Preliminary

Not for redistribution

First-Time Homebuyers: The Role of Thin Credit Files

Joseph Gyourko[1]

Donghoon Lee[2]

Joseph Tracy[2]

June 2015

Abstract

We investigate the performance of FHA first-time buyers and the impact that thin credit files have on their default risk. We find that, holding other factors constant, first-time borrowers with thin credit files have around a five percentage point higher likelihood of serious delinquency. This indicates that if underwriters are aware of thin credit files, then they are unsuccessful in fully compensating for this when making their credit decisions. The impact of thin credit files is concentrated in the early years of the mortgage. In contrast, we do not find that short credit files have any significant difference in default risk.

[1] The Wharton School, University of Pennsylvania and NBER

[2] Federal Reserve Bank of New York

A core mission of the Federal Housing Administration (FHA) is to support lending to first-time homebuyers and low- and moderate income households. For 2014, the FHA reported that 81 percent of their purchase mortgage originations went to first-time buyers.[1] However, the FHA's definition of a first-time buyer includes previous homeowners if there has been at least three years since their last ownership experience, as well as cases where a married couple co-signs for the mortgage and at least one has never previously signed for a mortgage. Using a panel dataset of credit files that we will describe later, we can recalculate the FHA's first-time ownership share restricting the definition to cases where the borrower(s) have never previously had a mortgage. For the three years from 2012 to 2014, we estimate that around 60 percent of the FHA's purchase mortgages go to first-time buyers.

While first-time home buyers are an important segment of the FHA borrower population, the FHA is not the only lender to first-time buyers. Figure 1 shows the volume of first-time purchase mortgages from 2000 to 2013 (using our strict definition) broken down by FHA, GSE and other. The FHA's relative share of the flow of new first-time purchase mortgages has changed over these years. During the height of the private label security lending, both the FHA's overall and the first-time buyer shares were significantly reduced. Following the onset of the financial crisis, the FHA's first-time share rose and even exceeded that of the GSEs. In the last couple of years, the FHA and GSE first-time buyer shares have been similar, while the FHA's "other" share remains well below its pre-crisis level.

An implication of a first-time buyer is that the mortgage must be underwritten taking into account that, by definition, the borrower has no previous track record of performance on a mortgage. In a number of these cases, the borrower may even have had no significant experience to date managing a credit balance – what we will call a "thin" credit file. While borrowers with thin credit files have an assigned credit score, what is this score likely to be and how reliable is this score at predicting their future performance on the mortgage? If credit scores are less reliable for thin files, do underwriters compensate by requiring a higher credit score and/or that other aspects of the application have to be even stronger than otherwise? How successful are underwriters at adjusting for thin files? Given the very high exposure of the FHA to first-time buyers, these questions may be of particular importance for the FHA.

---

[1] U.S. Department of Housing and Urban Development (2014), page 18.

We investigate the performance of FHA first-time buyers and the impact that thin credit files have on their default risk. We find that, holding other factors constant, first-time borrowers with thin credit files have around a five percentage point higher likelihood of serious delinquency. This indicates that if underwriters are aware of thin credit files, then they are unsuccessful in fully compensating for this when making their credit decisions. Our estimates suggest that to offset the worse performance by thin credit file borrowers, their credit scores on average would have to be higher by approximately 40 points. Credit losses associated with default by first-time buyers are also typically more serious in that these borrowers start out with higher loan to value (LTV) ratios than repeat buyers. Consequently, on average, loss given default will be more severe for first-time buyers.

The paper is organized as follows. In the next section, we discuss the importance of the reliability as well as the unbiasedness of credit scores as predictors of future loan performance. We describe the panel data on credit files that we use in our analysis. Using this data, we summarize the early life-cycle of the distribution of credit scores for a vintage of credit files. We introduce a definition of a thin credit file and show its prevalence among FHA first-time buyers. In section three, we present our estimation results for the performance by FHA first-time buyers. We then conduct a series of robustness checks to explore alternative definitions of thin files as well as additional dimensions of credit files. We discuss the implications of thin credit files for the pricing of mortgage insurance guarantee fees. The final section provides concluding remarks.

First-Time Buyers and Thin Credit Files

Credit scores have become an integral component of mortgage underwriting. Credit scores are data driven and are designed to provide an unbiased indicator of the likelihood that a borrower will default on a credit obligation over the next two years. While unbiasedness is important, it is not the only feature that is relevant for underwriting. The "reliability" of the credit score is also important – where reliability relates to the signal to noise ratio.

This is illustrated in Figure 2. Let θ denote the unknown credit worthiness of a borrower and $\hat{\theta}$ the estimated credit score for that borrower. The credit score is an estimate for the borrower's unknown credit worthiness. Associated with each credit score is a sampling distribution. We consider two borrowers with the same unbiased credit score equal to $\hat{\theta}$. Borrower A has a thin

credit file, whereas borrower *B* has a thick credit file. Consequently, the sampling distribution for the borrower *A*'s credit score has a larger variance. Let $\tilde{\theta}$ be the threshold value for θ such that lower values of θ are associated with default. The default probability for a borrower is given by the area of the sampling distribution associated with the borrower's credit score that is below $\tilde{\theta}$. Even though borrower *A* and *B* have the same credit score, the default probability is higher for borrower *A* due to the lower reliability of the thin file credit score.[2] To equalize the default probability across these two borrowers, we could increase the credit score for borrower *A* shifting the sampling distribution to the right sufficiently so that the areas of both sampling distribution below $\tilde{\theta}$ are the same. Alternatively, we could adjust some other aspect of the loan file such as increasing the downpayment percentage or lowering the borrower's debt-to-income (DTI).

The implication of thin credit files for default risk can be seen more clearly in Figure 3 where we redisplay the information from Figure 2 now showing the credit score on the horizontal axis and the probability of default on the vertical axis. We set the origin at $\tilde{\theta}$. At the origin, both thin (A) and thick (B) credit files have the same default risk of 50 percent. As we increase the credit score, the default risk falls faster for the thick file than for the thin file. Thus, at every credit score above $\tilde{\theta}$ the default risk is higher for the thin credit file borrower. In addition, for a given default risk the associated credit score is higher for the thin credit file borrower. An implication is that for modeling default the credit score is not a sufficient summary measure for the borrower's credit worthiness. Holding the credit score constant, borrowers with thin credit files are predicted to have a higher default risk.

