Tax Deductibility, Market Frictions, and Price Discrimination: Evidence from the Mortgage Market.*

Maxence Valentin†

December 8, 2020

PRELIMINARY - DO NOT CIRCULATE

Abstract

Although not the primary intended beneficiary of the Mortgage Interest Deduction (MID) program, lenders capture a significant fraction of this government subsidy. This paper shows that borrowers who benefit from the MID pay on average 14.9 basis points higher interest rate than similar borrowers who do not benefit from it. This interest-gap represents a present value of $6,600 (or 2.35% of the original loan amount) for the median borrower. Consistent with a model of first-degree price discrimination, the interest-gap increases (1) with borrowers' marginal tax rate, and (2) with the degree of market frictions such as lenders' concentration, search cost, and leverage in bargaining.

JEL classification: G51, H24, H31, R21, R31.

Keywords: Mortgage interest deduction, Predatory lending, Distortionary tax, Market power.

*I am grateful to my dissertation committee members Brent Ambrose, Jiro Yoshida, Paul Grieco, and Liang Peng for their constant advice and help throughout this project. I am also thankful to Andrea Heuson, Ruchi Singh, Daniel Feenberg, Terry O'Brien and the participants of the 2020 FMA Doctoral Student Consortium for helpful suggestions at different stage of this project. All errors are mine.

†Smeal College of Business at the Pennsylvania State University. Email: maxence@psu.edu

"What beats borrowing at 5.5% APR?- Deducting it" - Citibank.

The Wall Street Journal - January 10, 1994

Interest payments on primary home mortgages have been tax deductible in the United States (US) since the 1913 introduction of the individual income tax. Designed to benefit borrowers, it is unclear whether lenders see any benefit or disadvantage from this tax subsidy. They have however aggressively communicated on the Mortgage Interest Deduction (MID) benefit to their borrowers. Are lenders hence getting higher profit from dealing with itemizers?[1] In this paper, I show that mortgage prices for borrowers benefiting from the MID is systematically higher than for otherwise identical borrowers. I show that this interest-gap is independent of mortgage ex-ante and ex-post risk factors, market concentration, borrowers' characteristics, and search behavior. The empirical results of this paper are consistent with a price discrimination theoretical model within which lenders' extract additional markups from borrowers benefiting from the MID.

The effects of MID on households' behavior and welfare have drawn scrutiny of both scholars and policy makers (Gruber et al., 2017; Glaeser and Shapiro, 2003; Hanson, 2020) with a constant debate about its repeal.[2] Many papers showed the positive association between MID benefits and house prices, indicating that a share of this subsidy is capitalized in the housing market (Sommer and Sullivan, 2018; Rappoport, 2016). However, little is known about the behavior of lenders and whether the MID is capitalized in mortgage prices too. I propose to add to this debate by showing mortgage lenders, economic agents not intended to benefit from the MID, capture a significant fraction of the borrowers' surplus of the government-subsidized MID program.

Why are lenders able to capitalize the MID into prices? I develop a theoretical model of demand for deductible goods that emphasizes the importance of two economic features. On the one hand, tax deductibility reduces the effective mortgage interest rate and hence, increases demand for mortgage debt. On the other hand, the market frictions that characterized the credit markets such as search costs, lack of competition, and leverage in bargaining (e.g. Argyle et al. (2020); Bachas and Liu (2019); Degryse and Ongena (2005)), create room for lenders' markups. Thus, the combination of higher demand elasticity introduced by the MID, with some degree of non-competitiveness allow lenders to capitalize the MID into mortgage prices. The model generates two main predictions: mortgage prices increase (1) with borrowers' marginal tax rate, and (2) with the degree of market frictions. The mortgage market offers an ideal setting for investigating the effects of deductibility on sellers' surplus because of the homogeneous nature of the good,

[1] Throughout this paper, I refer to itemizers as households who deduct their mortgage interest payments from their annual taxable income.

[2] See, for example, Bourassa and Grigsby (2000) which evaluates the different preferential treatments for homeownership arguing for the repeal of mortgage interest and property tax deductibility.

the sellers' access to buyers' information, the impossibility of resale, and the sharp variation in borrowers' deductibility benefits.

I obtain this sharp variation in borrowers' deductibility benefits by relying on a key feature of the US federal tax system: the standard deduction level. Deducting mortgage interest is conditional on having itemizable deductions larger than the standard deduction.[3] Itemizable deductions include, in order of importance, state and local taxes, MID, charitable givings and medical expenses (See Figure A2 of the appendix). For each borrower in my sample, I estimate the mortgage interest payments, the property tax and use the taxsim model of the National Bureau of Economic Research (NBER) to compute state income tax (Feenberg and Coutts, 1993). By comparing the sum of the deductibles expenses to the standard deduction, I create two classes of distinct borrowers with heterogeneous mortgage demand elasticities: the itemizers and the non-itemizers.

Based on this categorization, I test whether itemizers end up with mortgages with higher contract interest rate, after controlling for mortgage risk factors. With a sample of homogeneous mortgage products from Fannie Mae (FNMA), I show that itemizers have mortgage interest rates that are 14.9 basis points higher than similar borrowers who cannot benefit from the MID. The results hold with the inclusion of highly saturated risk factors, borrowers' income, local market concentration, and lenders' fixed effects. Discounting the differences in mortgage payments resulting from this interest-gap represents a present value of $6,600 for the median itemizers in the sample (or 2.35% of the median loan amount). A back-of-the-envelope counterfactual indicates that this mechanism reduces federal and state income tax collection by approximately 1%.

The theoretical model predicts that the interest-gap increases with borrowers' marginal tax rate (MTR) and the level of market frictions, predictions that I test empirically. By interacting the dummy variable for itemizers with the MTR, I show that the interest-gap indeed increases with borrowers' marginal tax rate (MTR). The result suggests that an increase of MTR from 15% to 30% for itemizers is associated with an interest rate increase of 12 basis points. The results also hold by using the interaction of itemizer and income. I further test whether lower market frictions mitigates the interest gap. The results show that lenders' concentration increases the interest-gap, as predicted by the theoretical model. But the data does not reveal that higher search cost (proxied by the number of bank branches available to a borrower), nor the refinancing market (characterized by higher rate dispersion) increases the interest-gap.

There are possible alternative explanations that could explain the interest-gap between itemizers and

[3] I am aware of the standard deduction unprecedented increase for fiscal year 2018. This policy would certainly impact both the real estate and the mortgage industries. The investigation of the impact of the recent fiscal change on market outcome is left for future research endeavor.

non-itemizers. First, it could be that mortgages sold to itemizers are yielding higher losses. Using the ex-post performance data as of June 2020, I show that the mortgages issued to itemizers and to non-itemizers do not differ in ex-post losses such as prepayment, modification, or default. Second, one might be concerned that itemizers take systematically more negative points on their mortgage as documented by Bhutta and Hizmo (2020) for minorities borrowers. Using closing costs estimates as a proxy for the incentive to take negative points, I show itemizers and non-itemizers respond similarly to such incentives. Lastly, itemizers might differ in their leverage in bargaining and/or their search intensity. To answer this question, I use the data from the National Survey of Mortgage Origination (NSMO) which contains information about borrowers' demographics, their search activity and the characteristics of the mortgage product bought. I show that the interest-gap is not explained by differences in borrowers' characteristics that have been shown to be associated with higher interest rate such as education (Agarwal et al., 2017), sophistication (Malliaris et al., 2019; Guiso et al., 2018), age (Agarwal et al., 2009) nor search activity (Allen et al., 2019; Malliaris et al., 2019; Bhutta et al., 2019). After controlling for mortgage characteristics, borrowers' demographics as well as borrowers' search intensity, I find an interest-gap of 18.6 basis points. This point estimate is remarkably close to the interest-gap of 14.9 basis points using the FNMA data; despite the differences in the nature of the loans included in both datasets.

The paper adds to the literature in two ways. First, it complements the literature about preferential tax treatment for housing and its impact on residential real estate outcomes. In the US, homeowners' imputed rent is exempt from income tax which encourages households to invest in the housing stocks as opposed to other industries (Gervais, 2002; Floetotto et al., 2016; Poterba and Sinai, 2008). Additionally, property tax and mortgage interest on primary residence are deductible from households income tax that further promotes homeownership. Although the MID increases housing and mortgage demand through lower effective mortgage rate, it also leads to house price appreciation (Sommer and Sullivan, 2018; Rappoport, 2016; Martin and Hanson, 2016; Hilber and Turner, 2014; Blouri et al., 2019).Given the direct increase in housing demand and the indirect increase in housing prices, the net effects of the MID on housing intensive and extensive margins are ex-ante ambiguous. Binner and Brett (2015) and Hanson (2012) find no relationship between MID and homeownership, whereas Sommer and Sullivan (2018) find that the revocation of the MID would actually increase homeownership. The results of this paper add to these empirical papers by providing evidence of MID capitalization into lenders' surplus; an economic agent that is not intended to benefit from this federal tax subsidy.

Second, the paper complements the literature on credit market structure and predatory lending. Numerous academic studies have documented market failure in the market for credits leading to price dispersion despite the homogeneous nature of the products (Bhutta et al., 2019; Bachas and Liu, 2019; Hortaçsu and Syverson, 2004). In the mortgage market, this observed price dispersion has been shown to be the result of borrowers' search activity (Allen et al., 2014; Ambokar and Samaee, 2019), borrowers' financial literacy (Agarwal et al., 2009, 2017), market power of the borrowers' preferred-bank (Allen et al., 2019; Guiso et al., 2018), borrowers' bargaining (or lack of) ability (Allen et al., 2014; Malliaris et al., 2019; Guiso et al., 2018), borrowers' demand for mortgage (Ma, 2020), mortgage brokers distortionary incentives (Robles-Garcia, 2018), and/or lenders' vertical differentiation (Benetton, 2018; Allen et al., 2019). In this paper, I suggest that, because of these market frictions, the MID is capitalized in mortgage prices and, therefore, lenders' profits.

The paper is structured as follows. In Section 1, I presents the theoretical framework as well as the empirical setting of the empirical study. In section 2, I provide the main results of the paper using the conforming loans dataset. In Section 3, I provide additional empirical results using the NSMO data to convince the readership that first-degree price discrimination is the most likely mechanism at play. In Section 4, I conclude.

1 Conceptual framework, and testing ground

In this section, I develop a stylized theoretical model with simplifying assumptions to provides the economic intuition behind the empirical strategy. It highlights the effects of introducing deductibility in a market characterized by market frictions on equilibrium prices. Then, I discuss the testing framework that relies on price discrimination in the mortgage market based on borrowers' itemizing status.

1.1 Theoretical framework

Demand for deductible goods. Following a two-period additive separable utility model (DeFusco and Paciorek, 2017; Brueckner, 1994), a borrower maximizes utility from housing service h and numeraire good c. In order to show the economics at play, I set exogenously the quantity of housing service to one.[4] As a consequence, we can simplify the consumer problem by:

$$
\begin{aligned}
\max_{c_1, c_2} \quad & u(c_1, c_2) = u(c_1) + \delta u(c_2) \\
\text{s.t.} \quad & c_1 + p = y + m \\
& c_2 = p - (1 + r_i(1 - \tau_i)m) \\
& 0 \le m \le p
\end{aligned}
\tag{1}
$$

where $\delta \in (0, 1)$ is the discount factor, y is borrowers income, m is quantity of mortgage, p is the price of the house, and $r_i(1 - \tau_i)$ is the after tax interest rate on mortgage available to borrower i. I further

[4] It is evident from the literature that the MID increases demand for mortgage for borrowers and hence demand for housing for those who benefit from it (Rappoport, 2016). However, in accordance with the empirical analysis, I aim at comparing individual borrowing similar amounts.

assume that household preferences follow a constant elasticity utility function $u(c) = \frac{c^{1-\xi}}{1-\xi}$. Under these assumptions and taking all parameters as exogenous from the borrowers' perspective, we can solve for the optimal mortgage demand[5]:

$$m^*(r_i, \tau_i) = \frac{p - (\delta(1 + r_i(1 - \tau_i)))^{1/\xi}(y - p)}{(\delta(1 + r_i(1 - \tau_i)))^{1/\xi} + (1 + r_i(1 - \tau_i))} \tag{2}$$

Introducing deductibility (i.e. increasing from $\tau_i = 0$ to $\tau_i > 0$) strictly increases the mortgage demand. As $\frac{\partial m^*(r_i, \tau_i > 0)}{\partial r_i} > \frac{\partial m^*(r_i, \tau_i = 0)}{\partial r_i}$, consumers with a positive MTR are less price sensitive with respect to mortgage prices. The after-tax own-price elasticity decreases with the level of MTR τ_i.

Lemma 1: Tax deductibility (1) increases demand for deductible goods, and (2) decreases the price elasticity of demand, with the magnitude of these effects increasing with marginal tax rate.

Supply of deductible goods. Lenders are profit maximizers providing unlimited mortgage at constant marginal cost at their level of outputs. When a borrower with income y and house price p meets with a monopolistic lender, the latter solves the following profit maximization problem:

$$\max_{r_i} \quad (r_i - c) \times \frac{p - (\delta(1 + r_i(1 - E[\tau_i])))^{1/\xi}(y - p)}{(\delta(1 + r_i(1 - E[\tau_i])))^{1/\xi} + (1 + r_i(1 - E[\tau_i]))} = \max_{r_i}(r_i - c)E[m^*(r_i, \tau_i)|Z_i] \tag{3}$$

where the expectation is taken over i's marginal income tax given information available about borrower i: Z_i. I follow the literature testing for market structure in a homogeneous good market for incorporating market frictions in the monopolistic version of the model (Graddy, 1995; Porter, 1983). Denoting $j \in \{1, \dots, J\}$ as lender, the equilibrium price must follow:

$$r_i^*(c, E[\tau_i], \theta_{i,j}) = c + \theta_{i,j} \frac{m^*(r_i^*, E[\tau_i])}{-m_r^*(r_i^*, E[\tau_i])} \tag{4}$$

where $0 \leq \theta_{i,j} \leq 1$ is an index that measures bargaining leverage of lender j over borrower i. Note that bargaining leverage is thought of as a composite of local market structure, individual bargaining power, and search cost. For instance, in a market with infinite search cost, $\theta_{i,j} = 1$, describing the inability of borrower to bargain for a lower rate. Or, assuming that there is no variation of $\theta_{i,j}$ across borrowers, equilibrium prices would be solely determined by the degree of competition. Hence, introducing this composite market friction parameter provides more flexibility in the interpretation of the empirical results.

It follows from Equation 4, that the introduction of deductibility (i.e. increasing from $\tau_i = 0$ to $\tau_i > 0$) increases equilibrium prices (conditional on $\theta \neq 0$). This increase in prices is strictly increasing in borrow-

[5] Note that throughout this section, I only assume interior solutions for both itemizers and non-itemizers. That is to say that lenders cannot distinguish from the two borrowers based on the decision to borrow or not.

ers' MTR τ_i, and the degree of market frictions $\theta_{i,j}$. This relates to the main propositions of this paper.

Proposition 1: Tax deductibility increases equilibrium prices for deductible goods.

Proposition 2: The increase in prices due to deductibility (a) increases with marginal tax rate, and (b) increases with market frictions.

1.2 Testing ground and hypothesis

I use the US mortgage market to establish the relationship between deductibility and equilibrium prices. I believe the mortgage market provides an ideal setting for multiple reasons. First, mortgage products in the US are seen as homogeneous products after controlling for borrowers risk. The predominance of conforming loans, guaranteed by the Government Sponsored Enterprises (GSEs) makes differentiation scarce in this market. Second, the mortgage market is characterized by various market frictions which does not allow borrowers, despite the product homogeneity, to borrow at lenders' marginal cost. This market friction are well documented in the literature and include for instance search cost (Allen et al., 2014; Ambokar and Samaee, 2019), borrowers' financial literacy (Agarwal et al., 2009, 2017), market power of the household preferred-bank (Allen et al., 2019; Guiso et al., 2018), borrowers' bargaining ability (Allen et al., 2014; Malliaris et al., 2019). In the empirical sections of this paper, I will mainly test three main types of frictions namely search cost, bargaining leverage, and lenders' concentration.

Lastly, the US mortgage is an ideal market as there is non-continuous variation in MTR resulting from the households choice between deducting itemizable deductions, or claiming the standard deduction. We can therefore think of $\tau_i = 0$ for non-itemizers and $\tau_i > 0$ for itemizers.[6] Figure 1 depicts the identification strategy. For two borrowers in the same market, borrowing the same amount, and being as costly to serve, we would observe r^I, the rate for itemizers, and r^{NI}, the counterpart for non-itemizers. The difference between the two interest rates is the interest-gap that is hypothesized to increase with both the MTR of the itemizer, and the degree of market frictions.

Hypothesis 1: Itemizers pay higher interest rate than non-itemizers on similar mortgages.

Hypothesis 2: The interest-gap increases with the itemizers marginal tax rate.

Hypothesis 3: The interest-gap increases with (a) search costs, (b) market concentration, and (c) lender's leverage in bargaining.

As a consequence, this paper uses a price discrimination framework to test the propositions of the theoretical model. Two additional conditions are necessary in light with the literature on price discrimination

[6] The choice between claiming the standard deduction and itemizable deductions is not specifically modeled in this context and is taken as exogenous in the lender/borrower pair bargaining game. For example, households cannot change their state of residency nor their marital status to benefit from higher deductions.

(Stole, 2007; Graddy, 1995; Leslie, 2004). First, resale cannot occur. This condition is met in the mortgage market as a borrower cannot sell her lower-priced mortgage to another borrower. Second, it must be the case that sellers have information about the itemizing status of the borrowers. I posit that mortgage originators have a substantial level of information to be able to approximate the itemizing status of borrowers. Specifically, lenders have often access to the borrowers' income, their state of residency, their marital status, the value and location of the property. These pieces of information are critical and sufficient for a good approximation of borrowers MID benefits.[7]

2 Itemizing deductions, mortgage risks, and mortgage rates

2.1 Data, and summary statistics

To test the main hypotheses of the paper, I use individual conforming loans originated from December 2015 to March 2016 and securitized by Fannie Mae (FNMA).[8] The sample consists of about 350,000 individual mortgages covering 50 states and the District of Columbia. The mortgages are all conforming loans with similar features (loan amount, loan term, loan type etc.). To create a pool of homogeneous mortgages, I further restrict the sample to urban (within an MSA), fixed rate mortgages, collateralized by a single unit that are either use for refinancing or purchasing (construction loans are excluded). The sample used in this analysis consists of 240,815 individual mortgages. Despite not generalizable to non-conforming loans data, using homogeneous products would provide more convincing evidence about first-degree price discrimination. Hence, the effect depicted below would provide a lower bound of the effect on non-conforming loans.

