ÁKOS ACZÉL, LAJOS TAMÁS SZABÓ

BIRDS OF A FEATHER INDEBTED TOGETHER — SOLUTIONS TO THE INFORMATION PROBLEM IN THE CASE OF MORTGAGE LOANS

MNB WORKING PAPERS | 3

2025
DECEMBER

BIRDS OF A FEATHER INDEBTED TOGETHER — SOLUTIONS TO THE INFORMATION PROBLEM IN THE CASE OF MORTGAGE LOANS

MNB WORKING PAPERS | 3

2025

DECEMBER

MNB Working Papers 2025/3

Birds of a Feather Indebted Together — Solutions to the Information Problem in the Case of Mortgage Loans*

(Madarat tolláról embert eladósodott barátjáról avagy Megoldások a jelzáloghitelezésnél jelentkező információs problémára)

Written by Ákos Aczél** and Lajos Tamás Szabó***

Budapest, December 2025

Published by the Magyar Nemzeti Bank

Publisher in charge: Milán Farkas

Szabadság tér 8-9., H-1054 Budapest

www.mnb.hu

ISSN 1585-5600 (online)

*We would like to thank Ádám Zawadowski, Ibolya Schindele, Ádám Szeidl, Yann Bramoullé, Fernando Alvarez, Christian Bayer, Ágoston Reguly, Gábor Békés, Péter Gábriel, Alice Kuegler, Róbert Lieli, Gergely Lakos, Leonardo Melosi, Lili Márk, Anna Naszódi, Andrea Weber and Arieda Muco for their valuable comments. We are also grateful to the participants of the Quantitative Finance and Financial Econometrics Conference, the Conference in Applied Econometrics using Stata, the XXV Applied Economic Meeting, the Annual Meeting of the Association of Southern European Economic Theorists, the International Association for Applied Econometrics Annual Conference 2024 and the European Economic Association Annual Congress 2024.

**Central Bank of Hungary, Szabadság tér 8-9., 1054 Budapest, Hungary and Central European University, Quellenstraße 51. 1100 Wien, Austria

***Central Bank of Hungary, Szabadság tér 8-9., 1054 Budapest, Hungary

Contents

Abstract

We examine peer effects in mortgage borrowing decisions. We find that having financially literate colleagues improves the borrowing decisions of financially less literate co-workers. Interest rates on the mortgage loans of these co-workers are significantly lower than for similar employees, whose peers have lower financial literacy. The magnitude of the effect is economically significant, amounting to roughly 4 to 5 monthly instalments until maturity. The results are heterogeneous: advice is more valuable for borrowers with low mathematical skills, and the peer effect is considerably higher in districts, where competition is weaker among banks. We also find that introducing a standardised loan product can offset the impact of the peer effect by making the decision problem of borrowers less complex.

JEL Classification: J24, G21, G41

Keywords: peer effects, skills, borrowing decisions, mortgage loan, standardised loan product

Összefoglaló

Tanulmányunkban a társak hatását vizsgáljuk a jelzáloghitel-felvételi döntésekben. Azt találjuk, hogy egy pénzügyileg jártas kolléga tud segíteni egy pénzügyekben kevésbé jártas munkatársnak. Ezen munkatársak jelzáloghitelének kamata szignifikánsan alacsonyabb, mint azon alkalmazottaké, akiknek a munkatársai pénzügyileg kevésbé tájékozottak. A hatás mértéke közgazdaságilag szignifikáns, a lejáratig 4-5 havi törlesztőrészletnek felel meg. A hatás azonban heterogén: a tanács értékesebb azon hitelfelvevők számára, akiknek alacsony a matematikai képességük, és azon járásokban nagyobb a munkatársak hatása, ahol a bankok közötti verseny gyengébb. További eredményünk, hogy egy szabványosított hiteltermék bevezetése ellensúlyozhatja a munkatársak hatását, mivel a döntési probléma ezáltal kevésbé összetetté válik.

1 Introduction

Taking out a mortgage loan is one of the most important decision of people make from a financial perspective. Mortgage loans are complex products, and considerable effort and financial knowledge is required to collect understand and compare the information on the available bank offers[1] Moreover, professional help does not necessarily serve the borrowers due to conflict of interests of the intermediaries.[2] As a result consumers often make mistakes as demonstrated for example by (Agarwal, Chomsisengphet, Liu, & Souleles, 2015) and (Agarwal & Mazumder, 2013). Little is known about the factors that drive these mistakes.

In this paper we analyse two salient remedies to cope with financial mistakes: informal advice and regulation. In our special setup we are also able to account for the interaction of these two aspects. We show that the informal advice of colleagues is relevant in the decision on a mortgage loan and that the financial literacy of the colleagues matters as well. We proxy financial literacy by mathematical skills and by experience on the mortgage loan market.[3] To identify the impact of peers' financial literacy we use the variation in timing: upon taking out a mortgage, some borrowers have highly skilled colleagues who are already experienced with mortgage loans, while other borrowers do not have such colleagues. Therefore, we compare borrowers who are similar in most characteristics but differ in respect of the financial literacy of their peers. We find that borrowers who have low skills, but have the financially most literate colleagues tend to take mortgage loans with interest rates that around 0.2 percentage point lower than similar borrowers who have less literate colleagues, which translates into savings that amount to four to five monthly instalments when considering the total term of the mortgage. Finally, we show that peer effects fade with regulation – for example, with the introduction of a standardised loan product – suggesting that regulation may help alleviate the information problem of households.

To study this issue, we combine three administrative datasets that cover the entire universe of borrowers in Hungary. We merge the Hungarian credit registry database for the period 2015-2019 with the Hungarian administrative linked employer-employee database. In this manner, we can establish a co-worker network and see borrowing decisions as well. From the credit registry, we use key loan characteristics (e.g. loan amount, term of the loan, repricing period, bank of origination etc.) as well as demographic variables such as age and district of the property. Using data from the National Tax and Customs Administration we can observe balance sheet information as well as the four digit NACE code. We also use skill sets for occupational categories from the ONET database[4].

In our empirical strategy we use the interest rate of the loan as the main dependent variable. We look at that as the price of the loan while we control for most credit risk related demographic variables, including the borrower's own mathematical skill. In some specifications we include a large set of loan characteristics as well. Moreover, we also analyse our main specification with alternative outcomes where we replace the interest rate of the loan with other loan characteristics such as repricing period, maturity and total loan amount. In one specification we also include bank fixed effects to test if the choice of bank may be part of the informal advice. Our key independent variable is the average financial literacy of a borrower's peers. We proxy financial literacy with the level of mathematical skill from ONET and experience in the mortgage loan market, as it is well documented in the literature that these aspects are important in financial decisions (e.g. (Agarwal & Mazumder, 2013), (Lusardi & Mitchell, 2007), (Lusardi & Mitchell, 2014), (Folke, Gjorgjiovska, Paul, Jakob, & Ruggeri, 2021), (Frijns et al., 2014)).

If studying peer effects, at least two challenges arise that need to be addressed. First, the definition of peer groups may be ambiguous, as researchers rarely have detailed information on the subject's social network. We treat all of the colleagues of an

[1] For example, (Lusardi & Mitchell, 2014), (Gerardi, Goette, & Meier, 2010), (Campbell, 2016), (Barr, Mullainathan, & Shafir, 2008), (Heidhues, Johnen, & Kőszegi, 2021)

[2] e.g. (Inderst & Ottaviani, 2012), (Fan & Chatterjee, 2017)

[3] Based on, for example (Gerardi et al., 2010), (Lusardi & Mitchell, 2007), (Lusardi & Mitchell, 2014), (von Gaudecker, 2015) and (Frijns, Gilbert, & Tourani-Rad, 2014)

[4] ONET reports required level of skills for 873 occupations

employee as her peers, but we make a distinction between the relevance of different peers. Based on the homophily literature (e.g. (Currarini, Jackson, & Pin, 2009)) we assign a high weight to colleagues who share similar characteristics with the borrower and a low weight to colleagues who are less similar. In doing so we follow the approach of (Giorgi, Frederiksen, & Pistaferri, 2020). Our main results are robust to the variation of applied weighting scheme.

Second, following the terminology of (Manski, 1993) we face the reflection problem and the problem of correlated effects. We focus on the past decisions of peers, and can thus avoid the problem of simultaneity in peers' decisions (reflection problem). Regarding correlated effects, we pursue two strategies. We include a large set of control variables including demographic controls, company fixed effects and the interaction of industry and time fixed effects. Therefore, we compare borrowers within the same firm and same time period also taking into account industry level and local shocks. Moreover, we include the sales of the company to control for company level time-varying shocks that are related to economic conditions of the company. We argue that after including this large set of controls it is unlikely that plausible endogeneity problems could distort our results. Our identification is based on the change in the composition of borrowing colleagues over time. As a robustness check we also control for regional-level bank supply shocks by adding the interaction of bank, district and time fixed effects.

However, even including such a large set of fixed effects does not fully eliminate the threat of time-varying local shocks at the peer group level. We perform a placebo test: we replace the peer group that contains only borrowers who have taken out a mortgage loan in the *recent past* with a peer group that contains borrowers who take out a loan in the *near future*. We argue that if the peer effect functions according to our narrative then the peers taking out a loan in the near future are assumed to have no impact on current borrowers. However, if the impact that we measure is solely due to time-varying common shocks and not due to the peer effect, then that shock should have an impact on both the borrowers and the peers who take their loans in the future. Consequently, in that case our placebo test would result in significant coefficients. The fact that the key coefficient in our placebo test is insignificant and the point estimate is small suggests that time-varying common shocks do not distort our results.

We also analyse whether peers are related to the borrower's choice of bank. We model the borrowers' choice by running a conditional logit model where the borrower has not only two but eleven[5] choices. The key dependent variable in this analysis is the share of colleagues in the borrower's network who have chosen the given bank. We also run a placebo test, similar to the previous placebo test, to ensure that no time-varying common shocks distort our results.

How could regulation improve mortgage loan market outcomes? How could regulation be related to peer effects? To answer these questions we analyse the impact of introducing a standardised product, the Certified Consumer Friendly Housing Loan (CCFL) on the mortgage loan market at the end of 2017. One of the goal of the Central Bank of Hungary in introducing such a loan was to support borrowers by decreasing the complexity of the problem of finding a suitable loan. We exploit the exogenous variance in the spread of the product among districts over time. We argue that the spread of this product after using a large set of controls in the regression – including a bank competition proxy – is orthogonal to our key variable of interest: the average financial literacy of peers. As a robustness check we also carry out an IV strategy. The validity of the IV is based on the assumption that there was random variation in the spread of CCFL in the first few months after the introduction of this product.[6]

To complement all of the analysis above we examine small firms where we can confidently assume that all employees know each other and we do not need to impose any network structure. The number of mortgage borrowers at small firms is limited; therefore we test if the borrowers' interest rates are smaller if they have at least one colleague who has high mathematical skill and has also taken out a loan in the recent past. It is important that although the share of borrowers is similar in small and large firms, in the case of small firms the chance that there is at least one borrowing colleague is small. Consequently, we argue that whether or not a borrower has a borrowing colleague with high mathematical skill is essentially random. In this setup, the identification variation is across firms. We also run a similar placebo test as in the main specification. As robustness checks, we

[5] We have ten large banks plus one category for all the remainder small banks adding up to eleven institutions in the sample.

[6] We argue that the spread shortly after the introduction of the product was exogenous. Our instrument is based on the share of CCFL among all housing loans at the district level in the first half-year after its introduction of CCFL. The idea behind such an IV is that the introduction of the programme was based on the cooperation of CBH and the product development departments of banks. Therefore, depending on a given bank's product development capacity the bank could either come up with the product early, or spent more months to be able to finalise the exact details of the loan. As a result in some districts CCFL was already offered by most banks in the first few months after its introduction of CCFL, while in some others the product appeared later. The instrumented variable is CCFL share in housing loans per districts observed on a monthly basis.

also carry out matching and coarsened exact matching exercises and compare borrowers who are not only similar in their own characteristics and in their firm's characteristics, but also have similar chance of having a financially literate peer.

According to our main results a borrower who has financially literate peers tends to take out loans at a lower interest rate compared to one whose peers have low skills. There is considerable heterogeneity in this effect, and the lower the own skills of a borrower the stronger the peer effect is. Similarly, peer effect is stronger in districts where bank competition is low. Comparing a borrower who either has low skills or lives in districts where competition is low, the estimated effect of having financially literate peers is between 0.15 - 0.24 percentage point compared to a similar borrower who does not have such colleagues. These results are consistent with the narrative that peers' advice is most valuable if it is difficult for the borrower to find a good loan. The estimated effect accounts to four to five monthly installments taking into consideration all repayments until maturity.

