Out of Sight, Out of Mind: Financial Illiteracy and Sluggish Mortgage Refinancing

EMANUELE BAJO*
MASSIMILIANO BARBI*

This Draft: *November 2014*

Abstract

We analyze the effect of an exogenous shock to the Italian mortgage market and we show that most households do not act rationally when it comes to take mortgage-refinancing decisions. Thanks to a new legislation passed in 2007, borrowers have been allowed to refinance their loans at no cost. This reform—along with the drop of interest rates occurred between 2008 and 2009—has produced a unique opportunity to refinance fixed rate mortgages with substantial gains. However, only a minority of borrowers has shown this rational behavior. This sub-optimal choice is strongly associated with socio-demographic characteristics and with the level of financial illiteracy.

JEL Classification: G14; G32.

Keywords: Mortgage Refinancing; Financial Literacy; Household Finance.

---

* University of Bologna, Department of Management, via Capo di Lucca 34, 40126 Bologna, Italy.
Corresponding Author: Emanuele Bajo, e-mail address: emanuele.bajo@unibo.it.

1. Introduction

The detrimental effects of poor financial literacy on household's financial decisions have been widely documented. Several studies have shown that the inability to understand even simple financial problems leads to non-negligible losses and inefficient behavior. It has also been documented that the level of financial literacy is highly correlated with some socio-demographic characteristics. For instance, it is recognized that less educated and low-income individuals, women and immigrants are less able to correctly answer unsophisticated financial questions, and act irrationally when it comes to take financial decisions (Lusardi, 2008; Campbell, 2006; Calvet *et al.*, 2007, 2009; Jappelli and Padula, 2013; Lusardi and Mitchell, 2008; Santos and Abreu, 2013). Saving and investment behavior of households and their financial decisions are deeply affected. Insufficient accumulation of wealth before retirement (Lusardi and Mitchell, 2007), poor participation in stock markets (Rooij *et al.*, 2011), choice of high-cost financial instruments (Hastings and Mitchell, 2011), inadequate portfolio diversification (Guiso and Jappelli, 2009), borrowing at higher cost (Lusardi and Tufano, 2009), and over-indebtedness (Gathergood, 2012), are all examples of sub-optimal economic choices attributable to poor financial literacy.

In this paper we study the effects of household's financial literacy exploiting a natural experiment in the Italian mortgage market. In 2007 a new legislation was passed by the Italian government, granting the borrower the option to transfer her existing loan to a different financial institution at no cost, and without any authorization from the bank of origin. Moreover, the same legislation introduced greater refinancing flexibility imposing no prepayment penalties for new mortgages, and virtually nullifying those on existing contracts. This structural break in the previous legislation has considerably reduced refinancing costs, regardless of the contractual prepayment fee that were agreed upon at the time of mortgage origination. Obviously, the effects of this new regime have depended on the type of mortgage. While adjustable rate mortgages (ARMs) have only marginally benefitted from it, as the prepayment penalty was for these contracts generally modest, fixed rate mortgages (FRMs) have enjoyed the full potential of it.[1] In addition, while ARMs have had only a minor incentive to refinancing, being their instalments automatically adjusted to interest rate changes, FRMs have experienced a unique opportunity to cash in monetary refinancing gain from the interest rate drop between late 2008 and the beginning of 2009.[2] Using a large proprietary dataset of approximately 147,000 FRMs issued by a primary Italian bank, we estimate that the potential gain for fixed-rate borrowers can be quantified in about 8 percent of the average value of the mortgage loan. Despite this sizeable gain, and notwithstanding

[1] The legislator's intent was to liberalize the mortgage market, allowing borrowers to switch bank and relax the rigid preexistent setting. However, the economic rationale of mortgage prepayment fee is twofold. On the one hand, it discourages borrowers from modifying the structure of mortgage cash flows, thus facilitating bank's financial planning and efficient use of capital. On the other hand, the fee compensates banks from the potential loss originated by mortgage early termination, as in this case they are deprived of a contractual interest rate above current market conditions. Whilst the legislator's will was to enhance bank competition, we suspect the latter point has been completely unseen.

[2] Using financial option's jargon, the right to refinance a mortgage corresponds to a long put option written on the prevailing fixed market rate, with a strike price equal to the contractual fixed rate (minus transaction costs, mainly early termination fee and notary expenses). The combined effect of the new legislation (abolishing transaction costs) and the significant interest rate drop have contributed to push the refinancing option in-the-money.

the extensive media coverage provided to new legislation, we document that a substantial portion of fixed-rate borrowers did not refinance their mortgages. After controlling for mortgage characteristics and financial market conditions, we find that this sub-optimal behavior is strongly driven by socio-demographic characteristics and proxies of financial illiteracy. We show that less educated, poorer, immigrated, women and households living in less developed areas of Italy are more likely to miss this favorable refinancing opportunity. Also, financial literacy explains the optimal refinancing behavior. Borrowers holding specific financial knowledge (college degree in a finance related discipline), or a deeper financial experience (on banking products or other financial services), are more likely to timely exercise the refinancing option.

The literature on mortgage termination suggests that fixed-rate borrowers should refinance their loan to take advantage of lower interest rates (Bennett *et al.*, 2001). At the optimal refinancing rate, the present value of future interest savings should compensate refinancing costs, including the time value of the refinancing option (Agarwal *et al.*, 2013). In terms of the empirical behavior of fixed-rate borrowers, the evidence is less clear-cut. Green and LaCour-Little (1999) find that borrowers do not follow this refinancing rule, prepaying their loan when it is not optimal, and failing to prepay when it would be optimal. This evidence appears to be generally confirmed in the literature (e.g., Chang and Yavas, 2009, for early refinancing; Giliberto and Thibodeau, 1989, for late refinancing). Campbell (2006) shows that active refinancers are younger, better-educated, white households with high-valued houses. More recently, Keys *et al.* (2014) and Andersen *et al.* (2014) provide further evidence of late mortgage refinancing. The former study shows that 20 percent of unconstrained US households with a clear economic benefit from refinancing have not exploited this opportunity. Likewise, Andersen *et al.* (2014) find an evident sluggish behavior in the refinancing of Danish mortgages. They also show that age and wealth increase the refinancing inertia, whilst education and income operate in the opposite direction.

We contribute to both the literature on financial literacy and mortgage refinancing analyzing households' behavior and their sub-optimal choices in exercising the mortgage refinancing option. Different from existing literature, our study leverages on a natural experiment induced by an exogenous shock occurred in Italy in 2007, i.e. the introduction of a new legislation on mortgages. This event has to be considered as unique for three reasons. (a) The new legislation has imposed no prepayment penalty for all mortgages, contributing to push the refinancing option in-the-money at any time in which the current fixed interest rate falls below the contractual rate; (b) the concomitant 2008-2009 dramatic interest rate drop, combined with the absence of an "exercise price," offered fixed-rate borrowers a valuable (and free) deep in-the-money refinancing option; (c) the tumultuous hype that followed the introduction of the new law has made Italian borrowers aware of the potential financial gains from mortgage refinancing. These combined effects (i.e., sizeable refinancing gain, penalty fee abolition, and visibility shock), at the best of our knowledge, have produced a never-occurred and clean environment to test the determinants of sub-optimal refinancing decisions.

Our results are based on data taken from a primary Italian financial institution credit files. The sample used in our main empirical specifications comprises around 170,000 fixed rate mortgages (FRMs) issued between January 2003 and June 2009.[3] For

[3] We exclude ARMs from our main analysis, as the economic convenience to refinancing in this case is minor. However, we include them in our robustness checks to correct for potential selection bias. The number of ARMs in our final sample (after data screens) is about 147,000.

all loans we have a complete set of mortgage-specific variables, along with some socio-demographic characteristics of borrowers. We complement our dataset with information on more strictly related financial literacy proxies (such as, knowledge of different financial instruments, past experience on trading these products, subjective financial expertise and risk aversion) for approximately 18,000 borrowers.

The remainder of the paper is organized as follows. The next section provides a review of the literature. Section 3 overviews of the Italian market for mortgages and describes the new legislation passed in 2007. Section 4 details our research methodology. Section 5 and 6 present our empirical findings and some related robustness checks. Finally, section 7 concludes.

2. Review of the Literature

The decision to refinance a mortgage depends on a number of factors, of personal (endogenous) and financial (exogenous) nature. Among the formers, the decision to sell the house (due to an enlargement of the family, higher income making a more comfortable residence affordable, or the need to move elsewhere), or the decision to prepay the mortgage, for example thanks to an inheritance. In addition to personal drives, financial motivations also underpin rational refinancing decision. The refinancing option should be exercised in response to a significant interest rate drop, such to benefit of a lower borrowing cost. This interest rate drop should be large enough to outweigh refinancing costs (Downing *et al.*, 2005; Agarwal *et al.*, 2013). However, individuals show far from pure rational behavior in their lives, and economic choices are often driven by other than rational reasons. Likewise, irrational financial decisions are shown to be associated with a low level of financial literacy.

Recent literature has shown a surprising poor level of financial skills among households. Lusardi and Tufano (2009) document that only one-third of US population appears to be able to understand simple financial concepts such as interest compounding or the functioning of credit cards. Similarly, Klapper *et al.* (2013) report that, in spite of the massive growth in consumer's borrowing, only 41 percent of Russian households understands interest compounding, and 46 percent correctly answers simple questions on inflation. A slightly better scenario is depicted by van Rooij *et al.* (2011), who survey Dutch households. The authors report that a majority of respondents displays some basic financial knowledge on interest compounding, inflation and time value of money, but only a minority understands the difference between stocks and bonds, risk diversification, or the inverse relationship between bond prices and interest rates. Jappelli (2010) offers an international comparison of the level of financial literacy using a survey of executives in 55 countries and covering the period 1995-2008. His results show that financial literacy greatly varies among countries, and depends on educational achievement, social interactions, and financial development. In particular, among the first 10 European countries by GDP, Italy and Spain show the lowest score, while Sweden, Switzerland and Netherlands have the highest.

Financial literacy is strongly related to socio-demographic characteristics. Lusardi (2008) reports that financial illiteracy is higher among low-educated individuals, women, and ethnic minorities. Campbell (2006) finds that lower income and education are likely to lead to financial mistakes. Calvet *et al.* (2007, 2009) use Swedish data and document that lower wealth, income, and education are predictors of lesser financial sophistication and sub-optimal financial behavior. Jappelli and Padula (2013) support

most of the those findings, as they show that more efficient saving decisions are carried out by male, young, married, graduated and high income investors. There is also evidence that women display generally lower levels of financial literacy (Lusardi and Mitchell, 2008; Santos and Abreu, 2013). In Italy, a recent study by Fornero and Monticone (2011) documents that more than two third of individuals answer incorrectly to questions involving very basic financial concepts, such as interest rates and inflation, and men, more educated, and Centre-North-located individuals have higher financial literacy.

There is also indisputable evidence that lack of financial literacy induces sub-optimal economic choices. Saving and investment decisions, accumulation of wealth, access to financial markets, and portfolio choices are all deeply affected by financial literacy. Bernheim (1995) first points out that financial literacy is fundamental to individual's consumption and saving behavior, and investors with modest financial knowledge often act on the basis of rough rules of thumb. Evidence shows that personal finance investment decisions are driven by financial literacy. Lusardi and Mitchell (2011) study the relationship between financial literacy and individual's plans for retirement, finding that the relationship is positive and strongly significant. Similarly, Lusardi and Mitchell (2007) relate financial literacy to the accumulation of wealth, and find that higher financial literacy leads to greater wealth accumulation before retirement. A positive relationship between literacy and participation in employer-provided pension plans is also supported in Clark *et al.* (2012). Similarly, Behrman *et al.* (2012) show that less financial literate individuals are also less likely to accumulate wealth. Hastings and Tejeda-Ashton (2008) on Mexican data and Hastings and Mitchell (2011) for Chile show that financial knowledge helps investors to discern and choose low-fees investment funds. Financial literacy also affects access to financial markets and stockholdings. Christelis *et al.* (2010) find that cognitive abilities and financial literacy influence stockownership, and van Rooij *et al.* (2011) show that low literate investors are less likely to enter the stock market. Recently, Cole *et al.* (forthcoming) find that education increases financial market participation, and reduces the likelihood for individuals to go bankrupt. Also, Klapper *et al.* (2013) document that during the 2009 Russian crisis literate investors have more likely participated in financial markets, have saved a greater portion of unspent income, and have experienced less important losses. Cole *et al.* (2011) study economic literacy in developing countries, using Indian and Indonesian data, and find that it strongly predicts demand and access to formal financial markets. Finally, portfolio diversification positively depends on financial literacy and investors' sophistication (Guiso and Jappelli, 2009; Abreu and Mendes, 2010; Santos and Abreu, 2013).

Other than investment decisions, financial literacy affects household's financing decisions, in that borrowers with limited financial literacy are more likely to incur into higher cost of funding and larger fees. Campbell (2006) finds that in the mortgage market higher costs are associated with lower levels of financial literacy. The same evidence is found by Lusardi and Tufano (2009), as borrowing at higher cost and greater fees are associated to low levels of financial literacy. Lusardi and Mitchell (2009) document that one-third of charges and fees paid on credit cards are attributable to ignorance. Disney and Gathergood (2013) show that less financial knowledgeable UK borrowers hold larger portion of high cost credit, such as home collected credit, mail order catalogue debt and payday loans. Similar evidence is provided by Agarwal *et al.* (2009), as financial illiteracy drives the decision to use payday loans instead of cheaper financing alternatives. Poor financial literacy also leads to over-indebtedness (Stango and Zinman, 2009; Gathergood, 2012; Lusardi and Tufano, 2009), and enhance the risk

of mortgage delinquency. Gerardi *et al.* (2010) find a negative correlation between individual's numerical ability, delinquency, and default rates on mortgages. The magnitude of such effect is important, as those with the highest ability exhibit two-third less foreclosures relative to the lowest group. Financial fragility—especially among less educated, women, low-income and older people, and ethnic minorities—is the final negative social outcome of financial illiteracy (Bernheim, 1995; Lusardi and Mitchell, 2007, 2009; Lusardi and Tufano, 2009).

Within borrowing decisions, illiteracy may also drive inefficient mortgage refinancing choices. Financial literature has shown that households do not behave rationally in early termination or refinancing of their mortgages. Green and LaCour-Little (1999) find that borrowers irrationally prepay their loans when it is not optimal, and fail doing so when their prepaying option is deep-in-the-money. Although they do not relate the prepayment decision to either personal characteristics or to the level of financial literacy, they show that one-fourth of those borrowers act in response of a declining collateral constraint (i.e., the market value of their house). Similarly, Archer *et al.* (1996) analyze the role of post-origination income and collateral constraints effects upon the decision of mortgage termination, showing that constrained borrowers exhibit a markedly higher propensity to behave sub-optimally. A very similar sub-optimal behavior is found in other papers (e.g., Chang and Yavas, 2009, for early refinancing; Giliberto and Thibodeau, 1989, for late refinancing). Campbell (2006), using data from roughly 5,000 respondents to the American Housing Survey (AHS), shows that most active refinancers are younger, better educated, white households with higher-priced houses. Andersen *et al.* (2014) study the Danish mortgage market and find that younger, more educated, and wealthier households are more active refinancers, and are more likely to respond to refinancing incentives.

3. Legal Framework and Refinancing Decision

On 1st of February 2007 a new legislation was passed in Italy, with the aim of promoting competition in certain economics sectors and strengthening consumer rights. A significant innovation introduced by this decree—named "Decreto Bersani,"[4] after the name of the minister who proposed it—is greater flexibility in mortgage market. Before 2007, Italian mortgage market was extremely rigid if compared to other European countries, as both prepayments and renegotiations were exceptionally rare. Italian banks used to discourage borrowers from these practices through high fees for early redemption. A report published by the European Central Bank in 2009[5] shows that only 1 percent of loans for house purchase has been subject to early repayment in 2007, *vs.* 6 percent average for the Euro area. Since 2007 was the first year in which mortgage market was liberalized, this percentage is expected to be far smaller in the earlier years. Anecdotal evidence[6] suggests that before 2007 both early redemption of a mortgage (with the aim of achieving more favorable economic conditions with a new lender), or renegotiation of mortgage provisions with the original lender, were very unusual practices. For a standard fixed-rate mortgage of €100,000 (roughly the median amount in our sample), the total completion costs attached to mortgage refinancing—consisting

[4] Decree number 7/2007, ratified as Law number 40/2007.

[5] ECB Structural Issue Reports, "Housing Finance in the Euro Area," Table 2 on p. 27.

[6] We discussed this point with some commercial bankers, and we had full anecdotal confirmation of our conjecture.

of prepayment fee, application fee, mortgage registration tax, and notary fees—may have reached 5 to 8 percent of the value of the mortgage, making mortgage refinancing an economically unattractive choice.[7]

The advent of the new legislation in 2007 have simplified mortgage refinancing, eliminating prepayment fees, and introducing the so-called "subrogation." Thanks to this right, borrowers have now the opportunity to switch from the original bank to another financial institution without having to redeem the old mortgage and register a new one. The borrower is only required to notify the previous bank of this change, and the bank can neither oppose to this decision, nor apply any repayment fees. When the mortgage is subrogated, the new market conditions immediately apply, and the fixed-rate borrower can benefit of a decrease in interest rates with the new bank.[8] In this regard, the market for mortgage refinancing is now similar to the US, where borrowers are not subject to any prepayment fee. Instead, the usual practice in Europe (with the notable exception of Denmark), is that FRMs are always subject to a prepayment penalty (despite in some countries, such as France and Spain, this penalty fee is capped).

Thanks to this new legislation, the refinancing decision has been a valuable opportunity to Italian borrowers since 2007. Specifically, Italian households have had the possibility to undertake one of the following two actions: (a) switching at no cost from the original bank to another financial institution, without redeeming the old mortgage and register a new one, exploiting the drop in the interest rates (i.e., "subrogation"); (b) renegotiating the mortgage with the original bank, with the aim of obtaining more favorable conditions, leveraging on the increased competition among banks and the new borrower's bargaining power. Both these options allow the borrower to obtain a potential benefit at the expense of her counterpart.