The biggest challenge is that individual tax deduction status (itemizer or non-itemizer) is not observable. However, the higher the sum of the deductions, the higher the probability of itemizing. Given that state income tax, MID, and property taxes are the largest deductions claimed by American households (see Figure A2 of the Appendix), a good approximation of the three components is adequate for categorizing borrowers into one or the other category. Based on information in the dataset, I calculate the amount of mortgage interest paid during the first year of the mortgage, and approximate the property tax for each borrower based on the location of the property.[9] In order to derive the state income tax, and the itemizable status of the borrower, I rely on the NBER taxsim software (Feenberg and Coutts, 1993). The NBER taxsim calculates state and federal income taxes, marginal tax rates and itemizable status based on individual tax-

[7]Section A of the Appendix discusses in more details the effects of lenders information on equilibrium prices, markups and lender surplus. Overall, the higher the level of information, the stronger the relationship depicted in this theoretical section.

[8]The Fannie Mae loan performance data publicly provides individual loan information at origination as well as performance throughout the length of the mortgage.

[9]The median property tax rate is calculated by divided the median property value by the median property tax paid from the 2016 ACS 5-years provided at the zipcode level. Note that the tax rate is measured as percentage of home value, and not assessed value; which allows me to ignore the variation in assessment ratio across space.

payers characteristics. It classifies borrowers into itemizers and non-itemizers based on rational behavior (i.e. claiming the highest deduction). Note that several household characteristics that largely influence the itemizable status of borrowers are unobserved in the FNMA dataset. Of greater importance for this study is the absence of information about borrowers' marital status preventing the model from using different standard deduction levels. In the main analysis, I use the highest level of standard deduction to avoid miscategorizing non-itemizers as itemizers (therefore reducing the probability of a Type II error) and consequently set marital status for all borrowers as married. Therefore, in this setting, the identification relies more heavily on income and spatial variation than borrowers' characteristics' variation.

From this categorization, 53.7% of the borrowers in my sample are categorized as itemizers. In 2016, 30.4% of American households itemized their deductions. Three main reasons can rationalize this difference. First, I assume that borrowers are taking the itemizable deductions when it is more beneficial to them. It might be the case that households with deduction level closed to the standard deduction threshold decide to claim the standard deduction for simplicity. Second, my sample is conditioned on buying a house, which therefore mechanically increases the probability of itemizing. Lastly, I am restricting my sample to urban markets which again increases the probability of having high deductions. Therefore, 53.7% can be seem as a reasonable estimate of the number of itemizers within an urban homeowner sample. To confirm the appropriateness of my measure, I compare the share of itemizers from my categorization aggregated at the state and zip-code 3-digits level with data from the Statistics of Income dataset (SOI) of the IRS.[10] Figure 2 shows this relationship for all states (panel A) and all zipcode 3-digits (panel B). Though the level of itemizing is systematically higher with my measure, the positive correlation between the two series is evident.[11] Overall, we can be confident that the itemizing status computed at the borrower level aligns with the IRS data. This dummy variable is highly correlated with income (correlation of 0.50), which is consistent with statistics of the IRS (See Figure A3 in the Appendix).

Table 1 reports the summary statistics of the sample for both itemizers and non-itemizers (the correlation table is available in Table A6 of the appendix). We can observe that itemizers are contracting loan with longer term and larger LTV; consistent with the idea of maximizing mortgage interest deductions. On the other hand, itemizers are more credit worthy with a larger credit score on average. Interest rate averages 4.02% with a wide variation. Panel A of Figure 3 depicts the dispersion in raw interest rates. Despite the product homogeneity, the raw difference between the 95$^{\text{th}}$ and the 5$^{\text{th}}$ is 1.625 percentage points. After controlling for spatiotemporal risk factors (Panel B of Figure 3), the difference remain economically significant at 1.53 percentage points. After controlling for risk characteristics of the mortgage with highly saturated

[10]Figure A1 of the Appendix shows the geographical distribution of itemizers by plotting the proportion of itemizers in each zipcode-3 digits. Though the North-East and West Cost appear to have higher proportions of itemizers, there are within Zip-code variations in all regions.

[11]Controlling for homeownership rate increases the fit of the regression (R^2 increase to circa 0.75) which confirms the selection bias. Regression results are available upon request.

fixed effects interactions (see next section for details), the difference is still 0.90 percentage points (Panel C). This result is consistent with previous academic works depicting a high price dispersion in credit markets (Bhutta et al., 2019; Argyle et al., 2020).

The theoretical model in the previous section predicts a positive association between MTR and mortgage prices. Figure 4 plots the density distribution of interest rates for individual having low federal MTR (lower than 15%), and for individuals with high MTR (higher or equal to 15%). Although the distribution of rates for itemizers and non-itemizers are similar for low MTR, it becomes evident that for individuals facing high MTR, the itemizers' distribution is to the right of the non-itemizers' one. Lastly, given that MTR increases with income, the benefits of the MID should also increase with income conditional on itemizing (Sommer and Sullivan, 2018). Figure 5 plots the distribution of mortgage interest rates within each income decile for borrowers classified as itemizers or non-itemizers. Although the difference in distribution is not noticeable for lower income individuals, the rate distribution between itemizers and non-itemizers widens as income rises.

2.2 Methodology

In order to test whether itemizers pay higher mortgage interest rates, I rely on a semi-parametric methodological approach similar to Argyle et al. (2020) and Bhutta et al. (2019). Specifically, I run the following regression where coefficient of interest is δ which measures the additional interest rate paid by a borrower classified as itemizer. I expect this coefficient to be positive corroborating the first proposition of the theoretical framework.

$$rate_{i,m,t} = MSA_m \times Month_t + f(X_i) + \delta Itm_i + \epsilon_{i,m,t} \tag{5}$$

where $rate_{i,m,t}$ is the mortgage of interest paid by borrower i, in MSA m, at month t, X_i are mortgage level characteristics, Itm_i equals one for borrowers classified as itemizers, and $\epsilon_{i,m,t}$ is the error term. I control for mortgage level characteristics with a discrete grid in the mortgage characteristics space. A grid cell (i.e. a market) comprises of all mortgages with the same purpose (refinance or purchase), within a credit score bin (10 deciles), a DTI bin (5 categories), a loan amount bin (5 categories), a LTV bin (5 categories), and a given term (10, 15, 20, 25, 30 year). A market is therefore defined as the interaction of fixed effects $Purpose^{FE} \times Score^{FE} \times DTI^{FE} \times Amount^{FE} \times LTV^{FE} \times Term^{FE}$ resulting in potentially 12,500 markets. As a lot of borrowers end up in singleton cell after the highly saturated grid procedure, I only keep borrowers in market where at least one itemizer and one non-itemizers are represented (n=145,513 with 40.5% of itemizers). One can think of this procedure as a discrete matching procedure over the risk dimension of a mortgage. The identifying assumption of this procedure is that the macro-economic related risk are additive to the mortgage level risk factors. I include X_i, the vector of mortgage continuous characteristics

($Score$, DTI, $Amount$, LTV), to control for linear association not controlled in the discrete grid procedure (Bhutta et al., 2019).

2.3 Main results

The main results are presented in Table 2. Column (1) does not control for any spatio-temporal nor mortgage level characteristics. The coefficient δ is negative indicating that itemizers pay on average a lower interest rate. Considering that the dummy variable is highly correlated with income, the negative sign is consistent with the idea that higher income individual are less risky and hence are offered a lower interest rate. Column (2) adds the MSA times month fixed effects in the regression. The δ coefficient increases to -0.014 though still negative and significant.

From Column (3), I control for both spatial-temporal variation and mortgage risk factors. Coefficient δ on Column (3) is positive and significant at the 1% significance level. On average, an itemizer pays 0.163 percentage points more on a similar mortgage. Adding the market control significantly increases the R^2 to 0.69 indicating that mortgage risk characteristics highly influence mortgage rates. Given that itemizers have higher income than non-itemizers, this result could be driven by a lower price sensitivity of higher-income individuals. In column (4), I control for income by including income deciles fixed effects.[12] These fixed effects are significant and monotonically increasing with income. It indicates that higher income individuals do in fact pay higher mortgage interest after controlling for risk consistent with diminishing price sensitivity. After controlling for income, the coefficient δ decreases in magnitude to 0.147 further documenting price discrimination. As the identification of δ relies heavily on the state variation (through state income taxes), I add 51 state fixed effects in the baseline regression to mitigate endogeneity concern. The identification here relies on MSA that span multiple states. The coefficient δ increases to 0.149 after such inclusion (Column [5]). These results confirm that deductibility increase equilibrium prices, which supports Hypothesis 1.[13]

2.3.1 Counterfactual

Based on an interest-gap of 14.7 basis points between itemizers and non-itemizers, I compute the change in mortgage payments for each itemizers if they were charged the lowest rate. The results of this counterfactual exercise are shown in Table 3.[14] For the median itemizer in the sample, a 14.7 basis points reduction of interest rate represents an annual payment reduction of $304. Given that this cash flow would be free of

[12] Decile 1: [2.3 - 18.0K]; Decile 2: (18.0K - 23.8K]; Decile 3: (23.8 - 29.1K]; Decile 4: (29.1 - 34.6K]; Decile 5: (34.6 - 40.3K]; Decile 6: (40.3 - 46.8K]; Decile 7: (40.3 - 54.3K]; Decile 8: (54.3 - 64.3K]; Decile 9: (64.3 - 82.0K]; Decile 10: (82 - 2,430K].

[13] One might be worried that my results are driven by the sample period chosen. I replicate the analysis using data from 2011 to 2018 that is shown Table A4. We can observe that the interest-gap is persistent through the last decade.

[14] Given the endogeneity between MID and mortgage demand; this counterfactual analysis solely aims at providing an approximation of the magnitude of such effects for borrowers, lenders, and the collection of income taxes. A companion paper constructs a partial equilibrium model for properly investigating the welfare implications of this price discriminating behavior.

risk, I assume a discount rate of 1.92% (the average 10-years treasury bill over the sample period); which represents \$6,633 of present value for the median borrower (2.35% of her loan amount) and \$8,487 for the itemizer at the 75th percentile of that present value distribution. Even discounting this cash flow at 8% represents \$3,546 for the median itemizers; an amount economically significant for a median US household.[15]

Using the NBER taxsim, I compute the tax burden for all itemizers assuming a reduction of mortgage interest payment resulting from a lower interest rate of 14.7 basis points. The amount of itemizable deductions for all itemizers is therefore reduced leading to an increase in federal and state taxes.[16] Through this change, the median itemizers would pay an additional \$64.58 of federal taxes and \$9.42 of state taxes in 2016. Despite the increase in taxes, the decrease in mortgage payment for each borrower strictly exceeds the additional tax burdens. In sample, prohibiting price discrimination leads to an increase in taxes amounts \$13.95 millions; or an increase of 1.00% of taxes collected. Given that mortgages considered in this analysis are all conventional, one might expect a more pronounced effects if all loans were to be considered. This tax loss due to price discrimination represents 1.24% of the MID tax benefits in my sample. This share can be seen as the capitalized portion of the MID into lenders' surplus.

2.3.2 Robustness checks

Interest rate and points trade-off. One shortcoming of the current dataset is the absence of information about points. Mortgage borrowers can pay points at origination to lower their contract mortgage interest rate; where points can be positive (reducing interest rate) or negative (increasing interest rate). There exists therefore a menu of points/interest rates from which the borrower can choose from. If itemizers are more inclined to get negative points at origination, it would mechanically inflates the contract interest rate paid and therefore invalidate the first-degree price discrimination hypothesis. Focusing solely on contract interest rates can therefore generate misleading conclusions about price discrimination (Bhutta and Hizmo, 2020).

To mitigate this concern, I collected mortgage closing costs in each State and computed the closing cost for each borrower as the percentage of her mortgage loan amount.[17] If itemizers systematically offset their closing costs with negative points more than non-itemizers, we should expect a positive relationship between closing costs and the δ coefficient. Column (1) of Table 4 shows that the higher the closing costs a borrower face, the higher the contract interest rate; consistent with the idea of offsetting high closing

[15]Report on the Economic Well-Being of U.S. Households in 2017

[16]Note that, only 3.9% of the itemizers became non-itemizers after the change of interest rate. The choice between itemizing and non-itemizing seems secondary in this setting and can be thought as exogenous within the lender-borrower bargaining game for the magnitude of interest rate considered in this study.

[17]Closings costs are collected from a survey of closing costs per US state performed by ClosingCorp. I assume that some of these costs are fixed (title insurance, appraisal fees) whereas others are a function of the property value (transfer tax). Though facing obvious data limitation, this section aims at providing suggesting evidence that the effects found in the previous section is not solely driven by point behavior.

costs with negative points. In column (2), I add the interaction between closing cost and the itemizer dummy to verify whether the effect of closing cost on interest rates is heterogeneous. The coefficient on the interaction is negative and significant suggesting that, if anything, itemizers would pay lower interest rate when facing higher closing cost. More importantly, the δ coefficient magnitude and significance is unchanged; confirming that itemizers pay higher interest rate. These results show that there is a trade-off between the possibility of taking negative points (i.e. higher closing costs) and contract interest rate. Though the results are not more pronounced for itemizers which would have justified observing higher contract interest rate. In addition, as noted by Agarwal et al. (2017), small lenders usually do not offer the interest-point trade off. I use in Column (3) a sample of mortgages for which no lenders name is provided. The coefficient δ is slightly higher though of similar magnitude in this sub-sample adding evidence that the points-interest rate trade-off cannot explain the interest-gap.

Lender fixed effects. The interest-gap could result from itemizers going to higher quality lenders which would be reflected into mortgage prices. Similarly, we could think of mortgage lenders only operating in areas with higher proportion of itemizers to capture the higher willingness to pay. I include a lender fixed effects in the baseline regression with results shown in Column (4). As I only observe 20 lenders names, the sample is reduced to 60,945 individual mortgages. Controlling for mortgage lenders reduces the coefficient δ to 0.129; though still significant at the 1% level. This coefficient indicates that the interest gap between an itemizer and a non-itemizer (same DTI, LTV, Credit score, Income, Zip code, loan amount) going to the same lender is 13.5 basis points. I show in Figure 6, by interacting lenders fixed effect with the Itm_i dummy, that this discriminating behavior is prevalent for 19 out of 20 lenders in the sample. Despite some dispersion in these estimates, it suggests that most lenders are getting a higher markups from itemizers by price discriminating.

Standard deduction levels. As mentioned above, I estimate the itemizer dummy assuming that all borrowers file their income tax as couple. I relax this assumption here. In Column (5) of Table 4, I recompute taxes and itemizing status assuming that all borrowers are single households. The standard deduction for single household is about half the amount for married filers, which increases the number of itemizers in the sample tremendously (89% are considered itemizers). Given that I only keep market with at least one itemizer and one non-itemizer, it reduces the the number of observation to 53,170. The coefficient δ is positive (0.184) and significant at the 1% level. In Column (6) of Table 4, I reproduce the same analysis but removes from the analysis all borrowers that would become itemizers if I were to consider them as single. Hence, the identification compares individual that are itemizers regardless of marital status against non-itemizers regardless of their marital status. Note that the number of observations is reduced importantly as the matching algorithm requires to observe one of each in each market. The coefficient on Itm_i increases to 0.340 consistent with the expected measurement error bias.

Performance of loans and itemizers. If the loans originated to itemizers are more risky ex-post, the observed interest-gap could be justified by the accurate expectation of lenders' future loss. To rule out this possibility, I follow the literature testing ex-ante characteristics of the loans to ex-post performance (Keys et al., 2010; Ma, 2020), and run the following regression:

$$ExPostRisk_{i,m,t} = MSA_m \times Month_t + f(X_i) + \delta Itm_i + rate_{i,m,t} + lengthi, t + \epsilon_{i,m,t}$$

where $ExPostRisk_{i,m,t}$ are different measures of loan performance such as default, prepayment, modification, and foreclosure. I further include the contract interest rate as well as a month since origination variables as both as been showed to influence prepayment and delinquency. I base my definitions of default, prepayment and modification as of the second quarter of 2020 and present the results in Table 5. Column (1) to (4) present the estimates of linear probability models where the outcome variable are indicators for prepayment (column [1]), delinquency (Column [2]), modification (Column [3]) or liquidation (Column [4]). Column (5) and (6) verify that, conditional on modification (Column [5]) and foreclosure (Column [6]), itemizers are not leading to higher losses. Regardless of the dependent variable, I do not find evidence that mortgage originated to itemizers are performing differently than mortgages originated to non-itemizers. In all the regressions, the coefficient on Itm_i is not distinguishable from zero.

2.4 The effects of higher deductibility benefits and market frictions

The second Proposition of the theoretical framework indicates that the effects of deductibility on prices is enhanced (1) by higher marginal tax rate (MTR), and (2) by higher degree of market frictions. In order to do so, I interact the Itemizer dummy with proxy for MTR (τ_i in Equation 6) or with proxy for market frictions ($\theta_{i,j}$ in Equation 7).

$$rate_{i,m,t} = MSA_m \times Month_t + f(X_i) + \delta\ Itm_i + \gamma\ (Itm_i \times \tau_i) + \epsilon_{i,m,t} \tag{6}$$

$$rate_{i,m,t} = MSA_m \times Month_t + f(X_i) + \delta\ Itm_i + \gamma\ (Itm_i \times \theta_{i,j}) + \epsilon_{i,m,t} \tag{7}$$

Table 6 and 7 investigate in more details these effects by testing Hypotheses 2 and 3. All regressions in these table include spatio-temporal controls, mortgage risk factors controls, and borrowers' income fixed effects as Column (4) of Table 2.