There is another limitation of thin credit files as a predictor of mortgage default risk in addition to their higher sampling variance. Credit scores are estimated from a panel of credit files at two points in time – two years apart. Data from the credit file at the first point in time is used to predict the existence of a "negative credit event" at the second point in time. Many borrowers with thin files will continue to have thin files at both of these points in time. Consequently, they are unlikely to have a negative credit event at the second point in time. However, from the standpoint of underwriting a thin credit file for a purchase mortgage for a first-time buyer, we want to predict the likelihood that the borrower will have a serious delinquency over the subsequent two years

[2] This result would hold in the more general case where the default risk is a smoothly decreasing function of θ rather than a step function.

conditional on having a mortgage. This is quite different from the unconditional likelihood of a negative credit event.

The potential contrast in the information content between a credit score based on a thick or thin credit file raises the issue of the early life-cycle of credit scores. A credit file is started for an individual when the first credit "event" for that person takes place. This initial credit event could be an individual applying for a gas, department store or credit card. Alternatively, the initial credit event could be a collection notification, say on unpaid medical or utility bill. Depending on the scoring algorithm a credit score can be assigned immediately when the credit file is opened with an account. What credit score is typically assigned to a new credit file? The answer depends on what type of event led to the creation of the file.

Define a vintage of new credit files by the date when the credit files are opened and their initial credit scores are assigned. Figure 4 shows the distribution of end of quarter Equifax credit scores for the 2004 vintage of new credit files.[3] In constructing this figure, we exclude those files that were opened due to a collection action or authorized user credit account.[4] Credit files appear to be initialized with a credit score of around 650 – this is around the 30th percentile of the overall credit score distribution. These represent files that are opened and have no positive or negative subsequent credit events over their initial quarter. Files that have a positive credit event during their initial quarter can end the quarter with a score higher than 650. Similarly, files that have a negative credit event during their initial quarter can end the quarter with a score lower than 650. As time passes this vintage's credit distribution begins to spread out reflecting, in part, the accumulation of credit experiences by individuals in the vintage. An important thing to note is that credit models have a positive drift in the credit score (for Equifax it is around 10 points per year).

All newly opened credit files are by definition "thin" files in that the individual has had no meaningful experience yet managing credit. As time goes by, the individual may take on more credit in which case there will be more information with which to update the credit score. That is, the individual transitions from having a thin to a thick credit file. However, an individual may, instead, not take on any additional credit lines nor accumulate any sizeable credit balance. In this case, at

[3] This figure is based on the Federal Reserve Bank of New York Consumer Credit Panel data that we will describe in the next section.

[4] An example of an authorized user credit account would be a parent adding an adult child to one of the parent's credit cards. If a credit file did not exist for the adult child, this would result in a file being opened.

time progresses the individual's credit file remains thin. To operationalize this concept of a thin credit file, we will define a thin credit file to exist when the individual has to date not had a credit balance that exceeded $5,000. We exclude student loan balances from this calculation. As an example, an individual would transition from a thin to a thick credit file by taking out a car loan. Note that we expanded the notion of a thin credit file to include those with somewhat long histories but with relatively small balance activities.

There is a close correspondence between first-time buyers and thin credit files. By definition, a repeat-home buyer has a thick credit file since the file contains the performance by the individual on the prior mortgage(s). First-time buyers may or may not have a thin credit file. Using our definition of a thin file given above, between 20-30 percent of all FHA first-time buyers in our data also have thin credit files. Figure 5 shows by year the distribution of FHA first-time buyers for different measures of the thinness or thickness of their credit files. The percentage of thin files exceeded 30 percent in 2004 and was close to 50 percent if we also include recent thin files -- those borrowers who crossed our debt balance threshold less two years prior. The fraction of FHA thin and recent thin files has declined over the years reaching 30 percent in 2013. This decline is mainly due to a steadily increasing fraction of FHA first-time borrowers who crossed our credit balance threshold more than four years prior to taking out their mortgage.

A second dimension of a credit file is the duration of the file -- where the duration is measured as the elapsed time since the file was opened with an account (excluding authorized user credit account). Credit files can be "short" or "long." In terms of the subsequent performance of the borrower on the first-time mortgage, a question is whether the duration or the thickness of the credit file is more important? For our sample of FHA first-time buyers in 2013, the median duration of the borrower's credit file is 9 years. The inner-quartile range is from 5.6 to 13.4 years. We define a short credit file to have a duration that is less than 2 years.

On average, thin credit files also tend to be short, but there are differences between these two characteristics. Table 1 shows a cross-tab for first-time FHA homebuyers between short/long and thin/thick credit files. The data indicate that almost 85 percent of the cases fall on the main diagonal. However, 22 percent of short files are thick and 14 percent of long files are thin. Looking just at thin files, 55 percent are short files and 45 percent are long files. From the standpoint of the

variance of the sampling distribution shown in Figure 2, we expect that there is a closer correspondence of the sampling variance to the thin/thick vs the short/long dimensions.

First-time homebuyers using an FHA mortgage tend to be younger than those using non-FHA mortgages. For our sample of FHA first-time buyers, the median age at the transition to homeownership is 32 years. This is in contrast for all first-time buyers where the median age is slightly older at 35 years. The median age of FHA first-time buyers has been relatively constant over our sample period, whereas the median age of overall first-time buyers has declined from 37 in 2000 to 33 in 2013.

The Impact of Thin Files on Mortgage Performance

We use the Federal Reserve Bank of New York's Consumer Credit Panel (FRBNY CCP) data. This is a panel dataset of credit files on a five percent random sample of U.S. households. The data is quarterly starting in 1999. For each household, we observe the credit file information for each household member. We can follow these households over time and continue to track individuals even if they leave the household. For each individual, we see information on all of their credit lines. For mortgages, there are narrative codes that identify the mortgage type as FHA, VA and GSE (although the coverage of GSE is incomplete). At a quarterly frequency we see the balances and payment status on the credit lines. In contrast to mortgage servicing data, we do not see the origination loan-to-value (LTV) ratio or the borrower's DTI ratio. For more details on the FRBNY CCP see Lee and van der Klaauw (2010).