Among the alternative outcomes, we find a statistically and economically significant impact only in the case of the loan amount.The presence of financially literate peers leads to higher loan amounts, especially in the case of borrowers with low mathematical skills: the measured impact can be up to 14 per cent.[7] With the interest rate as the dependent variable and including all loan characteristics as controls, we obtain the same result in qualitative terms as the original regression without additional controls. Including bank fixed effects decreases the measured impact of peers by one-third, suggesting that one important channel of peer effect is the choice of bank.

Results on bank choice strengthen these observations. If all of the colleagues in a borrower's the network of a have chosen the same bank, then the borrower's chance of choosing that bank is more than 20 per cent higher[8] compared to a borrower whose colleagues in his network have not chosen that bank. Placebo tests again show no significant result, suggesting that no time-varying common shocks distort the results.

These results also highlight the unfairness of the mortgage loan market. Borrowers, who are not lucky enough to have a colleague who can help them tend to take out loans that are more expensive than the loans of borrowers who get help. We find that the spread of CCFL is negatively related to the peer effect and also negatively related to interest rates. IV results reinforce these findings. Therefore, we argue that CCFL has potentially helped borrowers to solve their information problem and also improved the fairness of the mortgage loan market.

Our research contributes to four strands in the literature. First, to the best of our knowledge, there is no research on how peers affect the choices of borrowers in the mortgage loan market. The closest paper is the research of (Maturana & Nickerson, 2018), which analyses the impact of co-workers on the decision of whether to refinance a mortgage. Although the topic may seem similar, their research is different in three relevant aspects. The problem that borrowers face in our research is mainly related to complexity and search costs, while in (Maturana & Nickerson, 2018) it is a well-defined decision that lacks complexity issues. Second, we also show results also on the skills of borrowers and peers. Third, we also analyse the impact of regulation on peer effect. There are also other papers that focus on peer effects in financial decisions (e.g. (Ouimet & Tate, 2020), (Duflo & Saez, 2002)), but again, the problem of the actor in our case (complexity and search costs) is strikingly different (to participate or not in company pension or investment programmes). Lastly, according to our current understanding it is unique in the finance literature that we have different peer groups for each individual. There are studies (see previous examples) which deal with peer groups in financial decisions, but in these specifications the peer groups do not vary on individual level, while in our setup they do.

Second, our results also add to the literature on financial advice. (Campbell, 2016) argues that research on *formal* financial advice shows discouraging results (see e.g. (Foerster, Linnainmaa, Melzer, & Previtero, 2017) and (Bhattacharya, Kumar, Visaria, & Zhao, 2024)). There is also rare evidence that *informal* financial advice can improve the decision on complex financial problems. In a recent paper, (Ambuehl, Bernheim, Ersoy, & Harris, 2022) argues that there is lack of evidence on whether informal advice improves or deteriorates financial decisions. According to our results, financially literate peers improve the borrowing decision, while peers with lower skills are not relevant in the decision.

[7] According to the results those borrowers with low mathematical skill whose experienced peers' average mathematical skills is at the 90th percentile of the distribution of mathematical skill tend to take 14 per cent higher amount of loan on average than those, whose experienced peers' average mathematical skill is at the 10th percentile.

[8] Regarding the baseline probability of choosing that bank. For example, if the baseline probability of that borrower i chooses bank A is 10 per cent, then it increases to 12 per cent.

Third, we show that peer effects contribute to inequality. (Chetty et al., 2022) argues that social capital may help in social mobility. We find similar results, as high social capital ('having smart colleagues') improves the borrowing decision. Conversely, our results also highlight that social capital may increase inequality.

Lastly, our research also contributes to the findings of the literature on financial regulation. The results of the literature are mixed regarding the optimal regulation. On the empirical side for instance (Nicholson, Skelton, & Tarr, 2019) show the drawbacks of regulation on the Australian mortgage loan market, while (Barr et al., 2008) argues why the Truth in Lending Act was inefficient to improve the US mortgage loan market and what other measures could help. (Agarwal, Chomsisengphet, Mahoney, & Stroebel, 2015) show how the 'CARD act' improved credit card market conditions. On the theoretical side, there are also several studies that try to pin down how financial markets work and then suggest adequate regulation. (Heidhues et al., 2021) argue that caps on hidden prices of products may help in achieving perfect competition and also mention that certificates are less likely to make a difference. Both (Campbell, 2016) and (Barr et al., 2008) call for special disclosure rules. We find that special regulation may be able to improve mortgage borrowers situation, decrease the peer effect and therefore work against inequality. The regulation that we analyse however, has several features that were implemented at the same time (e.g. cap on margin, issue certification by CBH, control over special terms of loans), and therefore we cannot pin down the impact of those features one-by-one, but only on aggregate.

Our research loosely connects to the herding literature, where agents change their mind in a financial decision and follow the crowd (see (Spyrou, 2013), for example). The peer effect can cause herding, but during our examined period the economic environment was calm and we did not observe herding behaviour. Furthermore, in herding models agents observe each others' previous financial choices and make their own decision (see e.g. (Cipriani & Guarino, 2014) or (Park & Sabourian, 2011)); in our setup, however, borrowers get informal advice and are not just simply copying others' decisions.

The study is organised as follows. In Section 2 we present the databases and descriptive statistics. In Section 3 we show the identification strategy. We carry out the estimation and discuss the results in Section 4. We develop an alternative specification using only small firms in Section 5. Conclusions are presented in Section 6.

2 Database

We use three different data sources to test our research hypothesis. We merge these three databases using individual identifiers and occupational codes.

2.1 HUNGARIAN STATE TREASURY DATABASE

The database of the Hungarian State Treasury (HST) contains the entire population, who had at least one day of legal work between 1997-2019. The following variables are used in the study:

- district level place of residence,
- four digit occupational code (FEOR),
- firm identifier.

There are 198 districts in Hungary, with an average population of 55 thousand. One district can be considered as a local labour market, because according to the 2011 Census, except for Budapest and its surroundings, 80% of the residents live and work in the same district. This ratio is 72% for Budapest and its commuting area.

For every debtor, we choose the firm, where he earns the most as an employer. Based on this definition we can identify the co-workers, which is an essential measure in our calculations.

2.2 ONET DATABASE

We define skills based on the ONET database, which contains detailed information on the level of skills that are needed for specific occupations. Based on the literature mathematical skill is one of the most important skills in financial decisions (e.g. (Lusardi & Mitchell, 2007), (Lusardi & Mitchell, 2014), (Folke et al., 2021)), and therefore we focus on the mathematics score. It is measured on a scale of 0 to 100 which we normalise to be between 0 and 1 in our estimations. The ONET dataset is based on the international system of occupations categories (ISCO-08). We can merge the international occupational categories with the four-digit Hungarian counterparts (FEOR-08) but there are cases, where we cannot find an exact match for some Hungarian occupational categories. The matching rate is 73%. To increase the number of observations we can use in our analysis, we average the mathematical score for every three-digit occupational category and use this average if the four-digit mathematical score is missing. In this case, we have mathematical scores for 87% of the total observations.

2.3 CREDIT REGISTRY DATABASE

The Credit Registry (CR) database contains all loans taken out by individuals. We use the 2015-2019 time period [9]. We observe the day the loan was taken out, the total loan amount, the interest rate, maturity and repricing period of the loan, the bank that granted the loan, the place of residence of the borrower, and also the age and wage of the borrower.

After merging all datasets we have almost 80 thousand observations. In the main estimations due to constraints on the minimum number of employees, we are able to use more than 73 thousand observations in most estimations. Table 1 shows the descriptive statistics of the merged dataset.

[9] The reason for this is that the earliest point in time where we have detailed information on loans is 2015 and we wanted to end our analysed period before the Covid crisis started

Table 1
Descriptive statistics of the merged dataset

	Mean	10th percentile	90th percentile	Standard deviation
Interest rate	5.1	3.5	6.7	1.3
Maturity (year)	15.3	7.0	25.0	6.5
Repricing period (month)	76.4	12.0	121.0	54.8
Principal (EUR)	24885.0	7142.9	47619	20419.4
Wage (EUR)	1323	519.9	2288.3	816.3
Age	36.6	27.0	48.0	8.3
Bank concentration	0.3	0.2	0.4	0.1
Mathematics	0.4	0.3	0.6	0.1
Mathematics average of colleagues	0.4	0.3	0.5	0.1
Observations	79135			

All mortgage loans are denominated in HUF; we calculated EUR values using an exchange rate of EUR/HUF 314.95, which was the average exchange rate in 2015-2019.

3 Empirical approach

3.1 NETWORK

We face two types of challenges when we want to identify the peer effect on borrowing decisions. One problem is related to the identification of peer groups, and several endogeneity-related concerns also arise regarding the estimation. We start by describing how we construct peer groups from the available data and then turn to discussing the endogeneity problems.

To define peer groups that may be relevant in helping to choose a good loan the researcher would ideally know the social network of borrowers. However, data on exact social network is rarely available, and therefore researchers often use other alternative data sources. Several studies exploit the fact that colleagues tend to know each other well and potentially share information on their own experiences regarding financial products ((Ouimet & Tate, 2020)), charity activities ((Lieber & Skimmyhorn, 2018)), consumption, parental leave ((Dahl, Loken, & Mogstad, 2014)), etc. We observe all of the colleagues of the borrowers, and therefore we can thus construct peer groups of colleagues.

We do not have precise information on the informal network of colleagues, and therefore we build on the homophily literature to define a realistic network structure. We cannot assume that a peer group consists of all of the colleagues of a borrower: for example in the case of large companies it is likely that a borrower knows only a small fraction of all employees. We do not have information regarding these ties. However, following (Giorgi et al., 2020) we assume that employees with similar characteristics have a high likelihood of knowing each other. Their approach is consistent with the literature on homophily: (McPherson, Smith-Lovin, & Cook, 2001) surveys the literature and argues that more similar people tend to be more influential. They argue that the most relevant characteristics regarding homophily are race, religion, gender, age, education and occupation. (Currarini et al., 2009) build a structural model that rationalises that similar people tend to make more friendships than dissimilar ones. (Stolper & Walter, 2019) also find that homophily plays and important role in financial decisions.

Following (Giorgi et al., 2020) we create a distance measure between every pair of employees. This measure is based on education and occupation, however as we do not have data on education we use wage as a proxy for it. We also use occupation and add age as well when calculating the distance metric. We define distance in the following way. First we evenly distribute all employees into three buckets of age (younger than 32 years, between 32-41 years, older than 41 years) and three buckets of wage (less than €953 , between €953 - €1587, more than €1587). We also split borrowers into three buckets based on their occupation (managerial, white collar and blue collar occupations). We assign integers 1,2 and 3 to the three buckets and then define the distance as the sum of the absolute deviations between the buckets of employee i and employee j:

$$d_{i,j} = (|age_i - age_j| + |wage_i - wage_j| + |occupation_i - occupation_j|) \tag{1}$$

We use the distance measure $d_{i,j}$ to assign weights to all colleagues of the borrowers. We continue to follow (Giorgi et al., 2020) and use the distance measure as weights when we calculate peers' average characteristics. Similarly to (Giorgi et al., 2020) we also use quadratic weighting scheme in order to add higher weights to the more similar colleagues and lower weights to the less similar colleagues. The weights are defined as:

$$\omega_{i,j} = (d_{i,j} + 1)^{-2} \tag{2}$$

The weighted average of any characteristic x (for instance mathematics score) of the colleagues of borrower i who works at company c and takes out a loan in period t is then calculated by the formula:

$$\overline{x_{i,c,t}} = (\sum_{i,j\neq i} \omega_{i,j})^{-1} \sum_{i,j\neq i} \omega_{i,j} x_{j,c,t} \tag{3}$$

As this weighting scheme is somewhat ad-hoc, we also apply linear weights and run the main specifications based on that weighting scheme as well. Changing the weighting scheme does not significantly change our results; the estimation results can be found in Appendix A.

3.2 BASELINE MODEL

Turning to the estimation of peer effects there are at least three challenges that should be handled to be able to identify peer effects. These are i) simultaneity, ii) those omitted variable biases that are due to the fact that peer groups are formed endogenously; and iii) common shocks. (Manski, 1993) refers to i) as reflection problem while ii) and iii) are known together as correlated effects. To avoid the problem of simultaneity we focus on peers who made their choice earlier than the borrower: at least three months and at most one year before. We assume that a borrower who is currently trying to find the appropriate loan has no direct effect in this search process on a peer who took out a loan in the aforementioned time window. This assumption is based on the observation that financial experience is an important building block in financial literacy ((Frijns et al., 2014), (Weiner, Baron-Donovan, Gross, & Block-Leib, 2005), (Mandell, 2006)). One could say that good understanding of the mortgage loan market requires not only good skills, but experience in taking out loan as well.