Is the interest-gap increasing with borrowers' deductibility benefits? We saw in the baseline regression, that income is positively associated with contract interest rate (Column (1) of Table 6 replicates this result without presenting the income decile fixed effects). Given that the MID disproportionally benefits higher income individuals conditional on itemizing (Sommer and Sullivan, 2018), I test whether higher income individuals pay larger interest rate by interacting the Itm_i dummy with the borrowers' income. The second coefficient of Column (2) supports this hypothesis by showing that higher income individuals who itemize their deductions pay higher interest rates (after controlling for income). Column (3) adds the borrower's federal MTR in the baseline regression. On average, individual with higher MTR pay higher mortgage rates. Column (4) shows that the effects is however heterogeneous with the interaction between MTR and Itm being positive. It suggests that itemizers' interest rates increases with MTR even more than for non-itemizers. This results supports hypothesis 2 of the theoretical section: the interest-gap increases with the itemizers marginal tax rate. The result suggests that an increase of MTR from 15% to 30% for itemizers is associated with an interest rate increase of 12.1 basis points.

To verify the non-linearity of these effects, I also interact the dummy Itm_i with each income decile, and four MTR categories. The coefficients of these interactions are presented in Figure 7. Although the effect is not different from 0 for low income borrowers or for borrower with low MTR, the difference in coefficients monotonically increase as borrower's income or MTR increases. For the highest income individuals (more than $82,200 of annual income), the results suggest that they pay 27.5 basis points more than borrowers with identical income that do not itemize, or 133.8 basis points more than a non-itemizer in the lowest decile.

Is the interest-gap increasing with lenders' market power? In this section, I test Hypothesis 3 by using three types of market frictions that prevents borrowers from borrowing at marginal cost. First, I consider the effect of access to lenders as a measure of search cost for borrowers. Consistent with the literature on search in lending markets (Degryse and Ongena, 2005; Allen et al., 2019), I compute the number of bank branches within each zicode 3-digits (the smallest geography included in the dataset) normalized by the population.[18]. Hence, a higher number indicates a lower search cost. Adding this variable in the baseline model (Column [1] of Table 7) shows that the higher the number of bank branches available, the lower the interest rates; though non-significant. It however becomes negative and significant once the interaction with Itm_i is included (Column [2]); consistent with a model of search. Interestingly, the interaction with the Itm_i dummy variable is positive and significant at the 10% level. It indicates that itemizers seems to benefit less from a lower search cost, which is inconsistent with Hypothesis 3 that higher market friction lead to higher prices for deductibles items. The coefficient δ is however unchanged in response to this inclusion.

Second, I consider market concentration as a measure of higher market frictions. I compute the level of market concentration by computing the Share of the top 4 lenders in each state using all the mortgage originated from The Home Mortgage Disclosure Act (HMDA) in 2016.[19] Adding the Top4Share variable in the baseline regression shows a non-significant association between concentration and price. By adding

[18] I use all the bank branches listed on the FIDC SOD database out of June 2016.

[19] I also use a Herfindahl-Hirschman Index which do not change the magnitude nor significance of the results.

the interaction of Top4Share with Itm_i, it appears that the effects of concentration is asymmetric. The coefficient on the interaction between Top4Share and Itm_i is positive and significant at the 1% level. A 10% increase of market share of the top 4 lenders is associated with an increase of contract interest rate of 0.023 and 0.046 percentage points for non-itemizers and itemizers respectively. It indicates that itemizers in State with a more concentrated mortgage market pay higher interest rate. This result supports Hypothesis 3 of the theoretical model.

Lastly, Fuster et al. (2013) noted that there is a larger price dispersion in the refinancing market than in the purchase market. They argue that this is consistent with the idea that the purchase market is more competitive; potentially due to borrowers' attachment to their current lender for instance (Allen et al., 2019; Guiso et al., 2018) or high switching cost. In my setting, I therefore associate the refinancing market with a higher level of lenders' market power (i.e. higher $\theta_{i,j}$). To support Hypothesis 3, we should expect the interest-gap to be larger in the refinancing market. The coefficient on the interaction between refinance and itemizers, which is positive and non-significant, does not support this hypothesis. There is no statistical evidence that the higher market power of lenders in the refinancing market allows them to extract additional surplus from the itemizers. Only one out of three tests of this section suggest that the higher the market frictions, the higher the interest gap. The lack of significance could be due to an overall low level of $\theta_{i,j}$ in the mortgage market which leave little room for lenders to price discriminate. In section **??** of the Appendix, I structurally estimate the level of market friction using the FNMA dataset. It shows that $\theta_{i,j} \approx 0.05$, which indeed depicts a low level of market power. Despite this low level on average, this section showed that it is sufficient for lenders to extract the some MID surplus from borrowers.

3 Itemizing deductions, borrower's characteristics, and mortgage rates

In this section, I turn to another dataset to provide more insights on the mechanism that leads to the results shown previously. By using a survey of borrowers, this section primarily aims at showing that itemizers are not paying higher mortgage rates because of behavioral reasons such as shopping, knowledge, or education. It shows that $\theta_{i,j}$, the bargaining leverage of lender j over borrower i, is not systematically higher for itemizers. It does also adds results showing the impact of higher market frictions on the equilibrium interest-gap (Hypothesis 3).

3.1 Data, and summary statistics

I use the National Survey of Mortgage Origination (NSMO) provided publicly by the Federal Housing Finance Agency (FHFA) in collaboration with the Consumer Financial Protection Bureau (CFPB).[20] The NSMO is a quarterly survey of a representative pool of borrowers in the US. The survey includes a rich

[20] National survey of mortgage origination (NSMO) public use file.

set of borrowers characteristics and documents their experience in relation to the mortgage recently originated. In addition, each respondent's set of answers is matched with her mortgage characteristics (Loan amount, interest rates, LTV etc.). Hence, it allows to document how borrowers' characteristics or behavioral activity impacts the mortgage originated (Malliaris et al., 2019; Bhutta et al., 2019). I further restrict the sample to FRM, conventional loans, that are either for purchasing or refinancing a single-unit dwellings not for investment purpose with loan terms of 10, 15, 20, 25, or 30 years.

As per the previous analysis, I infer the itemizing status of each respondent in the survey using the taxsim of the NBER (Feenberg and Coutts, 1993). However, the NSMO data does not include borrowers' geographical information. As state income tax is the major component of the federal deduction, it creates some room for miscategorizing borrowers as itemizer or as non-itemizers. Moreover, the property tax varies with location too. To overcome this issue, I compute the state and property taxes as if each respondent were living in either 50 state or the District of Columbia. For each respondent-state pair, I obtained whether the respondent would have been an itemizer or not. I then weight the outcome by the state population to get the probability that the respondent is an itemizer or not in the absence of state residency information. Hence, in this study $0 \leq ITM_i \leq 1$ is a continuous variable. Figure 8 shows the distribution of this variable. Despite this approximation, 63% of the respondents are, regardless of their state of residency, classified 100% of the time in either the non-itemizer or itemizer category. As opposed to the previous section where identification was heavily relying on spatial variation, this current study relies more heavily on borrowers demographics (age, marital status etc.) to identify the effect of itemizing deductions. As per the analysis with the FNMA dataset, the itemizer distribution presents an upward selection bias as compared to the IRS data.

Table 8 shows the summary statistics of the main variables of the study for the respondents that are classified surely in one or the other category out of which 69% are classified as itemizers (Correlation table is available in Table A7 of the Appendix). In the whole sample, the interest rate spread over the US 10-years treasury averages at 1.83%. I use the interest spread as measure of prices in this study as opposed to contract rate to control for time variant marginal cost fluctuations.[21] Unconditionally, the mean and median interest rates is shown to be higher for itemizers. Consistent with the FNMA data, itemizers take longer term mortgages with higher LTV but they have on average a higher credit score and lower DTI.

3.2 Itemizing deductions and borrowers' bargaining power

The literature has shown that borrowers' characteristics such as age (Agarwal et al., 2009; Bucks and Pence, 2008; Andersen et al., 2020), education (Agarwal et al., 2017), sophistication (Malliaris et al., 2019;

[21] One limitation of the NSMO dataset is that it provides the spread over the PMMS censored at -1.5 and 1.5%. Hence the interest rate spread used in this study also suffers from this data censoring.

Guiso et al., 2018) or minorities (Ambrose et al., 2020; Bucks and Pence, 2008; Bhutta and Hizmo, 2020) are non-risk based borrowers' characteristics that are related to mortgage rates. I see these characteristics as elements of that affect $\theta_{i,j}$, the bargaining leverage of lender j when meeting borrower i. If itemizers' characteristics were highly related to these, it would undermine the price discrimination (with respect to deductibility benefits) argument of the theoretical model.

Age. Agarwal et al. (2009) showed that financial mistakes follow a U-shape function with respect to age with the cost-minimizing age being 53. If the distribution of age for itemizers is thinner around mid-ages, we could attribute the price differential to financial mistakes made by the young and elderly borrowers. Plot A of Figure 9 shows the distributions of age for both itemizers and non-itemizers in the sample. The distribution suggests that itemizers are not over-represented by younger and/or older households as the density in the mid-age category is larger for the itemizers. Rationalizing itemizers discrimination with age-based discrimination seem unrealistic.

Income. Higher income individuals have lower price sensitivity though higher income individuals make less mistakes when choosing a mortgage (Bucks and Pence, 2008). Hence, the marginal effect of income on mortgage prices is ex-ante ambiguous. I however show in the previous section that after controlling for mortgage risk factors, higher income individuals pay more on average supporting the former statement. We can see in Panel B of Figure 9 that income is highly correlated with itemizing deductions. This figure highlights that thoroughly controlling for income is essential which the previous empirical tests do.

Education. Multiple studies have shown that individuals with higher education make less financial mistakes and face lower mortgages rates (Agarwal et al., 2017). If itemizers are on average less educated, the interest-gap could be explained by educational mistake or education-based discrimination. Panel C of Figure 9 shows the distribution of itemizers and non-itemizers across different education level (the lowest bin being the lowest educational attainment). It suggests a positive association between education and itemizers. This positive association goes against an education-based discrimination highlighted in the literature.

Sophistication. Not only formal education matters when it comes to lowering mortgage interest rates. As pointed by a contemporaneous paper that uses the NSMO dataset, confusion about the mortgage market process leads to higher interest rates on average (Malliaris et al., 2019). I use question 5 of the survey to compute the variable *sophistication*. The question asks for the level of familiarity (not at all, somewhat, and very) with elements of the mortgage origination process.[22] I summed the responses to create a *sophistication* variable that ranges from 0 (always "not at all" familiar) to 14 (always "very" familiar). I plot the distribution of this variable for both itemizers and non-itemizers in Panel D of Figure 9. If anything, the

[22] *When you began the process of getting this mortgage, how familiar were you with each of the following: (1) The mortgage interest rates available at times, (2) The different types of mortgage available, (3) The mortgage process, (4) The downpayment needed to qualify for a mortgage, (5) The income needed to qualify for a mortgage, (6) Your credit history or credit score, and (7) The money needed at closing.*

itemizers' distribution for sophistication is to the right of the non-itemizers' indicating that itemizers are more knowledgeable about the mortgage origination on average.

At first glance, it seems hard to believe that the additional interest rates paid by itemizers are due to differences in borrowers' characteristics that are systematically related to a higher $\theta_{i,j}$.[23] Furthermore, the price discrimination identification relies on the assumption that lenders are able to distinguish borrowers that itemize versus the ones that do not. This section provides evidence that itemizers and non-itemizers do differ on multiple dimensions, though not the ones impacting their bargaining leverage downward.[24]

3.3 Itemizing deductions and search intensity

The identification of the model relies on the fact that $\theta_{i,j}$, the bargaining leverage of lender j over borrower i is orthogonal to itemizing status. This bargaining leverage is thought of as a composite of local market structure, individual bargaining power, and/or search cost. In this section, I test whether itemizers differ in the search activity. If itemizers search less for their mortgages products, it would result in higher mortgage rates which, under traditional search paradigm, are not due to price discrimination but through lower search activity (Stigler, 1961).

I primarily use question 11 of the NSMO to investigate the relationship between itemizing and search intensity. The question asks *How many different lenders/mortgage brokers did you seriously consider before choosing where to apply for this mortgage?* with answers going from *One*, to *Five or more*.[25] Unconditionally, itemizers search more than non-itemizer with an average number of search of 1.71 (versus 1.61), and a median of 2 (versus 1). I also use question 12 for robustness checks that ask for the number of applications on the same scale.[26] The two variables are highly correlated (0.43). I also remove all respondents that declared to search more because of a first rejection to compare individuals who searched for a more suitable mortgage product.

To formally test whether itemizers search less conditional on borrowers and mortgage characteristics, I regress the number of search on the itemizers variable and additional controls.

$$Search_{i,t} = \alpha_t + \delta ITM_i + X_i\beta + Z_i\gamma + \epsilon_{i,t,m}$$

[23] The ethnicity is not observed in the NSMO data which prevents me from formally testing whether itemizers are more likely to be part of a minority.

[24] Table A5 of the Appendix tests more formally the hypothesis positing that higher lenders' market power induces a higher interest-gap using the borrowers' characteristics as variation in borrowers ability to bargain. The results confirm this hypothesis but the coefficients lack significance due to the low variation in the independent variables. The interest-gap is, regardless of the model specification, always positive and significant at the 1% level.

[25] I consider this latter category as 5 searches.

[26] *How many different lenders/mortgage brokers did you end up applying to?*

where $Search_{i,t,m}$ is the number of search made by borrower i for her mortgage originated in month t, ITM_i is the computed probability that borrower i itemizes her deductions, X_i are mortgage characteristics bin fixed effects (LTV, Loan amount, DTI, Credit score, loan term, purpose), and Z_i are borrowers characteristics. I control for month fixed effect α_t as the search activity varies depending on the phase of the economic cycle (Bhutta et al., 2019).

Results. The results are shown in Table 9. Column (1) shows the positive relationship between itemizing deduction and search unconditionally. By controlling for mortgage characteristics and time variation, the coefficient δ becomes negative though non-significant. It suggests that holding mortgage characteristics constant (searching for a given mortgage), itemizers do not search less than non-itemizers. When adding borrowers' characteristics, the coefficients δ remains non-significant. We cannot reject the hypotheses that itemizers have similar searching activity than non-itemizers. Consistent with prior studies relating borrowers' behavior and mortgage rates, Column (3) shows that sophisticated, educated, and young borrowers are positively related with search activity (Bhutta et al., 2019; Malliaris et al., 2019). The table also shows that refinance mortgage are associated with lower level of search consistent which confirms the higher level of lenders' market power in this submarket (Fuster et al., 2013).

Column (4) uses the number of applications as dependent variable and shows that when it comes to applying to multiple mortgages, itemizers and non-itemizers do not differ either. Column (5) and (6) use binary definitions of search (search is greater than 1) and application (application is greater than 1) as dependent variables. Estimating the binary response models with probit regressions shows that itemizers do search less (Column [5]) but do not differ in their number of applications. The coefficients δ of Column (5), significant at the 10% level, suggests that being classified as an itemizer reduces the probability of searching more than 1 lender by 3.4%.[27]

The different results of Table 9 do not suggest that itemizers have a lower search intensity nor that they do apply to less mortgages than non-itemizers. One might still be concerned about the effects of a lower search on interest rates and on the interest gap. Therefore, the following section controls for search behavior to further document the price discrimination mechanism.

3.4 Does borrowers' search activity mitigate the interest-gap?

To verify whether the price discrimination is mitigated by borrowers' characteristics or search intensity, I rely on the same methodology shown in Section 2.

$$spread_{i,t} = Month_t + f(X_i) + \delta Itm_i + \mu Search_i + Z_i\gamma + \epsilon_{i,t} \tag{8}$$

[27]Using only the observations for which we are certain about the itemizing status, increases the magnitude of the δ coefficients to 0.14, significant at the 10% level.

where $spread_{i,t}$ is the contract interest rate spread over the 10-years treasury bill, X_i are mortgage level characteristics, Itm_i is the probability that borrower i is an itemizer, $Search_i$ is the number of search, Z_i are borrowers characteristics, and $\epsilon_{i,t}$ is the error term.[28] I control for mortgage level characteristics with a discrete grid in the mortgage characteristics space. A grid cell (i.e. a market) is defined as the interaction of fixed effects $Purpose^{FE} \times Score^{FE} \times DTI^{FE} \times Amount^{FE} \times LTV^{FE} \times Term^{FE}$ resulting in 3,476 individual markets. I only keep observations that fall in a market with at least one borrower having a non-zero and one borrower having a non-one ITM_i.

Results. Results are shown in Table 10. Column (1) shows, consistent with the results of the previous section, that itemizers pay on average more for their mortgages after controlling for mortgage temporal and risk factors. Column (2) adds the borrowers' characteristics that could influence the mortgage rates. The coefficient δ remains positive and significant at the 1% level. Despite the evidence that itemizers and non-itemizers are different individuals on several aspects (as shown in Figure 9), the inclusion of borrowers characteristics influence only marginally the level δ. The interest-gap is robust to the inclusion of the borrowers' characteristics. We can observe that less educated, less sophisticated, and more than 60 years old borrowers pay higher interest rates ceteris paribus. The coefficient δ of Column (2) suggests that itemizers pay 21.6 basis points higher interest rates on similar mortgage. Despite the differences in sample periods covered, and the nature of the Itm_i calculation across the two studies, the magnitude of the coefficient δ is surprisingly consistent.

Column (3) adds the number of mortgage search in the baseline model. The coefficient on search is, as expected, negative and significant indicating that searching more reduces the average mortgage rates confirming results of previous studies (Bhutta et al., 2019; Malliaris et al., 2019; Allen et al., 2019). Yet, the coefficient δ does not change after the inclusion of search. Column (4) adds the interaction between the number of search with the itemizer variable. Hypothesis 3 suggests that in the presence of lower market frictions (i.e. more search), the interest-gap should be lower. Thus, we should expect a negative coefficient on the interaction term. The coefficient on the interaction is however non-significant implying that search intensity affects interest rates homogeneously across borrowers who itemize and those who do not. Column (5) and Column (6) repeats the analysis by using the number of filled applications. The results indicate that applying to more mortgage does not have any significant effect on interest rates. Hence, this section shows that search activity impacts borrowers similarly regardless of their itemizing status providing evidence that the interest-gap is not a result of borrowers' characteristics nor heterogeneous search behavior.

[28] Note that I do not observe any locational information in the NSMO data which prevents from using more granular spatio-temporal grid as in the above analysis.

