Bank of Israel **Research Department**

Does Location Matter?
Evidence on Differential Mortgage Pricing in Israel[1]

Natalya Presman[*] and Nitzan Tzur-Ilan[**]

Discussion Paper 2019.10
November 2019

Bank of Israel; http://www.boi.org.il

[1] The views expressed in the paper are those of the authors and do not necessarily reflect those of the Bank of Israel.

[*] Bank of Israel, Research Department, nitzan.tzur@boi.org.il

[**] Bank of Israel, Research Department, natalya.presman@boi.gov.il

חטיבת המחקר, בנק ישראל ת"ד 780 ירושלים 91007
Research Department, Bank of Israel, POB 780, 91007 Jerusalem, Israel

האם מיקום הנכס הנרכש משפיע על תמחור משכנתה?

נטליה פרסמן וניצן צור אילן

תקציר

המאמר בוחן את מידת ההשפעה של גורמים שונים על קביעת שיעור ריבית המשכנתאות בישראל. כדי לשקף את הסיכונים הנובעים ממאפייני שוקי הדיור האזוריים ואת מידת התחרות השוררת במערכת הבנקאית אנו משתמשות בבסיס נתונים ייחודי שמשלב נתונים פרטניים על המשכנתאות שהועמדו על ידי המערכת הבנקאית בשנים 2010—2013 ומאפייני הנכסים שמומנו באמצעות משכנתאות אלה ונתונים נוספים. המחקר מצביע על פערים מובהקים סטטיסטית בשיעורי ריבית המשכנתאות שניתנות למימון רכישת נכסים שממוקמים באזורים גיאוגרפיים שונים ובאזורים סטטיסטיים ברמות סוציואקונומיות שונות. בעוד שהריבית שרוכשי דירות בשכונות מבוססות כלכלית במרכז הארץ משלמים היא הנמוכה ביותר, הריבית שרוכשי דירות בשכונות מוחלשות כלכלית באזורי פריפריה משלמים היא הגבוהה ביותר.

התכונות הנצפות של הלווים, המשכנתאות והנכסים הממומנים - וכן מידת התחרות בין הבנקים, יכולות להסביר עד כשני שלישים לכל היותר מהפערים הבין-אזוריים בריבית המשכנתאות. גורמים נוספים שיכולים להסביר את הפערים, כוללים תכונות לא-נצפות של הלווים, כגון אוריינות פיננסית וכושר מיקוח, מאפייני תעסוקה לא-נצפים, ומידה מסוימת של אפליה סטטיסטית של קבוצות לווים מסוימות - מצב בו תמחור הסיכון של הלקוח מושפע מעצם היותו משתייך לקבוצת לווים שמתאפיינת בכך ששיעור המתקשים לעמוד בתשלומי המשכנתא בה גבוה יחסית.

Does Location Matter?

Evidence on Differential Mortgage Pricing in Israel

Natalya Presman and Nitzan Tzur-Ilan

Abstract

This paper explores the contribution of various factors to determining mortgage interest rates in Israel. We use a unique database combining loan-level data on mortgage loans originated by the Israeli banking system during 2010–13 with proprietary data on assets underlying mortgage origination as well as several additional variables designed to capture risk associated with regional real estate markets and the extent of competition prevailing in the banking system. We show that significant differences exist in real mortgage interest rates among different locations and neighborhood qualities. While homebuyers purchasing assets in the more prosperous central neighborhoods pay the lowest interest rates, those purchasing assets in the peripheral and economically weak neighborhoods pay the highest ones. Observable characteristics of the borrower, the mortgage and the underlying asset risk, and banking competition explain up to two thirds of the regional and socioeconomic differences in mortgage interest rates found in the raw data. Other factors that may explain remaining regional differences in the interest rates include unobservable borrower characteristics such as financial literacy and bargaining ability, unknown characteristics of borrower's employment and statistical discrimination of some groups of borrowers.

1. Introduction

Investment in housing is the main share of households' investment portfolio, especially among households in the lowest income deciles. Balance sheet data of the Israeli households sector for 2017 show that the value of real estate constitutes 51 percent of all households' wealth (real and financial assets combined).[1] According to data from Israel's Central Bureau of Statistics, in 2017, 71.8 percent of households owned at least one housing unit.[2]

Since housing investment usually cannot be financed solely by equity,[3] mortgage credit availability (that is, mortgage approval) and its affordability (that is, the cost of credit, i.e., the interest rate) are crucial to the ability of households to accumulate wealth. Given the large scale and extremely long duration of mortgage loans, even small differences in interest rates may add up to large amounts of money "overpaid" to the lender. Although differences in the interest rates are supposed to reflect differences in the risk stemming from borrowers and their investment and financing decisions, they may also incorporate some element of differential treatment towards certain types of borrowers—racial and religious minorities, women, immigrants, residents of specific regions, etc.

Despite the public and political importance of the issue, it is difficult to conduct empirical research of differential treatment in mortgage lending. This is mainly due to the lack of suitable databases that combine all the information needed for such research, including detailed loan conditions, borrower demographic and occupational characteristics and credit history, and influential characteristics of the underlying asset. The scarcity of the data is the main reason that the issue of discrimination in mortgage lending was explored mainly in the US, where collection of the data became mandatory in the early nineties, but to a much lesser extent in Europe, where such data were not collected due to non-existence of anti-discriminatory legislation.

We do not focus on the discrimination issue but rather we explore the contribution of various factors to determining mortgage interest rates, paying special attention to the role of the distance from the business center of the country (Tel Aviv) and

[1] For details see Financial Stability Report for the first half of 2019, Bank of Israel.
[2] Specifically, 61.8 percent owned one housing unit and 10 percent owned two or more housing units.
[3] In Israel, approximately 85 percent of home purchases are financed using a mortgage.

socioeconomic status of the neighborhood where purchased housing asset is situated. In other words, we examine if Israeli lenders price differently similar mortgage products offered to comparable households purchasing homes in different locations that differ in the liquidity of assets and their potential price growth, thus reflecting different costs of foreclosure. We use a unique database combining loan-level data on mortgage loans originated by the Israeli banking system during 2010–13 with proprietary data on assets underlying mortgage origination as well as several additional variables designed to capture risk associated with regional real estate markets and extent of competition prevailing in the banking system. These data provide an opportunity to conduct a large scale and relatively complete study of the potential differential treatment of mortgage borrowers. The uniqueness of the data means that, to the best of our knowledge, no previous research has considered the questions addressed in this paper.

Controlling for multiple factors that might affect interest rate pricing, we show that significant differences exist in real mortgage interest rates among different locations and neighborhood qualities. While homebuyers purchasing assets in the more prosperous central neighborhoods pay the lowest interest rates, those purchasing assets in the peripheral and economically weak neighborhoods pay the highest ones. Observable characteristics of the borrower, the mortgage and the underlying asset risk, and banking competition explain up to two thirds of the regional and socioeconomic differences in mortgage interest rates found in the raw data. Additional factors that may explain these differences—borrower's credit history, wealth, employment characteristics (in terms of occupation, seniority, tenure, stability and employment contract duration), financial literacy and bargaining ability—are unobservable in our data. We also assume that lenders incorporate the risk of asset foreclosure in the cost of credit, and the costs of foreclosure seem to be higher in regions with lower housing demand and lower prospective house price growth (or higher possibility of price depreciation). It also seems reasonable that lenders consider past experience concerning the incidence of loans in arrears of similar borrowers.[4]

The rest of the paper is organized as follows. Section 2 surveys relevant literature. Section 3 explains the data. Section 4 presents the descriptive statistics. Section 5

[4] Partly due to the absence of a Public Credit Registry in Israel in this period. Recently, the Credit Data System was founded in Israel and it began to operate in April 2019.

outlines basic empirical framework. Section 6 reports the results. Section 7 presents some robustness checks, and Section 8 concludes.

2. Related literature and basic econometric issues

Historically, research on differential mortgage pricing has dealt with discrimination issues. As noted above, we are not focusing on discrimination, but we refer to this research in the construction of our empirical framework.

Despite the existence of a great deal of work concerning discrimination in mortgage lending in the US, mainly against African-American and Hispanic borrowers, very little research has been done dealing directly with *pricing inequality*. The conventional explanation for the scarcity of such studies is lack of appropriate data on interest rates.[5] Studies that have succeeded in overcoming this issue are essentially case studies focusing on data from a single particular lending institution; the immediate and obvious disadvantage of such studies is that their findings cannot be generalized to the market level.

The majority of such case studies document significant differences in lending terms between minority borrowers and whites, but in most cases they cannot attribute these differences to solely racial issues. For example, researchers who analyzed the incidence of paying overages[6] and their size were not convinced that it was exactly minority status that shaped the differences. Courchane and Nickerson (1997) suggest that differences in bargaining and negotiating power of whites and minorities may have caused the observed racial differences. Similarly, Black, et al. (2003) conclude that the differences in overages have more to do with the market power of the lending institution and differential bargaining skills of borrowers and have less to do with the borrower's race. Moreover, Crawford and Rosenblatt (1999), who document significant price

[5] Under a credit rationing regime, discrimination of minorities could take place at the approval/denial stage; the US mortgage market was considered as a highly competitive one where lenders have little room for differential rate manipulations in the sense that long-term loans were made at a very thin spread over lenders' cost of funds (Holmes and Horvitz, 1994). Even in the 1990s, after the transition to a risk-based pricing regime (due to improvement of statistical models of individual risk assessment and substantial reductions in the data storage costs), the lenders were not required to report their lending terms for all individual loans, but only to indicate high-interest ones.

[6] An "overage" is a type of premium, the difference between the price at which a loan closes and the minimum price acceptable to the lending institution for specific loan products and for borrowers with particular credit attributes. Since the borrower is usually unaware of how the loan is priced, lack of financial information, severe liquidity constraint, risk aversion, or unwillingness to bargain could lead to an overage.

differences between individual borrowers, found these differences largely race-neutral, controlling for various borrower demographic and financial characteristics, differences in market rates, and rate-lock protection periods. Likewise, Courchane (2007) concludes that after controlling for individual and market characteristics, relatively little of the differences in the annual percentage rates (APR)[7] paid by minority compared to non-minority borrowers are attributable to the differential treatment of borrowers. Some studies used explicit data on mortgage interest rates of subprime mortgages, but their results are also inconclusive: Ghent, et al. (2014) document evidence of adverse pricing for African-American and Hispanic borrowers in subprime mortgage market in metropolitan areas of California and Florida during 2005, while Haughwout, et al. (2009) find no evidence of pricing discrimination against minority borrowers (possibly due to missing data on mortgage origination costs).

In France, Gary-Bobo and Larribeau (2004) show that lenders exercise "social discrimination" against workers vs. professionals due to the differences in elasticities of demand (to buy a home as opposed to rent) and in perceived default risks. In Spain, controlling for a large set of household, mortgage and market characteristics and using Oaxaca-Blinder decomposition, Diaz-Serrano and Raya (2011) find that between two-thirds to three-fourths of the gap in average mortgage interest rates between immigrant and native borrowers can be attributed to discrimination.

The only study that treats *interregional* disparities in mortgage rates is Eichengreen (1984). He aims to explain lower mortgage rates for agricultural land in the North and South Atlantic states (Eastern states) in the US compared to the rest of the country. In a simple framework of mortgage interest rate determination, he finds that when mortgage interest rates are adjusted for risk (expressed in the price of the land—with the consideration that the market prices risk—and types of agricultural crops) and for the effects of statutory interest rate ceilings, varying widely among the states, there remains no evidence of significant interregional differentials.

One of the drawbacks of the studies dealing with discrimination is an omitted variable bias caused by unobservable features, such as financial literacy and bargaining ability of a borrower, and leading to biased estimates of discrimination variable coefficients.

