International Review of Financial Analysis 71 (2020) 101542

Contents lists available at ScienceDirect

International Review of Financial Analysis

journal homepage: www.elsevier.com/locate/irfa

Institutional investment in online business lending markets

Mark Cummins[a], Ciarán Mac an Bhaird[b,*], Pierangleo Rosati[a], Theo Lynn[a]

[a] DCU Business School, Irish Institute of Digital Business, Dublin City University, Glasnevin, Dublin 9, Ireland
[b] Fiontar (Enterprise), Faculty of Humanities & Social Sciences, Irish Institute of Digital Business, Dublin City University, All Hallows, Drumcondra, Dublin 9, Ireland

ARTICLE INFO

JEL classification:
D 53
D81
D82
G11
G21
G23

Keywords:
Institutional investment
Online business lending
Default risk minimisation
Lending efficiency

ABSTRACT

We provide new insights into the business lending decisions of institutional investors in online credit markets by benchmarking their lending performance against that of retail investors.We find superior performance for loans financed by institutional investors, although large sized retail investor groups achieve equivalent performance. Lending decisions of institutional investors are not default risk minimising, and we quantify lending inefficiencies. From a platform perspective, we show that (i) the platform-administered loan allocation process is not biased in favour of institutional investors, (ii) institutional participation in the retail marketplace is not a distorting factor in loan performance, and (iii) the platform's move to a fixed rate system had detrimental effects on loan outcomes for institutional investors. The superior loan performance achieved by institutional investors is confined to the auction period, when institutional investors had autonomy over setting interest rates.

1. Introduction

Alleviation of information asymmetries is the greatest challenge in lending to private enterprise, and this is exacerbated in the small business environment by the relatively high cost of compiling information on individual firms, the limited and fragmented market for this information, and difficulties in signalling to the market (Boot, 2000; Mac An Bhaird & Lucey, 2010). Such information asymmetries are most pronounced for start-ups (Cassar, 2004; Mac an Bhaird & Lynn, 2015), and lead to borrower discouragement from conventional intermediated debt among younger, smaller and higher risk firms (Freel, Carter, Tagg, & Mason, 2012; Kon & Storey, 2003; Mac an Bhaird, Sanchez-Vidal, & Lucey, 2016). Firms are restricted in their access to investment finance due to the lack of diversification in private debt markets (Mac an Bhaird, 2010). Additionally, because of the pro-cyclical supply of debt finance, SMEs experience a reduction in credit availability in the wake of banking and financial crises.

Online credit markets enable borrowers to seek funding from lenders on a direct peer-to-peer basis using web-based platforms, without the traditional intermediation function provided by financial institutions. The rapid embrace of peer-to-peer lending in recent years indicates significant demand from firms and investors alike. Studies on online lending markets focus on equity based or debt based channels (e.g. Pierrakis & Collins, 2013). Equity based crowdfunding studies examine the issue from the perspective of the firm (Hornuf & Schwienbacher, 2017; Signori & Vismara, 2018), and the investor (Günther, Johan, & Schweizer, 2018; Hornuf & Schwienbacher, 2018). Debt based crowdfunding studies generally examine personal lending (e.g. Morse, 2015; Tang, 2019). A particular focus has been the use of hard and soft information in mitigating informational asymmetry (Dorfleitner et al., 2016; Ge, Feng, Gu, & Zhang, 2017). In a peer-to-business lending setting, Dorfleitner, Hornuf, and Weber (2018) examine a proposed 'default shock bias' effect, showing that subsequent to suffering loan defaults investors make decisions that lead to a worsening of the risk-return profile of investor portfolios.

While online lending markets were initially dominated by retail investors, a significant emerging trend is the increasing participation of institutional investors, including commercial and development banks, non-bank financial institutions and asset management firms. Many peer-to-peer lending platforms have opened up their marketplaces to institutional investors, including the largest business lending platform in the UK, Funding Circle, which is the basis of this study. From a platform perspective, we investigate the channels through which institutional investors provide business lending, while, from an investment perspective, we investigate loan outcomes based on institutional investors' appraisal of loan opportunities. Given this context, we seek to

* Corresponding author.
E-mail addresses: Mark.Cummins@dcu.ie (M. Cummins), Ciaran.Macanbhaird@dcu.ie (C. Mac an Bhaird), Pierangelo.Rosati@dcu.ie (P. Rosati), Theo.Lynn@dcu.ie (T. Lynn).

https://doi.org/10.1016/j.irfa.2020.101542
Received 7 February 2020; Received in revised form 15 May 2020; Accepted 27 May 2020
Available online 15 June 2020

add to a growing literature on this topic (Boubaker & Nguyen, 2019), by presenting a number of new insights into institutional investment in online business lending markets.

Institutional investment in venture capital (e.g. Bruton & Ahlstrom, 2003; Bruton, Fried, & Manigart, 2005) and private equity (e.g. Nielsen, 2007), whether direct or indirect (via funds), is well studied. In contrast, not much is known of institutional investment in crowdfunding markets. There is, however, an emerging literature addressing this topic, and it is to this discourse that we add insights from the perspective of online business lending.

Vallee and Zeng (2019) show that sophisticated investors systematically outperform unsophisticated investors in personal lending and this is explained by superior information processing. By contrast, Lin, Sias, and Wei (2017) find that institutional investors do not outperform retail investors, despite having larger and more diverse portfolios and avoiding home bias. We differ from Vallee and Zeng (2019) and Lin et al. (2017) in focusing on online business lending, which exposes lenders to higher likelihood of default than personal unsecured lending (Mach, Carter, & Slattery, 2014). Our study closely aligns with Mohammadi and Shafi (2017), although their testing does not take account of institutional investment in the retail marketplace on the Funding Circle platform.

Our study differs in a number of important respects, providing a number of contributions as follows. Firstly, we provide an important clarification on the participation of institutional investors in an online credit market, which requires careful consideration in the testing design. We account for two channels of institutional investor participation: the wholesale market place (wholeloans) where a single institutional investor provides the full funding for a loan; and the retail market place (partloans), consisting predominantly of retail investors, although one particular institutional investor has invested in all retail loans over our sample period of between 5 and 20% of the loan values. We explore the potential distortion from this institutional participation in the retail marketplace. In a series of tests, our results are consistent when the analysis is repeated on sets of partloans selected to either dilute or remove this potential distortion. We find that wholeloans outperform partloans, earning superior returns across likelihood of loan default and loan rates, indicating a superior ability of institutional investors to screen loan applications. We also find that institutional investors achieve higher realised returns upon loan repayment. However, we find that losses are substantially greater for wholeloans upon default.

Secondly, we consider the Funding Circle loan allocation process and explore if there is any bias in favour of institutional investors. Applying the Wald-Wolfowitz runs test of randomness (Wald & Wolfowitz, 1940), we find no evidence that the loan allocation process administered by the lending platform is non-random and biased. This assertion holds when tested on loans across each of the credit rating bands as assigned by the lending platform.

As a third contribution, we use a testing approach drawn from the banking literature (Greene, 1998; Jacobson & Roszbach, 2003), which allows us to examine lending efficiency and whether investment decisions of institutional investors are default risk minimising. This approach also provides a more complete mitigation of sample selection bias. We find no correlation between ex-ante loan granting decisions and ex-post default outcomes. While this validates a univariate estimation of default likelihood, it establishes that lending decisions of institutional investors are not consistent with an objective of default risk minimisation and, hence, there is a degree of lending inefficiency in the wholesale market. Exploiting our default modelling framework, we provide a number of new insights by examining lending efficiency through the lens of value-at-risk (VaR), following the simulation approach of Jacobson and Roszbach (2003). We find that institutional investors could have made more efficient lending decisions, manifesting as reduced value-at-risk exposure. We show that this lending inefficiency is consistent with a naïve strategy of randomly rejecting loans.

Fourthly, we consider the effect of investor group size in the retail marketplace, to ascertain whether institutional investors systematically outperform *small*, *medium* and *large* sized retail investor groups. Literature suggests that group size is positively related to better outcomes in group decision making, possibly being a proxy for group diversity and group experience (Bassamboo, Cui, & Moreno, 2015; Surowiecki, 2005). Using a novel tercile-based segmentation of the retail investor base, we find that the performance of wholeloans financed by institutions is superior to partloans financed by small- and medium-sized retail investor groups only, with no outperformance in the case of partloans financed by large-sized retail investor groups. We conclude that large-sized groups of retail investors appear to perform equivalently to institutional investors.

Finally, we present new evidence on the performance of loans funded by institutional investors through an important structural change in the lending platform, whereby the determination of interest rates changed from an auction-based system to one where interest rates are set by the lending platform. This change in systems has important implications for institutional investors, moving them from interest rate setters to interest rate takers. We document the performance of wholeloans between the two interest rate setting periods, finding that institutional investors show outperformance only in the auction period. The loss of autonomy from switching to the fixed rate system appears to have been detrimental for institutional investors.

The remainder of the paper is organised as follows. Section 2 describes the loan data and presents the technical details of our proposed testing. We also discuss important sample selection bias issues and how we address these within the testing framework. Section 3 provides a comprehensive discussion of our main findings, in addition to providing the results of the robustness checks that we conduct. Section 4 provides our analysis of lending efficiency. In Section 5, we perform our analysis of retail investor group size effects, while Section 6 investigates the platform's transition from an auction-based system to a fixed rate system of setting interest rates. Concluding remarks are made in Section 7.

2. Data and testing methodology

Our dataset consists of all loans provided through one of the largest global online lending platforms between 2014 and 2016 inclusive. Funding Circle facilitates lending to businesses. The platform initially launched in the UK in 2010 and subsequently set up operations in the US in 2013 and Germany and the Netherlands in 2015. While Funding Circle acquired market exposure in Spain, in addition to Germany and the Netherlands, through its acquisition of Zencap in 2015, it subsequently exited the Spanish market in 2016.

We focus on the UK lending market, which consists of £5.6bn in loans advanced to 56,000 businesses worldwide between its launch in 2010 and the end of November 2018. The lending platform charges the borrower an origination fee, and charges investors an annual servicing fee of 1% of the outstanding principal. Most loans are unsecured, although the platform facilitates secured loans, which constitutes a smaller segment of the market. The platform has a base of over 50,000 investors, consisting mostly of individuals, along with a number of institutional investors, including the European Investment Bank, the British Business Bank, as well as international asset management firms such as Aegon and KLS. Institutional investors were granted a dedicated wholesale marketplace on 6th May 2014. The wholesale market is confined to institutional investors, although institutional investment in the retail market is possible. Indeed, as confirmed in re with the lending platform, one large institutional investor passively invests in the retail market on a consistent basis, having funded between 5 and 20% of the value of *all* retail loans over our sample period. The wholesale market differs from the retail market in that for any loan financed, a single institution provides the full funding. Such loans are designated as ‘wholeloans’. By contrast, in the retail market, loans are financed by a number of, predominantly, retail investors. Such loans are designated 'partloans'.

Investment in wholeloans in the wholesale marketplace can be made in a number of different ways. Firstly, a direct lending route, whereby institutional investors purchase loans directly onto their own balances sheets or through special purpose vehicles. Secondly, an indirect lending route, whereby institutional investors provide funds through the Funding Circle SME Income Fund that is quoted on the London Stock Exchange, an investment company that passively invests in Funding Circle loans and pays dividends to investors. Thirdly, an additional indirect lending route, whereby institutional investors provide funds through an umbrella ICAV (Irish Collective Asset Management Vehicle) that facilitates the launch of private sub-funds of Funding Circle loans.

Our focus is on comparing the lending performance of wholeloans and partloans, from which we can, with the appropriate transition in interpretation, evaluate the decision making of institutional investors. We therefore focus on loans with observed outcomes, while we exclude outstanding loans for which outcomes are not realised. We do this as we consider performance jointly in respect of loan screening ability and realised loan payoffs. Our final dataset consists of 3791 loans in total; of which 3201 were repaid in full and 590 defaulted. Loans are assigned to the two marketplaces in a randomised fashion; a processed administered by the lending platform. The retail marketplace funded all 1946 loans it was initially offered. In contrast, institutional investors in the wholesale marketplace were selective, choosing to fund only 1489 of the 1845 loans it was offered. The 356 loans that were rejected by institutional investors were then subsequently offered to retail investors in a recycling process automated by the lending platform. All such recycled loans were funded in full in the retail marketplace.

