvious when the shock is related to random funding gaps that contain no "fundamental" information about default probabilities. Even if this is not the case, any information that is included should be absorbed into the price through the price-discovery process, leaving the funding gap with no predictive power.

Finally, to relax the risk-neutrality assumption used in the derivation of equation (5) above, and in order to account for aggregate macro risk, we include in equation (5) the systematic-risk variable, Sys. Since our SMEs are not listed, we have no company-level measures of systematic risk. Instead, we use industry-level betas as reported in Demoderan.[15]

4.2 Baseline results

Baseline regressions are presented in Table 5; columns (1) to (3) use the borrowing rate (i.e., the weighted average of all accepted bids), while columns (4) to (6) use the marginal interest rate at the close (i.e, the highest interest rate of all accepted bids), both adjusted for LGD, as derived in Section 3 above.

[14] See Hilscher and Wilson (2017) for a similar point.

[15] http://pages.stern.nyu.edu/~adamodar/New_Home_Page/datafile/Betas.html

[Table 5 About Here]

Columns (1) and (4) report baseline results, without including any measure of liquidity shocks. They already carry the main message of our analysis: while the evidence does not reject the hypothesis that the market aggregates disperse investors' information, information efficiency is rejected. That is, on the one hand, the 1% statistical significance of the interest rate coefficient implies that its inclusion in the regression enhances default predictability over and above the credit scores; on the other hand, the borrowing rate coefficient is significantly smaller than one, at 0.33. That is, a 1% increase in the borrowing rate predicts only a 33*bp* increase in the default probability. Similarly, the joint F test on the estimates of the credit score fixed effects rejects the null hypothesis that they lose all of their predictive power.

Both of these findings are consistent with sizable random deviations of the interest rate from its information-efficient level. Such deviations in the pricing of loans can lead to a standard errors-in-variables effect that flattens the regression line and delivers an interest rate coefficient smaller than one. Similarly, when credit score information is priced into the interest rate with an added random term that diverts the price away from the information efficiency, the credit score FEs regain their statistical significance, as they help in filtering out the deviation.[16] In Section 4.3.1 below, we derive a measure of the magnitude of these deviations as a proportion of the overall variance of the interest rate.

To track the origins of these deviations, we augment the baseline regression in columns (2) and (4) of Table 5. The first of these proxies, *Aggregate Weekly Borrowing,* captures variations in aggregate demand for funding. For any given loan i, we measure the total value of loans auctioned off during the seven days that the loan's auction is open, and we normalize it by the initial size of FC's loan book at the beginning of that week, i.e.,

[16]Formally, let π be the expected default probability, with mean normalized to zero and variance σ_π^2. π is derived via the price-discovery process by extracting the information in the credit score, s, and combining it with the information contained in the order flow. The credit score has a zero-mean noise υ, $s = \pi + \upsilon$. Liquidity shocks drive the price, p, away from the default probability by a zero-mean random term, ε, $p = \pi + \varepsilon$. There is zero correlation between π,ε and υ. Our EMH regression eastimates the default probability, $\widehat{\pi} = a^* p + b^* s$, $(a^*, b^*) = argmin_{a,b} E\left(\pi - \widehat{\pi}\right)^2$. It follows that: $a^* = \frac{\left(1 - \frac{\sigma_\pi^2}{\sigma_s^2}\right)\frac{\sigma_\pi^2}{\sigma_p^2}}{1 - \frac{\sigma_\pi^2}{\sigma_s^2}\frac{\sigma_\pi^2}{\sigma_p^2}}$ and $b^* = \frac{\left(1 - \frac{\sigma_\pi^2}{\sigma_p^2}\right)\frac{\sigma_\pi^2}{\sigma_s^2}}{1 - \frac{\sigma_\pi^2}{\sigma_s^2}\frac{\sigma_\pi^2}{\sigma_p^2}}$. Under the EMH, $\varepsilon = 0$, $\frac{\sigma_\pi^2}{\sigma_p^2} = 1$, $a^* = 1$ and $b^* = 0$. But with variance $\sigma_\varepsilon^2 > 0$, $a^* < 1$ and $b^* > 0$.

the variable is the growth rate of FC's loanbook.[17] The validity of our test does not rely on such fluctuations in aggregate borrowing being uncorrelated with loan i's default probability. In the event that such a correlation arises, under the EMH, investors should incorporate the relevant information into the interest rate.

The -0.002 coefficient implies that an auction running at a time when there is a one-standard-deviation (equals 1.3%) surge in platform-wide demand for funds is associated with a reduced annual default probability of $1.0\%(= 1.3 \times 0.002 \times 4)$. We use a simple "correspondence argument" in order to identify the relationship between the shock and the deviation of the price from information efficiency, in both its direction and magnitude: since a demand surge is associated with a lower probability of default, it follows that, in order to restore information efficiency, the interest rate should be adjusted downwards. Or, to put it differently, the demand surge drives up the interest rate on loans auctioned during that period over and above their information-efficient level.

We use another proxy to capture variations in the composition of the funding supply. By design, a drop in active funding automatically injects, more passive funding along the autobid supply curve (see Figure 2). We decompose loan i's autobid funding into its aggregate time series component and its auction-specific, cross-sectional component. Again, we measure loan i's aggregate component of autobid funding as the average share of autobid funding across loans that are open during the seven days that loan i is open. We then measure the cross-sectional component of autobid funding as the difference between loan i's share of autobid funding and the aggregate component as defined above. This results in two variables, *Aggregate Weekly Autobid* and the *Loan Level Autobid*.

The estimated coefficient of *Aggregate Weekly Autobid* suggests that interest rates are above the information-efficient price at times of high aggregate autobid activity. The -0.016 implies that an auction running at a time when there is a one-standard-deviation (equal to 7%) surge in aggregate autobid funding is associated with a lower annual default probability of $45bp(= 7 \times 0.016 \times 4)$. Using the correspondence argument above, we conclude that periods of high autobid activity– thus low active investing– are associated with high interest rates relative to the level of the information-efficient price. This result

[17] Starting from zero, the loan book had an extremely high growth rateinitally, flattening gradually, implying that the series has a downwards non-linear trend, which we filter out using a logistic fitted line.

also suggests that, even though the autobid channels in more passive funds, the interest rate still remains too high. It can be argued that the autobid is undersmoothing the interest rate.

In contrast, the estimated coefficient of *Loan Level Autobid* suggests that interest rates are below the information-efficient price at times of high loan level autobid activity. The 0.05 coefficient implies that an auction in which the autobid is one standard deviation (which stands at 18%) more active is associated with a higher annual default probability of $36bp(= 18 \times .005 \times 4)$. Again, we use the correspondence argument above to conclude that an auction with a low level of active participation tends to end up with a low interest rate relative to the information-efficient price. In terms of the cross-sectional allocation, the autobid channels in an excessive amount of passive funds, resulting in an interest rate that is too low. In that respect, the autobid is oversmoothing the interest rate.

Another source of mispricing on the platform originates in the borrower's ability to terminate the auction early. Table 5's *Early Termination* variable is a dummy that receives a value of one in the event that the borrower takes such an action (and zero otherwise). Early termination is associated with an annualized default probability $1.2\%(= 0.003 \times 4)$ higher relative to auctions that run their full course, so that the interest rate should have been set higher by a similar amount. Obviously, investors are not informed about early termination until the event actually takes place, when it is too late for them to adjust their orders. Nevertheless, the FC auction design does price in the event of early termination: interest rates can move only downward in auction time, so that by terminating early, the borrower increases his own cost of borrowing. While early termination is, therefore, priced in, the estimated coefficient 0.003 shows that the "penalty" to the borrower for early termination is not sufficiently high.

Columns (3) and (6) of Table 5 address a problem raised in Iyer et. al. (2015), whereby EMH regressions may not be sufficient evidence for information aggregation. Rather, one may argue that the extra predictive power obtained by augmenting regressions containing credit scores with the interest rate may result from investors recovering information lost when the platform converted a *continuous* credit mark into discrete credit scores, from A+ to D. If this is so, no new information is actually revealed by the auction's price-discovery process. Like Iyer et al. (2015), we reject this hypothesis but use a different approach. If the credit mark is the only source of information, all loans should be priced within the credit-score band, which, in our case, is $50bp$ around the band's midpoint (see Table 3). The variable *Above The Band* is a dummy that receives a value of one if loan i's auction closes above the band. Interacted with the interest rate, the variable is positive, 0.15, and significant at the 5% level for the borrowing rate. The implication is that the information content of prices is actually higher above the credit score band.

4.2.1 Discussion: optimizing the autobid

Given that the autobid allocates 48% of the platform's funding, it can be argued that it could be used more effectively to improve price efficiency. That is, the parameters of the autobid could be recalibrated so as to refine and optimize the design towards information efficiency.

Consider, first, aggregate liquidity shocks, such as cases of high loan demand or low supply of funds by active investors. Our analysis finds that a high demand for new loans drives interest rates up and away from the information-efficient level for *all* loans auctioned at the time. A similar result is found in weeks of active investors' lower participation in the supply of funds. In both cases, the autobid should inject more funding to drive down the interest rate towards information efficiency. However, this policy required a stock of liquidity on which FC could draw. To build up such a liquid stock, the autobid would have needed to defer the allocation of passive funds contributed at times of excess aggregate funding. It seems, however, that FC was reluctant to engage in such reallocation of liquidity over time, as it owed the passive investors a prompt allocation of their funds into income-generating loans.

Consider next, liquidity shocks with a cross-sectional effect. Our analysis of the the *Loan Level Autobid* variable reveals that the autobid channeled too much liquidity into auctions in which active investors held back funding, resulting in an interest rate too low relative to information efficiency. Another related result is that, while early termination is penalized with a higher interest rate relative to auctions that run their full course, the penalty is not sufficiently high to achieve information efficiency. Both results highlight an oversmoothing effect, whereby the autobid inhibits market forces from having their full impact on the interest rate. The solution would be to make the autobid more price sensitive to the order flow, say, by making the autobid supply function steeper. In this case, no buildup of liquid funds would be required as funds would be reallocated across auctions running simultaneously.

It seems that the FC investor community was particularly aware of cross-sectional faults in the design of the autobid. For example, on February 21, 2014 at 1:28 pm, a blogger named "aloanatlast" commented on an auction that attracted a low level of attention from active investors: "The autobidder will now be chucking every penny it can into that loan." He then articulated his own idea about design refinement: "If I were an Autobid user [i.e., a passive investor], I'd want it to buy me a random sample, like a sort of index tracker - not something programmed to soak up the [loans] that manual bidders don't want." Notice that this refinement is a neutral allocation of passive funds, without any attempt to use the autobid to deal with liquidity shortages.

It has to be emphasized that our finding of a suboptimal design is not an indication of incompetence on the part of FC. Unlike in a market microstructure textbook problem, the (algorithmic) market maker had no prior knowledge of the cross-correlations between default probabilities and funding flows, upon which the parameters of a policy could be calibrated. Rather, in a newly innovated market, refinement could rely only on a trial-and-error learning process. That is, the autobid had to be "trained" dynamically on a relatively small sample of just $7,500$ auctions at the end of the sample period. Moreover, the number to be estimated– i.e., the default probability– was relatively small, with default events being realized with a long delay. In addition, any change in the design of the autobid was likely to cause an adjustment of the active investors' bidding strategies along the lines of the Lucas Critique (1976).