4 Concluding remarks

Using two different sources of data, this paper shows that borrowers who benefit from the MID pay a higher mortgage rate on average. The interest-gap increases with marginal tax rates and lenders' market power. I further show that the point estimate is consistent across datasets, and robust to the inclusion of highly saturated risk factors, borrowers' income, closing costs, local market structure, market concentration, lenders' fixed effects, borrowers' characteristics, and borrowers' search intensity. All results point toward a price discriminating mechanism within which lenders extract surplus from borrowers benefiting from this federal subsidy. This paper is, to my knowledge, the first to show how differences in households' itemizing status can lead to capitalization of federal subsidy by market participants not intended to benefit from this fiscal policy. Therefore, if the market for deductible goods is not-competitive, in the sense that there are substantial markups, the transmission of fiscal policy is not fully capitalized into consumer surplus.

References

Agarwal, S., Ben-David, I., and Yao, V. (2017). Systematic mistakes in the mortgage market and lack of financial sophistication. *Journal of Financial Economics*, 123(1):42–58.

Agarwal, S., Driscoll, J. C., Gabaix, X., and Laibson, D. (2009). The age of reason: Financial decisions over the life cycle and implications for regulation. *Brookings Papers on Economic Activity*, 2009(2):51–117.

Allen, J., Clark, R., and Houde, J.-F. (2014). The effect of mergers in search markets: Evidence from the canadian mortgage industry. *American Economic Review*, 104(10):3365–96.

Allen, J., Clark, R., and Houde, J.-F. (2019). Search frictions and market power in negotiated-price markets. *Journal of Political Economy*, 127(4):1550–1598.

Ambokar, S. and Samaee, K. (2019). Inaction, search costs, and market power in the us mortgage market.

Ambrose, B. W., Conklin, J., and Lopez, L. A. (2020). Does borrower and broker race affect the cost of mortgage credit? *Review of Financial Studies (Forthcoming)*.

Andersen, S., Campbell, J. Y., Nielsen, K. M., and Ramadorai, T. (2020). Sources of inaction in household finance: Evidence from the danish mortgage market. *American Economic Review*, 110(10):3184–3230.

Argyle, B., Nadauld, T. D., and Palmer, C. (2020). Real effects of search frictions in consumer credit markets. Technical report, National Bureau of Economic Research.

Bachas, N. and Liu, E. (2019). Market power in small business lending: A two dimensional bunching approach.

Benetton, M. (2018). Leverage regulation and market structure: An empirical model of the uk mortgage market. *Available at SSRN 3247956*.

Bhutta, N., Fuster, A., and Hizmo, A. (2019). Paying too much? price dispersion in the us mortgage market. *Price Dispersion in the US Mortgage Market (July 19, 2019)*.

Bhutta, N. and Hizmo, A. (2020). Do Minorities Pay More for Mortgages? *The Review of Financial Studies*.

Binner, A. and Brett, D. (2015). Exploring mortgage interest deduction reforms: An equilibrium sorting model with endogenous tenure choice. *Journal of Public Economics*, 122:40–54.

Blouri, Y., Büchler, S. C., and Schöni, O. (2019). The geography of housing subsidies.

Bourassa, S. C. and Grigsby, W. G. (2000). Income tax concessions for owner-occupied housing.

Brueckner, J. K. (1994). The demand for mortgage debt: some basic results. *Journal of housing Economics*, 3(4):251–262.

Bucks, B. and Pence, K. (2008). Do borrowers know their mortgage terms? *Journal of urban Economics*, 64(2):218–233.

Chetty, R., Looney, A., and Kroft, K. (2009). Salience and taxation: Theory and evidence. *American economic review*, 99(4):1145–77.

DeFusco, A. A. and Paciorek, A. (2017). The interest rate elasticity of mortgage demand: Evidence from bunching at the conforming loan limit. *American Economic Journal: Economic Policy*, 9(1):210–40.

Degryse, H. and Ongena, S. (2005). Distance, lending relationships, and competition. *The Journal of Finance*, 60(1):231–266.

Feenberg, D. and Coutts, E. (1993). An introduction to the taxsim model. *Journal of Policy Analysis and management*, 12(1):189–194.

Floetotto, M., Kirker, M., and Stroebel, J. (2016). Government intervention in the housing market: Who wins, who loses? *Journal of Monetary Economics*, 80:106–123.

Fuster, A., Goodman, L. S., Lucca, D. O., Madar, L., Molloy, L., and Willen, P. (2013). The rising gap between primary and secondary mortgage rates. *Economic Policy Review*, 19(2).

Gervais, M. (2002). Housing taxation and capital accumulation. *Journal of Monetary Economics*, 49(7):1461–1489.

Glaeser, E. L. and Shapiro, J. M. (2003). The benefits of the home mortgage interest deduction. *Tax policy and the economy*, 17:37–82.

Graddy, K. (1995). Testing for imperfect competition at the fulton fish market. *The RAND Journal of Economics*, pages 75–92.

Gruber, J., Jensen, A., and Kleven, H. (2017). Do people respond to the mortage interest deduction? quasi-experimental evidence from denmark. Technical report, National Bureau of Economic Research.

Guiso, L., Pozzi, A., Tsoy, A., Gambacorta, L., and Mistrulli, P. E. (2018). The cost of steering in financial markets: Evidence from the mortgage market. *Available at SSRN 2951042*.

Hanson, A. (2012). Size of home, homeownership, and the mortgage interest deduction. *Journal of Housing Economics*, 21(3):195–210.

Hanson, A. (2020). Taxes and borrower behavior: Evidence from the mortgage interest deductibility limit. *Journal of Urban Economics*, page 103256.

Hilber, C. A. and Turner, T. M. (2014). The mortgage interest deduction and its impact on homeownership decisions. *Review of Economics and Statistics*, 96(4):618–637.

Hlavac, M. (2018). *stargazer: Well-Formatted Regression and Summary Statistics Tables*. Central European Labour Studies Institute (CELSI), Bratislava, Slovakia. R package version 5.2.2.

Hortaçsu, A. and Syverson, C. (2004). Product differentiation, search costs, and competition in the mutual fund industry: A case study of s&p 500 index funds. *The Quarterly Journal of Economics*, 119(2):403–456.

Keys, B. J., Mukherjee, T., Seru, A., and Vig, V. (2010). Did securitization lead to lax screening? evidence from subprime loans. *The Quarterly journal of economics*, 125(1):307–362.

Leslie, P. (2004). Price discrimination in broadway theater. *RAND Journal of Economics*, pages 520–541.

Ma, C. (2020). Per-customer quantity limit and price discrimination: Evidence from the us residential mortgage market. *International Journal of Industrial Organization*, page 102588.

Malliaris, S. G., Rettl, D. A., and Singh, R. (2019). Is competition a cure for confusion? evidence from the residential mortgage market. *Evidence from the Residential Mortgage Market (July 30, 2019)*.

Martin, H. and Hanson, A. (2016). Metropolitan area home prices and the mortgage interest deduction: Estimates and simulations from policy change. *Regional Science and Urban Economics*, 59:12–23.

Porter, R. H. (1983). A study of cartel stability: the joint executive committee, 1880-1886. *The Bell Journal of Economics*, pages 301–314.

Poterba, J. and Sinai, T. (2008). Tax expenditures for owner-occupied housing: Deductions for property taxes and mortgage interest and the exclusion of imputed rental income. *American Economic Review*, 98(2):84–89.

Rappoport, D. E. (2016). Do mortgage subsidies help or hurt borrowers?

Robles-Garcia, C. (2018). Competition and incentives in mortgage markets: The role of brokers.

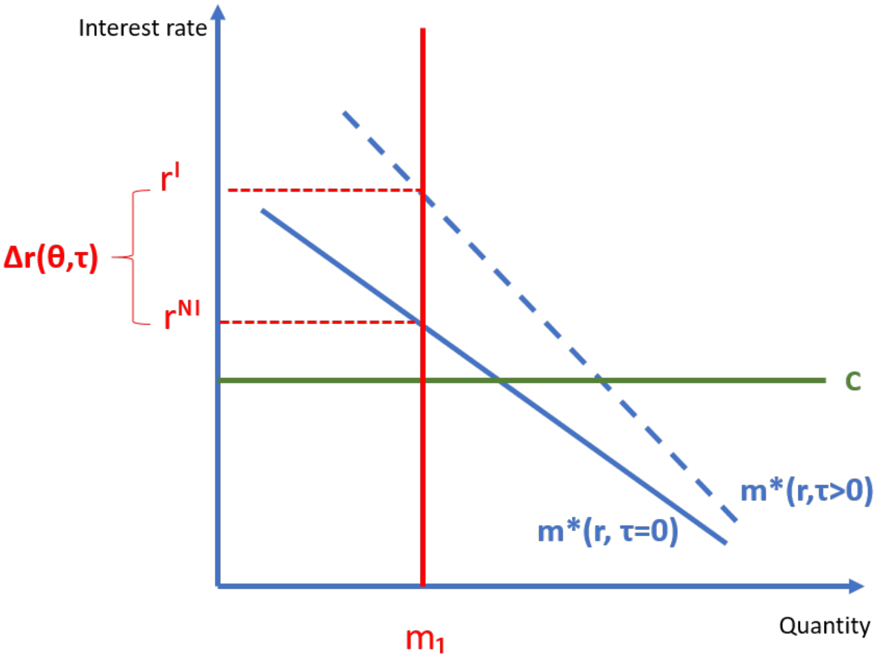
Saiz, A. (2010). The geographic determinants of housing supply. *The Quarterly Journal of Economics*, 125(3):1253–1296.

Sommer, K. and Sullivan, P. (2018). Implications of us tax policy for house prices, rents, and homeownership. *American Economic Review*, 108(2):241–74.

Stigler, G. J. (1961). The economics of information. *Journal of political economy*, 69(3):213–225.

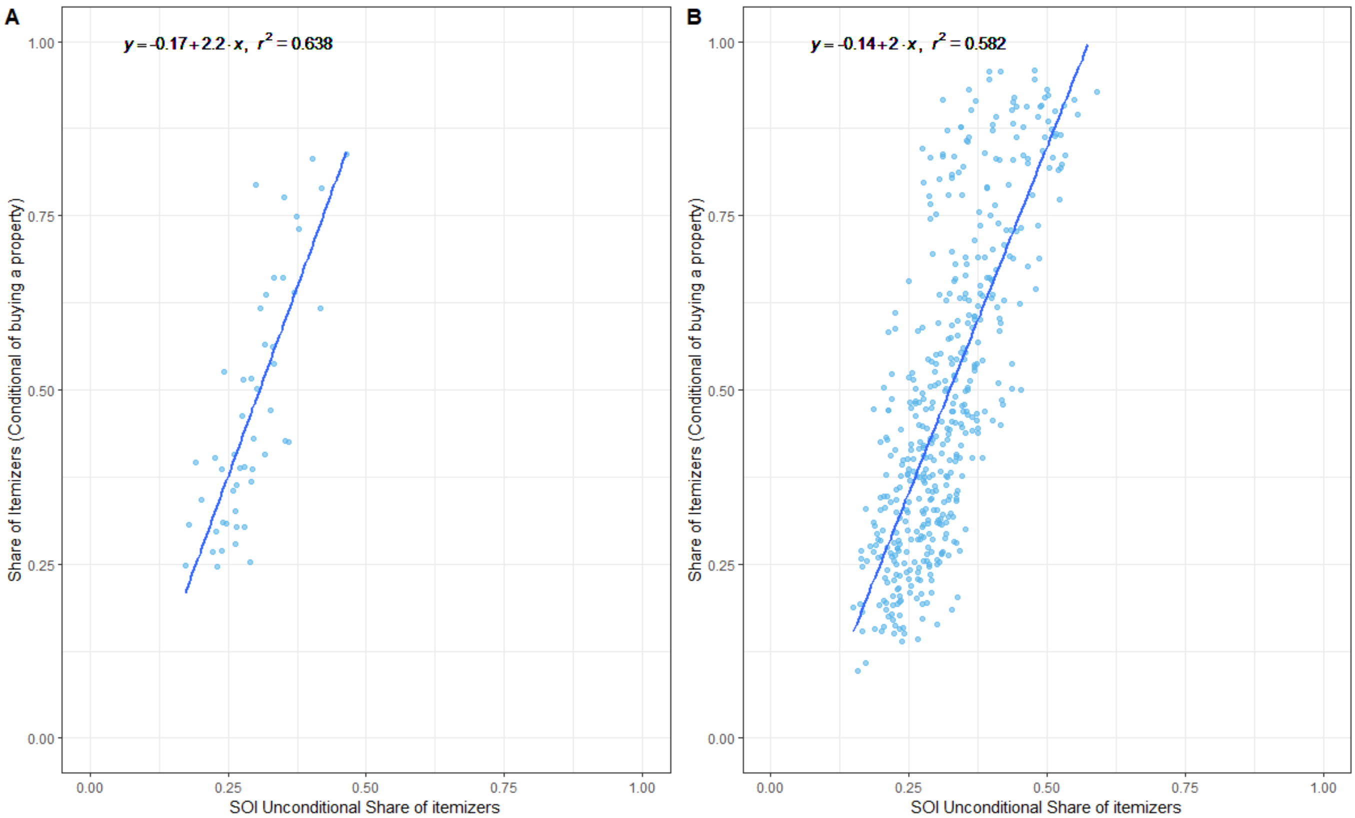
Stole, L. A. (2007). Price discrimination and competition. *Handbook of industrial organization*, 3:2221–2299.

Figure 1: Theoretical framework and empirical strategy



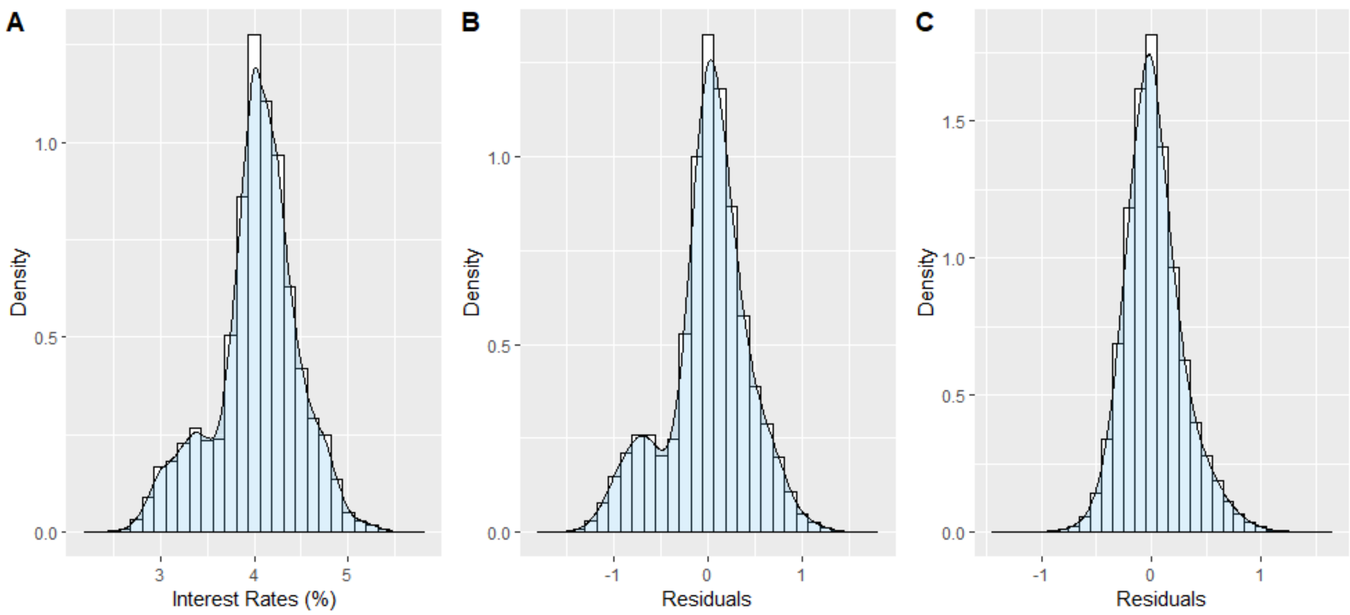
Note: This Figure graphically depicts the identification strategy employs to test the main proposition of the paper. The blue curves represent heterogeneous mortgage demand functions for itemizers (dotted line), and non-itemizers (solid line). The green line represents the marginal cost to serve this equally costly borrowers. For each mortgage demand level (vertical red line at m_1, rate charge to itemizers r^I and to non-itemizers it observed r^{NI}. The difference between the two is the interest-gap ($\Delta r(\theta, \tau)$) hypothesized to be positive for non-zero θ.

Figure 2: Itemizers share by State and Zipcode 3-digits relative to SOI data



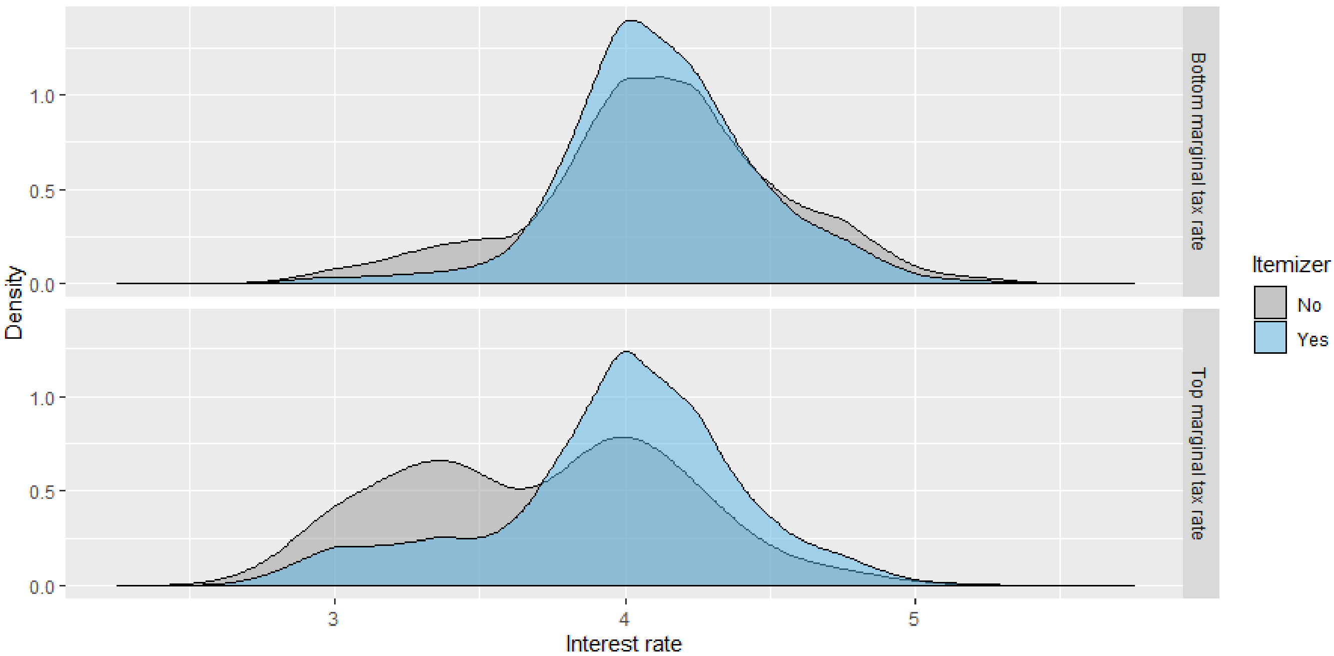
Note: This scatter plots compares the share of itemizers between the dataset of the Statistics of Income (SOI) of the IRS and the dataset used in the main analysis. In Panel A and B, each dot represents a state and a zip code 3-digits respectively.