Our key variables of interest are defined in Table I and descriptive statistics are provided in Table II. Analysis of investments, returns and defaults by loan type is presented in Table III. There are notable differences between the profile and performance of loans funded in the wholesale and retail markets, as evidenced by analysing across credit ratings throughout the investment period. Risk preferences within both marketplaces change over the period. Investors in the retail marketplace increasingly invest in lower risk loans, with half of their total investments in the lowest risk (A+) loans in 2016. Investors in the wholesale marketplace, by contrast, invest relatively less in the lowest risk loans, between 21% and 32% over the period. Such investors also invest relatively more in credit grade B and C loans relative to the retail marketplace, with this differential increasing over time. In 2016, over 31% of the total investments in the wholesale marketplace were in grade B and C loans, as opposed to < 9% in the retail marketplace.

It is notable that institutional investors in the wholesale market earn higher interest rates than investors in the retail marketplace across all credit grades in 2014 and 2015, apart from very slight lower differences in the highest and lowest grades in 2015 (0.02% and 0.16% respectively). In 2016, this was reversed, with investors in the retail marketplace earning higher interest rates across all credit grades apart from grade C (equal) and grade E (0.14% higher). Similar patterns are evident in the realised returns earned on repaid loans, as measured by the internal rate of return (IRR). Wholeloans earn a higher return than partloans when repaid in 2014 and 2015, with the exception of the highest and lowest credit grades in 2015. In 2016 this pattern is partially reversed, as investors in the wholesale market earn a slightly lower rate on the two highest credit grades (0.20% and 0.04% respectively), and credit grade D (0.11%). In respect of defaulted loans, wholeloans experience significantly greater losses than partloans across all credit grades in 2014. By contrast, in 2015 loans funded in the wholesale marketplace record lower losses than loans in the retail marketplace, across credit grades A+, C and E, losing on average 19.21%, 11.8% and 4.36% less. A discernible pattern is that the losses on defaulted wholeloans decline over time vis-a-vis partloans, although average losses on defaulted loans are increasing for both marketplaces.

Table II
Variable summary statistics.

Variable	Obs	Mean	Std. Dev.	Min	Max
Default	3791	0.156	0.363	0	1
LoanAmt	3791	73,016.580	83,313.980	5000.00	650,000.00
log LoanAmt	3791	10.724	0.974	8.52	13.39
IntRate	3791	9.853	1.977	6.00	18.37
Term	3791	36.762	18.737	3.00	60.00
log Age	3791	2.033	0.748	0.00	4.64
Ret	3791	−2.715	30.874	−100.00	79.01
Wholeloan	3791	0.393	0.488	0	1
Recycled	3791	0.094	0.292	0	1
CrBandA+	3791	0.249	0.433	0	1
CrBandA	3791	0.258	0.438	0	1
CrBandB	3791	0.223	0.416	0	1
CrBandC	3791	0.162	0.368	0	1
CrBandD	3791	0.093	0.291	0	1
CrBandE	3791	0.014	0.119	0	1

Variables defined in Table I.

Table I
Key variable definitions.

Variable	Variable type	Description
Default	Binary	Indicates loan default 1 = loan defaulted; 0 = loan repaid.
Term	Continuous	Duration of loan (in months)
log Age	Continuous	Natural log of the age of the borrowing firm (date of incorporation less date of loan origination)
log LoanAmt	Continuous	Natural log of the loan amount
Wholeloan	Binary	Wholeloan investment in the wholesale marketplace 1 = wholeloan; 0 = partloan
Recycled	Binary	Loan rejected by institutional investor in the wholesale marketplace and recycled into the retail marketplace 1 = recycled; 0 = otherwise
IntRate	Percentage	Interest rate payable on the loan
Ret	Percentage	Ex-post return on loan
CrBandA–CrBandE	Dummy	Dummy variables corresponding to the credit band categories: A, B, C, D, E. The reference category is credit band A+
Sector	Dummy	Dummy variable representing sectors: Agriculture, Arts and Entertainment, Automotive, Consumer Services, Education and Training, Health, Leisure, Manufacturing, Professional and Business support, IT and Telecoms, Retail, Transport, Wholesale, Property and Other. The reference category is IT and Telecoms.
Region	Dummy	Dummy variables representing UK regions: South East, South West, North East, North West, East Anglia, Midlands, London, Scotland, Wales, Northern Ireland. The reference category is London.
Purpose	Dummy	Dummy variables representing loan purpose: Expansion, Working Capital, Asset Finance, Property Finance, Debt Refinancing, Other Purpose. The reference category is Property.

Table III
Data descriptive statistics by retail and wholesale marketplaces.

	Retail marketplace			Wholesale marketplace		
Credit grade	2014	2015	2016	2014	2015	2016
Percentage investment						
A+ (very low risk)	31.34%	59.67%	50.53%	21.16%	31.71%	28.67%
A (low risk)	23.68%	23.88%	36.37%	27.79%	27.85%	33.12%
B (below average risk)	19.52%	7.66%	6.22%	28.32%	18.78%	17.34%
C (average risk)	17.20%	4.69%	2.68%	17.19%	13.00%	13.76%
D	8.26%	2.55%	1.90%	5.54%	7.86%	5.28%
E	0.00%	1.55%	2.30%	0.00%	0.80%	1.82%
Average default rates (as % of amount invested in credit band)						
A+ (very low risk)	0.41%	3.13%	0.77%	0.96%	6.22%	1.46%
A (low risk)	4.82%	4.81%	3.73%	0.83%	8.75%	2.40%
B (below average risk)	10.86%	5.78%	1. 81%	2.54%	9.42%	2.03%
C (average risk)	14.69%	5.53%	1.29%	2.88%	14.18%	6.58%
D	13.48%	5.11%	0.97%	3.11%	18.83%	4.25%
E		28.62%	17.90%		24.98%	16.33%
Average loan rates						
A+ (very low risk)	7.75%	7.87%	8.09%	8.26%	7.85%	7.94%
A (low risk)	9.03%	9.18%	9.40%	9.48%	9.26%	9.32%
B (below average risk)	10.07%	9.84%	10.44%	10.53%	10.23%	10.40%
C (average risk)	11.12%	10.99%	11.81%	11.53%	11.36%	11.81%
D	12.52%	12.92%	14.01%	12.98%	13.27%	13.95%
E		18.09%	17.79%		17.93%	17.93%
Average realised returns (IRR; defaulted loans)						
A+ (very low risk)	−42.54%	−83.21%	−99.92%	−78.24%	−64.00%	−96.35%
A (low risk)	−46.73%	−76.17%	−83.76%	−52.83%	−81.00%	−80.71%
B (below average risk)	−41.55%	−81.35%	−99.76%	−57.63%	−85.59%	−99.31%
C (average risk)	−50.86%	−81.58%	−99.23%	−67.58%	−69.78%	−98.88%
D	−41.61%	−84.52%	−90.27%	−53.14%	−85.57%	−99.89%
E		−88.91%	−99.59%		−84.55%	−99.96%
Average realised returns (IRR; repaid loans)						
A+ (very low risk)	6.96%	7.08%	7.34%	7.36%	7.04%	7.14%
A (low risk)	8.13%	8.53%	8.69%	8.71%	8.54%	8.65%
B (below average risk)	9.31%	9.10%	9.76%	9.89%	9.57%	9.82%
C (average risk)	10.30%	10.33%	11.28%	11.02%	10.86%	11.37%
D	11.99%	12.41%	13.84%	12.53%	13.01%	13.73%
E		18.46%	18.27%		18.45%	18.30%

2.1. Testing methodology

Our study comprises, in line with existing work, an analysis of loan screening ability, through the modelling of (i) loan rates and (ii) loan default probability; and, in an extension of existing work, an analysis of loan payoff, through the modelling of (iii) realised returns upon loan repayment and (iv) realised returns upon loan default. We align our loan performance measures with previous studies. Similar to Emekter, Tu, Jirasakuldech, and Lu (2015) and Miller (2015), we consider probability of loan default and the interest rate charged on approved loans. Consistent with Berkovich (2011), we quantify the payoffs achieved by investors through realised loan returns. We analyse defaulted and repaid loans separately to avoid information loss through averaging effects in the proposed regression analysis. In analysing defaulted loans, we differ from Berkovich (2011) in having access to recovery amounts and so we avoid the need to make a zero recovery rate assumption. This facilitates a more accurate measurement of loan losses. For repaid loans, there are instances where loans are repaid early, leading to differences in ex-post returns and ex-ante loan rates.

We employ the probit modelling framework to investigate potential determinants of the probability of loan default. We use realised loan outcomes, with a specific emphasis on the relationship between ex-ante loan rates and ex-post loan defaults. While alternative default probability modelling techniques are possible, such as Cox proportional hazard regression or stacked logit (DeYoung, Glennon, & Nigro, 2008), such a survival analysis is not deemed necessary as the prediction of default timing is not our particular concern. We seek to establish whether institutional investors in the wholesale marketplace exploit their professional sophistication to show strict dominance in approving loans at higher interest rates than loans approved in the retail marketplace, while not facing a higher frequency of loan defaults. The binary outcome of a loan default event is therefore sufficient. The probit model is specified as follows:

$$prob(Default^i = 1 \mid \Theta) = \Phi(\beta_0 + B'\Theta) \quad (1)$$

where

$$\Theta := \{Term, \log Age, \log LoanAmt, IntRate, Wholeloan, Recycled, CrBandA, CrBandB, CrBandC, CrBandD, CrBandE\}$$

is the vector of predictors, as defined in Table I;

$$\mathrm{B} := \{\beta_{Term}, \beta_{\log Age}, \beta_{\log LoanAmt}, \beta_{IntRate}, \beta_{Wholeloan}, \beta_{Rejected}, \beta_{CrBandA}, \beta_{CrBandB}, \beta_{CrBandC}, \beta_{CrBandD}, \beta_{CrBandE}\}$$

is the vector of regression coefficients; $\Phi()$ is the cumulative standard normal distribution function; and the binary variable $Default^i = 1$ designates a default on loan *i*.[1] In including the credit band dummy variables, we use the A+ credit rating as the reference case. Two variables are worth particular comment. Firstly, the principal variable of interest is the *Wholeloan* dummy variable, which identifies wholeloans funded by institutional investors in the wholesale marketplace. This variable provides insights into differences in the likelihood of default on wholeloans relative to partloans funded in the retail

[1] We suppress the i superscript on the independent variables, here and throughout, for convenience of notation.

marketplace. Secondly, *Recycled* is a dummy variable that is included to capture whether a loan has been recycled. We discuss later the appropriateness of including this variable.

We include interest rate along with the credit band categories in the probit model specification. The latter might seem redundant on the basis that the interest rate should fully incorporate default risk (Adams, Einav, & Levin, 2009). However, the immaturity of online lending markets motivates us to consider the interest rate as a function of the investor's noisy estimate of the probability of default. We therefore include the credit bands in the probit specification on this basis. For robustness, we extend the base probit model specification with time, region, sector and loan purpose fixed effects. Finally, while our database primarily includes unsecured loans, a total of 685 loans are secured. We do not include a control variable for secured loans as the secured loans are exclusively property-based loans. As property-based loans are identified through our loan purpose dummy variable, including a secured loan identifier is redundant.

To investigate the ex-ante perception of loan risk by market participants, we use the interest rate charged on loans. We estimate our model using ordinary least squares (OLS) regression estimation with robust standard errors. The model is formally defined as follows:

$$\begin{aligned} IntRate^i &= \beta_0 + B'\Theta \\ &= \beta_0 + \beta_{Term} Term + \beta_{\log\ Age} \log\ Age + \beta_{\log\ LoanAmt} \log\ LoanAmt \\ &+ \beta_{Wholeloan} Wholeloan + \beta_{\mathrm{Re}jected} \mathrm{Re}\,cycled + \beta_{CrBandA} \\ &\quad CrBandA + \beta_{CrBandB} CrBandB \\ &+ \beta_{CrBandC} CrBandC + \beta_{CrBandD} CrBandD + \beta_{CrBandE} CrBandE + \varepsilon^i \end{aligned} \tag{2}$$

where $IntRate^i$ represents the interest rate charged on loan i. Similar to the probit model, the *Wholeloan* dummy variable in this case facilitates an examination of differences in the interest rate charged on wholeloans funded in the wholesale marketplace relative to partloans funded in the retail marketplace, with the *Recycled* dummy variable providing information on recycled loans. We again extend the base regression specification to take account of potential fixed effects by including time, region, sector and loan purpose dummy variables.