4.3 Liquidity events

To deepen our understanding of the way that liquidity shocks can drive prices away from fundamental values, we analyze two such events: the first involves the termination of the BIS program mentioned above, and the second involves a shortage of funding in auctions closing at hours less convenient to retail investors. In both cases, we find evidence of price impact that is not consistent with the EMH.

Starting in March 2013, the UK government, via the Department of Business Innovation and Skills (BIS), created a £20 million fund to be allocated via the FC platform. The intention was to support the development of alternative models of financial intermediation, in order to complement and compete with the banking industry, which was perceived to have failed to provide adequate funding to the SME sector (see Breedon Review (2012)). While the scheme was running, 20% of each loan processed by the platform was BIS-funded, at a price that was determined by the auctioning of the remaining 80% of the loan. BIS funding was eventually exhausted, with the last auction to benefit opening on February 28,2014. The effect was a temporary platform-wide shortage of funding, similar to the effect of a surge in loan demand, as analyzed in the previous section. The termination of the program was triggered solely by the depletion of the original fund and was unrelated to any other policy decision or to the functioning of the platform. It is, therefore, highly unlikely that there was any difference in loan quality before and after the event.[18]

[Figure 4 About Here]

To capture price dynamics around the event, we run a regression similar to equation (1)– namely, the borrowing rate on credit scores, but with periodic FEs, each lasting ten days, replacing the quarterly FEs in equation (1). The new equation is estimated over a period starting 60 days before the BIS termination date and ending 60 days after. Non-parametric estimates of the ten-day FEs are plotted in Figure 4. For completeness, the figure also plots a polynomial time path for the borrowing rate. The figure reveals a surge in the interest rate of around 1%, starting about ten days after the termination of the BIS funding, lasting for about 20 days and then, gradually, tapering off. The delay in the response is consistent with investors, both human and algorithmic, providing the market with liquidity from their precautionary reserves, but once these reserves were exhausted, prices were affected. Figure 4 gives currency to the notion of slow-moving capital. It took

[18]For both liquidity events we report in Online Appendix A a set of balancing tests on pre-determined borrower characteristics. These characteristics include borrower credit risk, activity and location. For both events we find no economic or statistically significant difference in these characteristics.

some time before new liquidity flowed in, either from new investors or from already active investors, responding to opportunities created by the end of BIS funding. Crucially, the time it took for capital to move in response to such opportunities is measured in weeks rather than hours or even days.

[Table 6 About Here]

For a stronger test, which controls for the default information that is incorporated into the price, we draw, again, on the EMH specification. Columns (1) and (2) in Table 6 present EMH regressions with the default dummy as a dependent variable. On the right-hand side, we have the interest rate, the previous set of covariates from Table 5, plus FEs for ten-days consecutive intervals following the event. The main result is a 70*bp* drop in default probabilities (2.8% annualized), over and above any information that is priced into the interest rate 20 to 30 days after the event. The result is statistically significant at the 5% to 10% level, notwithstanding the sharp drop in sample size. Using the conversion argument once again, we conclude that interest rates surged in the short term following the withdrawal of BIS funding.

The second test exploits the auction's closing hour. We conjecture that, throughout the day, retail investors have relatively less time to engage with the FC platform during work hours or late at night, as opposed to afternoon and early evening hours. Therfore, we define the peak hours as those between 4pm and 7pm, (henceforth peak hours). If our conjecture is correct, loans closing off-peak would attract les attention from retail investors and, therefore, create a shortage of funds. As the closing hour is publicly known at the open, under the EMH, liquidity providers should inject extra funds to compensate for the shortage off-peak, so as to avoid any deviation from price efficiency. Note, also, that the allocation of the opening hour is random, so that there should be no systematic quality difference between loans, conditional on the closing hour.

[Figure 5 About Here]

To confirm the difference in the flow of funds on- and off-peak, Figure 5 plots mean order flows, normalized by loan size, during the six hours prior to the closing of the auction. Only auctions that run for the full seven days are included, so the closing hour is, indeed, perfectly foreseen. As expected, loans that close within peak hours have a significantly higher order flow during the last hour. Normalized by loan size, last-hour order flow is 0.25 for auctions closing in peak hours, compared with only 0.15 for auctions closing in off-peak hours.

Again, for a stronger test, we augment our EMH regressions with the usual covariates, with a dummy variable for the off-peak closing hour.[19] Auctions that closed off-peak have a quarterly default probability 10*bp* (0.4% annualized) lower than that of auctions that close during peak hours. Using the conversion argument once again, we conclude that the interest rate was hiked for auctions closing off-peak, relative to the information-efficient interest rate. The result is statistically significant at the 10% level.

4.3.1 Economic magnitude of deviation from information efficiency

The two events allowed us to clearly trace out the mechanism through which liquidity shocks drive prices away from fundamental values. We turn next to a quantification of mispricing. Intuitively, we ask how much of the variance in the lending rate can be explained by expectations about fundamentals.

In order to obtain an idea of the magnitude of the deviations from information efficiency, we report the R-square of an equation in which the dependent variable is the borrowing rate, and the right hand side variable is the loan's estimated default probability. The latter is derived as a fitted value from the EMH regression in column (2) of Table 4.2. It can be shown that the R-square provides an upper bound on the information content of the price, as a fraction of the overall variance of the interest rate. It is striking that information accounts for, at most, 20% of the price variance; or, to put it differently, *at least* 80% of the variance is non-fundamental "noise" in the form of deviations from the information-efficient price. For full technical detail about the results reported here see Online Appendix B.[20]

[19]Unlike in Figure 5, we use the entire sample ($N = 80,529$), but in order to control for off-peak closures due to early, therefore unanticipated, terminations, we include the early-termination dummy among the covariates.

[20]A related issue concerns the over reliance of investors on the credit scores. In theory, the estimated default probability should price in all the relevant information contained in the credit scores. Thus, absent any over reliance of investors on credit scores, their inclusion on the right-hand side of the price equation should not change the estimated fit. However, we find that their inclusion pushes the R-square

It is noteworthy that not all platform participants were equally affected by this amount of noise in the price. While investors can diversify it away and, indeed, even use it to generate arbitrage profits (see Section 5), borrowers are unlikely to be able to do so. The result is that a borrower who was unlucky enough to auction his loan at a hiked-up interest rate, had to bear the cost for the entire duration of the loan, typically for several years. This result echoes one of the reasons given by FC for abandoning the auction design: "The price of each loan will now be based on the risk and term of the loan (and term) rather than the availability of investor funds."

4.4 The declining information content of prices

We might expect that, with the expansion of the platform and the accumulation of data and expertise, the quality of pricing would improve: the autobid can be better trained to extract information from funding flows and to prevent uninformative liquidity shocks from affecting prices. The hypothesized result is that the magnitude of random deviations of interest rates from information-efficient prices would fall, and the interest rate coefficient in the estimated EMH equation (4) would approach one.

[Table 7 About Here]

To test this hypothesis, column (1) of Table 7 augments the baseline EMH regression with a linear time trend, interacted with the borrowing rate, while column (2) interacts the borrowing rate with year dummies. Columns (3) and (4) repeat the test using the marginal rate.[21] We find a secular deterioration in the quality of pricing over time. Moreover, column (2) reports a benchmark coefficient of 0.82 in 2011, which is not statistically different from one; the EMH cannot be rejected for the early phase of FC's activity. Over time, the interest rate coefficients fall sharply, so that by 2015, the year before FC canceled the auction design and moved to posted prices, the interest rate coefficient was no longer statistically different from zero. That is, towards the end of our sample, prices

up from 20% to 48%. In other words, over reliance on the credit scores led to a significant amount of added noise in the pricing of the loan. Most likely an important contribution to this pattern was made by the autobid, which was calibrated to rely excessively on FC's own credit ratings.

[21]The specification does not include level year dummies since they would be perfectly collinear with the quarter of loan issue FEs.

no longer carry information that can contribute to improved default predictability. We explore reasons for this result in the next section.

5 The vanishing crowd

The declining information quality of price is surprising since it might be expected that, with accumulated experience, FC designers could gather more information over time and use it to recalibrate the autobid and improve price efficiency. One possible impediment to such learning might be that the patterns of the funding flows were not stable over time, and were changing faster than FC's learning process. In this section, we provide evidence that there was a sizable shift from active to autobid investment and, equally significantly, from a large number of small and medium-sized investors to a relatively small number of large investors. That is, to a large extent, the "crowd" simply vanished. We link these new results to the results for the EMH regressions, discussed earlier.

We use our bidding data to measure investor characteristics. Since investors' entry into (and exit from) the platform is staggered, comparing an investor who was active for several years with one who entered towards the end of the sample may be problematic. To restore a measure of consistency, we track investors' activity within fixed time intervals. That is, for each investor, we aggregate the value of accepted bids across auctions within calendar quarters. Hence, for each investor quarter, we derive a portfolio, denoted as *Total Wealth*, which is used to sort investors into size groups of less than £100, £100-500 etc. (see Table 8). Each portfolio may be managed via the autobid (see column (3) of Table 8), or invested directly by the investor, which is denoted as the *Active Portfolio*. We classify investors as *Active* even if they partly use the autobid. Summary statistics on their portfolios can be found in columns (5) to (10) of Table 8.

[Figure 6 & Table 8 About Here]

Two important developments took place during the sample period. The first was a significant increase in autobid activity. The left panel in Figure 6 plots the ratio between *Autobid Allocation* and *Total Wealth* against logged *Total Wealth*, across years and across investor size classes. From 2011 to 2014, autobid activity increased across *all* wealth classes. Surprisingly, the effect was even stronger at the high end of the size distribution, which is an early indication that we need not take it for granted that all large investors are well-informed or even sophisticated. For example, within the highest wealth class– that is, among investors who put more than £$50,000$ per quarter into the platform the share of their autobid investment increased from 9% ($8,206/86,681$) in 2011 to 33% in 2014. This pattern is revealing, given the fact, reported in Table 1 above, that, on average, autobid orders are placed 1% below the marginal interest rate at the close, where active investors can bid just below the marginal close in order to receive an allocation. Overall, across size groups, the share of funding flows allocated by the autobid increased from 44% in 2011 ($(291 \times 32 + ... + 5 \times 8,206)/(291 \times 49 + ... + 5 \times 86,681)$) to 55% in 2014.

The second development is a very significant shift over time, towards the high end of the wealth distribution. In particular, within the over- £$50,000$ wealth class, the size of the active portfolio increased from £$83,100$ in 2011 to £$198,600$ in 2014. While in 2011, the over- £50,000 class contributed only 14% of the total funding, $(83.1 \times 4)/(83.1 \times 4 + ... + 0.04 \times 113)$, by 2014, that figure had increased to 44%. Thus, by 2014, 75% $(0.55 + 0.45 \times 0.44)$ of the funding was allocated by either FC's algorithm or by large active investors. Taken together, these two trends– greater autobid participation and the rise of large investors– indicate that the size of the crowd diminished significantly over the sample period.

