No Job, No Money, No Refi:
Frictions to Refinancing in a Recession

Online Appendices

Anthony A. DeFusco John Mondragon

A Data Appendix

A.1 Sample Selection

To conduct our empirical analysis we make use of a loan-level data set containing detailed information on individual mortgages active during the 6 months prior to the SLR policy change announcement date and 6 months after its implementation. The raw data was purchased from CoreLogic and is provided to them by a consortium of participating mortgage servicers.

Because the data are very large, we select our primary analysis sample from an initial 20 percent random sample of relatively "standard" mortgages. This initial sample includes 20 percent of all first-lien, fixed-rate, 30-year, single-family, owner-occupied mortgages with non-missing interest rates, LTVs, and FICO scores at origination that were active at some point between March, 2009 and July 2010. This initial sample contains 1,845,744 unique loans and 25,876,460 loan-month observations as reported in the first row of Table A.1.

Starting from this sample, we then impose a series of restrictions to make sure that we are able to accurately determine both whether a loan was paid off in a particular month and the reason for that payoff. First, we drop any loans for which the payoff status cannot be known with certainty during our entire sample period. This includes loans that were transferred to servicers not in our data set as well as loans with missing or unknown delinquency information.[1] We then drop any loan-month observations observed after the recorded payoff date. In the vast majority of cases, these observations report a zero remaining balance. Finally, we drop a small number of loans where the payoff information is inconsistently recorded across differing variables in the data.[2] The number of loans and loan-month observations dropped by each of these filters are reported in rows 2–4 of Table A.1. Together, they eliminate 86,138 loans and their associated 1,402,766 loan-month observations.

From the remaining sample of loans, we then impose a set of additional restrictions to ensure that all observations in our final analysis sample meet the payment history requirements imposed under the updated SLR program. Specifically, as described in Section II, to be eligible for the SLR program after the policy change a loan was required to be at least 6 months old and to have a satisfactory payment history. The payment history requirement

[1] Specifically, we drop any loan for which the `mba_delinquency_status` variable in the LLMA data ever takes the values: S (Servicing Sold/Released), T (Loan Status No Longer Provided/Available), X (Unknown/missing LPI Date), or Z (Incompatible).

[2] The data includes both a static variable indicating the date of the payoff (for loans that do pay off) as well as a dynamically updated variable that indicates whether a loan is paid off in a particular month. We drop any loans for which these two methods for determining the payoff status ever disagree.

depended on the age of the loan. For loans older than 12 months, all payments within the last 3 months must have been on time and no more than one payment in the last year may have been 30 days late. For loans between 6 and 12 months old, all payments since origination must have been on time. We drop any loan-month observations that do not meet these requirements. The fifth row of Table A.1 shows that this filter eliminates 237,550 loans, which is roughly 13 percent of the initial sample and is consistent with aggregate mortgage delinquency rates during our sample period.[3]

Finally, there are several important variables that we require to be non-missing for all loans in our final analysis sample. These include an estimate of current home equity, county-level unemployment rate changes, and an estimate of the potential interest rate available to each borrower if she were to refinance in a given month. To calculate current equity in each month, we need an estimate of house price appreciation between origination and the current month. Thus, we drop any loans in our data for which the county-level Zillow house price indices are unavailable either in the month of origination or during every month in our sample period. Similarly, because the remaining loan balance reported in the data set is an end of month number, we also drop any observations for which the lag of the remaining balance is missing since such observations would have a missing value for estimated current equity as well. Together these restrictions eliminate 148,579 loans. An additional 62,162 loans are dropped because they cannot be matched to county-level unemployment estimates from the ACS. Finally, we drop 160,412 loan-months for which we cannot calculate a potential interest rate due to a lack of observed refinances on similar loans in the candidate month. This leaves us with a final analysis sample containing 1,309,393 loans and 15,645,645 loan-month observations. Descriptive statistics for this sample are reported in Table II and discussed in Section III of the paper.