To model realised returns on loans that default/repay, we follow a similar regression structure to the loan rate case. The model is formally defined as follows:

$$\begin{aligned} \mathrm{Re}\,t^i &= \beta_0 + B'\Theta \\ &= \beta_0 + \beta_{Term} Term + \beta_{\log\ Age} \log\ Age + \beta_{\log\ LoanAmt} \log\ LoanAmt \\ &+ \beta_{Wholeloan} Wholeloan + \beta_{\mathrm{Re}jected} \mathrm{Re}\,cycled + \beta_{CrBandA} \\ &\quad CrBandA + \beta_{CrBandB} CrBandB \\ &+ \beta_{CrBandC} CrBandC + \beta_{CrBandD} CrBandD + \beta_{CrBandE} CrBandE + \varepsilon^i \end{aligned} \tag{3}$$

where Ret^i represents the realised return on loan i. Realised returns are estimated using internal rate of return (IRR), under the following assumptions: (i) the monthly repayment on each loan is that agreed when the loan is funded; (ii) if a loan is repaid early, the outstanding principal is repaid on the last payment; and (iii) if a loan defaults and there are recoveries, these are repaid in the month following default. These assumptions are necessary as borrowers can decide to increase or decrease the monthly payment at discretion but this is unknown, while the exact timing of recoveries is also unknown. The assumptions we make provide a conservative estimate of the IRR.

2.2. Addressing sample selection bias

2.2.1. Initial loan application screening

One potential source of sample selection bias comes from the initial loan application screening conducted by the lending platform, which selects loans to be presented to the marketplaces. Three key pillars are set out by the platform as underlying its credit assessment process: (i) policy criteria, which set out eligibility requirements, including a minimum of two years of trading history, at least one year of filed accounts, no court judgements in the past 12 months, and UK ownership and resident directors; (ii) statistical credit models, which are proprietary and use extensive data to assess creditworthiness; and (iii) expert judgement, which draws on the expertise of the credit assessment team. While it is likely that this screening protects the platform from 'bad' borrowers to some degree, Funding Circle does not make data available on loans that are rejected at this initial screening stage that would allow us to appraise this issue. This is, therefore, necessarily outside the scope of our study and we focus exclusively on investor investment decisions post this screening stage. As discussed next, the allocation of loans by the lending platform to the wholesale and retail marketplaces is random. There is, therefore, no reason to expect a concentration of adverse selection in either marketplace.

2.2.2. Remarks on the randomised loan allocation process

A statistical concern in our comparison of the wholesale and retail marketplaces is the randomness of the two loan groups being compared. This is directly addressed through the loan allocation process administered by the lending platform. Loans are allocated randomly by the platform between the wholesale and retail marketplaces. This randomisation eliminates sample selection bias in respect of loan group formation and resulting endogeneity concerns. In terms of statistical identification, we are dealing with a randomised experiment. However, we note that sample selection bias still exists given the selectivity shown by institutional investors in deciding to grant certain loans over others. Such decision making is not random. We address this particular point later in this section.

The statement that we are dealing with a randomised experiment is, of course, only true if we believe that the process administered by the lending platform is unbiased. It may be questioned as to whether this is truly the case and whether the lending platform is motivated in some way to direct higher quality loans to the wholesale marketplace. One may suggest that this can be probed through testing whether there are statistical differences in the characteristics of the wholeloan and partloan books. However, as informed by the literature on randomised control-treatment analysis in other disciplines, testing statistical differences in baseline characteristics between the subject groups is not appropriate in efforts to confirm the randomness of the group formation procedure (e.g. Altman, 1985; de Boer, Waterlander, Kuijper, Steenhuis, & Twisk, 2015; Harvey, 2018). On average, across randomised experiments, randomised groups will have the same characteristics, but if differences in characteristics do exist between groups in a given randomised experiment then this should not necessarily raise doubts about the randomisation procedure (Altman, 1985). On this basis, we take a different approach and exploit the fact that we know when each loan is accepted by the lending platform and presented to investors, and we know in which marketplace the loan is placed. This allows us to test whether the sequence of loan allocations to the wholesale and retail marketplaces follows a random process. Any obvious bias in this loan allocation procedure should become apparent. We use the Wald-Wolfowitz runs test of randomness (Wald & Wolfowitz, 1940) for this purpose.

For the implementation of the Wald-Wolfowitz runs test of randomness, we code each loan allocation as either "1" if it is first placed in the wholesale marketplace or "2" if it is first placed in the retail marketplace. We then order this sequence of loan allocations by the loan acceptance date, i.e. the date on which each loan is accepted by the lending platform and placed in either the wholesale marketplace or retail marketplace. We can then test this sequence of loan allocations on the basis of a null hypothesis that the sequence was produced in a random manner. A normally distributed test statistic, Z, can be obtained as follows:

$$Z = \frac{R - E(R)}{SD(R)}$$

where R is the observed number of runs and a "run" is defined as a sequence of consecutive "1's" or "2's", $E(R)$ is the expected number of runs, and $SD(R)$ is the associated standard deviation. $E(R)$ is defined as follows:

$$E(R) = \frac{2n_1n_2}{n_1 + n_2} + 1$$

and $SD(R)$ is derived from

$$VAR(R) = [SD(R)]^2 = \frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}$$

where n_1 is the total number of "1's" (or wholeloans) and n_2 is the total number of "2's" (or partloans).

Table A.I (Panel I) in Appendix A (online) details the results. We find no evidence that the loan allocation process administered by the lending platform is non-random and biased. When we repeat the analysis for each of the credit bands separately (Table A.I, Panels II–VII), we again find no evidence that the allocation process favours institutional investors in terms of loan quality.

2.2.3. Inclusion of recycled loans

As established, loans are randomly allocated by the lending platform between the wholesale and retail marketplaces. This is true when loans are first presented to market. However, wholeloans that are not financed by institutional investors are subsequently recycled and presented to the retail marketplace. The presentation of such loans is of course non-random. Some comment is therefore required on the inclusion of recycled loans in our testing set out. We show that incorporating this non-random information in our randomised experiment does not present a problem. We show later that when we repeat the tests with recycled loans excluded, our findings are consistent. Our motivation for including recycled loans though is that they provide interesting insights into the investments foregone by institutional investors.

2.2.4. Institutional investor selectivity

The focus of our paper is on comparing the wholesale and retail markets places in order to make statements on the decision making of institutional investors. A potential source of sample selection bias therefore comes from the selectivity shown by institutional investors in funding loans in the wholesale marketplace, i.e. institutional investors do not finance all loan applications. The random allocation of loans into the wholesale marketplace does not remove sample selection bias concerns in this respect as the decision to fund a given loan is not random. We address this explicitly and present a statistical approach motivated by existing literature.

We begin with a conventional bank lending setting and draw on the literature pertaining to credit score modelling (Greene, 1998; Jacobson & Roszbach, 2003). In the process of screening loan applications, sample selection bias arises from the observability of loan default outcomes only for those loans funded by the bank. Any estimation of default likelihood in isolation is therefore likely to be biased given the non-random nature of the loan pool used in the estimation. In line with Heckman (1979), recasting the problem as a bivariate probit selection model, comprising a loan granting process *and* a loan default process allows us to explicitly account for this source of sample selection bias (Boyes, Hoffman, & Low, 1989). Importantly, this allows us to link the decisions of investors to grant loans with the ultimate outcomes in respect of loan default and to make inferences on whether these decisions are consistent with an objective of default risk minimisation. Sample selection bias, as an inherent issue of correlation between observable and unobservable variables, manifests in this bivariate probit selection model setting as a significant correlation between the error terms of the two processes (Vella, 1998). This suggests that there are unobserved factors influencing the decision to lend that are related to the unobserved factors that are influencing the likelihood of loan default. In this case, only the joint estimation of the two processes delivers estimates that are consistent (Vella, 1998).

The likelihood of default in our setting may be influenced by the selectivity of investors in granting certain loans over others, such that an estimation of the univariate default probit model may be biased. Some observations are required. Firstly, over the sample period all loans presented to the retail marketplace are funded, and so no selectivity is observed in the lending decisions of the retail investor groups operating in this marketplace. This is not the case for institutional investors operating in the wholesale marketplace, where some loans are indeed rejected. Secondly, all rejected wholeloans are fully funded when recycled into the retail marketplace. Therefore, in our setting, we have full sight of all institutionally rejected loans, albeit such loans are ultimately funded by retail investors in the retail marketplace. It could be argued therefore that this mitigates sample selection bias, as we observe the outcomes of these rejected loans. However, we argue that the subsequent funding of recycled wholeloans in the retail marketplace is not a perfect counterfactual situation. For instance, during the auction period of our sample (1st January 2014–27th September 2015), the interest rates bid by investors in the retail marketplace on recycled wholeloans are likely to be different from the interest rates that would have been bid by institutional investors had the lending requests been initially successful in the wholesale marketplace.

It is therefore prudent for us to account for sample selection bias in a bivariate selection-outcome system and explore to what extent there is a relationship between the decision to grant funding for loans and subsequent loan default events. There is also an appealing economic interpretation to our approach in that it allows us to construct a narrative pertaining to the unobserved determinants of the decision to grant a loan. This is pertinent for our sample, as we have a limited set of observable loan and firm characteristics, and we are unaware of the borrower information being utilised by lenders.

We closely follow the approach taken in the credit scoring literature (Greene, 1998; Jacobson & Roszbach, 2003; Marshall et al., 2010; Roszbach, 2004) in structuring a bivariate selection-outcome (loan-granting-loan-default) probit model defined in general terms as follows:

$$\text{Granted\_FirstTime}^i = \mathbf{1}\{\beta_1' X_1^i + \varepsilon_1^i > 0\}$$

$$\text{Default}^i = \mathbf{1}\{\beta_2' X_2^i + \varepsilon_2^i > 0\}$$

For the loan default process, *Default* is again assumed to be driven by

$$X_2 := \{Term, \log Age, \log LoanAmt, IntRate, Wholeloan, Recycled. , CrBandA, ..., CrBandE\}'$$

For the loan granting process, we define *Granted_FirstTime = 1* if a loan presented to the market (whether wholeloan or partloan) is granted first time,[2] which is assumed to be driven by

[2] It is important to emphasise that when we say a loan has been funded first time, we mean that the loan has been funded by the investor cohort to which it is first presented. So a loan that is presented to the wholesale market place and funded fully is deemed to have been granted first time (and, hence, designated as a 'wholeloan' with one part). In contrast, a loan that is presented to the wholesale market place and not funded, i.e. declined by institutional investors, is deemed not to have been granted first time. This is not to be conflated with the internal procedure of Funding Circle, whereby if a loan is not funded on a first round by a *given* investor cohort then it is subsequently re-presented to the *same* investor cohort on a second round.

Table IV
Probability of loan default.