Following Bach, Calvet and Sodini (2016), we report in Column (7) of Table 8 the realized rates of return across size groups. For each active portfolio, we calculate a weighted average of realized returns of the loans in the portfolio.[22] Loans that are still open at the end of our sampling window are excluded since we do not know if they have defaulted. For that reason, we report the returns separately for each year, focusing the analysis on annual variations across wealth classes. The effect is economically significant: between the highest and the lowest wealth class, the annual gap in realized returns increased sharply over the years 2011 to 2014 , $56bp(= 8.68 - 8.12)$, $107bp$, $114bp$ and $166bp$, respectively.

There is no reason to suppose that *all* of that difference was due to large investors be-

[22]More accurately, we calculate the return on each loan part since investors may have accepted bids at different interest rates for the same loan. Returns on loans include loan repayments and, in the event of default, post-default recoveries.

ing better informed. Some might be attributed to their ability to better exploit arbitrage opportunities, perhaps by building their own algorithms.[23] Table 8 reports several additional investor characteristics. *Loan Stake*, reported in column (6), is the mean amount invested in a single loan, varying from around £30 (just above the £20 minimum bid that the platform allows) to around £2,000. The right panel of Figure 6 plots the ratio between the *Loan Stake/Active Portfolio* against logged *Active Portfolio.* The results show that better-diversified, large investors invested in riskier loans. Columns (8) and (9) in Table 8 report that the share of A-rated loans decreased with investor size, while the share of C rated loans increased with investor size.

Column (10) in Table 8 reports the weighted average of the spread between the marginal closing rate and the interest rate on the loans in the active portfolio. Uniformly, this spread is falling in wealth: the difference in spreads between the smallest and the highest size group is 29bp($= 0.48 - 0.19$), 46bp, 11bp and 33bp for the years 2011 to 2014, respectively. As noted above, in FC's open order-book auction, investors needed only slightly undercut the marginal closing rate in order to receive an allocation. Larger, more professional investors, possibly assisted by algorithms, could submit last-minute orders, just below the marginal closing rate.

[Table 9 About Here]

Table 9 applies regression analysis to the same data as in Table 8, using the realized rate of return on the active portfolio as the dependent variable. The estimates on the credit scores, as well as the spread between the bid and the marginal rate are consistent with the proposition that much of the extra return made by large active investors was due to their ability to exploit arbitrage opportunities. All coefficients are clustered at the investor level and are statistically significant at the 1% level. In columns (3) and (4), we calculate, for each portfolio, the share of investment in auctions with above-median *Aggregate Weekly Borrowing.* Investors who timed their bidding to coincide with surges in the demand for loans benefited from a higher rate of return. Yet the size effect, which *might* capture better-informed investors although it might also capture a "talent" to identify arbitrage opportunities– remains positive. Even when the regression is saturated with an investor

[23] While collecting our data, we came across some anecdotal evidence to that effect.

FE, the estimate on the *Active Portfolio* (logged) is 0.55. Hence, investor size, between the £50 and the £50,000 size groups explains only $17bp (= 0.55 \times [ln(50,000) - ln(50)]$ of the 1% difference reported in Table 8. These results are consistent with Bach, Calvet and Sodini (2016), in that risk taking is a dominant factor in explaining excess returns for large investors.

The changes in the composition of the investor population were likely to have conflicting effects on information efficiency. All else equal, the transition from active to autobid investing decreased the amount of information on the platform. At the same time, had the autobid been effectively calibrated, a better-funded autobid could have provided the market with more liquidity. Section 4.2 demonstrates that this was not the case. As for the large investors, even if they had been somewhat better-informed, the net effect of replacing a crowd of small investors with a small number of large investors remains ambiguous.

In short, the combined effect of the changes in investment composition is an empirical question. Therefore, we augment our reduced-form EMH regression with a proxy for the large investors' aggregate activity and interact it with the interest rate. Constructing the large-investor proxy, we follow the same logic as before: for each loan, we take a weighted average of large-investors' funding across all the auctions open in the week of the given loan. We also interact the *Aggregate Autobid Activity* with the interest rate. The results are presented in Table 10.

[Table 10 About Here]

The results suggest that large investor activity had a strong negative impact on the predictive power of prices. The interest rate coefficient becomes significantly smaller when large investors' activity increases. Quantitatively, we can approximate the contribution of the shift in investors across time to the loss of pricing efficiency: between 2011 and 2015, the share of large-investor funding grew by 12 percentage points and predicted a 28% drop in the information content of the interest rate.[24]

[24] The calculation is based upon estimates in column (1) of Table 10 ((.12*-1.48)/.642). It is also worth noting that the level effect of large investors is positive. That is, given the interest rate, loans auctioned off at times with higher large-investors activity tended to have a higher default probability. Hence, the interest rate tended to be depressed relative to the information-efficient interest rate. In that respect, the

6 Robustness Checks

6.1 Estimation Methods

We test the robustness of our findings on price efficiency relative to alternative non-linear estimators. Column (1) of Table 11 reports the OLS estimates for the benchmark EMH specification. Columns (2) and (3) estimate the default equation using Logit/Probit models with marginal effects evaluated at the mean. Column (4) in each panel estimates the relationship using a Cox Proportional Hazards model.

[Table 11 About Here]

The relationship between default and interest rates is unaffected by the estimation method, the point estimates being statistically indistinguishable from one another. The Cox Proportional Hazard model operates on a sample of loans rather than on the sample of performance periods– hence the sharp drop in sample size. In this setting, the benchmark is a hazard ratio of 1, indicating that the variable does not affect survival. Consistent with the previous estimates, the Proportional Hazard model predicts an increased risk of default of 0.35% associated with a 1% increase in the borrowing rate.

6.2 Repeat Borrowers

In Section 4.4, we documented that, despite the growth of the platform, the pricing of loans deteriorated over time. To corroborate this finding, we implement another learning test that exploits repeat borrowers on the platform. Among the 7455 loans originated in our sample period, 2127 loans correspond to returning borrowers. Among these repeat borrowers, three quarters obtained two loans, and the maximum number of loans obtained by the same borrower was six (three borrowers). We embed repeat borrowers in our EMH specification and test for learning effects.

[Table 12 About Here]

results do not lend support to the hypothesis that large investors monopolized the market or manipulated the price.

In column (1) of Table 12, we interact the interest rate with a dummy variable for repeat borrowers, *Repeat Borrower*. In column (2) of the same table we interact the interest rate with the count of the number of loans taken up by the borrower until that auction, *Borrower Freq.*. In both cases, we again reject an improvement in the pricing efficiency of loans of repeat borrowers with respect to non-repeat borrowers.

6.3 Early Termination

We next provide additional evidence of patterns of early termination and how they interacted with pricing efficiency. 2032 loans, or 27% of the sample, were terminated early. In columns (3) and (6) of Table 12, we include, in addition to the level dummy variable,an interaction between the interest rate and the *Early Termination* dummy.

The estimates contain two important insights. First, the interest rate coefficient is not significantly affected by the distinction. Second, apart from the level effect of early termination, there is no evidence that auctions running their full course have higher price efficiency. This pattern also holds with respect to the declining information content of prices (Tables available upon request). Thus, our results are not driven by the inclusion of these borrowers.

Quantitatively, the effect of early termination is about 100bps. Offsetting the cost of early termination, one has to consider the potential benefits of timely transactions where delays may carry significant penalties. For example, loans for which the declared purpose is to meet tax payments have an early termination rate of 47%, versus 30.2% for other purposes. Similarly, the share of early terminations is significantly larger for loans used for working capital, where firms can earn the suppliers' discounts for early cash payment. Consistent with the EMH evidence, the incidence of early terminations is larger for borrowers with lower credit ratings.

7 Conclusions

We provide evidence that dispersed, nonstandard, information in the hands of investors in FC's online SME loans auctions was incorporated into interest rates through a price-discovery process, enhancing default predictability over and above standard credit scores.

At the same time, our evidence is also consistent with the presence of liquidity shocks that drove prices away from information efficiency. Quantitatively, these deviations account for most of the variance of the interest rate. Liquidity shocks were not neutralized by large and liquid investors or by the autobid, as they should have been in an efficient market. The problem was exacerbated over time. Substantial changes in the composition of the investor population contributed to the deterioration in the quality of pricing, making a refined calibration of the autobid more difficult. These findings are consistent with the reasons given by FC for its decision to abandon the auction design of the platform in favor of posted prices.

The results shed light on the comparative advantages and disadvantages of the old model of financial intermediation, operated by commercial banks, compared with the new technology, operated by FC and other platforms. While P2Bs' ability to directly connect investors and corporate borrowers relieves them from the need to hold costly regulatory capital, it also exposes them to the effect of liquidity shocks. In contrast, the banks' liquid inventories allow them to bridge funding gaps, but at a cost of exposing them to credit risks and the resulting regulatory oversight, including capital requirements. The diversity in designs used in the P2B industry, as well as the changes that have taken place in FC, indicate that the industry is still in a stage of experimentation. It is, therefore, premature to pass judgment on whether the transparent, decentralized and disintermediated model of credit finance can displace the old model.

References

[1] Admati, Anat and Motty Perry, (1991) "Joint Projects without Commitment." *Review of Economic Studies*, 58, pp. 259-276.

[2] Allen, Franklin and Douglas Gale, (1995). "A Welfare Comparison of Intermediaries and Financial Markets in Germany and the US." *European Economic Review*, 39, pp. 179-209.

[3] Bach, Laurent, Laurent E. Calvet, and Paolo Sodini. "Rich pickings? Risk, return, and skill in the portfolios of the wealthy." (2016).

[4] Back, Kerry and Jaime F. Zender, (1993). "Auctions of Divisible Goods: On the Rationale for the Treasury Experiment." *Review of Financial Studies*, 6, pp. 733–764.

[5] Baeck, Peter, Liam Collins and Bryan Zhang (2014). *Understanding Alternative Finance, The UK Alternative Finance industry Report 2014.* NESTA, University of Cambridge.

[6] Bellia, Mario, Loriana Pelizzon, Marti G. Subrahmanyam, Jun Uno and Darya Yuferova, (2017). *Coming Early to the Party: High Frequency Traders in the Pre-Opening Phase and the Opening Auction of Euronext Paris.*

[7] Benveniste, Lawrence M. and William J. Wilhelm, (1990). "A comparative analysis of IPO proceeds under alternative regulatory environments." *Journal of Financial Economics*, 28, pp. 173-207.

[8] Biais, Bruno, and Anne Marie Faugeron-Crouzet, (2002). "IPO Auctions: English, Dutch, . . . French, and Internet." *Journal of Financial Intermediation*, 11, pp. 9–36.

[9] Biais, Bruno, Pierre Hillion, and Chester Spatt, (1999). "Price Discovery and Learning during the Preopening Period in the Paris Bourse." *Journal of Political Economy*, 107 (6), pp. 1218-1248.

[10] Brenner Menachem, Dan Galai and Orly Sade, (2009). "Sovereign debt auctions: Uniform or discriminatory?" *Journal of Monetary Economics*, 56, pp. 267–274.