An example for the second problem (sorting) is that if all colleagues have a high level of education, it obviously helps them in making a good choice among mortgage loan offers. It can happen that the researcher does not measure the impact of peers on each other but rather measures the similarity of peers regarding the unobserved factor of education. In this case, it is this omitted variable that makes peers' decisions similar and the estimated peer effect is biased. This problem can be partially solved by including company fixed effects. These can capture the impact of all company level time-invariant unobserved factors (such as a special offer of a bank to a given company). Time-variant company level shocks, such as the impact of sorting due to laying off (hiring) employees whose skills are low (high) during an economic downturn (upturn) are not captured by company fixed effects. We interact month fixed effects with industry fixed effects to control for time varying industry level shocks. These fixed effects also control for the downward trend in interest rates during the sample period. We believe that this control rules out most time-varying unobserved shocks. Nevertheless, company level, (or even peer group level) time-varying shocks could still distort our estimation. To show that such common shocks do not bias our result we carry out a placebo test that is discussed after the main results have been shown.

Following the terminology of Manksi what we want to identify is the contextual (exogenous) peer effect: the effect of the characteristics of peers on the choice of the borrower. The equation that we estimate by OLS regression is the following:

$$y_{i,c,t} = \beta X_{i,c,t} + \delta \overline{X_{i,c,t-1}} + \phi S_{c,t} + \Omega_c + \Theta_{g,t} + \varepsilon_{i,c,t} \tag{4}$$

where $y_{i,c,t}$ stands for the interest rate of the mortgage loan of borrower i, who works at company c and has taken out the loan in period t. As a mortgage loan is a product that is usually not taken out more than one or two times in a borrower's lifetime, in our narrow time window of five years we observe borrowers taking out loans only once. However, we use t subscript to indicate that our data has a time dimension as well; this is important to be able to estimate company and time fixed effects, as we discuss below. Similarly, we use c subscript to indicate that in our estimation we keep track of the company of the employee as well.

$X_{i,c,t}$ includes all of the borrower level explanatory variables that we include in the regression. These are wage, age and age squared of the borrower and the own mathematics score associated with the borrower's occupation. $\overline{X_{i,c,t-1}}$ includes the colleagues' weighted average of the same variables. The weighted average only includes the colleagues, who have taken out a mortgage loan in the time window of a minimum of three months and a maximum of twelve months before the borrower takes out a loan. $S_{c,t}$ stands for the logarithm of sales of company c in period t. This control variable can capture shocks that are related to the changing economic conditions of the company. Ω_c is a company fixed effect and $\Theta_{g,t}$ is an industry $\times$ time (month) fixed effect, while $\varepsilon_{i,c,t}$ is an individual level shock (error term). The company fixed effect absorbs all of the unobserved time-invariant company-level shocks, while the interacted industry-time fixed effect absorbs all of the industry level time-varying shocks. By including these fixed effects, we compare borrowers who take out their loan within the same firm also accounting for time-varying industry-level shocks.

Our identification comes from the random variation in the composition of colleagues who take out a mortgage loan in the given period of time. In some periods mortgage loans are taken out by people whose mathematics score is high, while in other periods they are taken out by people whose mathematics score is low. In theory the change of colleagues (some employees leave the firm and there can be new hires as well) could also add to the random variation, but the share of borrowers who change company is negligible in our sample (less than 1.5 per cent of all employees change company during this period, and this figure is lower inthe case of borrowers).

One threat to our estimation is that in addition to the impact of colleagues, borrowers' own credit risk has a strong impact on the interest rate of the mortgage loan. If borrowers with high credit risk are systematically surrounded by colleagues with low mathematics scores, for example, then our estimated parameter on colleagues' mathematics score would be biased. By including the age, wage, company and own mathematical skill of the borrower in the regression, we believe that credit risk is captured and that the aforementioned omitted variable bias does not distort our estimation.

3.3 PLACEBO TEST

We apply a placebo test to show suggestive evidence on the lack of time-varying common shocks. As our identification is based on the change over time in the composition of peer groups, we cannot directly control for time-varying shocks at the peer group level. For instance, by including company $\times$ time fixed effects, we would not only control for common shocks, but would also remove the aforementioned variation in the composition of peers over time (that serves as the source of identification). Consequently, instead of directly controlling for shocks, we implement a placebo test that should result in significant parameters if time-varying common shocks distort our results. Non-significant results mean that we cannot reject the hypothesis of no time-varying common shocks.

In our placebo test we focus on colleagues who take out a loan *after* the borrower. If it is true that experienced peers help current borrowers, then peers who take out their loans later than the borrower should have no impact on the borrower's interest rate. However, if our results are governed by common shocks, then these shocks affect both the current borrower and her peers, who borrow later. Therefore, the estimated coefficient should be significant. If it is not significant then it suggests that common shocks do not distort our results.

We estimate the following specification, where the only difference compared to our baseline model is the change in the timing of peers' borrowing:

$$y_{i,c,t} = \beta X_{i,c,t} + \delta \overline{X_{i,c,t+1}} + \phi S_{c,t} + \Omega_c + \Theta_{g,t} + \varepsilon_{i,c,t} \tag{5}$$

3.4 FURTHER SPECIFICATIONS

We also run specifications where we include several other loan related variables such as the repricing period, maturity and principal of the loan. These variables are determined simultaneously with the interest rate; therefore estimating a regression where all of these are included may lead to a bad control problem. However, we view these regressions more as a robustness check: if the results remain qualitatively the same it suggests that leaving out these variables from the original regression does not distort our results. Finally, we also run regressions with alternative outcomes.

3.5 BANK CHOICES

We suspect that bank choice may be a channel of the peer effect. We test if borrowers tend to choose the same bank as their peers. To do so we model the choice of borrower with a conditional logit model, where the borrower can choose from all of the banks that have at least one branch in the district where the borrower lives (see a more detailed description of the choice model in Appendix C).

The dependent variable is the choice of the bank which takes the value of 1 if a borrower has chosen a given bank and 0 otherwise[10]. The marginal effects of the choice model can be computed based on the relationship between the estimated coefficients and the probability of a borrower choosing a given bank:

$$\frac{\partial p_{i,j}}{\partial x_{i,k}} = \begin{cases} p_{i,j}(1-p_{i,j})\beta & j=k \\ -p_{i,j}p_{i,k}\beta & j \neq k \end{cases} \tag{6}$$

where $p_{i,j}$ is the probability of borrower i chooses bank j, $x_{i,k}$ is a vector that contains the variables that may affect the choice of borrower: the share of the borrower's loan taking colleagues have chosen the given bank, the share of all colleagues who have chosen a given bank, bank fixed effects for all alternatives (interact bank fixed effects with borrowers' other potential choices).

A one unit increase in $x_{i,j}$ increases the probability of choosing alternative j by $p_{i,j}(1-p_{i,j})\beta$ and decreases the probability of choosing any of the other alternatives by $p_{i,j}p_{i,k}\beta$. Therefore, to obtain the marginal effect of an increase in the share of peers choosing a given bank in the borrowers network both own and cross-effects should be considered.

As a final step, we run a placebo test to show that no time-varying common shocks distort our results.

3.6 CERTIFIED CONSUMER FRIENDLY HOUSING LOAN

At the end of 2017, The Central Bank of Hungary (CBH) introduced the Certified Consumer Friendly Housing Loan (CCFL) to support borrowers in finding a suitable product on the mortgage loan market. Working out together Hungarian banks and the CBH elaborated the details of these products[11] to ensure that they were standardised and favourable for borrowers. The margin above financing costs was capped at 3.5 percentage point. The disbursement fee was limited at 0.75 percentage point of the current principal while the prepayment fee was also capped at 1 percentage point of the prepaid amount. Repricing periods could be either 5/10/15 years or last to maturity (i.e. fully fixed interest rate, no repricing). Credit assessment could not exceed 15 working days while loan disbursement (if all requirements were met) was maximised at 2 working days. Monthly payments of CCFLs were determined according to annuities. CCFLs were developed by commercial banks and the CBH together. To obtain the certification that enables banks to use the logo of CCFL the offered product had to meet all of the criteria set by the CBH. Introducing CCFL decreased the complexity of choosing a suitable loan product as these loans were better in most characteristics compared other available loans on the market and it was also advertised by the CBH[12].

Banks were urged to participate in the programme by the CBH, but were not legally obliged to do so. The terms on these loans were different from bank-to-bank, but the limits shown above lead to similar products regardless of which bank offered them. In line with the capacities of banks' product development departments the introduction of the product differed at the bank level: some banks were able to introduce it early, while some banks had less capacity and were only able to roll out a CCFL product later. In theory, banks could also start to offer CCFLs in a strategic way: by offering it only if the competition they face made it necessary to do so and by continuing to offer more expensive products when low a level of competition allowed that. We consider these aspects in our empirical strategy concerning CCFL.

To understand what impact CCFL had on borrowing decisions, we focus on the first 2 years of the newly introduced product. We argue that the spread of this product is plausibly exogenous after controlling for bank competition, company fixed effects and the interaction of time and industry fixed effects. First, we measure if informal advice incentivised borrowers to choose CCFL. We also include an interaction term of the share of CCFLs and the average mathematical skill of colleagues. This variable helps to understand whether informal advice changes when CCFL becomes widely spread. We expect that high mathematical skill of colleagues is related to higher chance of exploiting the good opportunity of CCFLs, but, when the loan is widely spread this effect may change.

[10] All alternatives that could be chosen by the borrower appear as separate observations.

[11] https://www.mnb.hu/en/supervision/news/certified-consumer-friendly-housing-loans-to-reduce-interest-rates

[12] Financial Stability Report (November 2017) , box 2 and Financial stability Report (May 2018) chapter 8.

In a separate regression we have the interest rate as independent variable and by including the interaction of share of CCFL in a district and the average mathematical skill of colleagues we test how the spread of CCFL impacted informal advice with respect to the interest rate on the loan. Finally, although we think it is difficult to find any plausible endogeneity story, as a robustness check we carry out an IV strategy that is based on the random variation of the spread of CCFLs by banks in the first few months after official introduction of the product. The validity of the instrument relies on the assumption that how quickly banks were able to develop the new product was related solely to the capacity of their product development departments and is orthogonal to the problem of informal advice - especially after controlling for company and industry $\times$ time fixed effects and bank competition.

3.7 SMALL FIRMS

Lastly, we run specifications where we include only small firms. The advantage of such a special sample is that there is no need to assume any network structure in the case of small firms: it is a plausible assumption that everyone knows each other in a firm where the number of employees remains below 50. We also run robustness checks with alternative cutoffs for firm sizes and also carry out a similar placebo test as in the main regression. In this specification our key independent variable is whether the borrower has a colleague who has high mathematical skill and recent experience from the mortgage loan market. We cannot use the average mathematical skill of peers, as in several cases the borrowers have no loan-taking colleagues. With this specification we also highlight that not only the average skill of colleagues but also having at least one financially literate colleague is important in the borrowing decision.

As we cannot include company fixed effect for such small firms (there are not enough observations per company to do so), we carry out a coarsened exact matching exercise as a robustness check as well. This special type of matching allows us to efficiently find types of borrowers who are similar in many aspects except having financially literate colleagues or not.

4 Results

4.1 BASELINE RESULTS

In Table 2 we present 3 different specifications: first without company fixed effect in Column (1), and then with company fixed effects (Column (2) and (3)). In specification (1) the estimated peer effect is -0.63, which means if we compare two similar borrowers with the only difference between them being that one has peers whose average mathematics score is 0.5 (90^{th} percentile) and the other has peers whose average mathematics score is 0.2 (10^{th} percentile), then the first borrower's estimated interest rate is 19 basis points (0.186 percentage point) lower than the second borrower's interest rate. The own mathematics score coefficient is -0.945, meaning that for someone, whose mathematical knowledge is at the maximum point the interest rate is 95 basis points lower than for someone, whose maths knowledge is at 0. We include company fixed effects in Specification (2), and the peer effect halves and decreases to -0.29. This result shows that even if we control for company specific (time-invariant) shocks there remains a significant unexplained part, which we consider as peer effect. In Column (3) we include company fixed effect and industry $\times$ time fixed effect. In this specification we control for time-invariant company shocks and time dependent industry shocks. The coefficient of the peer effect is -0.309 in Specification (3), which does not differ significantly from the result in Specification (2). The coefficient of own mathematics score is stable across all three specifications. In Appendix B we show the results of a regression where we also control for regional level time-varying bank supply effects. We control for bank supply by including bank fixed effects interacted with district fixed effects and with time fixed effects. The results remain qualitatively the same.