	Base model		Ext. model	
	Coeff.	Marginal effects	Coeff.	Marginal effects
Term	0.0197***	0.0042***	0.0191***	0.0039***
	(0.0016)	(0.0003)	(0.0018)	(0.0004)
log Age	−0.1156***	−0.0246***	−0.1985***	−0.0408***
	(0.0384)	(0.0082)	(0.0427)	(0.0087)
log LoanAmt	−0.0458	−0.0097	−0.0081	−0.0017
	(0.0326)	(0.0069)	(0.0364)	(0.0075)
IntRate	0.0958***	0.0204***	0.1102***	0.0226***
	(0.0356)	(0.0076)	(0.0368)	(0.0075)
Wholeloan	−0.1165**	−0.0248**	−0.1529**	−0.0314**
	(0.0585)	(0.0124)	(0.0643)	(0.0132)
Recycled	0.0848	0.0180	0.0381	0.0078
	(0.0858)	(0.0182)	(0.0879)	(0.0181)
CrBandA	0.3373***	0.0718***	0.2709***	0.0557***
	(0.0997)	(0.0212)	(0.1051)	(0.0216)
CrBandB	0.3336***	0.0710***	0.2209*	0.0454*
	(0.1207)	(0.0257)	(0.1262)	(0.0259)
CrBandC	0.4127***	0.0878***	0.2949*	0.0606*
	(0.1517)	(0.0323)	(0.1587)	(0.0326)
CrBandD	0.2325	0.0495	0.0851	0.0175
	(0.2098)	(0.0446)	(0.2180)	(0.0448)
CrBandE	−0.0872	−0.0186	−0.2667	−0.0548
	(0.4080)	(0.0868)	(0.4230)	(0.0869)
Constant	−2.3096***		−7.2298	
	(0.3912)		(127.1547)	
Region			Yes	
Sector			Yes	
Purpose			Yes	
Year			Yes	
Observations	3791	3791	3791	3791
LR chi^2	355.51		471.40	
Prob > chi^2	0.0000		0.0000	
Pseudo R^2	0.1084		0.1437	

Table reports the results of the probit regression model of loan default probability. The base probit regression model specification is as per Eq. (1) (Section 2.1), with the extended model extending the base model with time, region, sector and purpose fixed effects. All variables are defined in Table I. *Coeff.* denotes regression coefficient. Standard errors are in parentheses. ***, **, * denote statistical significance at the 1%, 5% and 10% levels respectively.

$$X_1 := \{Term, \log Age, \log LoanAmt, Offered\_Wholeloan, CrBandA, \ldots, CrBandE\}',$$

where *Offered_Wholeloan* is a dummy variable introduced to indicate if a loan has been offered in the first instance to the wholesale marketplace. The error terms ${\varepsilon_1}^i$ and ${\varepsilon_2}^i$ capture the influence of unobservable factors driving the two processes; the former partly capturing the borrower information used by investors to inform their decisions to grant loans, and the latter partly capturing firm specific factors leading to default. The two error terms are assumed to be jointly normally distributed with correlation coefficient ρ. A statistically significant ρ establishes that there is a relationship between the loan granting and loan default processes, implying that the two must be estimated jointly. Estimating default probability in a univariate probit model would therefore be biased. In terms of economic interpretation, if the correlation is also negative then this indicates a pattern of investors minimising default risk.

The bivariate probit model is estimated using three approaches for robustness. We first consider wholeloans and partloans jointly and include loans rejected by institutional investors in the wholesale market in the default process, thus adopting identical dimensions for both the loan granting and loan default processes. We then run a Heckman selection probit model, where we exclude the rejected loans from the default process, thereby working as if these loan outcomes are unobservable. The default process therefore has a lower dimension than the granting process (by the number of loans rejected). For completeness, and to isolate the decision making of the institutional investors in the wholesale market, we implement the Heckman selection probit model once more, on this occasion including only loans offered to the wholesale marketplace, while continuing to exclude rejected loans from the default process. This latter specification only models the lending decisions of institutional investors within the wholesale marketplace. We therefore drop *Wholeloan* from the default process and *Offered_Wholeloan* from the granting process.

2.3. Institutional participation in the retail marketplace

In our testing framework, we have to carefully transition from a comparison of marketplaces (i.e. wholesale versus retail marketplaces) to a comparison of investor cohorts (i.e. institutional versus retail investors). This is due to the fact that we have confirmed, via direct communications with the lending platform, that there is a single large institutional investor that has invested consistently in the retail market on a passive basis, funding between 5 and 20% of the loan value of *all* retail loans covered in our sample. This creates a potential source of bias that may distort information from the retail marketplace, preventing us from directly interpreting the information as relating purely to retail investors. We therefore conduct a number of robustness checks to investigate whether this institutional activity in the retail marketplace is a distorting factor.

The first approach we take is to repeat our analysis using a database of loans that pre-appends our current sample of 2014–2016 with the 2010–2013 period, allowing us to include all years since the establishment of the platform. While the same large institutional investor participated in the retail marketplace in 2013, only retail investors operated in the years 2010–2012. This serves to dilute any potential distortion of information from the retail marketplace due to institutional participation. The second approach we take is to compare the wholesale marketplace over the 2014–2016 period with the retail marketplace over the 2010–2012 period. This serves to remove potential distortion due to institutional participation in the retail marketplace. In comparing marketplaces across different periods of time, we account for time fixed effects. The third approach we take is similar to this but we create on this occasion a one-to-one matched loan sample from the retail marketplace loans over 2010–2012. The matching procedure involves *exact* matching on the basis of sector and loan term and *nearest neighbour* matching on the basis of loan amount. This serves to remove potential institutional investment distortion in a more controlled manner than the previous approach, while we again account for time fixed effects.

If our findings from the three approaches prove to be broadly consistent with our main findings then this provides some confidence that institutional activity is not distorting the lending information emanating from the retail marketplace. After careful consideration of this issue, we may then consider the retail marketplace and retail investment interchangeably in our discussions.

3. Empirical results and discussion

3.1. Main findings

The main results of the default probability probit models are presented in Table IV, with the main results of the loan rate and realised returns regression models presented in Table V. We focus on the base model specifications, making reference to the extended model specifications with time, region, sector and loan purpose fixed effects as required. There is broad consistency between the alternative model specifications. The results of the robustness checks outlined in Section 2.3, pertaining to the presence of institutional investors in the retail marketplace, are presented in Tables B.I–B.VI in Appendix B (online). The findings of these robustness checks are consistent with the main findings we report below. We can, therefore, validly speak of the wholesale

Table V
Loan rate and realised return analysis.

Dep. Var.	Loan rate (%)		Realised return (defaulted) (%)		Realised return (repaid) (%)	
	Base model	Ext. model	Base model	Ext. model	Base model	Ext. model
Term	0.0093***	0.0103***	−0.7235***	−0.5752***	0.0099***	0.0112***
	(0.0007)	(0.0007)	(0.1143)	(0.1085)	(0.0008)	(0.0009)
log Age	−0.0873***	−0.0494***	7.7114***	7.1346***	−0.0857***	−0.0391**
	(0.0163)	(0.0176)	(2.2882)	(2.2141)	(0.0187)	(0.0198)
log LoanAmt	0.2536***	0.2160***	−1.3513	−1.6460	0.2755***	0.2313***
	(0.0129)	(0.0148)	(2.1723)	(1.8906)	(0.0154)	(0.0184)
Wholeloan	0.4042***	0.4473***	−15.3659***	−4.8149*	0.4664***	0.5160***
	(0.0259)	(0.0280)	(3.0511)	(2.7994)	(0.0292)	(0.0308)
Recycled	0.2936***	0.3105***	−6.4406*	−7.8892**	0.2855***	0.3100***
	(0.0436)	(0.0435)	(3.7165)	(3.6860)	(0.0630)	(0.0634)
CrBandA	1.3334***	1.3806***	3.7428	−0.1402	1.4060***	1.4603***
	(0.0344)	(0.0354)	(5.5632)	(5.4909)	(0.0399)	(0.0435)
CrBandB	2.3154***	2.3609***	5.7617	−2.6572	2.4759***	2.5329***
	(0.0363)	(0.0379)	(5.3892)	(5.2860)	(0.0422)	(0.0460)
CrBandC	3.4072***	3.4638***	6.9410	−2.9672	3.6457***	3.7110***
	(0.0398)	(0.0414)	(5.6679)	(5.5301)	(0.0466)	(0.0503)
CrBandD	5.1319***	5.1946***	5.6841	−1.6727	5.5643***	5.6395***
	(0.0477)	(0.0488)	(6.1820)	(5.7564)	(0.0541)	(0.0571)
CrBandE	10.1409***	10.2067***	−14.3245	−3.4966	11.3467***	11.3944***
	(0.1053)	(0.1055)	(9.2227)	(8.9707)	(0.0792)	(0.0869)
Constant	4.7448***	5.3945***	−31.5165	12.6367	3.6830***	4.4511***
	(0.1493)	(0.2256)	(21.6970)	(22.8260)	(0.1797)	(0.2618)
Region		Yes		Yes		Yes
Sector		Yes		Yes		Yes
Purpose		Yes		Yes		Yes
Year		Yes		Yes		Yes
Observations	3791	3791	590	590	3201	3201
Adj. R^2	0.8605	0.8640	0.1433	0.3275	0.8563	0.8614
F	2342.08	603.55	9.14	9.62	3202.57	796.38

Table reports the results of the loan rate and realised returns regression models. The base regression model specifications are as per Eqs. (2) and (3) (Section 2.1) respectively, with the extended model extending the base model with time, region, sector and purpose fixed effects. All variables are defined in Table I, along with the base cases for the dummy variable categories used. Standard errors are in parentheses. ***, **, * denote statistical significance at the 1%, 5% and 10% levels respectively.

(retail) marketplace and wholeloans (partloans) interchangeably when referring to institutional (retail) investment. The results reported in Appendix C (online) (Tables C.I–C.II) confirm that the presence of recycled loans in our testing is not a source of bias. We find broad consistency in our findings, whether recycled loans are included or excluded from the analysis.

In respect of our main findings, we begin with a discussion of loan screening ability. We observe that the probability of loan default is 2.48% lower for wholeloans funded in the wholesale marketplace compared to partloans funded in the retail marketplace (Table IV), while the interest rates on wholeloans are 0.40% higher on average (Table V). Hence, institutional investors appear to have a superior ability to screen loans than retail investors in the retail marketplace, achieving higher ex-ante loan rates on agreed loans while assuming lower ex-post default risk. Our results contrast with Lin et al. (2017) who find no evidence of a difference between institutional and retail investment in peer-to-peer personal lending. This suggests that peer-to-peer business lending may be different in this respect.

In respect of recycled loans, we observe that such loans bear interest rates 0.29% higher on average (Table V) without an attendant higher probability of default (Table IV) compared to *all* other loans (across both wholesale and retail marketplaces). This seems to suggest some missed opportunity on the part of institutional investors in the wholesale marketplace who appear to forego on this higher return potential. However, when we compare the recycled loans against wholeloans only, we make a different conclusion. We present the results of this analysis in Appendix D (online). It can be seen that recycled loans actually have a higher probability of default than wholeloans (Table D.I), while loans rates are actually lower than wholeloans on average (Table D.III). So institutional investors appear to have been able to offload loans that offer lower interest rates against higher default exposure. This contrasts to the case of recycled loans compared against non-recycled partloans. It is in this case that recycled loans bear higher interest rates on average (Table D.IV) without an attendant higher probability of default (Table D.II). So what we see in our main results is really being driven by the relationship between recycled and non-recycled partloans.

Moving beyond loan screening ability, our analysis of realised returns (Table V) provides some new insights. Firstly, we show that wholeloans funded in the wholesale marketplace achieve higher realised gains when loans are repaid, although the margin is narrow at 0.47%. By contrast, outcomes on defaulted loans in the wholesale marketplace are considerably worse than in the retail marketplace, with losses recorded to be 15.37% higher. We can conclude therefore that although the likelihood of default is lower for loans funded in the wholesale marketplace, such loans lose considerably more upon default. With regard to recycled loans, such loans appear to lose more upon default and earn more upon repayment (Table V) when compared to all other loans. Separating out the analysis as done in Appendix D (online), we observe that this result is again being driven by the relationship between recycled and non-recycled partloans (Table D.IV). When the comparison is confined to just recycled loans and wholeloans, we see instead that recycled loan lose in the case of loan default but earn less when loans are repaid.

We have therefore presented evidence that, relative to retail investors in the retail marketplace, institutional investors in the wholesale marketplace make lending decisions which are, on average, associated with lower probability of default and higher loan rates – showing superior loan screening ability - and higher realised gains upon loan repayment – showing superior loan payoffs when investors are repaid in

Table VI
Sample selection bias testing: Bivariate selection-outcome models.