We identify first-time buyers in the FRBNY CCP as individuals who are obtaining a purchase mortgage but have had no prior mortgage debt on their credit files.[5] In the case where two individuals co-sign for the mortgage, we apply this criterion to both individuals. In measuring the borrower(s) performance on the mortgage, we focus on default which we define as the borrower going 120-days delinquent on the mortgage. An important variable in mortgage default models is the current LTV. This is typically estimated by updating the origination LTV to account for any debt amortization and house price appreciation to date. One challenge is that the FRBNY CCP does not provide the origination LTV.

[5] We observe that it is a 1st mortgage. We infer that it is a purchase mortgage.

This data limitation is mitigated if we focus our empirical investigation on FHA first-time buyers. The reason is that the vast majority of first-time FHA purchase mortgages are originated with the maximum allowed LTV of 96.5 percent. In addition, nearly all of these FHA mortgages are 30-year fixed-rate mortgages. Consequently, there is little debt amortization over the initial years of these mortgages.

One challenge in modeling mortgage default for FHA mortgages is that the FHA has a streamline refinance program that provides FHA borrowers with a low cost option to refinance even if the borrower has negative equity (where the current house value is less than the mortgage balance). With the sharp reduction in mortgage rates following the housing crisis and recession, the FHA experienced a significant pick-up in streamline refinances.[6] In addition, an FHA borrower can streamline refinance multiple times. As noted by Caplin et al. (2012), the standard default model follows a mortgage and not a borrower. This is problematic due to the streamline refinances. To see this, consider the case of a first-time buyer who streamline refinances the FHA purchase mortgage and subsequently defaults on the refinanced mortgage. If we follow the traditional approach of analyzing the performance of first-time FHA purchase mortgages, we would record this outcome as a prepayment. The default would not enter the analysis since it would be assigned to the refinance mortgage which, by definition, is not a first-time mortgage. This would understate the default risk for our FHA first-time buyers.

Fortunately, the FRBNY CCP data allow us to address this problem. Since we can follow our FHA first-time buyers through time, we can observe any subsequent internal refinances.[7] Following Caplin et al. (2012), we use a data structure based on the borrower instead of the mortgage. That is, we will merge the internal refinances back to the original first-time purchase mortgage and track the borrower until the mortgage either defaults, prepays out of the FHA system or the outcome is censored due to the end of the data sample. Figure 6 shows the time pattern of the FHA internal refinances (the differences between the two lines). To illustrate the impact that internal refinances can have on the analysis, we will report our results both using the borrower and mortgage based data framework. We include first-time FHA borrowers who originated their purchase mortgages between 2004 and 2008, and follow the performance on these borrowers (mortgages) until the end of 2014.

---

[6] U.S. Department of Housing and Urban Development (2014), page 11, exhibit I-1.
[7] Internal refinances include both streamline refinances as well as standard FHA to FHA refinances.

We estimate static linear probability models for whether a first-time FHA borrower (or purchase mortgage) is ever observed as 120-days delinquent. Each FHA first-time borrower (or purchase mortgage) contributes a single observation. We control for the borrower's origination credit score (as the score reported at the same time that the mortgage loan appears on the credit file) and whether or not the borrower has a thin credit file. The credit score is entered as a series of indicator variables for different ranges of credit scores with 760 or higher as the left-out category (<560, 560-599, 600-639, 640-679, 680-719, 720-759). In addition, we control for the borrower's age with 30-39 as the left-out category (20-29, 40-49, 50+). To control for the maximum time that the mortgage is at risk of a default, we include indicators for the year that the first-time purchase mortgage was originated. We also control for state fixed effects.

Summary statistics for our sample of first-time FHA mortgages/borrowers is provided in Table 2. As discussed earlier, the fraction of first-time FHA purchase mortgages associated with either thin or short credit files has declined between 2001 and 2013. In addition, the data indicate that between 2007 and 2009 there was a shift away from borrowers with credit scores below 600. This change in the distribution of the FHA first-time buyer credit scores is also displayed along with the yearly origination volumes in Figure 7. The age composition of FHA first-time buyers has been roughly stable over the sample period.

We present our baseline estimates in Table 3. We report on three different specifications for how we control for thin/short credit files, and separately for mortgage-based and borrower-based samples. For each first-time FHA purchase mortgage (borrower) there are three possible outcomes: the mortgage defaults (reaches 120-days delinquent), the mortgage prepays (outside of the FHA in the borrower based case), or the outcome is censored due to the fact that the mortgage (borrower) is still active at the end of our sample period. There is a high default rate for FHA first-time purchase mortgages originated during our sample period. Switching from the mortgage to the borrower based data framework increases the measured default rate from 26.6 percent to 30.5 percent – an increase of 15 percent relative to the mortgage based average. Caplin et al (2014) report for a sample of all FHA mortgage originations between 2007 and 2009 a similar relative increase in measured default moving from a mortgage to a borrower framework.

We will first discuss the results for the borrower's credit score and age. The data indicate that default risk declines as a borrower's credit score increases. Caplin et al (2014) report for their

borrower based sample of FHA purchase mortgages (not restricted to first-time borrowers) that the default risk increases by a factor of 8 moving from a credit score of 680-719 to below 580. For our sample of FHA first-time purchase borrowers we find that the default risk increases by a factor of 3.4 moving from a credit score of 680-719 to below 560. So, while credit scores are statistically significant at predicting default for first-time FHA buyers, the sensitivity of default risk to the credit score for these borrowers appears to be lower than for all FHA borrowers. Looking at the age effects, younger (age 20-29) and older (age 50+) borrowers are less risky than borrowers age 30-39.

Turn now to the results for the thickness and length of the credit file. Again, if lenders are fully aware of these aspects of the credit file and compensate for them in their underwriting, then we would not expect that thin files would be associated with higher default risk. The data indicate that for both the mortgage and the borrower based frameworks that thin file FHA first-time borrowers are more likely to default by around 5 percentage points. Relative to the average default rate, this is an increase of 16 percent (borrower based) or 19 percent (mortgage based). Borrowers who transitioned from thin to thick files within the prior two years are not significantly more risky than borrowers who have had thick files for a longer period. Finally, controlling for whether the FHA first-time borrower has a thin file, there is no impact on the estimated default risk if the borrower also has a short versus a long file. So, for default risk the information content as measured by the thickness of the credit file is more important than the length of time since the file has been opened.