Figure 3: Mortgage interest rates distributions



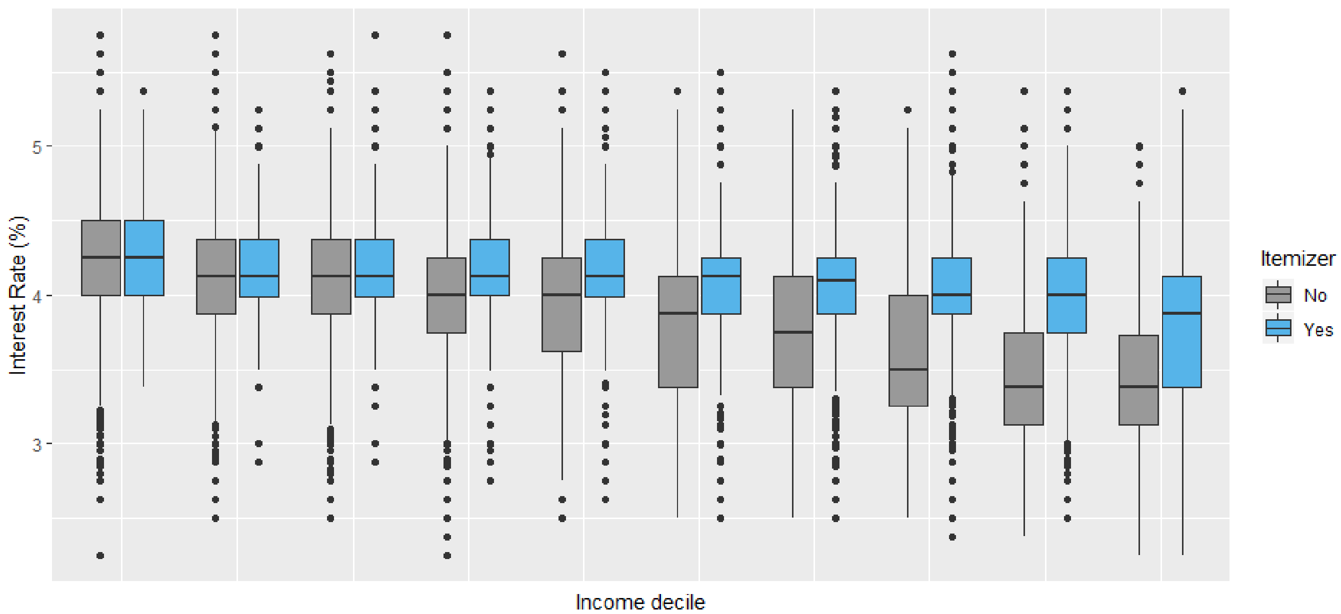
Note: these histograms show the dispersion of interest rates in the FNMA Q1-2016 originated mortgages. Panel A shows the raw distribution of interest rates for the 240,831 loans in the dataset. Panel B shows the distributions of the interest rates distribution after controlling for spatio-temporal risk factors. The sample is reduced to 240,143 loans by keeping market (MSA x months) with at least one itemizer and one non-itemizer. Panel C shows the distribution after controlling for mortgage characteristics by interacting 10 credit score deciles, 5 LTV bins, 5 DTI bins, 5 loan amount bins and two purpose categories. The residuals shown are the difference between the residual of Panel B and the median residuals of the interacted bin (6,925 individual markets). Loans in market with less than one itemizer or less than one non-itemizers are removed (149,335 observations).

Figure 4: Interest rate paid as a function of borrower's federal marginal tax rate, and itemizing status



Note: the two panels show the distribution of mortgage interest rates across heterogeneous households. The top panel shows the distribution for individuals with low federal marginal tax rates (less than 15%) while the bottom panel shows the distribution for individuals facing high marginal tax rates (equal of more than 15%). In both panels, the blue shaded distribution represents the distributions for individuals classified as itemizers.

Figure 5: Rate distributions per income group, and itemizers status

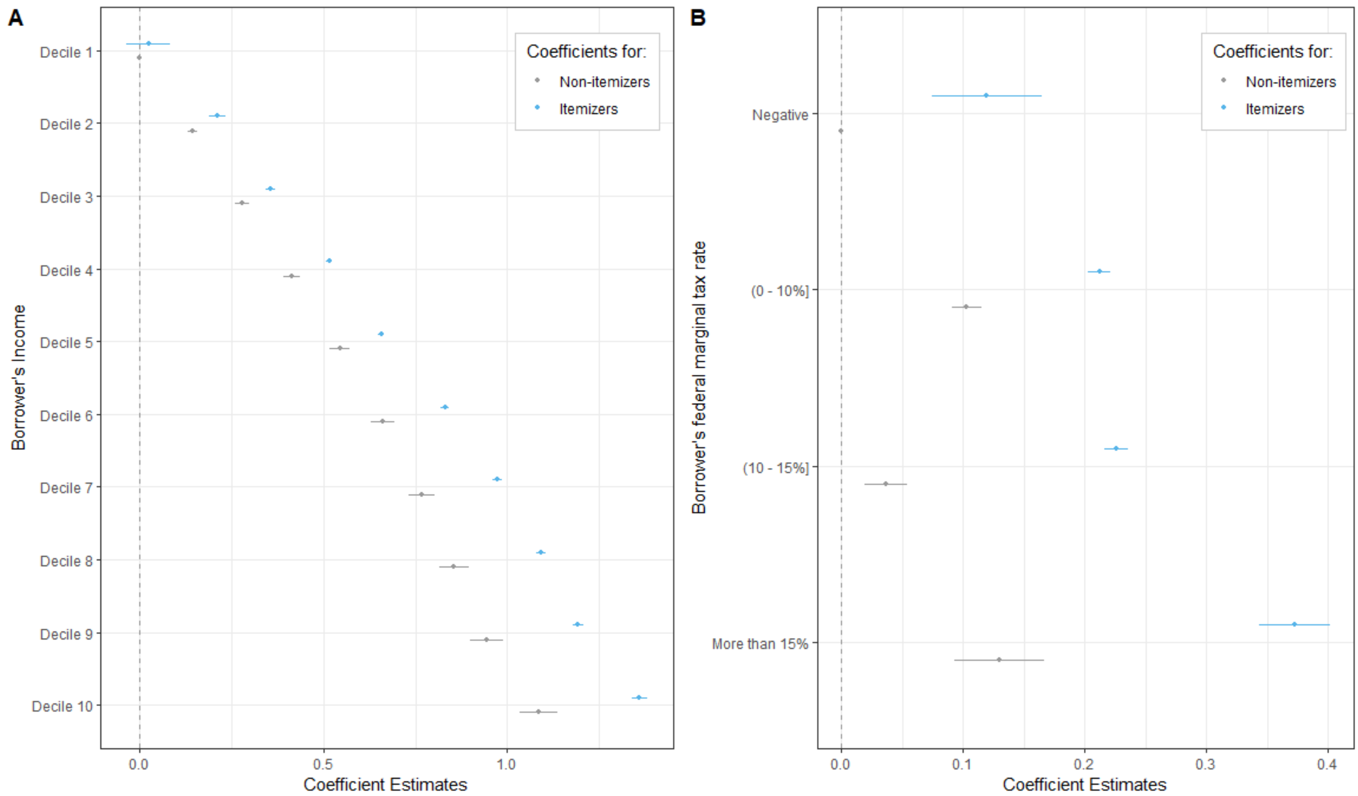


Note: this figure uses box-plots to show the distribution of mortgage interest rates across households with different income and itemizing status. Each box-plot pair shows the distribution of interest rates (vertical axis) within an income decile with the lowest income being on the left-hand side of the horizontal axis. Within all income decile, the blue shaded box-plot represents the distributions for individuals classified as itemizers.

Figure 6: Coefficients on the interaction terms between Itemizer dummy and lenders' fixed effect

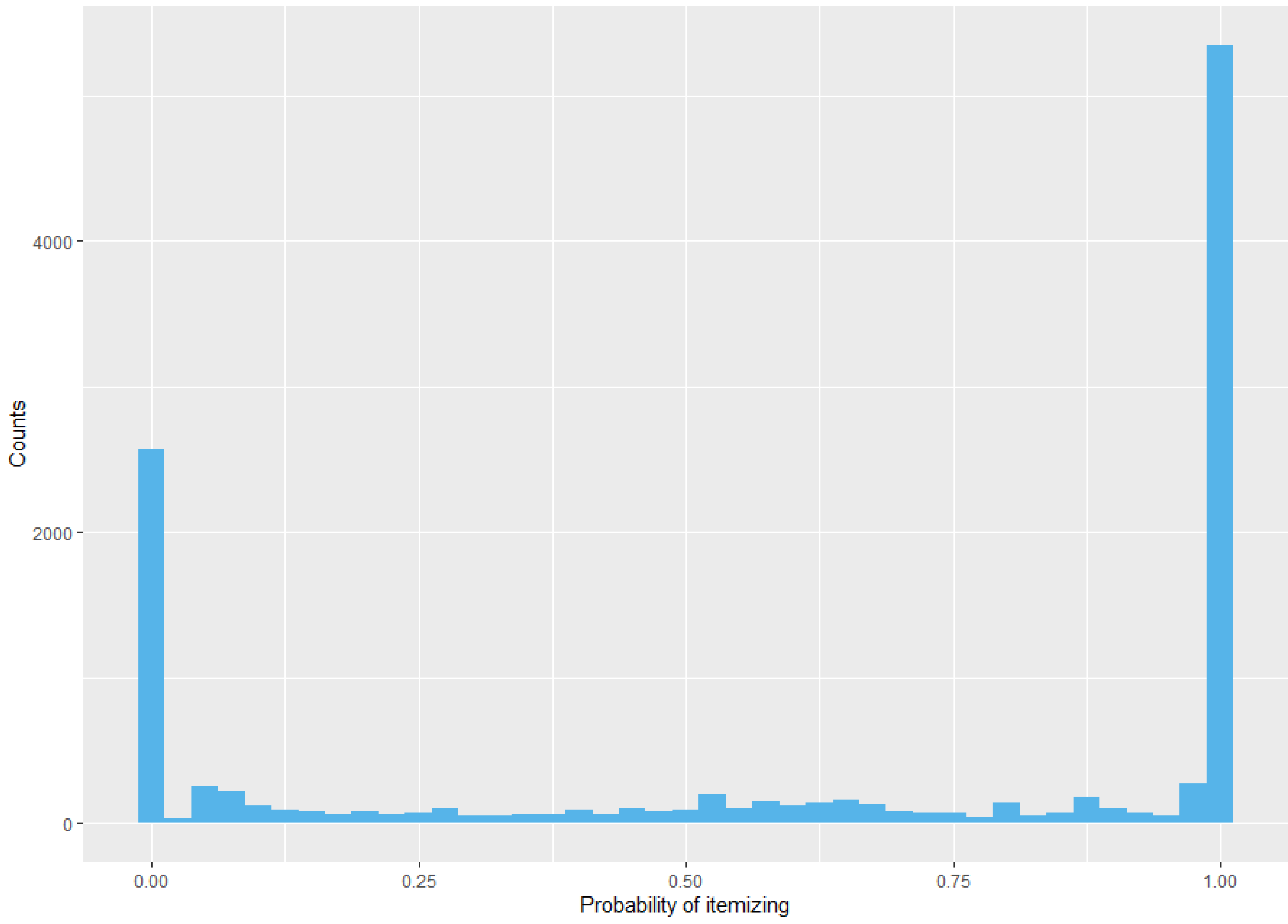
Note: this figure shows the regression coefficients on the interaction terms between the itemizer dummy and lenders' fixed effects. Only the mortgage for which a lender name is available in the FNMA is used reducing the sample to 63,098 loans. The results follow from the methodology outline in Equation 5 which controls for spatio-temporal risk factors, mortgage product risk factors, state fixed effects, income deciles and lenders' fixed effects.

Figure 7: Coefficients on the interaction terms of income and MTR with the Itemizer dummy



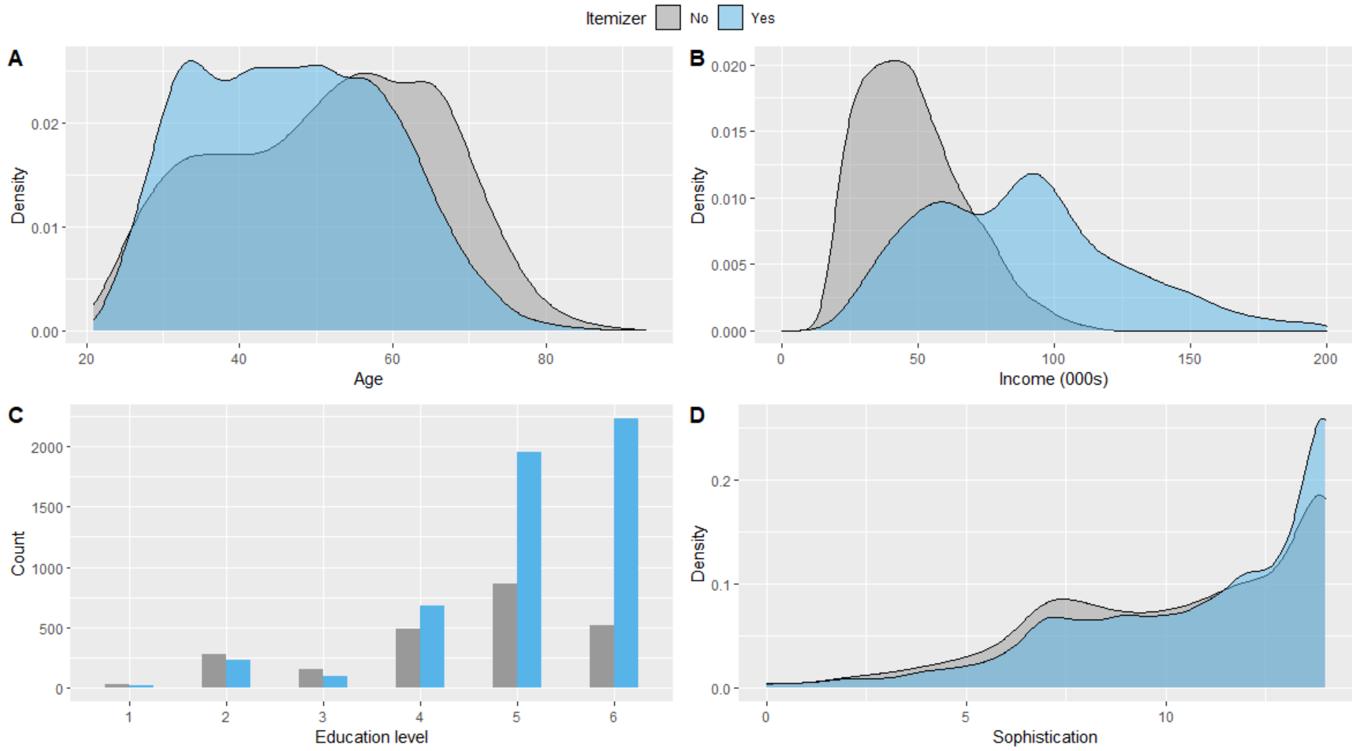
Note: these two panels show in blue the regression coefficients on the interaction terms between the itemizer dummy and income decile (Panel A) and federal marginal tax rate categories (Panel B). All coefficients are compared with the lowest non-itemizers income decile (Panel A), and lowest non-itemizer MTR. The results follow from the methodology outline in Equation 6 which controls for spatio-temporal risk factors, mortgage product risk factors, state fixed effects and income deciles. On both panel, the vertical axis represents income (or MTR) categories with the lowest being on the top.

Figure 8: NSMO respondents - probability of itemizing



Note: this histogram shows the distribution of the the variables "probability of itemizing" for all respondents in the NSMO sample. The probability of itemizing is computed with the taxsim by providing borrowers characteristics (age, number of dependents, income, marital status etc.) for each possible state of residency and by weighting the results by the state population. A value of one (zero) indicates that regardless of the state of residency, the respondent would always (never) be classified as itemizer.

Figure 9: Itemizing status and borrowers' characteristics distribution



Note: these panels show the distribution of key variables in the NSMO data for all respondents for households classified surely as itemizers (blue shaded) or non-itemizers (grey). Panel A shows the distribution of respondents' age, Panel B shows the income distribution, Panel C shows the level of educational attainment with the lowest level on the left hand-side, and Panel D shows the distribution of the constructed variable "sophistication" capturing the level of familiarity of each respondent with the mortgage process (the lowest number indicate low familiarity).

Table 1: Summary statistics of the originated loans

This table reports the summary statistics of the loan level origination data. All loans securitized by Fannie Mae from December 2015 to March 2016 are included. The itemizer dummy is determined by computing the main deductions faced by a borrower; namely the state taxes, the mortgage interest and the property taxes; and comparing it to the standard deduction level for married couple in tax year 2016. The marginal tax rate is computed with respect to borrower's income by finite difference by the NBER taxsim. Bank branches index is the number of FIDC insured bank branches divided by the number of loan originated at the Zipcode 3-digits level. Top 4 share is the market share (as the number of loan originated) of the biggest four mortgage originator in each US State computing with the HMDA origination dataset. Closing cost (%) is the estimated mortgage closing cost as a percentage of loan amount.