B Additional Results and Robustness Checks

B.1 Quadratic Time Trends in the Difference-in-Differences Regressions

One potential concern with our main difference-in-differences analysis is that changes in economic conditions prior to the policy change may have differentially affected the refinancing rate of FHA borrowers relative to conventional borrowers since FHA borrowers had easier access to refinancing through the SLR program during that period. This would violate the parallel trends assumption and introduce bias in our estimate of the effect of the policy

[3] Table 1193 in the 2011 Statistical Abstract of the United States reports that 13.9 percent of all mortgages by dollar volume were delinquent or in foreclosure in 2009.

change. As we discuss in Section IV.A, our main approach to dealing with this concern is to control for a set of linear time trends for FHA borrowers that are allowed to vary freely before and after the policy change. Here we show that our main results are also robust to more flexible quadratic trends, which may do a better job of capturing differential trends in the FHA market during the period surrounding the policy change.

In Figure A.1 we begin by replicating the results presented in Figure II but allowing for quadratic trends in time. As the figure makes clear, including these trends has no effect on the qualitative conclusions we draw from Figure II. Panel A. shows that there is still a large and discontinuous drop in refinancing in the FHA segment of the market that occurs in precisely the month of the policy change. The size of the drop in FHA refinancing implied by the quadratic trends is smaller than in the linear case but is nonetheless still visually apparent and economically similar in magnitude. Panel B. plots the same trends in the conventional market and, as before, shows little evidence of a meaningful drop in refinancing in that segment. Relative to the linear trends, the quadratic trends imply a larger drop in conventional refinancing; however, in both cases the size of this drop pales in comparison to the drop in the FHA market and is of similar magnitude to many other month-over-month changes in the raw averages shown by the blue dots.

Consistent with this evidence, Table A.2 shows that our main difference-in-differences estimates are also robust to including more flexible time trends. This table replicates Table IV from the main text, with the only difference being that here we include quadratic rather than linear time trends for FHA borrowers that are allowed to vary freely before and after the policy change. In all specifications, we continue to find a statistically significant and economically large fall in FHA refinancing after the policy change. The size of this drop is roughly 50 basis points, which is smaller than the 70 basis point fall we find in Table IV but still constitutes a 40 percent fall relative to average FHA refinancing rate in the month just prior to the policy change. Taken together, we view this as strong evidence that our main results are not affected by the specific manner in we attempt to control for potential differential trends.

B.2 Alternate Specifications for Flexible Difference-in-Differences

Figure III in the main text plots coefficient estimates from a version of equation (4) that is analogous to our most conservative difference-in-differences estimates from column 4 of Table IV. For robustness, in Figure A.2 we plot similar coefficient estimates from less conservative specifications that mirror those from columns 1–3 of Table IV. For completeness, Table A.3 also reports these coefficient estimates numerically.

Each set of coefficients reported in the figure comes from a different version of equation (4)

that includes an increasingly conservative set of controls. In "Model 1," we control only for the CBSA of the property. "Model 2" adds a further set of fixed effects for the current loan age (one-year bins), interest rate (one-percentage point bins) and the borrower's estimated home-equity ($10,000 bins), as well as the full pairwise interaction between the borrower's LTV (10-point bins) and FICO score (50-point bins). In "Model 3" we further interact all but the CBSA fixed effects with the dummies for the month of observation. This allows each control variable to have a fully non-parametric effect on refinancing over time. Finally, in "Model 4" we repeat our preferred estimates from Figure III for reference. This specification further interacts all of the control variables with the FHA dummy and produces results that are nearly identical to those from "Model 3." While the pre-trends differ somewhat across specifications, these trends are largely eliminated when we allow for our control variables to differentially affect refinancing rates over time. Moreover, across all four specifications, we find that the fall in FHA refinancing after the policy change is persistent and of roughly the same magnitude.

B.3 Pooled Full-Sample Triple-Difference Regressions

The triple-difference regressions we present in Section V are estimated in restricted samples. For example, when studying the role of the income documentation requirement, we restrict the sample to include only loan-months for which the borrower is in positive equity. Similarly, when estimating the effect of the upfront costs, we include only properties in counties that experienced below median increases in unemployment. The reason we restrict our samples in this way is because the two "treatments" are highly correlated. Counties that experienced large increases in unemployment also experienced on average significant declines in house prices during this period. By restricting the sample to include borrowers who are only likely to face one of the two constraints, we are able to plausibly isolate their separate effects. To be even more conservative, in our most stringent specifications we also control directly for any remaining effect of negative equity (in the income documentation analysis) or county-level unemployment (in the upfront costs analysis) on FHA refinancing after the policy change.