In Table 4 we restrict our attention to the borrower based data and focus on the first three years after an FHA first-time purchase mortgage is originated. One might expect that a thin file is more of a default risk early in the life of the mortgage. If a borrower has successfully paid on a mortgage for a few years, then this may be more relevant to the future performance on the mortgage than whether the borrower initially had a thin file or not. The data indicates that holding other factors constant thin files are associated with a 4 percentage point higher default risk over the first three years. This translates into a 30 percent increase in the relative three year cumulative default risk – nearly double the earlier relative impact.

Pricing Mortgage Guarantees

The FHA provides insurance on the full balance of the mortgages it guarantees. In return, borrowers pay an insurance premium that consists of two components: an up-front mortgage insurance premium (UFMIP) and an annual premium. Currently, the UFMIP is 175 basis points and the annual premium is 85 basis points applied against the origination mortgage amount. The FHA does not at present use risk-based pricing for it's mortgage insurance. Risk-based pricing was introduced in July 2008, but was subsequently cancelled by Congress in October 2008 with the passage of the Housing and Economic Recovery Act of 2008.[8] In contrast, the GSEs use risk-based pricing for their guarantee fees.[9]

The reliability of credit scores can impact the value from using risk-based pricing. This is illustrated in Figure 8. Here we make the more realistic assumption that default risk monotonically declines as we increase the creditworthiness of the borrower. This is shown by the dashed line, $\mathrm{E}P_D(\theta)$. If credit scores were a perfect indicator of θ, then the guarantee fee could be mapped directly to $\mathrm{E}P_D(\theta)$. However, credit scores are imperfect indicators of the borrower's θ. The degree of risk-based pricing will vary with the reliability of the credit score. For thin files, their lower reliability implies that the expected default risk varies less with the credit score than for thick credit files. This is illustrated by the flatter slope to $\mathrm{E}P_D(\hat{\theta}_A)$ as compared to $\mathrm{E}P_D(\hat{\theta}_B)$. This means that with thin credit files there is less scope to use risk-based pricing. In addition, less is lost by not using risk-based pricing at all.

A consequence of not using risk-based pricing of the guarantee fee is that the funding of the FHA's Mutual Mortgage Insurance Fund (MMIF) involves cross-subsidization between higher and lower credit worthy borrowers. This can be seen in Figure 8 where borrowers with a θ greater than $\hat{\theta}'$ subsidize borrowers of lower creditworthiness. These high θ borrowers would like to signal their creditworthiness but are prevented by the relatively large sampling distribution associated with their credit score due to their first-time buyer status. When these borrower's make a trade-up purchase at a later date, however, they will be in position to signal their higher θ enabling them to obtain cheaper financing outside of the FHA. That is, the FHA policy of no risk-based pricing works only while it is difficult for credit worthy borrowers to credibly signal their status.

[8] For insurance issued during this 3-month interval on a 30-year mortgage, the UFMIP/annual premium ranged from a low of 1.25/0.5 for LTVs less than or equal to 95 and credit scores from 680-850 to a high of 2.25/0.55 on LTV greater than 95 and credit scores of 500-599.

[9] See https://www.fanniemae.com/content/pricing/llpa-matrix.pdf for Fannie Mae's pricing schedule.

In addition to not charging differential UFMIP and annual fees, the FHA also cross-subsidizes poor credit borrowers though allowing the UFMIP to be financed into the balance of the mortgage. Charging an up-front as well as annual fee is a form of two-part pricing. The rationale for using a two-part pricing scheme for mortgage insurance is that the majority of defaults occur early in the life of the mortgage. If the guarantee fee consisted only of an annual fee, then defaulters would on average pay less for their mortgage insurance than borrowers who perform on the mortgage. To counteract this, an up-front fee is also charged. The role of the up-front charge is to make sure that even early defaulters end up paying a significant insurance fee.[10]

However, the FHA allows the UFMIP to be financed into the balance of the mortgage. While this helps to make the mortgage "affordable" by lowering the amount of cash that the borrower must bring to the closing, it counters the role of the UFMIP. To see this, consider the situation where a borrower takes out a 96.5 LTV FHA purchase mortgage and finances the 1.75 basis point UFMIP. This raises the borrower's effective LTV to 98.25. The FHA books the UFMIP as current revenue. Now assume that the borrower defaults three years later. Traditional models of default involve the borrower being in negative equity and then being hit with an adverse income shock. Since the borrower is in negative equity at the time of the default, the value of house does not cover the remaining mortgage balance including the financed UFMIP. Consequently, the FHA must return the UFMIP to the lender as part of its guarantee. This implies that the borrower in effect never paid the UFMIP. In contrast, borrowers who perform on their mortgage and fully pay off their balance end up paying for their UFMIP. Thus, the policy of enabling borrowers to finance the UFMIP involves another form of cross-subsidization between different credit quality borrowers.

Conclusions

We explore the role of thin credit files in the performance of first-time FHA purchase mortgages. Thin credit files are only a potential issue for underwriting mortgages for first-time buyers. First-time purchase mortgages, though, account for roughly sixty percent of FHA originations. Credit scores based on thin credit files are a noisier estimate of the borrower's underlying credit worthiness. Our data reject the hypothesis that thin credit files are recognized and compensated for in the underwriting process. The data indicate that thin file borrowers have a five

[10] For a detailed discussion, see Oi (1971).

percentage point highe default risk. This higher default risk is concentrated in the first three years of the mortgage.

This finding raises the question of why the lower reliability of thin files is not being offset in the underwriting process. This offset could occur through requiring a higher credit score, higher downpayment percentage or lower borrower DTI. However, using these offsets would imply that first-time purchase mortgages are either more difficult to obtain or more expensive. This could be perceived as going against the mission of the FHA to provide financing to first-time buyers. There is a tension between the FHA's requirement to break even in its mortgage guarantees and maintaining credit availability to a core constituency. Similar tensions existed for the GSE's between their mission goals and prudential supervision – see Frame et al. (2015).

