Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.

Regression Discontinuity Estimates of the Effects of the GSE Act of 1992

Neil Bhutta

2009-03

Regression Discontinuity Estimates of the Effects of the GSE Act of 1992

Neil Bhutta[♣]

November, 2008

[Preliminary; please do not cite or redistribute without permission of author]

Abstract: In this paper I estimate the effect of the Underserved Areas Goal (UAG) established under the "GSE Act", a 1992 law mandating that the housing government-sponsored enterprises Fannie Mae and Freddie Mac help promote credit access and homeownership opportunities for low-income households and in low-income and minority neighborhoods. I identify the goal's impact by taking advantage of a discontinuity in the census tract eligibility rule. Employing local linear and non-parametric regression discontinuity methods, I find that this goal has had a direct effect on GSE purchasing activity of 3-4% and increases overall GSE-eligible originations by 2-3% on average at the cutoff between 1997 and 2002. Changing eligibility status following the release of Census 2000 data provides another source of variation to identify the UAG's effect in 2005 and 2006, years of sharply increasing goals levels and years which have contributed heavily to current credit losses. I find that while the UAG affected GSE behavior in 2005 and 2006, GSE risk avoidance limited their response. Unlike previous research, I find no evidence that UAG-induced increases in GSE credit supply crowds-out FHA and subprime lending.

[♣] Board of Governors of the Federal Reserve System, Washington, DC 20551 (neil.bhutta@FRB.gov) This research draws partially from one chapter of my dissertation at MIT. Thanks to Christine Brickman, Glenn Canner and Michael Greenstone for helpful comments and suggestions. All mistakes are mine. The views expressed in this paper are solely those of the author. Thanks to the MIT Dewey Library staff and the MIT Shultz fund for help in acquiring data, and the MIT Department of Economics and National Science Foundation for support.

1. Introduction

Recent turmoil in the mortgage market has brought considerable attention to public policies aimed at increasing credit supply and homeownership among relatively low-income households and in lower-income and minority neighborhoods. In particular, many have raised concerns that the Community Reinvestment Act (CRA), which encourages bank and thrift institutions to extend credit to lower-income neighborhoods and households, and the "GSE Act", under which Fannie Mae and Freddie Mac are expected to meet regulator-established mortgage purchase goals for mortgages made to lower-income and minority households and neighborhoods, have pushed regulated institutions to take excessive risk. In this paper, I test empirically whether the GSE Act has affected the behavior of Fannie Mae and Freddie Mac.

Congress established Fannie Mae and Freddie Mac to connect distant borrowers and investors, improving the efficiency of mortgage markets. Before being conserved in 2008, the two institutions operated as private "government-sponsored enterprises" (GSEs), with implicit and explicit financial benefits from the government (CBO 2001). In return for their sponsored status, the GSEs were expected to help promote homeownership opportunities for low-income and minority households and neighborhoods. The 1992 Federal Housing Enterprise Financial Safety and Soundness Act (or GSE Act) formalizes this responsibility. The GSE Act instructs the Department of Housing and Urban Development (HUD) to create "Affordable Housing Goals" for the GSEs and monitor their progress towards those goals. These goals, adjusted periodically, specify that a certain fraction of GSE mortgage purchases should be of mortgages made to targeted households and neighborhoods.

The GSE Act is one of several local and federal government programs and mandates intended to increase credit access and homeownership. One factor motivating these policies are sociological and economic theories suggesting that credit access and homeownership generate positive externalities such as reduced crime, improved child outcomes, increased investment in local public goods and increased voter turnout[1].

These policies have also been motivated as ways to help overcome apparent market failures such as racial discrimination (e.g. Munnell et al 1996). Another often cited market imperfection is that some of the information lenders use to price mortgages is a public good and

[1] See Kubrin and Squires for an example from the sociology literature. See Haurin et al (2003) for a review of the economics literature.

so there may be underinvestment in information by lenders resulting in suboptimal credit supply (Lang and Nakamura 1993). Lang and Nakamura argue, for instance, that a positive shock to credit supply that increases neighborhood home sales may generate public information about local home values that leads to a higher level of credit supply in equilibrium. Bhutta (2008) provides some evidence for this hypothesis, finding that the CRA pushes regulated lenders to increase their lending in targeted lower-income neighborhoods, which in turn leads to an increase in lending by *unregulated* lenders in some targeted neighborhoods – specifically, those neighborhoods that have had relatively low loan volume in the recent past, consistent with an information spillover story.

Several studies have documented positive trends in GSE purchases of loans from targeted groups over time. Figure 1 shows how the three affordable housing goals created by HUD have changed over time and each of the GSEs' corresponding purchase shares over time.[2] The initial goals (thick black line) were finalized in December 1995, then were increased in 1997 and then again in 2001, 2005 and 2006. Numbers generated from analyses by Manchester (2002, 2008) of GSE data are also shown in Figure 1 and suggest that the GSEs have been raising their share of purchases towards the target populations.

However, Figure 1 does not indicate whether the affordable housing goals have actually *forced* the GSEs to behave differently than they would in the absence of the Act. In this paper, I use comprehensive mortgage application data collected under the Home Mortgage Disclosure Act (HMDA) and test whether the Underserved Areas Goal (top panel of Figure 1), which targets relatively low-income neighborhoods, has had a causal impact on GSE purchase activity and subsequent mortgage credit flow in targeted neighborhoods. Under this goal, a purchased loan counts towards the GSEs' goal if the loan is for an owner-occupied property in a census tract that has a median family income less than or equal to 90% of the MSA median family income.[3] I exploit this discontinuity in the selection rule to identify the impact of the housing goal, in essence comparing GSE purchase volume and overall loan volume in tracts just below the 90% cutoff to that in tracts just above the cutoff. More precisely, I employ local nonparametric and

[2] See notes underneath Figure 1 for description of GSE housing goals

[3] The Underserved Areas goal also targets low and moderate income, predominantly minority neighborhoods: census tracts where minorities make up at least 30% of the population and median family income is no greater 120% of the MSA median family income. Ongoing research by the author exploits this discontinuity as well as the discontinuity in the other two Affordable Housing Goals to more thoroughly evaluate the GSE Act's impact.

semi-parametric regression discontinuity methods similar to those described in Porter (2003) and Imbens and Lemieux (2008).

Discerning the causal effect of the GSE goals on GSE behavior and overall credit flow is important for several reasons. First, in order to prevent future catastrophes it is crucial to have a clear understanding of the factors contributing to the current one. For example, falsely attributing the GSEs' collapse to their affordable housing responsibilities might lead policy makers to ignore other, more important determinants of the GSEs' failure.

Policy makers should also be interested in whether increased GSE purchasing activity ultimately affects the mortgage supply – whether, in fact, the GSEs can be used to achieve social policy ends. One reason the affordable housing goals would have little or no impact is if increased GSE purchase activity in targeted census tracts crowded out non-GSE purchasers. Another reason, proposed by An and Bostic (2008, 2006), is that the increased supply of GSE funds in targeted neighborhoods may draw the highest quality FHA and subprime applicants into conventional, conforming loans (still, a boon to consumer welfare, the authors note).

Finally, identifying and quantifying the GSE Act's effects can provide insight into the role the Act played, as opposed to market forces and other public policies, in the surge in homeownership and household mortgage debt since the mid-1990's. While the homeownership rate and household mortgage debt grew sharply in the years following passage of the GSE Act, a number of other policy, regulatory and (supply side) market changes could have driven these trends as well (Li 2005). From the late 1980's to mid 1990's, Congress strengthened HMDA, helping regulators and consumer advocates detect discrimination by mortgage lenders, as well as the CRA. At the same time, information technology improvements greatly reduced the cost of mortgage lending during the 1990's and improved the ability of mortgage lenders to price risk, which is important for subprime lending (see Straka 2000).

Only a few papers test for causal effects of the GSE Act. Ambrose and Thibodeau (2003) try to estimate the impact of the GSE Act on credit supply using a credit supply-demand model that assumes price increases indicate excess demand (loan volume is on the supply curve) and vice versa. They conclude that the GSE Act has had little impact based on their finding that there is only a small (positive) relationship between MSA loan volume and the MSA population share living in GSE targeted census tracts. However, it is possible that their results are biased downward as this regressor is likely negatively correlated with many MSA-level socio-economic characteristics that predict mortgage volume.[4]

Bostic and Gabriel (2006) estimate the impact of being a GSE-targeted census tract on tract housing outcomes (e.g. median home value) in California using the 2000 Census. Similar to the identification strategy in this paper, they take advantage of the discontinuous eligibility rule, comparing outcomes for tracts within ten percentage points below the cutoff to those within the same distance above the cutoff.[5] They do not find evidence of a treatment effect, but the interval they use is large and they do not control for the "assignment variable" (i.e. tract to MSA median family income ratio). As the results in this paper will show, such a strategy can lead to severe bias despite controlling for many other covariates.

Finally, An and Bostic (2008, 2006) test whether increased GSE activity in targeted tracts crowds out FHA and subprime lending, motivating their study as an exploration of one reason the GSE Act appears to have had a limited impact on real outcomes as in Bostic and Gabriel above. In these two papers, their outcomes of interest are the change (between 1996 and 2000) in FHA and subprime shares of total tract originations, respectively. They regress these variables on the change in GSE share of originations.[6] Further, they propose instrumenting this regressor with a variable indicating for being a GSE-targeted census tract. Similar to Bostic and Gabriel's study described above, An and Bostic's main regressions use tracts within ten percentage points of the 90% cutoff, and again they do not control for the assignment variable.