The refinancing decision is a time-dependent optimization problem, where the fixed-rate borrower trades-off the refinancing costs with the benefits of paying less in future mortgage instalments (Aggarwal *et al.*, 2013). In terms of option-like payoff, the refinancing decision may be viewed as a long put option, where the underlying is the current level of fixed interest rates, and the exercise price is the contractual fixed interest rate minus the refinancing costs (as a percentage of the residual balance). In other words, if the borrower refinances the loan, she receives back the contractual rate, net of refinancing costs, in exchange of a current lower fixed interest rate. In the context of the new legislation, from February 2007 Italian mortgage borrowers were granted the flexibility to refinance their mortgage at no cost. This means that the exercise price of the refinancing put option has unexpectedly increased by an amount equal to the refinancing cost, thus pushing the option moneyness toward the in-the-money interest-rate interval. Despite substantial gains from refinancing were exploitable, the next section documents that only a scant minority of borrowers has actively refinanced their

[7] The very limited number of mortgage prepayments (29) in our sample of FRMs prior to year 2007 is an empirical confirmation of this conclusion.

[8] More precisely, according to the new legislation, (a) no penalty fee is applicable in case the old mortgage is "subrogated," i.e. moved to another financial institution without modifying principal amount and maturity. Also: (b) all new mortgage loans, originated after 2007, must be free of any prepayment penalty, and (c) old mortgage loans, originated before 2007, must have their maximum prepayment penalty automatically reduced to a maximum amount set forth by the law, with a further reduction for mortgages close to expiry. Maximum prepayment penalty is equal to 0.5 percent of the principal for loans originated before 2001, with a reduction to 0.2 percent and 0 percent for mortgages maturing in three years and two years or less, respectively. Loans originated after 2001 can be charged a maximum prepayment fee of 1.9 percent, but if expiring in three years and two years or less, the fee drops to 0.2 percent and 0 percent, respectively.

mortgages, being this sub-optimal behavior strongly associated with individual (socio-demographic and literacy) characteristics.

4. Data and Summary Statistics

We use two sources of proprietary data provided by Unicredit Group, the largest Italian commercial bank. The first set of data comprises loan-level data, providing information on both mortgage (contractual) and borrower (socio-demographic) characteristics. The second set profiles the level of knowledge and financial experience (i.e., financial literacy) of the borrowers. The former dataset includes full information, as of 30 June 2009 (in what follows, the "examination date"), on domestic mortgages provided to households in Italy from 2003, for a total of more than 630,000 loans. Variables covered can be classified into three categories.

(a) Mortgage-specific information, comprising the loan amount, its duration, the mortgage structure (essentially, adjustable-rate mortgage—ARMs, *vs*. fixed-rate mortgage—FRMs), the base rate (basically, the swap rate—Eurirs—for FRMs, and the 3-month Euribor for ARMs), the credit spread, the loan-to-value (i.e., the ratio between the mortgage principal and the appraised value of the real property), the number of guarantors other than the borrower, details on the location of the property (at the level of zip code), details on the location of the bank branch originating the mortgage (at the level of zip code), and full information on special clauses attached to the loan (i.e., cap rates, some advantageous conditions to bank employees, etc.).

(b) Borrower-level information, including gender, occupation, monthly net income (for a subset of around 27,000 observations),[9] wealth segmentation,[10] and details on date, place of birth and nationality. We do not possess from this source the level of education (*Graduate* dummy), but we do infer it from borrower's occupation.[11]

(c) Information on the status of the mortgage at the examination date. In particular, we detect whether (and when) the mortgage has been subrogated (i.e., transferred to another bank), or prepaid. From comparing the type of mortgage and its conditions (i.e., base rate and credit spread) at the inception and at the examination date, we also detect whether the mortgage has been renegotiated (at different conditions) within the same bank. Finally, we track whether the borrower has delayed her instalments, or if she has defaulted (we classify these loans as non-performing mortgages).

---

[9] Since information on the borrower's income is only available for a subset of observations, we impute this variable extending it to the full dataset. To do this, we run a median regression of the reported net income on 18 occupation dummies (describing the profession of borrower), the gender of the borrower, her age and geographic region, for the subset of about 27,000 observations for which net income is available. We then predict the net income for the complementary portion of our dataset for which the net income is not available. All coefficients from median regression are statistically significant at 1 percent level.

[10] According to the following scheme: individuals with total financial assets of less than €100,000 (mass market), owning a total wealth of €100,000 to €500,000 (affluent), or larger than €500,000 (private), respectively.

[11] The correspondence between occupation and education is one-to-one for some professions (e.g., physician, teacher, magistrate, etc.). For some other professions (e.g., lawyer, architect, public accountant, engineer, etc.), it is mandatory to join specific associations for practice, and a university degree is compulsory to obtain association membership. For other occupations, we infer the higher education level when the definition "director," "manager," or "executive" was present within the profession description. Residually, we set the *Graduate* dummy to zero. We argue that this procedure conservatively underestimates the number of individuals with a university degree.

Aside of our main dataset, we obtain information on the level of financial literacy of the borrowers from the "Markets in Financial Instruments Directive" (MiFID) questionnaire. In 2004 the European Parliament has passed a directive designed at regulating a number of aspects of European financial markets.[12] The aim of this directive is to protect retail investors, requiring banks to assess their knowledge and experience related to investment in financial instruments. As a consequence, every bank must require its clients to fill out a questionnaire before allowing her to submit any purchase or sale order.[13] The directive imposes the assessment of the level of financial literacy only for investment activities. Mortgage borrowers are not directly concerned by the MiFID, and they have not to fill out the questionnaire before taking on their loan. However, it is not uncommon that households keep their investments (if any) under management with the same bank where they have a borrowing relationship. Therefore, we collect from Unicredit information on MiFID questionnaires for those clients having both a mortgage and a financial portfolio, consisting of approximately 18,000 observations. From the questionnaire, we obtain information on:

(a) the self-declared level of knowledge of different asset classes which are related to the refinancing decision (fixed income, bank products, and financial derivatives);

(b) the self-declared level of experience (based on the number of past trades) of the same asset classes;

(c) personal characteristics, such as whether the borrower has a educational background in finance or other finance-related discipline, whether she holds or has held in the past a working position in the financial field, and her risk propensity.

The first two sets of information are also broken-out into further detail. Each of the three asset classes is divided into different types of instrument, and for every instrument the respondent has a four-notch scale to declare her level of knowledge and experience.[14] To construct a measure of financial knowledge (*Awareness*), for each of the three asset classes, we average out the responses provided by the borrower across the sub-levels. For instance, the debt (fixed income) asset class is divided into four sub-classes (Treasury, corporate, structured, and subordinated bonds). We measure knowledge of fixed income products (*Fixed Income Awareness*) with the average of the score attributed to each of the four sub-classes by the respondent. To measure borrower's experience on financial instruments (*Experience*), we use a dummy variable that takes 1 if the respondent has ever traded any of the considered product types. For instance, if the respondent has ever traded either Treasury, corporate, structured or subordinated bonds, the dummy *Fixed Income Experience* takes the value of 1.[15] Likewise, borrower's personal characteristics are described with dummy variables, such as *Economics*

[12] The MiFID (directive 2004/39/EC) is in force since November 2007.

[13] The questionnaire includes questions on client's investment experience, knowledge of financial products, and risk appetite. A sample of the questionnaire is reported in appendix. For brevity, we only report the fields related to the level of financial literacy we use in our analysis.

[14] Regarding her "knowledge," the respondent has to declare a level from 0 (none) to 3 (high), while for "experience" the respondent reports the number of past transactions of every financial instrument according the following clusters: 0, 1 to 3, 4 to 6, or higher.

[15] The average over the four type of instruments is not suitable here, since we do not know, for example, whether the "1 to 3" category means 1, 2, or 3 trades. The problem is even more severe for the "higher than 6" response (which could mean equivalently 7 or many more trades). Therefore, we use a dummy variable to detect financial experience, which prudentially classifies as having no experience if and only if the borrower has never traded any kind of instruments belonging to the given asset class.

Background (an educational background in economics, finance, or a related discipline), *Financial Expertise* (the current or past job is in the financial or related field), and *Risk Averse* (the risk propensity is defined as "prudent" or "cautious", instead of "balanced" or "dynamic").

We finally complement our dataset with market data (from Thomson Reuters Datastream) on interest rate swap (Eurirs) and Euribor yield curves, statistical data (from the Italian National Institute of Statistics, ISTAT) on population divided by geographical area (at the level of zip code), media attention on the new mortgage legislation through the number of articles published on the Italian newspapers (from the database Factiva), and household attention through the number of web searches on mortgage portability (from Google Trends).

In our main analysis, we use only a subset of the total number of mortgages. Since the aim of our empirical investigation is to analyze borrower's behavior in exploiting the (suddenly turned into) costless refinancing option, we restrict our attention to plain-vanilla mortgages, thus excluding other types of loans which embed optionalities (e.g., mortgages with interest rate caps or collars), and mortgages with favorable conditions to bank employees. We also exclude non-performing mortgages,[16] and loans transferred ("subrogated") from other institutions, as for these mortgages (a) we lack information at time of inception, and (b) they can be considered as mortgages already refinanced. We also filter out mortgages with incomplete information on basic mortgage characteristics (i.e., loan amount, maturity, loan-to-value ratio, and contractual interest rate). Data screens leave us with 146,975 FRMs and 137,369 ARMs. Out of these two subsamples, we mostly use the former, as the economic benefit of ARM refinancing is negligible.[17] To evaluate the impact of direct measures of financial literacy, we run some regressions jointly using the dataset of FRMs and that including MiFID variables. In these analyses, the maximum number of observations is reduced to 18,087.

[Insert Figure 1 about here]

Figure 1 graphs the number of mortgages included in our sample as a function of the quarter of initiation, and the distribution between FRMs and ARMs, compared to the yield curve slope (proxied with difference between the 20-year and the 3-month interbank rate). The number of mortgages registers an upward trend from the first quarter of 2003 to year 2007. Afterward, we document a steady trend, followed by a severe fall during the second part of 2008, and particularly in the first half of 2009. Both the upward and the downward trends can be mostly explained by the real estate bubble and the following subprime crisis that hit the global financial system in 2008. Figure 1 also splits the number of mortgages between ARMs and FRMs. As for the number of new

[16] The Italian regulation defines "non performing" loans very strictly. This category includes (a) bad loans (the borrower is insolvent), (b) sub-standard loans (the borrower is facing temporary difficulties in paying her installments), (c) restructured loans (the bank has agreed to reschedule deadlines or reduce interest rates due, accepting a loss due to the borrower's deteriorating conditions), and (d) past due (exposures other than those classified as bad, sub-standard or restructured, that are past due for more than 90 days on a continuous basis).

[17] We use the sample of ARMs in section 7, where we account for a potential selection bias due to the potential non-random choice of a FRM *vs*. an ARM.

mortgages, the ratio of ARMs to FRMs is time dependent. While the portion of FRMs is minor from 2003 to 2005, it increases afterward, and from 2007 it becomes prevalent. The severe rise of short-term interest rates has progressively induced new borrowers to prefer FRMs, as the difference between short and long-term rates (i.e., the slope of the yield curve, or *Term Spread*) is one of the most important determinants in the choice of the mortgage type (Campbell and Cocco, 2003; Campbell, 2006). Toward the end of 2008, as the yield curve flattens, the proportion of FRMs reaches its peak.

[Insert Figure 2 about here]

Figure 2 displays the behavior of long-term interest rates in the considered time period.[18] The 20y-swap rate is around 5 percent in early 2003, it slides down to 3.5 percent in mid-2005, then it reverts to the initial starting point during 2007, and finally drops towards the end of 2008. Looking at this Figure, we can argue that FRM borrowers have experienced two major profitable opportunities to refinance their loans: the first in mid-2005, and the second toward the end of 2008 and thereafter. However, as we will document later in the paper, we witness a material refinancing activity only following the second interest rate drop. This is not surprising, as the new legislation entered into force in between the two down peaks. Slashing down refinancing costs, the Bersani law has induced a profitable refinancing opportunity which was not present earlier. It is worth to notice that this new legislation is the only marginal difference between the two time periods, since the size of the interest rate drop is comparable.

[Insert Table 1 about here]

Table 1 shows the basic average mortgage-specific variables of our sample of FRMs by year of inception. The number of mortgages presents large variations across time. In particular, the first three years register a limited amount of loans. We can explain this pattern as the combined effect of (a) the lower number of mortgages issued during the initial period, and more than this, (b) the minor fraction of FRMs relative to ARMs due to the steepness of the yield curve. The average loan amounts to around €110,000, and its average duration is 21.3 years. Both these figures display an increasing trend over time. However, although the mortgage length does not exhibit a large variability, the loan amount ranges from a minimum of €70,000 in 2004, to a maximum of about €118,000 in 2008, i.e. roughly a 50 percent increase. This rise is attributable to both higher demand—due to the upward price trend in real estate market, in turn requiring higher investments to purchase a comparable house—and large liquidity available to financial institutions, which fuelled the mortgage market and the real estate bubble. After 2008 the average loan amount declines, and so does the number of new mortgages (Figure 1). The loan-to-value (LTV) does not considerably change during the 2003-2009 period, ranging from a minimum of 58.3 percent in 2009, to a maximum of 66.3 percent in 2003. Both mortgage amount and LTV are distant from figures documented in the US market right before the subprime crisis (Jiang *et al.*, 2014), where the average loan is about $230,000 and the average LTV is 80 percent, and

[18] Figure 2 shows the 20-year Eurirs rate, as 20 years is the median length of FRMs in our sample.

also from the average European mortgage market[19] (where the LTV is comparable to that of the US), reflecting a more prudent lending policy of Italian banks. Not surprisingly, the credit spread[20] is lower at the beginning of the period, and it increases thereafter, reflecting the deteriorating market conditions in the second part of 2008, and in 2009.[21] The last column reports the average number of mortgage guarantors. This number ranges from a minimum of 0.43 to a maximum of 1.00. The lowest figure is reached towards the end of the period, when banks were most concerned about the riskiness of their loans. At first sight, this evidence appears as counterintuitive. However, banks generally require guarantors for riskier mortgages, and these loans belong to less solvent borrowers. From the end of 2008, the credit crunch severely reduced the number of new risky mortgages, and this fact most likely explains the lower number of guarantors. From Table 1 it is also evident that the majority of FRMs originates between 2006 and 2008, as an effect of the higher demand for FRMs due to the rise of interest rates and the flattening of the yield curve. Finally, the sample reduction in the last year of analysis is due to the sharp contraction of mortgage supply.

[Insert Table 2 about here]

Table 2 depicts the descriptive statistics of our sample, divided by mortgage-specific variables, borrower's socio-demographic characteristics, breakdown of geographical areas, market conditions, and proxies of financial literacy. The dummy variable *Active* identifies borrowers who have taken advantage of the more favorable market conditions and have exercised their mortgage refinancing option, while the dummy variable *Efficient* indicates that for 81.2 percent of mortgages refinancing would be appropriate, as for these cases the refinancing rate at the end of our investigation period (June 2009) is lower than the contractual base rate. Interestingly, only a small fraction of borrowers (4.2 percent) has either decided to subrogate their loan, or has chosen to renegotiate the economic conditions of it with the original bank. The distance between these two numbers testifies this sizeable sub-optimal behavior.[22] The median mortgage in our sample amounts to €100,000, expires in 20 years, shows a 67 percent

[19] ECB Structural Issue Reports, "Housing Finance in the Euro Area," Table 2 on p. 27.

[20] Due to confidentiality reasons, we are not allowed to disclose the absolute amount of credit spread applied to mortgages. Credit spread in Table 1 is normalized to the average 2003 figure.

[21] This increase can also be explained by the new mortgage legislation, which caused a transformation of the fee structure of FRMs. Since commercial banks were no longer allowed to price early termination options within a specific contractual fee, they reacted increasing the general spread on FRMs.

[22] Previous literature (Agarwal *et al.*, 2013; Andersen *et al.*, 2014) has named this phenomenon as error of *omission*, in contrast to the error of *commission* (i.e., refinancing when it is not economically profitable). In our sample, out of 119,403 borrowers (i.e., 81.24 percent of 146,975) who have a potential economic gain at the end of our observation period, we observe only 6,223 refinanced mortgages (4.23 percent of 146,975). Hence, we register 113,180 errors of omission. Some of these errors might be due to (1) a modest refinancing gain, making refinancing activity little appealing for the borrowers (due to the time required for surrogating or renegotiating their loan), or (2) an increase of credit spread, offsetting some of the gain from the drop in base interest rates. In spite of these two explanations, the wedge between the number of active and passive borrowers remains important. As for the error of commission, this concept is little applicable in our setting. Since the new legislation has eliminated refinancing costs, a loss can only arise from the decision to contractually turn the mortgage into a more expensive loan, which is unlikely to happen even for very poor financial literate individuals.

LTV, and shows no guarantors other than the borrower.[23] In terms of demographic characteristics, Table 2 shows that the average borrower is about 39 years old, and earns an after-tax monthly income slightly below €1,500.[24] About one tenth of the borrowers in our sample is classified as "wealthy," i.e. owns a portfolio of total financial assets worth more than €100,000. Two third of our borrowers are male, 13 percent of them has reached college graduation, and about 10 percent are non-Italian citizens. The third panel shows the geographical dispersion of borrowers. More than half of mortgages are located in the North of Italy.[25] The third panel provides information on market conditions. The base (swap) rates at origination have experienced significant variability—ranging from a minimum of 2.7 percent to a maximum of 5.1 percent—with a negatively skewed distribution (mean 4.5 percent, median 4.7 percent). *Interest Differential* measures the difference between the swap rate at the end of our investigation period (June 2009) and the contractual rate at the origination of the mortgage. Hence, it represents a crude proxy of the potential benefit from mortgage refinancing as of June 2009. Despite the highest benefits are expected to occur earlier (at the end of 2008), fixed rates were still relatively low in June 2009. In fact, Table 2 shows that interest rate differential is negative, as in June 2009 fixed rates were 43 (56) basis points lower than contractual rates, in mean and median, respectively. For an average 22-year mortgage with a principal of €112,000, this figure corresponds to a difference of approximately €6.5 to €8.5 thousand in terms of total interest paid over the residual life. *Term Spread* proxies the slope of the swap yield curve at the mortgage origination, and it is computed as the difference between long-to-short rates (20-years Eurirs rate minus 3-month Euribor). The average difference is 66 basis points, and 51 basis points in median. Finally, the last panel relates to financial literacy variable, drawn from the MiFID questionnaires. The first three variables (*Awareness*), ranging from a minimum of 0 to a maximum of 3, provide information on the self-declared level of knowledge of fixed income instruments (such as Treasury and corporate bonds), bank products (such as certificates of deposit or repos), and financial derivatives (such as warrants, options and futures). As expected, the knowledge of sophisticated financial derivatives is quite low, as the average figure is 0.19 out of 3. On the contrary, households seem to be more familiar with fixed income and bank products, showing an average (median) score of 1.83 (2.00) and 1.46 (1.33), respectively. Looking at the quartiles, in both cases at least three-fourth of households declare some (low, since the score is close to 1) knowledge of these instruments. The next three variables (*Experience*) report the level of trading experience on the same asset classes. Unlike the proxies of awareness, these variables are dummies. However, insights are in line with the level of knowledge. Roughly 5 percent of respondents have traded at least once in financial derivatives, whilst approximately half of the sample has subscribed bank (investment) products, and almost two-third exhibit some experience in the bond market. The last two variables suggest that 9.2 and 3.9 percent of respondents possess an educational background (college degree) in economics or finance (*Economics Background*), and have worked in the financial field (*Financial Expertise*), respectively. Finally, only one third of individuals describe themselves as risk averse.