References

Caplin, Andrew, Anna Cororaton, and Joseph Tracy. "Is the FHA Creating Sustainable Homeownership?" NBER Working Paper No. 18190, June, 2012. http://www.nber.org/papers/w18190.

Frame, W. Scott, Andreas Fuster, Joseph Tracy, and James Vickery. "The Rescue of Fannie Mae and Freddie Mac." *Journal of Economic Perspectives* 39 (Spring 2015): 25-52.

Lee, Donghoon, and Wilbert van der Klaauw. "An Introduction to the Consumer Credit Panel." Staff Report No. 479. Federal Reserve Bank of New York, November, 2010. http://www.newyorkfed.org/research/staff_reports/sr479.html.

Oi, Walter Y. "A Disneyland Dilemma: Two-Part Tariffs for a Mickey-Mouse Monopoly." *Quarterly Journal of Economics* 85 (February 1971): 77-96.

U.S. Department of Housing and Urban Development. *Annual Report to Congress Regarding the Financial Status of the FHA Mutual Mortgage Insurance Fund Fiscal Year 2014*, November 14, 2014.

Table 1. Credit File Duration vs. Thickness: First-Time Buyers

	Thin	Thick	
Short	14.0	3.9	17.9
Long	11.6	70.5	82.1
	25.6	74.4	

Notes: A credit file is "thin" if the borrower has not had a credit balance that exceeds $5,000 excluding a student loan. A credit file is "short" if duration is less than 2 years.

Table 2. Summary Statistics

Variable	2001	2003	2005	2007	2009	2011	2013
Thin file	44.7	33.4	29.8	23.8	19.7	22.3	20.2
Thick < 2 years	23.9	17.1	10.5	12.9	12.5	9.2	10.3
Short	43.5	30.5	24.3	16.1	9.8	8.6	7.5
Credit score:							
<560	12.9	12.6	13.6	18.0	3.7	2.1	2.8
560-599	12.9	10.3	16.7	11.7	4.9	4.6	5.5
600-639	18.0	18.5	16.9	19.7	13.9	15.5	12.7
640-679	22.9	23.8	22.1	19.0	18.5	19.7	21.9
680-719	18.7	19.2	16.9	15.6	23.0	23.9	27.0
720-759	10.5	11.3	9.5	10.0	19.1	19.2	20.6
760+	4.1	4.3	4.3	6.1	16.8	15.0	9.6
Borrower age:							
20-29	41.0	44.9	48.0	44.9	46.6	43.7	44.6
30-39	35.1	31.3	27.5	26.5	31.7	32.1	33.0
40-49	14.0	14.5	13.2	16.9	11.2	13.7	12.2
50+	9.9	9.3	11.3	11.8	10.5	10.5	10.3

Notes: FHA first-time purchase mortgages from FRBNY CCP.

Table 3. Default risk

	Mortgage based			Borrower based		
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Thin file	0.052	0.056	0.060	0.049	0.053	0.069
	(0.019)	(0.019)	(0.024)	(0.020)	(0.020)	(0.025)
Thick & < 2 years		0.021	0.023		0.019	0.029
		(0.025)	(0.027)		(0.026)	(0.028)
Short			−0.006			−0.031
			(0.027)			(0.028)
Credit score:						
<560	0.458	0.456	0.456	0.468	0.466	0.467
	(0.040)	(0.040)	(0.040)	(0.041)	(0.041)	(0.041)
560-599	0.307	0.307	0.306	0.326	0.325	0.325
	(0.039)	(0.039)	(0.039)	(0.041)	(0.041)	(0.041)
600-639	0.207	0.205	0.206	0.219	0.218	0.220
	(0.037)	(0.037)	(0.037)	(0.038)	(0.038)	(0.038)
640-679	0.143	0.142	0.142	0.154	0.153	0.155
	(0.036)	(0.036)	(0.037)	(0.038)	(0.038)	(0.038)
680-719	0.048	0.046	0.047	0.054	0.052	0.054
	(0.037)	(0.037)	(0.037)	(0.039)	(0.039)	(0.039)
720-759	0.017	0.016	0.016	0.020	0.019	0.019
	((0.039)	(0.039)	(0.039)	(0.041)	(0.041)	(0.041)
Borrower age:						
20-29	−0.050	−0.053	−0.053	−0.051	−0.054	−0.055
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.021)
40-49	−0.018	−0.018	−0.018	−0.006	−0.006	−0.006
	(0.026)	(0.026)	(0.026)	(0.027)	(0.027)	(0.027)
50+	−0.048	−0.048	−0.048	−0.064	−0.064	−0.063
	(0.028)	(0.028)	(0.028)	(0.029)	(0.029)	(0.029)
Average default rate	0.266			0.305		

Notes: FHA first-time purchase mortgages from the FRBNY CCP data. Regression coefficients with standard errors given in parentheses. Number of observations is 2,657. State and vintage year fixed effects are included. The left-out category contains borrowers with a thick file, credit score of 760 or higher and between 30 and 39 years of age.