[11] Buchak, G., Matvos, G., Piskorski, T., and Amit Seru, (2018). "Fintech, regulatory arbitrage, and the rise of shadow banks." *Journal of Financial Economics*, 130(3), pp. 453-483.

[12] Butler, Alexander W., Jess Cornaggia and Umit G. Gurun (2017). "Do Local Capital Market Conditions Affect Consumers' Borrowing Decisions?" *Management Science*, 63, pp. 4175-4187.

[13] The Breedon Review, (2012). *Boosting Finance Options for Business.* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/32230/12-668-boosting-finance-options-for-business.pdf.

[14] Cameron, A. Colin, and Pravin K. Trivedi, (2005). "Microeconometrics: methods and applications." *Cambridge university press.*

[15] Chakraborty, Archishman and Bilge Yilmaz, (2004). "Information Manipulation." *Journal of Economic Theory*, 114, pp. 132-152.

[16] Collin-Dufresne, Pierre, Robert S. Goldstein and J. Spencer Martin, (2001). "The Determinants of Credit Spread Changes," *The Journal of Finance*, 56, pp. 2177-2207.

[17] Cespedes, Jacelly, (2019). "Heterogeneous Sensitivities to Interest Rate Changes: Evidence from Consumer Loans."

[18] D'Acunto, Francesco, Nagpurnanand Prabhala and Alberto G. Rossi, (2019). *"The Promises and Pitfalls of Robo-advising*." *The Review of Financial Studies*, 32(5), pp. 1983-2020.

[19] De Roure, Calebe, Loriana Pelizzon, and Anjan V. Thakor, (2019). *"P2P lenders versus banks: Cream skimming or bottom fishing?."* 18-13.

[20] Diamond, Douglas W., (1984). "Financial Intermediation and Delegated Monitoring." *Review of Economic Studies*, 51, pp. 393-414.

[21] Duffie, Darrell, (2010). “Asset Price Dynamics with Slow-Moving Capital.” *The Journal of Finance*, 65, pp. 1237-1267.

[22] Einav, Liran, Chiara Farronato, Jonathan D. Levin and Neel Sundaresan (2018). “Auctions versus posted prices in online markets” *Journal of Political Economy*, 126(1), pp. 178-215.

[23] Feldhütter, Peter and Stephen M. Schaefer, (2018). “The Myth of the Credit Spread Puzzle,” *Review of Financial Studies,* Forthcoming.

[24] Franks, Julian and Oren Sussman, (2005). “Financial Distress and Bank Restructuring of Small to Medium Size UK Companies.” *Review of Finance*, 9, pp. 65-96.

[25] Glosten, Lawrence R. and Paul Milgrom, (1985). “Bid, ask and transaction prices in a specialist market with heterogeneously informed traders.” *Journal of Financial Economics*, 14, pp. 71-100.

[26] Goldstein, Itay and Alexander Guembel, (2008). “Manipulation and the Allocational Role of Prices.” *Review of Economic Studies*, 75, pp. 133–164.

[27] Goldstein, Itay, Wei Jiang, and Andrew G. Karolyi, (2019). “To FinTech and Beyond.” *Review of Financial Studies*, 32, pp. 1647–1661,

[28] Grennan, Jullian and Roni Michaely, (2019). *FinTechs and the Market for Financial Analysis*, Swiss Finance Institute Research Paper Series 19-10.

[29] Gromb, Denis and Dimitri Vayanos, (2018). “The Dynamics of Financially Constrained Arbitrage.” *Journal of Finance*, 73, pp. 1713-1750.

[30] Hertzberg, Andrew, Andres Liberman, and Daniel Paravisini, (2018). "Screening on loan terms: evidence from maturity choice in consumer credit." *The Review of Financial Studies*, 31(9), pp. 3532-3567.

[31] Hilscher, Jens and Wilson, Mungo, (2017). Credit Ratings and Credit Risk: Is One Measure Enough? *Management Science*, 63, pp. 3414-3437.

[32] Hortaçsu, Ali and David McAdams (2010). “Choice and Strategic Bidding in Divisible Good Auctions: An Empirical Analysis of the Turkish Treasury Auction Market.” *Journal of Political Economy*, 118, pp. 833-865.

[33] Iyer, Rajkamal, Asim Ijaz, Khwaja, Erzo F. P. Luttmer and Kelly Shue, (2015). “Screening Peers Softly: Inferring the Quality of Small Borrowers,” *Management Science*, 62, pp. 1554-1577.

[34] Jackson, Andrew (2016). “Marketplace lending: a new strategy for dealing with distressed businesses.” *Corporate Rescue and Insolvency,* June, pp. 125-127.

[35] Kyle, Albert S. (1985). “Continuous Auctions and Insider Trading.” *Econometrica*, 53, pp. 1315-1335.

[36] Kyle, Albert (1989). “Informed Speculation with Imperfect Competition.” *Review of Economic Studies*, 56, pp. 317-355.

[37] Liskovich, Inessa, and Maya Shaton, (2017). "Borrowers in search of feedback: Evidence from consumer credit markets."

[38] Lucas, Robert, (1976). “Econometric Policy Evaluation: A Critique”. In Brunner, K. and A. Meltzer (eds.). The Phillips Curve and Labor Markets. Carnegie-Rochester Conference Series on Public Policy. 1. New York: American Elsevier. pp. 19–46.

[39] Levine, Ross and Sara Zervos, (1998). “Stock Markets, Banks, and Economic Growth.” *The American Economic Review*, 88 (3), pp. 537-558.

[40] Lin, Mingfeng, Nagpurnanand R. Prabhala and Siva Viswanathan, (2013). “Judging Borrowers by the Company They Keep: Friendship Networks and Information Asymmetry in Online Peer-to- Peer Lending.” *Management Science*, 59, pp.17-35.

[41] Lin, Mingfeng and Siva Viswanathan, (2016). “Home Bias in Online Investments: An Empirical Study of an Online Crowdfunding Market.” *Management Science*, 62, pp. 1393-1414.

[42] Morse, Adair, (2015). “Peer-to-peer crowdfunding: Information and the potential for disruption in consumer lending.” NBER Working Paper, 20899.

[43] Nyborg, Kjell and Suresh Sundaresan, (1996). “Discriminatory versus uniform Treasury auctions: Evidence from when-issued transactions.” *Journal of Financial Economics*, 42, pp. 63-104.

[44] Philippon, Thomas, (2016). “The Fintech Opportunity." NBER Working Paper, 22476.

[45] Ravina, Enricheta, (2019). “Love & Loans The Effect of Beauty and Personal Characteristics in Credit Markets.”

[46] Rock, Kevin, (1985). “Why New Issues are Underpriced.” *Journal of Financial Economics*, 15, pp. 187-212.

[47] Roth, Alvin E, and Axel Ockenfels, (2002). “Last-Minute Bidding and the Rules for Ending Second-Price Auctions: Evidence from eBay and Amazon Auctions on the Internet.” *The American Economic Review*, 92, pp. 1093-1103.

[48] Shleifer, Andrei (1986). “Do demand curves for stocks slope down?” *Journal of Finance*, 41, pp. 579-590.

[49] Soyeshi, Glenn, (1995). “A Class of Binary Response Models for Grouped Duration Data.” *Journal of Applied Econometrics*, 10, pp. 411-431.

[50] Tang, Huan, (2019). "Peer-to-peer lenders versus banks: substitutes or complements?." *The Review of Financial Studies,* 32(5), pp. 1900-1938.

[51] Treasury, U. S., (2016). "Opportunities and Challenges in Online Marketplace Lending." *US Treasury, Washington, DC.*

[52] Vallee, Boris, and Yao Zeng, (2018). Marketplace Lending: A New Banking Paradigm? *Review of Financial Studies* 31, pp. 1939-1982.

[53] Wei, Zaiyan and Mingfeng Lin (2017). “Market Mechanisms in Online Crowdfunding.” *Management Science,* 63(12), pp. 4236-4257.

[54] Yermack, David (2015). “Corporate Governance and Blockchains.” NBER Working Paper, 21802.

[55] Wilson, Robert , (1979). “Auctions of shares,” *Quarterly Journal of Economics*, 93, pp. 675–698.

[56] Zhang, Bryan, Tania Ziegler, Leyla Mammadova, Daniel Johanson, Mia Gray and Nikos Yerolemou, (2018). *The 5th UK Alternative Finance Industry Report*, University of Cambridge.

Table 1: Descriptive Statistics on Loans and Investors

Variable	Mean	Median	S.D.	Min	Max
Panel A: Loan Characteristics					
Loan Size (£000)	57	50	40	5	516
Maturity (months)	44	36	14	6	60
Age of SME (years)	12	9	10	0	107
Credit Score: AA	0.116	0	.321	0	1
A	0.298	0	.457	0	1
B	0.268	0	.443	0	1
C	0.232	0	.422	0	1
D	0.083	0	.276	0	1
Panel B: Investors' Characteristics					
Investors Per Loan (N)	571	459	446	2	5393
Number of Loan Parts	805	659	594	37	8117
Active Investors Per Loan (N)	200	176	127	2	985
Share of Top Lender (%)	8	10	7	0.2	83
Share of Top 5 Lender (%)	18	17	11	0.7	100
Share of Top 20 Lender (%)	29	27	14	0.7	100
Investor Stake Per Loan (%)	1.06	.48	2.51	0.5	79.85
Share of Autobid (%)	48	50	18	0	99
Active Bid to Marginal Rate Spread(%)	.37	.27	.36	0	4.48
Passive Bid to Marginal Rate Spread(%)	1.01	.61	1.06	0	7.41
Length of Auction (hours)	157	168	15	0.1	504.0

Descriptive statistics, 7,455 FC loans, issued between 2010Q4 and 2015Q1. Panel A presents descriptive statistics on the loans issued. *Loan Size* is measured in thousands of GBP; *Maturity* of the loan is measured in months, and *Age* of the borrowing firm in terms of years. *Credit Score* is a dummy variable equal to one for the credit category of the borrower. Panel B presents descriptive statistics on investor characteristics. *Investors Per Loan* counts all distinct investors with accepted bids in a loan. *Number of Loan Parts* counts all distinct bids accepted at the end of the auction. *Active Investors Per Loan* counts all distinct active investors with manually placed and accepted bids. *Share of Top (...) Lender* measures the contribution of the largest investors relative to the total amount of the loan. *Active Bid to Marginal Rate Spread* measures the difference between the interest rates of manually placed and accepted bids, relative to the marginal interest rate at the close–i.e., the highest interest rate accepted. *Passive Bid to Marginal Rate Spread* measures the difference between the interest rates of autobid placed and accepted bids, relative to the marginal interest rate at the close. *Length of the Auction* measures time, in terms of hours, between the opening and close of the auction.