[23] As we have previously mentioned, descriptive statistics on credit spread are only partially informative, and should be interpreted taking the 2003 figure as 100 percent.

[24] The reported net monthly income is marginally larger than the average net income of Italian households (€1,239, source: Italian National Institute of Statistics, ISTAT, 2008). However, this difference is expected, as mortgage borrowers are likely to dispose of larger wealth and income than the average population. ISTAT also reports differences in gender, as Italian women have on average 20 percent lower income.

[25] ISTAT shows that are important differences in terms of average income across the three areas of the country. North is the more industrially developed area of Italy, while the southern area is the least wealthy.

[Insert Table 3 about here]

Table 3 reports additional descriptive statistics of time-varying variables included in our sample. Since these variables will be employed in survival analysis to estimate the likelihood of mortgage refinancing, we focus our attention from year 2007 onward, i.e. the time period when borrowers have been granted effectively such a possibility.[26] *Refinancing Gain Base (RGB)* quantifies the potential benefit from mortgage refinancing, i.e. the present value of the difference between future instalments from the original mortgage (at the contractual IRS rate) and future instalments according to the IRS rate prevailing at each quarter (considering an amortized mortgage with constant monthly instalments), divided by the residual balance of the loan. In other words, this variable measures the potential refinancing benefit as a percentage of the residual principal of the loan. As Table 3 reports, mortgage refinancing is not profitable in most of years 2007 and 2008 (except for the first quarter of 2007 and the last quarter of 2008), but it delivers positive gains afterward. The maximum refinancing gain in our sample occurs in the last quarter of 2008 and in the first quarter of 2009 (corresponding to the sharp drop in interest rates), where it reaches almost 8 percent of the residual mortgage value.

For robustness, we compute an additional measure of refinancing gain, i.e. *Refinancing Gain Spread (RGS)*. It could be argued that the change in the (swap) base rate may be an incorrect measure of the refinancing profit, as credit spreads have increased as well over time, potentially offsetting the gain resulting from the drop of market interest rates. To account for that, Table 3 also reports the refinancing gains where both the change in interest rate and mortgage credit spreads are considered. Apart from the credit spread at the inception of the mortgage, we do not have information in our dataset on the potential credit spread for each borrower at each subsequent quarter. Therefore, to compute the time-varying credit spread that can be added to the base rate for refinancing old mortgages, we average the credit spread applied to new mortgages within each quarter.[27] The two measures of gain are, as expected, positively correlated, being the time-variation of interest rates a dominant factor over the change in credit spreads.[28]

The last two columns of Table 3 describe media and household attention paid to the new legislation on mortgages. For *Media Attention*, we analyse the headlines and body text of articles published in the Italian newspapers, in search of the term "subrogation". The variable counts the number of cumulative hits at each quarter, starting from the first quarter of 2007. The last column (*Household Attention*) refers to the average number of web searches of the word "subrogation," detected through Google Trends (standardized at 100 at the maximum during the period). Despite the two

[26] While a borrower who has stipulated the loan before 2007 enters into our survival analysis at the time of mortgage inception, she becomes "at risk" only after February 2007, when the new legislation has effectively eliminated the prepayment penalty (canceling the exercise price of the refinancing option).

[27] We are aware that this approach does not account for the cross-sectional differences among borrowers. However, different from the US market, Italian banks do not charge different credit spreads to different clients to reflect the creditworthiness of the borrower. Provided that the individual applying for a mortgage is eligible, Italian banks usually apply a standard credit spread prevailing on the market at that time. The different lending policy is mostly driven by the lack of a continuous credit score like the FICO score.

[28] The *RGS* can be considered as a conservative proxy of the refinancing benefit as we (anecdotally) observe that subrogated mortgages show lower credit spread if compared to new mortgages.

attention proxies (i.e., newspapers and web searches) are correlated, they do not coincide. Newspapers have progressively paid more attention to the effects of the new legislation up to the second quarter of 2008. Instead, the number of web searches presents a slightly different pattern. For the first three quarters, households have paid little interest to the legislative bill.[29] Starting from early 2008, individuals appear to be progressively more aware of the refinancing opportunity, up to the end of our observation period. The acceleration of household attention in the early part of our investigation period is explainable with the reaction of individuals to the dissemination of public information (Tetlock, 2007).

5. The Role of Socio-Demographic Characteristics

The combined effect of the new legislation and the drop in long-term interest rates has delivered significant (potential) gain to mortgage refinancing. This gain has been only marginally exploited by Italian borrowers. We argue that this sub-optimal behavior is largely explained by individual characteristics, namely socio-demographic attributes and financial literacy. To test this hypothesis, we first investigate the effect of household's socio-demographic characteristics (in this section), and in the next section we explore the incremental explanatory power (with respect to the previous attributes) of financial literacy proxies. Results will show that socio-demographic variables are themselves good predictors of the refinancing behavior of households. Since socio-demographic attributes are highly correlated with the level of financial literacy, these insights will allow to indirectly explore the role of financial illiteracy over the optimal refinancing decision. Also, direct measures of financial literacy have an incremental explanatory power over the optimal refinancing decision. However, since financial literacy proxies cover only a limited portion of our dataset, and sample selection may deliver a potential estimation bias, we initially focus on the whole sample, and then refine our analysis to the sub-sample for which financial literacy proxies is available.

We compare the subsamples of refinancers *vs.* non-refinancers in Table 4. As our interest is on the effects of individual attributes on refinancing decision, we also present a propensity score matching (PSM) analysis, as the choice of mortgage specifications may be non-random across socio-demographic characteristics. Next, we present multivariate analyses and we run a probit regression (Table 5) for the likelihood of mortgage refinancing as a function of socio-demographic characteristics, after controlling for mortgage-specific attributes and exogenous market conditions. We are aware that cross-sectional probit may underestimate the refinancing probability, as our sample ends in June 2009, and some borrowers may have well refinanced thereafter. For this reason, we complement our investigation through a survival analysis, estimating the instantaneous hazard of mortgage refinancing after accounting for data censoring and including time-varying covariates (such as increasing media attention paid to the new legislation).

5.1 Refinancers vs. *Non-Refinancers*

Table 4 compares the subsamples of refinancers *vs.* passive borrowers. Refinanced mortgages are larger in terms of amount, slightly longer-dated, exhibit higher LTV, lower number of guarantors and appear to be to some extent more

[29] In the first two months after the introduction of the law, Google Trends reports a virtually insignificant number of web searches.

expensive. Although these variables are statistically significant, the difference between the two subsamples does not appear markedly large once compared to total sample means. Only the average amount (€120,000 *vs.* €112,000) and the number of guarantors (0.36 *vs.* 0.79) are sensibly different in the sample of active borrowers.

[Insert Table 4 about here]

The next set of variables focuses on borrower's personal characteristics. Although statistically significant, mortgage refinancers do not appear to be very different in terms of age and income, but they are less wealthy, more likely men (71.3 percent of active *vs.* 67.9 percent of passive borrowers), more educated (*Graduate* dummy shows 15.0 percent mean for refinancers *vs.* 12.5 percent for passive borrowers), and have Italian citizenship (immigrants are 7.9 percent in the refinancers subsample *vs.* 10.0 percent in the complement). These differences are in line with our expectations and previous literature, where it is shown that financial mistakes are more common among women, less-educated people and immigrants (Agarwal *et al.*, 2009; Calvet *et al.*, 2007, 2009).

The last two sets of variables relate to geographical and market conditions. Active borrowers live in more densely populated cities, and they are more concentrated in the northern part of the country. These results are consistent with the evidence that larger cities and northern areas of Italy exhibit higher financial literacy (Fornero and Monticone, 2011). The last panel of Table 4 shows that refinanced mortgages are associated with a higher decrease in interest rate relative to their non-refinanced counterparties. This finding is expected, as refinancing gain increases as interest rates decrease. Finally, numbers show that refinancers have stipulated their mortgages when the yield curve was steeper (the difference between 20-year swap rate and 3-month Euribor is 0.92 percent *vs* 0.65 percent, for active and passive borrowers, respectively).

Since borrower's characteristics may self-select mortgage-specific figures, our next step is to compare the two subsamples after controlling for this effect. We match mortgage characteristics using a propensity score matching (PSM) procedure (Rosenbaum and Rubin, 1983; Imbens, 2004), using the following matching variables: *Amount, Maturity, Spread, LTV, Number of Guarantors, IRS at Origination* and *Interest Differential*, and we compare personal and geographical characteristics. The significance of ATTs[30] from Table 4, despite slightly weaker, confirms our previous insights. After controlling for mortgage and market characteristics between the two groups, we find evidence that socio-demographic characteristics affect the propensity to optimal mortgage refinancing. Younger and better educated men, living in large cities and richer areas of the country, are more likely to exploit the new refinancing opportunity.

5.2. Refinancing Probability

The univariate analysis provides preliminary evidence on personal characteristics being important factors in explaining the active response of individuals to declining market rates. In multivariate analysis we investigate the role of these personal attributes, controlling for other determinants, such as mortgage features, geographical areas, market conditions and year fixed effects. Table 5 reports probit average marginal

[30] ATT stands for "average effect of treatment on treated," and measures the average difference between treated (refinancers) and untreated (non-refinancers) subjects in the matched sample for each variable.

effects (AMEs) of covariates on the likelihood of mortgage refinancing, i.e. the dependent variable is a dummy variable (*Active*) equal to 1 if the borrower has refinanced her mortgage during the observation period. More precisely, *Active* includes borrowers who have either decided to move their mortgage to another bank ("subrogation"), or have renegotiated mortgage conditions with the original bank, in order to decrease the borrowing cost.[31]

[Insert Table 5 about here]

In the first model we include mortgage characteristics, and they are all statistically significant at 1 percent level. The refinancing probability is positively associated with the size and the time-to-maturity of the loan, and with a decrease of the refinancing rate, and negatively related to the LTV and the number of mortgage guarantors. The signs of these determinants are expected, as refinancing larger loans and loans with a longer maturity conveys higher benefits to the borrower. The benefit itself depends on *Interest Differential*. A negative interest rate differential corresponds to a potential lower borrowing cost, and hence higher refinancing profit.[32] The magnitude of this effect is important, as one standard deviation of interest rate drop (40 basis points) increases the refinancing probability by 1.15 percent.[33] Both *LTV* and *Number of Guarantors* have a negative effect on the likelihood to refinance. The sign of *LTV* could appear as counterintuitive, as we might believe that a larger (and not a smaller, as we find) proportion of debt coverage makes the borrower more willing to reduce her financing cost. However, as the size of the mortgage has been controlled for, a less pronounced recourse to debt necessarily means that the equity stake is larger. Mortgages showing a higher LTV are therefore associated with a lower creditworthiness. To some extent, the same conjecture can be extended to the number of guarantors. Banks are likely not to request guarantors other than the borrower if credit standing is high and the equity stake is sizeable. It is also plausible to assume that higher credit quality is associated with greater financial literacy, and therefore a larger LTV and number of guarantors reduce the likelihood of mortgage refinancing.

As an additional comment, we could reasonably argue that some of the passive borrowers—who did not either subrogated or renegotiated her mortgage—have not be able to refinance their mortgage, simply because the changed market conditions inhibited it. After the sub-prime crisis, the appetite for mortgages among banks has widely reduced. Accordingly, some borrowers were no longer eligible to be refinanced given the new standards. For instance, after the crisis immigrants have become largely unattractive for banks and likely not refinanceable. *LTV* and *Number of Guarantors*

[31] Out of 6,223 active borrowers, we record 5,441 subrogations and 782 renegotiations.

[32] It might be argued that the economic convenience to mortgage refinancing does not only depend on the current level of market interest rates relative to that prevailing at the inception of the mortgage (change in the base rate). An increase of the credit spread can potentially offset the change in the base interest rates. However, two comments are in order. First, although credit spreads have effectively increased with the arrival of the subprime crisis, this increase is modest if compared with the size of the shift of the swap curve. Second, in the next paragraph we will show that the inclusion of the credit spread does not alter both the pattern of the refinancing gain and the effects of socio-demographic characteristics on the refinancing probability.

[33] Unconditional probability of refinancing is below 5 percent. Hence, a change in *Interest Differential* equal to one standard deviation increases the likelihood of active borrowers by roughly 27 percent.

explain the creditworthiness of the borrower, partially correcting for the potential problem of over-estimating refinanceable mortgages. As a robustness check, we will present later in the paper the results of a two-stage regression, where we first estimate the probability of being eligible for mortgage refinancing, and then the conditional probability to refinance the mortgage. Our dataset contains detailed information on delayed or unpaid mortgage instalments, and we can model the likelihood of observing these non-performing loans as a function of both mortgage and personal characteristics.

Model 2 adds socio-demographic characteristics of the borrower, which are in line with results of the univariate analysis. Men exhibit 0.67 percent higher probability to refinance; higher-educated individuals are 0.97 percent more likely to be active borrowers, and immigrants 1.10 percent less incline to refinance their mortgages.[34] The relation with *Age*, *Ln(Income)* and *Wealth* is negative, suggesting that older and richer borrowers tend to miss the refinancing opportunity offered by the new law. These results are in line with those on refinancing inertia found in a very recent study by Andersen *et al.* (2014) for the Danish market.

The third model includes geographical dummies,[35] and the (log of the) population of the city where the bank branch is located. The rationale for these inclusions relates to the effect of financial literacy on optimal financial decisions. More populated cities—especially if located in the northern part of the country—are expected to be associated with a higher refinancing probability, for four reasons. First, households living in larger centers are expected to exhibit higher financial literacy, regardless of the level of school education. Second, in larger cities spreading out of information (through words of mouth) is more likely, and so it is the event that less financially educated borrowers are advised to refinance their mortgages. Third, with the sole exception of Rome, important financial centers are mostly located in the northern part of the country. Fourth, despite unrelated to financial literacy, larger cities (and the northern and richer part of Italy) present higher bank competition, and this may ease borrowers to either subrogate or renegotiate their loans. Results confirm our conjectures, as the AME of *Ln(Population)* is positive (even if only weakly significant), and so are the *North* and *Center* dummies. Table 5 shows that households living in the northern and central area of the country show—relative to the ones living in the southern area—1.76 percent and 1.46 percent higher likelihood of being active, respectively.

The last model in Table 5 contains the same set of variables with the inclusion of squared effects of *Age* and *Ln(Income)*, to capture possible non-linearity.[36] Results confirm our findings, as signs, statistical significance and magnitude of covariates are generally maintained. We do find non-linearity in both variables, showing a positive convexity. Combining the magnitude of the marginal effects with the variable range, we note that *Age* shows a U-shaped relationship, with younger borrowers being more active. The lower refinance likelihood is reached at the age of 36 years, where the conditional probability is reduced by 4 percent. In terms of income, an increase of the variable decreases the likelihood of active refinancing but with a decreasing pace.

[34] The magnitude of these marginal effects corresponds to a change on refinancing likelihood of 16 percent, 25 percent and 27 percent with respect to the unconditional probability (4.23 percent).

[35] Italy is usually divided into the following areas: North, Center and South. The last area also includes the two major islands (Sicily and Sardinia).

[36] In the last model we run a linear probability model (LPM) instead of a probit. This is because Table 5 reports average marginal effects, and they would not be available for quadratic variables separately from their first-order counterparties. The number of observations assures that the coefficients of LPM consistently estimate probit AMEs.

5.3. Media and Household Attention

In this section we refine our previous findings adding two additional checks. First, we run a survival analysis, where we model the hazard rate of mortgage refinancing (i.e., the "failure" variable). This analysis allows to address right censoring and to consider the time at which mortgages become "at risk". In our sample we follow the behavior of borrowers up to June 2009, and previous analyses have implicitly considered as passive borrowers: (a) individuals who have never refinanced their mortgages, and (b) individuals who possibly have refinanced it, but after the end of our observation period (right censoring). Likewise, not all borrowers enter into our sample before the introduction of the new legislation. Survival analysis allows to: (a) distinguish the refinancing hazard rate for borrowers who effectively exercised their option from those who did not due to the effect of censoring; (b) take into account that any mortgage in place before February 2007 was not virtually allowed to be refinanced, in spite of a potential interest rate advantage. A second enhancement concerns the inclusion of a time-varying driver for household's behaviour, that is media and household attention. It has been shown that media hype has an impact on individuals in terms of a wide spectrum of phenomena.[37] Among others, Tetlock (2007) finds that media pessimism predicts downward pressure on market prices and high market trading volume. Gurun and Butler (2012) document that local media slant affects the value of the firm.[38] Along these lines, media attention to the new legislation is likely to be important, and should be controlled for. We proxy the level of attention paid to the new legislation using a variable (*Media Attention*) which counts the number of articles published in the Italian newspapers reporting the keyword "subrogation". This proxy captures newspaper coverage, but does not consider the effective impact of it on the attention paid by households to the new legislation. To measure this additional effect (*Household Attention*), we use the standardized number of web searches (in Google Trends), based on the same keyword.[39] Figure 3 displays the time evolution of the two variables from January 2007 to June 2009. The number of web searches accelerates after an initial very low level. Comparing this pattern with that of *Media Attention* suggests a delayed effect of media over the household attention. Borrowers might not immediately react to the information reported into the newspapers, as they need a strong (repeated) signal before forming their beliefs.