An and Bostic find statistically significant negative second-stage relationships between GSE share and the two outcome variables. But since they fail to find a first stage relationship, these results likely represent the equilibrium cross-sectional relationship one would expect – that tracts with more conforming, prime loans purchased by the GSEs will tend to have fewer subprime loans.[7]

[4] Another slight criticism of their study is that it spans 1995 to 1998 despite 1995 being a "transition period" year when goal levels were quite low and the definition of "underserved census tract" differed from that for years after 1995.

[5] An et al (2007) show results of the same set of tests for the nation as a whole. They also fail to find any impact in targeted tracts.

[6] GSEs typically purchase conventional, conforming loans and not government-insured loans (i.e. FHA and VA) or subprime loans. See Table 2.

[7] Methodologically speaking, An and Bostic do not appear to execute a formal instrumental variables strategy. They generate predicted GSE share in a first stage equation, and use these predicted values in a second stage equation where the outcome variable is either FHA loan share or subprime loan share, but they do not appear to include the same set of covariates that they used in the first stage regression.

Regardless of these potential shortcomings, analyzing loan shares may generate misleading results as well. Take for example (1.1) below, which regresses tract-level FHA loan share on tract-level GSE-purchased loan share (a la An and Bostic, 2008):

$$(1.1) \qquad \left(\frac{FHA}{GSE\text{-}Ineligible\ +\ GSE\text{-}Eligible}\right)_i = \alpha + \beta\left(\frac{GSE\text{-}Purchased}{GSE\text{-}Ineligible\ +\ GSE\text{-}Eligible}\right)_i + \mathbf{X}_i\boldsymbol{\theta} + \varepsilon_i$$

Instrumenting the GSE-purchase share variable with a treatment dummy is not valid since the treatment may affect *GSE-Eligible* (indeed, the ultimate goal of the Act is to increase overall credit flow), which enters in the dependent variable. As such one cannot interpret a negative estimate of β as evidence that FHA lending falls in response to increased GSE purchase activity as An and Bostic do.

In contrast to An and Bostic, I estimate separately the impact of the Underserved Areas Goal (UAG) of the GSE Act on (1) the number of GSE purchases, (2) the total number of GSE-eligible originations and (3) the number of GSE-*in*eligible loans in targeted tracts. I also use data for all years between 1994 and 2002 in order to generate more precise estimates and because 2001 and 2002 follow a sharp increase in the UAG level. And, as mentioned before, I implement local regression discontinuity methods to help mitigate identification concerns.

As a further extension, I use the change in target status of some tracts following the release of Census 2000 income data to identify the UAG's effect in 2005 and 2006, years when the GSEs met increasing goal levels and also years contributing heavily to current credit losses.[8]

To preview, I find evidence of a direct effect of the UAG on GSE purchasing activity and "GSE-eligible" originations, but of modest magnitude (3-4% and 2-3%, respectively). At the same time, I find no evidence of a reduction in "GSE-ineligible" lending, which is comprised mainly of FHA and subprime loans (see Section 2.2).

These results are similar for different bandwidths and control function specifications. I also find that the relationship between GSE-eligible loan volume and the assignment variable (conditional on covariates) is generally smooth – that is, I find almost no other discontinuities at points away from the cutoff. However, the results for GSE purchases and GSE-eligible lending

[8] For instance, Fannie Mae's 2008 2Q results indicate that 2005 and 2006 vintage mortgages account for more than 50% of their credit losses

just mentioned are conditional on tract-level covariates, including a lagged ("pre-treatment") value of the outcome variable. Baseline regression discontinuity estimates excluding covariates are in general slightly smaller but quite noisy.

With respect to more recent outcomes, I find that while tracts that had deteriorating income during the 1990's and thus became targeted by the GSEs under the UAG in 2004 experienced relatively *greater* loan growth between 2001 and 2006, a careful comparison of tracts that fell just below the UAG threshold in 2004 against those that fell just above fails to provide support for the notion that this goal was an important driver of that trend. Allowing for the effect of crossing the threshold to be a function of the magnitude of a tract's deterioration, I find that while the GSEs do respond positively to the UAG, their response seems to be limited considerably by risk avoidance.

In the next section, I discuss the data and the regression discontinuity strategy that I use to estimate the effect of the UAG between 1997 and 2002, and in Section 3 I describe these results. In Section 4, I develop the estimation strategy to evaluate the UAG's impact in 2005 and 2006, discuss data issues related to estimation and present the results. Finally, I conclude in Section 5.

2. Data & Empirical Strategy

2.1. Overview

I take advantage of a sharp discontinuity in the UAG eligibility rule to identify the goal's impact on GSE purchase volume and overall credit flow between. Again, GSE purchases of mortgages for owner-occupied properties in census tracts with a median family income no greater than 90% of MSA median family income count towards meeting the UAG. In the regression discontinuity (RD) analysis that follows, this ratio of tract median family income to MSA median family income is the "assignment" (or "running") variable and will be referred to as TM. Thus, census tracts with $TM \leq 0.90$ are targeted by the GSEs and the impact of the UAG *at the cutoff* will be identified by measuring the jump in purchase and loan volume at $TM = 0.90$.

Tract and MSA median family income are based on decennial Census income data and MSA definitions, which change periodically. Between 1994 and 2002 almost all census tracts had a constant value of *TM* based on 1990 Census data and 1993 MSA definitions (by OMB).[9] In 2004, tracts' *TM* was re-calculated based on 2000 Census data and new MSA definitions, I will discuss in more detail in Section 4 how I use this change to identify the UAG's effect in recent years.

2.2. Data & Summary Statistics

Data on census tract level mortgage activity comes from information submitted by lenders under the Home Mortgage Disclosure Act (HMDA 1977). Since 1990, lenders covered by HMDA have been required to compile and submit detailed information on the individual mortgage applications they receive. And in 1993 HMDA started including a large number of independent mortgage companies previously exempt from HMDA reporting.[10]

A number of variables available in the HMDA will be important to this analysis. One is the census tract of the loan: for each loan application lenders are asked to report the census tract of the property.[11] HMDA also provides a unique lender ID, which I will combine with data from HUD to identify loans in HMDA as being originated by lenders that specialize in subprime lending.[12] Other loan-level variables I will use are the loan amount, the disposition of the loan (e.g. approved, originated, denied, etc.), the loan purpose (e.g. refinance) and the type of loan (e.g. FHA, VA or conventional). Finally, lenders also report whether the loan was purchased and to whom it was sold, including Fannie Mae and Freddie Mac. This variable will be used to count up and compare the number of GSE-purchased loans in targeted versus not-targeted tracts. See Table 1 for a full list and description of the variables available in HMDA.

Census tract-level characteristics from the 1990 Census (and 2000 Census for the analysis in Section 4), compiled and distributed by Geolytics, will also be used. These characteristics will be merged to the HMDA loan data using the 1990 census tract code to create a tract-level dataset with yearly mortgage activity variables from HMDA and tract demographic and housing characteristics from the Census.

---

[9] The exception is tracts that are part of the few newly formed MSAs between 1994 and 2002. I only use tracts that are in the 1993 set of MSAs.
[10] See Avery et al (2007) for more details on HMDA data. Also see FFIEC website (www.ffiec.gov) for rules defining which lending institutions are required to report under HMDA.
[11] See Avery et al (2007) for details and reliability of geographic reporting requirements. 1990 census tract definitions apply to HMDA data between 1992 and 2002, and 2000 census tract definitions have been used since 2003.
[12] Each year, HUD releases a "subprime lender list". See www.huduser.org .

Although the GSE Act covers rural and urban areas, the analysis focuses on census tracts in MSAs as HMDA data are unreliable in rural areas (Avery et al 2007). I also exclude census tracts in MSAs in Hawaii and Alaska as well as those in MSAs formed between 1993 and 1999 in order to maintain a constant set of geographies for the period under study, 1997-2002. HMDA reporters used new census tract and MSA definitions starting in 2003 and 2004, respectively, which necessitates a separate study of the GSE Act using HMDA data in years after 2002.

Some outlier census tracts were dropped as well. Those with fewer than 100 housing units, those with zero "specified" owner-occupied units, those with more than 30% of the population living in group quarters and those with fewer than one home purchase or refinance origination per (1990) owner-occupied between 1997 and 2002 or more than ten originations per (1990) owner-occupied unit between 1997 and 2002 were dropped. In the vicinity of the GSE-eligibility cutoff (i.e. the census tracts of interest for the regression discontinuity design described below), just over 97% of census tracts are retained in the sample.

Table 2 provides means of tract-level mortgage activity and tract characteristics. Panel A shows average mortgage volume between 1997 and 2002. Column 1 provides means and standard deviations for all tracts that meet the sample selection criteria. Columns 2 and 3 provide means and standard deviations for tracts just below and above the UAG cutoff, respectively, and column 4 provides the p-Value for a test of the equality of the means in columns 2 and 3. The top row of Panel A shows total refinance and home purchase, owner-occupied originations. Tracts around the cutoff had loan volume below the average, while those just above the cutoff had about 8% more than those just below.