Repeating the calculation of changing from peers with a low mathematics score peers to peers with a high mathematics score results in an interest rate which is 8 basis point lower according to Specification (3). This is an economically moderate impact: considering all payments until the loan is fully repaid, the total saving amounts to one and a half monthly instalments. Two things should be taken into account regarding this result. Firstly, although the total impact is low compared to the total loan amount, it is not small compared to the cost of obtaning the effect. Unfortunately, we cannot measure how much discussion is needed between the peers in order to realise this gain. However, we imagine that exchange of information about loan taking experience happens during non-working periods such as coffee breaks or for instance when peers have lunch together. Therefore, the costs of obtaining the information is low and the total saving compared to this low costs is relatively high. Secondly, the next set of results show that the impact of peers is heterogeneous with respect to own mathematical skills and bank concentration, and therefore some borrowers gain much more from the peer effect.

4.2 HETEROGENEITY IN PEER EFFECTS

The results of the specification containing heterogeneous peer effects are shown in Table 3. The only difference compared to the previous regression is that we add an interaction term of the average mathematics score of colleagues with own mathematical skill in Specification (1) and with bank concentration in Specification (2). The results suggest that there is a considerable heterogeneity with respect to these factors. Peers' average mathematics score results in a lower interest rate only in the case of borrowers whose own mathematical skill is low: the estimated impact is twice as large (-15 bp) in the case of borrowers with mathematical skill at the 10^{th} percentile of the distribution of mathematical skills as compared to the results in the baseline specification (Column (3) in Table 2).

In the second specification we include a measure of bank concentration and we also interact this measure with peers' average mathematics score. This measure is the Herfindahl-Hirschmann index (HHI) and we normalise it to take values between 0 and 1: the higher the HHI, the stronger the bank concentration. The index is calculated taking into account the 10 large banks in Hungary and treating all of the small ones as one institution. Bank market shares are calculated based on the observed loan disbursements in our sample and calculated at the district-year level.

$$HHI_{d,t} = \sum_{b=1}^{11} bankshare^2_{d,t,b} \tag{7}$$

Table 2
Baseline specification

	(1)	(2)	(3)
	Interest rate	Interest rate	Interest rate
maths avg col	-0.630***	-0.290***	-0.309***
	(0.0929)	(0.0937)	(0.0956)
maths own	-0.945***	-0.856***	-0.847***
	(0.0475)	(0.0531)	(0.0538)
Observations	75962	73018	73016
R^2	0.094	0.176	0.184
Covariates	Yes	Yes	Yes
Time FE	Yes	Yes	No
Company FE	No	Yes	Yes
Industry x Time FE	No	No	Yes

Dependent variable: mortgage loan interest rate. Other independent variables: age, age^2, own wage, peers' weithted average wage, log(sales). Standard errors clustered on company level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

where

- d district, t time, b bank.
- bank share - sum of loans in a given district at a given time for a given bank.

We use this index as a proxy for competition among banks. Similarly to the previous result, the estimated impact is small in districts where HHI is at its average, and much larger where bank concentration is high. We show the level of heterogeneity with respect to own mathematical skill in Figure 1 and to bank concentration in Figure 2.

In Figure 1 we compare three individuals with different mathematics knowledge scores of: 0.2 (around the 10th percentile), 0.4 (around the median) and 0.6 (around the 90th percentile). The results show that for employees with high mathematics knowledge there are no peer effects, while in the case of employees with low mathematical skill the peer effect can be large. Comparing two borrowers (A and B) with low mathematical skill (10th percentile of the distribution of mathematical skill in our sample) and with the only difference between them being that borrower A has peers with low average mathematics score (10th percentile of the distribution of peers' mathematical skill in our sample) and borrower B has peers with high mathematics score (90th percentile of the distribution of peers' mathematical skill in our sample) borrower B's interest rate is expected to be 15 basis point lower than borrower A's interest rate.

Repeating this back of the envelope calculation with borrowers who live in districts where bank concentration is high (90th percentile of the HHI distribution) and the only difference between borrowers again being the average mathematics score of the peers then the borrower whose peers have the high average mathematics score is expected to have an interest rate that is 24 basis points lower than the other borrower.

Considering a typical loan with 20-year maturity (the maturity that is taken out most often) and an interest rate of 5.1 per cent (the mean interest rate in our sample) and comparing a borrower whose peers have a high mathematics score and one whose peers have a low mathematics score, the estimated effect of having peers with a higher mathematics score calculated by the baseline regression (Table 2 Specification (3)) is 8 basis points which is consistent with a saving amounting to one and a half monthly instalment[13]. The similar impact based on the regression that takes into account the heterogeneity with respect to the own mathematics score of the borrower (Table 3 Specification (1)) is 15 basis points[14]. Finally, the estimated impact based

[13] The calculated impact is the result of cumulating the difference between the two loans' monthly instalments over the 20-year period until maturity.

[14] That is consistent with savings amounting to three monthly instalments.

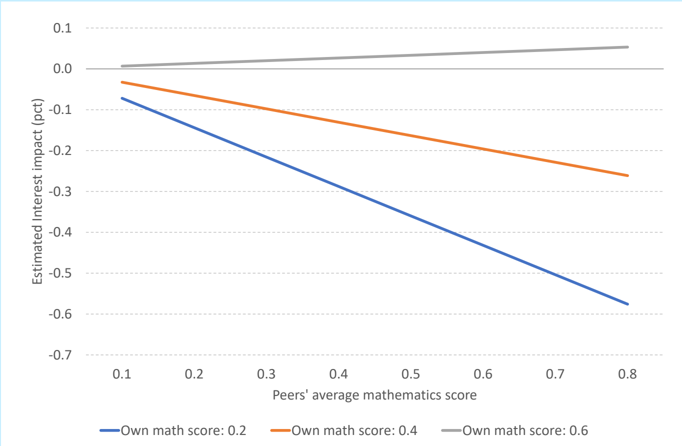
on the regression that takes into account the heterogeneity with respect to bank concentration (Table 3 Specification (2)) is 24 basis points, which is consistent with a saving amounting to four to five monthly instalments.

Table 3
Heterogeneity based on mathematics and competition

	(1) Mathematics	(2) Competition
maths avg col	-1.113***	0.745***
	(0.212)	(0.249)
maths avg col x maths own	1.966***	
	(0.461)	
maths own	-1.668***	-0.836***
	(0.201)	(0.0522)
maths avg col x bank concentration		-3.858***
		(0.865)
bank concentration		2.864***
		(0.379)
Observations	73016	66991
R^2	0.184	0.202
Covariates	Yes	Yes
Time FE	No	No
Company FE	Yes	Yes
Industry x Time FE	Yes	Yes

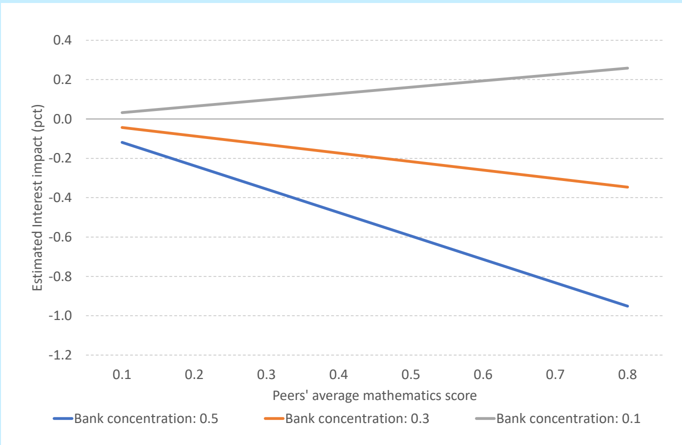
Dependent variable: mortgage loan interest rate. Main independent variables: maths avg col: weighted average math score of colleagues, maths own: own maths score, bank concentration: measured as a Herfindahl-Hirschman-index. Other independent variables: age, age^2, own wage, peers' weithted average wage, log(sales). Standard errors clustered at company level. *** p$<$0.01, ** p$<$0.05, * p$<$0.1

Figure 1
Estimated impact of the change in the average mathematics score of loan-taking peers



The three lines are selected based on the distribution of math scores: 0.2 is around the 10th percentile, 0.4 is around the median and 0.6 is around the 90th percentile. The calculation is based on Table 3 Column (1).

Figure 2
Estimated impact of the change in the average mathematics score of borrower's peers



The three lines are selected based on the distribution of bank concentration scores: 0.1 is around the 10th percentile, 0.3 is around the median and 0.5 is around the 90th percentile. The calculation is based on Table 3 Column (2).

4.3 PLACEBO TEST

The last specification controls for various borrower (demographic variables), company (company fixed effect and sales) and time-varying industry (industry $\times$ time fixed effect) related shocks. Including this large set of controls considerably decreases potential endogeneity problems, leaving time-varying peer group level common shocks as the main source of potential biases. As there is some lag between the peers' and the borrower's decision even time-varying common shocks threaten our results only if they are autocorrelated. As we cannot rule out the possibility of autocorrelated common shocks we carry out a placebo test to show suggestive evidence that such shocks do not bias our results.

To conduct this exercise we replace the peers' average mathematics score with the average of the peers who took out a loan after the borrower. Nevertheless, there is one more challenge in this setup. Namely, if there is any peer effect, then the borrower herself may have an impact on colleagues who take out a loan in the following period. In this case, the placebo test may result in significant coefficient not because of the presence of common shocks, but simply due to the peer effect that we are after. To overcome this issue, we restrict our analysis to low skilled borrowers. The reason we use low-skilled borrowers is that, according to our result in Section 4.2, low-skilled borrowers have a limited impact on their peers because peer impact increase in conjunction with mathematical skill. Therefore, a low mathematics score points to a low peer effect (as in Table 3).

The results of the placebo test are shown in Table 4. Here, we restrict our sample to those whose mathematics score is below the median. According to our results, those who have low mathematical skill have a low impact on the peer's interest rate. In this case, the parameter of the peer effect is -0.458 in the original setup.[15] It is significant at all conventional levels. In the placebo case the parameter is -0.154 and it is not significant. These estimations suggests that common shocks do not distort our results.

Table 4
Placebo test

	(1) Before$<$50%	(2) After$<$50%
maths avg	-0.458***	
	(0.147)	
maths avg after		-0.154
		(0.146)
maths	-0.349**	-0.362**
	(0.155)	(0.155)
Observations	33462	33462
R^2	0.168	0.168
Covariates	Yes	Yes
Time FE	No	No
Company FE	Yes	Yes
Industry x Time FE	Yes	Yes

Only those borrowers' interest rate is included in the estimation whose own math score is below the median. Dependent variable: mortgage loan interest rate. Main independent variables: maths avg col: weighted average math score of colleagues, maths own: own maths score, maths avg col after: weighted average math score of colleagues who take a mortgage after. Other independent variables: age, age^2, own wage, peers' weithted average wage, log(sales). Standard errors clustered at company level. *** p$<$0.01, ** p$<$0.05, * p$<$0.1

[15] This result is somewhat different from the one in Table 2 Specification (3) due to the fact the sample is different (narrower). We include only those observations that both have peers in the before and after period and we also restrict for low skilled borrowers.

4.4 ALTERNATIVE OUTCOMES

As we discussed earlier, a mortgage loan is a complex financial product and it is not only the interest rate that defines its features. We test alternative outcomes (Table 5) to determine whether there are peer effects in respect of such features. First we examine the repricing period.[16], i.e. the period for which the interest rate of a loan is fixed. For instance, if the repricing period is 10 years, this means that the interest rate is fixed for the first 10 years and is then subsequently recalculated based on the current market interest rate. After the new interest rate is calculated it remains fixed for the next 10-year period. We see a positive effect for the repricing period, which means that ceteris paribus if the mathematics score of a borrower's peers is higher then the repricing period of the borrower is shorter. An increase in peers' mathematics score from the 10^{th} to the 90^{th} percentile implies a decrease in the repricing period by around 1.5%, which is a rather small impact.

Maturity is the period of time during which the consumer should repay the loan. We cannot measure any peer effects in respect of this feature.

The principal is the amount of money, which the consumer borrowed from the bank. According to our results changing the mathematics score of peers from the 10^{th} to the 90^{th} percentile implies an increase in principal of roughly 12% (in the case of low-skilled borrowers). These results may indicate some increase in risk-taking behaviour if peers' mathematics score increases.