[Insert Figure 3 about here]

[Insert Figure 4 about here]

[37] Dyck *et al.* (2008) analyze the impact of media on legislative outcomes, documenting the effect of magazines on the voting patterns of US representatives and senators in the early part of the 20th century. Miller (2006) shows how the press has a role as a monitor or "watchdog" for accounting frauds. Johnson *et al.* (2005) document that publishing ratings on boards of directors by business press induces significant market reaction.

[38] Local media slant refers to the use of fewer negative words when local media report news about local companies. The authors also document that the main reason for the positive slant is the firms' local media advertising expenditures.

[39] The Italian keyword we use is: "surroga."

Attention variables are time-varying, and they are measured on a quarterly basis. Including the time dimension in our analysis allows to employ a dynamic measure for the economic benefit of mortgage refinancing. In the previous section we have proxied the potential profit considering the size and the duration of the loan, and the change in the borrowing cost (*Interest Differential*). A lesser refinancing rate, larger loan amount, and longer mortgage are associated with a higher refinancing propensity—as the resulting potential benefit is larger. In this section we measure this potential gain as the present value of interest savings captured by the borrower in the event of mortgage refinancing. We compute this gain (as a percentage of the residual loan amount) at each quarter from February 2007 to June 2009.[40] Figure 4 exhibits the time evolution of the refinancing gain (right axis) and the number of active borrowers (left axis). The refinancing gain sharply rises from September 2008 (in line with the shape of swap rates), and it slightly reverts in the second quarter of 2009. In this period the refinancing gain is sizeable, reaching almost 8 percent of the loan. Instead, the number of active borrowers registers a delay, and it peaks up in the first and second quarter of 2009. We register less than 500 active borrowers in the third quarter of 2008, but the same figure is more than twice as much in the next quarter, three times and six times in the two subsequent quarters, respectively. The visual correlation between active borrowers and refinancing gain is not as strong as we might expect. Toward the end of the period, we observe an opposite behavior of the two variables. The possible explanation is linked to the delayed effect of media hype. We conjecture that the beneficial effects of the new legislation have been neglected when first introduced, due to limited initial gain and thin media coverage. Only afterwards, media hype and the simultaneous interest rate drop have triggered the (still modest) refinancing wave, whose effects have protracted for some time after the peak (highest media coverage and contemporaneous maximum refinancing gain).

[Insert Table 6 about here]

Table 6 reports the results of a Cox proportional-hazard regression, where the hazard rate of mortgage refinancing is reported.[41] The first model considers the effects of refinancing gain, mortgage characteristics and *Media Attention*. Coefficients are consistent with our expectations and previous findings. A larger loan is associated with a higher conditional probability of refinancing, while we observe the opposite relationship when considering the role of *LTV* and *Number of Guarantors*. More interesting are the insights from *Refinancing Gain (Base)*. The effect of this variable is positive, significant, and also large in magnitude. Cox regression allows us to measure the incremental effect of *Media Attention*. As suspected, *Media Attention* has a positive and significant coefficient. This result confirms the conjecture that higher media coverage mitigates the sluggish refinancing behavior, as households become more aware of the profitable opportunity. Model 2 adds personal characteristics and geographical information. Socio-demographic characteristics maintain their expected signs: men and more educated individuals show about 30 and 37 percent higher refinancing probability, respectively, while immigrants are 30 percent less likely to take this decision. Likewise, wealth, income and age show a negative impact, although the last two also exhibit a non-linear

[40] As we previously pointed out (footnote 7), the refinancing activity is nearly absent before February 2007.

[41] To interpret the signs of the covariates, the table reports regression coefficients and not the hazard ratios. The marginal effects are computed exponentiating these coefficients and subtracting 1.

effect. Apart from the result on population, survival analysis confirms that residents in the northern and central part of the country have higher propensity to refinance. Model 3 and 4 test the same regressors, but *Refinancing Gain (Spread)* (*RGS*) includes the (estimated) credit spread. All variables are robust to this different specification. Model 5 and model 6 consider the effect of *Household Attention*. As for *Media Attention*, the number of Google queries significantly predicts the hazard of mortgage refinancing. The base measure of refinancing gain is robust to this different specification, whilst *RGS* loses most of its predictive power. The remaining explanatory variables confirm their signs and significance.

6. The Role of Financial Literacy

Previous analyses have shown that socio-demographic characteristics are important drivers in explaining the refinancing behavior of Italian mortgage borrowers. To some extent, these findings allow us to infer a causal relationship between a poor financial knowledge and sub-optimal refinancing decisions. However, as it has been shown in the literature, proxies of financial literacy may have additional explanatory power over socio-demographic characteristics. In this section we aim at addressing this point, shedding light on how the level of financial knowledge and expertise contributes to driving optimal refinancing decisions.

[Insert Table 7 about here]

To investigate this question we leverage on the information obtained from the MiFID questionnaire. Thanks to this questionnaire, we profile each household in terms of self-declared level of knowledge and past experience over a number of financial products. Furthermore, this source allows us to control whether the borrower has an educational background in economics or finance (college degree), if she has ever worked in the financial industry, and her level of risk aversion. Unfortunately, the sample of individuals who filed out the MiFID questionnaire cannot be considered as representative of the population of mortgage borrowers. MiFID profiling is mandatory for those bank clients who have financial investments under management, and therefore it does not apply to the majority of mortgage borrowers. Accordingly, borrowers in the MiFID sample are expected to be wealthier than those in our overall sample, and most likely differences arise also on other personal characteristics. Including financial literacy variables into our analysis reduces the number of observations, and the resulting inference possibly suffers from sample selection bias. Panel A of Table 7 contrasts the two samples. Households belonging to the MiFID sample are wealthier, as about one third of them belong to the affluent or private segment (i.e., net worth under management larger than €100,000), *vs.* 8.3 percent of borrowers in the residual sample. In terms of socio-demographic characteristics, borrowers who belong to the MiFID sample are older (age is 41.3 *vs.* 38.4), earn a larger income (€1,520 *vs.* €1,486), are more likely women (men are 65.1 percent *vs.* 68.3 percent), have a college degree (14.5 percent *vs.* 12.4 percent) and are domestic (foreign borrowers are 6.4 percent *vs.* 10.3 percent). All these differences are statistically significant at 1 percent level.[42]

[42] Table 7 also shows the differences between the two sets of household are strongly significant (with the sole exception of the *Graduate* dummy) even after controlling for mortgage specifications. The last two columns report the ATTs (and their t-statistics) obtained through the 10-nearest neighbors propensity score matching

If the sample of MiFID borrowers is not a random subset of the population, any inference based on this sub-sample is questionable, as the marginal effects of variables are biased and inconsistent. To mitigate selection bias we follow the following approaches: (a) we use propensity score matching (PSM), contrasting the financial literacy of refinancers with comparable individuals (in terms of mortgage specifications and personal attributes) that have not refinanced their loan; (b) we run a two-stage probit regression, where the first stage estimates the probability of the borrower belonging to the MiFID subsample, and the second stage estimates the refinancing probability, conditional on the first-stage predicted likelihood. Both these approaches will show that the potential sample selection is not large enough to drive our results, which are robust to different model specifications. Leveraging on this check, for completeness we also report the results of a Cox regression where we add proxies of financial literacy from MiFID questionnaires.

6.1. Propensity Score Matching

PSM controls for selection based on observable mortgage and socio-demographic characteristics. The variables used for matching are: *Amount*, *Maturity*, *LTV*, *No. of Guarantors*, *IRS at Origination* and *Interest Differential*, and socio-demographic characteristics, i.e. *Age*, *Income*, *Wealthy*, *Man* and *Foreign*. The second panel of Table 7 reports simple and propensity-score matched differences on the level of financial literacy between active and passive borrowers. Proxies of financial literacy include the self-declared level of (a) knowledge, (b) past experience of financial products, along with information on (c) educational background in economics/finance, and (d) professional involvement in a financial-related industry. From the MiFID questionnaire we register also the (self-declared) level of risk aversion. For the level of knowledge and past experience we consider investment in fixed income, bank products and derivatives. The MiFID questionnaire is broader, and reports knowledge and experience also on other products. However, we select those financial instruments which are more closely related to the mortgage refinancing decision, i.e. fixed income (as a FRM may be considered as a short coupon bond), bank products (which proxy the familiarity of the investor to the banking practice, broadly including both investment and financing products), and financial derivatives (as the refinancing decision may be viewed as the exercise of an option). In particular, being complex financial instruments, it could be argued that knowledge of financial derivatives should help borrowers to correctly assess the economic benefit from mortgage refinancing.

Starting with the level of knowledge (*Awareness* variables), the difference between the two samples denotes a higher financial literacy among active borrowers for all variables. Differences are also statically significant, at least at 5 percent level. However, as discussed, these differences might be the result of some selection bias. To mitigate this problem, the last two columns of Table 7 (lower panel) report the ATT difference and its t-statistic. As suspected, most of the distance between the two subsets weakens once that the differences are corrected for contractual and socio-demographic attributes between groups. Although these differences remain positive, their statistical significance holds only for the knowledge of derivative products. The second block of variables covers borrower's past experience on fixed income, bank products and derivatives (*Experience* variables). As before, simple differences reveal that refinancers possess larger experience on financial products, and are likely to be more financial

---

methodology, where the following variables are used for matching: *Amount, Maturity, LTV, No. of Guarantors, IRS at Origination and Interest Differential.*

literate. Again, when passing to ATT, differences lose some statistical power, but two out of our three proxies (i.e., experience on bank products and financial derivatives) remain positive and significant. The last two proxies of financial literacy are *Economics Graduate* (the borrower has obtained a college degree in an economics or finance related field), and *Financial Expertise* (i.e., the borrower's current or past occupation is related to finance or financial markets). Table 7 also reports the level of borrower's risk aversion. The effect of *Economics Graduate* is particularly strong (13.8 percent of active borrowers *vs.* 9.2 percent of passive borrowers has an educational background in economics or finance), and this effect holds after controlling for mortgage and socio-demographic matching variables. *Financial Expertise* has also a positive effect on active borrowers (5.1 percent *vs.* 3.7 percent), but ATT loses statistical significance. Surprisingly, we note no difference in terms of risk aversion between active and passive borrowers. We could argue that less risk averse households are more likely to refinance their loans, and react more promptly to external signals, such as the new regulation in place, or the drop in market interest rates. Our findings show instead that this attribute does not discriminate between active and passive borrowers.

6.2. Two-Stage Regression

To mitigate potential selection bias we run a two-stage regression, where we first estimate the probability for a borrower to be present in our MiFID subsample, and then we run a probit for the determinants of the borrower's refinancing decision. To carry out such approach, we need an identification strategy, i.e. one or more variables driving the probability of having a MiFID questionnaire, but potentially not correlated with the mortgage refinancing decision (i.e., one or more identification restrictions). Looking for valid instruments is never an easy task, and least of all in this case. From Table 7 it is natural to use socio-demographic variables as the main drivers of the selection equation, but these variables are obviously correlated with the level of financial literacy, which in turns, according to our hypothesis, explains the household's refinancing behavior. However, among socio-demographic attributes, household's wealth seems to greatly discriminate between MiFID and non-MiFID subsamples, as almost one third of MiFID borrowers are classified as *Wealthy*, relative to a scant 8 percent in the residual sub-sample. Moreover, survival analysis of Table 6 shows that the dummy *Wealthy* loses most of its statistical significance in explaining the active behavior of borrowers. Therefore, including in the selection equation socio-demographic attributes, along with the dummy *Wealth*, should contribute to mitigate the endogenous selection concern on the non-random nature of the MiFID subsample.

[Insert Table 8 about here]

Table 8 reports the results of the second stage selection model. With the sole exception of *Maturity,* the set of variables capturing mortgage characteristics, as well as market conditions, confirm their role in explaining the refinancing behavior. Borrowers are more likely to actively manage their loans in response of an interest rate drop. The refinancing decision is positively associated with the size of the loan, and negatively related to the *LTV* ratio and the number of guarantors. Interesting insights result from the inspection of the financial literacy proxies. The self-declared level of knowledge (*Awareness*) of financial products does not have a significant effect on the refinancing decision. None of the three *Awareness* variables has an incremental explanatory power. Turning to the level of *Experience*, the evidence is different. Despite familiarity with trading fixed income products does not affect the refinancing probability, the experience

with bank products has a positive and statistically significant coefficient. Also, this variable is economically important, as it increases the likelihood of mortgage refinancing by over 1 per cent. The experience on financial derivatives is positive, but statistically insignificant. We find no significance for the role of a working experience in a financial-related field, whilst the educational background in economics or a related field materially increases the refinancing probability. Finally, as for the PSM, we find no statistical significance for the self-declared level of risk-aversion.

[Insert Table 9 about here]

6.3 Survival Analysis

To corroborate the insights on the role of financial literacy, we complement Cox proportional-hazard regression of the previous section with financial literacy variables (Table 9). For sake of comparison to the previous analyses, we do not include socio-demographic characteristics among the explanatory variables (as they were previously used for identification purpose).[43] However, we do take into account the two measures of refinancing gain. Model 1 to model 3 use *Refinancing Gain (Base)*, while model 4 to model 6 use *Refinancing Gain (Spread)*, to reflect the deteriorating market conditions toward the end of our observation period. Focusing on financial literacy variables, we have a clear confirmation of the insights offered by probit regression. The experience on bank-related products (significant at 1 percent) increases the instantaneous refinancing probability by 40 percent, and the educational background in economics or finance (significant at 5 percent) leads to a 46 percent increase. The other control variables (with the sole exception of *Media Attention*) maintain their explanatory power.

7. Robustness Checks

In this final section, we present two robustness checks. First, we verify whether the choice of FRMs (*vs.* ARMs) drives any of our insights. If the characteristics of individuals choosing a FRM are correlated to the determinants of active refinancing a bias may arise. Second, we check whether the explanatory power of both socio-demographic characteristics and financial literacy variables are affected by the inclusion of subjects potentially willing to renegotiate their mortgages, but not eligible due to low creditworthiness.

7.1. Choosing between FRM and ARM

Since ARMs enjoy an automatic drop of future instalments in response to a decrease in interest rates, we have focussed our empirical investigation exclusively on FRMs. In fact, only FRMs exhibit an unquestionable refinancing benefit.[44] Although economically motivated, the choice to exclude ARMs may produce a bias if the *ex-ante*

[43] However, including socio-demographic variables does not alter signs and the statistical significance of the financial literacy proxies.

[44] We acknowledge that the new legislation has increased the degree of competition in the mortgage market, producing indirect non-negligible benefits also for ARMs.

decision to choose a FRM instead of an ARM is non-random. To control for this potential bias, we run a selection model. In the first stage, the selection equation models the decision to opt for a FRM, adding the instrument *Term Spread*—i.e., the slope of the swap yield curve at the inception of the contract—among the explanatory variables.[45] It is widely documented that the slope of the yield curve largely explains the kind of mortgage (i.e., FRM *vs.* ARM) chosen by the borrower. On the contrary, the refinancing probability (second stage) should not be driven by this variable, as a shift (and not a twist) of the yield curve is likely to be important for this purpose.[46] In first-stage regressions *Term Spread* is negative and highly significant.[47] Second-stage regressions correct for this bias estimating a correlation coefficient between the error terms of the two equations (Van de Vend and Van Pragg, 1981).

[Insert Table 10 about here]

Table 10 reports the coefficients of second-stage regressions. Model 1 includes the average marginal effects of mortgage-specific variables only. Models 2 to 4 include also socio-demographic characteristics of borrowers. We notice no appreciable differences in terms of signs and statistical significance of all covariates relative to Table 5, thus confirming the robustness of prior results. A greater refinancing propensity is associated with an interest rate drop, larger and longer-dated mortgages with lower LTV ratios and number of guarantors, non-immigrant men holding a college degree and living in more densely populated areas of central-northern Italy. Both *Age* and *Ln(Income)* show a non-linear U-shaped relationship.

For the role of financial literacy, we likewise run a two-stage model, where we first estimate a probit for the choice between FRMs *vs.* ARMs (using *Term Spread* and the socio-demographic characteristics as selection variables), and then we run a probit for the refinancing probability. As for socio-demographic attributes, financial literacy proxies are robust to this specification. *Bank Products Experience* and *Economics Background* maintain their positive and statistically significant effect.[48]

7.2. Ineligibility to Refinance

A possible caveat in interpreting our results could consists in the complexity to discern, among the non-refinancers, those who did not take advantage of the refinancing opportunity because of their inability to foresee the potential gain, from those who instead were not eligible due their poor creditworthiness.

[Insert Table 11 about here]

[45] Other than this variable, the selection model also includes the same covariates that we use in each of the outcome (second-stage) models.

[46] In other words, a borrower who has chosen a FRM (*vs.* an ARM) when the yield curve was steeper (less steep), should exhibit the same refinancing likelihood when perceiving a potential gain.

[47] We do not present first-stage models since, apart from *Term Spread*, they are little informative. Results are available upon request.

[48] We do not report these results. However, they are available upon request.

To settle this legitimate objection, we run a two-stage regression, where we first estimate the probability to be eligible for refinancing, and then we use the predicted values to estimate the refinancing propensity (conditional to be refinanceable). For the first stage regression we incorporate into our sample 19,426 non-performing FRMs previously excluded from our main analyses. The dependent variable is a dummy taking one if the mortgage is not classified as non-performing loan. Independent variables are mortgage and socio-demographic characteristics, market conditions (interest rate change), along with loan-to-value (*LTV*) and number of guarantors (*No. Guarantors*), which we use as instruments. In the second stage we run a probit for the refinancing probability, conditional on the information from the first stage. In other words, the first stage provides a rating to each observation in our sample (i.e., an estimated default probability), and the second stage re-estimates the refinancing probability (as in Table 5) conditional to the rating obtained from the first stage. Table 11 reports the results from the second stage regressions. With the only exception of the net income, the other variables confirm both their sign and statistical significance. This result suggests that the sluggish refinancing behavior is explained by individual (socio-demographic and financial literacy) attributes, though correcting for potential refinancing ineligibility.