[7] The APR is based on the full cost of the loan, including both the interest or note rate on the loan and additional charges and fees, amortized over the full loan term.

This issue is essentially insoluble, because banks do not collect data that can be used as a proxy. Even the borrower's income is not a perfect proxy, because it is not necessarily correlated with financial skills. Studies that attempt to overcome this issue are scarce. Cheng, et al. (2015) use data from the Survey of Consumer Finance (SCF), instead of administrative data provided by lending institutions, exploiting the advantages of more detailed household information[8] and a unique variable of shopping behavior.[9] They document persistent statistically and economically significant racial differences in mortgage rates, but while individual risk factors affect differential pricing of mortgages, shopping behavior appears to be a weak explanatory factor. More accurate analysis (using residual analysis and quantile regressions) reveals that racial discrimination is concentrated in the most vulnerable demographic subgroups, such as younger and lower-educated minority borrowers. These results are quite the opposite of their previous findings concerning gender gaps in mortgage pricing.[10] In the earlier study, Cheng, et al. (2011) find that the traditional explanatory variables (mortgage features, borrower characteristics, market conditions, etc.) can only partially explain such gender disparity, while a behavioral variable that captures how men and women differ in shopping for mortgages can completely explain the rest of the gap. In particular, they claim that men are likely to pay lower rates on mortgages because they simply tend to search more for the lowest rates.[11,12]

A much more severe critique relates to the modelling of the functional relationship between mortgage outcome (rejection/interest rate/default rate) and the discriminatory feature (race in most cases). As Yezer (2010) states, there is no fully developed *theoretical model* of the mortgage lending process. Single-equation *empirical models* of mortgage outcomes all relate mortgage outcomes to a variety of "causal" variables, which include various loan terms: loan amount, loan to value ratio (LTV), payment to

[8] The SCF collects both detailed loan information, including type of mortgage, loan amount, term, interest rate, time of origination, as well as borrower's characteristics, including age, race, education level, information on wealth and debts, past bankruptcy and credit applications that were rejected in the past five years.

[9] This information distinguishes between those who primarily rely on recommendations by people they trust and those who make an effort in searching and comparing among multiple loan offers.

[10] There is some empirical evidence that women are more likely than men of the same race to be subprime mortgage borrowers (Fishbein and Woodall, 2006).

[11] In their sample, 42.1 percent of men selected their lenders based on the search for the lowest rate offer, whereas only 20.5 percent of women behaved in the same way.

[12] Moreover, the results suggest that searching for the lowest rates is much more beneficial for borrowers who choose adjustable rate mortgages, compared to those who choose fixed rate mortgages, supporting the hypothesis that the impact of search is greater for more complex mortgage products.

income ratio (PTI), term to maturity, cosigner, financial characteristics of the applicant, characteristics of the real property collateral, and variables reflecting demographic factors, including minority status of the applicant or borrower. The assumption underlying such single-equation models is that the mortgage outcome variables have no role in causing the loan terms. In other words, applicants have no knowledge of the relation between the loan terms that they request and loan outcomes. As those, they never behave strategically and mortgage terms are selected by lender. However, a much more reasonable assumption is that loan applicants recognize that mortgage interest rate[13] is a function of the mortgage terms, such as loan amount, down payment/LTV, monthly PTI, and term to maturity, and that they may influence these variables to lower the interest rate. Furthermore, sometimes lenders do not allow the initial terms required by the applicants; instead, they demand to increase equity, to change loan maturity or to underwrite a cosigner.

From the econometric point of view, the "independent" right-hand side variables are jointly determined with the dependent variable, leading to an endogeneity and identification problem, causing estimates obtained using single-equation techniques to be biased and inconsistent. For example, as Yezer (2010) claims, higher-risk applicants can self-select into loan programs with higher mortgage rates and higher rejection and default rates. These outcomes are due to applicants' self-selection into particular loan programs, not to differential treatment by lenders. As a result, discrimination tests tend to produce false positive indicators of discrimination when none exists.

But even without such opportunistic behavior the problem of simultaneous equations bias occurs when applicants increase their down payment, i.e., lower the LTV, to gain more favorable interest rates. Households with more resources are better able to avoid high interest rates, supplying additional equity. The conventional solution of the identification problem is utilization of some instrumental variable that can explain the LTV but is not associated with the interest rate. Unfortunately, the standard set of variables in mortgage files does not include such variables. The econometric issue is even more complicated since not only LTV is simultaneously determined with the interest rate, but also other variables that can be used to influence the loan terms, such

[13] Although the reasoning is implemented equally on rejection and default rates, from now on we will refer to only interest rate, as this is the focus of our research.

as PTI and term to maturity. This means that each of these variables is to be modeled separately to obtain unbiased estimates.

We are aware of econometric problems with single-equation estimation of mortgage interest rate, but we have no suitable data to overcome these problems. In any case, we assert that in the Israeli case the severity of the endogeneity problem is much weaker than in the US. We claim that in Israel the self-selection problem resulting from borrowers' opportunistic behavior and moral hazard is not an issue because of the recourse nature of mortgage loans, i.e., loans that allow the lender to take action above and beyond the foreclosure of housing asset securing the mortgage. In Israel, mortgages are mostly originated by banks: 94 percent of mortgage stock is in the banking system, while only 4 percent of the mortgage stock is in nonbanking financial institutions such as pension funds, with savings as a lien. Mortgage origination at banking institutions is possible only in person at the bank branches and operated only by bank employees; requirements of employment documentation are uniform (wage stubs for the three previous months for all employed members of the household); the lock-in period is uniform for all banks and is defined by the Supervisor of Banks; interest rates do not incorporate ex-ante prepayment penalty, such penalty is charged to borrowers only at the time of actual prepayment, mostly due to switching to an improved mortgage contract[14]. Furthermore, there is no sub-prime market for mortgages in Israel.

Even if we had some data to attempt to model LTV, PTI or time to maturity, we would face an additional econometric issue; we don't know what the functional form of the relationship between the interest rate and other loan terms is. Anyway, it doesn't seem to be linear and continuous. It makes sense that there are some levels of these risk variables that signal that the risk associated with a loan has climbed to another, higher level. Such significant change has to cause the interest rate to jump. To extract such signals we can utilize macroprudential policy tools that were implemented in Israel and their timing. We consider those macroprudential tools that raise the cost of lending for homebuyers. It is probable that these signals are commonly recognized to be the main risk-increasing factors and they were always considered by the banks in their decision-making in housing credit. Several macroprudential tools were implemented during our

[14] This is different from the US experience, where the cost of future prepayment is incorporated in the mortgage interest rate as an option while a borrower pays for the option at loan origination.

research period. First, in May 2010, banks were required to make additional provisions for housing loans with LTV higher than 60 percent. Later that year, in October 2010, banks were required to make higher capital provision for loans above NIS 800,000, LTV higher than 60 percent, and share of variable-rate loans higher than 25 percent. In May 2011, the share of the loan with variable rate based on Bank of Israel interest rate was limited to one-third. In February 2013, the risk weights for capital adequacy requirements rose for loans with LTV higher than 45 percent. In August 2013, the PTI ratio was limited to 50 percent while risk weights for capital adequacy requirements on loans with PTI exceeding 40 percent were raised to 100 percent, and loan duration was limited to 30 years.

To lessen the econometric hardship, we will not use the loan terms in their continuous form; instead, we will divide the mortgages into less risky and more risky ones with the levels of these variables that we regard as risk-increasing: LTV more than 60 percent, PTI above 30 percent, and duration above 20 years.

3. Data

For the purposes of the macroprudential policy, mainly in order to perform stress tests, during 2015, Banking Supervision Department at the Bank of Israel required all banking corporations to report retroactively on all approved mortgage originations during each calendar year, beginning from 2010. These reports contain diverse data on loan characteristics, including approved loan amount and duration, loan to value ratio, mortgage monthly payment to net income ratio, purpose of the purchase (first home, upgrading or investment), interest rates set at mortgage origination, the type of interest rate (fixed or variable, adjustable or not adjustable), and type of the benchmark for adjustable rates. In addition, the data contain indication of the bank branch where the mortgage was originated.

In Israel, mortgage borrowers often decide to take out combination loans, consisting of several types of interest rates, including fixed and variable rates, real and nominal rates and adjustable rates with different benchmarks.[15] This makes the procedure of

[15] For example, one of the most popular combinations in recent years has been a mortgage that is one-third adjustable based on the Bank of Israel Rate, one-third fixed rate (CPI-indexed or not) and one-third

transformation of the reported interest rates to uniform ones somewhat tricky. We start with transforming all interest rates to real terms by subtracting the inflation expectations rate at mortgage origination (with data on banks' expectations[16] for 1, 2, 5 and 10 years, according to mortgage duration[17]) from all CPI non-adjusted rates in our sample. Then we calculate the weighted average of the real interest rates on all parts of each mortgage, weighted by the shares of all parts. Although the final price of a mortgage includes also two other components in addition to the interest rate – the mortgage opening fee and obligatory purchase of property appraisal – we don't have information on the extent of these expenses. The lack of these data is not crucial for our analysis due to uniformity of these costs – 0.25 percent of the loan amount, with minimum commission of NIS 500. Although some qualitative borrowers can receive a discount or even total exemption from paying the commission, more risky borrowers who are obliged to pay the full commission will pay approximately 0.01 percentage point in addition to yearly mortgage interest rate (with an average loan duration of 21 years in our data), while the average real interest rate is 1.75 percent in our data.

The banks also report some features of the borrowers, including after-tax monthly household income and certain monthly expenses (such as other loan repayments), number of borrowers (single or couple), age of all borrowers, if there was a guarantor (bank's requirement for more risky borrowers), and if the borrower manages current (wage) account at the same banking institution.

The data on mortgage originations include accurate property location, date of purchase and location of the bank branch where the loan was originated. The mortgage database was merged with the home-sale transactions database hold by the Israel Tax Authority, containing information on housing unit characteristics (CARMEN). This procedure left us with approximately one third of the observations from the mortgage database, because of omitted and partial information issues[18], but it enabled us to locate the neighborhood where purchased property is situated and to account for the distance from

variable rate changing every 2 or 5 years. This composition resulted from the restriction to at most one-third of the share of mortgage bearing interest rate adjusted to the Bank of Israel Rate, effective from May 2011 (as a measure of the macroprudential policy).

[16] We use the series of bank's inflation expectations, calculated by the Bank of Israel from data on banks' interest rates on CPI-indexed and unindexed loans and deposits.

[17] Since there are no data on longer ranges of inflation expectations, we use data on inflation expectations for 10 years for longer term mortgages as well.

[18] For a detailed explanation of the combined database construction see Tzur-Ilan (2017).

the neighborhood to the center of Tel Aviv and for the neighborhood socioeconomic status.[19] The Socioeconomic index serves as a proxy for neighborhood risk.

Unfortunately, we do not have information on borrowers' credit history, employment characteristics or net wealth. Additionally, some important factors influencing mortgage interest rates are unobservable, for example, borrowers' financial literacy, bargaining skills and shopping behavior.

4. Descriptive Statistics

Our data include 88,914 mortgage originations set between January 1, 2010 and December 31, 2013. Figure 1 presents the distribution of average real mortgage interest rates by distance from the center of Tel Aviv (the business center of Israel), divided into three groups: less than 40 km, 40 to 80 km, and 80 km and above. Figure 1 shows that as the distance from the center increases, the entire distribution moves to the right, meaning that the incidence of higher-priced mortgages in the peripheral neighborhoods is higher than in the central ones. According to the Kolmogorov-Smirnov test, differences between the three distributions are significant. However, not only distance matters. We add a socioeconomic status dimension by defining three classes—low, middle and high, with approximately one third of observations in each class. Figure 2 shows the distribution of average real interest rates by socioeconomic status for each group of distance from the center of Tel Aviv. The distributions for high socioeconomic class are shifted left for all groups of distance from the center, but most obviously for the most distant group. Generally, in the group of the most peripheral neighborhoods the differences between socioeconomic classes are the most striking. The Kolmogorov-Smirnov test shows significant differences between the distributions.