Statistic	Itemizer	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Interest rate (%)	1	129,417	4.01	0.42	2.25	3.88	4.00	4.25	5.75
	0	111,398	4.03	0.47	2.25	3.75	4.10	4.38	5.75
Loan amount (000s)	1	129,417	317.43	92.90	46	249	302	376	721
	0	111,398	140.23	48.07	8	104	140	176	360
Loan term (months)	1	129,417	332.34	63.51	120	360	360	360	360
	0	111,398	314.68	77.92	120	240	360	360	360
LTV (%)	1	129,417	78.26	15.18	5	72	80	90	105
	0	111,398	76.65	17.04	5	70	80	90	105
DTI (%)	1	129,417	33.76	8.78	1	28	35	41	51
	0	111,398	33.07	8.94	1	27	34	40	51
Credit Score	1	129,417	750.88	44.57	620	720	760	787	832
	0	111,398	749.33	47.50	620	715	759	789	839
Monthly payment	1	129,417	1,638.42	521.01	233.65	1,255.24	1,550.88	1,921.93	5,872.91
	0	111,398	759.95	284.47	40.53	559.20	740.98	930.34	2,881.41
Income (000s)	1	129,417	63.85	37.46	9.58	43.38	55.95	74.57	2,425.19
	0	111,398	30.15	16.48	2.35	19.81	26.91	36.56	971.39
Property value (000s)	1	129,417	435.32	199.97	58.75	310.67	389.80	500.00	6,944.44
	0	111,398	193.46	81.77	18	139.3	183.2	233.9	1,067
Mortgage interest (year 1)	1	129,417	12,717.65	3,993.49	1,955	9,897.5	12,035	15,015	36,951
	0	111,398	5,619.24	1,948.92	360	4,160	5,596.2	7,055	11,812
Property tax	1	129,417	4,944.06	2,797.07	623.92	3,105.19	4,267.83	5,997.20	82,405.38
	0	111,398	2,142.37	1,108.52	131.28	1,343.79	1,929.56	2,727.99	20,601.34
State taxes	1	129,417	1,403.49	2,456.20	−2,543.23	86.45	918.99	1,942.01	278,417.30
	0	111,398	357.75	681.30	−2,803.40	0.00	59.52	611.02	8,714.41
Federal deduction	1	129,417	19,301.77	5,894.37	3,254.84	14,933.95	18,069.22	22,246.47	234,675.40
	0	111,398	12,600.00	0.00	12,600	12,600	12,600	12,600	12,600
Marginal tax rate (%)	1	129,417	14.00	3.73	0.00	9.93	14.29	14.87	40.68
	0	111,398	9.66	4.18	−8	7.7	10	10	40
State Top 4	1	129,417	0.19	0.03	0.15	0.18	0.20	0.20	0.39
	0	111,398	0.19	0.04	0.15	0.16	0.19	0.21	0.39
Bank branches	1	129,417	0.29	0.21	0.07	0.16	0.23	0.37	3.85
	0	111,398	0.34	0.25	0.07	0.18	0.27	0.42	4.67
Closing cost (%)	1	129,417	2.15	1.31	0.45	1.32	1.77	2.45	36.42
	0	111,398	3.53	2.14	0.73	2.13	2.97	4.23	46.52

Table 2: The relationship between the mortgage interest rate, and the itemizing status of borrowers

This table reports the estimates of the regressions $rate_{i,m,t} = MSA_m \times Month_t + f(X_i) + \delta Itm_i + \epsilon_{i,m,t}$ *using the loans originated between December 2015 and Mach 2016 securitized by FNMA. Spatio-temporal control is MSA x Months fixed effects (1,414) and risk factors control is a grid of 2,165 unique markets defined as the interaction of Credit score, LTV, DTI, Loan amount, purpose and loan term in which at least one itemizer and one non-itemizer were granted a mortgage. The itemizer dummy is determined by computing the main deductions faced by a borrower; namely the state taxes, the mortgage interest and the property taxes; and comparing it to the standard deduction level for married couple in tax year 2016. Standard errors, presented in parentheses, are clustered at the MSA x Months level. Estimates followed by ***, **, and * are statistically significant at the 1%, 5%, and 10% level, respectively.*

	Dependent variable:				
	Contract interest rate (%)				
	(1)	(2)	(3)	(4)	(5)
Itemizer	-0.022^{***}	-0.014^{***}	0.163^{***}	0.147^{***}	0.149^{***}
	(0.002)	(0.004)	(0.005)	(0.005)	(0.005)
Income - decile 2				0.137^{***}	0.138^{***}
				(0.007)	(0.007)
Income - decile 3				0.262^{***}	0.262^{***}
				(0.010)	(0.010)
Income - decile 4				0.387^{***}	0.386^{***}
				(0.011)	(0.011)
Income - decile 5				0.507^{***}	0.507^{***}
				(0.013)	(0.013)
Income - decile 6				0.646^{***}	0.646^{***}
				(0.015)	(0.015)
Income - decile 7				0.766^{***}	0.765^{***}
				(0.018)	(0.018)
Income - decile 8				0.867^{***}	0.866^{***}
				(0.020)	(0.020)
Income - decile 9				0.966^{***}	0.964^{***}
				(0.023)	(0.023)
Income - decile 10				1.136^{***}	1.135^{***}
				(0.026)	(0.026)
Spatiotemporal control		X	X	X	X
Risk factors control			X	X	X
State FE					X
Observations	240,815	240,207	145,513	145,513	145,513
R^2	0.001	0.064	0.690	0.709	0.709
Adjusted R^2	0.001	0.058	0.683	0.701	0.701

Table 3: The economic significance for borrowers, lenders, and the collection of federal and state taxes

This table reports the changes in payment and taxes if all borrowers who itemize their deductions receive a reduction of interest rates of 0.149 percentage points (Column [5] of Table 2). The change in federal and state taxes is computed with the taxsim software. The discount rate is set at the average of the 10-years treasury bill over the sample period.

	Itemizers			*In-sample total*	
	Pctl(25)	Median	Pctl(75)	Sum (Millions)	% change
Δ Yearly Payment	\$−250.74	\$−304.42	\$−376.35	\$-54.36	−1.09%
PV(Δ PMT) @ 1.92%	-5,161.43	-6,633.35	-8,487.38	1,164.25	1.59
Δ Federal tax	45.29	64.58	84.54	11.50	1.07
Δ State tax	0.00	9.42	84.54	2.45	0.78
Δ Total tax	52.47	75.65	101.65	13.95	1.00
Δ Yearly CFs Borrower	−190.48	−229.74	−279.02	−40.41	−0.63

Table 4: Robustness checks results of the relationship between mortgage rates, and itemizing status

This table reports the estimates of the regressions $rate_{i,m,t} = MSA_m \times Month_t + f(X_i) + \delta Itm_i + \epsilon_{i,m,t}$ *using the loans originated between December 2015 and Mach 2016 securitized by FNMA. Spatio-temporal control is MSA x Months fixed effects (1,414) and risk factors control is a grid of 2,165 unique markets defined as the interaction of Credit score, LTV, DTI, Loan amount, purpose and loan term in which at least one itemizer and one non-itemizer were granted a mortgage. The itemizer dummy is determined by computing the main deductions faced by a borrower; namely the state taxes, the mortgage interest and the property taxes; and comparing it to the standard deduction level for married couple in tax year 2016. Closing cost (%) in Column (1) and (2) is the estimated closing cost as a percentage of the loan amount. Column (3) keeps only small lenders while Column (4) adds lender fixed effects in the regression . Column (5) assumes that all individuals filed their income taxes as single household. Column (6) removes observations for which changing the the marital status from couple to single would change the itemizing status. Standard errors, presented in parentheses, are clustered at the MSA x Months level. Estimates followed by ***, **, and * are statistically significant at the 1%, 5%, and 10% level, respectively.*

	Dependent variable:					
	Contract interest rate (%)					
	(1)	(2)	(3)	(4)	(5)	(6)
Itemizer	0.146***	0.159***	0.153***	0.129***	0.184***	0.340***
	(0.005)	(0.006)	(0.005)	(0.005)	(0.005)	(0.035)
Closing cost (%)	0.019***	0.020***				
	(0.002)	(0.002)				
Itemizer x Closing cost		−0.005***				
		(0.002)				
Spatiotemporal control	X	X	X	X	X	X
Risk factors control	X	X	X	X	X	X
Income decile FE	X	X	X	X	X	X
Lenders FE				X		
Observations	145,513	145,513	84,568	60,945	53,170	12,982
R^2	0.710	0.710	0.710	0.749	0.698	0.712
Adjusted R^2	0.702	0.703	0.697	0.734	0.682	0.676

Table 5: Do itemizers differ in ex-post risks?

This table reports the estimates of $ExPostRisk_{i,m,t} = MSA_m \times Month_t + f(X_i) + \delta Itm_i + rate_{i,m,t} + lengthi, t + \epsilon_{i,m,t}$ *using the loans originated between December 2015 and Mach 2016 securitized by FNMA.* $Ex-postRisk_{i,m,t}$ *is a ex-post performance associated with mortgage i (e.g. default, mortgage loss). The performance data is censored in the second quarter of 2020. Spatio-temporal control is MSA x Months fixed effects (1,414) and risk factors control is a grid of 2,165 unique markets defined as the interaction of the remaining risk factors: Credit score, LTV, DTI, Loan amount, purpose and loan term in which at least one itemizer and one non-itemizer were granted a mortgage. The itemizer dummy is determined by computing the main deductions faced by a borrower; namely the state taxes, the mortgage interest and the property taxes; and comparing it to the standard deduction level for married couple in tax year 2016.* $ratei, m, t$ *is the contract interest rate on mortgage i, and* $lenght_{i,t}$ *controls for the length since origination in months. Standard errors, presented in parentheses, are clustered at the MSA x Months level. Estimates followed by ***, **, and * are statistically significant at the 1%, 5%, and 10% level, respectively.*

	Dependent variable:					
	Prepaid	Ever delinquent	Modified	Foreclose	Modification x log(loss)	Foreclosure x log(loss)
	(1)	(2)	(3)	(4)	(5)	(6)
Itemizer	0.003	−0.001	0.0005	0.00005	0.002	0.001
	(0.003)	(0.002)	(0.001)	(0.0002)	(0.003)	(0.001)
Contract rate (%)	0.028***	0.018***	0.004***	0.0002	0.015***	0.0002
	(0.003)	(0.002)	(0.001)	(0.0002)	(0.004)	(0.001)
Months since origination	−0.025***	0.001***	0.0002***	−0.00002***	0.0004***	−0.0001***
	(0.0001)	(0.00004)	(0.00002)	(0.00000)	(0.00005)	(0.00002)
Spatiotemporal control	X	X	X	X	X	X
Risk factors control	X	X	X	X	X	X
Income decile FE	X	X	X	X	X	X
Observations	145,373	145,373	145,373	145,373	145,370	145,343
R^2	0.576	0.072	0.046	0.038	0.037	0.029
Adjusted R^2	0.565	0.048	0.022	0.014	0.013	0.005

Table 6: Deductibility benefits, and equilibrium interest rates

This table reports the estimates of the regressions $rate_{i,m,t} = MSA_m \times Month_t + f(X_i) + \delta\ Itm_i + \gamma\ (Itm_i \times \tau_i) + \epsilon_{i,m,t}$ *using the loans originated between December 2015 and Mach 2016 securitized by FNMA. Spatio-temporal control is MSA x Months fixed effects (1,414) and risk factors control is a grid of 2,165 unique markets defined as the interaction of Credit score, LTV, DTI, Loan amount, purpose and loan term in which at least one itemizer and one non-itemizer were granted a mortgage. The itemizer dummy is determined by computing the main deductions faced by a borrower; namely the state taxes, the mortgage interest and the property taxes; and comparing it to the standard deduction level for married couple in tax year 2016. Federal MTR is the marginal tax rate with respect to income of borrower i computed by finite differences by the NBER taxsim. Standard errors are presented in parentheses. Standard errors, presented in parentheses, are clustered at the MSA x Months level. Estimates followed by ***, **, and * are statistically significant at the 1%, 5%, and 10% level, respectively.*

	Dependent variable:			
	Contract interest rate (%)			
	(1)	(2)	(3)	(4)
Itemizer	0.147***	0.056***	0.153***	0.048***
	(0.005)	(0.020)	(0.005)	(0.009)
Itemizer x Income (000's)		0.002***		
		(0.0004)		
Federal MTR(%)			0.011***	0.009***
			(0.001)	(0.001)
Itemizer x MTR				0.008***
				(0.001)
Spatiotemporal control	X	X	X	X
Risk factors control	X	X	X	X
Income decile FE	X	X	X	X
Observations	145,513	145,513	145,513	145,513
R^2	0.709	0.711	0.710	0.711
Adjusted R^2	0.701	0.704	0.703	0.704

Table 7: Lenders' market power, and equilibrium interest rates

This table reports the estimates of the regressions $rate_{i,m,t} = MSA_m \times Month_t + f(X_i) + \delta \; Itm_i + \gamma \; (Itm_i \times \theta_{i,j}) + \epsilon_{i,m,t}$ *using the loans originated between December 2015 and Mach 2016 securitized by FNMA. Spatio-temporal control is MSA x Months fixed effects (1,414) and risk factors control is a grid of 2,165 unique markets defined as the interaction of Credit score, LTV, DTI, Loan amount, purpose and loan term in which at least one itemizer and one non-itemizer were granted a mortgage. The itemizer dummy is determined by computing the main deductions faced by a borrower; namely the state taxes, the mortgage interest and the property taxes; and comparing it to the standard deduction level for married couple in tax year 2016. Bank branches index is the number of FIDC insured bank branches scaled by the population at the Zipcode 3-digits level. Top 4 share is the market share (as the number of loan originated) of the biggest four mortgage originator in each US State computing with the HMDA origination dataset. Column (5) adds the interaction between the itemizer dummy and a dummy if the loan originated is a refinancing loan. Standard errors, presented in parentheses, are clustered at the MSA x Months level. Estimates followed by ***, **, and * are statistically significant at the 1%, 5%, and 10% level, respectively.*

	Dependent variable:				
	Contract interest rate (%)				
	(1)	(2)	(3)	(4)	(5)
Itemizer	0.147***	0.137***	0.147***	0.105***	0.145***
	(0.005)	(0.008)	(0.005)	(0.017)	(0.005)
Bank branches	−0.022	−0.041*			
	(0.021)	(0.024)			
Itemizer x bank branches		0.042*			
		(0.025)			
Top 4 share			−0.067	−0.164	
			(0.114)	(0.123)	
Itemizer x Top4				0.223***	
				(0.081)	
Itemizer x Refinancing dummy					0.006
					(0.005)
Spatiotemporal control	X	X	X	X	X
Risk factors control	X	X	X	X	X
Income decile FE	X	X	X	X	X
Observations	145,513	145,513	145,513	145,513	145,513
R^2	0.709	0.709	0.709	0.709	0.709
Adjusted R^2	0.701	0.701	0.701	0.701	0.701

Table 8: Summary statistics of the NSMO data

This table reports the summary statistics of the variables used from the NSMO data. All respondents from the first quarter of 2014 until the third quarter of 2017 are included. The itemizer dummy is computed with the taxsim by providing borrowers characteristics (age, number of dependents, income, marital status etc.). Only borrowers that would be itemizers or non-itemizers in all of the 50 states and Washington D.C. are kept in this table (65% of the full sample). Interest rate spread is the spread over the 10-years treasury bill of the origination month. Sophistication is an index characterizing the degree of familiarity of each borrowers with the mortgage origination based on question 5 of the survey.