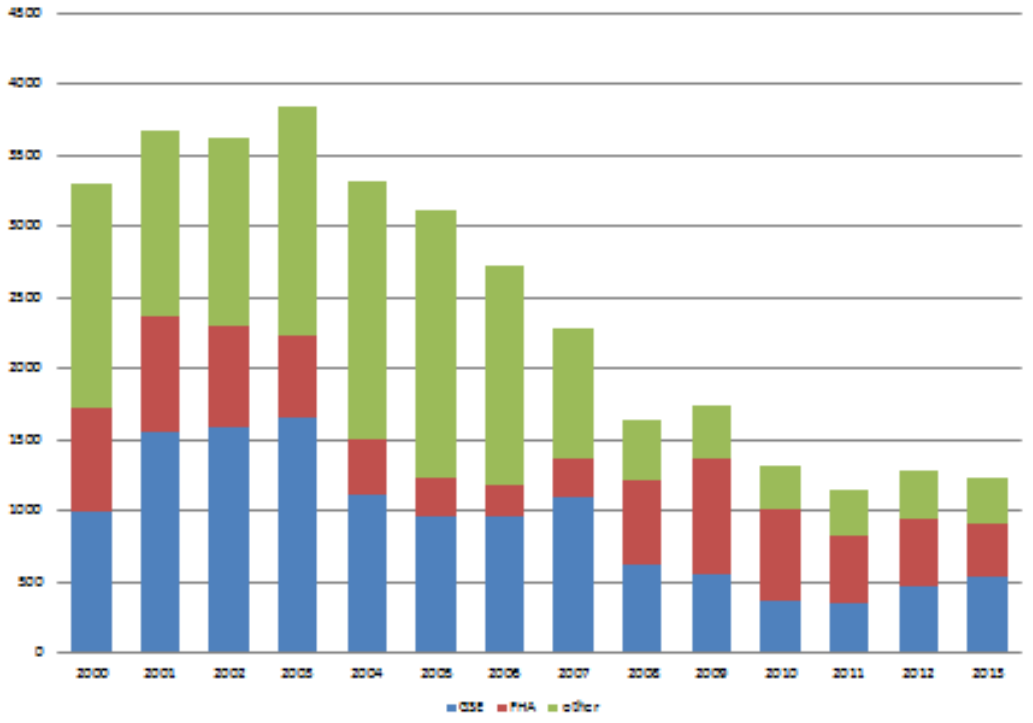
Table 4. Default risk – restrict to 3 years after origination

	Borrower based		
Variable	(1)	(2)	(3)
Thin file	0.040	0.038	0.037
	(0.019)	(0.015)	(0.019)
Thick & < 2 years		−0.007	−0.008
		(0.020)	(0.021)
Short			0.002
			(0.021)
Average default rate	0.132		

Notes: FHA first-time purchase mortgages from the FRBNY CCP data. Regression coefficients with standard errors given in parentheses. Number of observations is 2,657. State and vintage year fixed effects are included. The left-out category contains borrowers with a thick file, credit score of 760 or higher and between 30 and 39 years of age.

Figure 1. Volume and Composition of First-Time Purchase Mortgages – by year

(Thousands of mortgages)