Since differences in the mortgage interest rates have two-dimensional nature (location and neighborhoods quality), for further analysis we use nine combinations of distance and socioeconomic status (interactions). Table 1 presents median values of weighted

[19] The Israeli Central Bureau of Statistics constructs a socioeconomic index of neighborhoods, consisting of 16 different variables, including demography, education, employment, income, and standard of living. The 16 variables are combined into a single index, and all neighborhoods in Israel are classified into one of twenty clusters, 1 being the lowest socioeconomic status and 20 being the highest.

average real interest rates, LTV, weighted average duration (different parts of the mortgage may have different durations), PTI, size of the loan, monthly net (after taxes) income of household, age of the borrower (average age for borrower couples), loan to income ratio (LTI, loan size divided by yearly net income), weighted average real interest rates for investors and first-time home buyers along with investors' share in mortgages, percentage of borrowers who took out a mortgage outside the locality of purchased property and number of observations (mortgages), for nine interactions of distance and socioeconomic status. It is obvious that there are non-negligible differences in the real interest rates, while interest rates vary by both dimensions: they rise with the distance from the center and with deterioration of socioeconomic status. As a result, the lowest median real interest rate is found in the prosperous neighborhoods near the center of Tel Aviv, while the highest median real interest rate is found in the economically week peripheral neighborhoods. The data also show that households purchasing housing assets close to the center, independent of the socioeconomic class of the neighborhood, have lower LTV ratios and higher incomes, but take out larger mortgages with longer durations and higher PTI and LTI ratios. Also, investors' share is higher in the distant locations, especially in the distant and weak neighborhoods, where properties are relatively low-priced. However, investors receive consistently more favorable mortgage rates than first-time homebuyers in the same locations, probably due to their more favorable financial status. The share of those taking out mortgages outside the locality of purchased property is high in all regions, but it is lower in the remote regions compared to the central ones. Taking the size of locality into consideration makes it clear that borrowers negotiating for mortgages outside the locality of the purchased property are those purchasing assets in smaller localities, where the competition among banking institutions is anticipated to be lower. A higher percentage of mortgages that were originated outside the locality of purchased property in the central regions is possibly explained by the geographical proximity and territorial continuity between cities in Gush Dan (the metropolitan Tel Aviv area).

5. Basic empirical framework

A simple competitive loan-pricing model is based on the hypothesis that a financial institution's lending decisions are a function of risk and return factors that affect the

expected net present value of the loan. To maximize profits, financial institutions are assumed to accept those loan applications that generate a net present value that exceeds zero.

The model assumes that the interest rate charged on any loan includes four components: (1) the cost of funds incurred by the bank to raise funds to lend, while such funds are obtained either through customer deposits or through capital markets; (2) operating costs associated with servicing the loan, including application processing, monitoring, personnel remuneration and other current expenses; (3) a profit margin on capital; and (4) a risk premium to compensate the bank for the degree of default risk inherent in the loan.

The first three components may vary among lending institutions and over time. The risk for the lender arises mainly from the possibility that the borrower might remain in arrears, forcing the lender to foreclose. In the countries where the extent of the liability is limited to the value of collateral, for example in the US, the lender can suffer losses if sales proceeds are insufficient to cover the principal, interest, legal fees, and transaction costs of reselling the property. In Israel, this risk is minimized because a borrower's liability is not limited to the value of collateral and lenders can pursue other assets of the borrower to mitigate default-related losses.

According to the risk-based pricing theory, the risk premium is determined individually for each borrower and is influenced by a variety of factors linked to borrower characteristics, loan characteristics, and collateral characteristics.

A number of financial and nonfinancial characteristics of individual borrowers are systematically related to creditworthiness. Higher wages of family members and higher household income tend to reduce the chance of missed payments or default. However, higher obligations-to-income (for example, repayment of other loans or other fixed family expenses, some of which grow with family size) leaves less money available to service the mortgage. Financial stability and wealth, including liquid asset holdings, tend to increase with age, reducing the probability of loan delinquency. Higher education level of borrower guarantees not only higher present wage but also better employment stability and prospects. It also indicates higher financial sophistication. Borrowers buying homes for investment purposes are usually older and wealthier than

first-time homebuyers, and probably have other assets to back the loan. Home upgraders have a history of repayment of a mortgage.

Given all the abovementioned features, it should be noted that borrower's bargaining skills definitely affect interest rate determination. Some individuals might be better at bargaining than others and they may obtain better terms than would similar borrowers who lack these skills. Bargaining skills cannot be measured directly but they are probably correlated with other borrower's characteristics, including education, age, prior property ownership (in the case of upgraders and investors), i.e., those characteristics that increase expertise, experience, confidence and reasoning ability.

But even taking under consideration various objective borrower characteristics, predicting future loan delinquency, from the perspective of the lender, is problematic since many credit problems arise from events that are difficult to foresee, such as illness or disability, divorce, and job loss. Concerning the latter, lenders may believe that higher unemployment rates in the peripheral regions make the income of periphery residents more volatile, on average, over the economic cycle, compared to that of residents of the central regions, even controlling for the type of job, and hence increasing the probability of delinquency of borrowers in the periphery.

Concerning loan characteristics, there are several factors that can obviously increase the risk of default. A higher LTV ratio means that less collateral backs the loan, which implies greater risk for a lender. Longer maturities imply increased probability of default, since it increases the chances that borrowers will encounter a situation affecting their ability to repay a loan. Variable-rate mortgages have higher default risk if interest rates move upward. A higher loan amount subject to approved LTV[20] means better property in upscale neighborhoods and also a higher extent of down payment; as such it may be less risky for the lender since the borrower will work harder to keep the property. Higher PTI increases the risk of default, especially for low-income households; in Israel, the PTI has been constrained to 50 percent since August 2013. As previously discussed, there is an econometric difficulty in including mortgage characteristics such as LTV, PTI and duration in the regression explaining the interest rate because of the endogeneity problem. Therefore, to weaken the severity of this

[20] It is crucial to examine the effect of loan amount when LTV ratio is in the regression as well, to ensure that higher loan amount does not suggest higher LTV.

problem, we do not include these variables as continuous, but we only define more risky mortgages as those with LTV higher than 60 percent, PTI higher than 30 percent and time to maturity above 20 years.

Some no-negligible risk stems from the collateral itself. This risk originates mainly from future price movements and assets turnover since these factors affect the potential of selling the property (by borrowers themselves or by the lenders) to settle a loan in the case of personal financial distress. Lender expectations of home price appreciation affect the assessment of mortgage risk; it may be assumed that the quality of a neighborhood and a socioeconomic level of its residents are interlinked with home prices and potential price appreciation. Neighborhoods with low turnover will tend to have more uncertain housing values (independent of housing prices) and, hence, represent greater risk for a lender (Ling and Wachter, 1998; Lang and Nakamura, 1993; Calem, 1996). As an additional risk factor we consider a substantial, all at-once, increase of housing supply in the way of building great number of housing units (building starts, by year and by municipality).

Using this simple model requires caution, since mortgage markets may not be fully competitive. Despite substantial competition on the supply side, mortgages are rather complex products and most of the consumers lack information about mortgage pricing, so regional disparities may reflect the limited ability of consumers to shop for the best products available in the marketplace. In our case, it may appear that there is a specialization by size of the lender, since larger banks have more branches and are represented in more localities including small and remote ones; they could be a main provider of financial services, including mortgages, in these locations. In this case, some lenders may have more market power than others in the peripheral regions.

Our empirical model exploits measures of borrower and location characteristics that are presumed to affect the loan's risk through their expected impact on the probability of default. We estimate the following reduced form linear regression:

$$R_i = \alpha + \beta_1 A_{1i} + \beta_2 A_{2i} + \beta_3 A_{3i} + \beta_4 Competition_i + \beta_5 LenderID_i + \beta_6 Time_i + \varepsilon \quad (1)$$

where R_i is a real weighted average interest rate on a mortgage of borrower i, A_{1i} is a vector of borrower i characteristics, A_{2i} is a vector of mortgage characteristics of borrower i, A_{3i} is a vector of collateral characteristics of borrower i, $Competition_i$ is

a vector of variables measuring the extent of banking competition that borrower *i* faces, $LenderID_i$ is a vector of bank fixed effects, $Time_i$ is the month and year of borrower *i* mortgage origination, and ε is an error term. The variables are introduced in Table 2.

We expect that single borrower, younger age, lower family net income, high LTV ratio, high PTI ratio, long duration, lower socioeconomic index of the neighborhood, lower housing market turnover and fast expansion of housing supply are associated with higher risk and therefore with higher interest rates. We expect that further distance from the center is also associated with higher interest rates but we allow a nonlinear relationship between interest rates and distance. Realizing that there may be several regions that are quite independent from Tel Aviv, we add *Potential Accessibility Index*[21] to our regression; we expect that higher accessibility is associated with low mortgage risk and thus lower interest rate. On the other hand, we expect that home upgraders and those who manage a current account at the same bank are supposed to gain more favorable terms of credit. Home upgraders have not only had loans originated in the past but have been paying them for a period of time. Banks also have more prolonged acquaintance with customers who manage their salary account within the bank, receiving higher precision signals of their creditworthiness.

A larger number of all banking institutions providing mortgage services in the area is expected to be associated with higher competition and therefore lower interest rates. We don't have clear expectations for *Investor* variable; although investors are usually more financially mature, the standard approach ascribes higher property risk to non-owner occupied properties, since those buying secondary homes for investment purposes are less eager to invest money in property maintenance. Sometimes banks demand guarantors to underwrite more risky mortgages, but it is hard to assume if this procedure completely neutralizes the extra risk. *Log of loan amount*, for a given LTV ratio, may signal better property with higher prospects of price appreciation, but in contrast, it means higher total loss for a lender in case of default. The sign of *Bank in*

[21] The index is calculated by the Central Bureau of Statistics according to the gravity model and reflects the proximity of the given locality to each of the localities in Israel, weighted by the size of their populations, with the size of the population indicating the intensity of the opportunities, activities, and assets in each locality. The value of the index ranges from -1.487 (the most remote and least accessible town) to 6.318 (the most central and accessible town). The Potential Accessibility Index is a part of the Peripheriality Index, which also includes the distance from the boundary of the Tel Aviv district.

the same location dummy is also ambiguous, because we do not know the considerations behind such choice. For example, investors are likely to negotiate with lenders close to the place of residence and not close to purchased property which may be located in another locality; commuters may choose to negotiate with lenders located close to their workplace; upgraders may change place of residence, moving to larger homes outside central cities while negotiating for mortgages at their current locations. This variable may also proxy shopping behavior.

The use of the time dummies should control for macroeconomic factors (changes in the basic price of credit) and the effect of macroprudential policies concerning mortgage lending (LTV limits and higher capital requirements for more risky loans). Bank fixed effects control for the differences in the cost of funds, operating costs and business strategies among lending institutions.

The main variables of interest are the interactions of distance and socioeconomic status. Holding constant all available measures of household, loan and neighborhood risk and accounting for degree of banking competition in the area, the econometric analysis is aimed at examining the role of purchased property location in the mortgage interest rate determination.