ρ	Standard error	Likelihood ratio test of ρ = 0	
		$\chi^2(1)$	p-Value
Bivariate probit model (all loans; recycled wholeloans included in loan default process)			
−0.1886	0.3533	0.35	0.5557
Heckman selection probit model (all loans; recycled wholeloans excluded from loan default process)			
0.0230	0.3017	0.01	0.9395
Heckman selection probit model (wholeloans only; recycled wholeloans excluded from loan default process)			
−0.9999	0.0000	0.78	0.3768

Table reports the results of the bivariate selection-outcome (loan-granting-loan-default) probit model, with recycled loans included in the loan default process, and the associated Heckman selection probit models, with recycled loans excluded from the loan default process. The loan default process is defined with dependent variable *Default* (=1 if loan defaults) related to the set of independent variables $\{Term, \log Age, \log LoanAmt, IntRate, Wholesale, Recycled, CrBandA, \ldots, CrBandE\}$. The loan granting process is defined with dependent variable *Granted_FirstTime* (=1 if loan granted first time) related to the set of independent variables $\{Term, \log Age, \log LoanAmt, Offered\_Wholeloan, CrBandA, \ldots, CrBandE\}$. See Section 2.2 for further details. For comparative purposes, we also implement the Heckman selection probit model on loans offered to the wholeloan segment of the market, excluding recycled loans from the default process. This latter specification models the lending decisions of institutional investors only.

full. Furthermore, we are able to ascertain that institutional investors generally make smart decisions with respect to the loans that they reject, as evidenced when we confine our analysis to a comparison of wholeloans and recycled loans. The evidence we present broadly supports the conclusion that the sophistication of institutional investors leads to superior lending outcomes in the wholesale market. This aligns with the conclusions of Vallee and Zeng (2019), albeit their study concerns personal lending.

3.2. Institutional investor selectivity

We address the potential source of sample selection bias resulting from the non-random selectivity shown by institutional investors in their lending decisions in the wholesale marketplace. We re-consider our loan default estimation and implement the bivariate selection-outcome (loan-granting-loan-default) model proposed in Section 2.2. Estimated correlation coefficients are presented in Table VI. The bivariate probit model confirms that the correlation between the loan granting and loan default processes is indistinguishable from zero, and so we conclude that this potential source of sample selection bias is not a concern in our calculation of default probability. The Heckman selection probit specification, which excludes recycled loans from the default process, confirms this conclusion. Our univariate estimation of the probit model as per Eq. (1) is therefore not biased, but the economic interpretation of this is notable.

The unobservable factors driving decisions to fund loans, which in part capture the use of borrower information in appraising loan applications, are not associated with the unobservable factors driving loan default outcomes. This points to an inability of institutional investors in the wholesale marketplace to manage information asymmetries in such a way as to minimise default risk. Indeed, when we isolate the lending decisions of institutional investors within the wholesale marketplace, we find no link between decisions to lend and the default outcomes of loans.

4. Lending efficiency

The last section showed that institutional investors make lending decisions that do not reflect a practice of default risk minimisation. That is, institutional investors are not making decisions to grant loans that lead to lower default exposure. This is evidence of lending inefficiency. It is therefore of interest to quantify the level of inefficiency in the lending decisions of institutional investors. To this end, we implement the simulation exercise of Jacobson and Roszbach (2003), which uses value-at-risk (VaR) as a measure of lending efficiency.

We conduct a Monte Carlo simulation of the estimated probit model, where for each simulation we calculate an expected default probability for each loan request. We then define a default risk decision rule whereby loan requests with an expected default probability that equal or exceed a *chosen* default likelihood threshold δ are rejected, while those below the threshold are funded. To ensure robustness, we conduct 1000 simulations in total. Repeating the exercise for a range of δ-threshold values allows us to imply the threshold level that corresponds to the actual level of financing provided by institutional investors. In total, institutional investors extended £85.66 m in loans, which implies a δ-threshold of 26.71%, which is a notably high likelihood of default upon which to base lending decisions.[3] Table E.I in Appendix E (online) presents the loan loss distribution characteristics for the wholesale marketplaces. The δ-threshold of 26.71% is found by a simple linear interpolation between the adjacent total lending amounts corresponding to the 25% and 30% δ-thresholds.

To quantify the level of lending inefficiency, we perform a VaR analysis based on the simulated distribution of loan losses using a 26.71% default likelihood decision rule, following the design of Jacobson and Roszbach (2003). We generate results at the usual 1%, 5% and 10% VaR levels, which are reported in Table VII (Panel A). To distinguish between the sample and efficient institutional investor lending reported, note first that in the former sample case, the VaR estimates are calculated based on the actual loans chosen by the institutional investors in the sample. That is, for each simulation, the portfolio is formed from the 1489 loans that institutional investors actually funded. In the latter efficient case, for each simulation the portfolio is formed by selecting from the full pool of wholeloans, i.e. the same 1489 loans *plus* the 356 loans that were originally rejected by institutional investors. This allows us to assess whether the institutional investors could have made more efficient lending decisions, i.e. rejected a different set of wholeloans to reduce VaR exposure. We see that more efficient decision making could indeed have been made and would have

[3] The δ range we are consider is {0.05, 0.06, 0.07, 0.08, 0.09, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 1.00}. For each of these threshold levels, we calculate the mean level of funding across the 1000 simulations. We then use simple linear interpolation to infer the δ-threshold that corresponds to the actual level of lending extended by institutional lenders, i.e. £85.66 m.

Table VII
Efficiency of wholeloan marketplace lending.

		1% VaR		5% VaR		10% VaR	
	Total lending	Value	Frac. of total lending	Value	Frac. of total lending	Value	Frac. of total lending
Panel A: lending efficiency: observed loan allocation							
Sample	85.66	13.34	0.16	13.27	0.15	13.24	0.15
Efficient	85.66	10.77	0.13	10.68	0.12	10.62	0.12
Panel B: lending efficiency: randomised loan allocation (random loan rejection)							
Sim sample	108.76	15.56	0.14	15.49	0.14	15.45	0.14
Sim efficient	108.76	12.91	0.12	12.77	0.12	12.70	0.12

Table is organised as follows. Panel A presents a comparison of the estimated value-at-risk exposure of the sample lending in the wholeloan marketplace and the efficient lending in the wholeloan marketplace as derived from the Monte Carlo simulation exercise detailed in Section 4, where the total lending amount under the simulation is constrained to equal the observed total lending in the wholeloan marketplace (£85.66m). To distinguish between the sample and efficient institutional investor lending, note that in the former sample case the VaR estimates are calculated based on the actual loans chosen by the institutional investors in the sample, i.e. for each simulation, the portfolio is formed from the 1489 loans that the institutional investors actually backed. In the latter efficient case, for each simulation, the portfolio is formed by selecting from the full pool of wholeloans, i.e. the same 1489 loans plus the 356 loans that were originally recycled by institutional investors. Panel B presents the results of the randomised loan allocation experiment detailed in Section 4. The randomised loan allocation procedure works by randomly redefining each loan in the sample as being either a wholeloan or partloan. This generates a set of simulated wholeloans that differ from the set of observed wholeloans. Lending efficiency can then be determined by comparing the value-at-risk exposure for a simulated sample derived from a process of random loan rejection in proportion to the observed rejection rate in the observed sample data and an efficient sample derived by following a default risk based decision rule consistent with the implied δ-threshold of 27% reported for the wholeloan marketplace. This procedure is repeated for a large set of simulations (1000 simulation runs of 1000 simulations each). The reported figures are averages across the simulations. Monetary amounts are in units of one million pounds.

reduced the institutional investor VaR exposure by approximately £2.5 m.

These findings may of course be an artefact of the observed outcome of the random loan allocation process administered by the lending platform. We therefore conduct a randomised loan allocation simulation designed as follows. We randomly redefine each loan as either wholeloan or partloan, which generates a counterfactual scenario. We obtain a set of *simulated* wholeloans that differ from the set of *observed* wholeloans. We then proceed to randomly reject a number of wholeloans, in the same proportion to the observed data. This allows us to test whether the lending inefficiency of institutional investors in the wholeloan marketplace is consistent with or different from a naïve strategy of randomly rejecting loans. We repeat this procedure for a large set of simulations, whereby we generate 1000 counterfactual scenarios of 1000 simulations each.

Table VII (Panel B) reports the results of this particular analysis, where it can be seen that, across the alternative counterfactual loan allocation scenarios, had institutional investors followed a default risk based decision rule then greater efficiencies would have been achieved relative to case of randomly rejecting loans. The magnitude of VaR reductions is in the order of £3m, and 2% on a relative basis, which is remarkably similar to the efficiencies reported in Table VII (Panel A). This suggests that our main finding is not an artefact of the observed outcome of the random loan allocation process, and further that the observed decision making of institutional investors does not appear to be distinct from a strategy of randomly rejecting loans.

5. Retail investor group size effects

In a unique contribution to existing work, we extend our analysis thus far and argue for a more nuanced consideration of the retail investor base. Surowiecki (2005) identifies that the 'wisdom of crowds' is premised on three important conditions: independence, diversity, and decentralisation. These are all properties that apply in the crowdfunding setting. A somewhat surprising conclusion of Surowiecki (2005) is that large groups of untrained people are often better at decision making than small groups of experts. Drawing on these theories, group size and group diversity have been found to be positively related to better outcomes in a number of non-financial applications (Bassamboo et al., 2015). In addition, Robert and Romero (2017) identify group experience as another important component of the wisdom of crowds.

On the flip side, however, there is a counter view to the above logic that must be acknowledged. Olson (1965) argues that unless a group is small or unless there is some form of coercion or some other channel through which to unify constituent members' objectives then even rational, self-interested individuals will not work towards the common interests of the group. The logic game based analysis of Hardin (1982) confirms that collective action is generally bound to fail unless the scenario is such that collective action is not actually needed. By contrast, the anthropological analysis conducted by the same author suggests that collective action can be effective (ineffective), even in large (small) groups, but this is subject to a complex set of interactions. Oliver and Marwell (1988) provide a counter argument to Olson (1965), however, suggesting that when groups are large and diverse, fewer contributors are needed to provide an overall contribution to the group. In our online lending market context, an interpretation of this idea is that a subgroup of financially experienced investors may be sufficient in order for large groups, in aggregate, to make better financial decisions. Oliver and Marwell (1988) explain that this relationship holds on the basis of there being high *jointness of supply* to the good, meaning that the cost of the good has a low dependency on the number of participants in the group, i.e. the cost of the good does not increase with the number of participants. This is the case for loan funding applications, whereby the overall cost of funding the loan, i.e. the loan amount sought by the borrower, does not vary depending on the number of funders.

While our database does not provide information on the diversity and experience of retail investor groups, we can readily identify group size by means of the number of loan parts. We propose that group size is positively related to lending outcomes, primarily through better aggregation of information that exploits the diversity and experience of constituent members. Consequently, we suggest that a large group is more likely to have a broader, more diverse span of opinion, knowledge and skills across its constituent members.

To investigate if there are group size effects in our setting, we repeat our comparative analysis, segmenting the partloan book into terciles, whereby we establish small (*Low_Parts*), medium (*Med_Parts*) and large (*High_Parts*) retail investor group size identifiers. This segmentation of the partloan book implies an upper threshold for the small group size category as 498 parts (or individual investors) and an upper threshold for the medium group size category as 1221 parts (or individual investors). Repositioning wholeloans as the baseline group, we adapt our testing models to now include the *Low_Parts*, *Med_Parts* and *High_Parts* dummy variables. Results are reported in Tables VIII and IX. While the likelihood of default for loans financed by small- and medium-sized retail investor groups is statistically significant and higher (by over 3%; Table VIII) than that for loans financed by institutional investors in the wholesale marketplace, this is not the case for loans financed by large-sized retail investor groups, where there is no statistically significant difference evidenced relative to institutional investors. So in respect of loan default, there appears to be no difference in default risk exposure between institutional investors and large-sized retail investor groups.

Table VIII
Probability of loan default: Size-based retail investor groups.