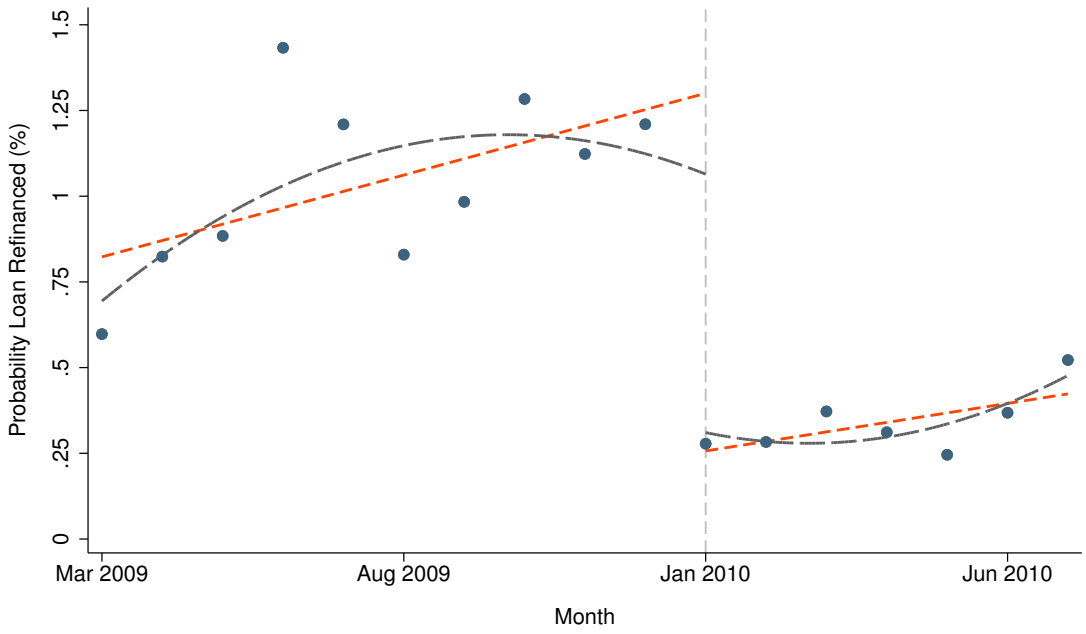
An alternative approach to isolating the separate effects of each constraint would be to estimate a pooled regression in the full sample that allows for both county-level unemployment changes and negative equity to separately effect FHA refinancing. This regression would be similar to our most conservative specifications (column 5 in Table V and Table VI) but be estimated in a non-restricted sample. In Table A.4 we present results from various versions of this type of regression. In columns 1–3 we estimate specifications that are directly analogous to those in Table V and Table VI. Broadly speaking, the estimates for the triple interaction coefficients in rows 3 and 4 are qualitatively similar to our main estimates.

For example, the results in column 3 indicate that the differential fall in FHA refinancing after the policy change was 46 basis points larger among FHA borrowers in negative equity and grows by roughly 3.1 basis points with each additional one percentage point increase in the county-level unemployment rate. These estimates are slightly smaller but economically similar to the 51 and 4.7 basis point estimates from the most directly analogous regressions in column 5 of Table V and Table VI.