Table 2: Mean Daily Order Flows, Normalized by Loan Size

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Autobid		Active Investors								
			Small			Medium			Large		
Day	Flow	Exec.	Flow	Exec.	New	Flow	Exec.	New	Flow	Exec.	New
1	0.65	0.25	0.05	0.01	0.05	0.18	0.01	0.18	1.11	0.01	1.11
2	0.06	0.04	0.03	0.00	0.02	0.08	0.01	0.03	0.17	0.00	0.03
3	0.06	0.04	0.03	0.00	0.02	0.08	0.01	0.03	0.18	0.00	0.03
4	0.05	0.03	0.03	0.00	0.02	0.08	0.01	0.03	0.13	0.00	0.02
5	0.04	0.03	0.04	0.01	0.03	0.08	0.01	0.03	0.10	0.01	0.02
6	0.04	0.03	0.05	0.01	0.03	0.10	0.02	0.06	0.11	0.01	0.02
7	0.04	0.04	0.14	0.08	0.11	0.36	0.18	0.22	0.47	0.18	0.23
Total	0.94	0.46	0.37	0.11	0.28	0.96	0.25	0.56	2.27	0.21	1.46

Daily order flows for auctions that ran for seven days. Auction time is split into 24-hours interval, each defining a "Day" in auction time. Daily orders are normalized by loan size and then averaged across auctions. Passive investors place orders via the autobid. Active investors are sorted by the total amount of their daily order: small, less than £100 a day; medium, £100-1000 a day; and large, more than £1000. "Exec." informs of the amount eventually accepted (when the auction is closed). Orders are classified as "New" on the first day that a certain investor participates in a certain auction.

Table 3: Loan Interest Rates, Default and Post-default Recovery Rates

	Interest Rates Regressions		Default Regressions	
	(1)	(2)	(3)	(4)
	Borrowing Rate	Marginal Rate	Default dDummy	Recoveries Post Default / Balance Remaining
Constant	8.472***	8.967***		-0.522*
	(0.100)	(0.165)		(0.295)
Credit Scores FEs				
AA Score	-1.164***	-1.096***	-0.004***	-0.074
	(0.032)	(0.053)	(0.001)	(0.071)
B Score	0.976***	1.002***	0.003***	0.002
	(0.024)	(0.040)	(0.001)	(0.038)
C Score	1.987***	1.986***	0.003***	-0.041
	(0.025)	(0.042)	(0.001)	(0.040)
D Score	3.713***	3.423***	0.007***	-0.003
	(0.036)	(0.060)	(0.002)	(0.051)
Loan Life Cycle FEs				
Early			0.007***	
			(0.001)	
Mid-Early			0.010***	
			(0.001)	
Mid-Late			0.008***	
			(0.001)	
Late			0.004***	
			(0.001)	
Log(MRecovery Period)				0.217***
				(0.077)
Quarterly of Loan Issue FEs	YES	YES	YES	YES
R^2	0.787	0.618	0.0105	0.142
N	7,455	7,455	81,049	674

The table reports estimates from OLS regressions using interest rates, default dummies and post-default recoveries as dependent variables. Columns (1) and (2) report estimates of equation (1) using the sample of $7,455$ FC loans issued between 2010Q4 and 2015Q1. The dependent variables are the borrowing rate in column (1) and the marginal closing rate in column (2). The specifications include fixed effects for credit rating, FE_Score, and the quarter when the loan was issued, FE_QIssue. Column (3) estimates the default equation (3) using $81,049$ performance quarters of the same 7,455 loans. The dependent variable, $Default$, is set equal one if loan i defaults during performance quarter t, and zero otherwise. The specification also includes fixed effects for the four stages of the loan's life cycle, FE_SLife. Column (4) reports estimates for the recovery equation (3) using all 671 default events realized in the same 7,455 loans. The dependent variable is the post-default recovery, as a fraction of the balance outstanding at the point of default, $RRecoveryPost$. $MRecovery$ measures the length of the recovery period, in months, from the point of default to the end of the sample, logged. Heteroscedasticity robust standard errors, clustered at the borrower level, are reported in brackets. One, two or three starts denote significance at the 10%, 5% or 1% level, respectively.

Table 4: Annualized Default Probabilities, Recoveries and LGD (%)

Credit score	Default prob.	Recovery Rates		LGD
		To default	Post default	
	(1)	(2)	(3)	(4)
A+	1.3	40.1	29.5	42.2
A	2.9	45.7	37.0	34.2
B	3.9	46.8	37.1	33.4
C	4.0	46.9	32.8	35.7
D	5.7	47.8	36.6	33.1

Column (1) reports annualized default probabilities based on estimates of quarterly default probabilities reported in column (3) of Table 3. Column (2) uses default probabilities conditional on credit scores and the various stages in the life cycle of the loan, as reported in Figure 3, in order to compute recovery rates up to the point of default. Column (3) extrapolates estimates reported in columns (4) of Table 3 in order to derive recovery rates up to five years post-default. Column (4) combines columns (2) and (3) to derive LGD conditional on credit scores.

Table 5: Baseline EMH Regressions

	Borrowing Interest Rate			Marginal Interest Rate		
	(1)	(2)	(3)	(4)	(5)	(6)
Borrowing Interest Rate	0.333***	0.280***	0.247**			
	(0.067)	(0.070)	(0.107)			
Marginal Interest Rate				0.179***	0.134***	0.236***
				(0.038)	(0.041)	(0.073)
Industry Asset Beta	0.007**	0.008***	0.008***	0.007**	0.008***	0.008***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Aggregate Weekly Borrowing		-0.002*	-0.002		-0.002	-0.002
		(0.001)	(0.001)		(0.001)	(0.001)
Aggregate Weekly Autobid		-0.016**	-0.015**		-0.016**	-0.016**
		(0.007)	(0.007)		(0.007)	(0.007)
Loan Level Autobid		0.005**	0.006**		0.004*	0.004
		(0.002)	(0.003)		(0.002)	(0.003)
Early Termination		0.003***	0.003***		0.003***	0.003***
		(0.001)	(0.001)		(0.001)	(0.001)
Above Band Deviation			-0.011**			0.003
			(0.005)			(0.004)
Interest Rate*Above Band Deviation			0.150**			-0.063
			(0.075)			(0.063)
Credit Score FEs	Yes	Yes	Yes	Yes	Yes	Yes
Loan Life Cycle FEs	Yes	Yes	Yes	Yes	Yes	Yes
Quarter of Loan Issue FEs	Yes	Yes	Yes	Yes	Yes	Yes
Joint Significance of Credit Scores FEs	***	***	***	***	***	***
R-squared	0.011	0.011	0.011	0.011	0.011	0.011
N	80,529	80,529	80,529	80,529	80,529	80,529

The table reports OLS estimates of the baseline equation (4) using over 80,000 performance quarters of the 7,455 FC loans issued between 2010Q4 and 2015Q1. The dependent variable, *Default*, is set equal to one if loan i defaults during performance quarter t, and zero otherwise. In columns (1) and (2), we use the LGD adjusted borrowing rate, and in columns (3) and (4) the LGD adjusted marginal interest rate. *Industry Asset Beta* is the company's asset beta measured at the industry level (Sys). We control for various proxies of liquidity shocks (Liq). *Aggregate Weekly Borrowing* is the total value of the loans auctioned off during the week that the auction for loan i is open, normalized by the value of the loan book at the beginning of the week detrended by a logistic line. *Aggregate Weekly Autobid* is the average share of autobid funding across loans that are open during the week that loan i is open. *Loan Level Autobid* is the difference between loan $i's$ share of autobid funding and *Aggregate Weekly Autobid*. *Early Termination* is a dummy variable that receives a value of one if the borrower terminates the auction terminated prematurely, and zero otherwise. *Above The Band* is a dummy variable that receives a value of one if loan i's auction closes 50*bp* above the midpoint of the credit score band, and zero otherwise. We also include FEs for the credit scores, the quarter when the loan was issued, and the four stages of the loan's life cycle (FE_Score, FE_QIssue, and FE_SLife, respectively in equation (4)). Heteroscedasticity robust standard errors, clustered at the borrower level, are reported in brackets. One, two or three starts denote significance at the 10%, 5% or 1% level, respectively, including an F test for the joint significance of the credit score FEs.

Table 6: EMH Regressions - Liquidity Events

	Off-Peak Closure		BIS Funding Ends	
	(1)	(2)	(3)	(4)
Borrowing Interest Rate	0.583***		0.287***	
	(0.194)		(0.071)	
Marginal Interest Rate		0.241***		0.141***
		(0.093)		(0.043)
Off-Peak Closure			-0.001*	-0.001*
			(0.001)	(0.001)
Days 0 to +10	-0.003	-0.002		
	(0.005)	(0.005)		
Days +10 to +20	-0.006	-0.005		
	(0.005)	(0.005)		
Days +20 to +30	-0.009**	-0.007*		
	(0.005)	(0.004)		
Days +30 to +40	-0.009*	-0.008		
	(0.005)	(0.005)		
Days +40 to +50	-0.006	-0.005		
	(0.005)	(0.005)		
Days +50 to +60	-0.006	-0.004		
	(0.006)	(0.006)		
Covariates	Yes	Yes	Yes	Yes
Credit Scores FEs	Yes	Yes	Yes	Yes
Loan Life Cycle FEs	Yes	Yes	Yes	Yes
Quarter of Loan Issue FE	No	No	Yes	Yes
Joint Significance of Credit Scores FEs	NS	NS	***	***
R-squared	0.012	0.012	0.011	0.011
N	13,588	13,588	80,529	80,529

Table 5 reports OLS estimates of the baseline EMH equation (4) adapted to test the impact of two liquidity events. The dependent variable, *Default*, is set equal to one if loan i defaults during performance quarter t, and zero otherwise. In columns (1) and (2), we use the LGD adjusted borrowing rate, and in column (3) and (4) the LGD adjusted marginal interest rate. Columns (1) and (2) use $13,588$ performance quarters of loans auctioned off within a ± 60 days window around the date when BIS funding ended, February 28, 2014. The specification includes six fixed effects, *Days(..)*, one for each ten-day interval starting at the BIS ending date, the omitted baseline being the period before the ending date. Columns (3) and (4) use over $80,000$ performance quarters of the 7,455 FC loans issued between 2010Q4 and 2015Q1. *Off-Peak Closure* is a dummy variable that receives a value of one if the auctions close outside of the 4PM to 7PM window. Additional control variables include *Aggregate Weekly Borrowing, Industry Asset Beta*, *Aggregate Weekly Autobid*, *Loan Level Autobid*, and *Early Termination*. All variables are defined as in Table (5). We also include FEs for the credit scores, the quarter when the loan was issued, and the four stages of the loan's life cycle. Heteroscedasticity robust standard errors, clustered at the borrower level, are reported in brackets. One, two or three starts denote significance at the 10%, 5% or 1% level, respectively, including an F test for the joint significance of the credit score FEs.