Statistic	Itemizer	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Interest rate spread (%)	1	5,228	1.94	0.66	−0.29	1.58	1.89	2.24	3.64
	0	2,340	1.67	0.73	−0.95	1.16	1.70	2.12	3.64
Loan amount (000s)	1	5,228	260.20	77.66	75	225	275	325	375
	0	2,340	99.49	27.25	75	75	75	125	225
Loan term (years)	1	5,228	27.13	5.69	10	30	30	30	30
	0	2,340	22.34	7.67	10	15	20	30	30
LTV (%)	1	5,228	74.47	17.45	5	67	79	87	125
	0	2,340	71.66	17.76	18	60	75	83	124
DTI (%)	1	5,228	34.98	12.15	3	27	35	42	100
	0	2,340	35.16	14.50	8	26	34	42	100
Credit Score	1	5,228	756.34	55.07	487	724	769	800	838
	0	2,340	749.88	59.47	476	711	759	801	839
Income (000s)	1	5,228	90.13	40.74	17.41	59.73	86.95	110.46	426.93
	0	2,340	48.12	19.06	16.83	33.18	45.61	59.83	115.25
Age	1	5,228	46.98	12.46	21	36	47	57	93
	0	2,340	51.64	14.31	21	40	53	63	90
Number of search	1	5,228	1.71	0.83	1	1	2	2	5
	0	2,340	1.61	0.79	1	1	1	2	5
Number of application	1	5,228	1.27	0.58	1	1	1	1	5
	0	2,340	1.24	0.53	1	1	1	1	5
Sophistication score	1	5,228	10.94	3.22	0	9	12	14	14
	0	2,340	10.40	3.34	0	8	11	14	14

Table 9: The relationship between search intensity and borrowers' itemizing status

This table reports the regressions results testing the relationship between mortgage search and itemizing status. The main specification takes the form: $Search_{i,t,m} = \alpha_t + \delta ITM_i + X_i\beta + Z_i\gamma + \epsilon_{i,t,m}$*; where* $Search_{i,t,m}$ *is the number of search,* ITM_i *is the computed probability that borrower i is itemizing her deductions,* α_t *are month fixed effects,* X_i *are mortgage characteristics fixed effects (LTV, Loan amount, DTI, Credit score, loan term, refinance), and* Z_i *are borrowers characteristics. The itemizer probability is computed with the taxsim by providing borrowers characteristics (age, number of dependents, income, marital status etc.) for each possible state of residency and weighting the results by the state population. Sophistication is an index characterizing the degree of familiarity of each borrowers with the mortgage origination based on question 5 of the survey. Column (5) and (6) presents the results of a binomial probit regressions. Standard errors are presented in parentheses. Estimates followed by ***, **, and * are statistically significant at the 1%, 5%, and 10% level, respectively.*

	Dependent variable:					
	Search			Application	search>1	application>1
	(1)	(2)	(3)	(4)	(5)	(6)
Itemizer probability	0.108***	−0.038	−0.032	0.006	−0.085*	−0.012
	(0.018)	(0.028)	(0.028)	(0.019)	(0.044)	(0.050)
Age (> 60)			−0.065***	−0.016	−0.122***	−0.050
			(0.022)	(0.015)	(0.031)	(0.037)
Age (< 40)			0.030*	0.024*	0.045	0.066**
			(0.018)	(0.013)	(0.029)	(0.033)
College degree			0.112***	0.005	0.203***	0.010
			(0.026)	(0.018)	(0.040)	(0.045)
Sophistication			0.006***	−0.005***	0.009**	−0.012***
			(0.002)	(0.002)	(0.004)	(0.004)
Refinance		−0.044**	−0.038**	−0.080***	−0.106***	−0.218***
		(0.018)	(0.018)	(0.013)	(0.028)	(0.032)
Income FE			X	X	X	X
Month FE		X	X	X	X	X
Mortgage characteristics FE		X	X	X	X	X
Observations	11,938	11,938	11,938	11,938	11,938	11,938
R^2	0.003	0.016	0.021	0.029		
Log Likelihood					−8,136.569	−5,839.100

Table 10: The relationship between itemizing status, interest rate and borrowers' search intensity

This table reports the estimates of the residual regressions $spread_{i,t} = Month_t + f(X_i) + \delta Itm_i + \mu Search_i + Z_i\gamma + \epsilon_{i,t}$ *using the loans information in the NSMO data.* $spread_{i,t}$ *is the contract interest rate spread over the 10-years treasury bill,* X_i *are mortgage level characteristics,* Itm_i *is the probability that borrower i is an itemizer,* $Search_i$ *is the number of search,* Z_i *are borrowers characteristics, and* $\epsilon_{i,t}$ *is the error term. Macro control are months fixed effects (48) and risk control is a grid of 3,476 unique markets defined as the interaction of Credit score, LTV, DTI, Loan amount, purpose and loan term. The itemizer probability is computed with the taxsim by providing borrowers characteristics (age, number of dependents, income, marital status etc.) for each possible state of residency and weighting the results by the state population. Standard errors are presented in parentheses. Estimates followed by ***, **, and * are statistically significant at the 1%, 5%, and 10% level, respectively.*

	Dependent variable:					
	Contract Interest Rate Spread (%)					
	(1)	(2)	(3)	(4)	(5)	(6)
Itemizer probability	0.225***	0.216***	0.215***	0.186***	0.216***	0.248***
	(0.015)	(0.015)	(0.015)	(0.029)	(0.015)	(0.032)
Nb. of search			−0.015**	−0.024**		
			(0.006)	(0.010)		
Itemizer x Search				0.018		
				(0.015)		
Nb. of application					−0.003	0.011
					(0.009)	(0.015)
Itemizer x Application						−0.026
						(0.022)
Age (> 60)		0.016	0.015	0.014	0.016	0.016
		(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Age (< 40)		−0.019	−0.018	−0.019	−0.019	−0.019
		(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
College degree		−0.041***	−0.040**	−0.039**	−0.041***	−0.041***
		(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Sophistication score		−0.008***	−0.008***	−0.008***	−0.008***	−0.008***
		(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Macro control	X	X	X	X	X	X
Market control	X	X	X	X	X	X
Linear Risk control	X	X	X	X	X	X
Income decile FE		X	X	X	X	X
Observations	7,748	7,748	7,748	7,748	7,748	7,748
R^2	0.612	0.615	0.615	0.616	0.615	0.615

APPENDIX

A Information, lenders' markups, and the capitalization of MID

Lenders are profit maximizers providing unlimited mortgage at constant marginal cost. When a borrower with income y and house price p meets with a lender, the latter solves the following profit maximization problem:

$$\max_{r_i} \quad (r_i - c) \times \frac{p - (\delta(1 + r_i(1 - E[\tau_i])))^{1/\xi}(y - p)}{(\delta(1 + r_i(1 - E[\tau_i])))^{1/\xi} + (1 + r_i(1 - E[\tau_i]))} = \max_{r_i}(r_i - c)E[m^*(r_i, \tau_i)|Z_i] \tag{9}$$

where the expectation is taken over i's marginal income tax given information available about borrower i: Z_i.

Perfect information. Consider a monopolist that have perfect information about τ_i (i.e. $E[\tau_i] = \tau_i \, \forall i$), the first order condition specifies:

$$r_i^*(c, \tau_i) = c + \frac{m^*(r_i^*, \tau_i)}{-m_r^*(r_i^*, \tau_i)} \tag{10}$$

where $m_r^*(r_i^*, \tau_i)$ is the partial derivative of mortgage demand evaluated at the optimal offered rate. Given that $m^*(r_i^*, \tau_i)$ increases with τ_i and that $-m_r^*(r_i^*, \tau_i)$ decreases with τ_i; the mortgage rate charged by the monopolist must increase once goods are deductibles under perfect information. The monopolist price schedule also increase with τ_i, allowing him to price discriminate (under specific condition outlined below).

No information. If the monopolist has no available information to distinguish borrowers, he would maximize the profit function with respect to its prior about the distribution of marginal tax rate in the market ($E[\tau_i] = \bar{\tau}$). The price schedule would be constant (with respect to marginal tax rate) at:

$$r_i^*(c) = c + \frac{m^*(r_i^*, \bar{\tau})}{-m_r^*(r_i^*, \bar{\tau})} \tag{11}$$

where $m_r^*(r_i^*, \bar{\tau})$ is here a constant. Despite the impossibility to price discriminate in this situation, note that the monopolist still capitalize the tax benefits into its pricing schedule. As long as lenders do not assume $E[\tau_i] = 0 \, \forall i$, there would be tax capitalization in mortgage markups once mortgage interest is deductible. MID capitalization is therefore not conditional on price discrimination as uniform pricing also allows capitalization. Note however, that in this setting, the heterogeneous elasticities emerge from a tax instrument, rather than heterogeneous preferences.

Informative signal. The above two cases represent bounds on the mortgage pricing schedule for the monopolist after the introduction of deductibility. Given that lenders have access to borrowers' information in the underwriting process, we can expect that $E[\tau_i|Z_i] = \bar{\tau} + \nu_i$ where $\nu_i = f(Z_i) \in [0, \tau_i - \bar{\tau}]$. As long as $\nu_i \neq \nu_k$ for at least one {i, k} pair, there will be price discrimination from the monopolist. The price schedule would therefore be:

$$r_i^*(c, \nu_i) = c + \frac{m^*(r_i^*, \bar{\tau} + \nu_i)}{-m_r^*(r_i^*, \bar{\tau} + \nu_i)} \tag{12}$$

where the monopolist's pricing schedule is based on the borrowers' signals. Therefore, lenders must have additional information about borrowers deductibility benefits.

From the monopolist case analysis, we can see that $\eta^I \neq \eta^{NI}$ is sufficient for MID capitalization into the mortgage lending industry. However, price discrimination is conditional on having informative information about borrowers' income tax status orthogonal to risk-based information.

The monopolist surplus over N consumer with identical marginal cost can therefore be shown to be:

$$PS(E[\tau_i]) = \sum_{i=1}^{N} \frac{m^*(r^*(c, E[\tau_i]), \tau_i)}{-m_r^*(r^*(c, E[\tau_i]), \tau_i)} \cdot m^*(r^*(c, E[\tau_i]), \tau_i) \tag{13}$$

where $E[\tau_i] \in \{0\,;\, \bar{\tau}\,;\, \bar{\tau} + \nu_i\,;\, \tau_i\}$ depending on the level of lenders' information.

It follows that the MID capitalization into the monopolist's surplus is increasing in the degree of informativeness:

$$\begin{aligned} &\underbrace{\sum_{i=1}^{N} \frac{m^*(r^*(c,0),\tau_i)}{-m_r^*(r^*(c,0),\tau_i)} \cdot m^*(r^*(c,0)),\tau_i)}_{\text{No capitalization}} \quad \leq \quad \underbrace{\sum_{i=1}^{N} \frac{m^*(r^*(c,\bar{\tau}),\tau_i)}{-m_r^*(r^*(c,\bar{\tau}),\tau_i)} \cdot m^*(r^*(c,\bar{\tau})),\tau_i)}_{\text{No discrimination but capitalization}} \\ \leq \quad &\underbrace{\sum_{i=1}^{N} \frac{m^*(r^*(c,\bar{\tau}+\nu_i),\tau_i)}{-m_r^*(r^*(c,\tau \bar{+} \nu_i),\tau_i)} \cdot m^*(r^*(c,\bar{\tau}+\nu_i),\tau_i)}_{\text{Imperfect price discrimination}} \quad \leq \quad \underbrace{\sum_{i=1}^{N} \frac{m^*(r^*(c,\tau_i),\tau_i)}{-m_r^*(r^*(c,\tau_i),\tau_i)} \cdot m^*(r^*(c,\tau i)),\tau_i)}_{\text{First degree price discrimination}} \end{aligned} \tag{14}$$

Introducing market friction parameter as in the theoretical framework, the total producer surplus is:

$$PS(E[\tau_i], \theta_{i,j}) = \sum_{j=1}^{J} \sum_{i=1}^{N} \theta_{i,j} \frac{m^*(r^*(c, E[\tau_i]), \tau_i)}{-m_r^*(r^*(c, E[\tau_i]), \tau_i)} \cdot m^*(r^*(c, E[\tau_i]), \tau_i) \tag{15}$$

which is also increasing in the degree of information lenders have with respect to the marginal tax rate of borrowers.

B Estimating the degree of market frictions

The above section shows that the mortgage market does not follow the paradigm of perfect competition. This section aims at estimating the degree of market power (market frictions) that allows price discrimination with respect to itemizing status. This section follows closely Graddy (1995).

Given that the consumers elasticity with respect to salient price has been shown higher than the elasticity with respect to non-salience taxes (Chetty et al., 2009), one would need to compute the elasticity for both itemizers and non-itemizers (i.e. we cannot assume that the elasticity of itemizers is a function of the elasticity of the non-itemizers and their MTR). Assuming that the mortgage market for itemizers and non-itemizers are distinct, we can estimate separately the demand functions such as:

$$\log q^I_{m,t} = \alpha_m + \alpha_t + \alpha_1 \log r^I_{m,t} + \alpha_2 \log Inc_{m,t} + \epsilon^I_{m,t} \tag{16}$$

$$\log q^{NI}_{m,t} = \omega_m + \omega_t + \omega_1 \log r^{NI}_{m,t} + \omega_2 \log Incm,t + \epsilon^{NI}_{m,t} \tag{17}$$

where $q^t_{m,t}$ is the quantity of mortgage amounts originated to consumer of type $t \in \{I, NI\}$ in market m at time t, $r^I_{m,t}$ is the mean contract mortgage rate in this market and X is a matrix a demand shifter.

Following Porter (1983) and the theoretical model above, in a non-competitive industry, we know that:

$$r^I_{m,t}(1 + \theta\tilde{\alpha_1}) = MC^I_{m,t} \tag{18}$$

$$r^{NI}_{m,t}(1 + \theta\tilde{\omega_1}) = MC^{NI}_{m,t} \tag{19}$$

where $\tilde{\alpha_1}$ denotes the inverse demand elasticity (i.e. $1/\alpha_1$). With marginal cost between the two market differing by an additive constant(i.e. $MC^I_{m,t} = MC_{m,t} + \nu^I_{m,t}$), we have:

$$r^I_{m,t} = \underbrace{\frac{1 + \theta\tilde{\omega_1}}{1 + \theta\tilde{\alpha_1}}}_{\beta} r^{NI}_{m,t} + \pi_{m,t} \tag{20}$$

where $\pi_{m,t} = (\nu^{NI}_{m,t} - \nu^I_{m,t})/(1 + \theta\tilde{\alpha_1})$. One could estimate β from the prices observed in the markets. With the estimates of the elasticity of demand from the demand functions, we can retrived the degree of market frictions θ such that:

$$\theta = \frac{1 - \beta}{\beta\tilde{\alpha_1} - \tilde{\omega_1}} \tag{21}$$

Estimation.

- I aggregate market at the MSA/month level

- Demand shifters include the mean income in the market as well as a month and MSA fixed effects to control for potential spatial and temporal demand shifts.

- Endogeneity: Because the rate offered in each market is endogenously determined with quantity, instruments are needed for the variations of prices across markets and time. I use supply side price shifter as instrument; namely the avarage level of Loan to Value (LTV), Debt to Income (DTI), and the average credit score (FICO).[29]

- Non-spherical error: Given some potential correlation structure between the error terms of the demand equations, I use 3SLS with iterative procedure to control for the simultaneity bias.

- Equation beta is estimated using the market definition in ....

Table A1: Estimates of the demand function for mortgage

	3SLS iterated		2SLS		SUR		OLS	
	$\log q^I_{m,t}$	$\log q^{NI}_{m,t}$	$\log q^I_{m,t}$	$\log q^{NI}_{m,t}$	$\log q^I_{m,t}$	$\log q^{NI}_{m,t}$	$\log q^I_{m,t}$	$\log q^{NI}_{m,t}$
$\log r^I_{m,t}$	-1.277		-1.307		-0.634		-0.928	
	(0.808)		(0.809)		(0.402)		(0.413)	
$\log r^I_{m,t}$		-2.775		-2.705		-2.816		-2.846
		(1.214)		(1.214)		(0.494)		(0.508)
$\log Incm, t$	0.114	0.567	0.109	0.560	0.126	0.538	0.119	0.553
	(0.084)	(0.109)	(0.084)	(0.109)	(0.079)	(0.092)	(0.081)	(0.094)
MSA FE	X	X	X	X	X	X	X	X
Month	X	X	X	X	X	X	X	X
Observations	1,342	1,342	1,342	1,342	1,342	1,342	1,342	1,342
R^2	0.955	0.936						
Adjusted R^2	0.937	0.911						

- Results are presented in Table A1 for different estimation methods with the 3SLS being the preferred specification. First stage results are shown in Table A3.

- The results show that itemizers are less price sensitive than non-itemizers as expected.

- The income elasticity of demand is greater for non-itemizers which is also consistent with the idea of diminishing share of housing consumption.

- The first stage coefficients are in accordance with the expectations. More risky mortgage leads to higher prices.

[29] In other specification not presented here, I include the Saiz (2010) measures of supply elasticity for MSA variations and the 10-years treasury rates as time-series variations. However, because of the introduction of MSA and month fixed effects, these instruments are not identified.

Table A2: Estimates of of market friction parameters

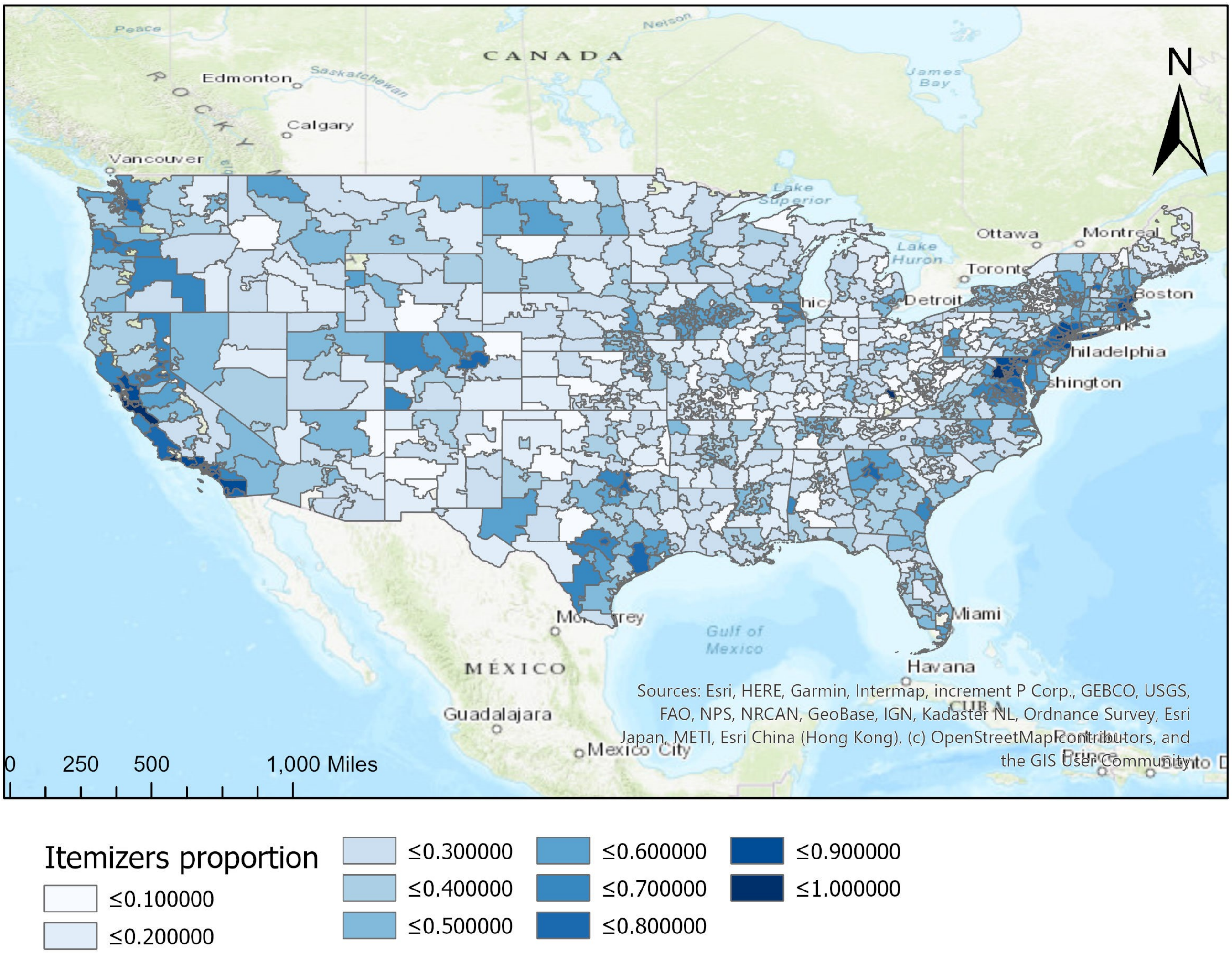
	3SLS	2SLS	SUR	OLS
θ	0.058	0.062	0.020	0.034