6. Results

We estimate regression model (1) by OLS in three variations to control for the impact of including some explanatory variables on the effects of interactions between distance and socioeconomic status on the real interest rate. Regression (1) includes the main measures of borrower, mortgage and property risk, and banking competition, according to equation (1). Although we examine the influence of the distance from the business center of the country, utilizing the monocentric model, there are several large urban centers that may be more relevant for small peripheral towns.[22] Following this consideration, we add the Potential Accessibility Index to get Regression (2), along with the number of building starts in the locality as a percentage of existing housing units to control for supply expansion. Regression (3) adds control for the number of

[22] Generally speaking, there are 4 metropolitan centers in Israel: Tel Aviv, Jerusalem, Haifa in the north and Beersheva in the south.

banking institutions in the location of purchased housing assets. All three versions include bank fixed effects and month and year fixed effects (coefficients are not reported,[23] but they are mostly statistically significant). Despite a relatively large number of explanatory variables and very large number of observations, the explanatory power of the model seems to be disappointingly low with R^2 of only 28 percent.

As Table 3 shows, most of the coefficients are statistically significant and have the expected signs. Other things being equal, households with higher incomes pay lower interest rates on mortgages; a ten percent increase in net income is associated with a 0.02 percentage point decrease in the mortgage interest rate. Those who take out mortgages from the banks where they manage their current account pay, on average, 0.13 percentage points less than similar borrowers managing their current account at another bank. Home upgraders pay interest rates that are lower, on average, by almost 0.03 percentage points compared to similarly situated first-time home buyers (our reference group), while investors pay 0.07 percentage points less than the reference group. Those borrowers who face guarantor requirements pay, on average, almost 0.06 percentage point higher interest rates. A higher loan amount is indeed associated with a lower interest rate, as predicted by the literature. Other things being equal, mortgages with LTV above 60 percent bear an interest rate higher by almost 0.06 percentage points. The effect of a PTI ratio greater than 30 percent is not statistically significant, while duration longer than 20 years increases the interest rate by 0.22 percentage points, on average, other things being equal. Households buying assets in neighborhoods with higher socioeconomic status pay lower interest rates, while the effect of the distance from Tel Aviv alone (not in the interaction with socioeconomic status) depends on regression specification. Further on, higher real estate market turnover operates in favor of mortgage borrowers; other things being equal, a 10 percentage point increase in the turnover is associated with a 0.08 percentage point decrease in the interest rate. Rapid expansion of housing supply contributes to an increase in the interest rate, but its influence is statistically significant only in specification (2). Other things being equal, households taking out mortgages from bank affiliates in the purchased asset's locality, pay, on average, 0.04 to almost 0.08 percentage points more, depending on the regression specification; the effect strengthens after inclusion of the control for bank competition in the property locality. It can be argued that those who exhibit shopping

[23] Bank fixed effects are not reported because of confidentiality issues.

behavior and exert effort looking for better deals can obtain lower interest rates. Such conduct can be beneficial since higher competition among mortgage lending institutions both in the town of mortgage origination and in the purchased property location is associated with lower interest rates; in specification (3), each additional banking corporation in the property location is expected to lower the average interest rate by 0.016 percentage points and each additional banking corporation in the town of mortgage origination is expected to lower the average interest rate by another 0.009 percentage points, other things equal.

Only three variables have unexpected signs. Contrary to our expectations, the *number of borrowers* variable has a positive sign, meaning that couples pay higher interest rates than singles, other things equal. Single borrowers are relatively rare (only 12 percent of all mortgage borrowers in our data), and perhaps have solid enough economic background[24] to gain favorable mortgage terms. The sign on *age* variable is positive meaning that older borrowers pay higher interest rates. However, the age variable is apparently correlated with several other variables including *upgrader* and *investor* dummies and with *net income*, so that it only partly captures the effect. The influence of the *Potential Accessibility Index* is positive, meaning that those purchasing assets in the central and accessible towns pay higher interest rates, but it is not economically sizable in any case.

Now we turn to understanding the effects of interactions of distance and socioeconomic status on mortgage pricing. Inspection of estimation results in Table 3 shows that all coefficients on interaction terms are highly statistically significant and all have an expected positive sign, since the omitted category is prosperous and close to the center neighborhoods. One can also observe that except for two coefficients (*Dmid_SElow, Dmid_SEmid*) the order of coefficients' magnitude varies in the expected way: keeping the distance category the same, the coefficients decrease with the improvement of the socioeconomic status; while keeping the socioeconomic status the same, the coefficients increase with the distance from the center.

Table 4 shows the unconditional means of the calculated real interest rates by nine interactions of distance and socioeconomic status (Panel A), the differences between

[24] In our data, average net monthly income of single borrowers is 71 percent of the average net monthly income of couples (approximately NIS 11,400 and NIS 16,000, respectively).

the means in each group relative to the group of the shortest distance and the highest socioeconomic status—categories omitted in the regression analysis (Panel B) and the estimated coefficients (Panel C).

Comparison of unconditional and conditional differences reveals that, on average, above one half of the unconditional differences in means are explained by the characteristics of borrower, mortgage and underlying asset risk, along with banking competition, included in the regression. Interestingly, the extent of explanation varies from only 40 percent in the most distant regions to as much as two thirds in the rest. The rest of the differences probably may be explained by unobservable regional and borrower characteristics, asset risk that is not captured by included variables and/or differential treatment of borrowers.

7. Robustness checks

7.1 The endogeneity issue

Since we are mostly concerned about the endogeneity problem in our econometric setting, we start with examining this issue and the possibility that our estimation results are distorted. We begin with examining simple pairwise correlations between the mortgage interest rate and its main characteristics: LTV, PTI and loan duration (in their continuous form). Table 5 demonstrates that these correlations are fairly low, meaning that linear relationships between the interest rate and each one of the variables that are expected to be endogenous is relatively weak.

We also estimate our regression model without the measures of loan risk (size of the loan, LTV, PTI and loan duration). Of course we realize that this will strengthen the omitted variables bias but this test is crucial for understanding the existence and severity of the endogeneity bias in our estimates.

Table 6 presents estimation results of the regression (4) without endogenous variables next to the regression (3) which is our basic model. As can be seen, all estimation results are very close and even the R^2 of the regressions are similar. Specifically, the magnitudes of the estimates of interaction coefficients are very close in these two regressions. All of them remain positive and statistically significant at the same level

(besides the coefficient of *DmidSEhigh,* which is statistically significant at the 5 percent level in regression (4)). Moreover, even the ranking of the coefficients (according to distances and socioeconomic levels) remains the same as in the basic regression. We can conclude that even if some endogeneity bias exists, it does not distort our main result that mortgage pricing depends on the asset location.

7.2 Inclusion of mortgaged asset price

It is possible that inter-regional interest rate differentials reflect the risk premium charged by a competitive market for the greater uncertainties associated with lending to borrowers purchasing assets in peripheral and poor neighborhoods. Under the assumption that the market prices such risk, we would like to include, say, rates of return on housing assets (calculated as rent divided by price of the asset), but we do not have such data—neither for neighborhoods nor for cities. It may be argued that the price of the asset mainly reflects the socioeconomic status of the borrower. We find that correlation between the income of the household and the price of the asset is 0.46 for first-time home buyers and for home upgraders (who intend to live in the asset) and 0.32 for investors. Although these correlations are relatively high, they definitely leave some room for other factors to be incorporated in the price of an asset.

We run regression model (3) with the addition of the *log of price* variable (regression model 5) and rerun it without the *loan amount* variable, because of the relation between these two variables (regression model 6). Table 7 presents the estimation results, compared to regression (3). The influence of *log of price* variable is statistically significant and has the expected sign, indicating that higher-priced assets are regarded as less risky, and therefore mortgages originated to finance purchasing such assets are cheaper. Inclusion of the variable does not influence most of the estimation results, and the explanatory power of the regression does not really increase. However, the coefficients of the interactions of distance and socioeconomic status decrease in magnitude (except one, *DmidSEhigh*). In other words, these interaction terms are supposed to incorporate some share of housing asset risk that is not expressed in the rest of the variables. However, the distance-socioeconomic status differentials remain.

7.3 Restriction on the share of "prime"-adjusted interest rate

In May 2011, the Banking Supervision Department at the Bank of Israel imposed a restriction on the share of variable interest rate indexed to the BOI interest rate ("prime"[25]) and set this share to one third of the mortgage at the most. "Prime" adjustment was quite popular before the restriction became effective (Figure 3). The "Prime"-adjusted rate is in fact the lowest rate of interest that a borrower can get (it is also not CPI indexed) since it is thought of as the riskiest one, especially when the BOI interest rate is on a rising path, because it can change every month. Therefore, we divide all observations into two sub-samples, according to the incidence of the restriction.[26]

We can hypothesize that, since the choice of a larger share of "prime"-adjusted rate demands better awareness of the mortgage market, more intensive shopping and perhaps stronger negotiation ability, we expect that in the period before the restriction these borrower's qualities could be beneficial to attain the most convenient mortgage setting. Table 8 confirms this hypothesis. In the pre-restriction period (regression model 7), the effect of distance-socioeconomic status interactions dissolve for only high socioeconomic status groups, which is also consistent with the relatively high share of "prime"-adjusted rate in the mortgage within these groups (Figure 3). Furthermore, the influence of banking competition after the restriction (regression model 8) weakened both for the locality of mortgage origination and of the mortgaged asset. The effect of the *bank in the same location* variable weakened, meaning a decrease in the return to shopping behavior after the restriction. In contrast, the influence of high LTV and PTI ratios and local housing market situation (*turnover* variable) strengthened after the restriction. We can also mention differences in the distance from the center and the socioeconomic status of neighborhood influence before and after the restriction, with

[25] "Prime" is the Bank of Israel interest rate + 1.5 percentage points. Banks lend mortgages indexed to this rate, usually with negative increment.

[26] We also included the share of "prime"-adjusted interest rate into our basic regressions (see Table A1 in the Appendix), but we do not think that its inclusion has economically reasonable results. Regression results show that it decreases the average interest rate both statistically and economically significant. However, while in the period before the restriction we can hypothesize that the share of "prime"-adjusted rate is correlated with borrower's financial sophistication, it will not be true in the period after the restriction. We also claim that the large effect of inclusion of the share of "prime"-adjusted interest rate into the regression is generally technical and stems mainly from its large magnitude. In the before-restriction period, 63.5 percent of mortgages included one-third and more "prime"-adjusted interest rate share; after the restriction, 57 percent of mortgages included at least one-third "prime"-adjusted interest rate share.

the former influencing in the pre-restriction period and the latter influencing in the post-restriction one.

We also find the evidence of substantial weakening of the role of net income, being a home upgrader or investor (all of them signaling better negotiation ability) in the interest rate determination in the post-restriction period, as the composition of mortgage interest rates became more uniform and less risky.

7.4 Propensity-score matching approach application

While the OLS approach is completely valid, there are a few concerns about its implementation, the major one being that by using OLS, we make strong assumptions about normality or the linear relationship between the covariates of interest. In contrast, equivalent non-parametric statistical methods make no assumptions about the population distribution from which the data are sampled. In addition, the OLS approach allows for extreme outliers in the estimation, which can bias the interest rate estimates substantially.

Therefore, we use the Propensity Score-Matching (PSM) estimation strategy. The PSM is less parametric and more closely related to the notion of a randomized estimation[27] that deals with the self-selectivity problem that may bias the estimates of interest rate gaps. The PSM was developed as part of the selection on observables approach (Rubin, 1973; Rosenbaum and Rubin, 1983; and Heckman, et al. 1998). The propensity score is the probability of treatment assignment conditional on observed baseline characteristics. The propensity score is a balancing score: conditional on the propensity score, the distribution of observed baseline covariates will be similar between treated and untreated subjects.

In the following set of tests, we use the PSM method, and examine the difference in the average mortgage interest rates between three types of borrowers: those who purchase assets within 40 kilometers from the center of Tel Aviv, those who purchase assets within the distance of 40 to 80 kilometers from the center of Tel Aviv and those who purchase remote assets situated 80 kilometers and more from the center of Tel Aviv.

[27] See Angrist and Lang (2004) for a review.