The second row of Panel A provides the number of "GSE-eligible" originations per tract. GSE-eligible originations are those with loan amounts within the conforming loan limit set by Congress, are conventional (i.e. not FHA or VA insured), and are not originated by a subprime lender as defined by HUD. Eligible loans account for about two-thirds of the market, while the number of eligible loans in just-targeted tracts is about 10% lower relative to just-not-targeted tracts. These figures also show that the GSEs purchase about two-thirds of the eligible loans that are purchased. About one-third of loans are not purchased in the year that they are originated. If these loans are sold in later years – called 'seasoned' loans – they will not generally be reported in HMDA. As such, HMDA does not account for all GSE purchase activity (see Scheessele 1998).

In terms of "GSE-*in*eligible" loans, almost none of these are purchased by the GSEs. The majority of these loans are relatively high-priced, high-risk FHA-insured and subprime loans, while the rest are VA-insured and "jumbo" loans (above the single-family conforming loan amount). The few ineligible loans purchased by the GSEs may be attributed to reporting error and to the fact that identifying eligible loans in the HMDA data is not perfect – for instance, subprime lenders may make prime loans and vice versa.[13] Interestingly, in contrast to the case for eligible loans there is no statistical difference in ineligible lending for the two groups around the cutoff.

Panel B shows tract averages of housing and demographic characteristics measured in the 1990 Census. Similar to the pattern for overall origination volume, the number of owner-occupied housing units is significantly lower in GSE-targeted tracts relative to those not targeted, but the difference (5%) is not as great as the difference in loan volume (8%). At the same time, tract-size is not substantively different across the cutoff (second row), which is not surprising since tracts are designed to be of similar size.

Most of the other housing and demographic characteristics are (statistically) significantly different across the cutoff. These results suggest that these two groups of tracts are not similar despite their being relatively near the cutoff. In other words, tract characteristics appear to change quickly with the running variable, *TM*. As such, it will be important to employ a strategy that will adequately control for these differences.

2.3. Empirical Strategy: Regression Discontinuity

Consider the following tract-level regression of potential outcomes such as mortgage origination volume on a treatment indicator variable, $D_i = \mathbf{1}[TM_i \leq 0.90]$:

$$Y_i = \alpha + \beta D_i + e_i \tag{1.2}$$

The following expression captures the intuition behind the regression discontinuity design:

[13] Overall, though, mortgage lenders tended to specialize in either prime or subprime lending during this period. (see Nichols et al 2005). Another source of discrepancy is the fact that conforming loan limits are higher for 2, 3 and 4 unit dwellings, but the HMDA do not provide information on the number of units for 1-4 family properties.

(1.3) $$\lim_{h\to 0}\left\{\mathrm{E}[e_i \mid 0.90-h \le TM_i \le 0.90]-\mathrm{E}[e_i \mid 0.90 < TM_i \le 0.90+h]\right\}=0$$

(1.3) implies that GSE targeted and not-targeted tracts arbitrarily close to the cutoff (TM = 0.90) are identical in expectation with the exception of their eligibility status. As such, any substantive difference in outcomes across the cutoff for tracts 'near' the cutoff can be attributed to a GSE treatment effect.

With that in mind, a natural approach to estimating β is to simply compare the mean of Y_i for tracts "just below" the cutoff to that for tracts "just above" the cutoff. This can be done in a single regression, weighting observations with kernel weights $W_i = K([TM_i - 0.90]/h)$, where h is some chosen bandwidth. I will show estimates using both a simple rectangular kernel ($K(a) = \mathbf{1}[0\le |a| \le 1]$) as well as a triangular kernel ($K(a) = \mathbf{1}[0\le |a|\le 1]*[1-|a|]$) that gives the most weight to tracts closest to the cutoff.

Porter (2003) shows that the bias of this nonparametric approach increases (for a given h) in the slope of the relationship between the outcome and running variables. More concretely, since mortgage activity is negatively correlated with tract income (see Table 2) nonparametric estimates of β from will be downward biased (i.e. against finding a positive impact of the UAG). I try to mitigate this bias by using a small bandwidth (h = 0.02), triangular kernel weights and adding tract-level covariates into the regression, including a lagged ('pretreatment') value of the outcome variable.

Following Imbens and Lemieux (2007), I also perform local linear regression to estimate β. This strategy allows one to use data further away from the cutoff by controlling for the slope of the relationship between the outcome and assignment variables on either side of the cutoff. Simply stated, in this approach I fit a line to the data within a distance h on either side of the cutoff and β is calculated as the difference between the intercepts of these two estimated lines. This will be done in a single, kernel weighted regression:

(1.4) $$Y_i = \alpha + \beta D_i + TM'_i + TM'_i * \mathbf{1}[TM_i < 0.90] + \mu_i$$

where $TM'_i = TM_i - 0.90$. The term $TM'_i + TM'_i * \mathbf{1}[TM_i < 0.90]$ in (1.4) is often referred to as the "control function". For all regressions, robust standard errors clustered at the MSA level are reported.

3. Results

3.1. Testing the Identification Assumption

The RD identification assumption is that all tract characteristics affecting tract loan volume change smoothly across the cutoff except for tract eligibility status, which changes sharply at the cutoff. Although this assumption is not fully testable, one may be more confident it holds if observable, "pre-treatment" tract characteristics change smoothly across the cutoff.

Figure 2 suggests that observable tract characteristics change quite smoothly across the cutoff. Figure 2 plots the predicted values from a regression of the (log) number of tract originations between 1997 and 2002 on the set of tract characteristics listed in Panel B of Table 2, with the exception that I log-transform the number of owner-occupied units, total number of housing units and median home value before entering them into the regression. The regression also includes MSA fixed effects but the estimated fixed effects are not used to generate the predicted values.

Each data point shown in Figure 2 represents a mean predicted value for tracts in a one percentage point interval of *TM*. Also shown are local linear regression estimated lines on either side of the cutoff. These fitted lines are estimated using the underlying tract-level predicted values. The two lines meet close to each other at the cutoff, indicating that the set of tract characteristics used in the regression do not change discretely at the cutoff since the tract characteristics just to the left of the cutoff predict nearly same volume of loans as the tract characteristics just to the right of the cutoff (conditional on MSA).

3.2. Effects on Secondary Market Purchasing Activity

Figure 3 illustrates the basic idea of the RD design, and suggests that the effect of the UAG on GSE purchase activity is small (in percentage terms). Figure 3 is analogous to Figure 2, except that the Y-axis variable in this case is (log) GSE-purchases of originations that are eligible for GSE purchase per (1990) owner-occupied unit. The gap between the two lines implies an effect at the cutoff of less than 5%.

Table 3 provides local linear regression and nonparametric RD estimates of the effect of the UAG on GSE purchase activity between 1997 and 2002. In all the regressions, I include MSA fixed effects so that the discontinuity is identified only from variation across the cutoff within MSAs. I also control for two tract scale variables – (log) number of owner-occupied units and (log) total housing units, both measured in 1990. Columns 1-7 show estimates using a bandwidth of five percentage points, while columns 8-11 show estimates using a bandwidth of two percentage points.

The point estimate in column 1 basically provides the difference in mean GSE purchase activity across the cutoff after adjusting for tract size and MSA, and shows the GSEs purchase about 10% fewer loans in tracts with *TM* between 85 and 90 relative to those between 90 and 95. Even after adjusting for the remainder of the covariates listed in Table 2, the GSEs purchase about 3% fewer loans (column 2).

Column 3 institutes the local linear approach described earlier, controlling only for MSA and tract size. The coefficient on *TM* is significant and the discontinuity estimate now is positive (1.5%) though not statistically different from zero. The results of the first three columns illustrate the difficulties of a strategy that tries to identify the GSE effect simply by controlling for observables.

In column 4 I add in the set of covariates used in column 2. Under the assumptions (1.3) and that the linear *TM* terms correctly control for the trend in GSE purchases on either side of the cutoff, covariates are not needed in theory to identify the discontinuity and including covariates might be viewed as deviating from the spirit of the quasi-experimental RD design. But including covariates can improve precision and help correct small sample bias in the basic specification (Imbens and Lemieux 2007). Of course, if the inclusion of covariates results in "large" differences in the point estimates, one might be concerned that the identification assumption (1.3) does not hold. But Figure 2 suggests this should not be the case. The standard error in column 4 does fall and the point estimate rises only slightly in absolute terms – about 0.007 – again suggesting that observables change quite smoothly across the cutoff.

In column 5 I include the (log) number of GSE-purchases between 1994 and 1996 as a regressor that will control for unobservable tract characteristics and likely improve the precision of the estimates further. I consider 1994-1996 a pre-treatment period as these years come before the goal level increases in 1997 and 2001. Nevertheless, if the GSEs responded to knowledge of the future goals during these early years, then including this lagged value will net out this earlier impact.

The estimate (standard error) in column 5 is about 0.031 (0.017). While marginally significant, I view this result with some caution as the point estimate has now roughly doubled relative to the baseline specification in column 3. Still, the absolute value of the difference between the point estimates in columns 3 and 5 is small (about 0.015).

Columns 6 and 7 use triangular kernel weights. These estimates are slightly larger than those using rectangular weights. And again, including the lagged value reduces the standard error considerably. These estimates suggest a discontinuity in GSE-purchase volume of just under 4%.