	Base model		Ext. model	
	Coeff	Marginal effects	Coeff	Marginal effects
Term	0.0197***	0.0042***	0.0191***	0.0039***
	(0.0016)	(0.0003)	(0.0018)	(0.0004)
log Age	−0.1153***	−0.0245***	−0.1996***	−0.0410***
	(0.0385)	(0.0082)	(0.0428)	(0.0087)
log LoanAmt	−0.0157	−0.0033	0.0281	0.0058
	(0.0427)	(0.0091)	(0.0463)	(0.0095)
IntRate	0.1043***	0.0222***	0.1218***	0.0250***
	(0.0367)	(0.0078)	(0.0378)	(0.0078)
Low_parts	0.1659*	0.0353*	0.2285**	0.0469**
	(0.0848)	(0.0180)	(0.0899)	(0.0184)
Medium_parts	0.1427*	0.0304*	0.1471*	0.0302*
	(0.0758)	(0.0161)	(0.0813)	(0.0167)
High_parts	0.0404	0.0086	0.0795	0.0163
	(0.0855)	(0.0182)	(0.0905)	(0.0186)
Recycled	0.0778	0.0166	0.0319	0.0066
	(0.0860)	(0.0183)	(0.0881)	(0.0181)
CrBandA	0.3219***	0.0685***	0.2557**	0.0525**
	(0.1005)	(0.0214)	(0.1058)	(0.0217)
CrBandB	0.3080**	0.0655**	0.1933	0.0397
	(0.1226)	(0.0261)	(0.1280)	(0.0263)
CrBandC	0.3798**	0.0808**	0.2565	0.0527
	(0.1545)	(0.0329)	(0.1614)	(0.0331)
CrBandD	0.1844	0.0392	0.0273	0.0056
	(0.2142)	(0.0456)	(0.2225)	(0.0457)
CrBandE	−0.1488	−0.0316	−0.3506	−0.0720
	(0.4127)	(0.0878)	(0.4279)	(0.0878)
Constant	−2.8070***		−7.8515	
	(0.5333)		(126.9492)	
Region			Yes	
Sector			Yes	
Purpose			Yes	
Year			Yes	
Observations	3791	3791	3791	3791
LR chi^2	356.98		472.98	
Prob > chi^2	0.0000		0.0000	
Pseudo R^2	0.1089		0.1442	

Table reports the results of the probit regression model of loan default probability with the retail market investor base stratified by group size. The base model, as discussed in Section 5, includes *Low_Parts*, *Med_Parts* and *High_Parts* as dummy variables that identify respectively partloans funded by small-, medium- and large-sized retail investor groups respectively. *Wholeloan*, representing loans funded by institutional investors in the wholesale marketplace, is omitted and used as the reference case. All variables are defined in Table I, along with the base cases for the dummy variable categories used. The extended model includes region, sector and purpose fixed effects. *Coeff.* denotes regression coefficient. Standard errors are in parentheses. ***, **, * denote statistical significance at the 1%, 5% and 10% levels respectively.

Furthermore, no identifiable difference in interest rates is observed between wholeloans approved by institutional investors and partloans approved by large-sized retail investor groups, which contrasts with the statistically significant and lower interest rates on loans financed by small- and medium-sized groups (0.80% and 0.34% lower respectively, Table IX). A similar finding emerges in relation to realised returns on repaid loans, where small- and medium-sized retail investor groups underperform both large-sized retail investor groups and institutional investors (by 0.92% and 0.41% respectively; Table IX). Furthermore, while our main findings suggest that institutional investors lose more on defaulted wholeloans than retail investors lose on defaulted partloans, we find no statistically significant difference (Table IX; extended model) in the losses incurred by institutional investors and large- and medium-sized retail investor groups.

The results provide a more nuanced interpretation of our main conclusion that institutional investors outperform retail investors. We find that this is generally true only for small- and medium-sized retail investor groups, with comparable performance, in fact, being evidenced in the lending decisions of institutional investors and large-sized retail investor groups. This is an important new observation about institutional investment in peer-to-business lending.

6. Platform change from auction to fixed rate system

We reported earlier that institutional investors fund loans at ex-ante higher interest rates without facing higher ex-post default risk. It is necessary to consider this finding in the context of an important change to the lending platform that occurred during the sample period. On the 28th September 2015, the mechanism through which interest rates are set on loans was changed from an auction-based system to one of fixed interest rates set by the platform. The Funding Circle auction process allowed for competitive bidding, where an investor bid the gross interest rate required and, at that level of interest rate, the amount of funds willing to be advanced. A business had a degree of flexibility in the process in that it could accept a business loan early, before the auction ended, once 100% of the funding was secured.

The purpose of the move to a fixed rate system was to address inherent flaws with the auctioning system as identified by the lending platform, including cash drag for investors, loan rate uncertainty for borrowers, and the overall complexity of the auction process for its users. In setting fixed interest rates, the lending platform employs its own credit assessment process and sets out a schedule of fixed interest rates across risk band and loan term; property-related loans are dealt with on a case-by-case basis.[4] Fixed interest rates are set based on a number of inputs including macroeconomic trends, expected loss rates, volatility of returns, and comparisons with the wider market for pricing business loans. Similar transitioning away from the auction model has occurred in other marketplaces. For example, Prosper.com moved to a fixed rate system in December 2010. Wei and Lin (2016) study this change of regime and find that the advantages of doing so include a faster deployment of funds and increased interest rates for lenders, although loans are found to have a higher likelihood of defaulting.

The loss of autonomy over the setting of interest rates has considerable implications for institutions. The theory on interest rate setting behaviour suggests that various factors influence the setting of interest rates by institutions. Macro economy-level factors include GDP, inflation, the level and volatility of market interest rates, monetary policy (interest rate channel) and industry structure (competition), while micro firm-level factors include institutional characteristics (size, liquidity, capitalisation, deposit strength, long-term business), maturity mismatch between assets and liabilities, the cost of intermediation and the riskiness of credit portfolios held (Gambacorta, 2008). Demirgüç-Kunt and Huizinga (1999) likewise examine interest margins and show evidence of the role of macroeconomic indicators, the characteristics of the taxation system, the financial structure and institutional characteristics in the setting of interest rates by institutions. Kitamura, Muto, and Takei (2015) summarise the literature on interest pass-through as commonly reporting that high proportions of relationship lending lead to lower interest rates, while capital buffers or liquidity buffers also lead to lower interest rates. This is confirmed by Gambacorta and Mistrulli (2014) in the post financial crisis period. Kitamura et al. (2015) also argue that the balance sheets of borrowing firms are an important driver of interest rates, while Weth (2002) suggests similar in respect of firms' risk characteristics.

The above literature suggests a complex system within which institutions set interest rates, while it is notable in the online lending market setting that institutions cannot work to a relationship lending model and usually have much more limited information on prospective borrowers. Therefore, this provides an important context to the structural change in the interest rate regime underlying the Funding Circle platform. Given this move from lenders as interest rate setters to

[4] The current schedule of fixed interest rates (as of September 2019) is available at https://www.fundingcircle.com/uk/fixedrate/.

Table IX
Loan rate and realised returns analysis: Size-based retail investor groups.

Dep. Var.	Loan rate (%)		Realised return (defaulted) (%)		Realised return (repaid) (%)	
	Base model	Ext. model	Base model	Ext. model	Base model	Ext. model
Term	0.0091***	0.0098***	−0.7212***	−0.5716***	0.0097***	0.0105***
	(0.0007)	(0.0008)	(0.1149)	(0.1090)	(0.0008)	(0.0009)
log Age	−0.0811***	−0.0468***	7.4343***	6.5748***	−0.0767***	−0.0373*
	(0.0163)	(0.0172)	(2.3170)	(2.2371)	(0.0182)	(0.0193)
log LoanAmt	0.0647***	0.0377**	−0.7617	1.1097	0.0677***	0.0377*
	(0.0162)	(0.0171)	(3.0293)	(2.7097)	(0.0185)	(0.0196)
Low_parts	−0.7998***	−0.8296***	17.6040***	10.8026***	−0.9159***	−0.9521***
	(0.0356)	(0.0363)	(4.6108)	(4.0147)	(0.0410)	(0.0422)
Medium_parts	−0.3408***	−0.3661***	12.5837***	0.2457	−0.4036***	−0.4290***
	(0.0382)	(0.0381)	(3.6884)	(3.3709)	(0.0458)	(0.0458)
High_parts	0.0111	−0.0322	15.9229***	2.1555	0.0106	−0.0410
	(0.0377)	(0.0391)	(4.4603)	(4.2242)	(0.0442)	(0.0463)
Recycled	0.3090***	0.3215***	−6.3062*	−7.8371**	0.3107***	0.3292***
	(0.0494)	(0.0491)	(3.7017)	(3.6617)	(0.0610)	(0.0607)
CrBandA	1.3535***	1.3913***	4.1278	0.5636	1.4316***	1.4716***
	(0.0343)	(0.0371)	(5.5556)	(5.5107)	(0.0385)	(0.0419)
CrBandB	2.3479***	2.3840***	6.2561	−1.6880	2.5265***	2.5672***
	(0.0360)	(0.0387)	(5.3968)	(5.3271)	(0.0409)	(0.0441)
CrBandC	3.4316***	3.4788***	7.5013	−1.9970	3.6874***	3.7365***
	(0.0375)	(0.0406)	(5.6756)	(5.5496)	(0.0448)	(0.0479)
CrBandD	5.1533***	5.2083***	6.2910	−0.6138	5.5963***	5.6561***
	(0.0440)	(0.0463)	(6.2304)	(5.8292)	(0.0522)	(0.0547)
CrBandE	10.0016***	10.0626***	−13.4847	0.0456	11.2100***	11.2485***
	(0.0668)	(0.0700)	(9.2474)	(9.1736)	(0.0939)	(0.0970)
Constant	7.1222***	7.6100***	−53.1629*	−21.3326	6.3074***	6.8715***
	(0.1732)	(0.2330)	(30.6468)	(30.7939)	(0.1978)	(0.2693)
Region		Yes		Yes		Yes
Sector		Yes		Yes		Yes
Purpose		Yes		Yes		Yes
Year		Yes		Yes		Yes
Observations	3791	3791	590	590	3201	3201
Adj. R^2	0.8701	0.8742	0.1454	0.3346	0.8670	0.8716
F	3083.10	930.12	7.60	9.30	2246.58	691.66

Table reports the results of the loan rate and realised returns regression models with the retail market investor base stratified by group size. The base model, as discussed in Section 5, includes *Low_Parts*, *Med_Parts* and *High_Parts* as dummy variables that identify respectively partloans funded by small-, medium- and large-sized retail investor groups respectively. *Wholeloan*, representing loans funded by institutional investors, is omitted and used as the reference case. All variables are defined in Table I, along with the base cases for the dummy variable categories used. The extended model includes time, region, sector and purpose fixed effects. Standard errors are in parentheses. ***, **, * denote statistical significance at the 1%, 5% and 10% levels respectively.

interest rate takers, it is pertinent and prudent to test the average difference in loan rate changes vis-à-vis institutional loans in the wholesale marketplace and retail loans in the retail marketplace. To this end, we perform separate loan rate regressions on the auction period (pre-28th September 2015) and the fixed rate period (post-27th September 2015), and then conduct formal model comparison tests thereafter. Mohammadi and Shafi (2017), in their study of the same platform, use a difference-in-difference method to show that interest rates generally increased by 31 basis points on average for *all* lenders with this transition to fixed interest rates. This corroborates the finding of Wei and Lin (2016) that interests increase to the advantage (disadvantage) of lenders (borrowers) with the transition to fixed rates. Our analysis differs and provides insights into the differential effects of this transition on institutional and retail investors. For completeness, we first consider retail investors as a single collective (as in Section 3), and subsequently stratify by retail investor group size categories (as in Section 5). Results are reported in Table X. Note that as property related loans have always worked off a fixed rate system, we exclude this type of loan from our analysis. Descriptive statistics for the loan observations across both periods are presented in Table F.I of Appendix F (online).

Considering retail investors as a single collective first, we find that during the auction period institutional investors achieve an interest rate which is 0.51% higher than retail investors. In the fixed rate period, however, this completely disappears. This indicates that institutional investors were better placed relative to retail investors when setting interest rates via the auction process. When we stratify by retail investor group size, we again find superior performance on the part of institutional investors relative to small- and medium-sized retail investor groups in the auction period, with such loans achieving 0.92% and 0.40%, respectively, below the average loan rates approved by institutional investors, with no significant difference relative to large-sized retail investor groups. This contrasts with the fixed rate period, where we find that interest rates on institutionally-financed wholeloans are indistinguishable from partloans financed by any of the retail investor groups. When we do a similar analysis on the probability of default probit model, we find again that for the fixed rate period there is no difference in the likelihood of default between loans financed by institutional investors and loans financed by retail investor groups of any size (Table XI), while in the auction period loans financed by small- and medium-sized retail investor groups have a higher probability of default relative to wholeloans financed by institutional investors.

When we couple the loan rate and default probability results above, we conclude that the ability of institutional investors to generally screen loans better than retail investors is an observation confined to the auction period.