If we include the latter three dependent variables into the original estimation as controls we obtain similar results. This result suggests that omitting these variables from the original regression does not distort our results, and we do not face the problem that we leave out key variables that have a strong impact on the pricing of loans. We refer to these estimation as being exposed to the so-called 'bad control' problem, because these control variables may be treated as outcome variables as well (all of these variables together with the interest rate are determined simultaneously). The inclusion of these variables does not affect the main result. If we add the bank dummies the parameter estimates decreases by one third. This suggests that one part of the advice is probably about which bank to choose. The estimated parameters are very similar to what we obtained in the original estimation in Subsection 4.2.

4.5 SIMILAR BANK CHOICES

In this section we present additional suggestive evidence that one channel of advice is the choice of bank. We test if borrowers tend to choose the same bank as peers have chosen. We model the borrowers' choice of bank in a random utility model framework estimating a conditional logit model.

Our key independent variable is the weighted number of peers who have chosen the given bank in the borrower's network[17]. Our previous results showed that peers whith mathematics scores have an impact on borrowers, therefore, when we calculate the share of peers choosing a given bank we include only those peers whose mathematics score is higher than the median mathematics score. We also include the share of the bank chosen in the borrower's company (again, including only peers with above median mathematics scores) over the entire time frame (time-invariant variable). This variable captures shocks related to the company for instance if a bank has a branch close to the company and therefore most employees use that bank. Finally, we add bank fixed effects for all alternatives[18] which pins down how often a specific bank is chosen in general during the observed period[19].

The main concern with this approach could be that, although we control for company shares and bank fixed effects, time-varying local shocks may distort our results. To overcome this problem we once again run a placebo test where we replace the peers who have already taken out a loan to with the peers who take out a loan later than the borrower (from 3 to 12 months later), including only peers with higher-than-median mathematics scores. The peers who take out their loan later are supposed to

[16] The repricing period is calculated as fraction of the time the interest rate is fixed to the total maturity of the loan. For instance, it is 1 if the interest rate is fixed for the whole period. It is 0.5 if the interest rate is fixed for the first half of the loan.

[17] We restrict the choice set of borrowers to those banks that are active in the district of the borrower in the given year. We treat a bank as active if the share of loans granted is at least 5 per cent of all the loans granted in the district.

[18] We interact bank fixed effects with fixed effects of alternatives

[19] Bank fixed effects capture all bank related, time-invariant characteristics that are not observed.

Table 5
Alternative outcomes

	(1)	(2)	(3)	(4)	(5)
	Repricing	Maturity	Principal	Bad controls 1	Bad controls 2
	Repricing period	Maturity	Ln(Principal)	Interest rate	Interest rate
maths avg col	-0.101**	-1.110	0.668***	-0.605***	-0.395**
	(0.0510)	(0.981)	(0.125)	(0.191)	(0.178)
maths avg col	0.194*	1.811	-1.006***	1.134***	0.797**
x maths own	(0.113)	(2.260)	(0.299)	(0.414)	(0.379)
repricing period				1.446***	1.447***
				(0.0243)	(0.0255)
maturity				0.0461***	0.0381***
				(0.00113)	(0.00110)
ln(principal)				-0.466***	-0.291***
				(0.00875)	(0.00860)
maths own	-0.0517	-2.669***	0.954***	-1.025***	-0.690***
	(0.0487)	(0.969)	(0.121)	(0.181)	(0.165)
Observations	73016	73016	73016	73016	73016
R^2	0.124	0.156	0.285	0.330	0.426
Covariates	Yes	Yes	Yes	Yes	Yes
Time FE	No	No	No	No	No
Company FE	Yes	Yes	Yes	Yes	Yes
Ind. x Time FE	Yes	Yes	Yes	Yes	Yes
Bank FE	No	No	No	No	Yes

Dependent variables: (1) repricing period, (2) maturity, (3) ln(principal), (4)-(5) mortgage interest rate. Independent variables: maths avg col: weighted average math score of colleagues, maths own: own maths score. Other independent variables: age, age^2, own wage, peers' weithted average wage, log(sales). Standard errors clustered at company level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

have no impact on the borrowers if the mechanism is indeed informal advice. To put it more precisely, if there are time-varying common shocks, then we expect the placebo test to be significant. The insignificance of the placebo test would suggest that there are no time-varying common shocks distorting our results.

Table 6
Conditional logit model on bank choice

	(1)	(2)	(3)
	Baseline	Before<50%	After<50%
Weighted avg share of bank in network	0.129***	0.158***	
	(0.0363)	(0.0369)	
Weighted avg share of bank in network, after			-0.00685
			(0.0383)
Share in company	5.317***	4.871***	5.020***
	(0.213)	(0.205)	(0.211)
Observations	469804	212409	212409

Column (1) contains the full sample. Column (2)-(3) contains those observations, where the own math score is below the median. One observation is one potential choice of the borrower. Dependent variable: choice of bank – indicator variable. Main independent variables: Weighted avg share of bank in network: weighted number of peers who have chosen the given bank divided by weighted number of all peers. Bank chosen in network (weighted) after: weighted number of peers who have chosen the given bank during 3 to 12 month after the borrower has taken her loan divided by weighted number of all peers in the given period. Share in company: share of colleagues who have chosen the same bank (time-invariant). Share in total sample (not shown): share of borrowers who have chosen the same bank (time-invariant). Standard errors clustered at company level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

As summarised in Table 6, the results show a significant impact of peers on the borrower's choice as the variable *Weighted avg share of bank in network* is strongly significant. The economic relevance of this impact can be shown by an example: comparing the chance that a borrower chooses bank A if 0 per cent of his peers has chosen bank A versus 100 per cent having chosen the bank. In the case of that bank A has a 10 per cent share in the sample there is 2.3 per cent higher chance that the borrower chooses this bank[20]. That is equivalent to a 23 per cent impact compared to the original chance of choosing the bank. Although this figure varies with the share of the banks ($p_{i,j}$) the impact of this variation is limited regarding our analysis: taking into account the shares of the banks in our sample the weighted average impact of the change outlined above in the share of peers choosing a certain bank (and not choosing others) is 20 per cent.

The results suggest that part of the informal advice is the choice of bank. Our key variable is significant even after controlling for the share of borrowers choosing a certain bank at the same company. The placebo test, where only those are included in the sample whose mathematics score is below the median and are therefore unlikely to give advice to borrowers (similar as in the main specification) shows no significant results, suggesting that no time-varying common shocks are distorting the results.

[20] Estimated impact can be calculated the following way. The own effect of choosing the bank will increase the chance that borrower i chooses bank A by $p_{i,j}(1-p_{i,j})0.129 = 0.1 \cdot 0.9 \cdot 0.129 = 0.0116$. The cross-effects also increase the chance of choosing bank A by $\sum_k (p_{i,j} p_{i,k} 0.129) = \sum_k (0.1 p_{i,k} 0.129) = 0.0116$, as the share of choosing other banks decreases in the network of the borrower. The total impact is the sum of own and all cross-effects.

4.6 CERTIFIED CONSUMER FRIENDLY HOUSING LOAN

The reason for the presence of peer effect is likely to be an information problem of the borrower. There are at least three potential problems that makes the decision of the borrower difficult. First, mortgage loan is a complex product as when a borrower decides about a mortgage loan should decide about several things in the same time. These are for instance the loan amount, maturity, acceptable interest rate, repricing period, which bank to choose etc. Second, it is difficult to collect information about all the loans that are available to a borrower. It is not enough to go to all the banks but should collect all offers of the bank as banks have potentially multiple offers. Bank offer comparison websites may help borrowers, however those not necessarily contain all offers, and all details of the offers that are relevant. Finally, even having all the information on the offer level it is difficult to digest all that information and make a good comparison of the offers. Moreover, there is an additional layer of these problems as borrowers may not trust in the information provided by banks.

Based on our results so far we argue that a colleague may help in all of these problems. A skilled peer can share own experience and explain the important aspects of the loan based on own experience. Then need for more information and for comparison of offers may substantially decreases. Moreover, borrowers can also trust more their peers.

What would be then the policy implication of our results? The simplest implication would be to place peers who has taken a loan recently and has high mathematical skill to all borrowers. This is obviously not a feasible solution. The closest feasible solution to this would be to run a free advisory that supports borrowers in this complex problem. One drawback of that solution is that running such an advisory on the country level may be expensive.

Controlling the offers of banks by regulation can also support borrowers. Nevertheless, strong control may also decrease the options that borrowers have and may have negative effects as well. Forcing banks to offer standardised loans (among other loans as well) that are easy to compare and borrowers can trust them would solve most of the information problems of borrowers. Certified Consumer Friendly Housing Loan that was introduced in Hungary at the end of 2017 was such a product.

Implementation of new products needs time. Some banks with efficient product development could introduce CCFL quickly, while other banks were slower with the introduction. At the end of 2017 around 10-15 per cent of all mortgage loans were CCFL and almost one third of fixed repricing period loans. From 2020 two out of three newly disbursed mortgage loans were CCFL (see Figure 3), and this share stabilised.

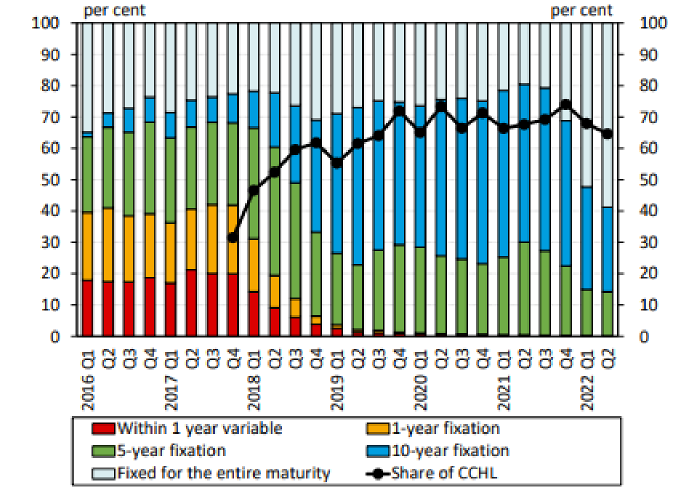
To understand how the introduction of CCFL affected peer effect we use the spread of CCFL that is the period from the second half of 2017 to end 2019 for analysis. If the introduction of CCFL helps to decrease the information problem, then impact of peer effect should decrease after the introduction. Nevertheless, it is also possible that the peer effect remains relevant but narrows down to suggesting choosing a CCFL product. That would be a sign that product standardisation cannot solve the information problem, as remaining presence of peer effect would indicate remaining presence of the information problem.

First we estimate the effect of peers in choosing CCFL product after the introduction of it. We run a similar regression as in Table 3, however we model whether the borrower has chosen CCFL by changing the dependent variable to an indicator variable that takes the value of one if a borrower has chosen CCFL product. Moreover, we analyse the heterogeneity of the peer effect with respect to the spread of CCFL. We do so by interacting the average mathematics score of peers by CCFL share among newly disbursed loans in the district in the given month. We also include the share of CCFL as a control variable, it varies between zero and one. We argue that the spread of this product was partially random, as it depended on the product development of the bank - how quickly it could introduce a product that meets the CCFL criteria. We keep including company fixed effect and industry $\times$ time fixed effect in the regressions. Moreover, in this regression we also include the bank concentration measure (HHI) to control for the change in bank competition in different districts.

The results of the estimation are in Table 7 Specification (1). Average mathematics score of peers is large and significant suggesting that peers advise choosing CCFL. However, the interaction term of peers' mathematics score and CCFL share in the district in the given period is negative and (depending on the share of CCFL) it may fully offset the peer effect. This result supports that the introduction and spread of CCFL decreases peer effect. It is suggestive evidence that it also decreases the information problem.

We run a similar regression, however now we use interest rate as dependent variable. Explanatory variables are the same as previously. Results that are in the second column of Table 7 are also similar. peer effect decreases the interest rate of borrowers,

Figure 3
Distribution of new housing loan volume by interest rate fixation, and the share of Certified Consumer Friendly Housing Loan products



Note: Share of CCHL products compared to new issues with at least 3 years of interest rate fixation (at least 5 years since Q4 2018) excluding disbursements by building societies and without FGS GHP. Source: MNB

Source: Trends in Lending 2022 September, https://www.mnb.hu/letoltes/hitelezesi-folyamatok-2022-szeptember-en.pdf

however, if CCFL share is large, then peer effect vanishes. Again, we conclude that this result is suggestive evidence that the introduction of a standardised product can decrease the information problem.

We think that the spread of CCFL is driven by such effects that are related to peer effect, i.e. our results are not biased. Although, endogeneity concerns regarding the spread of such a product could arise, we include a wide set of control variables (company fixed effect, industry $\times$ time fixed effect, demographic controls, bank concentration) that captures most of the shocks that we could imagine would bias our estimate. Although we think that it is difficult to find a plausible endogeneity story, in theory any company or district level time-varying common shock can distort the results. As a robustness check we implement an IV strategy.