Columns 8-11 show nonparametric estimates (in the sense that they do not include parametric controls for *TM*) and the bandwidth (*h*) has been cut to 0.02. Columns 8 and 9 use rectangular weights while 10 and 11 use triangular weights. The column 8 specification is analogous to that in column 2. These estimates are similar in magnitude to those discussed earlier and the specifications that include the lag (columns 9 and 11) are statistically significant.[14]

Overall, Table 2 provides evidence of a 3-4% effect of the UAG on GSE purchase activity at the cutoff, although this is conditional on including tract covariates. Next, I look for evidence of whether increased GSE purchase activity in targeted tracts crowds out non-GSE purchases of GSE-eligible originations in targeted tracts.

Table 4 provides estimates of the discontinuity in non-GSE purchase activity in targeted tracts. The eight specifications in Table 4 correspond to the specifications in columns 4-11 in Table 3. Table 4 provides no evidence of crowd out. On the contrary some of the point estimates, in particular those specifications controlling for a lag of the outcome variable, are similar in magnitude to those in Table 3 although only one of the estimates is statistically significant. One reason that non-GSE purchases might increase in targeted tracts is that GSE-eligible mortgages in these areas now represent more liquid assets because of the UAG and so may be more attractive to non-GSE secondary market participants.

[14] It is worth noting that the point estimates in columns 9 and 11 rise slightly with the inclusion a linear control function.

Reporting error by HMDA respondents is another reason non-GSE purchases might *appear* to increase. In particular, the *purchaser* variable in HMDA may be reported with considerable error (Scheeseele 1998). If lenders do not accurately report to whom they sell their loans, and if the UAG does in fact lead to an increase in credit flow, this combination could result in an observed increase in non-GSE purchases that simply reflects the underlying expansion in GSE-eligible credit flow.

3.3. Effect on Overall Credit Flow

The ultimate objective of the GSE Act is to increase credit flow to targeted groups. Instead of looking at direct purchases by the GSEs, I now look at the effect of the UAG on all GSE-eligible originations from 1997-2002 (3rd row of Table 2). To be more concise, I show only the specifications corresponding to those in columns 5, 6 and 8 in Table 4. Other specifications shown earlier generate similar results for the outcome variables in the tables to follow. The estimates in Table 5 follow a pattern similar to those in Tables 3 and 4. In particular, including the lag as an independent variable yields a slightly higher and statistically significant estimate of about 2.7% that is robust to using a triangular kernel (column 3).

One useful exercise is to show discontinuity estimates at points other than *TM* = 0.90. If the same specification(s) yields significant discontinuities elsewhere, that would indicate that the discontinuity found at *TM* = 0.90 may be spurious. Figure 4 shows estimated discontinuities in (log) GSE-eligible originations using the specification in column 2 from Table 5 (i.e. $h = 0.02$, no control function used and includes the lag) at 30 different values of *TM* between 0.75 and 1.05. Other than the discontinuity at 0.90, there is only a negative discontinuity at 0.88 and a positive discontinuity at 0.81. The discontinuity at 0.88 is likely related to the discontinuity at 0.90. This estimate represents the conditional mean for tracts in the *TM* interval [86, 88] relative to that in the interval (88, 90] and therefore suggests the treatment effect may be concentrated in tracts right near the cutoff. There is no clear reason for a discontinuity at 0.81, but when testing for a discontinuity at so many values, it should not be surprising in a statistical sense to find at least one significant jump. Overall, Figure 4 suggests the empirical relationship between loan volume and *TM* is smooth and that the discontinuity at 0.90 is not likely to be spurious.

Table 6 breaks the estimates of Table 5 into three two-year period estimates. The results suggest that the UAG was not binding until the 1999-2000 period, and that the discontinuity just before and after the sharp increase in the UAG level from 24% to 31% was basically the same. That the discontinuity is similar before and after the 2001 change is surprising. Perhaps expectations of a substantive increase in the UAG level led to an anticipatory adjustment prior to the actual enactment. But the new goal levels were released in October, 2000 and I have not found evidence of considerable public debate about raising the goal level prior to that date.

The results thus far suggest two reasons why one might fail to find substantive differences in housing outcomes across the cutoff measured in the 2000 Census as in Bostic and Gabriel's (2006) previously mentioned study. One is simply that the measured effects are fairly small. And two, these effects do not show up until the 1999-2000 period, while the 2000 Census measures outcomes in April of 2000.

3.4. Testing for Crowd-out of Subprime and FHA Loans

As hypothesized in two papers by An and Bostic (2006, 2008), if prime lenders are more aggressive in pursuing applicants in targeted areas they may attract the highest quality FHA and subprime applicants. Along these lines, other research suggests that many borrowers that obtain high cost FHA or subprime loans likely could have qualified for a prime loan (e.g. Pennington-Cross et al 2000, Temkin et al 2002). Alternatively, new GSE flexible lending programs could also crowd-out FHA and subprime lenders (see Quercia et al 2002). As An and Bostic argue, while these dynamic interactions might limit the impact of the GSE Act on aggregate credit flow or even reduce total flow, it may improve consumer welfare by helping get borrowers into cheaper mortgage contracts.

In Table 7 I test for crowd-out by estimating the discontinuity in GSE-*in*eligible lending, which consists primarily of FHA loans and loans by subprime lenders (as determined by HUD). Also included in this set of loans are VA loans and loans that are above the conforming loan amount limit ('jumbo' loans). Table 7 provides no evidence of crowd-out as argued by An and Bostic. In fact, all of the point estimates are positive, but relatively small (about 1%) and statistically insignificant.[15]

---

[15] The results are essentially identical when focusing just on loans originated between 1999 and 2002, the years when the discontinuity in GSE-eligible lending is largest (see Table 6).

4. The Underserved Areas Goal and GSE Activity in 2005 and 2006

In addition to cross-sectional variation in treatment status, *TM* also varies over time. After the release of 2000 Census data, *TM* was revised to reflect the new data and new MSA definitions in 2004[16] (I will refer to the new value as TM_{new} and the old value as TM_{old}). This change yields a set of tracts targeted under the UAG after 2004 and were not previously targeted (i.e. $TM_{old} > 0.90$ and $TM_{new} \leq 0.90$). I will use these tracts and exploit the change in target status for some of them to identify the UAG's effect on neighborhood credit supply in 2005 and 2006.

For a sample of tracts *not* targeted prior to 2004, consider the following regression model:

$$\Delta Y_i = \alpha + \beta \Delta D_i + \Delta e_i \tag{4.1}$$

where Δ represents the change across 2004. Equation (4.1) is not generally identified since unobserved factors will likely drive both changes in treatment status and changes in mortgage activity (i.e. unobserved deterioration in neighborhood quality could cause both treatment status and mortgage activity to change). However, the regressor of interest, ΔD, is a deterministic function of TM_{new} (given that the sample is comprised of tracts with $TM_{old} > 0.90$). This observation leads to the following estimating equation:

$$\Delta Y_i = \alpha + \beta \Delta D_i + \mathrm{E}[\Delta e_i \mid TM_{i,new}] + \eta_i \tag{4.2}$$

where $\Delta D_i = \mathbf{1}[TM_{i,new} \leq 0.90]$ and $\eta_i = \Delta Y_i - \mathrm{E}[\Delta Y_i \mid TM_{i,new}]$. The third term in (4.2) represents the control function. Again, this term is a function of the assignment variable, TM_{new}, that controls for *all* variables correlated with ΔD and ΔY not explicitly included in the regression. As in the prior section, I implement (4.2) using local linear and non-parametric methods.

Intuitively, this strategy aims to compare the change in lending in tracts that *just switched* treatment status to those that *almost switched*. In other words, (4.2) merges a difference-in-difference (DD) identification strategy with an RD strategy.

[16] Tracts' median family income based on the 2000 Census was first used in 2003. However, MSA definitions changed considerably between 2003 and 2004, altering many tracts' *TM* again. I ignore outcomes based on 2003 values of *TM* since these figures were temporary.

Table 8 provides group means of various housing and credit flow variables for tracts that switched treatment status between 2002 and 2004 ("switchers") versus tracts that did not switch ("non-switchers"). I limit the sample to tracts used in the earlier analysis (see Section 2.2) with TM_{old} between 0.90 and 1.10 because most switchers (1850 of 1999 tracts) come from this group. I also exclude tracts with minority population share in 1990 above 0.30 since these would have been targeted under the UAG between 1997 and 2002. Finally, I limit the sample to those tracts with only minor boundary changes between 1990 and 2000 so that I am using a set of tracts that are comparable before and after 2004.[17]

Table 8 shows that TM_{old} is slightly lower for the switchers (first row, panel A) and TM_{new} differs considerably, indicating that the two groups had opposing income trends between 1990 and 2000. Other characteristics in panel A provide further evidence of these divergent trends, with the exception of minority population share, although switching tracts minority share grew more sharply. Cleary, these two groups are not comparable in terms of their pre-treatment trends, invalidating a simple DD strategy. The RD strategy just discussed, however, will attempt to control for these divergent trends and isolate the effect of switching into the treatment group.