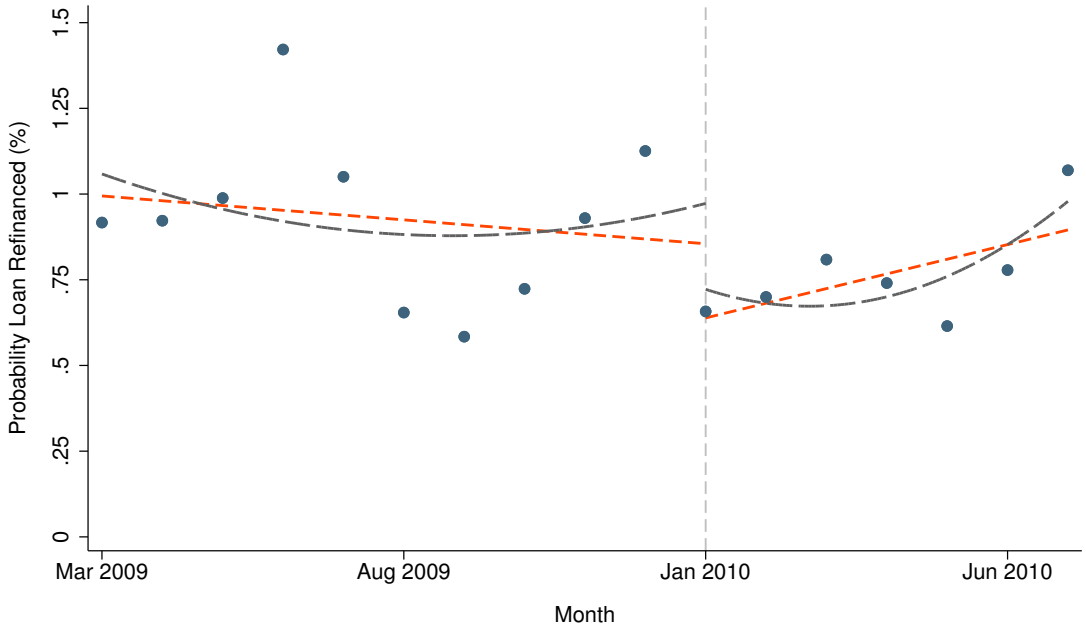
The largest difference between the estimates in full sample and those from the restricted samples in the main text are for the coefficient estimates on the triple interaction with unemployment changes in columns 2 and 3 of Table A.4. These effects are roughly half the size of their counterparts from Table V and are no longer statistically significant at conventional levels. However, this attenuation is exactly what would be expected given the fact that county-level unemployment is both highly correlated with home equity in the full sample and a very weak proxy for individual employment status. As a result, much of the "signal" from the county-level unemployment change variable is likely being picked up by the borrower-level measure of negative equity, leading to an attenuation of the coefficient on the triple interaction with unemployment. Nonetheless, at 2.3 basis points, these coefficients are still economically quite large and indicate that employment documentation requirements were likely a significant barrier to FHA refinancing after the policy change.

In columns 4–6, we present a slightly different version of the specification. Instead of entering the county-level unemployment change linearly, in these specifications we simply include a dummy for whether the borrower lives in a county experiencing an above-median increase in unemployment. This coarser measure of unemployment helps to address some of the issues with measurement error and correlation between unemployment and negative equity mentioned above. The results in the bottom row are statistically significant at the one percent level in all three specifications and indicate that living in a high unemployment county is associated with a roughly 20 basis point larger fall in FHA refinancing subsequent to the policy change. As a way of comparing these estimates to those from columns 1–3 and our main estimates in Table V we can divide this coefficient by the difference between the average increase in unemployment in low and high unemployment counties, which is 3.025 percentage points. Doing this suggests that the FHA refinancing rate falls by roughly $0.22/3.025 = 0.07$ percentage points more for each one percentage point increase in the county-level unemployment rate. This estimate is slightly larger but economically similar in magnitude to our preferred estimate of 4.7 basis points from Table V. Together, we think these results provide strong evidence that our main estimates are robust to alternative methods for isolating the independent effect of each of the two policy changes and indicate that both played an important role in contributing to the overall fall in FHA refinancing

after the policy change.



A. FHA Loans



B. Conventional Loans

FIGURE A.1
FHA and Conventional Refinancing Trends, Allowing for a Quadratic

NOTE.—This figure plots monthly unconditional refinancing rates between March 2009 and July 2010. Each dot represents the percent of outstanding mortgages of a given type that refinanced in the indicated month. Refinancing rates are calculated separately for FHA (Panel A) and Conventional loans (Panel B). The vertically dashed grey line in January 2010 marks the first month that the SLR policy changes went into effect. The dashed orange lines are the predicted values from a regression of the plotted refinancing rates on a linear time trend fit separately on either side of the cutoff date. The dark grey dashed lines are the predicted values from a similar regression that allows for a quadratic trend on either side of the cutoff.