[Insert Table 12 about here]

To further support this point, we sort out our observations by the synthetic rating obtained through first-stage probit regression, and then we run four probit regressions excluding the lowest first, fifth, tenth and twenty-fifth percentiles in terms of credit quality, respectively. The rationale for this is to verify whether our results are robust after excluding an increasing portion of poor quality loans, which could be simply ineligible to refinancing, given the tighter mortgage supply. Model 1 of Table 12 shows the first-stage probit regression for the likelihood of a performing loan (i.e., credit rating attribution), and models 2 to 5 present second-stage probit regressions for the refinancing likelihood, after excluding an increasing portion of potentially non-performing loans. In terms of mortgage characteristics, model 1 shows that higher probability of mortgage delinquency is positively associated with the time-to maturity, the loan-to-value ratio, and the number of guarantors, and negatively correlated to the size of the loan. Model 1 also suggests that poorer men, immigrants, and elder borrowers are more likely to default, whilst a college diploma reduces mortgage delinquency.[49] Model 2 to model 5 are robust with our previous findings after excluding lower quality mortgages, and the statistical significance of coefficients hold. The only exception is the dummy *Foreign*, that turns from negative and significant (at 1 percent cut-off), to positive and significant (at 10 and 25 percent cut-off). This evidence indicates that excluding the poorest bracket of borrowers is akin to take out of our sample the majority of foreign individuals, and the remaining foreign borrowers exhibit an even larger refinancing probability than the average population.[50]

[49] The effect of these explanatory variables over mortgage delinquency are consistent with the literature on the topic (e.g., Magri and Pico, 2011, for Italy and some selected European countries).

[50] After excluding the left tail of the rating distribution we are left out with 3,182 (10 percent cut-off) and 1,883 (25 percent cut-off) out of 14,401 immigrants in our full sample.

We carry out the same analyses when incorporating our proxies of financial literacy. Results show that *Bank Products Experience*, as well as *Economics Background,* are strong predictors of borrower's refinancing likelihood.[51]

8. Conclusions

In 2007 a new legislation was passed in Italy, allowing mortgage borrowers to refinance their loans at no cost. This reform, along with the drop of interest rates occurred between 2008 and 2009, has produced a unique opportunity to refinance fixed-rate mortgages with substantial gains (about 8 percent of the average loan value). In spite of fundamental economic profitability, only a scant minority of borrowers has exercised this right, and this behavior is strongly associated with socio-demographic individual characteristics and financial literacy proxies. After controlling for mortgage characteristics, market conditions, media and household attention, we show that women, less-educated, immigrants, and borrowers located in the poorer and less developed areas of the country are less likely to take advantage of the exercise gain from the refinancing option. Controlling also for these individual characteristics, we find that the (self-declared) experience on bank products, as well an educational background in economics or finance, positively affects the optimal mortgage refinancing decision.

Our results are robust to two potential biases. As we focus on FRMs, we account for the endogenous choice of the mortgage type, under the conjecture that passive borrowers might more likely select FRMs instead of ARMs. Our two-stage selection models reject this concern. Also, we consider the fact that some borrowers did not refinance because they were not eligible, due to their poor creditworthiness. Leveraging on data on non-performing loans, we first estimate the probability of being eligible to mortgage refinancing, and then we estimate the refinancing probability conditioning to first stage results. Our insights are robust to this effect.

[51] For brevity we do not report the Tables here, and they are available upon request.

References

Abreu, M. and V. Mendes, 2010, Financial literacy and portfolio diversification, *Quantitative Finance* 10, No. 5, 515-528.

Agarwal, S., Skiba, P.M., and J. Tobacman, 2009, Payday Loans and Credit Cards: New Liquidity and Credit Scoring Puzzles?, *NBER Working Paper* No. 14659.

Agarwal, S., Driscoll, J.C., and D.I. Laibson, 2013, Optimal Mortgage Refinancing: A Closed Form Solution, *Journal of Money, Credit and Banking* 45, No. 4, 591-622.

Andersen, S., Campbell, J.Y., Meisner-Nielsen, K., and T. Ramadorai, 2014, Inattention and Inertia in Household Finance: Evidence from the Danish Mortgage Market, *SSRN Working Paper*.

Archer, W.R., Ling, D.C., and G.A. McGill, 1996, The effect of income and collateral constraints on residential mortgage terminations, *Regional Science and Urban Economics* 26, No. 3, 235-261.

Behrman, J.R., Mitchell, O.S., Soo, C.K., and D. Bravo, 2012, How financial literacy affects household wealth accumulation, *American Economic Review* 102, No. 3, 300-304.

Bennett, P., Peach, R., and S. Peristiani, 2001, Structural Change in the Mortgage Market and the Propensity to Refinance, *Journal of Money, Credit and Banking* 33, No. 4, 955-975.

Bernheim, D., 1995, Do households appreciate their financial vulnerabilities? An analysis of actions, perceptions, and public policy, *Tax Policy and Economic Growth, American Council for Capital Formation*, Washington, DC, 1-30.

Calvet, L., Campbell, J., and P. Sodini, 2007, Down or out: assessing the welfare costs of household investment mistakes, *Journal of Political Economy* 115, No. 5, 707-747.

Calvet, L., Campbell, J., and P. Sodini, 2009, Measuring the financial sophistication of households, *American Economic Review* 99, No. 2, 393-398.

Campbell, J.Y., 2006, Household finance, *Journal of Finance* 61, No. 4, 1553-1604.

Campbell, J. and J. Cocco, 2003, Household Risk Management and Optimal Mortgage Choice, *Quarterly Journal of Economics* 118, No. 4, 1449-1494.

Chang, Y. and A. Yavas, 2009, Do Borrowers Make Rational Choices on Points and Refinancing?, *Real Estate Economics* 37, No. 4, 638-658.

Christelis, D., Jappelli, T., and M. Padula, 2010, Cognitive abilities and portfolio choice, *European Economic Review* 54, No. 1, 18-38.

Clark, R.L., Morrill, M.S., and S.G. Allen, 2012, Effectiveness of employer-provided financial information: Hiring to retiring, *American Economic Review* 102, No. 3, 314-318.

Cole, S., Paulson, A., and G.K. Shastry, forthcoming, Smart Money? The Effect of Education on Financial Outcomes, *Review of Financial Studies*.

Cole, S., Thomas, S., and B. Zia, 2011, Prices or knowledge? What drives demand for financial services in emerging markets?, *Journal of Finance* 66, No. 6, 1933-1967.

Disney, R. and Gathergood, J., 2013, Financial literacy and consumer credit portfolios, *Journal of Banking and Finance* 37, No. 7, 2246-2254.

Downing, C., Stanton, R., and N. Wallace, 2005, An Empirical Test of a Two-Factor Mortgage Valuation Model: How Much Do House Prices Matter?, *Real Estate Economics* 33, No. 4, 681-710.

Dyck, A., Moss, D., and L. Zingales, 2008, Media versus special interests, *NBER Working Paper* No. 14360.

Dyck, A., Volchkova, N., and L. Zingales, 2008, The corporate governance role of the media: Evidence from Russia, *Journal of Finance* 63, No. 3, 1093-1135.

European Central Bank, 2009, Housing Finance in the Euro Area, *Structural Issue Report*.

Fornero, E. and C. Monticone, 2011, Financial literacy and pension plan participation in Italy, *Journal of Pension Economics and Finance* 10, No. 4, 547-564.

Gathergood, J., 2012, Self-control, financial literacy and consumer over-indebtedness, *Journal of Economic Psychology* 33, No. 3, 590-602.

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

Giliberto, S.M., and T.G. Thibodeau, 1989, Modeling Conventional Residential Mortgage Refinancing, *Journal of Real Estate Finance and Economics* 2, No. 1, 285-299.

Green, R.K., and M. LaCour-Little, 1999, Some truths about ostriches: Who doesn't prepay their mortgages and why they don't, *Journal of Housing Economics* 8, No. 3, 233-248.

Guiso, L. and T. Jappelli, 2009, Financial Literacy and Portfolio Diversification, *Working Paper*, Centre for Studies in Economics and Finance, Naples, Italy.

Gurun, U.G. and A. W.Butler, 2012, Don't Believe the Hype: Local Media Slant, Local Advertising, and Firm Value, *Journal of Finance* 67, No. 2, 561-598.

Hastings, J.S. and L. Tejeda-Ashton, 2008, Financial literacy, information, and demand elasticity: Survey and experimental evidence from Mexico, *NBER Working Paper*, No. 14538.

Hastings, J.S. and O.S. Mitchell, 2011, How financial literacy and impatience shape retirement wealth and investment behaviors, *NBER Working Paper* No. 16740.

Heckman, J., 1979, Sample Selection Bias as a Specification error, *Econometrica* 47, 153-161.

Keys, B., J, Pope, D.G., and J.C. Pope, 2014, Failure to Refinance, *Working Paper*.

Imbens, G.W., 2004, Nonparametric estimation of average treatment effects under exogeneity: A review, *Review of Economics and Statistics* 86, No. 1, 4-29.

Jappelli, T., 2010, Economic Literacy: An International Comparison, *Economic Journal* 120, No. 548, F429-F451.

Jappelli, T. and M. Padula, 2013, Investment in financial literacy and saving decisions, *Journal of Banking & Finance* 37, No. 8, 2779-2792.

Jiang, W., Nelson, A.A., and E. Vytlacil, 2014, Securitization and Loan Performance: Ex Ante and Ex Post Relations in the Mortgage Market, Review *of Financial Studies* 27, No. 2, 454-483.

Johnson, J.L., Ellstrand, A.E., Dalton, D.R., and C.M. Dalton, 2005, The influence of the financial press on stockholder wealth: The case of corporate governance, *Strategic Management Journal* 26, No. 5, 461-471.

Lusardi, A., 2008, Household saving behavior: The role of financial literacy, information, and financial education programs, *NBER Working Paper* No. 13824.

Lusardi, A. and O.S. Mitchell, 2007, Baby boomer retirement security: The roles of planning, financial literacy, and housing wealth, Journal *of monetary Economics* 54, No. 1, 205-224.

Lusardi, A. and O.S. Mitchell, 2008, Planning and Financial Literacy: How Do Women Fare?, *American Economic Review* 98, No. 2, 413-417.

Lusardi, A. and O.S. Mitchell, 2009, How ordinary consumers make complex economic decisions: Financial literacy and retirement readiness, *NBER Working Paper* No. 15350.

Lusardi, A. and O.S. Mitchell, 2011, Financial literacy and planning: Implications for retirement wellbeing, *NBER Working Paper* No. 17078).

Lusardi, A. and P. Tufano, 2009, Debt literacy, financial experiences, and overindebtedness, *NBER Working Paper* No. 14808.

Magri, S. and R. Pico, 2011, The rise of risk-based pricing of mortgage interest rates in Italy, *Journal of Banking & Finance* 35, 1277-1290.

Klapper, L., A. Lusardi, and G. A. Panos, 2013, Financial literacy and its consequences: Evidence from Russia during the financial crisis, *Journal of Banking & Finance* 37, No. 10, 3904-3923.

Rosenbaum, P.R. and D.B. Rubin, 1983, The central role of the propensity score in observational studies for causal effects, *Biometrika* 70, No. 1, 41-55.

Santos, E. and M. Abreu, 2013, Financial literacy, financial behaviour and individuals' over-indebtedness, *Working Paper*, Technical University of Lisbon, Portugal.

Shiller, R.J., 1999, Human Behavior and the Efficiency of the Financial System, in J. Taylor, and M. Woodford, eds., *Handbook of Macroeconomics*, Elsevier, Amsterdam.

Stango, V. and J. Zinman, 2009, Exponential growth bias and household finance, *Journal of Finance* 64, No. 6, 2807-2849.

Tetlock, P., 2007, Giving Content to Investor Sentiment: The Role of Media in the Stock Market, *Journal of Finance* 62, No. 3, 1139-1168.

Tversky, A. and D. Kahneman, 1974, Judgment under Uncertainty: Heuristics and Biases, *Science* 185, No. 4157, 1124-1131.

Van de Ven, W.P.M.M. and B.M.S. Van Pragg, 1981, The Demand for Deductibles in Private Health Insurance: A Probit Model with Sample Selection, *Journal of Econometrics* 17, No. 2, 229-252.

van Rooij, M., Lusardi, A., and R. Alessie, 2011, Financial literacy and stock market participation, *Journal of Financial Economics* 101, No. 2, 449-472.

Appendix – MiFID questionnaire

Knowledge and experience in investing

	Knowledge What do you think is the level of your knowledge with respect to the following financial instruments? (0 = none, 1 = low, 2 = medium, 3 = high)				**Experience** How many purchases/ subscriptions have you made with the following investment instruments?			
	0	**1**	**2**	**3**	**0**	**1-3**	**4-6**	**>6**
Debt Products								
a) Treasury bonds								
b) Corporate bonds and convertible bonds								
c) Structured bonds								
d) Subordinated bonds								
Equity Products								
a) Stocks								
b) Others								
Bank Products								
a) Certificate of deposits								
b) Repurchase agreements								
c) Securities lending								
Funds								
a) Mutual funds								
b) Exchange traded funds (ETFs)								
c) Real estate funds								
d) Hedge funds								
Financial Derivatives								
a) ETC								
b) Certificates								
c) Equity and covered warrants								
d) Options and futures								

Level of education, professional experience in finance

	YES	**NO**
Have you studied a field that deals with financial services?		
Do you have working experience in financial services?		

Risk propensity

Which of the following goals corrispond to your desired risk propensity:

a) Very high return with the risk of large losses	
b) High return with the risk of a medium loss	
c) Medium return with the risk of a medium/low loss	
d) Low return with the risk of a low loss	

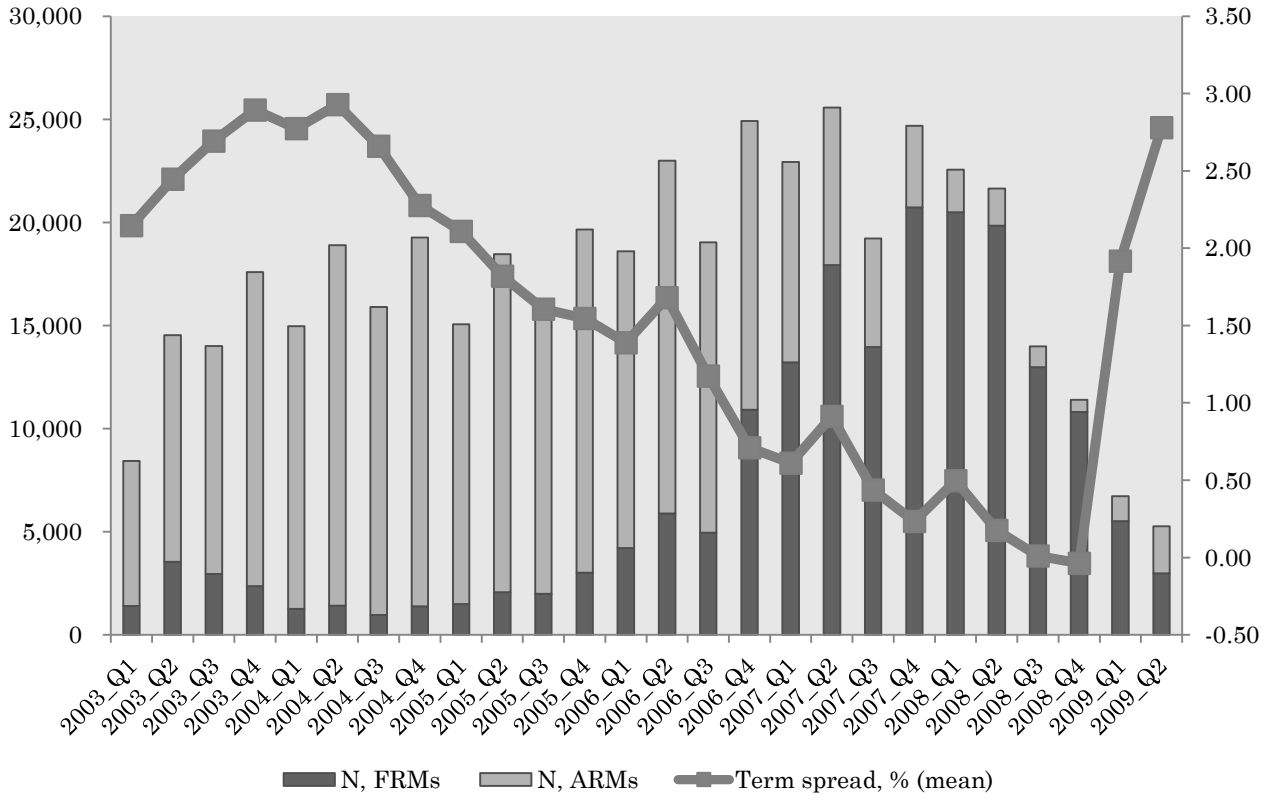

Figure 1 – *Distribution of mortgages over time*. The figure depicts the number of mortgage loans (rectangles, left axis), divided by FRMs *vs*. ARMs, and the percent *Term Spread* (solid line, right axis), that is the slope of the yield curve (i.e., the difference between 20y-swap rate and 3m-Euribor) at the initiation of the mortgage, as a function of the quarter of inception, from 2003_Q1 (year 2003, first quarter), to 2009_Q2 (year 2009, second quarter).