Table 7: EMH Regressions - Price Quality Over Time

	Borrowing Interest Rate		Marginal Interest Rate	
	(1)	(2)	(3)	(4)
Borrowing Interest Rate	0.691***	0.822***		
	(0.181)	(0.188)		
Marginal Rate			0.398***	0.525***
			(0.147)	(0.166)
Interest Rate*Trend	-0.009**		-0.006*	
	(0.004)		(0.003)	
Interest Rate*2012		-0.455***		-0.338**
		(0.169)		(0.159)
Interest Rate*2013		-0.438**		-0.352**
		(0.171)		(0.161)
Interest Rate*2014		-0.553***		-0.407**
		(0.173)		(0.161)
Interest Rate*2015		-0.652***		-0.482***
		(0.180)		(0.167)
Covariates	Yes	Yes	Yes	Yes
Credit Scores FEs	Yes	Yes	Yes	Yes
Loan Life Cycle FEs	Yes	Yes	Yes	Yes
Quarter of Loan Issue FEs	Yes	Yes	Yes	Yes
Joint Significance of Credit Scores FEs	***	***	***	***
R-squared	0.012	0.012	0.011	0.011
N	80,529	80,529	80,529	80,529

Table 5 reports OLS estimates of the baseline EMH equation (4) extended to test for changes in the quality of pricing over time. The estimates are based on over 80,000 performance quarters of the 7,455 FC loans issued between 2010Q4 and 2015Q1. The dependent variable, *Default*, is set equal to one if loan i defaults during performance quarter t, and zero otherwise. In columns (1) and (2), we use the LGD adjusted borrowing rate, and in column (3) and (4) the LGD adjusted marginal interest rate. Interest rates are interacted in columns (1) and (3) with a linear monthly time trend, and in columns (2) and (4) with year dummies. Year fixed effects are excluded to avoid collinearity with quarter fixed effects. Additional control variables include *Aggregate Weekly Borrowing, Industry Asset Beta*, *Aggregate Weekly Autobid*, *Loan Level Autobid*, and *Early Termination*. All variables are defined as in Table (5). We also include FEs for the credit scores, the quarter when the loan was issued, and the four stages of the loan's life cycle. Heteroscedasticity robust standard errors, clustered at the borrower level, are reported in brackets. One, two or three starts denote significance at the 10%, 5% or 1% level, respectively, including an F test for the joint significance of the credit score FEs.

Table 8: Dynamics of Investors' Characteristics, Strategies and Performance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	All investors			Active Investors (i.e., Exlucding 100% Autobid investors)							
								Share of active portfolio in			
	Investors	Total Wealth	Autobid Allocation	Investors	Active Portfolio	Loan Stake	Return on Active Portfolio	Score A Loans	Score C Loans	Bid to Marginal Rate Spread	Auctions
	N	mean, £	mean, £	N	mean, £	mean, £	mean, %	mean, %	mean, %	mean, %	N
						Year: 2011					
Less than £100	291	49	32	113	44	28	8.12	0.59	0.03	0.48	
£100 to £500	649	240	144	317	196	59	8.15	0.61	0.04	0.45	
£500 to £1k	303	712	418	170	524	97	8.23	0.58	0.04	0.38	
£1k to £5k	454	2,349	1,203	292	1,782	183	8.33	0.57	0.06	0.33	107
£5k to £10k	106	6,904	3,417	69	5,351	363	8.42	0.53	0.06	0.26	
£10k to £50k	91	19,941	8,561	66	15,674	636	8.38	0.52	0.07	0.22	
More than £50k	5	86,681	8,206	4	83,092	2,124	8.68	0.56	0.08	0.19	
						Year: 2012					
Less than £100	1,167	48	30	427	47	27	8.89	0.38	0.29	0.69	
£100 to £500	1,939	235	125	1,017	206	55	8.89	0.4	0.27	0.63	
£500 to £1k	902	707	402	492	554	94	9.02	0.39	0.29	0.54	
£1k to £5k	1,297	2,212	1,019	854	1,800	180	9.18	0.36	0.31	0.47	219
£5k to £10k	272	7,054	3,123	186	5,761	322	9.36	0.36	0.32	0.39	
£10k to £50k	222	19,523	7,673	160	16,372	618	9.43	0.34	0.34	0.35	
More than £50k	22	98,047	10,052	20	98,147	2,181	9.96	0.3	0.36	0.23	
						Year: 2013					
Less than £100	2,637	46	27	906	47	27	8.12	0.39	0.36	0.29	
£100 to £500	4,590	232	131	2,060	215	54	8.2	0.38	0.35	0.26	
£500 to £1k	1,968	707	410	968	591	92	8.25	0.39	0.34	0.23	
£1k to £5k	3,215	2,230	1,320	1,632	1,775	157	8.38	0.38	0.34	0.21	547
£5k to £10k	672	7,038	4,406	332	5,297	281	8.47	0.39	0.32	0.22	
£10k to £50k	457	19,276	10,609	248	15,931	569	8.69	0.36	0.33	0.21	
More than £50k	46	98,452	38,430	31	90,276	1,713	9.26	0.31	0.32	0.18	
						Year: 2014					
Less than £100	4,686	45	32	840	50	26	10.11	0.47	0.18	0.7	
£100 to £500	7,041	237	158	2,141	228	47	10.15	0.48	0.17	0.63	
£500 to £1k	3,092	708	485	1,045	622	77	10.19	0.5	0.17	0.6	
£1k to £5k	4,626	2,162	1,401	1,835	1,865	127	10.37	0.49	0.17	0.53	865
£5k to £10k	943	7,009	4,530	395	5,840	237	10.63	0.48	0.18	0.45	
£10k to £50k	702	19,182	11,746	317	16,409	485	10.89	0.47	0.19	0.42	
More than £50k	91	157,959	52,345	48	198,614	1,663	11.77	0.43	0.22	0.37	

Descriptive statistics on a population of investor quarters. An investor quarter is a portfolio of loans assembled by an individual investor over a calendar (three-months) quarter. Total Wealth is sorted to size groups of less than £100, £100-500 etc.. Column (2) reports the mean size of that portfolio. Column (3) reports the mean amount allocated algorithmically through autobid. Discarding investors who allocate 100% of their investment via the autobid, we are left with Active Investors (who still use the autobid to *some extent*); column (5) reports their mean non-autobid investment. Column (6) reports the mean amount invested in a single loan within Active Portfolios. Column (7) reports the weighted average of the realized returns on the loans in the active portfolio. Realized returns are the submitted rates of return multiplied by the realized recovery; only mature loans are included. Columns (8) and (9) report the share of the active portfolio invested in loans with credit scores A and C, respectively. Column (10) reports the weighted average spread between between the marginal closing rate and the submitted rate of each order. Columns (1), (4) and (11) report the average number of investors and auctions per quarter within each year.

Table 9: Realized Returns on Active portfolios Against Investor characteristics

	(1)	(2)	(3)	(4)
Log Active	0.100***	0.084***	0.071***	0.056***
	(0.006)	(0.004)	(0.004)	(0.005)
Position/Active	0.015	-0.013	0.054***	0.089***
	(0.029)	(0.021)	(0.020)	(0.024)
Rating: A		-1.372***	-1.366***	-1.226***
		(0.018)	(0.017)	(0.020)
Rating: C		0.881***	0.854***	0.764***
		(0.018)	(0.018)	(0.020)
Rating: D		2.788***	2.777***	2.558***
		(0.035)	(0.033)	(0.037)
Spread			-0.388***	-0.177***
			(0.011)	(0.012)
Aggregate Weekly Borrowing			0.534***	0.412***
			(0.013)	(0.013)
Duration FE	Yes	Yes	Yes	Yes
Quarter of Investment FE	Yes	Yes	Yes	Yes
Investor FE	No	No	No	Yes
R-squared	0.380	0.647	0.672	0.817
N	73,210	73,210	73,210	66,051

The table reports OLS estimates of realized returns on active porfolios against investor characteristics. The sample consists of 73,210 investor-quarter observations based on 13,126 active investors. The dependent variable is *Realized Returns*, defined as the weighted average of realized returns of the loans in the active portfolio during the quarter. *Log Active* is measured as the log total active investment by the lender during the quarter. *Position/Active* is the average amount allocated to each loan, divided by the total active investment during the quarter. *Aggregate Weekly Borrowing*, is computed as the share of investment in auctions with above-median *Aggregate Weekly Borrowing* during the quarter. *Spread* is measured as the weighted average difference between the interest rate of accepted bids and the marginal closing rate of the auction. We include shares for the amount invested in each credit score and omit the B category as the benchmark. The specifications also include fixed effects for the investor (*Investor FE*), the quarter when the returns realized (*Quarter of Investment FE*), and the number of quarters the investor has been active on the marketplace (*Duration FE*). Heteroscedasticity robust standard errors, clustered at the investor level, are reported in brackets. One, two or three starts denote significance at the 10%, 5% or 1% level, respectively, including an F test for the joint significance of the credit score FEs.

Table 10: Large Investors and the EMH Regression

	Borrowing Rate	Marginal Rate
	(1)	(2)
Borrowing Rate	0.642**	
	(0.283)	
Marginal Rate		0.433*
		(0.239)
Aggregate Large Investor Funding	0.086**	0.073**
	(0.035)	(0.032)
Aggregate Bot Funding	-0.002	-0.004
	(0.032)	(0.028)
Rate*Aggregate Large Investor	-1.481**	-1.208**
	(0.593)	(0.521)
Rate*Aggregate Bot Funding	-0.274	-0.230
	(0.511)	(0.426)
Covariates	Yes	Yes
Credit Score FEs	Yes	Yes
Loan Life Cycle FEs	Yes	Yes
Quarter of Loan Issuance FE	Yes	Yes
Joint Significance of Credit Scores FEs	***	***
R-squared	0.011	0.011
N	79,521	79,521

The table reports OLS estimates of the base line equation (4) using over 79,000 performance quarters of the 7,354 FC loans issued between 2010Q4 and 2015Q1. The dependent variable, *Default*, is set equal to one if loan i defaults during performance quarter t, and zero otherwise. *Aggregate Large Investor* is the weighted average of large-investors funding across all the auctions open during the week that theloan i auction is open. Note that the lower sample size with respect to the previous specifications is due to incomplete auction data on investor composition in the last quarter of our sample period. In column (1), we use the LGD adjusted borrowing rate, and in column (2) the LGD adjusted marginal interest rate. Additional control variables include *Aggregate Weekly Borrowing, Industry Asset Beta*, *Aggregate Weekly Autobid*, *Loan Level Autobid*, and *Early Termination*. All variables are defined as in Table (5). We also include FEs for the credit scores, the quarter when the loan was issued, and the four stages of the loan's life cycle. Heteroscedasticity robust standard errors, clustered at the borrower level, are reported in brackets. One, two or three starts denote significance at the 10%, 5% or 1% level, respectively, including an F test for the joint significance of the credit score FEs..