Source: Federal Reserve Bank of New York Consumer Credit Panel / Equifax

Figure 2. Credit Score and Reliability

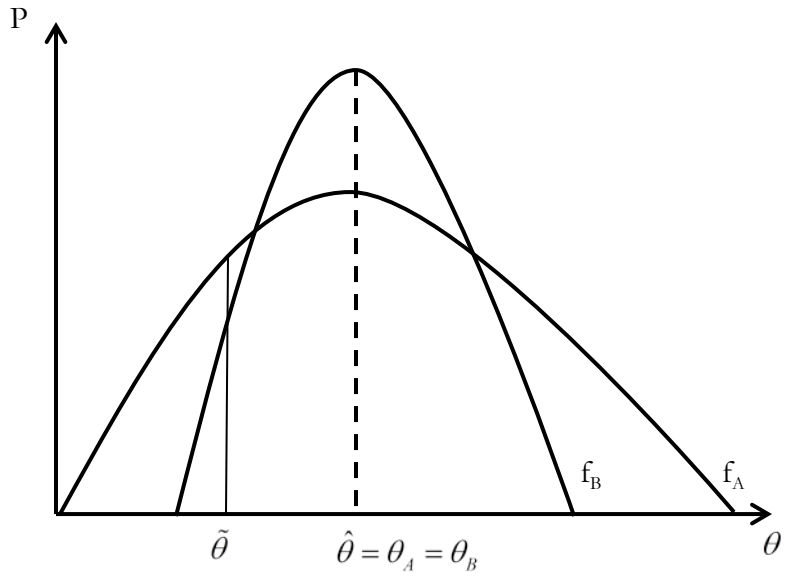


Figure 3. Credit Score and Default Risk

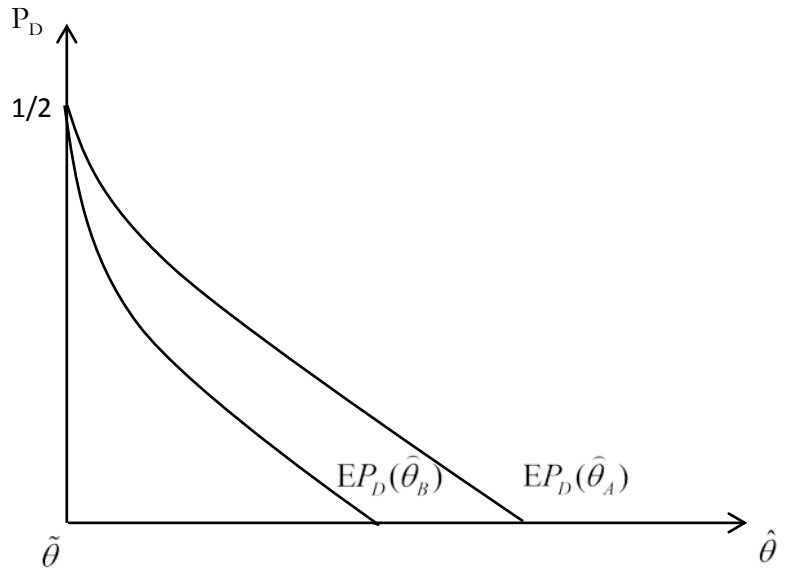
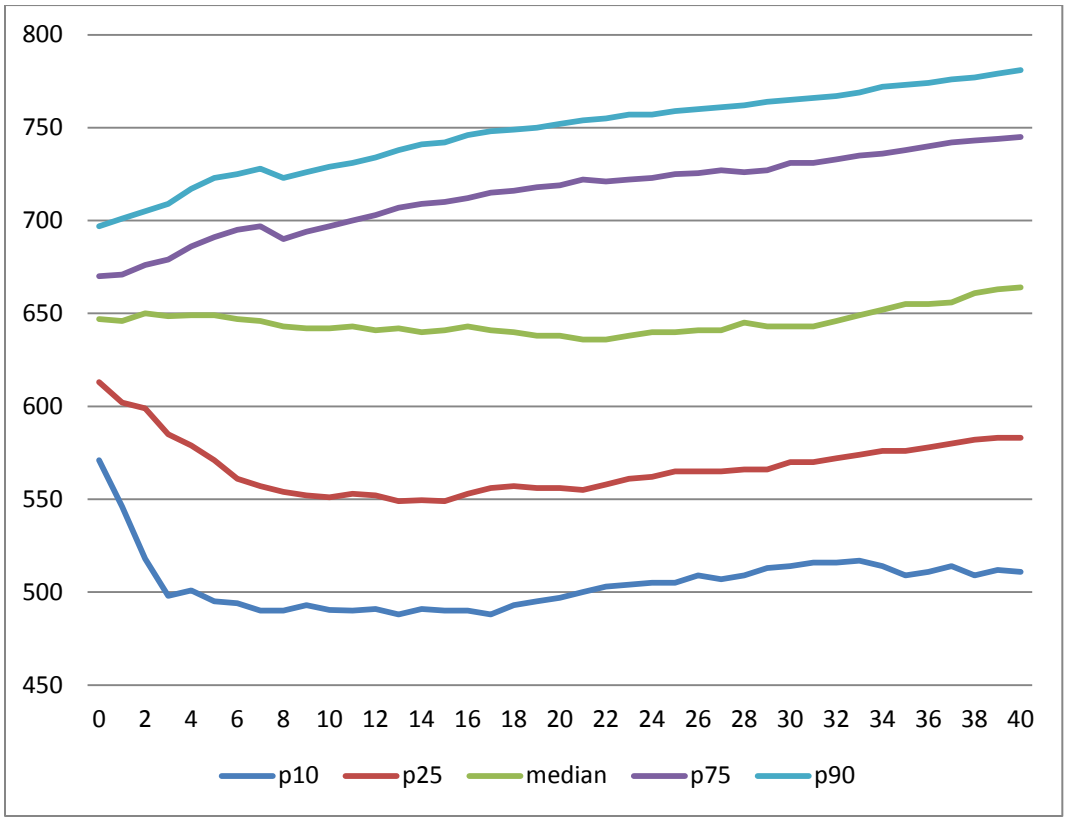
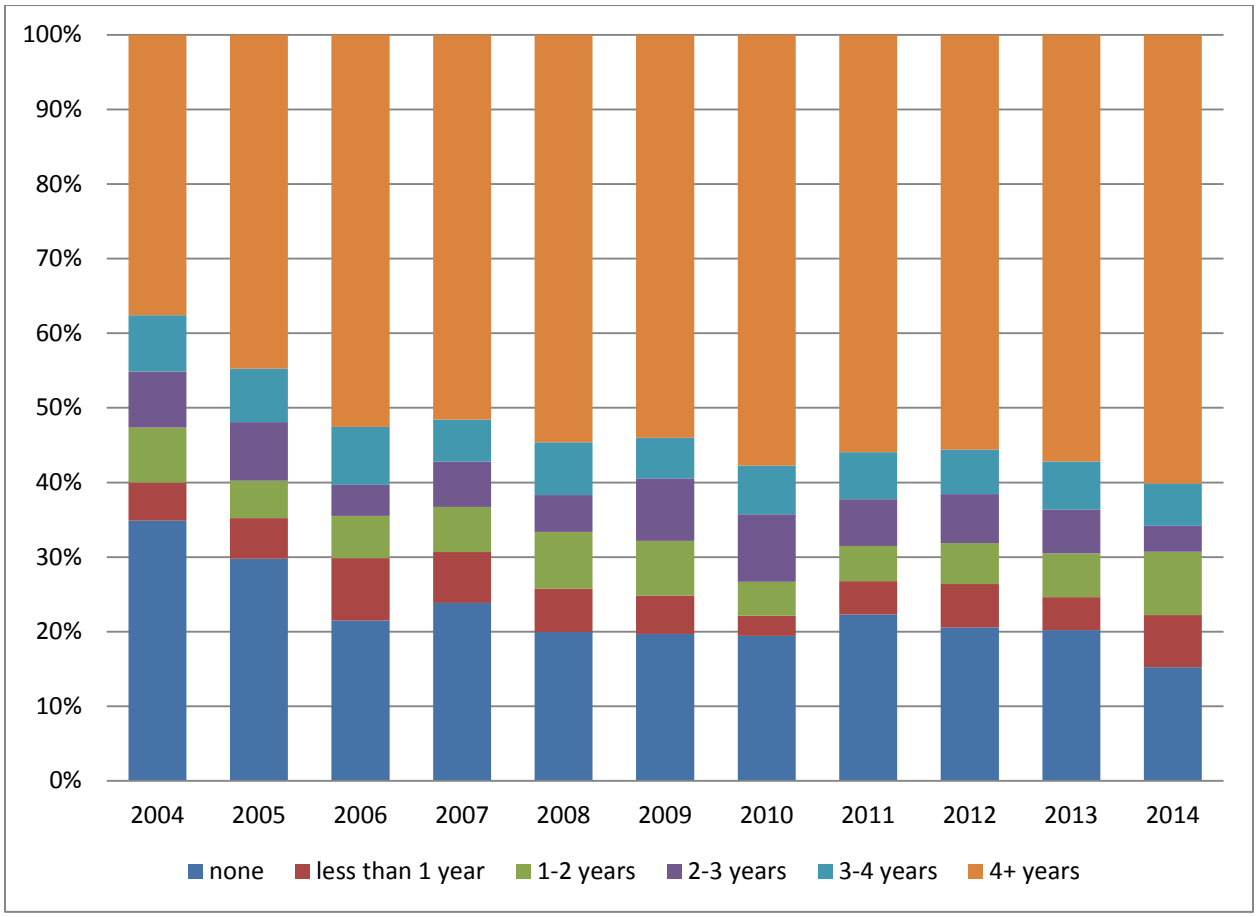