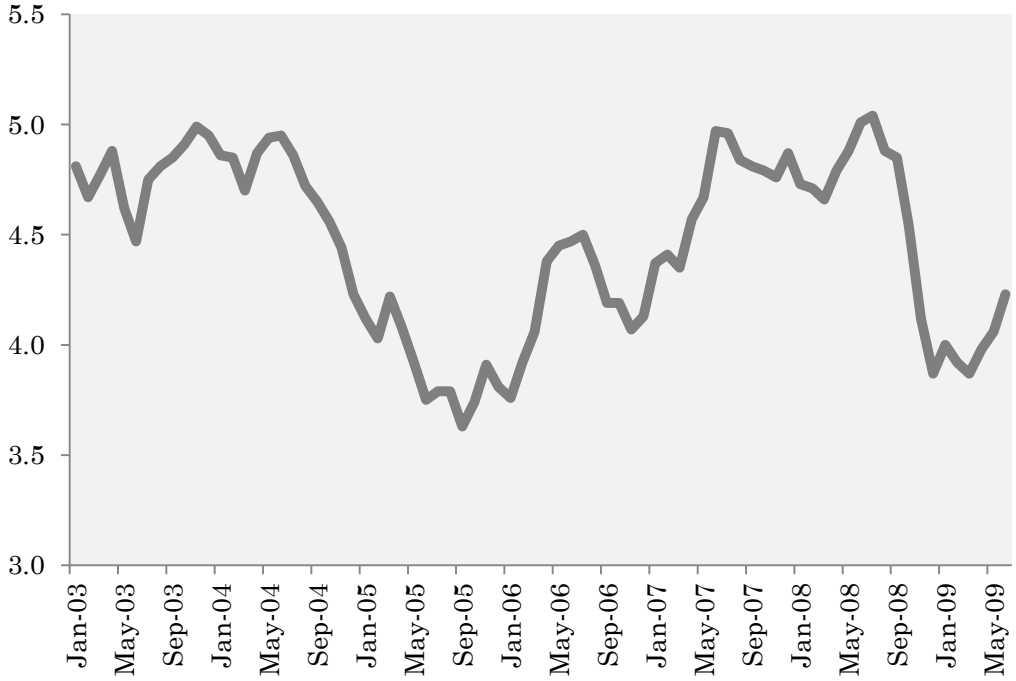

Figure 2 – *Pattern of interest rates*. The figure depicts the pattern of 20-year swap rate (Eurirs), from January, 2003 to June, 2009. Numbers are expressed in percentage.

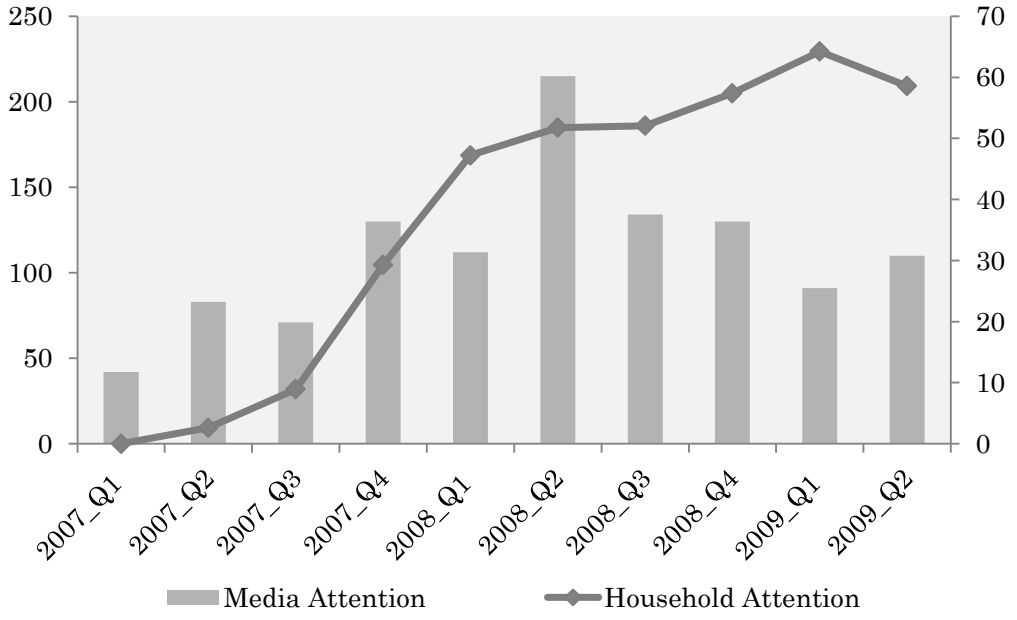

Figure 3 – *Pattern of media and household attention*. The figure depicts two proxies for media and household attention relative to the new legislation on mortgage refinancing passed on 1st of February, 2007, for each quarter, from 2007_Q1 (year 2007, first quarter) to 2009_Q2 (year 2009, second quarter). *Media Attention* (left axis) refers to the number of hits on the headlines and the body text of all Italian newspapers (source: Factiva) of the following term (in Italian): "subrogation"; *Household Attention* (right axis) refers to the number of web searches (standardized at 100 at the maximum during the period) of the word "subrogation" (in Italian).

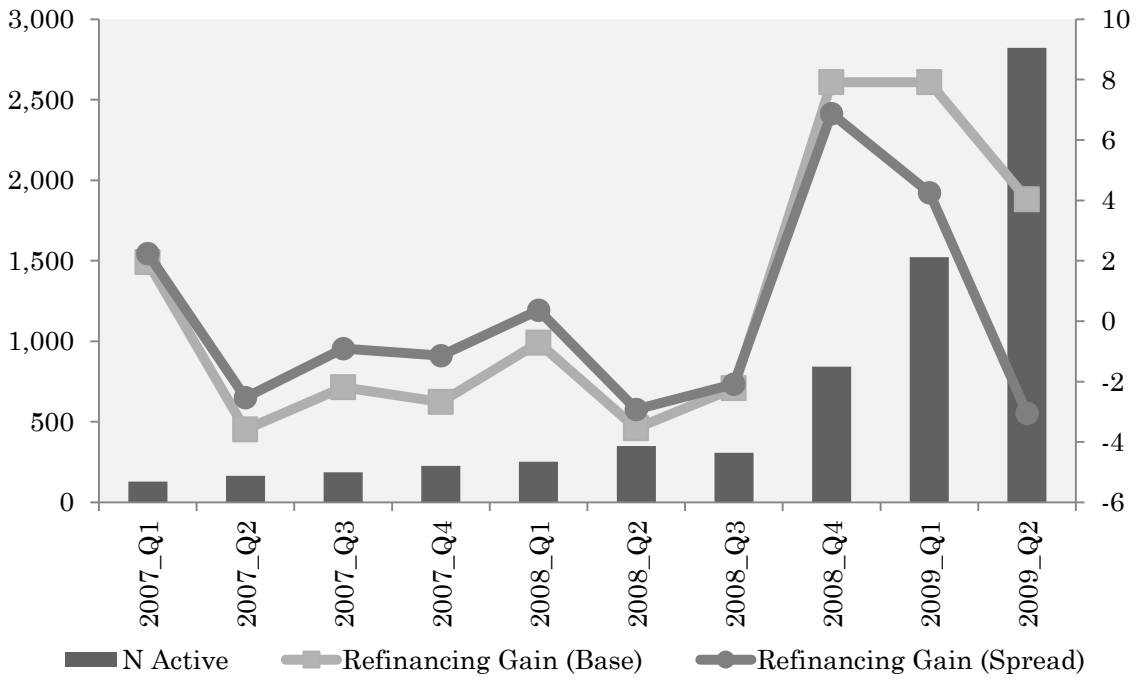

Figure 4 – *Refinancing gains*. The figure depicts the number of active borrowers (left axis), and the potential gains from refinancing, expressed as a percentage of the loan amount (right axis). Potential gains from refinancing are computed as the present value of the difference between future instalments from the original mortgage (at the contractual IRS rate) minus future instalments computed according to the IRS rate prevailing at each quarter (considering an amortized mortgage with constant instalments), divided by the residual amount of the loan. Figures start from 2007_Q1 (year 2007, first quarter) and end in 2009_Q2 (year 2009, second quarter).

Year	*N*	*Amount* , €	*Maturity* , years	*Spread* , % (*)	*LTV* , %	*No. Guarantors*
2003	2,891	80,433.59	19.17	100.00	66.25	0.51
2004	2,553	74,566.69	16.23	109.87	60.87	0.43
2005	5,171	88,110.15	17.48	99.87	58.43	0.75
2006	22,177	103,885.30	19.85	85.67	58.52	1.00
2007	53,382	115,967.90	22.46	79.17	61.51	0.92
2008	52,793	119,071.80	23.07	85.66	61.00	0.59
2009	8,008	112,021.80	22.36	129.94	58.34	0.45
Total	146,975	112,646.40	21.93	86.66	60.68	0.77

Table 1 – *Distribution of mortgages by year of inception*. The table reports the distribution of fixed-rate mortgage loans by year of origination. *Amount* is the average principal of the loan (in Euros), *Maturity* is the average length of the loan at inception (in years), *Spread* is the average difference between the loan rate and the swap rate (Eurirs) for the same maturity, *LTV* is the average loan-to-value, i.e. the ratio between the principal of the loan and the estimated value of the real property, *No. Guarantors* is the average number of guarantors of the loan. (*) For confidentiality reasons, the spread over the base rate (*Spread*) has been normalized at 100 at the beginning of the time period (year 2003). Hence, *Spread* has to be interpreted as the percentage increase or decrease of the credit spread over time relative to year 2003 figure.

Variable	*N*	*Mean*	*SD*	*Min*	*Q1*	*Median*	*Q3*	*Max*
Active , %	146,975	4.23	20.14	.	.	.	.	.
Efficient , %	146,975	81.24	39.04	.	.	.	.	.
Amount , €	146,975	112,646	68,416	1,422	70,000	100,000	140,000	4,000,000
Maturity , y	146,975	21.93	7.06	1.50	15.00	20.00	30.00	40.00
Spread (*)	124,623	86.6	28.2	5.4	66.0	77.6	100.9	275.6
LTV , %	146,975	60.7	20.4	11.9	45.3	66.8	78.4	100.0
No. Guarantors	146,975	0.8	1.0	0.0	0.0	0.0	1.0	5.0
Age	145,622	38.7	10.0	21	31	37	44	67
Income (pred.), €	144,107	1,490	332	939	1,350	1,488	1,504	3,569
Wealthy, %	144,655	10.89	31.15	.	.	.	.	.
Man , %	145,618	68.00	46.65	.	.	.	.	.
Graduate , %	144,164	12.63	33.22	.	.	.	.	.
Foreign , %	145,178	9.92	29.89	.	.	.	.	.
Population	146,930	679,446	780,225	98	29,034	205,535	1,324,110	2,761,477
North , %	146,948	57.68	49.41	.	.	.	.	.
Center , %	146,948	15.05	35.76	.	.	.	.	.
South , %	146,948	27.26	44.53	.	.	.	.	.
IRS at Origination , %	146,975	4.51	0.40	2.71	4.25	4.66	4.82	5.10
Interest Differential , %	146,975	-0.43	0.40	-2.06	-0.72	-0.56	-0.18	0.64
Term Spread , %	146,975	0.66	0.71	-0.82	0.19	0.51	0.92	3.12
Fixed Income Awareness	15,294	1.83	0.72	0.00	1.25	2.00	2.00	3.00
Bank Products Awareness	15,294	1.46	0.66	0.00	1.00	1.33	2.00	3.00
Derivatives Awareness	15,294	0.19	0.51	0.00	0.00	0.00	0.00	3.00
Fixed Income Experience, %	15,294	63.20	48.23	.	.	.	.	.
Bank Products Experience, %	15,294	50.95	49.99	.	.	.	.	.
Derivatives Experience, %	15,294	5.61	23.01	.	.	.	.	.
Economics Background, %	15,092	9.29	29.03	.	.	.	.	.
Financial Expertise, %	15,092	3.70	18.89	.	.	.	.	.
Risk Averse, %	15,294	36.50	48.14	.	.	.	.	.

Table 2 – *Descriptive statistics*. The table reports descriptive statistics for the whole sample of fixed-rate mortgage loans. *Active* is a dummy taking the value of 1 if the mortgage has been actively refinanced, *Efficient* is a dummy taking 1 if the mortgage exhibits positive refinancing efficiency (i.e., the difference between the IRS rate at the end of June, 2009 and the IRS rate at the inception of the mortgage is negative), *Amount* is the principal of the loan (in euros), *Maturity* is the length of the loan at inception (in years), *Spread* is the spread over the IRS rate paid by the borrower, (*) for confidentiality reasons, all numbers have been rescaled by the average spread in year 2003 and multiplied by 100, *LTV* is the loan-to-value, i.e. the ratio between the principal of the loan and the estimated value of the real property (in percentage), *No. Guarantors* is the number of guarantors of the loan, *Age* is the age of the borrower at the inception of the mortgage, *Income (pred.)* is the predicted monthly income of the borrower from median regression of income on borrower's occupation, age and gender (please refer to the body of the paper for details), *Wealthy* is a dummy variable taking 1 if the borrower does not belong to the segment "mass market" (< €100,000 financial assets), *Man* is a dummy variable taking 1 if the borrower's gender is male, *Graduate* is a dummy variable taking 1 if the borrower has graduated from a university, *Foreign* is a dummy variable taking 1 if the borrower's nationality is other than Italian, *Population* is the resident population in the urban area where the bank is located (by zip code), *North* (resp. *Center* and *South*) is a dummy variable taking 1 if the bank branch is located in a region of the North (resp. Center and South) of Italy, *IRS at Origination* is the fixed base rate of the loan mortgage, *Interest Differential* is the difference between the (average monthly) IRS rate at the end of June, 2009 and the IRS rate at the inception of the mortgage, *Term Spread* is the slope of the yield curve (i.e., the difference between 20y-swap rate and 3m-Euribor) at the inception of the mortgage, *Fixed Income Awareness*, *Bank Products Awareness*, and *Derivatives Awareness* are financial literature indicators and measure the degree of awareness (0 = none, 1 = low, 2 = medium, 3 = high) of fixed income instruments, bank products, and derivatives instruments, respectively, *Fixed Income Experience*, *Bank Products Investment*, and *Derivatives Experience* are dummy variables taking 1 if the borrower has ever invested in fixed income instruments, bank products, and derivatives instruments, respectively, *Economics Background* is a dummy variable taking 1 if the borrower has an educational background (college diploma) in a field related to Economics or Finance, *Financial Expertise* is a dummy variable taking 1 if the borrower's current or past job is related to finance or financial markets, *Risk Averse* is a dummy variable taking 1 if the borrower's profile can be described as "prudent" or "cautious" (*vs*. "balanced" and "dynamic").

Quarter	*Refinancing Gain (Base)*, %	*Refinancing Gain (Spread)*, %	*Media Attention*	*Household Attention*
2007_Q1	1.95	2.24	42	0
2007_Q2	-3.59	-2.53	83	3
2007_Q3	-2.19	-0.91	71	9
2007_Q4	-2.68	-1.14	130	29
2008_Q1	-0.71	0.37	112	47
2008_Q2	-3.55	-2.92	215	52
2008_Q3	-2.22	-2.09	134	52
2008_Q4	7.91	6.87	130	57
2009_Q1	7.91	4.25	91	64
2009_Q2	4.03	-3.05	110	59

Table 3 – *Descriptive statistics of time-varying variables*. The table reports average descriptive statistics of time-varying variables for the whole sample of fixed-rate mortgage loans. *Refinancing Gain (Base)* is the potential refinancing gain as a percentage of the loan amount, computed as the present value of the difference between future instalments from the original mortgage (at the contractual IRS rate) minus future instalments according to the IRS rate prevailing at each quarter (considering an amortized mortgage with constant instalments), divided by the residual principal of the loan; *Refinancing Gain (Spread)* is the same variable as before but the contractual fixed rate includes the credit spread, and the potential refinancing rate includes the average credit spread of the quarter under consideration; *Media Attention* (left axis) refers to the number of hits on the headlines and the body text of all Italian newspapers (source: Factiva) of the following term (in Italian): "subrogation"; *Household Attention* is the number of web (Google) searches (standardized at 100 at the maximum during the period) of the word "subrogation" (in Italian).

	Refinancers		*Non-Refinancers*		*Mean Difference*	*t-Statistic*	*ATT Difference*	*t-Statistic*
	N	*Mean*	*N*	*Mean*				
Amount , €	6,223	120,025	140,752	112,320	7,704	7.69 ***	1,657	1.52
Maturity , y	6,223	22.54	140,752	21.90	0.64	6.87 ***	-0.04	-0.35
Spread (*)	5,546	91.53	119,077	86.40	5.13	13.86 ***	-0.46	-1.17
LTV , %	6,223	0.63	140,752	0.61	0.03	10.05 ***	0.00	-0.30
No. Guarantors	6,223	0.38	140,752	0.78	-0.40	-38.61 ***	-0.01	-0.44
Age	6,196	37.6	139,426	38.7	-1.1	-8.76 ***	-1.0	-7.25 ***
Income (pred.), €	6,118	1,511.1	137,989	1,489.4	21.7	4.80 ***	11.4	2.25 **
Wealthy, %	6,086	8.87	138,569	10.98	-2.11	-5.63 ***	-1.6	-4.00 ***
Man , %	6,196	71.34	139,422	67.85	3.48	5.92 ***	1.56	2.40 **
Graduate , %	6,127	14.95	138,037	12.53	2.42	5.22 ***	2.28	4.33 ***
Foreign , %	6,177	7.90	139,001	10.01	-2.11	-5.98 ***	-4.58	-11.20 ***
Population	6,214	840,842	140,716	672,319	168,524	16.74 ***	56,096	5.14 ***
North , %	6,214	67.1	140,734	57.3	9.80	16.06 ***	2.24	3.38 ***
Center , %	6,214	14.9	140,734	15.1	-0.14	-0.31	1.37	2.74 ***
South , %	6,214	18.0	140,734	27.7	-9.66	-19.25 ***	-3.61	-6.78 ***
IRS at Origination , %	6,223	4.55	140,752	4.51	0.04	7.04 ***	0.00	-0.06
Interest Differential , %	6,223	-0.47	140,752	-0.43	-0.05	-9.84 ***	0.00	0.00
Term Spread , %	6,223	0.92	140,752	0.65	0.26	24.42 ***	0.33	26.85 ***

Table 4 – *Mean statistics of refinancers* vs. *non-refinancers*. The table reports the mean of the considered variables for the whole sample of fixed-rate mortgage loans, distinguishing between refinanced *vs*. non-refinanced mortgages. *Amount* is the principal of the loan (in euros), *Maturity* is the length of the loan at inception (in years); *Spread* is the spread over the IRS rate paid by the borrower, (*) for confidentiality reasons, all numbers have been rescaled by the average spread in year 2003 and multiplied by 100; *LTV* is the loan-to-value, i.e. the ratio between the principal of the loan and the estimated value of the real property (in percentage); *No. Guarantors* is the number of guarantors of the loan; *Age* is the age of the borrower at the inception of the mortgage; *Income (pred.)* is the predicted monthly income of the borrower from median regression of income on borrower's occupation, age and gender (please refer to the body of the paper for details); *Wealthy* is a dummy variable taking 1 if the borrower does not belong to the segment "mass market" (< €100,000 financial assets); *Man* is a dummy variable taking 1 if the borrower's gender is male; *Graduate* is a dummy variable taking 1 if the borrower has graduated from a university; *Foreign* is a dummy variable taking 1 if the borrower's nationality is other than Italian; *Population* is the resident population in the urban area where the bank is located (by zip code); *North* (resp. *Center* and *South*) is a dummy variable taking 1 if the bank branch is located in a region of the North (resp. Center and South) of Italy; *IRS at Origination* is the fixed base rate of the loan mortgage; *Interest Differential* is the difference between the (average monthly) IRS rate at the end of June, 2009 and the IRS rate at the inception of the mortgage; *Term Spread* is the slope of the yield curve at the inception of the mortgage (i.e., the difference between the swap rate with the same maturity as the mortgage and 3m-Euribor). Mean Difference (and its t-statistic) refer to the difference between refinancers and non-refinancers borrowers. ATT (and its t-statistic) represents the average treatment effect on treated, obtained through the 10 nearest neighbours (with replacement) propensity score matching methodology (variables used for matching are mortgage-specific characteristics: *Amount*, *Maturity*, *Spread*, *LTV*, *No. Guarantors*, *IRS at Origination* and *Interest Differential*) (Rosenbaum and Rubin, 1983; Imbens, 2000). ***, **, *, denote statistical significance at 1, 5 and 10 percent level, respectively.