Table 11: Estimation Methods

	Ordinary Least Squares (1)	Logit (2)	Probit (3)	Cox (4)
Average Interest Rate	0.333***	0.328***	0.324***	1.353***
	(0.067)	(0.057)	(0.057)	(0.062)
Industry Asset Beta	0.007**	0.008**	0.008**	2.53**
	(0.003)	(0.003)	(0.003)	(1.027)
Covariates	Yes	Yes	Yes	Yes
Credit Score FEs	Yes	Yes	Yes	Yes
Loan Life Cycle FEs	Yes	Yes	Yes	No
Quarter of Loan Issuance FEs	Yes	Yes	Yes	Yes
Joint Significance of Credit Scores FEs	***	***	***	***
R-squared (Pseudo)	0.011	0.027	0.027	.015
N	80529	80529	80529	7455

	Ordinary Least Squares (1)	Logit (2)	Probit (3)	Cox (4)
Marginal Rate	0.179***	0.175***	0.176***	1.178***
	(0.038)	(0.033)	(0.034)	(0.033)
Industry Asset Beta	0.007**	0.008**	0.008**	2.532**
	(0.003)	(0.003)	(0.003)	(1.028)
Covariates	Yes	Yes	Yes	Yes
Credit Score FEs	Yes	Yes	Yes	Yes
Loan Life Cycle FEs	Yes	Yes	Yes	No
Quarter of Loan Issuance FEs	Yes	Yes	Yes	Yes
Joint Significance of Credit Scores FEs	***	***	***	***
R-squared (Pseudo)	0.011	0.026	0.026	.015
N	80,529	80,529	80,529	7,455

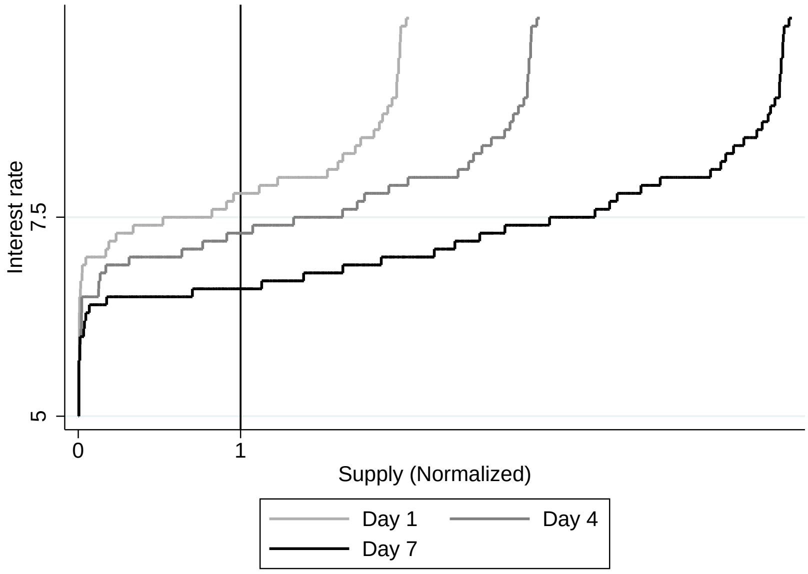
The table reports robustness checks of OLS estimates of the baseline equation (4) using over $80,000$ performance quarters of the 7,455 FC loans issued between 2010Q4 and 2015Q1. In columns (1) to (3), the sample is constituted by the panel of loan performance quarters used in the main analysis. The dependent variable, $Default$, is set equal to one if loan i defaults during performance quarter t, and zero otherwise. In column (1), the specification is estimated using OLS as in Table (5), in column (2) using a Logit model, and in column (3) using a Probit model. In columns (2) and (3), marginal effects are reported at the mean of covariates. In column (4), the sample is collapsed at the loan level and estimated using a Cox Proportional Hazard Model. For the dependent variable $Default$, defined at the loan level at the end of our performance data, survival time is expressed annually, andthe columns report hazard ratios. In the upper panel, we use the LGD adjusted borrowing rate, and in the lower panel the LGD adjusted marginal interest rate. Additional control variables include *Aggregate Weekly Borrowing, Industry Asset Beta*, *Aggregate Weekly Autobid*, *Loan Level Autobid*, and *Early Termination*. All variables are defined as in Table (5). We also include FEs for the credit scores, the quarter when the loan was issued, and the four stages of the loan's life cycle. Heteroscedasticity robust standard errors, clustered at the borrower level, are reported in brackets. One, two or three starts denote significance at the 10%, 5% or 1% level, respectively, including an F test for the joint significance of the credit score FEs.

Table 12: Repeat Borrowers and Early Termination

	(1)	(2)	(3)	(4)	(5)	(6)
Borrowing Rate	0.269***	0.310***	0.218***			
	(0.071)	(0.105)	(0.072)			
Marginal Rate				0.133***	0.103*	0.098**
				(0.043)	(0.057)	(0.045)
Repeat Borrower	-0.001			0.002		
	(0.005)			(0.004)		
Rate*Repeat Borrower	0.031			-0.003		
	(0.088)			(0.065)		
Borrower Freq.		0.003			-0.001	
		(0.004)			(0.004)	
Rate*Borrower Freq.		-0.029			0.037	
		(0.063)			(0.059)	
Rate*Early Termination			0.134*			0.072
			(0.071)			(0.053)
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Credit Score FEs	Yes	Yes	Yes	Yes	Yes	Yes
Loan Life Cycle FEs	Yes	Yes	Yes	Yes	Yes	Yes
Quarter of Loan Issuance FEs	Yes	Yes	Yes	Yes	Yes	Yes
Joint Significance of Credit Scores FEs	***	***	***	***	***	***
R-squared	0.011	0.011	0.011	0.011	0.011	0.011
N	80,529	80,529	80,529	80,529	80,529	80,529

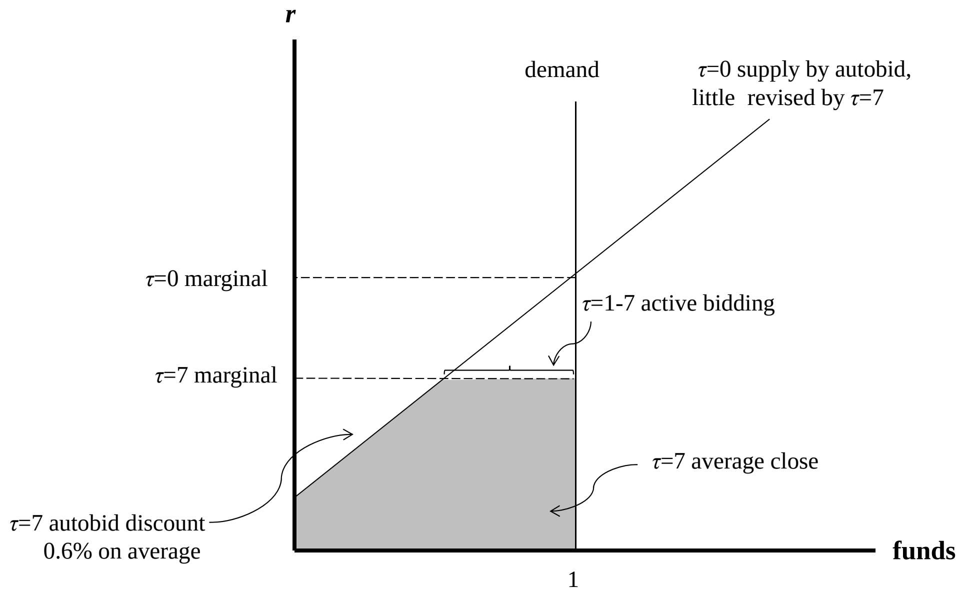
The table reports robustness checks of OLS estimates of the baseline equation (4) using over 80,000 performance quarters of the 7,455 FC loans issued between 2010Q4 and 2015Q1. The dependent variable, $Default$, is set equal to one if loan i defaults during performance quarter t, and zero otherwise. *Repeat Borrower* is a dummy variable that receives a value of one if the loan is taken up by a returning borrower on the platform, and zero otherwise. *Borrower Freq.* is the number of loans taken up previously by the borrower on the platform. In columns (1) to (3), we use the LGD adjusted borrowing rate, and in columns (4) to (6) the LGD adjusted marginal interest rate. Additional control variables include *Aggregate Weekly Borrowing, Industry Asset Beta*, *Aggregate Weekly Autobid*, *Loan Level Autobid*, and *Early Termination*. All variables are defined as in Table (5). We also include FEs for the credit scores, the quarter when the loan was issued, and the four stages of the loan's life cycle. Heteroscedasticity robust standard errors, clustered at the borrower level, are reported in brackets. One, two or three starts denote significance at the 10%, 5% or 1% level, respectively, including an F test for the joint significance of the credit score FEs.

Figure 1: Supply curves, Normalized by Loan Size, 24, 96 and 168 Hours After Open



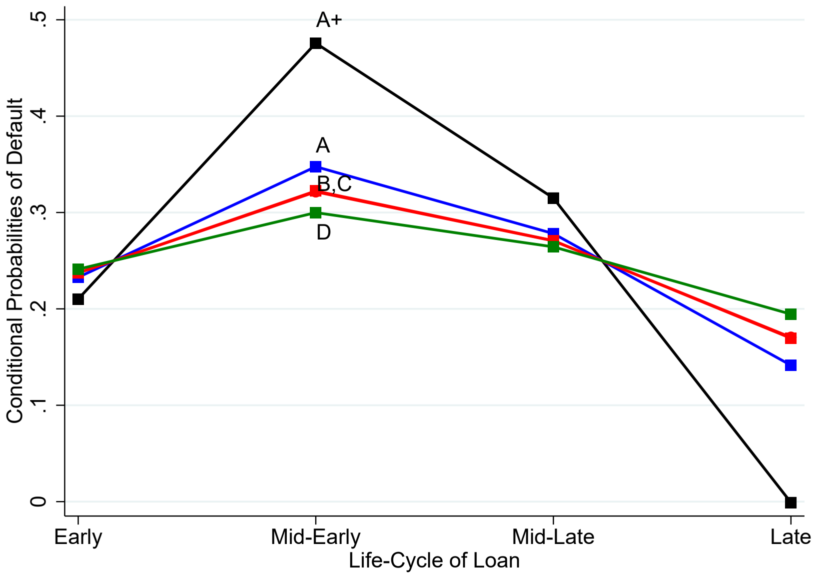
Supply curves for an arbitrary A-scored loan, 24, 96 and 168 hours from the open. Orders, submitted up to the respective points in auction time, are normalized by loan amount, £15,000 in this case, are sorted and then plotted against the submitted interest rate. By construction, demand is a vertical line at one unit.

Figure 2: FC Auction Design



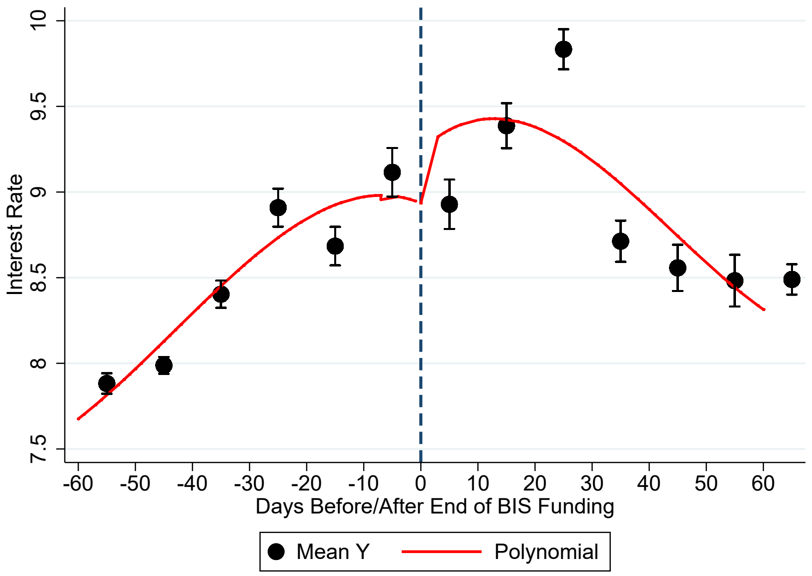
The diagram provides a stylized exposition of FC's platform design. Table 2 documents that autobid activity is largely concentrated on day 1, that it submits orders at different prices so as to form an upwards sloping supply curve. Normalizing all volumes by the size of the loan, the price insensitive demand curve is a vertical line at one unit. Orders submitted up to any point in auction time, by passive investors (via the autobid) as well as by active investors, can be crossed against demand to derive the marginal interest rate for that point in time. The shaded area represent the borrowing rate - a weighted average of all accepted orders. Since the order book is open, active investors who want to compete for an allocation have an incentive to undercut the marginal rate by just a few basis point. For more detail see Section 2.