In our IV estimations the endogeneous variable is the spread of CCFL. We use the spread of CCFL in the sixth month of introduction as an instrument[21]. To also account for the fact that CCFL share is increasing through time we interact CCFL share in the sixth months after introduction with a time trend[22]. The exogenous variation comes from the noise around the introduction of CCFL. As we mentioned earlier, depending on banks how quickly could develop the new product they granted either in large or in small scale after the first half year of the introduction of CCFL. Whether those banks are present in a district that were successful in implementing the CCFL or not is not related to the information problem that we analyse. Moreover, using our large set of controls in the first stage we cannot find any plausible endogeneity issue that would question the validity of the instrument.

The results of the IV estimation are similar to the OLS results (see in Table 8). Although point estimates of both peer effects and the interacted terms are somewhat larger in absolute terms in the IV estimation, they are not significantly different from the OLS results and they also tend to offset each other if CCFL share is large. This robustness check strengthens the results of the OLS estimates.

[21] The rationale for using the sixth month is that we see that in the first few months the product was not present in most districts, therefore our IV would not be strong enough. We also do not want to use much later periods as that is not consistent with the exogenous shock that roots in the variation of time needed for banks to implement the new product. Nevertheless, using seventh or eights month instead of the sixth month does not change our results significantly.

[22] It is important to do so as company fixed effects would fully absorb CCFL share if that is not changing through time.

Table 7
Spread of CCFL

	(1)	(2)
	Choice	Interest
maths avg col	0.475***	-0.998***
	(0.0848)	(0.245)
maths avg col × CCFL share	-1.044***	1.217**
	(0.183)	(0.513)
CCFL share	0.963***	-0.960***
	(0.0767)	(0.214)
bank concentration	0.0139	1.615***
	(0.0399)	(0.126)
maths own	0.144***	-0.909***
	(0.0239)	(0.0639)
Observations	30022	30022
R^2	0.477	0.313
Covariates	Yes	Yes
Company FE	Yes	Yes
Industry x Time FE	Yes	Yes

In column (1) the dependent variable is a dummy variable indicating whether the borrower choose the CCFL or not. In column (2) the dependent variable is the mortgage interest rate. Main independent variables: maths avg col: weighted average math score of colleagues, CCFL share: the share of CCFL in a given district at a given month, bank concentration: measured as a Herfindalh-Hirschman-index, maths own: own maths score. Other independent variables: age, age^2, own wage, peers' weithted average wage, log(sales). Standard errors clustered at company level. *** p$<$0.01, ** p$<$0.05, * p$<$0.1

Table 8
Spread of CCFL - IV estimation

	(1)	(2)
	Choice	Interest
maths avg col	0.650***	-1.376***
	(0.153)	(0.385)
mathematics avg col x CCFL share	-1.485***	2.136**
	(0.386)	(0.972)
CCFL share	0.787***	-1.370***
	(0.201)	(0.523)
bank concentration	-0.138**	1.621***
	(0.0612)	(0.168)
maths own	0.149***	-0.906***
	(0.0237)	(0.0645)
Observations	30022	30022
R^2	0.039	0.064
Covariates	Yes	Yes
Company FE	Yes	Yes
Industry x Time FE	Yes	Yes

In column (1) the dependent variable is a dummy variable indicating whether the borrower choose the CCFL or not. In column (2) the dependent variable is the mortgage interest rate. Main independent variables: maths avg col: weighted average math score of colleagues, CCFL share: the share of CCFL in a given district at a given month, bank concentration: measured as a Herfindalh-Hirschman-index maths own, own maths score. Other independent variables: age, age^2, own wage, peers' weithted average wage, log(sales). Standard errors clustered at company level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

5 Alternative specification: using only small firms

In the weighted network approach, which is described in Section 3, those firms are not included in the specification, where there are not any loan taker peers. Employees at these firms still can be used as controls. To overcome this caveat, we implement a different specification.

5.1 BASIC SPECIFICATION

In this setup we focus only on small firms, with employees less than 50. In this way we can suppose that every employee know each other personally. We change the parameter of interest because in case of small firms there are only few employees who took a credit that is why the calculation of weighted average is not meaningful here. We define a dummy variable for those who have at least one colleague, whose mathematical skill is above the 90th percentile and took a credit in the past one year. We define similar dummies for loan-taker colleagues with medium mathematical knowledge (between 50th and 90th percentile) and with low mathematical knowledge (below median). We estimate the following equation:

$$y_{i,c,t} = \alpha + \beta_1 M_1 + \beta_2 M_2 + \beta_3 C_{Hi,c,t} + \beta_4 C_{Ni,c,t} + \beta_5 C_{Mi,c,t} + \gamma X_{i,c,t} + \varepsilon_{i,c,t} \tag{8}$$

- i - individual, c - company, t - period,
- y - the interest of the loan
- M_1 own math skill: 1 if math skill is above 90th percentile, 0 otherwise.
- M_2 own math skill: 1 if math skill is between 50th and 90th, 0 otherwise.
- C_{Hit} - financially highly literate colleague dummy (1 if have a colleague, with high math skills, who took a credit in the past year)
- C_{Nit} - financially non-literate colleague dummy (1 if have a colleague, with low math skills, who took a credit in the past year)
- C_{Mit} - financially moderately literate colleague dummy (1 if a colleague took a credit in the past year and his math score is between the 50th and 90th percentile)
- X_{it} - other variables (wage, age, age^2, district, time and sector FE)

In this setup the parameter of interest is the coefficient of the financially literate colleague dummy (C_{Hit}). We expect this coefficient to be significant and negative. The financially non-literate colleague dummy (C_{Nit}) should be non-significant because we think that the mathematical skills and the mortgage experience together makes someone proficient enough to give a good advice. We expect the magnitude of the financially moderately literate colleague dummy (C_{Mit}) to be between 0 and the parameter of C_{Hit} variable.

The main specification (Table 9, Column (1)) shows that if someone has high mathematical skills the interest rate is lower by 0.27 percentage point. If someone has a colleague, who is good at math and have recent experience in mortgage loan taking, than his interest rate is 0.21 percentage point lower. This coefficient is somewhat lower than the own mathematical skill, which is intuitive. The magnitude of this peer effect is similar to what we estimated in the original specification (Section 4).

For the small firm specification we also run a similar placebo test as in Section 3. In case of the placebo we define those financially literate, who took a mortgage *after* their colleague had taken. If this variable is significant that would mean that the

Table 9
Estimations using only small firms

VARIABLES	(1) Main interest	(2) Placebo interest
Highfinlit colleague	-0.214***	-0.00107
	(0.0805)	(0.151)
Lowfinlit colleague	0.0412	-0.00815
	(0.0356)	(0.0438)
Midfinlit colleague	-0.0513	-0.0363
	(0.0369)	(0.0444)
Own math:High	-0.273***	-0.299***
	(0.0473)	(0.0562)
Own math:Mid	-0.177***	-0.184***
	(0.0235)	(0.0259)
Observations	16,007	13,415
R-squared	0.153	0.148
Industry × Time FE	Yes	Yes
District FE	Yes	Yes
Mean of dep. var.	4.817	4.817

The sample contains only those firms, where the number of employees is less than 50. Dependent variable: mortgage loan interest rate. Finlit colleague - shows if there is a colleague with mathematics score above the 90^{th} percentile who took a mortgage in the past year (in case of placebo: in the *following* year). Nonfinlit (Midfinlit) colleague - same as Finlit colleague, before but mathematics score is below the 50^{th} (between the 50^{th} and 90^{th}) percentile. Other independent variables: age, age^2, log(sales). Standard errors clustered at company level. *** p$<$0.01, ** p$<$0.05, * p$<$0.1

peer effect, which we estimate is some kind of firm specific effect, which is not related to the mechanism which we described. This firm specific effect can be, for instance, if the employees of a firm get some special offer from a certain bank. In the placebo estimation (see Column (2) in Table 9) the parameter of interest is not significant.

We make other robustness checks as well. First, we change the threshold for small firms. We find that by lowering the number of employees the peer effect increases. This probably shows that at smaller firms the employees know and trust each other better so their advise is more valuable. By increasing the firm size the parameter of interest decreases gradually. At a larger firm the chance of knowing a financially literate is smaller because by increasing the firm size the number of financially literate colleagues does not increase that much. We also make robustness check with the bank competition. We find that in those districts, where the number of banks is lower than the country median, the peer effect is larger. This is a similar result as in the main specification in Section 3. If we include bank fixed effects the parameter of peer effect decreases by one third. This suggests that one channel through which co-workers give advise is which bank to choose. All tables for these regression are in the Appendix D.

5.2 ROBUSTNESS CHECK: COARSENED EXACT MATCHING

The caveat of the small firm specification is that we cannot include firm fixed effects. If we include these the most of the identifying variation would be absorbed. Number of observations decrease a lot because the number of loan-takers is limited at a small firm. That is why it seems reasonable to estimate a coarsened exact matching on the small firm sample. In this setup we would like to compare an employee who has a high mathematical skilled colleague, who took a mortgage in the past one year and another, who has a same colleague but did not take a credit. For this we define a new variable, the number of employees at a given firm, whose math skill is above the 90^{th} and younger than 42 years. The second condition comes from the fact that younger people are more likely to take a mortgage because they are at the beginning of family foundation and they are more in need to have a bigger house or a flat. This number will show how many colleagues are at a firm, who are perspective client for a mortgage loan. We control for the employee's mathematical skill, the wage, the age, the sector of the firm, the total sales of the firm and the number of employees whose math skill is above the 90^{th} and younger than 42 years. Comparing based on these variables we can assume that taking a mortgage is random.

The coarsened exact matching (CEM) algorithm creates bins from the covariates in a way that there is at least one treated and one control observation in every bin. If there are observations for which there are not any pairs it is going to be dropped (see details for (Blackwell, Iacus, King, & Porro, 2009)).

Based on (Blackwell et al., 2009) we first run a simple OLS regressing the interest rate on our treatment variable, which is a dummy, showing whether there is a financially literate peer or not. The parameter of interest is -0.409, which means that for those who has a financially literate peer the interest rate is 0.4 percentage point lower (Table 10). Of course this result is biased but it shows a simple difference in means.

In the next step we run the CEM algorithm based on the following variables: the employee's mathematical skill, the wage, the age, the sector of the firm, the total sales of the firm and the number of employees whose math skill is above the 90^{th} and younger than 42 years. Using the weights coming from this estimation we run a weighted OLS. Based on the results the average treatment effect is -0.173. The number of observation is roughly halved. This shows that coarsening procedure did not find exact match for some observations on the coarsened variables.

As another robustness check we run a simple propensity score matching with the same covariates. In Table 11 Column (1) shows the result for propensity score matching. The estimated parameter is very similar to the previous results. We also estimate the propensity score for only those observations to which the CEM method find at least one treated and a control observation (Column (2)). The estimated parameter is similar to the previous one and to what we get in Table 9.

Table 10
OLS and coarsened exact matching

	(1)	(2)
	interest	interest
Highfinlit colleague	-0.409***	-0.173***
	(0.0311)	(0.0386)
Constant	4.957***	4.719***
Observations	82,693	48,832
R-squared	0.002	0.001
Type	OLS	CEM
Other variables	No	No

The table contains all mortgage loan takers, who work at a firm, where the number of employees is less than 50. The Highfinlit colleague dummy shows is there a colleague, for whom the mathematical score is above the 90^{th} percentile and took a mortgage at most 1 year before. The dependent variable is the mortgage loan interest rate. Column (2) is a weighted regression, where the weights come from a coarsened exact matching procedure on covariates (see text). Standard errors are clustered at company level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 11
Simple propensity score matching and with coarsened exact matching

	(1)	(2)
	interest	interest
ATE	-0.221**	-0.187***
	(0.104)	(0.0577)
Observations	60,034	32,739
Type	PS match	PS match with CEM

The table contains all mortgage loan takers, who work at a firm, where the number of employees is less than 50. The dependent variable is the mortgage loan interest rate. Column (1) is a propensity score matching based on the following variables: the employee's mathematical skill, the wage, the age, the sector of the firm, the total sales of the firm and the number of employees whose math skill is above the 90^{th} and younger than 42 years. Column (2) differs from Column (1) that only those observations are included, where the coarsened exact matching procedure find pairs. Robust standard errors are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

5.3 PEER EFFECT ALONG THE EXTENSIVE MARGIN OF LENDING

It is a natural question whether peers affect the probability of taking a mortgage loan. To examine this question we take all of the employees, who are employed at firms with less than 50 employees. We mark those, who has a mortgage loan. We also mark those who has a mortgage loan and has mathematical skills above the 90th percentile. We estimate the following regression:

$$l_{i,c,t} = \alpha + \beta_1 L_{i,c,t} + \gamma X_{i,c,t} + \varepsilon_{i,c,t} \tag{9}$$

- i - individual, c - company, t - period,
- $l_{i,c,t}$ - 1 if the individual i has a mortgage loan in period t at company c, 0 otherwise,
- $L_{i,c,t}$ 1 if the individual has a colleague, who took a credit in the past year, 0 otherwise,
- X_{it} - other variables (wage, age, age^2, number of colleagues, log sales of the company, district, time and sector FE).