Dependent variable: *Active*	(1)	(2)	(3)	(4)
Interest Differential	-0.0288***	-0.0294***	-0.0288***	-0.0270***
	(0.002)	(0.002)	(0.002)	(0.002)
Ln(Amount)	0.0208***	0.0218***	0.0210***	0.0217***
	(0.001)	(0.001)	(0.001)	(0.001)
Maturity	0.0007***	0.0005***	0.0003***	0.0005***
	(0.000)	(0.000)	(0.000)	(0.000)
LTV	-0.0211***	-0.0214***	-0.0209***	-0.0233***
	(0.003)	(0.003)	(0.003)	(0.003)
No. Guarantors	-0.0199***	-0.0201***	-0.0192***	-0.0165***
	(0.001)	(0.001)	(0.001)	(0.001)
Man		0.0067***	0.0059***	0.0100***
		(0.001)	(0.001)	(0.002)
Wealthy		-0.0065***	-0.0056***	-0.0045***
		(0.002)	(0.002)	(0.002)
Ln(Income)		-0.0107**	-0.0095**	-0.3062**
		(0.004)	-0.0288***	(0.128)
Age		-0.0003***	-0.0003***	-0.0028***
		(0.000)	(0.000)	(0.000)
Foreign		-0.0110***	-0.0126***	-0.0116***
		(0.002)	(0.002)	(0.002)
Graduate		0.0097***	0.0091***	0.0119***
		(0.002)	(0.002)	(0.002)
North			0.0176***	0.0182***
			(0.001)	(0.001)
Center			0.0146***	0.0146***
			(0.002)	(0.002)
Ln(Population)			0.0005*	0.0005*
			(0.000)	(0.000)
Ln(Income) Squared				0.0196**
				(0.008)
Age Squared				0.0000***
				(0.000)
Constant				1.1925**
				(0.481)
Origination Year FEs	Yes	Yes	Yes	Yes
Observations	146,975	141,799	141,782	141,782
Mc-Fadden Pseudo-R^2	0.0809	0.0826	0.0867	
Adjusted-R^2				0.0355

Table 5 – *Likelihood of active borrower.* The table reports in models 1 to 3 the average marginal effects (AMEs) of a probit regression of *Active*, i.e. a dummy taking 1 if the mortgage has been refinanced by the borrower, on the chosen explanatory variables. *Interest Differential* is the difference between the (average monthly) IRS rate at the end of June 2009 and the IRS rate at the inception of the mortgage; *Ln(Amount)* is the natural logarithm of the principal of the loan; *Maturity* is the length of the loan at inception; *LTV* is the loan-to-value, i.e. the ratio between the principal of the loan and the estimated value of the real property; *No. Guarantors* is the number of guarantors of the loan; *Man* is a dummy variable taking 1 if the borrower's gender is male; *Wealthy* is a dummy variable taking 1 if the borrower owns €100,000 financial assets or more; *Ln(Income) (Ln(Income) Squared)* is the natural logarithm (squared natural logarithm) of the predicted monthly income of the borrower from median regression of income on borrower's occupation, age and gender; *Age* (*Age squared*) is the age (squared age) of the borrower at the inception of the mortgage; *Foreign* is a dummy variable taking 1 if the borrower's nationality is other than Italian; *Graduate* is a dummy variable taking 1 if the borrower has graduated from a university; *North* (resp. *Center*) is a dummy variable taking 1 if the bank branch is located in a region of the North (resp. Center) of Italy; *Ln(Population)* is the natural logarithm of the resident population in the urban area where the bank is located (by zip code). Model 4 reports the results of a linear probability model (LPM) of *Active* on the chosen explanatory variables, including quadratic values of *Age* and *Ln(Income)*. Standard errors, in parentheses, are bootstrapped based on 1,000 replications. ***, **, * indicate statistical significance at 1, 5 and 10 percent level, respectively.

Failure variable: *Active*	(1)	(2)	(3)	(4)	(5)	(6)
Refinancing Gain (Base)	8.8889***	2.9334***			6.4094***	
	(0.194)	(0.328)			(0.215)	
Refinancing Gain (Spread)			1.1809***	0.9552**		0.1965
			(0.255)	(0.474)		(0.263)
Media Attention	0.5061***	0.7762***	0.2377***	0.6507***		
	(0.029)	(0.038)	(0.023)	(0.042)		
Household Attention					0.5212***	0.7087***
					(0.024)	(0.035)
Ln(Amount)	0.6501***	0.7360***	0.7941***	0.5891***	0.5912***	0.6662***
	(0.030)	(0.036)	(0.028)	(0.050)	(0.034)	(0.033)
LTV	-0.8668***	-1.0759***	-0.6380***	-1.0314***	-0.7487***	-0.5903***
	(0.079)	(0.094)	(0.078)	(0.127)	(0.088)	(0.087)
No. Guarantors	-0.5962***	-0.6241***	-0.6145***	-0.6612***	-0.6285***	-0.6458***
	(0.021)	(0.023)	(0.021)	(0.034)	(0.022)	(0.022)
Man		0.2879***		0.2478***	0.2575***	0.2606***
		(0.055)		(0.071)	(0.050)	(0.050)
Wealthy		-0.0841		0.0164	-0.1120**	-0.1539***
		(0.052)		(0.073)	(0.049)	(0.049)
Ln(Income)		-10.9165***		-13.8733***	-9.2306***	-7.7853**
		(3.634)		(4.722)	(3.316)	(3.268)
Ln(Income) Squared		0.7029***		0.9016***	0.5931***	0.4946**
		(0.238)		(0.310)	(0.218)	(0.215)
Age		-0.0909***		-0.1071***	-0.0872***	-0.0896***
		(0.010)		(0.013)	(0.010)	(0.010)
Age Squared		0.0010***		0.0012***	0.0009***	0.0009***
		(0.000)		(0.000)	(0.000)	(0.000)
Foreign		-0.3605***		-0.3317***	-0.3700***	-0.3185***
		(0.061)		(0.075)	(0.056)	(0.056)
Graduate		0.3059***		0.3654***	0.2919***	0.2692***
		(0.055)		(0.072)	(0.051)	(0.051)
North		0.3614***		0.5309***	0.5606***	0.6052***
		(0.040)		(0.058)	(0.039)	(0.039)
Center		0.3811***		0.7559***	0.5454***	0.5667***
		(0.053)		(0.074)	(0.050)	(0.050)
Ln(Population)		-0.0645***		-0.0347***	-0.0230***	-0.0194***
		(0.008)		(0.011)	(0.007)	(0.007)
Observations	908,154	878,598	908,154	878,598	878,598	878,598
Number of ID	143,374	138,411	143,374	138,411	138,411	138,411
McFadden Pseudo-R^2	0.03002	0.02228	0.01785	0.02324	0.04004	0.03414

Table 6 – *Determinants of debt refinancing over time.* The table reports the coefficients of a Cox proportional-hazards regression for the probability to refinance the mortgage loan. The failure variable is *Active*, i.e. a dummy taking 1 if the mortgage has been refinanced by the borrower. *Refinancing Gain* is the potential refinancing gain as a percentage of the loan amount, computed as the present value of the difference between future instalments from the original mortgage (at the contractual IRS rate) minus future instalments according to the IRS rate prevailing at each quarter (considering an amortized mortgage with constant instalments), divided by the residual amount of the loan; *Refinancing Gain (Spread)* is the same variable as before, but the contractual fixed rate includes the credit spread, and the potential refinancing rate includes the average credit spread of the quarter under consideration; *Media Attention* refers to the (log of the) number of hits on the headlines and the body text of all Italian newspapers (source: Factiva) of the following term (in Italian): "subrogation;" *Household Attention* is the (log of 1 plus the) number of web (Google) searches (standardized at 100 at the maximum during the period) of the word "subrogation" (in Italian); *Ln(Amount)* is the natural logarithm of the principal of the loan; *LTV* is the loan-to-value, i.e. the ratio between the principal of the loan and the estimated value of the real property; *No. Guarantors* is the number of guarantors of the loan; *Man* is a dummy variable taking 1 if the borrower's gender is male; *Wealthy* is a dummy variable taking 1 if the borrower does not belong to the segment "mass market" (< €100,000 financial assets); *Ln(Income) (Ln(Income) Squared)* is the natural logarithm (squared natural logarithm) of the predicted monthly income of the borrower from median regression of income on borrower's occupation, age and gender; *Age* (*Age squared*) is the age (squared age) of the borrower at the inception of the mortgage; *Foreign* is a dummy variable taking 1 if the borrower's nationality is other than Italian; *Graduate* is a dummy variable taking 1 if the borrower has graduated from a university; *North* (resp. *Center*) is a dummy variable taking 1 if the bank branch is located in a region of the North (resp. Center) of Italy; *Ln(Population)* is the natural logarithm of the resident population in the urban area where the bank is located (by zip code); Standard errors, in parentheses, are bootstrapped based on 250 replications. ***, **, * indicate statistical significance at 1, 5 and 10 percent level, respectively.

	MIFID Subsample		*Non-MIFID Subsample*					
Socio-Demographic Variables	*N*	*Mean*	*N*	*Mean*	*Mean Difference*	*t-Statistic*	*ATT Difference*	*t-Statistic*
Age	15,219	41.3	130,403	38.4	2.9	32.7 ***	1.1	11.7 ***
Income (pred.), €	14,994	1,520.4	129,113	1,486.8	33.5	9.7 ***	14.9	4.1 ***
Wealthy, %	15,236	32.69	129,419	8.32	24.37	62.9 ***	22.36	47.8 ***
Man , %	15,219	65.07	130,399	68.34	-3.27	-8.0 ***	-2.12	-3.7 ***
Graduate , %	14,996	14.54	129,168	12.41	2.13	7.0 ***	0.16	0.5
Foreign , %	15,201	6.41	129,977	10.33	-3.92	-18.2 ***	-1.51	-6.6 ***
Population	15,290	413,220	131,640	710,368	-297,148	-49.2 ***	-275,483	-42.4 ***
North , %	15,294	49.99	131,654	58.58	-8.58	-20.1 ***	-4.14	-9.2 ***
Center , %	15,294	14.62	131,654	15.10	-0.48	-1.6	-2.22	-6.9 ***
South , %	15,294	35.39	131,654	26.32	9.07	22.4 ***	6.36	14.9 ***
	MIFID subsample: Refinancers		*MIFID subsample: Non-Refinancers*					
Financial Literacy Variables	*N*	*Mean*	*N*	*Mean*	*Mean Difference*	*t-Statistic*	*ATT Difference*	*t-Statistic*
Fixed Income Awareness	442	1.90	14,852	1.83	0.07	2.1 **	0.05	1.3
Bank Products Awareness	442	1.53	14,852	1.46	0.07	2.2 **	0.05	1.4
Derivatives Awareness	442	0.27	14,852	0.19	0.08	2.9 ***	0.06	2.1 **
Fixed Income Experience, %	442	65.84	14,852	63.12	2.71	1.2	1.55	0.6
Bank Products Experience, %	442	57.69	14,852	50.75	6.94	2.9 ***	5.59	2.2 **
Derivatives Experience, %	442	8.14	14,852	5.53	2.61	2.0 **	2.46	1.8 *
Economics Background, %	435	13.79	14,657	9.16	4.64	2.8 ***	3.41	2.0 **
Financial Expertise, %	435	5.06	14,657	3.66	1.39	1.3	0.86	0.7
Risk Averse, %	442	35.29	14,852	36.53	-1.24	-0.5	0.33	0.1

Table 7 – *Mean financial literacy statistics of refinancers* vs. *non-refinancers, and sub-sample matching variables*. The table reports the mean of socio-demographic variables for the whole sample of fixed-rate mortgage loans, distinguishing observations for which the MIFID questionnaire was available *vs.* those for which it was unavailable. *Age* is the age of the borrower at the inception of the mortgage; *Income (pred.)* is the predicted monthly income of the borrower from median regression of income on borrower's occupation, age and gender (please refer to the body of the paper for details); *Wealthy* is a dummy variable taking 1 if the borrower does not belong to the segment "mass market" (< €100,000 financial assets); *Man* is a dummy variable taking 1 if the borrower's gender is male; *Graduate* is a dummy variable taking 1 if the borrower has graduated from a university; *Foreign* is a dummy variable taking 1 if the borrower's nationality is other than Italian; *Population* is the resident population in the urban area where the bank is located (by zip code); *North* (resp. *Center* and *South*) is a dummy variable taking 1 if the bank branch is located in a region of the North (resp. Center and South) of Italy. ATT (and its t-statistic) represents the average treatment effect on treated, obtained through the 10 nearest neighbours (with replacement) propensity score matching methodology (variables used for matching are mortgage-specific characteristics: *Amount*, *Maturity*, *LTV*, *No. Guarantors*, *IRS at Origination* and *Interest Differential*) (Rosenbaum and Rubin, 1983; Imbens, 2000). The table also reports the mean of the financial literacy variables for the subsample of fixed-rate mortgage loans for which a MIFID questionnaire was available, distinguishing between refinanced *vs.* non-refinanced mortgages. *Fixed Income Awareness*, *Bank Products Awareness*, and *Derivatives Awareness* are financial literature indicators and measure the degree of awareness (0 = none, 1 = low, 2 = medium, 3 = high) of fixed income instruments, bank products, and derivatives instruments, respectively, *Fixed Income Experience*, *Bank Products Experience*, and *Derivatives Experience* are dummy variables taking 1 if the borrower has ever invested in fixed income instruments, bank products, and derivatives instruments, respectively, *Economics Background* is a dummy variable taking 1 if the borrower has an educational background (college diploma) in a field related to Economics or Finance, *Financial Expertise* is a dummy variable taking 1 if the borrower's current or past job is related to finance or financial markets, *Risk Averse* is a dummy variable taking 1 if the borrower's profile can be described as "prudent" or "cautious" (*vs.* "balanced" and "dynamic"). Mean Difference (and its t-statistic) for financial literature variables refer to the difference between active and passive borrowers. ATT (and its t-statistic) represents the average treatment effect on treated, obtained through the 10 nearest neighbours (with replacement) propensity score matching methodology (variables used for matching are mortgage-specific characteristics: *Amount, Maturity*, *LTV*, *No. Guarantors*, *IRS at Origination* and *Interest Differential*, and socio-demographic characteristics, i.e. *Age*, *Income*, *Wealthy*, *Man* and *Foreign*) (Rosenbaum and Rubin, 1983; Imbens, 2000).. ***, **, *, denote statistical significance at 1, 5 and 10 percent level, respectively.

Dependent variable: *Active*	(1)	(2)	(3)	(4)	(5)	(6)
Interest Differential	-0.0280***	-0.0269***	-0.0275***	-0.0265***	-0.0256***	-0.0260***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)
Ln(Amount)	0.0159***	0.0147***	0.0152***	0.0160***	0.0152***	0.0155***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Maturity	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LTV	-0.0306***	-0.0290***	-0.0295***	-0.0295**	-0.0282**	-0.0285**
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
No. Guarantors	-0.0230***	-0.0222***	-0.0227***	-0.0222***	-0.0216***	-0.0221***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
North	0.0145***	0.0139***	0.0143***	0.0135**	0.0130**	0.0134**
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Center	0.0087	0.0082	0.0086	0.0076	0.0072	0.0076
	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Ln(Population)	0.0027*	0.0025*	0.0027**	0.0028**	0.0027*	0.0029**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Fixed Income Awareness		0.0016			0.0020	
		(0.004)			(0.004)	
Bank Products Awareness		0.0013			0.0019	
		(0.004)			(0.004)	
Derivatives Awareness		0.0056			0.0048	
		(0.004)			(0.004)	
Fixed Income Experience			-0.0065			-0.0067
			(0.005)			(0.005)
Bank Products Experience			0.0107**			0.0111**
			(0.005)			(0.005)
Derivatives Experience			0.0080			0.0072
			(0.007)			(0.007)
Economics Background				0.0154**	0.0134**	0.0148**
				(0.007)	(0.007)	(0.007)
Financial Expertise				-0.0012	-0.0041	-0.0018
				(0.010)	(0.010)	(0.010)
Risk Averse				0.0010	0.0031	0.0015
				(0.004)	(0.004)	(0.004)
Origination Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations (first stage)	141,782	141,782	141,782	141,585	141,585	141,585
Observations (second stage)	14,924	14,924	14,924	14,727	14,727	14,727
Rho	-0.1156	-0.1060	-0.1123	-0.1154	-0.1071	-0.1132
Wald-Chi^2	217.63***	219.13***	224.11***	222.00***	222.53***	228.95***

Table 8 – *Likelihood of active borrower with financial literacy variables.* The table reports the average marginal effects of a probit regression of *Active*, i.e. a dummy taking 1 if the mortgage has been refinanced by the borrower, on the chosen explanatory variables, where a two-stage probit selection model has been estimated for the probability of having a MIFID questionnaire. Selection equation includes socio-demographic borrower's characteristics (*Man* is a dummy variable taking 1 if the borrower's gender is male; *Wealthy* is a dummy variable taking 1 if the borrower does not belong to the segment "mass market," i.e. < €100,000 financial assets; *Ln(Income)* is the natural logarithm of the predicted monthly income of the borrower from median regression of income on borrower's occupation, age and gender; *Age* is the age of the borrower at the inception of the mortgage; *Foreign* is a dummy variable taking 1 if the borrower's nationality is other than Italian) as instruments (along with the same explanatory variables included into the second-stage equation). Rho is the correlation coefficient between errors of the two regression models and its significance relates to a Wald/LR-Chi^2 statistic of independent equations. *Interest Differential* is the difference between the (average monthly) IRS rate at the end of June 2009 and the IRS rate at the inception of the mortgage; *North* (resp. *Center*) is a dummy variable taking 1 if the bank branch is located in a region of the North (resp. Center) of Italy; *Ln(Population)* is the natural logarithm of the resident population in the urban area where the bank is located (by zip code); *Fixed Income Awareness*, *Bank Products Awareness*, and *Derivatives Awareness* are financial literature indicators and measure the degree of awareness (0 = none, 1 = low, 2 = medium, 3 = high) of fixed income instruments, bank products, and derivatives instruments, respectively; *Fixed Income Experience*, *Bank Products Experience*, and *Derivatives Experience* are dummy variables taking 1 if the borrower has ever invested in fixed income instruments, bank products, and derivatives instruments, respectively; *Economics Background* is a dummy variable taking 1 if the borrower has an educational background (college diploma) in a field related to Economics or Finance; *Financial Expertise* is a dummy variable taking 1 if the borrower's current or past job is related to finance or financial markets; *Risk Averse* is a dummy variable taking 1 if the borrower's profile can be described as "prudent" or "cautious" (*vs.* "balanced" and "dynamic"). Standard errors, in parentheses, are bootstrapped based on 1,000 replications. ***, **, * indicate statistical significance at 1, 5 and 10 percent level, respectively.