Panels B and C show 1-4 family owner-occupied, home purchase and refinance mortgage origination and GSE purchase volume in 2005-2006 and 2001-2002, respectively. For the 2001-2002 period, GSE-eligible and ineligible mortgages are defined as in previous sections. For 2005-2006, since the HUD subprime lender list is not available for 2006, I use instead the *APR Spread* variable to identify subprime mortgages.[18] Although the 2005-2006 HMDA provides information on lien status, I ignore this information in order to maintain comparability with the 2001-2002 data.

Panels B and C indicate that GSE-eligible origination volume fell between 2001 and 2006, and more so for *non*-switching tracts. At the same time, ineligible lending rose in general suggesting a sharply growing subprime share of originations during this period. Also interestingly, the GSE-purchased share of eligible originations fell significantly for both groups.

[17] I use the census tract population-based relationship file to link 1990 census tracts to their 2000 counterpart. Just over 85% of the census tractsI use in the 1997-2002 analysis are linked to a single 2000 census tract, where a positive link requires at least 90% of the 2000 tract's population to reside within the 1990 tract boundary and vice versa. Most of the remaining tracts were relatively large in 1990 and were split up for the 2000 Census.

[18] See Avery et al (2006) for a detailed discussion of this variable and the 2005 HMDA data.

One reason may be that an increasing number of “eligible” loans are in fact not conforming, consistent with an increasing subprime share of the mortgage market during this period.[19]

The first column of Table 9 provides DD estimates of the effect of switching on GSE purchases of eligible originations (Panel A), GSE-eligible originations (Panel B) and GSE-ineligible originations (Panel C). The results, although they lack a causal interpretation, are nevertheless interesting as they show that mortgage lending grew relatively more in switching tracts, that is, those tracts that became poorer between 1990 and 2000.[20]

In column 2, I exclude tracts more than 20 percentage-points away from the cutoff, but the point estimates do not change substantially, implying that the DD estimates are not driven by tracts that experienced large income changes between 1990 and 2000. Column 3 maintains the same bandwidth and includes a linear control function. The new estimates, which arguably have a causal interpretation, fall considerably relative to those in column 2. Assuming the control function correctly controls any important heterogeneity across the cutoff, these estimates indicate that tracts that *just* fell into the treatment group experienced 3.6% more GSE-purchases and 2.8% more GSE-eligible originations relative to tracts that just missed falling into the treatment group.

Columns 4 and 5 use narrower bandwidths (0.10 and 0.05, respectively). Column 4 maintains the same specification as in column 3, while in column 5 I do not include a control function, but do include a set of covariates (see table notes for list). In terms of GSE purchases, the new point estimates are about half the size of that in column 3 and no longer statistically significant. The estimates in column 5 for eligible and ineligible originations are also small and statistically insignificant. Columns 4 and 5 therefore suggest that those in column 3 are biased *upward*. In sum, while Table 9 indicates that tracts that had deteriorating income during the 1990’s experienced relatively *greater* loan growth between 2001 and 2006, a careful comparison of tracts around the UAG threshold fails to provide support for the notion that this goal was an important driver of that trend.

Of course, these estimates represent the average effect for all tracts falling just below the cutoff relative to those falling just above it. If the GSEs hesitate to lend in targeted tracts that have deteriorated considerably, there will be heterogeneity in the UAG’s effect on credit supply.

---

[19] These share shifts may be affected (in an unknown way) by the change in the way I identify subprime loans in HMDA across the two periods.

[20] To be more precise, the estimates imply that loan and purchase volumes *shrank* less quickly in switching tracts relative to non-switching tracts.

In other words, if the GSEs weigh safety and soundness considerations against their social commitments, then the effect of falling below the UAG threshold will be a function of how far the tract has fallen.

Table 10 shows the results from regressions (h = 0.05) that allow for the effect of switching into the treatment group to vary by TM_{old} :

(4.3) $$\Delta Y_i = \alpha + \beta_1 \Delta D_i + \beta_2 \Delta D_i * TM'_{i,old} + \lambda TM'_{i,old} + \mathbf{X}_i \boldsymbol{\delta} + \eta_i$$

where $TM'_{i,old} = TM_{i,old} - 0.90$, so that the coefficient on ΔD represents, in essence, the estimated effect of the UAG for tracts that fell from just above the threshold to just below between 2002 and 2004.

The estimates in column 1 imply that the UAG has a 6% effect on GSE purchases for tracts that fell just below the threshold in 2004 from just above, and that this effect trails off by 0.5% per unit increase in TM_{old}, consistent with the notion that the GSEs do react to the UAG, but weigh safety and soundness considerations against these social obligations.

Columns 2 and 3 show results for GSE- eligible and ineligible originations, respectively. Although the relatively large standard errors limit confidence in the point estimates, their magnitudes – 2.7% for eligible originations compared to 0.4% for ineligible originations – are consistent with increased GSE purchases due to the UAG having a positive effect on overall eligible credit supply and no effect on FHA and subprime lending.

4. Conclusions

In this paper, I identify the impact of the Underserved Areas Goal (UAG) established under the GSE Act by taking advantage of a discontinuity in the census tract eligibility rule.

I find that this goal has a direct effect on GSE purchasing activity of 3-4% and increases overall GSE-eligible originations by 2-3% between 1997 and 2002. Unlike previous research, I find no evidence of an offsetting reduction in FHA and subprime loans. I do not find discontinuities at points away from the cutoff (Figure 4), buttressing the causal interpretation of these results. However, the results do rely on including covariates to make the results precise and thus lack the interpretive confidence associated with a "true" RD design.

One important limitation of the regression discontinuity strategy is that it only identifies the goal's impact at the eligibility cutoff. The results in this paper say little about the UAG's effect on census tracts away from the cutoff. Of course, this limitation is countered by the RD design's ability to overcome a challenging identification problem as well as the ability to identify general equilibrium effects (e.g. crowd-out) within this framework.

The 2.7% estimated effect on GSE-eligible lending (Table 5) translates into about 18 extra home purchase and refinance originations, or about $2.2 million (in 2007 dollars) in credit, per tract *at the cutoff*.[21] Applying this number to the roughly 1100 sample tracts just below the cutoff establishes a lower bound on the aggregate impact of the UAG of about $2.4 billion between 1997 and 2002, which seems considerably smaller than the subsidy the GSEs maintained at the time. For instance, CBO (2001) estimated at over $12 billion (in 2007 dollars) in 2000 alone.

I also test for an effect of the UAG in 2005 and 2006, years when the GSEs met increasing goal levels and also years that contributed heavily to current credit losses. I find that while tracts that that fell into the group of targeted tracts in 2004 experienced 10% *greater* loan growth between 2001 and 2006 relative to tracts remaining not-targeted, a more careful comparison of tracts that fell *just below* the UAG threshold in 2004 against those that fell to *just above* the threshold fails to provide support for the notion that this goal was an important driver of that trend. Allowing for the effect of crossing the threshold to be a function of the magnitude of a tract's deterioration suggests that while the GSEs do respond positively to the UAG, their response is tempered by safety and soundness considerations.

Overall, the results of this research provide little support for the notion that the affordable housing goals pushed heavily on the GSEs to overexpose themselves to risky mortgages and that adherence to the goals fueled the current subprime crisis. Similarly, Bhutta (2008a and 2008b) argues that the "Special Affordable Goal" (bottom panel of Figure 1) has not distorted credit supply measurably. Nevertheless, the research on the effect of the affordable housing goals remains incomplete without detailed analyses of all the goals. It is also important to keep in mind the inability of the RD design to provide credible information about the goals' effect on GSE behavior away from the cutoff.

---

[21] Tracts with $88 < TM \leq 90$ had on average 679 originations between 1997 and 2002. Deflating that amount by 2.7% equals 18 originations, and loan amounts are about $120,000 on average over the period in 2007 dollars.

References

Ambrose, Brent W., and Thomas G. Thibodeau (2004). "Have the GSE Affordable Housing Goals Increased the Supply of Mortgage Credit?" *Regional Science and Urban Economics,* vol. 34, no. 3, pp. 263-273.

An, Xudong, and Raphael Bostic (2008). "GSE Activity, FHA Feedback, and Implications for the Efficacy of the Affordable Housing Goals," *The Journal of Real Estate Finance and Economics,* vol. 36, no. 2, pp. 207-231.

An, Xudong, and Raphael W. Bostic (2006). "Have the Affordable Housing Goals been a Shield Against Subprime? Regulatory Incentives and the Extension of Mortgage Credit," USC Lusk Center for Real Estate.

An, Xudong, Raphael W. Bostic, Yongheng Deng, and Stuart A. Gabriel (2007). "GSE Loan Purchases, the FHA, and Housing Outcomes in Targeted, Low-Income Neighborhoods," *Brookings-Wharton Papers on Urban Affairs,*.

Avery, Robert B., Kenneth P. Brevoort, and Glenn B. Canner (2007). "Opportunities and Issues in using HMDA Data," *Journal of Real Estate Research,* vol. 29, no. 4, pp. 351-380.

Avery, Robert B., Kenneth P. Brevoort, and Glenn B. Canner (2006). "Higher-Priced Home Lending and the 2005 HMDA Data," *Federal Reserve Bulletin,*pp. A123-A166.

Bhutta, Neil. 2008a. Giving Credit where Credit is Due? The Community Reinvestment Act and Mortgage Lending in Low Income Neighborhoods. Ph.D., MIT.