The matching procedure uses a logistic model to predict each borrower propensity score using covariates *Number of borrowers, Age, Age squared, Log of net income, Wage account, Upgrader, Investor, Guarantor, Log of loan amount, LTV*[28]*, PTI30, Dur20, Socioeconomic* and *Turnover* as well as bank identity and date (month and year).

The results are presented in Table 9. On average, borrowers purchasing assets within 40 to 80 kilometers from the center of Tel Aviv pay mortgage interest rates that are higher by 0.1 percentage points than similar borrowers purchasing assets within 40 kilometers from the center of Tel Aviv, while borrowers purchasing assets situated 80 kilometers and more from the center of Tel Aviv pay mortgage interest rates that are 0.2 percentage points higher than those of similar borrowers purchasing assets within 40 kilometers from the center of Tel Aviv. There is also a statistically significant gap of approximately 0.1 percentage points between two "peripheral" groups of borrowers. The magnitude of these gaps is in line with our OLS estimates.

8. Discussion

In this paper, we explore the contribution of various risk factors to the mortgage interest rate determination, paying particular attention to location-based differentials, including two dimensions—distance from the center and socioeconomic status of the neighborhood where the mortgaged asset is situated. Empirical evidence based on more than 80,000 mortgage loans originated during 2010–13 indicates that location does matter. It appears that borrowers purchasing housing assets in the prosperous central regions are perceived by lenders as preferred customers, receiving the best interest rate terms, while borrowers purchasing assets in the poorer peripheral neighborhoods are compelled to pay the highest interest rates. The ranking remains unchanged and statistically significant after controlling for various factors of borrower, mortgage and asset risk as well as for the extent of banking competition.

However, we cannot attribute these location-based interest rate differentials to discrimination against poorly situated borrowers. Some crucial factors of interest rate determination that are unobservable are likely to be correlated to some degree with the location of the purchased asset, including borrower's credit history, wealth, employment

[28] The exact value of the LTV ratio.

characteristics (in terms of occupation, seniority, tenure, stability, contract duration), financial literacy and bargaining ability. For example, Haran Rosen and Sade (2018) find that individuals living in central locations with a higher socioeconomic index demonstrated more active beneficial financial behavior.

Canner (1981), Stiglitz and Weiss (1981) and Williamson (1986, 1987) argued that creditors may apply binding credit constraints to loan applicants due either to their ability to repay a loan or because of factors that may adversely affect the collateral value of the property. In other words, lenders should apply tighter credit conditions (including higher down payment requirements, shorter terms to loan maturity, and higher interest rates) to riskier loan applicants, irrespective of whether that risk is related to the attributes of the borrower or to those of the neighborhood where the property is located. It is reasonable that lenders incorporate the risk of asset foreclosure in the interest rates, while the costs of foreclosure seem to be higher in the regions with lower housing demand (at least partly captured by *turnover* variable) and lower prospective house price growth (or higher possibility of price depreciation). Since in Israel the most demanded residential regions are central ones, where land reserves for residential building are limited, the prospects of price growth there are much more promising than in the periphery where abundance of land suitable for residential building keeps its price low.

Even when discrimination occurs, it is not feasible to test whether it is prejudice-driven or statistical. However, we assume that discrimination in mortgage pricing, at least against those purchasing assets in the peripheral regions (but not necessarily those buying assets in poor but close to the center neighborhoods), is unlikely to be a *prejudiced discrimination*. Peripheral borrowers approaching local banks' affiliations meet loan officers who are also local residents, such that we should not expect them to have personal prejudices against their neighbors. Several studies examined group identity effect on the credit market outcomes. For example, Beck, et al. (2012) tested the influence of shared gender identity and showed that in Albania, borrowers assigned to opposite-sex officers received lower loan amounts and paid higher interest rates, although, ex post, they did not experience higher arrears. Fishman, et al. (2017) report that shared ethnicity and religion between borrowers and loan officers in India increased access to credit and loan size dispersion and reduced collateral requirements, while improving future repayment.

Statistical discrimination occurs when individual members of a particular group are treated differently based on the use of empirical (statistical) correlations of this group's distinctive observable characteristics with its economic performance or outcomes. In the context of our study, statistical discrimination would occur if peripheral residents *as a group* have a higher average statistical risk of default or of being in arrears (say, because of less favorable conditions in the local labor markets), and a lender uses this past experience to charge a higher interest rate to all loan applicants living in the periphery, independent of their individual characteristics, while *particular applicants* belonging to this group may or may not cause higher risk. Since lenders have imperfect information about potential customers, they only use available statistical data as a type of screening device to assess a risk premium for a potential borrower based on the assumption that his or her group affiliation is correlated with socioeconomically relevant characteristics and likelihood of repayment/default (or being in arrears). As a result, an individual borrower could be adversely affected because of his group affiliation.

Unfortunately, we have no information on default rates by locality of residence or long-run statistics on regional distribution of loans in arrears. However, since our data was collected retroactively and not at the mortgage origination we know the status of the mortgages originated in 2010–13 (regularly repaid or in arrears) in the year 2015. The share of the mortgages in arrears is generally low, but there are some differences among distance-socioeconomic status groups (Table 10), suggesting that there is some economic rationality for differential treatment of certain groups of borrowers.

Given large loan amount and long duration, even a relatively small interest rate increment may cause a substantial increase in the total repayments over the mortgage life span. Our findings indicate that interest rate differentials hurt mostly the weakest borrowers; higher mortgage prices increase the economic burden on borrowers purchasing housing assets in inferior locations, possibly raising their probability of default, and even contributing to exacerbating inequality at the economy-wide level.

One of the causes of getting less convenient interest rates is inadequate financial literacy and shopping behavior. In recent years, provision of private mortgage counselling services has become increasingly common and perhaps it is successful in obtaining

more favorable mortgage terms for those lacking financial knowledge and bargaining ability. However, we cannot examine this hypothesis.

Commercial banks are profit maximizing firms. Their lending decision is based on their prediction of the probability that a borrower will repay the loan successfully. However, defaults generally occur due to some unexpected adverse life event. To minimize expected mortgage loss, a lending institution must, ex ante, predict whether the future value of the underlying property will exceed the outstanding debt. Naturally, properties situated in more demanded neighborhoods have better prospects of price stability and growth.[29]

Further research has to focus on enriching the list of independent variables to include dimensions of risk that are unavailable at this point of time. The Credit Data System at the Bank of Israel, which began to operate in April, 2019, will provide in the near future households' credit history data and measures of non-housing indebtedness which could be incorporated in the model of mortgage interest rate determination.

[29] For example, Haughwout, et al. (2009) find that mortgage rates are lower in locations that experienced higher past rates of house price appreciation, probably because lenders have expectations for such trend continuation.

References

Angrist, J.D. and K. Lang (2004). "Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program." *American Economic Review* 94 (5), pp. 1613-1634.

Beck, T., P. Behr and A. Madestam (2012). "Sex and Credit: Is There a Gender Bias in Lending?", European Banking Center Discussion Paper No. 2011-02J, CentER Working PaperSeries No. 2011-101.

Black, H.A., T.P. Boehm and R.P. DeGennaro (2003). "Is There Discrimination in Mortgage Pricing? The Case of Overages", *Journal of Banking and Finance*, vol. 27, pp. 1139-1165.

Calem, P.S. (1996). "Mortgage Credit Availability and Low- and Moderate-Income Minority Neighborhoods: Are Information Externalities Critical?", *The Journal of Real Estate Finance and Economics*, vol. 12 (1), pp. 71-89.

Canner, G.B. (1981). "Redlining and Mortgage Lending Patterns" in J. V. Henderson (ed.) *Research in Urban Economics*, Conn.: JAI Press Inc. pp. 67-101.

Cheng, P., Z. Lin and Y. Liu (2011). "Do Women Pay More for Mortgages?", *The Journal of Real Estate finance and Economics*, vol. 43, pp. 423-440.

Cheng, P., Z. Lin and Y. Liu(2015). "Racial Discrepancy in Mortgage Interest Rates", *The Journal of Real Estate finance and Economics*, vol. 51, pp. 101-120.

Courchane, M. (2007). "The Pricing of Home Mortgage Loans to Minority Borrowers: How Much of the APR Differential Can We Explain?", *Journal of Real Estate Research*, vol. 29 (4), pp. 399-440.

Courchane, M. and D. Nickerson (1997). "Discrimination Resulting from Overage Practices", *Journal of Financial Services Research*, vol. 11, pp. 133-152.

Crawford, G.W. and E. Rosenblatt (1999). "Differences in the Cost of Mortgage Credit Implications for Discrimination", *Journal of Real Estate Finance and Economics*, vol. 19 (2), pp. 147-159.

Diaz-Serrano, L. and J.P. Raya (2011). "Is There Discriminatory Mortgage Pricing against Immigrants in the Spanish Lending Market?", IZA Discussion Paper No. 5578.

Eichengreen, B. (1984). "Mortgage Interest Rates in the Populist Era", *The American Economic Review*, vol. 74 (5), pp. 995-1015.

Fishbein, A. and P. Woodall (2006). "Women are Prime Targets for Subprime Lending: Women are Disproportionately Represented in High-Cost Mortgage Market. Working Paper, Consumer Foundation of America.

Fishman, R., D. Paravisini and V. Vig (2017). "Cultural Proximity and Loan Outcomes", *American Economic Review*, vol. 107 (2), pp. 457-492.

Gary-Bobo, R.J. and S. Larribeau (2004). "A Structural Econometric Model of Price Discrimination in the Mortgage Lending Industry", *International Journal of Industrial Organization*, vol. 22 (1), pp. 101-134.

Ghent, A., R. Hernandez-Murillo and M.T. Owyang (2014). "Differences in Subprime Loan Pricing Across Races and Neighborhoods, *Regional Science and Urban Economics*, vol. 48, pp. 199-215.

Haran Rosen, M. and O. Sade (2018). "Does Financial Regulation Unintentionally Ignore Less Privileged Populations? The Investigation of a Regulatory Fintech Advancement, Objective and Subjective Financial Literacy", Discussion Paper No. 2017.10, Bank of Israel, Research Department.

Haughwout, A., C. Mayer and J. Tracy (2009). "Subprime Mortgage Pricing: The Impact of Race, Ethnicity, and Gender on the Cost of Borrowing", Federal Reserve Bank of New York Staff Reports, Staff Report no. 368.

Heckman, J., H. Ichimura and P. Todd (1998). "Matching as an Econometric Evaluation Estimator", *Review of Economic Studies*, 65 (2), pp. 261-294.

Holmes, A. and P. Horvitz (1994). "Mortgage Redlining: Race, Risk, and Demand", *Journal of Finance*, vol. 49 (1), pp. 81-99.

Lang, W.W. and L.I. Nakamura (1993). "A Model of Redlining", *Journal of Urban Economics*, vol. 33, pp. 371-379.

Ling, D.C. and S.M. Wachter (1998). "Information Externalities and Home Mortgage Underwriting", Journal of Urban Economics, vol. 44 (3), pp. 317-332.

Rosenbaum, P.R. and D.B. Rubin (1983). "The Central Role of the Propensity Score in Observational Studies for Causal Effects", *Biometrika*, 70 (1), pp.41-55.

Rubin, D.B. (1973). "The Use of Matched Sampling and Regression Adjustment to Remove Bias in Observational Studies", *Biometrics*, pp.185-203.

Stiglitz, J.E. and A. Weiss (1981). "Credit Rationing in Markets with Imperfect Information", *American Economic Review*, vol. 71, pp. 393-410.

Tzur-Ilan, N. (2017). "The Effect of Credit Constraints on Housing Choices: The Case of LTV Limit", Bank of Israel, Research Department, Discussion Paper 2017.03.