Figure 4. Life-Cycle of Credit Scores for Vintage of Borrowers



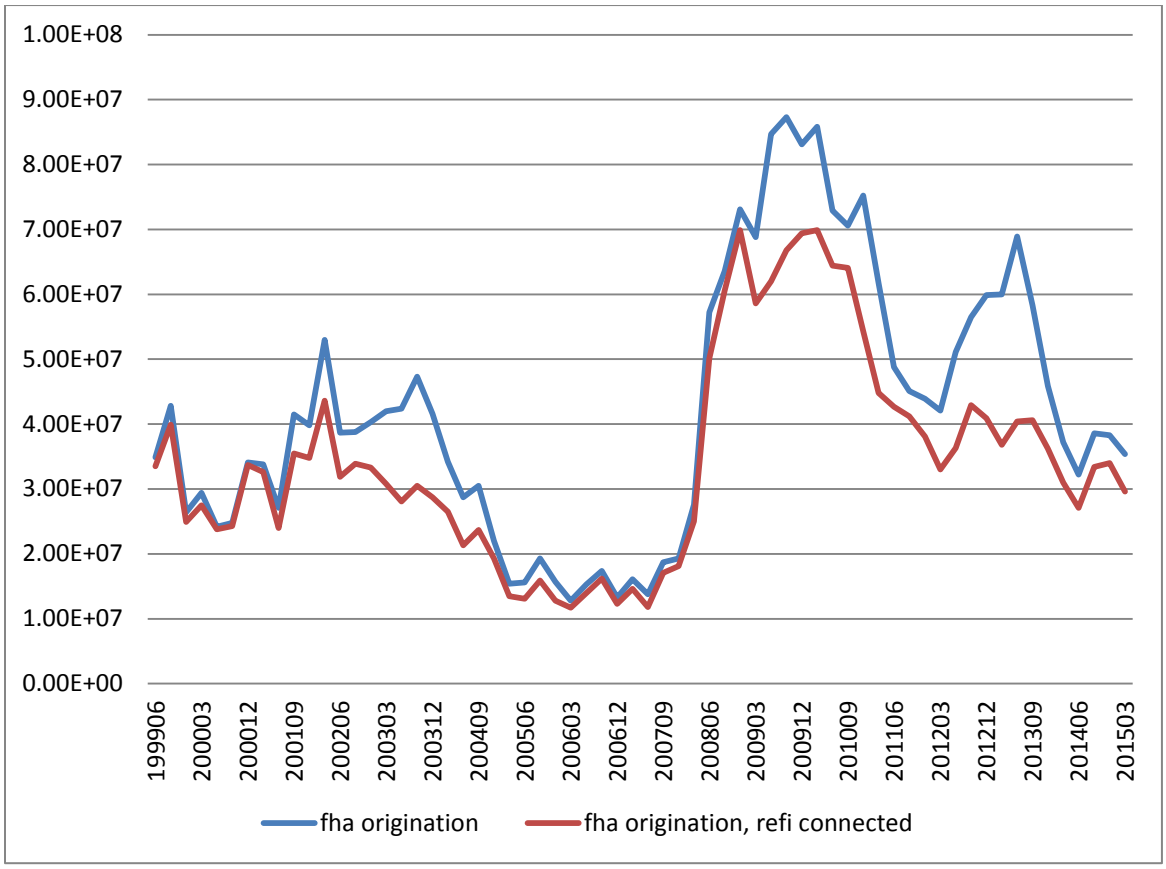
Source: Federal Reserve Bank of New York Consumer Credit Panel / Equifax

Figure 5. Distribution of FHA First-Time Buyers – by thickness of credit file



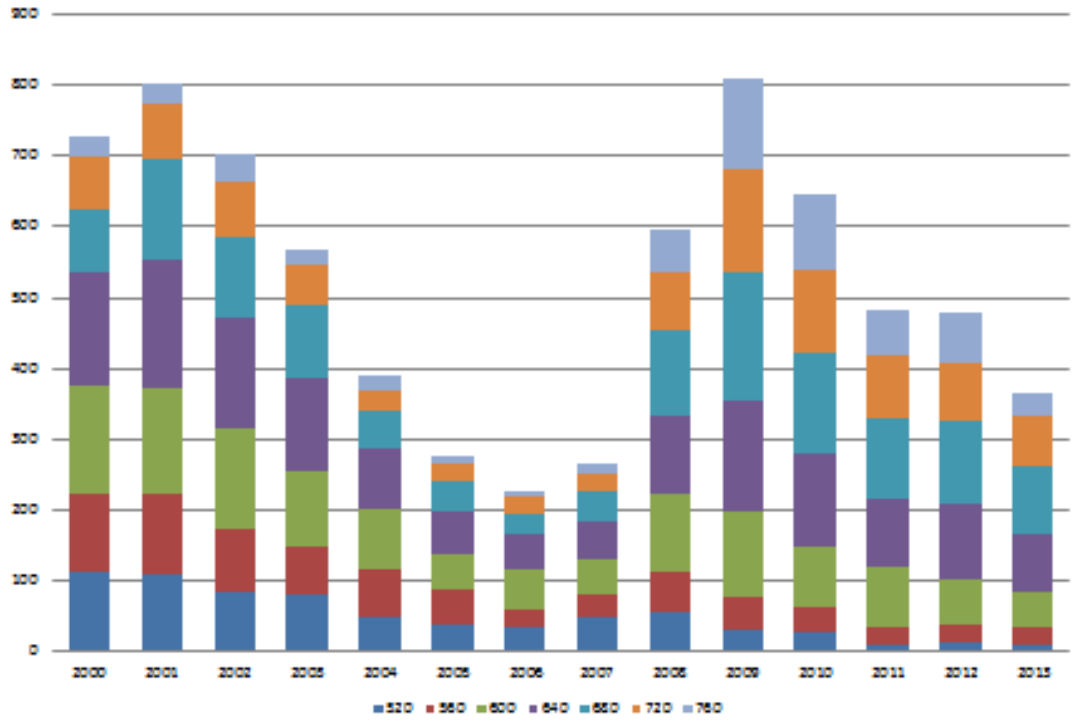
Source: Federal Reserve Bank of New York Consumer Credit Panel / Equifax

Figure 6. FHA Internal Refinancing



Source: Federal Reserve Bank of New York Consumer Credit Panel / Equifax

Figure 7. First-Time FHA Purchase Originations – by credit score and year



Source: Federal Reserve Bank of New York Consumer Credit Panel / Equifax

Figure 8. Thin vs Thick Files and Risk-Based Pricing