Bhutta, Neil. 2008b. The Effect of the Community Reinvestment Act on Mortgage Lending to Low Income Borrowers. Ph.D., MIT.

Bostic, Raphael W., and Stuart A. Gabriel (2006). "Do the GSEs Matter to Low-Income Housing Markets? an Assessment of the Effects of the GSE Loan Purchase Goals on California Housing Outcomes," *Journal of Urban Economics,* vol. 59, no. 3, pp. 458-475.

Congressional Budget Office (CBO) (May, 2001). "Federal Subsidies and the Housing GSEs," .

Garmaise, Mark J., and Tobias J. Moskowitz (April, 2006). "Bank Mergers and Crime: The Real and Social Effects of Credit Market Competition," *The Journal of Finance,* vol. LXI, no. 2, .

Haurin, Donald R., Robert D. Dietz, and Bruce A. Weinberg (2003). "The Impact of Neighborhood Homeownership Rates: A Review of the Theoretical and Empirical Literature," *Journal of Housing Research,* vol. 13, no. 2, .

Imbens, Guido W., and Thomas Lemieux (2008). "Regression Discontinuity Designs: A Guide to Practice," *Journal of Econometrics,* vol. 142, no. 2, pp. 615-635.

Kubrin, Charis E., and Gregory D. Squires "The Impact of Capital on Crime: Does Access to Home Mortgage Money Reduce Crime Rates?" In *Annual Meeting of the Urban Affairs Association.* Washington, DC.

Lang, William W., and Leonard I. Nakamura (1993). "A Model of Redlining," *Journal of Urban Economics,* vol. 33, pp. 223-234.

Li, Wenli (2005). "Moving Up: Trends in Homeownership and Mortgage Indebtedness," *Business Review,*pp. 26-34.

Manchester, Paul B. (August, 2008). "Overview of the GSEs' Housing Goal Performance, 2000-2006," *Housing Finance Working Paper Series*.

Manchester, Paul B. (May, 2002). "Goal Performance and Characteristics of Mortgages Purchased by Fannie Mae and Freddie Mac, 1998-2000," *Housing Finance Working Paper Series*.

Munnell, Alicia H., M. B. Tootell Geoffrey, Lynn E. Browne, and James McEneaney (1996). "Mortgage Lending in Boston: Interpreting HMDA Data," *American Economic Review,* vol. 86, no. 1, pp. 25-53.

Nichols, Joseph, Anthony Pennington-Cross, and Anthony Yezer (2005). "Borrower Self-Selection, Underwriting Costs, and Subprime Mortgage Credit Supply," *The Journal of Real Estate Finance and Economics,* vol. 30, no. 2, pp. 197-219.

Pennington-Cross, Anthony, Anthony Yezer, and Joseph Nichols (2000). "Credit Risk and Mortgage Lending: Who Uses Subprime and Why?" Research Institute for Housing America Working Paper No. 00-03.

Porter, Jack (2003). "Estimation in the Regression Discontinuity Model," Department of Economics, University of Wisconsin.

Quercia, Roberto G., George W. McCarthy, and Susan M. Wachter (2003). "The Impacts of Affordable Lending Efforts on Homeownership Rates," *Journal of Housing Economics,* vol. 12, no. 1, pp. 29-59.

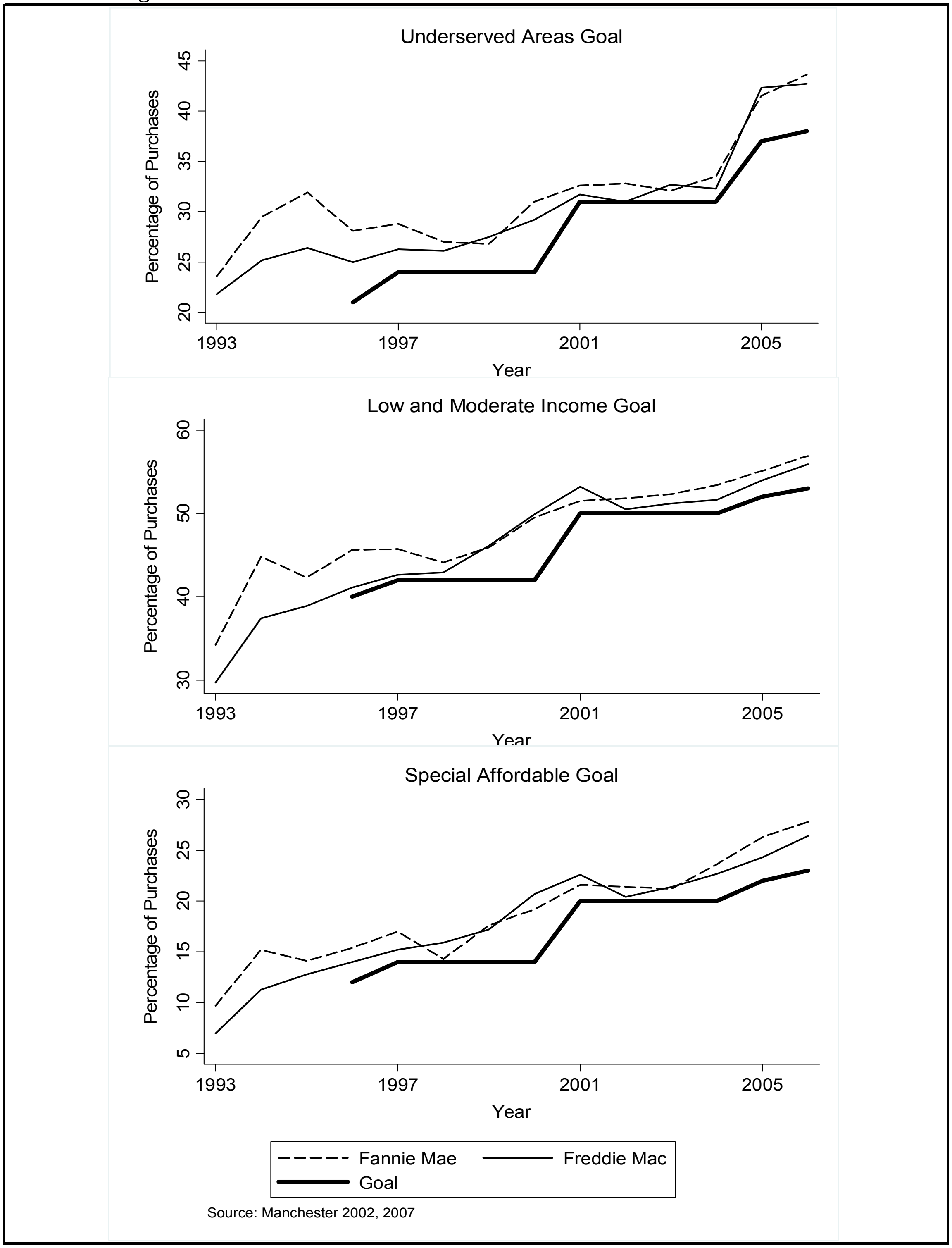
Quigley, John M. (2006). "Federal Credit and Insurance Programs: Housing," *Review,*pp. 281-310.

Scheessele, Randall M. (1998). "HMDA Coverage of the Mortgage Market," *Housing Finance Working Paper Series*.

Straka, John W. (2000). "A Shift in the Mortgage Landscape: The 1990's Move to Automated Credit Evaluations," *Journal of Housing Research,* vol. 11, no. 2, .

Temkin, Kenneth, Jennifer E. H. Johnson, and Diane Levy (March, 2002). "Subprime Markets, the Role of GSEs, and Risk-Based Pricing," .