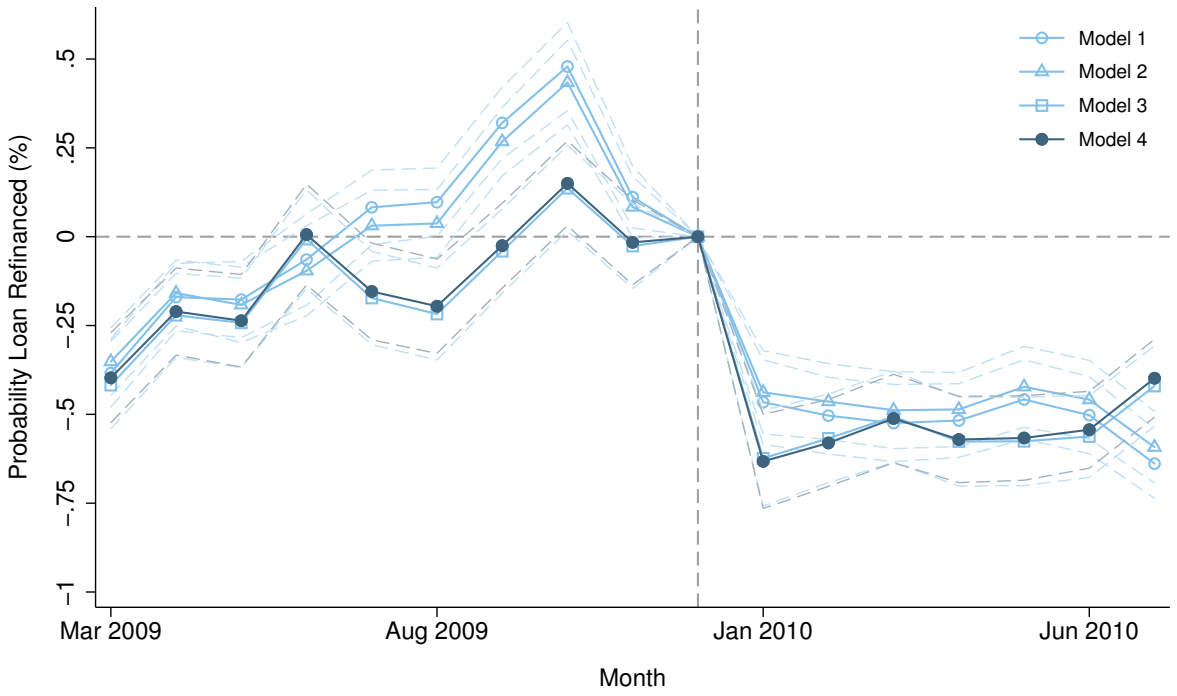


FIGURE A.2

Flexible Diff-in-Diff Estimates of the Effect of the FHA Policy Changes on Refinancing

NOTE.—This figure reports estimates of the effect of the change in FHA policies on FHA refinancing derived from a flexible difference-in-differences specification that allows the effect to vary freely by month of observation. Estimates were constructed by regressing an indicator for whether or not a loan refinances in a given month on a dummy variable denoting whether the loan was FHA insured and the interaction of that FHA dummy with a series of dummy variables indicating the month of observation. The coefficient for December 2009 is normalized to zero, so that all estimates can be interpreted as the change in the monthly probability of refinancing relative to the month prior to when the policy changes went into effect, which is marked by the vertically dashed grey line. In "Model 1" the regression also included fixed effects for the CBSA of the property. The regression in "Model 2" adds a further set of fixed effects for the current loan age (one-year bins), interest rate (one-percentage point bins) and the borrower's estimated home-equity ($10,000 bins), as well as the full pairwise interaction between the borrower's LTV (10-point bins) and FICO score (50-point bins) at origination. With the exception of the CBSA fixed effects, all of these controls are then separately interacted with the dummies for the month of observation in "Model 3" and the FHA dummy in "Model 4." The 95% confidence intervals are based on standard errors that are clustered at the CBSA level.

TABLE A.1
NUMBER OF OBSERVATIONS DROPPED DURING SAMPLE SELECTION

	Loans	Loan-Months
Original Count	1,845,744	25,876,460
Dropped Observations:		
Unknown Payoff Status	48,924	705,074
Payment Date after Payoff Date	34,077	663,771
Inconsistent Payoff Information	3,137	33,921
Payment History Ineligible for SLR	237,550	6,000,497
Missing Zillow Price Index	115,760	1,431,789
Missing Current Equity	32,819	506,016
Missing County Unemployment	62,162	729,335
Missing Potential Interest Rate	1,922	160,412
Final Count	1,309,393	15,645,645

NOTE.— This table reports the number of observations dropped during each stage of our sample selection procedure. Column 1 reports observation counts at the loan level, whereas column 2 is at the loan-month level.