Williamson, S.D. (1986). "Costly Monitoring, Financial Intermediation, and Equilibrium Credit Rationing", *Journal of Monetary Economics*, pp. 159-179.

Williamson, S.D. (1987). "Costly Monitoring, Loan Contracts, and Equilibrium Credit Rationing", *Quarterly Journal of Economics*, pp. 135-145.

Yezer, A.M. (2010). "A Review of Statistical Problems in the Measurement of Mortgage Market Discrimination and Credit Risk", Research Institute for Housing America.

Figure 1

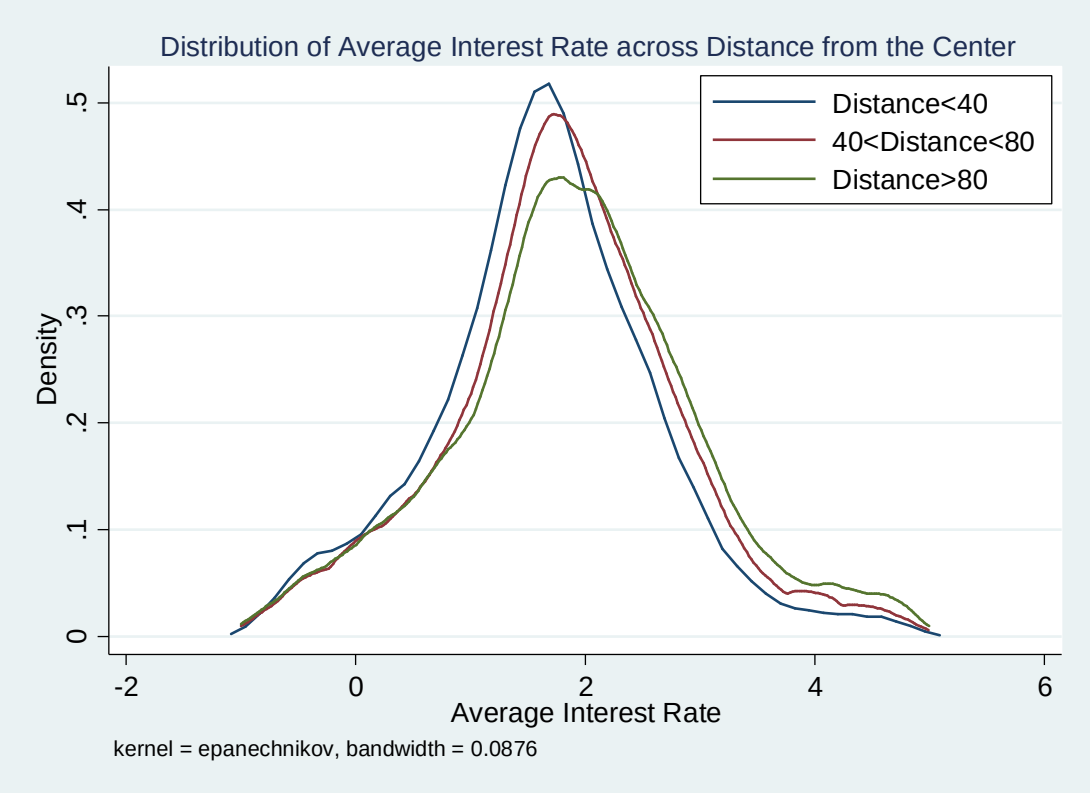
Distribution of Average Interest Rate across Distance from the Center
Distance<40
40<Distance<80
Distance>80
Density
0
.1
.2
.3
.4
.5
-2
0
2
4
6
Average Interest Rate
kernel = epanechnikov, bandwidth = 0.0876

Figure 2

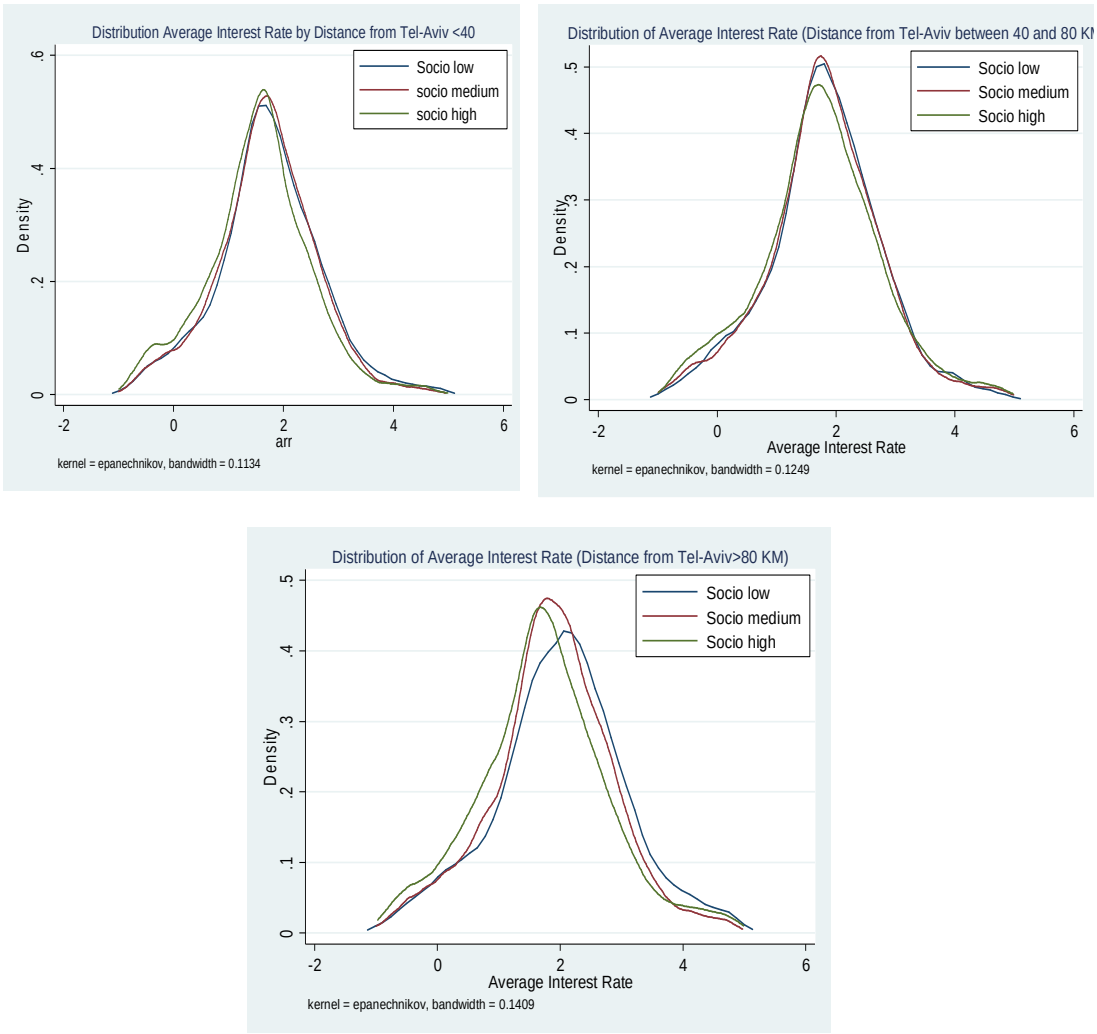
Distribution Average Interest Rate by Distance from Tel-Aviv <40
Socio low
socio medium
socio high
Density
arr
kernel = epanechnikov, bandwidth = 0.1134
Distribution of Average Interest Rate (Distance from Tel-Aviv between 40 and 80 KM
Socio low
Socio medium
Socio high
Density
Average Interest Rate
kernel = epanechnikov, bandwidth = 0.1249
Distribution of Average Interest Rate (Distance from Tel-Aviv>80 KM)
Socio low
Socio medium
Socio high
Density
Average Interest Rate
kernel = epanechnikov, bandwidth = 0.1409

Figure 3. Mean Share of "Prime"-Adjusted Interest Rate Before and After the Restriction

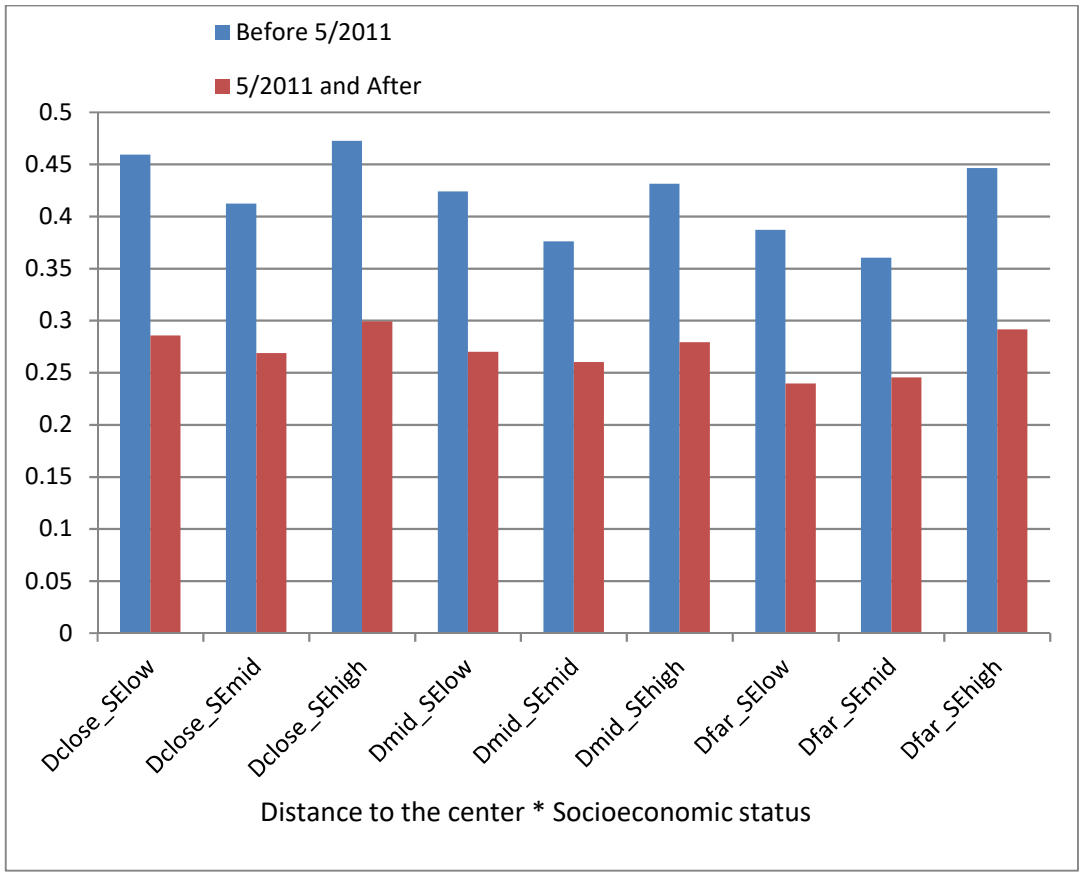

Table 1. Median Values of Main Mortgage and Borrower Characteristics, Broken out by Combinations of Distance and Socioeconomic Status of Neighborhood