Failure variable: *Active*	(1)	(2)	(3)	(4)	(5)	(6)
Refinancing Gain (Base)	10.9123***	10.9272***	10.8228***			
	(0.715)	(0.720)	(0.714)			
Refinancing Gain (Spread)				3.5118***	3.5200***	3.3654***
				(0.987)	(0.989)	(0.995)
Media Attention	0.1845*	0.1909**	0.1933**	0.0546	0.0597	0.0553
	(0.095)	(0.095)	(0.095)	(0.075)	(0.075)	(0.076)
Ln(Amount)	0.6164***	0.6212***	0.6530***	0.7569***	0.7634***	0.7927***
	(0.106)	(0.105)	(0.107)	(0.100)	(0.099)	(0.101)
LTV	-1.7966***	-1.7784***	-1.8067***	-1.6166***	-1.6059***	-1.6208***
	(0.283)	(0.280)	(0.289)	(0.279)	(0.278)	(0.287)
No. Guarantors	-0.7641***	-0.7656***	-0.7467***	-0.7677***	-0.7689***	-0.7500***
	(0.090)	(0.090)	(0.090)	(0.090)	(0.090)	(0.090)
North	0.4956***	0.4928***	0.4728***	0.6015***	0.6003***	0.5774***
	(0.126)	(0.126)	(0.126)	(0.123)	(0.124)	(0.124)
Center	0.3213*	0.3200*	0.2913	0.4225**	0.4245**	0.3884**
	(0.176)	(0.175)	(0.177)	(0.176)	(0.175)	(0.178)
Ln(Population)	-0.0062	-0.0039	-0.0004	-0.0195	-0.0182	-0.0134
	(0.025)	(0.025)	(0.026)	(0.025)	(0.025)	(0.025)
Fixed Income Awareness	-0.0013			-0.0125		
	(0.101)			(0.101)		
Bank Products Awareness	0.0623			0.0620		
	(0.108)			(0.109)		
Derivatives Awareness	0.0835			0.0783		
	(0.092)			(0.091)		
Fixed Income Experience		-0.1196			-0.1593	
		(0.129)			(0.128)	
Bank Products Experience		0.3372***			0.3329***	
		(0.123)			(0.123)	
Derivatives Experience		0.2080			0.2103	
		(0.183)			(0.183)	
Economics Background			0.3805**			0.4129**
			(0.164)			(0.163)
Financial Expertise			-0.1287			-0.1482
			(0.265)			(0.265)
Risk Averse			0.0494			0.0881
			(0.107)			(0.107)
Observations	96,667	96,667	95,457	96,667	96,667	95,457
Number of ID	14,944	14,944	14,745	14,944	14,944	14,745
Pseudo-R^2	0.05249	0.05374	0.05270	0.03453	0.03575	0.03484

Table 9 – *Determinants of debt refinancing over time with financial literacy variables.* The table reports the coefficients of a Cox proportional-hazards regression for the probability to refinance the mortgage loan. The failure variable is *Active*, i.e. a dummy taking 1 if the mortgage has been refinanced by the borrower. *Refinancing Gain* is the potential refinancing gain as a percentage of the loan amount, computed as the present value of the difference between future instalments from the original mortgage (at the contractual IRS rate) minus future instalments according to the IRS rate prevailing at each quarter (considering an amortized mortgage with constant instalments), divided by the residual amount of the loan; *Refinancing Gain (Spread)* is the same variable as before but the contractual fixed rate includes the credit spread, and the potential refinancing rate includes the average credit spread of the quarter under consideration; *Media Attention* refers to the (log of the) number of hits on the headlines and the body text of all Italian newspapers (source: Factiva) of the following term (in Italian): "subrogation;" *Ln(Amount)* is the natural logarithm of the principal of the loan; *LTV* is the loan-to-value, i.e. the ratio between the principal of the loan and the estimated value of the real property; *No. Guarantors* is the number of guarantors of the loan; *North* (resp. *Center*) is a dummy variable taking 1 if the bank branch is located in a region of the North (resp. Center) of Italy; *Ln(Population)* is the natural logarithm of the resident population in the urban area where the bank is located (by zip code); *Fixed Income Awareness*, *Bank Products Awareness*, and *Derivatives Awareness* are financial literature indicators and measure the degree of awareness (0 = none, 1 = low, 2 = medium, 3 = high) of fixed income instruments, bank products, and derivatives instruments, respectively; *Fixed Income Experience*, *Bank Products Experience*, and *Derivatives Experience* are dummy variables taking 1 if the borrower has ever invested in fixed income instruments, bank products, and derivatives instruments, respectively; *Economics Background* is a dummy variable taking 1 if the borrower has an educational background (college diploma) in a field related to Economics or Finance; *Financial Expertise* is a dummy variable taking 1 if the borrower's current or past job is related to finance or financial markets; *Risk Averse* is a dummy variable taking 1 if the borrower's profile can be described as "prudent" or "cautious" (*vs.* "balanced" and "dynamic"). Standard errors, in parentheses, are bootstrapped based on 250 replications. ***, **, * indicate statistical significance at 1, 5 and 10 percent level, respectively.

Dependent variable: *Active*	(1)	(2)	(3)	(4)
Interest Differential	-0.0266***	-0.0271***	-0.0262***	-0.0244***
	(0.002)	(0.002)	(0.002)	(0.002)
Ln(Amount)	0.0205***	0.0214***	0.0207***	0.0215***
	(0.001)	(0.001)	(0.001)	(0.001)
Maturity	0.0006***	0.0004***	0.0003**	0.0004***
	(0.000)	(0.000)	(0.000)	(0.000)
LTV	-0.0207***	-0.0209***	-0.0205***	-0.0232***
	(0.003)	(0.003)	(0.003)	(0.003)
No. Guarantors	-0.0198***	-0.0200***	-0.0192***	-0.0165***
	(0.001)	(0.001)	(0.001)	(0.000)
Man		0.0067***	0.0059***	0.0099***
		(0.001)	(0.001)	(0.002)
Wealthy		-0.0065***	-0.0056***	-0.0047***
		(0.002)	(0.002)	(0.002)
Ln(Income)		-0.0107**	-0.0095**	-0.2990**
		(0.004)	(0.004)	(0.131)
Age		-0.0003***	-0.0003***	-0.0028***
		(0.000)	(0.000)	(0.000)
Foreign		-0.0109***	-0.0124***	-0.0116***
		(0.002)	(0.002)	(0.002)
Graduate		0.0096***	0.0090***	0.0118***
		(0.002)	(0.002)	(0.002)
North			0.0173***	0.0180***
			(0.001)	(0.001)
Center			0.0142***	0.0141***
			(0.002)	(0.002)
Ln(Population)			0.0005*	0.0005*
			(0.000)	(0.000)
Ln(Income) Squared				0.0191**
				(0.009)
Age Squared				0.0000***
				(0.000)
Constant				1.1589**
				(0.491)
Origination Year FEs	Yes	Yes	Yes	Yes
Observations (first stage)	284,344	269,670	269,646	269,646
Observations (second stage)	146,975	141,799	141,782	141,782
Rho	0.0891***	0.1004***	0.1158***	0.1141***

Table 10 – *Likelihood of active borrower with selection (FRM* v. *ARM).* The table reports the average marginal effects of a probit regression of *Active*, i.e. a dummy taking 1 if the mortgage has been refinanced by the borrower, on the chosen explanatory variables, where a two-stage probit selection model has been estimated for the probability of *ex-ante* choosing a FRM instead of an ARM. Selection equation includes *Term Spread*, i.e. the slope of the yield curve at the inception of the mortgage (the difference between the swap rate with the same maturity as the mortgage and 3m-Euribor), as an instrument (along with the same explanatory variables included into the second-stage equation). Rho is the correlation coefficient between errors of the two regression models and its significance relates to a Wald/LR-Chi^2 statistic of independent equations. *Interest Differential* is the difference between the (average monthly) IRS rate at the end of June 2009 and the IRS rate at the inception of the mortgage; *Man* is a dummy variable taking 1 if the borrower's gender is male; *Wealthy* is a dummy variable taking 1 if the borrower does not belong to the segment "mass market" (< €100,000 financial assets); *Ln(Income)* is the natural logarithm of the predicted monthly income of the borrower from median regression of income on borrower's occupation, age and gender; *Age* is the age of the borrower at the inception of the mortgage; *Foreign* is a dummy variable taking 1 if the borrower's nationality is other than Italian; *Graduate* is a dummy variable taking 1 if the borrower has graduated from a university; *North* (resp. *Center*) is a dummy variable taking 1 if the bank branch is located in a region of the North (resp. Center) of Italy; *Ln(Population)* is the natural logarithm of the resident population in the urban area where the bank is located (by zip code). Model 4 reports the results of a linear probability model (LPM) of *Active* on the chosen explanatory variables, including quadratic values of *Age* and *Ln(Income)*, where a two-stage Heckman (1979) selection model has been estimated. Standard errors, in parentheses, are bootstrapped based on 1,000 replications. ***, **, * indicate statistical significance at 1, 5 and 10 percent level, respectively.

Dependent variable: *Active*	(1)	(2)	(3)	(4)
Interest Differential	-0.0267***	-0.0274***	-0.0268***	-0.0264***
	(0.002)	(0.002)	(0.002)	(0.002)
Ln(Amount)	0.0142***	0.0141***	0.0133***	0.0133***
	(0.001)	(0.001)	(0.001)	(0.001)
Maturity	0.0008***	0.0007***	0.0005***	0.0005***
	(0.000)	(0.000)	(0.000)	(0.000)
Man		0.0064***	0.0055***	0.0071***
		(0.001)	(0.001)	(0.002)
Wealthy		-0.0052***	-0.0033**	-0.0031*
		(0.002)	(0.002)	(0.002)
Ln(Income)		-0.0058	-0.0044	-0.1873
		(0.005)	(0.005)	(0.129)
Age		-0.0003***	-0.0002***	-0.0015***
		(0.000)	(0.000)	(0.000)
Foreign		-0.0097***	-0.0121***	-0.0119***
		(0.002)	(0.002)	(0.002)
Graduate		0.0112***	0.0103***	0.0110***
		(0.002)	(0.002)	(0.002)
North			0.0190***	0.0190***
			(0.001)	(0.001)
Center			0.0126***	0.0129***
			(0.002)	(0.002)
Ln(Population)			0.0022***	0.0021***
			(0.000)	(0.000)
Ln(Income) Squared				0.0122
				(0.009)
Age Squared				0.0000***
				(0.000)
Constant				0.7436
				(0.487)
Origination Year FEs	Yes	Yes	Yes	Yes
Observations (first stage)	163,499	158,323	158,306	158,306
Observations (second stage)	146,975	141,799	141,782	141,782
Rho	-0.0141**	-0.0080	-0.0107*	-0.0129*

Table 11 – *Likelihood of active borrower with selection (refinanceable* v. *non-refinanceable).* The table reports the average marginal effects of a probit regression of *Active*, i.e. a dummy taking 1 if the mortgage has been refinanced by the borrower, on the chosen explanatory variables, where a two-stage probit selection model has been estimated for the probability of the mortgage to be eligible for refinancing. Selection equation includes *LTV*, i.e. the ratio between the principal of the loan and the estimated value of the real property, and *No. Guarantors*, i.e. the number of guarantors of the loan, as instruments (along with the same explanatory variables included into the second-stage equation). Rho is the correlation coefficient between errors of the two regression models and its significance relates to a Wald/LR-Chi^2 statistic of independent equations. *Interest Differential* is the difference between the (average monthly) IRS rate at the end of June 2009 and the IRS rate at the inception of the mortgage; *Man* is a dummy variable taking 1 if the borrower's gender is male; *Wealthy* is a dummy variable taking 1 if the borrower does not belong to the segment "mass market" (< €100,000 financial assets); *Ln(Income)* is the natural logarithm of the predicted monthly income of the borrower from median regression of income on borrower's occupation, age and gender; *Age* is the age of the borrower at the inception of the mortgage; *Foreign* is a dummy variable taking 1 if the borrower's nationality is other than Italian; *Graduate* is a dummy variable taking 1 if the borrower has graduated from a university; *North* (resp. *Center*) is a dummy variable taking 1 if the bank branch is located in a region of the North (resp. Center) of Italy; *Ln(Population)* is the natural logarithm of the resident population in the urban area where the bank is located (by zip code). Model 4 reports the results of a linear probability model (LPM) of *Active* on the chosen explanatory variables, including quadratic values of *Age* and *Ln(Income)*, where a two-stage Heckman (1979) selection model has been estimated. Robust standard errors (clustered by origination year) are given in parentheses. ***, **, * indicate statistical significance at 1, 5 and 10 percent level, respectively.

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	*Refinanceable*	*Active*	*Active*	*Active*	*Active*
Non-Refinanceable at:		*1%*	*5%*	*10%*	*25%*
Interest Differential	-0.0298***	-0.0293***	-0.0305***	-0.0312***	-0.0327***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Ln(Amount)	0.0111***	0.0211***	0.0211***	0.0207***	0.0194***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Maturity	-0.0035***	0.0003***	0.0004***	0.0005***	0.0006***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LTV	-0.1081***	-0.0208***	-0.0199***	-0.0173***	-0.0106***
	(0.005)	(0.003)	(0.003)	(0.004)	(0.004)
No. Guarantors	-0.0018**	-0.0193***	-0.0196***	-0.0200***	-0.0193***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Man	-0.0027	0.0061***	0.0062***	0.0061***	0.0049***
	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Ln(Income)	-0.0072	-0.0097**	-0.0100**	-0.0092**	-0.0098**
	(0.007)	(0.004)	(0.004)	(0.005)	(0.005)
Age	-0.0011***	-0.0003***	-0.0003***	-0.0003***	-0.0002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Foreign	-0.0620***	-0.0110***	-0.0004	0.0119***	0.0173***
	(0.002)	(0.002)	(0.003)	(0.004)	(0.004)
Graduate	0.0616***	0.0091***	0.0088***	0.0078***	0.0062***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Wealthy		-0.0056***	-0.0058***	-0.0060***	-0.0057***
		(0.002)	(0.002)	(0.002)	(0.002)
North		0.0178***	0.0184***	0.0190***	0.0183***
		(0.001)	(0.001)	(0.002)	(0.002)
Center		0.0147***	0.0146***	0.0149***	0.0123***
		(0.002)	(0.002)	(0.002)	(0.002)
Ln(Population)		0.0005*	0.0006**	0.0006**	0.0010***
		(0.000)	(0.000)	(0.000)	(0.000)
Origination Year FEs	No	Yes	Yes	Yes	Yes
Observations	159,823	140,397	134,932	127,912	106,785
McFadden Pseudo-R^2	0.0471	0.0868	0.0878	0.0868	0.0891

Table 12 – *Likelihood of active borrower excluding potentially non-refinanceable mortgages.* The table reports the average marginal effects of a probit regression of *Active*, i.e. a dummy taking 1 if the mortgage has been refinanced by the borrower, on the chosen explanatory variables. To account for the fact that mortgages may not be equally eligible for refinancing, regressions exclude mortgages whose refinancing probability (as modelled in model (1)) is less than a given percentile. For example, model (2) excludes 1% of the full sample of mortgages, based on the lowest 1% of refinancing probability (as predicted by model (1)). *Interest Differential* is the difference between the (average monthly) IRS rate at the end of June 2009 and the IRS rate at the inception of the mortgage; *Ln(Amount)* is the natural logarithm of the principal of the loan; *Maturity* is the length of the loan at inception; *LTV* is the loan-to-value, i.e. the ratio between the principal of the loan and the estimated value of the real property; *No. Guarantors* is the number of guarantors of the loan; *Man* is a dummy variable taking 1 if the borrower's gender is male; *Wealthy* is a dummy variable taking 1 if the borrower owns €100,000 financial assets or more; *Ln(Income)* is the natural logarithm of the predicted monthly income of the borrower from median regression of income on borrower's occupation, age and gender; *Age* is the age of the borrower at the inception of the mortgage; *Foreign* is a dummy variable taking 1 if the borrower's nationality is other than Italian; *Graduate* is a dummy variable taking 1 if the borrower has graduated from a university; *North* (resp. *Center*) is a dummy variable taking 1 if the bank branch is located in a region of the North (resp. Center) of Italy; *Ln(Population)* is the natural logarithm of the resident population in the urban area where the bank is located (by zip code. Standard errors, in parentheses, are bootstrapped based on 1,000 replications. ***, **, * indicate statistical significance at 1, 5 and 10 percent level, respectively.