Figure 1:
GSE Housing Goals and Performance

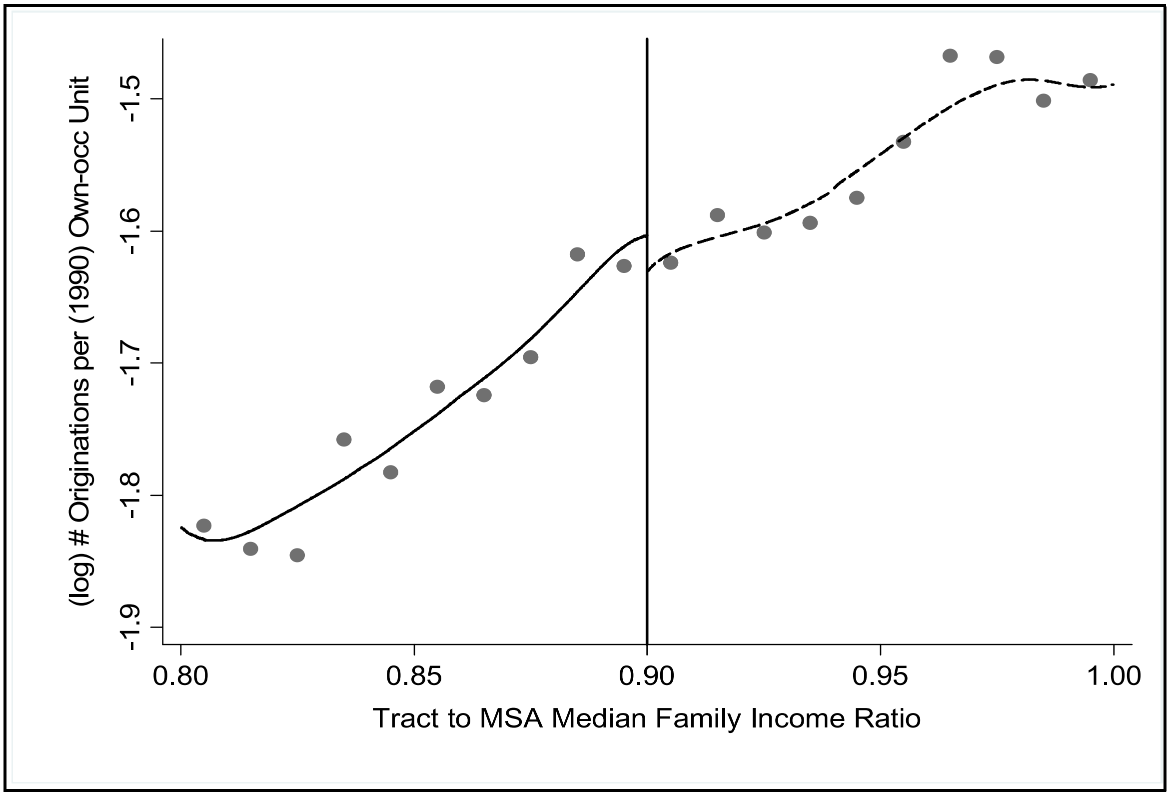


Notes: The "Underserved Areas" goal targets GSE purchases in low-income & minority neighborhoods: tracts with median family income less than or equal to 90% of the MSA median family income, or tracts with a minority population share of at least 30% and a median family income no greater than 120% of the MSA median family income. The "Low and Moderate Income" goal specifies a target share of purchases to borrowers with income below the MSA median family income. The "Special Affordable" goal targets borrowers with income below 60% of the MSA median family income and borrowers with income below 80% of the MSA median family income in a census tract that has a median family income less than or equal to 80% of the MSA median family income.

Figure 2:
Test of Identification Assumption - Mortgage Originations Predicted by 1990 Tract Characteristics

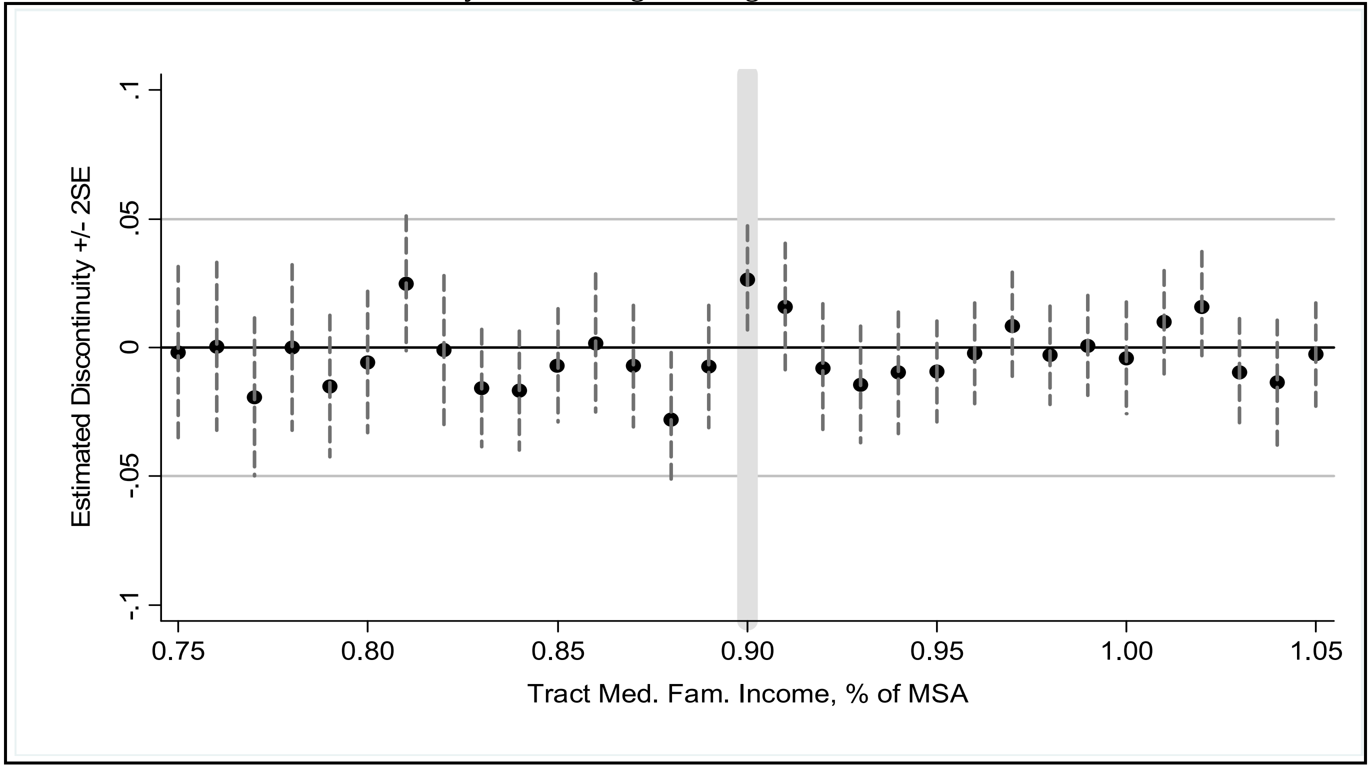
Notes: Y-axis values are predicted using coefficients on census tract characteristics (see Table 1) estimated from a tract-level regression of (log) number of refinance and home purchase originations on tract characteristics and MSA fixed-effects (estimated fixed-effects not used to generate predicted values). Each data point represents the mean of Y-axis variable within one percentage point bins of the X-axis variable. Also shown are local linear regression generated fits of the underlying predicted values, created seperately on either side of the cutoff.

Figure 3:
Discontinuity in (log) GSE-Purchased Originations in Census Tracts with *TM* ≤ 0.90, 1997-2002



Notes: Includes owner-occupied home purchase and refinance mortgages eligible (i.e. conforming, conventional loan originated by prime lender) and purchased by GSEs. Each data point represents the mean of (log) purchases in a tract between 1997 and 2002 for tracts with one percentage point bins of the X-axis variable. Also shown are local linear regression generated fitted lines of the underlying tract-level data, created seperately on either side of the cutoff.

Figure 4:
Falsification Exercise: Discontinuity in GSE-Eligible Originations at GSE and non-GSE Cutoffs



Notes: Each point represents the estimated discontinuity in (log) number of GSE-eligible originations in a tract between 1997 and 2002 from a seperate regression. Each regression includes tract covariates and scale variables and MSA-fixed effects, and bandwidth (*h*) is set to 0.02. Standard errors are clustered at MSA-level.

Table 1:
Primary Variables Available in HMDA Dataset, 1992-2005

Variable	Availability	Description
Year	all years	Year of mortgage application or purchase
Institution ID	all years	10 Character Lender Identifier
Regulatory Agency ID	all years	Code indicating OCC, Fed, FDIC, OTS, NCUA (credit unions) or HUD as supervisory agency
Loan Type	all years	Conventional or government insured (e.g. FHA, VA)
Loan Purpose	all years	Home purchase, refinance, home improvement or multifamily (i.e. 5+ family property)
Property Type	2004-2005	1-4 Family, manufactured housing or multifamily structure
Occupancy	all years	Owner-occupied or investment property/second home
Loan Amount	all years	Dollar amount of loan
Lien Status	2004-2006	Loan secured by first or subordinate lien
Rate Spread	2004-2006	APR spread above comparable Treasury, conditional on the spread exceeding 3 percentage points for first-lien mortgages and 5 percentage points for subordinate-lien loans
HOEPA Status	2004-2006	Indicator for high-cost loan (e.g. APR at consummation on first-lien mortgage exceeds yield for comparable Treasury by more than 8 percentage points.)
Action Taken	all years	Six possibilities: (1) Loan originated, (2) Borrower rejects lender offer (3) Application denied, (4) Application withdrawn by applicant (5) Application incomplete, (6) Loan purchased by the institution
Denial Reason (optional)	all years	Institution can provide primary reason(s) for denial (e.g. credit history, insufficient collateral, debt load, etc)
Geography	all years	State, county and census tract of property
Income	all years	Gross annual family income, rounded to the nearest thousand dollar
Applicant(s) Ethnicity	2004-2006	Indicator for being Hispanic/Latino; may not be provided if telephone/internet application. "Hispanic" is a choice under *Race* variable in prior years
Applicant(s) Race	all years	Race of primary applicant; race of co-applicant if applicable. May not be provided if telephone/internet application
Applicant(s) Sex	all years	Sex of primary applicant; sex of co-applicant if applicable. May not be provided if telephone/internet application
Purchaser	all years	For loans sold at time of origination, specifies purchaser of loan (e.g. Fannie Mae, commercial bank, etc.)