Figure 3: Conditional Default Probabilities Over the Loan's Life



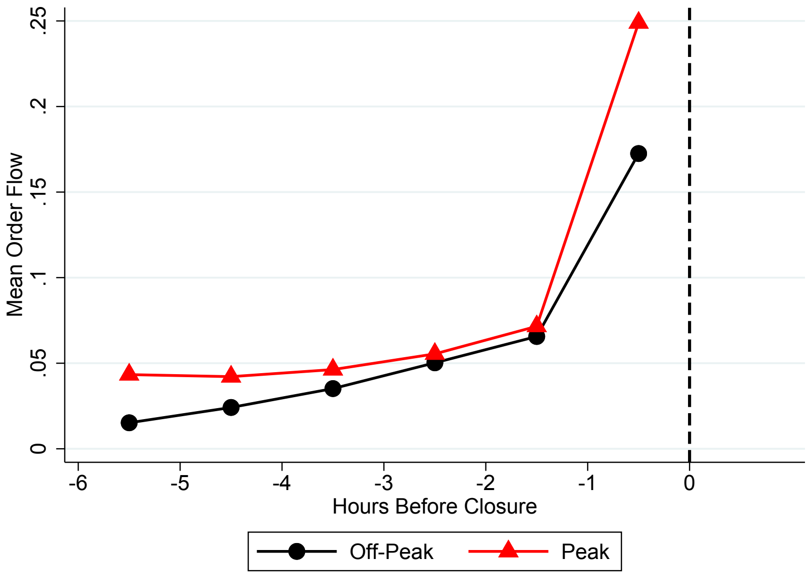
Conditional default probabilities, given the stage of the loan's life cycle (Early, Mid-Early, Mid-Late and Late) and its credit score, are derived from the unconditional probabilities as estimated in Table 3, using Bayes' Law. For more detail see the discussion in relation to 4 in the body text.

Figure 4: Interest Rates 60 Days Before and After BIS Funding



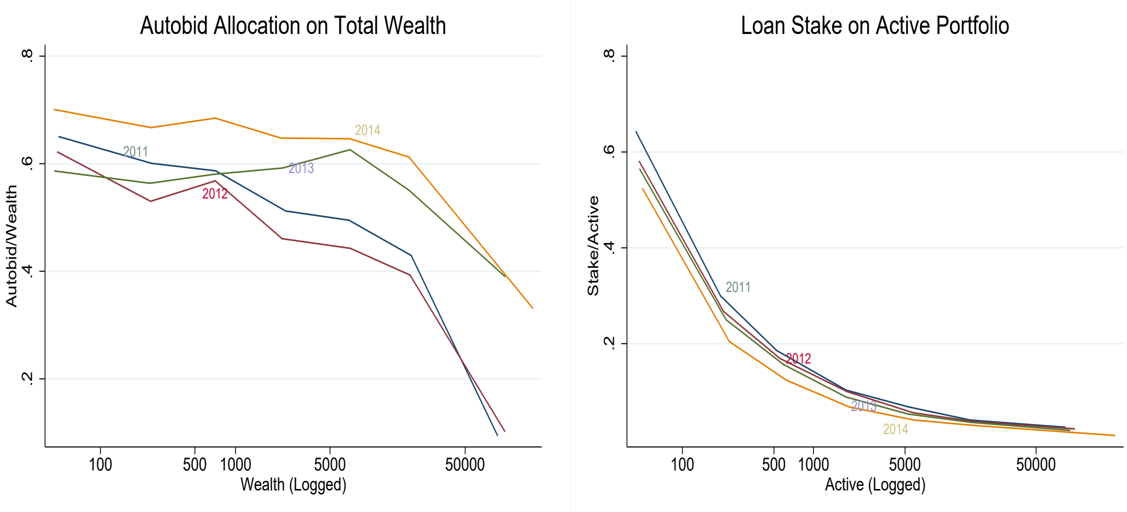
The figure plots borrowing rates of auctions held over a period starting 60 days before the BIS termination date (February 28, 2014) and ending 60 days after. During that time window 1,292 auctions were held with 653 held before the termination, and 639 afterwards. We divide the sample into mutually exclusive bins of ten days. For each bin, we compute the average and 90% confidence interval of the borrowing rate, and plot these values at the bin's mid-point. The red line plots a third-order polynomial separately before and after the end of the program.

Figure 5: Mean Inflow of Funds, According to Closing Hour



The figure plots hourly mean order flow normalized by loan size, during the six hours prior to the closing of the auction. The sample is constituted of auctions that run the full seven days. *Off-Peak* denotes auctions that closed outside of the 4PM to 7PM window, while *Peak* denotes auctions that closed between 4PM to 7PM.

Figure 6: Investor Characteristics Over Time



The left panel plots the ratio between the Autobid Allocation and Total Wealth against logged Total Wealth. The right panel plots the ratio between the Loan Stake and Active Portfolio against logged Active portfolio.

Online Appendix A

Table 13: Balancing Tests of Borrower Characteristics Around Liquidity Events

	Panel A: Peak and Off-Peak Closing Hour					
	Off Peak	Peak	Difference	Standard Error	N	P-Value
Rating: A+	0.123	0.130	0.00723	0.00811	6715	0.373
Rating: A	0.309	0.306	-0.00304	0.0113	6715	0.787
Rating: B	0.273	0.267	-0.00644	0.0108	6715	0.552
Rating: C	0.234	0.229	-0.00536	0.0103	6715	0.603
Rating: D	0.0608	0.0684	0.00761	0.00600	6715	0.205
Activity: IT	0.0695	0.0713	0.00183	0.00625	6715	0.770
Activity: Manufacturing	0.138	0.126	-0.0122	0.00825	6715	0.138
Geography: London	0.126	0.133	0.00717	0.00819	6715	0.382
Geography: South East	0.214	0.229	0.0151	0.0101	6715	0.137
	Panel B: BIS Funding Stop, 60 Day Window					
	Funding Stop	Funding	Difference	Standard Error	N	P-Value
Rating: A+	0.0803	0.0740	-0.00630	0.0150	1270	0.674
Rating: A	0.296	0.302	0.00630	0.0257	1270	0.807
Rating: B	0.291	0.294	0.00315	0.0256	1270	0.902
Rating: C	0.216	0.222	0.00630	0.0232	1270	0.786
Rating: D	0.117	0.107	-0.00945	0.0177	1270	0.594
Activity: IT	0.0567	0.0646	0.00787	0.0134	1270	0.557
Activity: Manufacturing	0.124	0.102	-0.0220	0.0178	1270	0.216
Geography: London	0.131	0.135	0.00472	0.0191	1270	0.804
Geography: South East	0.211	0.211	0	0.0229	1270	1

The table reports the mean of borrower characteristics according to two liquidity events. In Panel A, we consider auctions according to their closing hour. *Peak* refers to auctions closing between 3pm and 7pm, while *Off Peak* refers to all other auctions. In Panel B, we consider auctions within 60 days of the BIS funding stop. *Funding Stop* refers to auctions closing in the 60 days following the end of the BIS allocated funding. *Funding* refers to auctions closing in the 60 days before the BIS funding ran out.

Online Appendix B

In Section 4.3.1 we report that (at most) 20% of the price variance is due to information; the rest is just “noise”. In this appendix we provide the technical derivation of this figure. To begin with, its worth explaining why, though the EMH regression allows us to estimate the default probability, the R-square in this regression does not measure the quality of the estimation. For the simple reason that the EMH regression predicts the *event* of default;

the *probability* of default is then derived from that estimation. For that very reason, the R-square in the EMH regression is that small: even if we derive a perfect estimator of the default probability of, say, 2%, a regression where the dependent variable is a default dummy still generates 98 "errors" per 100 observations.

To derive loan i's probability of default we use the fitted value, $\widehat{\pi}_i$, of the estimated EMH regression. Intuitively, we can then regress the interest rate on the predicted default probability, $\widehat{\pi}$:

$$r_i = \alpha + \beta\widehat{\pi}_i + \varepsilon_i, \tag{7}$$

and use the R-squared from that equation to measure r's information content: the share of the variance explained by the predicted default probability. The prediction is based on public information such as credit scores, but also private information priced in during the price discovery process.

Unfortunately, this is not a perfect solution. The reason is that in equation (7), the lending rate, r, is the dependent variable, but also a covariate in the EMH equation used in the derivation of $\widehat{\pi}$. We can show, analytically and by simulations, that the result is biased: it over-states the amount of information (and understates the amount of noise) contained in the price. Intuitively, this is because any random shock in the default probability, $\widehat{\pi}$, is passed through to r, moving both in the same direction to create a false appearance of information "explaining" the lending rate. It is still the case that our method provides a useful upper bound to the amount of information contained in the lending rate. Another analytical result is that the more saturated the EMH equation with signals that predict the default probability, the less biased is the upper bound for the information content.

Table 14 presents the results. Moving down column (1), we saturate the EMH equation that produces the $\widehat{\pi}$ predictor with additional covariates, thereby improving its quality. In line with the argument above, that reduces the bias in the measured quality of price and tightens the upper bound. The bottom-line conclusion is that of the information content of the price is, at most 20%, of its variance.

Column (2) of Table 14 pushes the argument one step further by augmenting equation (7) with the credit scores. Remember that the credit scores are already included in the

Table 14: Upper-bound share of price variance explained by default probabilities

	Estimated from	Plus credit scores
	(1)	(2)
The above plus price	0.23	0.48
The above plus liquidity variables	0.21	0.48
The above plus early termination	0.20	0.48

Regression R-squares using 7,455 auctions. The dependent variable is the borrowing rate. In column (1) the only independent variable is the loans' estimated default probability, to which we add, in column (2), the credit scores. Default probabilities are fitted values from the Tables 5 and 6 EMH-regressions (using 80,529 performance quarters) augmenting credit scores (top row) with the borrowing rate and industry asset betas (second row) Table-6 liquidity measures (third row) and the early-termination dummy (bottom row); duration FEs excluded.

EMH regression that generated $\widehat{\pi}$. It follows that any credit-scores information that was relevant to the prediction of the default probability was already extracted and priced in. Whatever is left, is irrelevant "noise" that should not be correlated with r. In fact it is, pushing the R-squared of equation (7) up from 20% to 48%. That is, too much of the credit score noise was priced in, by a platform that over estimated the quality of its own credit scores.