TABLE A.2
Difference-in-Differences Estimates Allowing for Quadratic FHA Trends

	(1)	(2)	(3)	(4)
FHA	0.161***	0.551***	0.555***	1.796***
	(0.058)	(0.060)	(0.055)	(0.319)
FHA × Post	−0.490***	−0.485***	−0.484***	−0.490***
	(0.060)	(0.060)	(0.051)	(0.051)
Month FEs	X	X	X	X
CBSA FEs	X	X	X	X
Quadratic FHA Time Trends	X	X	X	X
Loan Age FEs		X	X	X
Interest Rate FEs		X	X	X
LTV × FICO FEs		X	X	X
Equity FEs		X	X	X
Controls × Post			X	X
Controls × FHA				X
Number of Observations	15,645,645	15,645,645	15,645,645	15,645,645

Note.—This table reports difference-in-differences estimates of the effect of the change in FHA policies on the monthly probability that an FHA loan refinances. Each column reports estimates from a separate regression where the dependent variable is an indicator for whether or not a loan refinances in the month of observation. The outcome is multiplied by 100, so that all coefficients can be interpreted as percentage point changes. Coefficients are reported for the FHA "treatment" dummy as well as its interaction with an indicator for whether the month of observation was after the implementation of the policy changes (Post), which occurred in January 2010. All specifications include fixed effects for the month of observation and the CBSA of the property as well as quadratic time trends for FHA borrowers that are allowed to vary freely before and after the policy change. Column 2 adds fixed effects for the current loan age (one-year bins), interest rate (one-percentage point bins) and the borrower's estimated home-equity ($10,000 bins), as well as the full pairwise interaction between the borrower's LTV (10-point bins) and FICO score (50-point bins) at origination. Column 3 further interacts all of the additional fixed effects contained in column 2 with the Post dummy. Column 4 adds an additional set of interactions between these fixed effects and the FHA dummy. Standard errors are reported in parentheses and are clustered at the CBSA level. Significance levels 10%, 5%, and 1% are denoted by *, **, and ***, respectively.

TABLE A.3
The Effect of the Policy Changes on FHA Refinancing: Flexible Difference-in-Differences

	(1)	(2)	(3)	(4)
Mar. 2009	−0.384***	−0.352***	−0.418***	−0.397***
	(0.049)	(0.048)	(0.063)	(0.065)
Apr. 2009	−0.170***	−0.159***	−0.221***	−0.211***
	(0.048)	(0.047)	(0.060)	(0.063)
May. 2009	−0.177***	−0.192***	−0.242***	−0.236***
	(0.055)	(0.054)	(0.064)	(0.066)
Jun. 2009	−0.064	−0.096	−0.007	0.006
	(0.066)	(0.065)	(0.071)	(0.072)
Jul. 2009	0.083	0.031	−0.173**	−0.154**
	(0.053)	(0.051)	(0.067)	(0.069)
Aug. 2009	0.097**	0.037	−0.218***	−0.196***
	(0.049)	(0.049)	(0.066)	(0.068)
Sep. 2009	0.320***	0.268***	−0.041	−0.025
	(0.051)	(0.049)	(0.060)	(0.061)
Oct. 2009	0.479***	0.434***	0.137**	0.150**
	(0.063)	(0.061)	(0.061)	(0.061)
Nov. 2009	0.111**	0.084*	−0.026	−0.016
	(0.044)	(0.044)	(0.061)	(0.061)
Jan. 2010	−0.466***	−0.438***	−0.624***	−0.632***
	(0.061)	(0.060)	(0.069)	(0.068)
Feb. 2010	−0.504***	−0.464***	−0.568***	−0.580***
	(0.055)	(0.054)	(0.064)	(0.062)
Mar. 2010	−0.525***	−0.488***	−0.506***	−0.512***
	(0.055)	(0.055)	(0.066)	(0.064)
Apr. 2010	−0.517***	−0.486***	−0.576***	−0.571***
	(0.053)	(0.053)	(0.064)	(0.062)
May. 2010	−0.458***	−0.422***	−0.576***	−0.566***
	(0.057)	(0.058)	(0.064)	(0.061)
Jun. 2010	−0.502***	−0.459***	−0.562***	−0.543***
	(0.055)	(0.056)	(0.059)	(0.055)
Jul. 2010	−0.639***	−0.592***	−0.420***	−0.398***
	(0.050)	(0.051)	(0.058)	(0.056)
Month FEs	X	X	X	X
CBSA FEs	X	X	X	X
Loan Age FEs		X	X	X
Interest Rate FEs		X	X	X
LTV × FICO FEs		X	X	X
Equity FEs		X	X	X
Controls × Month FEs			X	X
Controls × FHA				X
Number of Observations	15,645,645	15,645,645	15,645,645	15,645,645