	Distance<40			40<=Distance<80			Distance>=80		
socioeconomic	**low**	**medium**	**high**	**low**	**medium**	**high**	**low**	**medium**	**high**
interest rate (%)	**1.72**	**1.70**	**1.57**	**1.83**	**1.80**	**1.75**	**2.04**	**1.89**	**1.71**
LTV (%)	58.3	59.0	51.9	60.0	59.8	55.0	60.0	60.9	57.0
duration (months)	260	274	251	253	264	245	240	240	240
PTI (%)	27.9	27.5	28.0	26.0	27.0	27.0	23.1	24.4	24.5
loan size (000' NIS)	480	582	700	417	500	530	260	370	450
net monthly income (NIS)	12,000	13,200	16,500	11,800	12,706	14,805	11,900	12,380	14,900
age	36.3	38.5	40.1	35.2	38.6	40.3	38.6	39.4	41.0
LTI	3.4	3.7	3.5	3.2	3.3	3.1	2.0	2.7	2.6
investors' share (%)	15	13	15	15	13	14	29	17	16
investors' interest rate (%)	1.61	1.55	1.50	1.76	1.79	1.73	1.93	1.90	1.68
first-time home buyers' interest rate (%)	1.90	1.85	1.69	1.96	1.91	1.92	2.26	2.00	1.88
mortgage outside asset location	55.7	51.9	58.7	57.6	43.8	68.4	50.2	37.5	46.1
mortgage outside asset location for investors	59.7	55.5	59.1	60.0	51.0	62.5	59.2	49.6	47.6
mortgage outside asset location for first-time home buyers	56.4	52.9	62.3	59.5	45.2	73.4	49.5	38.7	47.3
mortgage outside asset location for upgraders	53.2	49.6	55.7	54.0	40.2	66.0	42.2	31.5	44.7
mortgage outside asset location in localities 100,000+ residents (%)	45.5	41.2	48.6	27.9	23.2	26.4	36.7	22.1	16.4
number of observations	11,833	13,366	20,608	7,507	7,984	5,380	8,201	9,022	5,013

Table 2. Variable Names, Definitions and Expected Relation to Mortgage Interest Rate

Variable	Description	Expected sign
R	Dependent variable; weighted average of real interest rates on all parts of the mortgage	
Borrower risk		
Number of borrowers	1 for a single borrower, 2 for a couple	-
Age	For multiple borrowers – average age	-
Age squared	For multiple borrowers – average age squared	?
Log of net income	Log of family after-tax monthly income net of fixed monthly payments	-
Wage account	Dummy, 1 for household with current account within the bank, 0 otherwise	-
Upgrader	Dummy, 1 for upgrader, 0 otherwise	-
Investor	Dummy, 1 for investor, 0 otherwise	?
Guarantor	Dummy, 1 for loan with guarantor requirement, 0 otherwise	+
Loan risk		
Log of loan amount	Log of approved mortgage	-
LTV60	Dummy, 1 for mortgages with LTV>60%	+
PTI30	Dummy, 1 for mortgages with PTI>30%	+
Dur20	Dummy, 1 for mortgages with maturity of more than 20 years	+
Collateral risk		
Socioeconomic	Socioeconomic index of neighborhood, 1 (the lowest) to 20 (the highest)	-
Distance	Distance from the neighborhood to the center of Tel Aviv, in km	+
Distance squared	Square of distance from the neighborhood to the center of Tel Aviv	?
Dclose_SElow, Dclose_SEmid, Dclose_SEhigh, Dmid_SElow, Dmid_SEmid, Dmid_SEhigh, Dfar_SElow, Dfar_SEmid, Dfar_SEhigh	System of dummies for interactions of distance and socioeconomic status as described in Section 4. Dclose_SEhigh variable is omitted in the regression analysis.	all +
Turnover, %	Number of transactions in the housing market divided by the number of existing housing units, by municipality[1]	-
Potential Accessibility Index	Continuous variable, by municipality, see footnote 21 in the text	-
Building starts, %	Number of housing units which construction had begun in the given year divided by the number of existing housing units, by municipality	+
Competitiveness		
Bank in the same location	Dummy, 1 if the loan was originated in the locality of purchased property, 0 otherwise	?
Number of banks in mortgage location	Number of different banking institutions providing mortgage services in the locality of loan origination[2]	-
Number of banks in property location	Number of different banking institutions providing mortgage services in the locality of purchased property[3]	-

Notes for Table 2

[1] To estimate turnover variable, we divide the number of transactions in each locality (not neighborhood) by the number of housing units in each municipality, officially reported by the Israeli Central Bureau of Statistics. For mortgage taken out in year *t* we use an average of turnover in years *t* and *t-1*. Since number of housing units in some small localities is not reported we lose 4.6% of observations due to use of the variable.

[2] Using the full list of bank affiliations, we count the number of banking institutions in the locality where the mortgage was originated. We also tried the number of different bank brances engaged in mortgage credit and the results of estimation were similar. Because of distortions in the data we were unable to identify correct location of mortgage origination in 4.1% of observations.

[3] The same as the previous variable, but for the locality where the mortgaged asset is situated.

Table 3. Basic Regression Estimation Results

Variable	(1)		(2)		(3)	
number of borrowers	0.050***	(0.009)	0.049***	(0.009)	0.046***	(0.009)
age	0.030***	(0.002)	0.030***	(0.002)	0.029***	(0.002)
age squared	-0.000***	(0.000)	-0.000***	(0.000)	-0.000***	(0.000)
log of net income	-0.223***	(0.008)	-0.225***	(0.009)	-0.225***	(0.009)
wage account	-0.126***	(0.008)	-0.125***	(0.008)	-0.126***	(0.008)
upgrader	-0.030***	(0.008)	-0.028***	(0.008)	-0.030***	(0.008)
investor	-0.075***	(0.010)	-0.072***	(0.010)	-0.069***	(0.010)
guarantor	0.059***	(0.012)	0.057***	(0.012)	0.058***	(0.012)
log of loan amount	-0.078***	(0.006)	-0.078***	(0.006)	-0.075***	(0.006)
LTV60	0.057***	(0.007)	0.058***	(0.007)	0.059***	(0.007)
PTI30	0.002	(0.007)	0.002	(0.007)	0.003	(0.007)
Dur20	0.221***	(0.008)	0.220***	(0.008)	0.219***	(0.008)
Dclose_SElow	**0.070*** **	(0.012)	**0.074*** **	(0.013)	**0.075*** **	(0.013)
Dclose_SEmid	**0.053*** **	(0.011)	**0.050*** **	(0.011)	**0.056*** **	(0.011)
Dmid_SElow	**0.068*** **	(0.018)	**0.063*** **	(0.019)	**0.089*** **	(0.019)
Dmid_SEmid	**0.080*** **	(0.017)	**0.081*** **	(0.017)	**0.109*** **	(0.017)
Dmid_SEhigh	**0.047** **	(0.020)	**0.041** **	(0.020)	**0.054*** **	(0.020)
Dfar_SElow	**0.179*** **	(0.025)	**0.200*** **	(0.026)	**0.248*** **	(0.026)
Dfar_SEmid	**0.105*** **	(0.025)	**0.123*** **	(0.025)	**0.175*** **	(0.026)
Dfar_SEhigh	**0.067*** **	(0.025)	**0.075*** **	(0.026)	**0.125*** **	(0.026)
distance	0.000	(0.000)	0.002***	(0.001)	0.001*	(0.001)
distance squared	0.000**	(0.000)	-0.000	(0.000)	0.000	(0.000)
socioeconomic	-0.005***	(0.001)	-0.004***	(0.001)	-0.003***	(0.001)
turnover	-0.008***	(0.002)	-0.011***	(0.002)	-0.007***	(0.002)
bank in the same location	0.044***	(0.006)	0.049***	(0.007)	0.076***	(0.007)
number of banks in mortgage location	-0.016***	(0.002)	-0.016***	(0.002)	-0.009***	(0.002)
potential accessibility index			0.001***	(0.000)	0.001***	(0.000)
building starts			0.007***	(0.002)	-0.002	(0.002)
number of banks in property location					-0.016***	(0.002)
Banks fixed effects	+		+		+	
Month & Year fixed effects	+		+		+	
Constant	3.580***	(0.101)	3.336***	(0.121)	3.343***	(0.120)
Observations	81,143		80,539		80,539	
R-squared	0.282		0.282		0.283	

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 4. Inter-Regional Interest Differentials

Panel A. Unconditional means, standard deviations in parentheses

	distance<40	<=40distance<80	distance>80
socio high	1.536	1.741	1.719
	(1.022)	(1.121)	(1.168)
socio middle	1.684	1.800	1.866
	(0.961)	(1.074)	(1.091)
socio low	1.730	1.807	2.016
	(1.031)	(1.041)	(1.213)

Panel B. Differences in unconditional means, relative to basic category

	distance<40	40<=distance<80	distance>80
socio high	base	0.205	0.183
socio middle	0.148	0.263	0.330
socio low	0.194	0.271	0.480

Panel C. Estimated coefficients (conditional differences in means)

	distance<40	40<=distance<80	distance>80
socio high	omitted	0.054	0.125
socio middle	0.056	0.109	0.175
socio low	0.075	0.089	0.248

Table 5. Correlations Between Suspected Endogenous Variables

	LTV	PTI	Duration	Loan size	Interest rate
LTV	1.0000				
PTI	0.0893	1.0000			
Duration	0.4662	0.0387	1.0000		
Loan size	0.3947	0.1789	0.3963	1.0000	
Interest rate	0.0393	0.0247	0.1282	-0.0556	1.0000

Table 6. Robustness Check 1: Endogeneity Issue Test

Variable	(3)		(4)	
number of borrowers	0.046***	(0.009)	0.057***	(0.009)
age	0.029***	(0.002)	0.033***	(0.002)
age squared	-0.000***	(0.000)	-0.000***	(0.000)
log of net income	-0.225***	(0.009)	-0.255***	(0.008)
wage account	-0.126***	(0.008)	-0.122***	(0.008)
upgrader	-0.030***	(0.008)	-0.041***	(0.008)
investor	-0.069***	(0.010)	-0.087***	(0.010)
guarantor	0.058***	(0.012)	0.053***	(0.012)
log of loan amount	-0.075***	(0.006)	-	
LTV60	0.059***	(0.007)	-	
PTI30	0.003	(0.007)	-	
Dur20	0.219***	(0.008)	-	
Dclose_SElow	**0.075*** **	(0.013)	**0.091*** **	(0.013)
Dclose_SEmid	**0.056*** **	(0.011)	**0.075*** **	(0.011)
Dmid_SElow	**0.089*** **	(0.019)	**0.092*** **	(0.019)
Dmid_SEmid	**0.109*** **	(0.017)	**0.118*** **	(0.017)
Dmid_SEhigh	**0.054*** **	(0.020)	**0.048** **	(0.020)
Dfar_SElow	**0.248*** **	(0.026)	**0.246*** **	(0.026)
Dfar_SEmid	**0.175*** **	(0.026)	**0.175*** **	(0.026)
Dfar_SEhigh	**0.125*** **	(0.026)	**0.110*** **	(0.026)
distance	0.001*	(0.001)	0.001**	(0.001)
distance squared	0.000	(0.000)	-0.000	(0.000)
socioeconomic	-0.003***	(0.001)	-0.003***	(0.001)
turnover	-0.007***	(0.002)	-0.006***	(0.002)
bank in the same location	0.076***	(0.007)	0.077***	(0.007)
number of banks in mortgage location	-0.009***	(0.002)	-0.010***	(0.002)
potential accessibility index	0.001***	(0.000)	0.001***	(0.000)
building starts	-0.002	(0.002)	-0.002	(0.002)
number of banks in property location	-0.016***	(0.002)	-0.017***	(0.002)
Banks fixed effects	+		+	
Month & Year fixed effects	+		+	
Constant	3.343***	(0.120)	2.749***	(0.110)
Observations	80,539		80,539	
R-squared	0.283		0.273	
Standard errors in parentheses				
*** p<0.01, ** p<0.05, * p<0.1				