Table A3: First stage regression estimates

	Dependent variable:			
	$\log r^{I}_{m,t}$	$\log r^{NI}_{m,t}$	$\log r^{I}_{m,t}$	$\log r^{NI}_{m,t}$
SAIZ			0.0003^{***}	0.0002^{**}
			(0.0001)	(0.0001)
GS10			0.051^{***}	0.057^{***}
			(0.007)	(0.007)
LTV	0.001^{***}	0.001^{***}	0.001^{**}	-0.00002
	(0.0002)	(0.0002)	(0.0003)	(0.0002)
DTI	0.0005	-0.002^{***}	0.002^{***}	0.001^{***}
	(0.0004)	(0.0004)	(0.0004)	(0.0005)
FICO	-0.001^{***}	-0.001^{***}	-0.001^{***}	-0.001^{***}
	(0.00005)	(0.0001)	(0.0001)	(0.0001)
log(Income)	0.001	-0.063^{***}	0.034^{***}	-0.006
	(0.009)	(0.007)	(0.008)	(0.007)
Constant			-2.903^{***}	-2.883^{***}
			(0.080)	(0.074)
MSA FE	X	X	X	X
Month	X	X X	X	X
Observations	1,342	1,342	853	853
R^2	0.683	0.766	0.294	0.152
F Statistic	5.385^{***}	8.169^{***}	58.688^{***}	25.231^{***}
Note:				$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

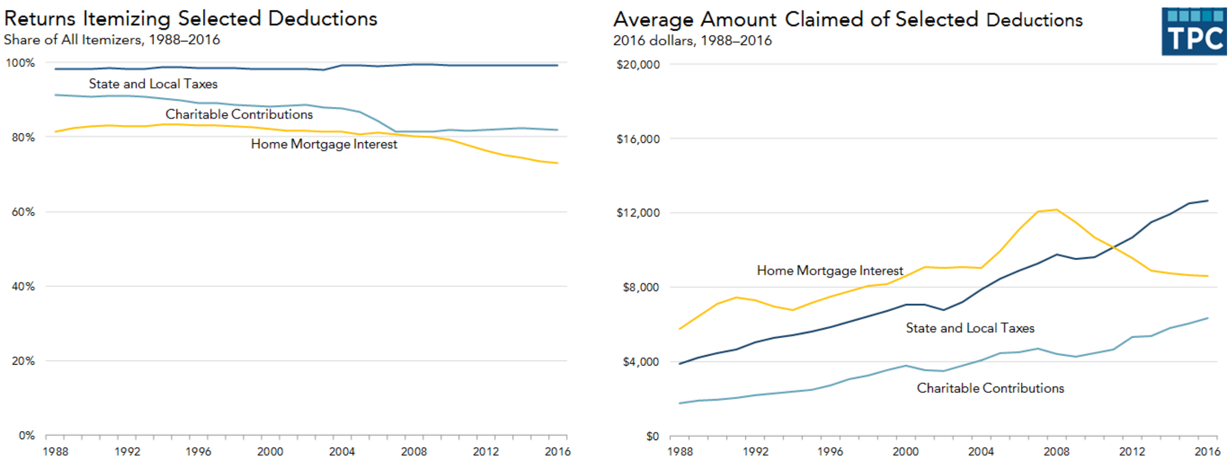
C Additional figures

Figure A1: Continental United States share of Itemizers per Zipcode 3-digits



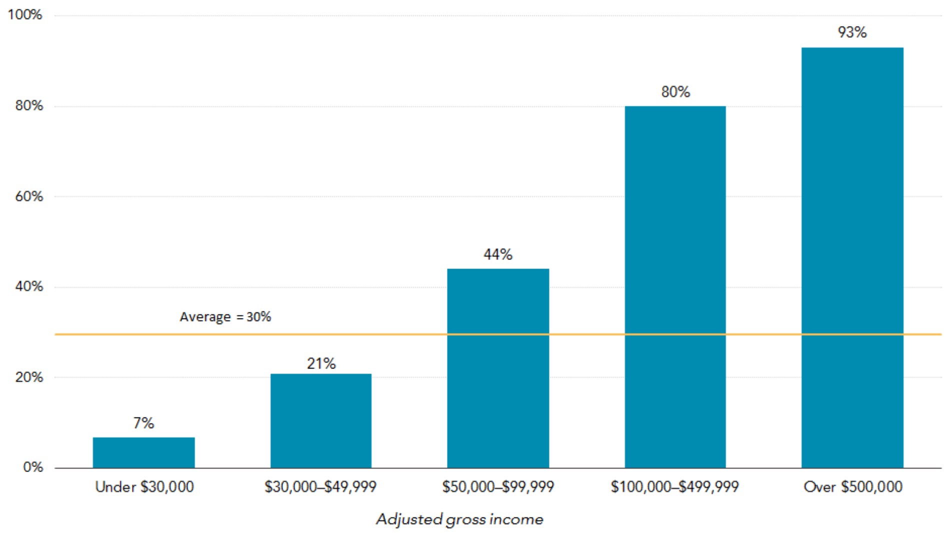
Note: this map shows the geographical distribution of the borrowers classified as itemizers using the loans originated between December 2015 and Mach 2016 securitized by FNMA. The itemizer dummy is determined by computing the main deductions faced by a borrower; namely the state taxes, the mortgage interest and the property taxes; and comparing it to the standard deduction level for married couple in tax year 2016. For each zipcode 3-digits, this map reports the proportion of borrower-itemizers over the total number of borrowers.

Figure A2: Itemizable deductions share, 1988-2016



Source: Internal Revenue Service. Statistics of Income. Basic Tables: Exemptions and Itemized Deductions. Table 2.1. "Returns with Itemized Deductions: Sources of Income, Adjustments, Itemized Deductions by Type, Exemptions, and Tax Items by Size of Adjusted Gross Income," Tax Year 2016.
Note: Real amounts are calculated using GDP deflators (Federal Reserve Bank of St. Louis, "Gross Domestic Product: Implicit Price Deflator," https://fred.stlouisfed.org/series/GDPDEF#0).

Figure A3: Share of itemizers per income brackets, 2016 fiscal year



Source: Internal Revenue Service. Statistics of Income. Table 1.2. "All Returns: Adjusted Gross Income, Exemptions, Deductions, and Tax Items, by Size of Adjusted Gross Income and by Marital Status," Tax Year 2016.

D Additional tables

Table A4: The itemizer interest-gap over time

Panel A of this table reports the estimates of the regressions $rate_{i,m,t} = MSA_m \times Month_t + f(X_i) + \delta Itm_i + \epsilon_{i,m,t}$ *by year. Panel B includes in addition 10 income fixed effects. The data comprises of all loans originated during the first quarter of 20111 to 2018. Spatio-temporal control is MSA x Months fixed effects and risk factors control is a grid of unique markets defined as the interaction of Credit score, LTV, DTI, Loan amount, purpose and loan term in which at least one itemizer and one non-itemizer were granted a mortgage. The itemizer dummy is determined by computing the main deductions faced by a borrower; namely the state taxes, the mortgage interest and the property taxes; and comparing it to the standard deduction level for married couple in each tax year. Standard errors, presented in parentheses, are clustered at the MSA x Months level. Estimates followed by ***, **, and * are statistically significant at the 1%, 5%, and 10% level, respectively.*

PANEL A: Without income fixed effects

	Dependent variable: Contract interest rate (%)							
	2011	2012	2013	2014	2015	2016	2017	2018
Itemizer	0.190***	0.152***	0.123***	0.124***	0.157***	0.163***	0.225***	0.218***
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.007)	(0.007)
Spatiotemporal control	X	X	X	X	X	X	X	X
Risk factros control	X	X	X	X	X	X	X	X
Observations	143,847	222,573	275,026	99,835	164,981	145,513	181,051	94,774
R^2	0.689	0.695	0.690	0.786	0.711	0.690	0.676	0.543
Adjusted R^2	0.682	0.690	0.686	0.779	0.704	0.683	0.669	0.532

PANEL B: With income fixed effects

	Dependent variable: Contract interest rate (%)							
	2011	2012	2013	2014	2015	2016	2017	2018
Itemizer	0.173***	0.138***	0.108***	0.113***	0.140***	0.147***	0.191***	0.201***
	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.005)	(0.006)	(0.009)
Spatiotemporal control	X	X	X	X	X	X	X	X
Risk factros control	X	X	X	X	X	X	X	X
Income decile FE	X	X	X	X	X	X	X	X
Observations	143,847	222,573	275,026	99,835	164,981	145,513	181,051	94,774
R^2	0.703	0.706	0.699	0.795	0.729	0.709	0.705	0.583
Adjusted R^2	0.696	0.701	0.695	0.789	0.723	0.701	0.699	0.573

Table A5: The effects of higher borrowers' bargaining power on the interest-gap

This table reports the estimates of the residual regressions $Spread_i = \alpha + \delta ITM_i + \gamma_1 z_i + \gamma_2 z_i \gamma \times ITM_i + X_i\beta + \epsilon_i$ *using the loans information in the NSMO data.* $Spread_i$ *is the spread of mortgage i over the 10-years treasury bill after controlling for macro and market specific risks. Macro control are months fixed effects (48) and risk control is a grid of 3,425 unique markets defined as the interaction of Credit score, LTV, DTI, Loan amount, purpose and loan term.The itemizer probability is computed with the taxsim by providing borrowers characteristics (age, number of dependents, income, marital status etc.) for each possible state of residency and weighting the results by the state population.* z_i *is the measure of borrowers' bargaining leverage namely . The* R^2 *are the overall* R^2 *and not the* R^2 *from the residuals regressions. Standard errors are presented in parentheses. Estimates followed by ***, **, and * are statistically significant at the 1%, 5%, and 10% level, respectively*

	Dependent variable:							
	Rate spread (%)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Itemizer	0.220***	0.259***	0.217***	0.227***	0.221***	0.225***	0.218***	0.277***
	(0.016)	(0.055)	(0.016)	(0.044)	(0.016)	(0.017)	(0.016)	(0.067)
College degree	−0.013***	−0.009					−0.013***	−0.008
	(0.005)	(0.007)					(0.005)	(0.007)
Itemizer x College		−0.008						−0.010
		(0.011)						(0.011)
Sophistication score			−0.007***	−0.007***			−0.008***	−0.007***
			(0.002)	(0.003)			(0.002)	(0.003)
Itemizer x Sophistication				−0.001				−0.001
				(0.004)				(0.004)
Age (> 60)					0.032**	0.042**	0.035**	0.045**
					(0.014)	(0.021)	(0.014)	(0.021)
Itemizer x Age (> 60)						−0.021		−0.021
						(0.034)		(0.034)
Observations	7,403	7,403	7,403	7,403	7,403	7,403	7,403	7,403
R^2	0.580	0.580	0.581	0.581	0.581	0.581	0.582	0.582

Table A6: Correlation table of the FNMA dataset

This table reports the correlation coefficients of the main variables used from the loan level origination data. All loans securitized by Fannie Mae from December 2015 to March 2016 are included. The itemizer dummy is determined by computing the main deductions faced by a borrower; namely the state taxes, the mortgage interest and the property taxes; and comparing it to the standard deduction level for married couple in tax year 2016. The marginal tax rate is computed with respect to borrower's income by finite difference by the NBER taxsim. Bank branches index is the number of FIDC insured bank branches divided by the number of loan originated at the Zipcode 3-digits level. Top 4 share is the market share (as the number of loan originated) of the biggest four mortgage originator in each US State computing with the HMDA origination dataset. Closing cost (%) is the estimated mortgage closing cost as a percentage of loan amount.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
(1) Interest rate (%)	1	-0.05	0.69	0.34	0.19	-0.28	-0.22	-0.26	-0.19	0.16	-0.20	-0.20	-0.002	-0.02	-0.29	-0.03	-0.05	0.01
(2) Loan amount (000s)	-0.05	1	0.15	0.10	0.11	0.03	0.94	0.60	0.80	0.97	0.57	0.31	0.85	0.76	0.59	0.18	-0.14	-0.48
(3) Loan term (months)	0.69	0.15	1	0.35	0.15	-0.10	-0.15	-0.20	-0.03	0.29	-0.07	-0.17	0.09	0.12	-0.22	0.02	-0.04	-0.14
(4) LTV (%)	0.34	0.10	0.35	1	0.13	-0.16	0.001	-0.09	-0.38	0.17	-0.24	-0.07	-0.03	0.05	-0.04	-0.09	0.03	-0.29
(5) DTI (%)	0.19	0.11	0.15	0.13	1	-0.18	0.07	-0.45	0.03	0.14	0.02	-0.36	-0.01	0.04	-0.46	0.04	-0.05	-0.05
(6) Credit Score	-0.28	0.03	-0.10	-0.16	-0.18	1	0.04	0.12	0.09	-0.03	0.07	0.10	0.01	0.03	0.13	0.03	-0.01	0.004
(7) Monthly payment	-0.22	0.94	-0.15	0.001	0.07	0.04	1	0.68	0.80	0.87	0.58	0.37	0.81	0.72	0.65	0.16	-0.13	-0.44
(8) Income (000s)	-0.26	0.60	-0.20	-0.09	-0.45	0.12	0.68	1	0.57	0.53	0.42	0.79	0.66	0.49	0.77	0.09	-0.07	-0.28
(9) Property value (000s)	-0.19	0.80	-0.03	-0.38	0.03	0.09	0.80	0.57	1	0.73	0.68	0.31	0.77	0.61	0.51	0.20	-0.11	-0.26
(10) Mortgage interest (year 1)	0.16	0.97	0.29	0.17	0.14	-0.03	0.87	0.53	0.73	1	0.51	0.25	0.84	0.74	0.52	0.17	-0.15	-0.47
(11) Property tax	-0.20	0.57	-0.07	-0.24	0.02	0.07	0.58	0.42	0.68	0.51	1	0.20	0.73	0.54	0.36	0.10	0.04	-0.14
(12) State taxes	-0.20	0.31	-0.17	-0.07	-0.36	0.10	0.37	0.79	0.31	0.25	0.20	1	0.48	0.27	0.48	0.07	0.03	-0.14
(13) Federal deduction	-0.002	0.85	0.09	-0.03	-0.01	0.01	0.81	0.66	0.77	0.84	0.73	0.48	1	0.61	0.48	0.16	-0.03	-0.27
(14) Itemizer dummy	-0.02	0.76	0.12	0.05	0.04	0.03	0.72	0.49	0.61	0.74	0.54	0.27	0.61	1	0.48	0.11	-0.10	-0.37
(15) Marginal tax rate (%)	-0.29	0.59	-0.22	-0.04	-0.46	0.13	0.65	0.77	0.51	0.52	0.36	0.48	0.48	0.48	1	0.07	-0.12	-0.39
(16) State Top 4	-0.03	0.18	0.02	-0.09	0.04	0.03	0.16	0.09	0.20	0.17	0.10	0.07	0.16	0.11	0.07	1	-0.01	0.05
(17) Bank branches	-0.05	-0.14	-0.04	0.03	-0.05	-0.01	-0.13	-0.07	-0.11	-0.15	0.04	0.03	-0.03	-0.10	-0.12	-0.01	1	0.20
(18) Closing cost (%)	0.01	-0.48	-0.14	-0.29	-0.05	0.004	-0.44	-0.28	-0.26	-0.47	-0.14	-0.14	-0.27	-0.37	-0.39	0.05	0.20	1

Table A7: Correlation table of the NSMO data

This table reports the correlation coefficients of the main variables used from the NSMO data. All respondents from the first quarter of 2014 until the third quarter of 2017 are included. The probability of itemizing is computed with the taxsim by providing borrowers characteristics (age, number of dependents, income etc.) for each possible state of residency and weighting the results by the state population. Interest rate spread is the spread over the 10-years treasury bill of the origination month. Sophistication is an index characterizing the degree of familiarity of each borrowers with the mortgage origination based on question 5 of the survey.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
(1) Itemizer probability	1	0.13	0.72	0.24	0.06	-0.02	0.04	0.41	0.39	-0.12	0.05	0.02	0.07
(2) Interest rate spread (%)	0.13	1	0.03	0.57	0.26	0.13	-0.21	-0.16	-0.13	-0.05	-0.05	0.02	-0.14
(3) Loan amount (000s)	0.72	0.03	1	0.22	0.16	0.05	0.04	0.52	0.31	-0.14	0.08	0.03	0.11
(4) Loan term (months)	0.24	0.57	0.22	1	0.34	0.10	-0.11	-0.18	-0.15	-0.14	0.03	0.06	-0.10
(5) LTV (%)	0.06	0.26	0.16	0.34	1	0.10	-0.20	-0.07	-0.05	-0.34	0.01	0.07	-0.13
(6) DTI (%)	-0.02	0.13	0.05	0.10	0.10	1	-0.22	-0.36	-0.31	0.09	-0.01	0.04	-0.07
(7) Credit Score	0.04	-0.21	0.04	-0.11	-0.20	-0.22	1	0.13	0.11	0.09	0.02	-0.05	0.15
(8) Income (000s)	0.41	-0.16	0.52	-0.18	-0.07	-0.36	0.13	1	0.59	-0.05	0.03	-0.04	0.20
(9) Marginal tax rate (%)	0.39	-0.13	0.31	-0.15	-0.05	-0.31	0.11	0.59	1	-0.07	0.02	-0.01	0.11
(10) Age	-0.12	-0.05	-0.14	-0.14	-0.34	0.09	0.09	-0.05	-0.07	1	-0.05	-0.07	0.15
(11) Number of search	0.05	-0.05	0.08	0.03	0.01	-0.01	0.02	0.03	0.02	-0.05	1	0.43	0.03
(12) Number of application	0.02	0.02	0.03	0.06	0.07	0.04	-0.05	-0.04	-0.01	-0.07	0.43	1	-0.05
(13) Sophistiction score	0.07	-0.14	0.11	-0.10	-0.13	-0.07	0.15	0.20	0.11	0.15	0.03	-0.05	1