In Equation 9 we look for whether a colleague has an effect on the probability of taking a mortgage loan. We also check whether a high mathematical skill colleague has an effect on loan taking decisions. To implement this we use another dummy variable, which is 1 if the individual has a colleague with high math skills, who took a credit in the past year, 0 otherwise.

Our results show that colleagues' mortgage loan decision does not have an effect on the probability of taking a mortgage loan. This is true with OLS and with logit estimation with and without controls. In some cases the coefficient is significant but it is less than 0.1% in absolute value. In case of a financially literate colleague the coefficient is somewhat larger. In all type of estimations the coefficients are economically not significant.

Table 12
Extensive margin of lending, OLS estimates

	(1)	(2)	(3)	(4)
VARIABLES	Choice	Choice	Choice	Choice
Colleague	0.000402***	0.000276		
	(0.000136)	(0.000170)		
Highfinlit colleague			-0.000938*	-0.00149**
			(0.000570)	(0.000641)
Observations	6,345,370	3,702,795	6,345,370	3,702,795
R-squared	0.0000	0.0000	0.0000	0.0000
Mean of dept. var.	0.006	0.006	0.006	0.006
Time FE	No	Yes	No	Yes
District FE	No	Yes	No	Yes
Industry FE	No	Yes	No	Yes
Controls	No	Yes	No	Yes

The table contains all employees at firms with less than 50 employees. The dependent variable is 1 if the individual took a mortgage loan in the given year, 0 otherwise. The Colleague dummy shows is there a colleague who took a mortgage at most 1 year before. The Highfinlit colleague dummy shows is there a colleague, for whom the mathematical score is above the 90th percentile and took a mortgage at most 1 year before. Controls are number of colleagues, wage, age, age^2, log(sales) of the company, year, district and industry FE's. Standard errors are clustered at company level. *** $p<0.01$, **

Table 13
Extensive margin of lending, logit estimates (margins)

	(1)	(2)	(3)	(4)
VARIABLES	Choice	Choice	Choice	Choice
Colleague	0.00039***	0.000276*		
	(0.00012)	(0.000170)		
Highfinlit colleague			-0.00101	-0.0017**
			(0.000668)	(0.000843)
Observations	6,345,370	3,702,795	6,345,370	3,702,795
R-squared	0.0000	0.0000	0.0000	0.0000
Mean of dept. var.	0.006	0.006	0.006	0.006
Time FE	No	Yes	No	Yes
District FE	No	Yes	No	Yes
Industry FE	No	Yes	No	Yes
Controls	No	Yes	No	Yes

The table contains all employees at firms with less than 50 employees. The dependent variable is 1 if the individual took a mortgage loan in the given year, 0 otherwise. The Colleague dummy shows is there a colleague who took a mortgage at most 1 year before. The Highfinlit colleague dummy shows is there a colleague, for whom the mathematical score is above the 90th percentile and took a mortgage at most 1 year before. Controls are number of colleagues, wage, age, age^2, log(sales) of the company, year, district and industry FE's. Standard errors are clustered at company level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

6 Conclusion

In our paper we used the whole universe of mortgage loans granted between 2015-2019 in Hungary for identifying peer effect of co-workers. As taking a loan is a complex task borrowers often make mistakes. We analyzed two salient strategies to cope with the difficulty of taking a mortgage loan: informal advice and regulation. In our special setup we could also account to the interaction of these two factors.

First, we tested whether informal advice play an important role in the borrowing decision. We find that if a mortgage loan-taker has low mathematical skill then if he has peers, who have high mathematical skill and experience in the mortgage loan market, the interest rate of the borrower tends to be smaller compared to similar borrowers with no smart and experienced peers. We showed that there is heterogeneity in peer effects depending on the strength of competition among banks. If the magnitude of the competition is lower than the peer effect is considerably higher. According to our results, considerable part of the peer effect happens through the choice of the bank.

Second, we find that the introduction of a standardised loan can ease the information problem of the borrower. After the introduction of Certified Consumer Friendly Housing Loans the relevance of informal advice decreased. We show evidence that in the first part of the program peers helped borrowers to find the CCFL products, while later when CCFL was widespread peers become less important in the borrowing decision.

Lastly, we also experiment with a special sample - small firms - as in this sample we do not have to make assumptions on the network structure of colleagues. Our results remain qualitatively the same.

The policy implication of our results connects to the social capital literature (see e.g. (Chetty et al., 2022)). We see that banks offer different rates not just based on the risk of individuals (which we measure by several control variables) but also on skills and experience. This can happen due to information problems, namely some people are unable to make comparison between different offers or even not aware of the existence of other offers. There are several ways how to solve this information asymmetry. First, policy makers can introduce general education programmes to increase the average financial literacy of the society. This can be a long and expensive process. Based on our results peers can help to find better bank offers. We argue that peers can considerably decrease the complexity of the problem and serve also as a trustful source of information. Consistent with this narrative, we also found that introducing a standardised product can decrease peer effect. The standardised product takes the place of peers in the information problem. It improves the utility of those borrowers who do not have peers that could help them. The introduction of standardised products can be fast and cheap.

References

Agarwal, S., Chomsisengphet, S., Liu, C., & Souleles, N. S. (2015). Do Consumers Choose the Right Credit Contracts? *Review of Corporate Finance Studies*, *4*(2), 239-257.

Agarwal, S., Chomsisengphet, S., Mahoney, N., & Stroebel, J. (2015). Regulating consumer financial products: Evidence from credit cards. *The Quarterly Journal of Economics*, *130*(1), 111-164.

Agarwal, S., & Mazumder, B. (2013). Cognitive abilities and household financial decision making. *American Economic Journal: Applied Economics*, *5*(1), 193-207.

Ambuehl, S., Bernheim, B. D., Ersoy, F., & Harris, D. (2022, 11). Peer Advice on Financial Decisions: A Case of the Blind Leading the Blind? *The Review of Economics and Statistics*, 1-45. doi: 10.1162/rest_a_01269

Barr, M. S., Mullainathan, S., & Shafir, E. (2008). Behaviorally informed financial services regulation. *Working paper*.

Bhattacharya, U., Kumar, A., Visaria, S., & Zhao, J. (2024). Do women receive worse financial advice? *The Journal of Finance*, *79*(5), 3261-3307. doi: https://doi.org/10.1111/jofi.13366

Blackwell, M., Iacus, S., King, G., & Porro, G. (2009, December). cem: Coarsened exact matching in Stata. *Stata Journal*, *9*(4), 524-546.

Cameron, A., & Trivedi, P. (2010). *Microeconometrics using stata, revised edition*. StataCorp LP.

Campbell, J. (2016). Restoring rational choice: The challenge of consumer financial regulation. *American Economic Review*, *106*(5), 1-30.

Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., … Wernerfelt, N. (2022). Social capital I: measurement and associations with economic mobility. *Nature*, *608*, 108-121. doi: 10.1038/s41586-022-04996-4

Cipriani, M., & Guarino, A. (2014, January). Estimating a structural model of herd behavior in financial markets. *American Economic Review*, *104*(1), 224-51. doi: 10.1257/aer.104.1.224

Currarini, S., Jackson, M. O., & Pin, P. (2009). An economic model of friendship: Homophily, minorities, and segregation. *Econometrica*, *77*(4), 1003-1045. doi: https://doi.org/10.3982/ECTA7528

Dahl, G. B., Loken, K. V., & Mogstad, M. (2014, July). Peer Effects in Program Participation. *American Economic Review*, *104*(7), 2049-2074.

Duflo, E., & Saez, E. (2002). Participation and investment decisions in a retirement plan: the influence of colleagues' choices. *Journal of Public Economics*, *85*(1), 121-148.

Fan, L., & Chatterjee, S. (2017). Borrowing decision of households: An examination of the information search process. *Journal of Financial Counseling and Planning*, *28*(1), 95-106. doi: 10.1891/1052-3073.28.1.95

Foerster, S., Linnainmaa, J. T., Melzer, B. T., & Previtero, A. (2017). Retail financial advice: Does one size fit all? *The Journal of Finance*, *72*(4), 1441-1482. doi: https://doi.org/10.1111/jofi.12514

Folke, T., Gjorgjiovska, J., Paul, A., Jakob, L., & Ruggeri, K. (2021). Asset: A new measure of economic and financial literacy. *European Journal of Psychological Assessment*, *37*(1), 65-80. doi: 10.1027/1015-5759/a000575

Frijns, B., Gilbert, A., & Tourani-Rad, A. (2014, April). Learning by doing: the role of financial experience in financial literacy. *Journal of Public Policy*, *34*(1), 123-154.

Gerardi, K., Goette, L., & Meier, S. (2010). *Financial literacy and subprime mortgage delinquency: evidence from a survey matched to administrative data* (FRB Atlanta Working Paper No. 2010-10). Federal Reserve Bank of Atlanta.

Giorgi, G. D., Frederiksen, A., & Pistaferri, L. (2020). Consumption Network Effects. *Review of Economic Studies*, *87*(1), 130-163.

Heidhues, P., Johnen, J., & Kőszegi, B. (2021). Browsing versus studying: A pro-market case for regulation. *Review of Economic Studies*, *88*(2), 708-729.

Inderst, R., & Ottaviani, M. (2012). Financial advice. *Journal of Economic Literature*, *50*(2), 494-512.

Lieber, E. M., & Skimmyhorn, W. (2018). Peer effects in financial decision-making. *Journal of Public Economics*, *163*(C), 37-59. doi: 10.1016/j.jpubeco.2018.05

Lusardi, A., & Mitchell, O. S. (2007, January). Baby Boomer retirement security: The roles of planning, financial literacy, and housing wealth. *Journal of Monetary Economics*, *54*(1), 205-224.

Lusardi, A., & Mitchell, O. S. (2014, March). The economic importance of financial literacy: Theory and evidence. *Journal of Economic Literature*, *52*(1), 5-44. doi: 10.1257/jel.52.1.5

Mandell, L. (2006, April). *Financial Literacy: If It's So Important, Why Isn't It Improving?* (NFI Policy Briefs No. 2006-PB-08). Indiana State University, Scott College of Business, Networks Financial Institute.

Manski, C. (1993). Identification of endogenous social effects: The reflection problem. *Review of Economic Studies*, *60*(3), 531-542.

Maturana, G., & Nickerson, J. (2018, 12). Teachers Teaching Teachers: The Role of Workplace Peer Effects in Financial Decisions. *The Review of Financial Studies*, *32*(10), 3920-3957. doi: 10.1093/rfs/hhy136

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, *27*, 415-444. doi: https://doi.org/10.3982/ECTA7528

Nicholson, G., Skelton, R., & Tarr, J. (2019). An exploratory study of regulatory failure in the australian home mortgage market. *Journal of Consumer Affairs*, *53*(1), 126-166.

Ouimet, P., & Tate, G. (2020). Learning from coworkers: Peer effects on individual investment decisions. *The Journal of Finance*, *75*(1), 133-172. doi: https://doi.org/10.1111/jofi.12830

Park, A., & Sabourian, H. (2011). Herding and contrarian behavior in financial markets. *Econometrica*, *79*(4), 973-1026. doi: https://doi.org/10.3982/ECTA8602

Spyrou, S. (2013, 11). Herding in financial markets: a review of the literature. *Review of Behavioral Finance*, *5*. doi: 10.1108/RBF-02-2013-0009

Stolper, O., & Walter, A. (2019). Birds of a Feather: The Impact of Homophily on the Propensity to Follow Financial Advice. *Review of Financial Studies*, *32*(2), 524-563.

von Gaudecker, H.-M. (2015). How does household portfolio diversification vary with financial literacy and financial advice? *Journal of Finance*, *70*(2), 489-507.

Weiner, R., Baron-Donovan, C., Gross, K., & Block-Leib, S. (2005). Debtor Education, Financial Literacy and Pending Bankruptcy Education. *Behavioral Sciences and the Law*, *23*, 347-366.

Appendix A Alternative specification: main results using linear weights

In this section we run the same regressions as in Section 4, except for that we change Equation 2 to linear weights as follows:

$$\omega_{i,j} = (d_{i,j} + 1)^{-1} \tag{10}$$

This modification does not change the results much (see Table 14 and 15).