Table 7. Robustness Check 2: Inclusion of *log of price* Variable

Variable	basic regression (3)		(5)		(6)	
number of borrowers	0.046***	(0.009)	0.051***	(0.009)	0.051***	(0.009)
age	0.029***	(0.002)	0.029***	(0.002)	0.029***	(0.002)
age squared	-0.000***	(0.000)	-0.000***	(0.000)	-0.000***	(0.000)
log of net income	-0.225***	(0.009)	-0.202***	(0.009)	-0.218***	(0.008)
wage account	-0.126***	(0.008)	-0.126***	(0.008)	-0.124***	(0.008)
upgrader	-0.030***	(0.008)	-0.012	(0.008)	-0.012	(0.008)
investor	-0.069***	(0.010)	-0.085***	(0.010)	-0.081***	(0.010)
guarantor	0.058***	(0.012)	0.060***	(0.012)	0.054***	(0.012)
log of loan amount	-0.075***	(0.006)	-0.047***	(0.006)		
LTV60	0.059***	(0.007)	0.038***	(0.008)	0.024***	(0.007)
PTI30	0.003	(0.007)	0.011	(0.007)	0.003	(0.007)
Dur20	0.219***	(0.008)	0.220***	(0.008)	0.204***	(0.007)
Dclose_SElow	**0.075*** **	(0.013)	**0.051*** **	(0.013)	**0.054*** **	(0.013)
Dclose_SEmid	**0.056*** **	(0.011)	**0.037*** **	(0.011)	**0.038*** **	(0.011)
Dmid_SElow	**0.089*** **	(0.019)	**0.067*** **	(0.019)	**0.073*** **	(0.019)
Dmid_SEmid	**0.109*** **	(0.017)	**0.090*** **	(0.017)	**0.096*** **	(0.017)
Dmid_SEhigh	**0.054*** **	(0.020)	**0.056*** **	(0.020)	**0.060*** **	(0.020)
Dfar_SElow	**0.248*** **	(0.026)	**0.189*** **	(0.026)	**0.197*** **	(0.026)
Dfar_SEmid	**0.175*** **	(0.026)	**0.143*** **	(0.026)	**0.150*** **	(0.026)
Dfar_SEhigh	**0.125*** **	(0.026)	**0.110*** **	(0.026)	**0.115*** **	(0.026)
distance	0.001*	(0.001)	0.001	(0.001)	0.001	(0.001)
distance squared	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)
socioeconomic	-0.003***	(0.001)	-0.002***	(0.001)	-0.002***	(0.001)
turnover	-0.007***	(0.002)	-0.007***	(0.002)	-0.007***	(0.002)
bank in the same location	0.076***	(0.007)	0.080***	(0.007)	0.081***	(0.007)
number of banks in mortgage location	-0.009***	(0.002)	-0.009***	(0.002)	-0.009***	(0.002)
potential accessibility index	0.001***	(0.000)	0.001***	(0.000)	0.001***	(0.000)
building starts	-0.002	(0.002)	-0.001	(0.002)	-0.001	(0.002)
number of banks in property location	-0.016***	(0.002)	-0.016***	(0.002)	-0.016***	(0.002)
log of price			-0.107***	(0.007)	-0.122***	(0.006)
Banks fixed effects	+		+		+	
Month & Year fixed effects	+		+		+	
Constant	3.343***	(0.120)	3.505***	(0.121)	3.159***	(0.112)
Observations	80,539		80,539		80,539	
R-squared	0.283		0.285		0.284	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 8. Robustness Check 3: Estimation Results Before and After the Restriction on "Prime"-Adjusted Interest Rate Share

Variable	basic regression (3)		before restriction (7)		after restriction (8)	
number of borrowers	0.046***	(0.009)	0.069***	(0.018)	0.036***	(0.010)
age	0.029***	(0.002)	0.054***	(0.005)	0.018***	(0.003)
age squared	-0.000***	(0.000)	-0.000***	(0.000)	-0.000***	(0.000)
log of net income	-0.225***	(0.009)	-0.365***	(0.017)	-0.165***	(0.010)
wage account	-0.126***	(0.008)	-0.113***	(0.015)	-0.095***	(0.009)
upgrader	-0.030***	(0.008)	-0.065***	(0.016)	-0.015*	(0.009)
investor	-0.069***	(0.010)	-0.196***	(0.020)	0.001	(0.012)
guarantor	0.058***	(0.012)	0.068***	(0.025)	0.053***	(0.014)
log of loan amount	-0.075***	(0.006)	0.019	(0.012)	-0.105***	(0.006)
LTV60	0.059***	(0.007)	0.056***	(0.014)	0.072***	(0.008)
PTI30	0.003	(0.007)	-0.055***	(0.014)	0.026***	(0.008)
Dur20	0.219***	(0.008)	0.233***	(0.015)	0.208***	(0.008)
Dclose_SElow	**0.075*** **	(0.013)	**0.100*** **	(0.026)	**0.069*** **	(0.014)
Dclose_SEmid	**0.056*** **	(0.011)	**0.103*** **	(0.021)	**0.029** **	(0.012)
Dmid_SElow	**0.089*** **	(0.019)	**0.077** **	(0.039)	**0.102*** **	(0.021)
Dmid_SEmid	**0.109*** **	(0.017)	**0.172*** **	(0.035)	**0.083*** **	(0.019)
Dmid_SEhigh	**0.054*** **	(0.020)	**-0.037**	(0.041)	**0.095*** **	(0.023)
Dfar_SElow	**0.248*** **	(0.026)	**0.265*** **	(0.054)	**0.234*** **	(0.029)
Dfar_SEmid	**0.175*** **	(0.026)	**0.182*** **	(0.052)	**0.169*** **	(0.028)
Dfar_SEhigh	**0.125*** **	(0.026)	**0.058**	(0.053)	**0.150*** **	(0.029)
distance	0.001*	(0.001)	0.003**	(0.002)	0.000	(0.001)
distance squared	0.000	(0.000)	-0.000	(0.000)	0.000	(0.000)
socioeconomic	-0.003***	(0.001)	0.001	(0.002)	-0.005***	(0.001)
turnover	-0.007***	(0.002)	0.001	(0.003)	-0.013***	(0.003)
bank in the same location	0.076***	(0.007)	0.114***	(0.014)	0.062***	(0.008)
number of banks in mortgage location	-0.009***	(0.002)	-0.020***	(0.004)	-0.005***	(0.002)
potential accessibility index	0.001***	(0.000)	0.001**	(0.001)	0.001***	(0.000)
building starts	-0.002	(0.002)	-0.006	(0.005)	0.000	(0.003)
number of banks in property location	-0.016***	(0.002)	-0.021***	(0.003)	-0.014***	(0.002)
Banks fixed effects	+		+		+	
Month & Year fixed effects	+		+		+	
Constant	3.343***	(0.120)	2.699***	(0.248)	3.939***	(0.132)
Observations	80,539		25,303		55,236	
R-squared	0.283		0.170		0.241	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 9. Robustness Check 4: Propensity-Score Matching Estimation Results

Comparison categories	Coefficient	Std. Err.	z	P>|z|	Number of obs.
40<=Distance<80 vs Distance<40	0.103	0.010	10.18	0.000	66,678
Distance>=80 vs Distance<40	0.202	0.012	16.42	0.000	68,043
Distance>=80 vs 40<=Distance<80	0.098	0.013	7.29	0.000	43,107

Table 10. Percentage of Mortgages in Arrears, by Distance and Socioeconomic Status, for Mortgages Originated in 2010–13 (%)

Distance<40			40<=Distance<80			Distance>=80		
SE low	SE middle	SE high	SE low	SE middle	SE high	SE low	SE middle	SE high
2.34	1.55	1.30	3.13	2.61	2.12	2.65	1.66	1.52

Appendix

Table 1A. Estimation results before and after the restriction on "prime"-adjusted interest rate share

Variable	before restriction (7)		after restriction (8)		before restriction (9)		after restriction (10)	
number of borrowers	0.069***	(0.018)	0.036***	(0.010)	0.057***	(0.012)	0.037***	(0.009)
age	0.054***	(0.005)	0.018***	(0.003)	0.015***	(0.003)	0.007***	(0.002)
age squared	-0.000***	(0.000)	-0.000***	(0.000)	-0.000**	(0.000)	-0.000	(0.000)
log of net income	-0.365***	(0.017)	-0.165***	(0.010)	-0.071***	(0.012)	0.010	(0.009)
wage account	-0.113***	(0.015)	-0.095***	(0.009)	-0.026***	(0.010)	-0.061***	(0.008)
upgrader	-0.065***	(0.016)	-0.015*	(0.009)	-0.039***	(0.011)	-0.002	(0.007)
investor	-0.196***	(0.020)	0.001	(0.012)	-0.090***	(0.014)	-0.027***	(0.010)
guarantor	0.068***	(0.025)	0.053***	(0.014)	0.057***	(0.017)	0.023*	(0.012)
log of loan amount	0.019	(0.012)	-0.105***	(0.006)	-0.139***	(0.008)	-0.285***	(0.006)
LTV60	0.056***	(0.014)	0.072***	(0.008)	0.049***	(0.010)	0.075***	(0.007)
PTI30	-0.055***	(0.014)	0.026***	(0.008)	0.046***	(0.010)	0.080***	(0.007)
Dur20	0.233***	(0.015)	0.208***	(0.008)	0.133***	(0.010)	0.248***	(0.007)
Dclose_SElow	**0.100*** **	(0.026)	**0.069*** **	(0.014)	**0.063*** **	(0.018)	**0.023***	(0.012)
Dclose_SEmid	**0.103*** **	(0.021)	**0.029** **	(0.012)	**0.023**	(0.014)	**-0.008**	(0.011)
Dmid_SElow	**0.077** **	(0.039)	**0.102*** **	(0.021)	**0.087*** **	(0.027)	**0.062*** **	(0.018)
Dmid_SEmid	**0.172*** **	(0.035)	**0.083*** **	(0.019)	**0.093*** **	(0.024)	**0.036** **	(0.017)
Dmid_SEhigh	**-0.037**	(0.041)	**0.095*** **	(0.023)	**0.014**	(0.028)	**0.090*** **	(0.020)
Dfar_SElow	**0.265*** **	(0.054)	**0.234*** **	(0.029)	**0.209*** **	(0.036)	**0.116*** **	(0.025)
Dfar_SEmid	**0.182*** **	(0.052)	**0.169*** **	(0.028)	**0.101*** **	(0.035)	**0.086*** **	(0.025)
Dfar_SEhigh	**0.058**	(0.053)	**0.150*** **	(0.029)	**0.100*** **	(0.036)	**0.128*** **	(0.025)
distance	0.003**	(0.002)	0.000	(0.001)	-0.001	(0.001)	-0.001	(0.001)
distance squared	-0.000	(0.000)	0.000	(0.000)	0.000**	(0.000)	0.000***	(0.000)
socioeconomic	0.001	(0.002)	-0.005***	(0.001)	-0.003***	(0.001)	-0.005***	(0.001)
turnover	0.001	(0.003)	-0.013***	(0.003)	-0.008***	(0.002)	-0.020***	(0.002)
bank in the same location	0.114***	(0.014)	0.062***	(0.008)	0.046***	(0.010)	0.044***	(0.007)
number of banks in mortgage location	-0.020***	(0.004)	-0.005***	(0.002)	-0.002	(0.002)	0.002	(0.002)
potential accessibility index	0.001**	(0.001)	0.001***	(0.000)	-0.000	(0.000)	0.001**	(0.000)
building starts	-0.006	(0.005)	0.000	(0.003)	-0.004	(0.003)	0.003	(0.002)
number of banks in property location	-0.021***	(0.003)	-0.014***	(0.002)	-0.012***	(0.002)	-0.009***	(0.002)
share of "prime"					-2.372***	(0.014)	-1.794***	(0.014)
Banks fixed effects	+		+		+		+	
Month & Year fixed effects	+		+		+		+	
Constant	2.699***	(0.248)	3.939***	(0.132)	4.362***	(0.168)	5.510***	(0.116)
Observations	25,303		55,236		25,303		55,236	
R-squared	0.170		0.241		0.619		0.422	

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1