To complete our assessment of the change to the fixed rate system, we conduct similar analysis this time on realised returns. Results for the realised returns on defaulted and repaid loans are presented in Tables XII and XIII respectively. It is interesting to note that during the auction period, institutional investors consistently lost more upon default (Table XII), while achieving higher returns on the upside (Table XIII), although this latter observation is true only for the small- and medium-

Table X
Loan rate analysis: change over from auction to fixed rate system.

	Base model (retail group collective)		Base model (retail group size categories)	
	Auction period	Fixed rate period	Auction period	Fixed rate period
Term	0.0064***	0.0207***	0.0063***	0.0207***
	(0.0009)	(0.0012)	(0.0009)	(0.0012)
log Age	−0.0379*	−0.0557*	−0.0381*	−0.0559*
	(0.0229)	(0.0314)	(0.0218)	(0.0315)
log LoanAmt	0.2950***	0.0298	0.0592***	0.0190
	(0.0188)	(0.0300)	(0.0199)	(0.0307)
Wholeloan	0.5089***	0.0256	−0.9190***	−0.0487
	(0.0292)	(0.0447)	(0.0380)	(0.0635)
Low_Parts			−0.9190***	−0.0487
			(0.0380)	(0.0635)
Medium_Parts			−0.3989***	−0.0651
			(0.0437)	(0.0511)
High_Parts			−0.0377	0.0309
			(0.0458)	(0.0730)
Recycled	0.3907***	−0.0353	0.4000***	−0.0395
	(0.0565)	(0.0513)	(0.0550)	(0.0519)
CrBandA	1.2410***	1.6648***	1.2353***	1.6668***
	(0.0491)	(0.0503)	(0.0461)	(0.0500)
CrBandB	2.2583***	2.6526***	2.2666***	2.6520***
	(0.0502)	(0.0445)	(0.0475)	(0.0446)
CrBandC	3.3294***	3.9963***	3.3242***	3.9951***
	(0.0510)	(0.0528)	(0.0477)	(0.0530)
CrBandD	4.9828***	6.1572***	4.9660***	6.1583***
	(0.0542)	(0.0475)	(0.0517)	(0.0478)
CrBandE	10.3639***	10.1226***	10.1512***	10.1132***
	(0.1039)	(0.0493)	(0.1073)	(0.0515)
Constant	4.3024***	6.6301***	7.3049***	6.7682***
	(0.2021)	(0.2473)	(0.2056)	(0.2780)
# Obs.	2758	734	2758	734
Adj R^2	0.8186	0.9678	0.8325	0.9679

Equaltiy of regression coefficients test (Clogg, Petkova, & Haritou, 1995)

Coeff	z-Score	p-Value
Wholeloan	8.9974	0.0000
Low_Parts	−11.7525	0.0000
Med_Parts	−4.9901	0.0000
High_Parts	−0.7995	0.2120

Table reports the results of the loan rate regression analysis around the change in the Funding Circle platform from an auction-based system of setting interest rates (‘Auction period’) to a fixed rate system (‘Fixed rate period’). The date of the system change is the 28th September 2015. *Wholeloan* represents loans funded by institutional investors. *Low_Parts*, *Med_Parts* and *High_Parts* are dummy variables that identify respectively partloans funded by small-, medium- and large-sized retail investor groups. All variables are defined in Table I, along with the base cases for the dummy variable categories used. The equality of regression coefficients test follows the approach of Clogg et al. (1995), with statistical significance of the z-score indicating a difference in the regression coefficient between the two sample periods. Standard errors are in parentheses. ***, **, * denote statistical significance at the 1%, 5% and 10% levels respectively. Property related loans are excluded on the basis that this loan type worked to a fixed rate system during the auction period.

Table XI
Probability of loan default: change over from auction to fixed rate system.

	Base model (retail group collective)		Base model (retail group size categories)	
	Auction period	Fixed rate period	Auction period	Fixed rate period
Term	0.0205***	0.0083**	0.0205***	0.0084**
	(0.0017)	(0.0036)	(0.0017)	(0.0036)
log Age	−0.1700***	−0.1981**	−0.1691***	−0.2017**
	(0.0459)	(0.0941)	(0.0459)	(0.0937)
log LoanAmt	−0.0251	0.0697	0.0011	0.0961
	(0.0387)	(0.0618)	(0.0482)	(0.0746)
IntRate	0.1158***	0.1209	0.1219***	0.1247
	(0.0373)	(0.1214)	(0.0387)	(0.1181)
Wholeloan	−0.1675**	−0.1183		
	(0.0685)	(0.1293)		
Low_Parts			0.2043**	0.2513
			(0.0958)	(0.2050)
Medium_Parts			0.1861**	0.0290
			(0.0853)	(0.1916)
High_Parts			0.1057	0.0684
			(0.0987)	(0.1927)
Recycled	0.0275	0.2217	0.0217	0.2269
	(0.0933)	(0.2444)	(0.0936)	(0.2441)
CrBandA	0.2411**	0.2342	0.2322*	0.2239
	(0.1197)	(0.2483)	(0.1206)	(0.2461)
CrBandB	0.2394*	0.0379	0.2227	0.0315
	(0.1376)	(0.3617)	(0.1406)	(0.3569)
CrBandC	0.2714*	0.1802	0.2507	0.1620
	(0.1637)	(0.5287)	(0.1676)	(0.5180)
CrBandD	0.0353	0.0635	0.0048	0.0478
	(0.2206)	(0.7757)	(0.2267)	(0.7595)
CrBandE	−0.2877	−0.5352	−0.3266	−0.5535
	(0.4862)	(1.2543)	(0.4914)	(1.2274)
Constant	−2.5041***	−2.9627***	−2.9975***	−3.3881***
	(0.4242)	(1.0311)	(0.5767)	(1.0444)
# Obs.	2762	734	2762	734
LR chi^2	243.33	54.06	244.66	56.72
Prob > chi^2	0.0000	0.0000	0.0000	0.0000
Pseudo R^2	0.0937	0.0778	0.0940	0.0791

Table reports the results of the probit regression model of loan default probability around the change in the Funding Circle platform from an auction-based system of setting interest rates (‘Auction period’) to a fixed rate system (‘Fixed rate period’). The date of the system change is the 28th September 2015. *Wholeloan* represents loans funded by institutional investors. *Low_Parts*, *Med_Parts* and *High_Parts* are dummy variables that identify respectively partloans funded by small-, medium- and large-sized retail investor groups. All variables are defined in Table I, along with the base cases for the dummy variable categories used. Standard errors are in parentheses. ***, **, * denote statistical significance at the 1%, 5% and 10% levels respectively. Property related loans are excluded on the basis that this loan type worked to a fixed rate system during the auction period.

sized retail investor groups. With the move to the fixed rate system, the upside advantage is no longer evidenced, while there is also no statistical difference in default losses between the two investor cohorts during the fixed rate period. As with the loan screening ability, we find that loan payoffs are only better (worse) for institutional investors on repaid (defaulted) loans during the auction period.

7. Conclusion

Our study contributes to our knowledge of the lending decisions of institutional investors operating in online business lending markets, and to the emerging discourse on institutional involvement in online credit extension more generally. We account for two channels of institutional investment in the online business lending market we examine; namely, the wholesale marketplace dedicated to institutional investors and the retail marketplace dominated by retail investors. We find that institutional investors perform better than retail investors in screening loan applications, evidenced through charging higher ex-ante loan rates without assuming higher ex-post default risk. We additionally find that institutional investors earn higher returns on repaid loans than retail investors. However, we find no evidence that the lending decisions of institutional investors are consistent with an objective of default risk minimisation, raising questions over the ability of institutional investors to process borrower information and alleviate information asymmetries. We leverage this observation to quantify the level of lending efficiency in the wholesale marketplace.

Extending our main findings, we firstly show that a more nuanced consideration of the retail investor base is required. We provide evidence of the influence of group size in assessing loan applications. We find that institutional investors outperform small- and medium-sized

Table XII
Realised return analysis (defaulted loans): change over from auction to fixed rate system.

	Base model (retail group collective)		Base model (retail group size categories)	
	Auction period	Fixed rate period	Auction period	Fixed rate period
Term	−0.7985***	−0.4063*	−0.7984***	−0.4311*
	(0.1334)	(0.2253)	(0.1350)	(0.2280)
log Age	8.5759***	2.4893	8.3072***	3.2865
	(2.5073)	(5.6327)	(2.5357)	(5.5934)
log LoanAmt	−2.7146	0.9794	−3.8820	2.1962
	(2.3241)	(4.8531)	(3.4445)	(5.9188)
Wholeloan	−13.1300***	1.4133		
	(3.5531)	(5.4440)		
Low_Parts			13.7284**	−0.5772
			(5.3457)	(8.6292)
Medium_Parts			8.9821**	4.4604
			(4.1348)	(8.4534)
High_Parts			17.4588***	−6.0646
			(4.8681)	(10.6633)
Recycled	−8.4477**	2.3525	−8.0244**	2.5016
	(3.9146)	(8.0417)	(3.8713)	(8.1350)
CrBandA	3.7255	2.5698	4.3663	1.8548
	(6.3319)	(10.0999)	(6.3559)	(10.7515)
CrBandB	6.0135	−10.3221	6.5625	−10.5446
	(6.2875)	(7.5816)	(6.4090)	(7.6930)
CrBandC	5.8480	−7.9650	6.4770	−8.1167
	(6.6457)	(8.7647)	(6.7880)	(8.9595)
CrBandD	6.4444	−10.4296	7.1853	−10.8494
	(7.3222)	(8.1633)	(7.5293)	(8.4931)
CrBandE	−22.4315***	−1.2943	−22.4319***	−0.7284
	(7.8848)	(13.8142)	(8.3455)	(13.5423)
Constant	−11.3341	−81.5201**	−12.1570	−93.2068**
	(25.2480)	(38.8228)	(38.4118)	(46.9802)
# Obs.	475	115	475	115
Adj R^2	0.1508	0.1001	0.1571	0.1071

Equality of regression coefficients test (Clogg et al., 1995)

Coeff	z-Score	p-Value
Wholeloan	−2.2371	0.9874
Low_Parts	1.4093	0.9206
Med_Parts	0.4764	0.6831
High_Parts	2.4114	0.9921

Table reports the results of the realised return regression analysis on defaulted loans around the change in the Funding Circle platform from an auction-based system of setting interest rates ('Auction period') to a fixed rate system ('Fixed rate period'). The date of the system change is the 28th September 2015. *Wholeloan* represents loans funded by institutional investors. *Low_Parts*, *Med_Parts* and *High_Parts* are dummy variables that identify respectively partloans funded by small-, medium- and large-sized retail investor groups. All variables are defined in Table I, along with the base cases for the dummy variable categories used. Standard errors are in parentheses. ***, **, * denote statistical significance at the 1%, 5% and 10% levels respectively. Note that no property related loans defaulted in either period.

Table XIII
Realised return analysis (repaid loans): change over from auction to fixed rate system.