Table 14
Baseline results with linear weights

	(1)	(2)	(3)
	Interest rate	Interest rate	Interest rate
maths avg	-0.631***	-0.178*	-0.195*
	(0.100)	(0.0988)	(0.101)
maths	-0.969***	-0.873***	-0.865***
	(0.0473)	(0.0542)	(0.0550)
Observations	75962	73018	73016
R^2	0.094	0.176	0.184
Covariates	Yes	Yes	Yes
Time FE	Yes	Yes	No
Company FE	No	Yes	Yes
Industry x Time FE	No	No	Yes

Dependent variable: mortgage loan interest rate. Other independent variables: age, age^2, own wage, peers' weithted average wage, log(sales). Standard errors clustered at company level. *** p$<$0.01, ** p$<$0.05, * p$<$0.1

Table 15
Heterogeneity with linear weights

	(1) Mathematics	(2) Competition
maths avg	-1.034***	0.812***
	(0.228)	(0.274)
maths avg x maths	2.061***	
	(0.499)	
maths	-1.724***	-0.856***
	(0.216)	(0.0556)
maths avg x bank concentration		-3.848***
		(0.940)
bank concentration		2.895***
		(0.409)
Observations	73016	67309
R^2	0.184	0.226
Covariates	Yes	Yes
Time FE	No	No
Company FE	Yes	Yes
Industry x Time FE	Yes	Yes

Dependent variable: mortgage loan interest rate. Main independent variables: maths avg col: weighted average math score of colleagues, maths own: own maths score, bank concentration: measured as a Herfindalh-Hirschman-index. Other independent variables: age, age^2, own wage, peers' weithted average wage, log(sales). Standard errors clustered at company level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Appendix B Alternative specification: controlling on regional level bank supply with additional fixed effects

In this section we run the same regressions as in Section 4, except for that we add district level bank fixed effects that are interacted with time fixed effects. By doing so, we control on district level bank supply changes from month-to-month. Although, this specification rules out several endogeneity problems it may also prone to overcontrolling as any advice on the choice of bank may be part of the mechanism that we are after. Any impact through this channel is likely to be absorbed by including bank fixed effects. Therefore, true impact may be larger than the estimations below. We do not estimate again the specification with bank contentration as the additional fixed effect included in these regressions would absorb the variation in bank concentration.

Table 16
Main specifications controlling on regional level bank supply with fixed effects

	(1)	(2)	(3)	(4)
	Basic_spec	Before	After	Interaction
maths avg	-0.221**	-0.395**		-0.800***
	(0.0918)	(0.175)		(0.206)
maths avg x maths				1.418***
				(0.442)
maths avg after			0.0201	
			(0.176)	
Observations	66023	27162	27162	66023
R^2	0.487	0.517	0.516	0.487
Covariates	Yes	Yes	Yes	Yes
Time FE	No	No	No	No
Company FE	Yes	Yes	Yes	Yes
Industry x Time FE	Yes	Yes	Yes	Yes
Bank x District x Time FE	Yes	Yes	Yes	Yes

Dependent variable: mortgage loan interest rate. Other independent variables: age, age^2, own wage, peers' weithted average wage, log(sales). Standard errors clustered at company level. *** p$<$0.01, ** p$<$0.05, * p$<$0.1

Appendix C Description of the choice model

Following (Cameron & Trivedi, 2010), we show the main assumptions and the model that we estimate below. The cornerstone of the model is the assumption that borrowers choose the alternative that gives them the highest utility.

$$U_{i,j} > U_{i,k}, \forall j \neq k \tag{11}$$

where $U_{i,j}$ stands for the utility of borrower i choosing alternative j.

Assuming additive random utility, the utility of borrower i can be written in the form:

$$U_{i,j} = V_{i,j}(x_{i,j}) + \varepsilon_{i,j} \tag{12}$$

where $V_{i,j}()$ is a function of observable characteristics of alternatives and $\varepsilon_{i,j}$ is an idiosyncratic taste shock that is not observable by the researcher and is assumed to follow an iid extreme value distribution.

The probability of choosing bank j related to the observable characteristics can then be expressed as below.

$$p_{i,j} = \frac{e^{V_{i,j}}}{\sum_k e^{V_{i,k}}} = \frac{e^{x_{i,j}'\beta}}{\sum_k e^{x_{i,k}'\beta}} \tag{13}$$

where $x_{i,j}$ contains the following explanatory variables: our key independent variable is what share of the borrower's loan taking colleagues have chosen the given bank. We keep the weighting scheme from the original regressions. We also include what share of all colleagues (no weighting) has chosen a given bank - a variable that is similar to a company fixed effect. Lastly, we include bank fixed effects for all alternatives (interact bank fixed effects with borrowers' options (i.e. alternatives)).

Appendix D Alternative specification: using only small firms

In this section we present the placebo and robustness check results for the small firm specification (Section 5).

We make several robustness checks. First, we change the threshold for small firms. If we increase the number of employees the parameter of interest decreases gradually (Table 17). This is intuitive since at a larger firm the chance of knowing a financially literate is smaller because by increasing the firm size the number of financially literate colleagues does not increase that much (see Table 18).

Table 17
Robustness check: changing the firm size threshold

Nr of colleagues	<60	<70	<80	<90	<100
Highfinlit colleague	-0.124	-0.118*	-0.137**	-0.134**	-0.129**
	(0.0778)	(0.0686)	(0.0642)	(0.0612)	(0.0579)
Nonfinlit colleague	0.0148	0.0151	-0.00302	0.0145	0.0227
	(0.0324)	(0.0308)	(0.0290)	(0.0279)	(0.0270)
Midfinlit colleague	-0.0701**	-0.0707**	-0.0573*	-0.0515*	-0.0542*
	(0.0337)	(0.0318)	(0.0301)	(0.0289)	(0.0280)
Skill:High	-0.287***	-0.311***	-0.318***	-0.323***	-0.323***
	(0.0453)	(0.0434)	(0.0420)	(0.0409)	(0.0401)
Skill:Mid	-0.176***	-0.189***	-0.193***	-0.193***	-0.190***
	(0.0224)	(0.0217)	(0.0213)	(0.0208)	(0.0206)
Observations	17,309	18,441	19,357	20,187	20,848
R-squared	0.152	0.153	0.152	0.150	0.150
Industry × Time FE	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Mean of dep. var.	4.81	4.81	4.80	4.81	4.81

Dependent variable: mortgage loan interest rate. Finlit colleague – shows if there is a colleague with mathematics score above the 90th percentile who took a mortgage in the previous year. Nonfinlit (Midfinlit) colleague – same as Finlit colleague, before but mathematics score is below the 50th (between the 50th and 90th) percentile. Other independent variables: age, age square, log(sales). Standard errors clustered at company level. *** p<0.01, ** p<0.05, * p<0.1

Table 18
Number of financially literate colleagues by firm size (descriptive statistics)

Firm size	p90	mean	max
<50	0	0.007	4
<60	0	0.008	4
<70	0	0.011	6
<80	0	0.012	6
<90	0	0.013	6
<100	0	0.014	6

If we decrease the number of employees the coefficient increases (Table 19). This probably shows that at smaller firms the employees know and trust each other better so their advise is more valuable.

Table 19
Robustness check: changing the firm size threshold

Nr of colleagues	<10	<20	<30	<40	<50
Highfinlit colleague	-0.406*	-0.371**	-0.230**	-0.177*	-0.214***
	(0.230)	(0.150)	(0.116)	(0.0982)	(0.0805)
Nonfinlit colleague	0.0339	0.0419	0.0290	0.0524	0.0412
	(0.103)	(0.0623)	(0.0475)	(0.0399)	(0.0356)
Midfinlit colleague	-0.0997	-0.0135	-0.0220	-0.0634	-0.0513
	(0.102)	(0.0587)	(0.0473)	(0.0412)	(0.0369)
Skill:High	-0.255***	-0.241***	-0.237***	-0.297***	-0.273***
	(0.0905)	(0.0650)	(0.0558)	(0.0506)	(0.0473)
Skill:Mid	-0.148***	-0.138***	-0.145***	-0.167***	-0.177***
	(0.0399)	(0.0298)	(0.0264)	(0.0246)	(0.0235)
Observations	5,586	9,911	12,490	14,363	16,007
R-squared	0.186	0.165	0.154	0.153	0.153
Industry × Time FE	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Mean of dep. var.	4.83	4.84	4.84	4.83	4.82

Dependent variable: mortgage loan interest rate. Finlit colleague – shows if there is a colleague with mathematics score above the 90th percentile who took a mortgage in the previous year. Nonfinlit (Midfinlit) colleague – same as Finlit colleague before but mathematics score is below the 50th (between the 50th and 90th) percentile. Other independent variables: age, age square, log(sales). Standard errors clustered at company level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Second, we focus on two bank related issues. To see whether the bank concentration also matters in the small firm specification we separate our sample based on the number of banks in each district. Here we use an alternative measure for bank concentration to see how robust is our results for changing this definition as well. We split our sample based on the number of banks in each district. The threshold is the median number of banks.

The first column of Table 20 shows the estimated peer effect if all borrowers and districts are included in the sample and no bank fixed effects are estimated. In Column (2) we restrict our sample to those districts where the number of banks is lower than the country median, i.e where competition is lower. The peer effect is more than 50% higher than in the baseline. We suspect that the reason is that in districts where the competition is low it is more difficult to find the good offer than in other districts, as there can be more expensive products on the local market as well. Lower competition corresponds to higher prices and wider gap between the banks' offers, therefore finding a better solution with the advice of peers can be more valuable. If we run the regression for those districts, where the number of banks is higher than the country median the peer effect disappears. These results are in-line with the main specification in Section 4.2, where we found that the peer effect is higher, where the bank concentration is lower.

In Columns (4)-(6) of Table 20 we included bank fixed effects. Intuitively, if the estimated coefficients decrease (in absolute terms) that indicates that the advice of peers includes which bank to choose. However if the coefficients do not change considerably that is a sign that advice is not related to the choice of banks. By including bank fixed effects the estimated parameters of peer effects decreases by one third. This result highlights that although the choice of bank is an important part of the peer effect it is not the only channel through which colleagues help each other.

All specifications with the small firm setup shows qualitatively and quantitatively similar results as in the main regressions.

Table 20
Heterogeneity according to bank competition

	(1)	(2)	(3)	(4)	(5)	(6)
Nr. of banks	Full smpl	≤p50	>p50	Full smpl	≤p50	>p50
Highfinlit colleague	-0.214***	-0.360***	-0.00524	-0.135*	-0.240**	0.0224
	(0.0805)	(0.115)	(0.115)	(0.0725)	(0.104)	(0.104)
Nonfinlit colleague	0.0412	0.0363	0.0341	0.0188	0.0384	-0.0215
	(0.0356)	(0.0439)	(0.0642)	(0.0320)	(0.0397)	(0.0567)
Midfinlit colleague	-0.0513	-0.0378	-0.0814	-0.0171	-0.0209	-0.0309
	(0.0369)	(0.0471)	(0.0588)	(0.0326)	(0.0416)	(0.0529)
Skill:High	-0.273***	-0.279***	-0.241***	-0.174***	-0.140**	-0.190***
	(0.0473)	(0.0627)	(0.0742)	(0.0419)	(0.0551)	(0.0665)
Skill:Mid	-0.177***	-0.172***	-0.168***	-0.120***	-0.115***	-0.112***
	(0.0235)	(0.0301)	(0.0384)	(0.0208)	(0.0266)	(0.0343)
Observations	16,007	10,281	5,681	15,995	10,268	5,677
R-squared	0.153	0.157	0.180	0.322	0.331	0.341
Industry × Time FE	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Bank FE	No	No	No	Yes	Yes	Yes
Mean of dep. var.	4.817	4.90	4.66	4.82	4.90	4.66

Dependent variable: mortgage loan interest rate. Finlit colleague – shows if there is a colleague with mathematics score above the percentile 90th who took a mortgage in the past year. Nonfinlit (Midfinlit) colleague – same as Finlit colleague, before but mathematics score is below the 50th (between the 50th and 90th) percentile. Column (2) and (5) are restricted to those districts, where the number of banks is less than or equal to the country median. Column (3) and (6) are restricted to those districts, where the number of banks is more than the country median. Other independent variables: age, age square, log(sales). Standard errors clustered at company level. *** p<0.01, ** p<0.05, * p<0.1

MNB Working Papers 2025/3
Birds of a Feather Indebted Together — Solutions to the Information Problem in the Case of Mortgage Loans
Budapest, December 2025

mnb.hu

1054 BUDAPEST, SZABADSÁG TÉR 8-9.