Note.—This table reports flexible difference-in-differences estimates of the effect of the change in FHA policies on the monthly probability that an FHA loan refinances. Each column reports estimates from a separate regression where the dependent variable is an indicator for whether or not a loan refinances in the month of observation. The outcome is multiplied by 100, so that all coefficients can be interpreted as percentage point changes. Coefficients are reported for the interaction between the FHA "treatment" dummy and a dummy for the month of observation. The coefficient for December 2009 is normalized to zero, so that all estimates can be interpreted as the change in the monthly probability of refinancing relative to the month prior to when the policy changes went into effect. All specifications include fixed effects for the month of observation and the CBSA of the property. Column 2 adds fixed effects for the current loan age (one-year bins), interest rate (one-percentage point bins) and the borrower's estimated home-equity ($10,000 bins), as well as the full pairwise interaction between the borrower's LTV (10-point bins) and FICO score (50-point bins) at origination. Column 3 further interacts all of the additional fixed effects added in column 2 with the dummies for the month of observation. Column 4 adds an additional set of interactions between these fixed effects and the FHA dummy. Standard errors are reported in parentheses and are clustered at the CBSA level. Significance levels 10%, 5%, and 1% are denoted by *, **, and ***, respectively.

TABLE A.4
POOLED FULL-SAMPLE TRIPLE-DIFFERENCE ESTIMATES

	(1)	(2)	(3)	(4)	(5)	(6)
FHA	0.078	0.504***	0.689***	0.078	0.463***	0.640**
	(0.080)	(0.095)	(0.254)	(0.052)	(0.064)	(0.270)
FHA × Post	−0.481***	−0.397***	−0.380***	−0.462***	−0.385***	−0.375***
	(0.096)	(0.088)	(0.081)	(0.071)	(0.066)	(0.060)
FHA × Low Equity × Post	−0.679***	−0.689***	−0.462***	−0.657***	−0.674***	−0.458***
	(0.065)	(0.068)	(0.061)	(0.064)	(0.067)	(0.061)
FHA × ΔUR × Post	−0.023	−0.023	−0.031*			
	(0.019)	(0.018)	(0.018)			
FHA × High ΔUR × Post				−0.220***	−0.200***	−0.220***
				(0.075)	(0.069)	(0.071)
Month FEs	X	X	X	X	X	X
CBSA FEs	X	X	X	X	X	X
FHA Time Trends	X	X	X	X	X	X
Loan Age FEs		X	X		X	X
Interest Rate FEs		X	X		X	X
LTV × FICO FEs		X	X		X	X
Equity FEs		X	X		X	X
ΔUR FEs		X	X		X	X
Controls × Post			X			X
Controls × FHA			X			X
Number of Observations	15,645,645	15,645,645	15,645,645	15,645,645	15,645,645	15,645,645

NOTE.—This table reports estimates of the effect of the changes in SLR income documentation requirements and upfront costs on the monthly probability that an FHA loan refinances. The outcome variable is an indicator for whether or not a loan refinances in the month of observation and is multiplied by 100, so that all coefficients can be interpreted as percentage point changes. The Post dummy takes the value one if the month of observation is after the implementation of the policy changes (January 2010). In columns 1–3 the county-level unemployment change is entered linearly. In columns 4–6 it is entered as a dummy indicating whether the borrower lives in a county that experienced an above median increase in unemployment. All specifications include fixed effects for the month of observation and the CBSA of the property as well as linear time trends for FHA borrowers that are allowed to vary freely before and after the policy change. Columns 2 and 4 add fixed effects for the current loan age (one-year bins), original interest rate (one-percentage point bins), the borrower's estimated home-equity ($10,000 bins), and county-level unemployment change decile as well as the full pairwise interaction between the borrower's LTV (10-point bins) and FICO score (50-point bins) at origination. Columns 3 and 6 further interact the loan age, interest rate, LTV, and FICO fixed effects with the Post dummy and the FHA indicator. Standard errors are reported in parentheses and are clustered at the CBSA level. Significance levels 10%, 5%, and 1% are denoted by *, **, and ***, respectively.