	Base model (crowd)		Base model (crowd size categories)	
	Auction period	Fixed rate period	Auction period	Fixed rate period
Term	0.0067***	0.0227***	0.0064***	0.0228***
	(0.0010)	(0.0013)	(0.0010)	(0.0013)
log Age	−0.0155	−0.0763*	−0.0189	−0.0752*
	(0.0248)	(0.0396)	(0.0237)	(0.0391)
log LoanAmt	0.3102***	0.0418	0.0573***	0.0347
	(0.0216)	(0.0356)	(0.0221)	(0.0359)
Wholeloan	0.6086***	0.0070		
	(0.0342)	(0.0543)		
Low_Parts			−1.0662***	−0.0102
			(0.0431)	(0.0764)
Medium_Parts			−0.4895***	−0.0645
			(0.0521)	(0.0629)
High_Parts			−0.0669	0.0457
			(0.0542)	(0.0926)
Recycled	0.3918***	−0.0563	0.4077***	−0.0598
	(0.0698)	(0.0683)	(0.0676)	(0.0684)
CrBandA	1.2978***	1.8078***	1.2943***	1.8091***
	(0.0546)	(0.0574)	(0.0511)	(0.0571)
CrBandB	2.4146***	2.8787***	2.4381***	2.8781***
	(0.0569)	(0.0506)	(0.0536)	(0.0509)
CrBandC	3.5480***	4.3755***	3.5596***	4.3735***
	(0.0591)	(0.0668)	(0.0552)	(0.0670)
CrBand_D	5.4047***	6.7628***	5.3951***	6.7638***
	(0.0628)	(0.0550)	(0.0589)	(0.0559)
CrBandE	11.5648***	11.2941***	11.3719***	11.2870***
	(0.1477)	(0.0572)	(0.1683)	(0.0597)
Constant	3.2665***	5.6744***	6.5489***	5.7512***
	(0.2298)	(0.2927)	(0.2295)	(0.3258)
# Obs.	2287	619	2287	619
Adj. R^2	0.8179	0.9637	0.8337	0.9638

Equality of regression coefficients test (Clogg et al., 1995)

Coeff	z-Score	p-Value
Wholeloan	12.2048	0.0000
Low_Parts	−12.0291	0.0000
Med_Parts	−5.2297	0.0000
High_Parts	−1.0532	0.1461

Table reports the results of the realised return regression analysis for repaid loans around the change in the Funding Circle platform from an auction-based system of setting interest rates ('Auction period') to a fixed rate system ('Fixed rate period'). The date of the system change is the 28th September 2015. *Wholeloan* represents loans funded by institutional investors. *Low_Parts*, *Med_Parts* and *High_Parts* are dummy variables that identify respectively partloans funded by small-, medium- and large-sized retail investor groups. All variables are defined in Table I, along with the base cases for the dummy variable categories used. Standard errors are in parentheses. ***, **, * denote statistical significance at the 1%, 5% and 10% levels respectively. Property related loans are excluded on the basis that this loan type worked to a fixed rate system during the auction period.

retail investor groups in lending performance, but find no discernible difference in performance against large-sized retail investor groups. This raises further questions over the ability of institutional investors to manage information asymmetries in this setting. We then proceed to establish that the auction process is critical to successful outcomes in the lending decisions of institutional investors. Indeed, we show that the superior performance relative to retail investors appears to be confined to the auction period.

Our knowledge of institutional investment in online business lending markets lags well behind our knowledge of institutional investment through venture capital and private equity channels. There are a couple of useful directions of future research. Firstly, our evidence raises questions as to how and why financial institutions engage with online lending markets. An interesting direction for future research would be to explore the motivation for institutional participation in such lending activity. Part of this motivation may be diversification in a wider investment portfolio context and so a joint analysis of conventional and alternative lending activities on the part of institutional investors would be welcome. Secondly, a limitation of the lending efficiency analysis we conduct is that it concentrates on the question of default risk minimisation. The question of risk-return trade-off is perhaps more pertinent. Such an analysis of Markowitz mean-variance efficiency in online business lending markets would provide useful insights. Indeed, we have seen a move among peer-to-peer lending platforms of removing autonomy from retail investors and investing instead on behalf of investors in diversified portfolios of loans. This is seen as a way to mitigate risk for lenders, particularly inexperienced lenders, thereby bettering their experience of and retaining their engagement with the lending platforms. As an example, Funding Circle, announced in 2017 an important change to its platform offering. Retail investors must now invest in one of two portfolios: a 'balanced' portfolio that lends to businesses across the full range of credit bands, offering higher project return for higher bad debt exposure; and a 'conservative' portfolio for that lands to business in the upper credit bands, with a lower project return for a lower bad debt exposure. Understanding this new market mechanism is a natural next step to our work.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.irfa.2020.101542.

References

Adams, W., Einav, L., & Levin, J. (2009). Liquidity constraints and imperfect information in subprime lending. *American Economic Review, 99*(1), 49–84.

Altman, D. G. (1985). Comparability of randomised groups. *Journal of the Royal Statistical Society: Series D (The Statistician), 34*(1), 125–136.

Bassamboo, A., Cui, R., & Moreno, A. (2015). *The wisdom of crowds in operations: Forecasting using prediction markets.* (Available at SSRN, 2679663).

Berkovich, E. (2011). Search and herding effects in peer-to-peer lending: Evidence from prosper.com. *Annals of Finance, 7*(3), 389–405.

Boot, A. W. (2000). Relationship banking: What do we know? *Journal of Financial Intermediation, 9*(1), 7–25.

Boubaker, S., & Nguyen, D. K. (2019). Financial decisions in a global uncertain context. *International Review of Financial Analysis, 63*(May 2019), 355–356.

Boyes, W. J., Hoffman, D. L., & Low, S. A. (1989). An econometric analysis of the bank credit scoring problem. *Journal of Econometrics, 40*(1), 3–14.

Bruton, G. D., & Ahlstrom, D. (2003). An institutional view of China's venture capital industry: Explaining the differences between China and the West. *Journal of Business Venturing, 18*(2), 233–259.

Bruton, G. D., Fried, V. H., & Manigart, S. (2005). Institutional influences on the worldwide expansion of venture capital. *Entrepreneurship Theory and Practice, 29*(6), 737–760.

Cassar, G. (2004). The financing of business start-ups. *Journal of Business Venturing, 19*(2), 261–283.

Clogg, C. C., Petkova, E., & Haritou, A. (1995). Statistical methods for comparing regression coefficients between models. *American Journal of Sociology, 100*(5), 1261–1293.

de Boer, M. R., Waterlander, W. E., Kuijper, L. D., Steenhuis, I. H., & Twisk, J. W. (2015). Testing for baseline differences in randomized controlled trials: an unhealthy research behavior that is hard to eradicate. *International Journal of Behavioral Nutrition and Physical Activity, 12*(1), 4.

Demirgüç-Kunt, A., & Huizinga, H. (1999). Determinants of commercial bank interest margins and profitability: Some international evidence. *The World Bank Economic Review, 13*(2), 379–408.

DeYoung, R., Glennon, D., & Nigro, P. (2008). Borrower–lender distance, credit scoring, and loan performance: Evidence from informational-opaque small business borrowers. *Journal of Financial Intermediation, 17*(1), 113–143.

Dorfleitner, G., Hornuf, L., & Weber, M. (2018). Paralyzed by shock and confused by glut: The portfolio formation behavior of peer-to-business lending investors (May 27, 2018). Max Planck Institute for Innovation & Competition research paper no. 18-11; CESifo working paper series no. 7092. Available at SSRN https://ssrn.com/abstract=3186093.

Dorfleitner, G., Priberny, C., Schuster, S., Stoiber, J., Weber, M., de Castro, I., & Kammler, J. (2016). Description-text related soft information in peer-to-peer lending – Evidence from two leading European platforms. *Journal of Banking & Finance, 64*, 169–187.

Emekter, R., Tu, Y., Jirasakuldech, B., & Lu, M. (2015). Evaluating credit risk and loan performance in online Peer-to-Peer (P2P) lending. *Applied Economics, 47*(1), 54–70.

Freel, M., Carter, S., Tagg, S., & Mason, C. (2012). The latent demand for bank debt: Characterizing "discouraged borrowers". *Small Business Economics, 38*(4), 399–418.

Gambacorta, L. (2008). How do banks set interest rates? *European Economic Review, 52*(5), 792–819.

Gambacorta, L., & Mistrulli, P. E. (2014). Bank heterogeneity and interest rate setting: What lessons have we learned since Lehman Brothers? *Journal of Money, Credit and Banking, 46*, 753–778.

Ge, R., Feng, J., Gu, B., & Zhang, P. (2017). Predicting and deterring default with social media information in peer-to-peer lending. *Journal of Management Information Systems, 34*(2), 401–424.

Greene, W. (1998). Sample selection in credit-scoring models. *Japan and the World Economy, 10*(3), 299–316.

Günther, C., Johan, S., & Schweizer, D. (2018). Is the crowd sensitive to distance?—How investment decisions differ by investor type. *Small Business Economics, 50*(2), 289–305.

Hardin, R. (1982). With regret: Comment on rational models, collective goods and nonelectoral political behavior. *Western Political Quarterly, 35*(2), 181–188.

Harvey, L. A. (2018). Statistical testing for baseline differences between randomised groups is not meaningful. *Spinal Cord, 56*, 919.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica, 47*(1), 153–161.

Hornuf, L., & Schwienbacher, A. (2017). Internet based entrepreneurial finance: Lessons from Europe. *California Management Review, 60*(2), 150–175.

Hornuf, L., & Schwienbacher, A. (2018). Market mechanisms and funding dynamics in equity crowdfunding. *Journal of Corporate Finance, 50*, 556–574.

Jacobson, T., & Roszbach, K. (2003). Bank lending policy, credit scoring and value-at-risk. *Journal of Banking and Finance, 27*(4), 615–633.

Kitamura, T., Muto, I., & Takei, I. (2015). *How do Japanese banks set loan interest rates?: Estimating pass-through using bank-level data (no. 15-E-6).* Bank of Japan.

Kon, Y., & Storey, D. J. (2003). A theory of discouraged borrowers. *Small Business Economics, 21*(1), 37–49.

Lin, M., Sias, R., & Wei, Z. (2017). Institutional investors in online crowdfunding. Working paper. Available at http://www.fmaconferences.org/Boston/II_LSW.pdf.

Mac an Bhaird, C. (2010). *Resourcing small and medium sized enterprises.* Heidelberg: Springer-Verlag.

Mac An Bhaird, C., & Lucey, B. (2010). Determinants of capital structure in Irish SMEs. *Small Business Economics, 35*, 357–375.

Mac an Bhaird, C., & Lynn, T. (2015). Seeding the cloud: Financial bootstrapping in the computer software sector. *Venture Capital, 17*(1–2), 151–170.

Mac an Bhaird, C., Sanchez-Vidal, J., & Lucey, B. (2016). Discouraged borrowers: Evidence for Eurozone SMEs. *Journal of International Financial Markets Institutions and Money, 44*(Sept 2016), 46–55 2016.

Mach, T., Carter, C., & Slattery, C. (2014). Peer-to-peer lending to small businesses. *Working paper*. Available at https://www.federalreserve.gov/pubs/feds/2014/201410/201410pap.pdf.

Miller, S. (2015). Information and default in consumer credit markets: Evidence from a natural experiment. *Journal of Financial Intermediation, 24*(1), 45–70.

Mohammadi, A., & Shafi, K. (2017, January). How wise are crowd? A comparative study of crowd and institutional investors. *Academy of management proceedings. Vol. 2017, No. 1. Academy of management proceedings* (pp. 13707–). Academy of Management.

Morse, A. (2015). Peer-to-peer crowdfunding: Information and the potential for disruption in consumer lending. *Annual Review of Financial Economics, 7*, 463–482.

Nielsen, K. M. (2007). Institutional investors and private equity. *Review of Finance, 12*(1), 185–219.

Olson, M. (1965). *The theory of collective action: Public goods and the theory of groups.* Cambridge: Harvard University Press.

Pierrakis, Y., & Collins, L. (2013). *Banking on each other: Peer-to-peer lending to business: Evidence from funding circle.* Project ReportLondon, UK: NESTA.

Robert, L. P., & Romero, D. M (2016). The influence of diversity and experience on the effects of crowd size. *Journal of the Association for Information Science and Technology, 68*, 321–332. https://doi.org/10.1002/asi.23653.

Signori, A., & Vismara, S. (2018). Does success bring success? The post-offering lives of equity-crowdfunded firms. *Journal of Corporate Finance, 50*, 575–591.

Surowiecki, J. (2005). *The wisdom of crowds.* New York: Anchor Books.

Tang, H. (2019). Peer-to-peer lenders versus banks: Substitutes or complements? *The Review of Financial Studies, 32*(5), 1900–1938.

Vallee, B., & Zeng, Y. (2019). Marketplace lending: A new banking paradigm? *The Review of Financial Studies, 32*(5), 1939–1982.

Vella, F. (1998). Estimating models with sample selection bias: A survey. *Journal of Human Resources,* 127–169.

Wald, A., & Wolfowitz, J. (1940). On a test whether two samples are from the same population. *The Annals of Mathematical Statistics, 11*(2), 147–162.

Wei, Z., & Lin, M. (2016). Market mechanisms in online peer-to-peer lending. *Management Science, 63*(12), 4236–4257.

Weth, M. A. (2002). The pass-through from market interest rates to bank lending rates in Germany. *Deutsche Bundesbank, economic research Centre discussion paper* (11/02).