Table 2:
Census Tract Summary Statistics

	(1) All Tracts	(2) 85 ≤ *TM* ≤ 90	(3) 90 < *TM* ≤ 95	(4) p-Value[1]
# of Census Tracts[2]	42,381	2,728	2,798	-
		A. Loans per Tract per Year (1997-2002)		
# Total Originations[3]	190.0	162.4	175.3	< 0.01
	(217.6)	(142.1)	(154.8)	
# GSE-Eligible Originations[4]	126.9	104.4	115.3	< 0.01
	(154.4)	(95.5)	(106.3)	
Purchased by GSEs	55.3	43.3	48.5	< 0.01
	(74.7)	(45.6)	(51.1)	
Purchased by non-GSE	27.9	22.4	24.4	< 0.01
	(38.7)	(24.4)	(26.9)	
# GSE-Ineligible Originations[5]	63.0	58.0	60.0	0.21
	(79.5)	(57.9)	(62.5)	
Purchased by GSEs	2.8	1.6	1.8	< 0.01
	(5.9)	(2.6)	(3.3)	
Purchased by non-GSE	42.8	42.2	43.3	0.41
	(55.5)	(45.5)	(48.7)	
		B. Census Tract Characteritics (1990)		
# Owner-Occupied Units	1050.6	1044.5	1105.4	< 0.01
	(686.0)	(585.6)	(596.2)	
# Total Housing Units	1814.2	1856.0	1870.4	0.58
	(1018.8)	(966.2)	(981.3)	
Med. Home Value ($, 1990)	112,246.20	89,945.23	95,425.02	< 0.01
	(86,830.49)	(59,366.45)	(64,444.07)	
Prop. Units Detached	0.584	0.559	0.597	< 0.01
	(0.278)	(0.252)	(0.235)	
Prop. Units Mobile Home	0.047	0.077	0.069	< 0.05
	(0.100)	(0.128)	(0.116)	
Prop. Units Built 1980-1989	0.16	0.143	0.146	0.53
	(0.169)	(0.147)	(0.146)	
Prop. Units Built 1940-1969	0.426	0.442	0.447	0.34
	(0.226)	(0.217)	(0.221)	
Prop. Units Built pre-1940	0.204	0.217	0.200	< 0.01
	(0.225)	(0.220)	(0.213)	
Prop. Units in Multifamily Bldg	0.177	0.169	0.157	< 0.05
	(0.209)	(0.199)	(0.181)	
Prop. Population Age >65	0.128	0.139	0.138	0.58
	(0.072)	(0.072)	(0.073)	
Prop. Population Black	0.137	0.103	0.088	< 0.01
	(0.250)	(0.194)	(0.172)	
Prop. Population Hispanic	0.088	0.079	0.071	< 0.05
	(0.162)	(0.146)	(0.131)	
Prop. Population in Group Qtrs	0.016	0.015	0.015	0.67
	(0.036)	(0.032)	(0.034)	

Notes: Standard deviations in parentheses. (1) p-Value testing the difference between columns 2 and 3. (2) See text for discussion of sample selection. (3) Originations includes owner-occupied home purchase and refinance loans. (4) GSE-Eligible loans are conventional, conforming loans originated by lenders not classified as a subprime lender (as determined by HUD). (5) Loans not likely to be eligible for purchase by GSEs: loan amount above the conforming limit, FHA or VA loans, or loans originated by subprime lenders.

Table 3:
Discontinuity Estimates of the Effect of the GSE Act on GSE Purchases of GSE-Eligble Originations in Census Tracts with $TM \leq 0.90$, 1997-2002

Dependent variable: *(log) # GSE-eligible & purchased originations, 1997-2002*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	$h = 0.05$							$h = 0.02$			
$\mathbf{1}[TM \leq 0.90]$	-0.1045***	-0.0309**	0.0154	0.0226	0.0307*	0.0382	0.0370**	0.0180	0.0329**	0.0305	0.0376**
	(0.0160)	(0.0133)	(0.0279)	(0.0232)	(0.0166)	(0.0257)	(0.0181)	(0.0216)	(0.0154)	(0.0244)	(0.0166)
TM'			0.0159**	0.0095*	0.0071*	0.0113*	0.0064				
			(0.0065)	(0.0053)	(0.0043)	(0.0066)	(0.0053)				
$(TM')*\mathbf{1}[TM \leq 0.90]$			0.0163*	0.0029	-0.0020	0.0077	0.0028				
			(0.0097)	(0.0078)	(0.0059)	(0.0106)	(0.0075)				
(log) # GSE-eligible & purchased originations, 1994-96					0.6264***		0.6294***		0.6302***		0.6374***
					(0.0218)		(0.0238)		(0.0343)		(0.0343)
Kernel Weights	rect	rect	rect	rect	rect	tri	tri	rect	rect	tri	tri
Covariates		Y		Y	Y	Y	Y	Y	Y	Y	Y
R-Squared	0.766	0.839	0.768	0.839	0.910	0.841	0.913	0.851	0.919	0.859	0.926
N	5525	5525	5525	5525	5509	5515	5500	2208	2202	2203	2197

Notes: Standard errors, clustered at MSA-level, shown in parentheses. * p<0.10, ** p<0.05, *** p<0.01. All regressions include MSA fixed effects and two tract-level scale variables measured in 1990: (log) owner-occupied units and (log) total housing units. For list of covariates included, see Table 1. Dependent variable is the (log) number of GSE-eligible refinance and home purchase mortgages purchased by GSE's between 1997 and 2002 at the census tract level.

Table 4:
Discontinuity Estimates of the Effect of the GSE Act on Non-GSE Purchases of GSE-Eligble Originations in Census Tracts with $TM \leq 0.90$, 1997-2002

Dependent variable: *(log) # GSE-eligible purchased by non-GSEs, 1997-2002*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	$h = 0.05$				$h = 0.02$			
$\mathbf{1}[TM \leq 0.90]$	0.0100	0.0288	0.0066	0.0269	0.0099	0.0314**	0.0016	0.0270
	(0.0234)	(0.0184)	(0.0246)	(0.0188)	(0.0202)	(0.0157)	(0.0233)	(0.0178)
TM'	0.0054	0.0080*	0.0039	0.0067				
	(0.0053)	(0.0045)	(0.0065)	(0.0057)				
$(TM')*\mathbf{1}[TM \leq 0.90]$	0.0039	-0.0027	0.0033	-0.0024				
	(0.0079)	(0.0069)	(0.0101)	(0.0092)				
(log) # GSE-eligible originations purchased by Non-GSEs, 1994-96		0.4019***		0.4050***		0.3918***		0.4036***
		(0.0200)		(0.0215)		(0.0280)		(0.0304)
Kernel Weights	rect	rect	tri	tri	rect	rect	tri	tri
R-Squared	0.818	0.866	0.822	0.870	0.835	0.879	0.846	0.889
N	5521	5504	5511	5494	2207	2196	2202	2191

Notes: Standard errors, clustered at MSA-level, shown in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. All regressions include MSA fixed effects, covariates (see Table 1 for list) and two tract-level scale variables measured in 1990: (log) owner-occupied units and (log) total housing units. Dependent variable is the (log) number of GSE-eligible refinance and home purchase mortgages purchased by Non-GSE's between 1997 and 2002 at the census tract level.

Table 5:
Discontinuity Estimates of the Effect of the GSE Act on GSE-Eligble Originations in Census Tracts with $TM \leq 0.90$, 1997-2002

Dependent variable: *(log) # GSE- eligible originations, 1997-2002*

	(1)	(2)	(3)
1[$TM \leq 0.90$]	0.0211	0.0266**	0.0272**
	(0.0167)	(0.0104)	(0.0112)
TM'			
(TM')***1**[$TM \leq 0.90$]			
(log) # GSE-Eligible originations, 1994-96		0.7045***	0.7139***
		(0.0332)	(0.0328)
Kernel Weights	rect	rect	tri
R-Squared	0.863	0.934	0.939
N	2208	2207	2202

Notes: Standard errors, clustered at MSA-level, shown in parentheses. * p<0.10, ** p<0.05, *** p<0.01. All regressions include MSA fixed effects, covariates (see Table 1 for list) and two tract-level scale variables measured in 1990: (log) owner-occupied units and (log) total housing units. Dependent variable is the (log) number of GSE-eligible refinance and home purchase mortgages originated between 1997 and 2002 at the census tract level.

Table 6:
Discontinuity Estimates of the Effect of the GSE Act on GSE-Eligible Originations in Census Tracts with $TM \leq 0.90$,
Estimates by 2-Year Period ($h = 0.02$)

Dependent variable: *(log) # GSE-eligible originations, 1997-2002*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	1997-1998			1999-2000			2001-2002		
$\mathbf{1}[TM \leq 0.90]$	0.0077	0.0138	0.0109	0.0290*	0.0347***	0.0314**	0.0263	0.0312**	0.0345**
	(0.0183)	(0.0114)	(0.0124)	(0.0164)	(0.0114)	(0.0129)	(0.0183)	(0.0128)	(0.0143)
TM'									
$(TM')*\mathbf{1}[TM \leq 0.90]$									
(log) # GSE-Eligible originations, 1994-96		0.7826***	0.7963***		0.6785***	0.6877***		0.6793***	0.6869***
		(0.0340)	(0.0318)		(0.0350)	(0.0367)		(0.0357)	(0.0366)
Kernel Weights	rect	rect	tri	rect	rect	tri	rect	rect	tri
R-Squared	0.848	0.933	0.939	0.842	0.913	0.915	0.854	0.913	0.918
N	2208	2207	2202	2208	2207	2202	2208	2207	2202

Notes: Standard errors, clustered at MSA-level, shown in parentheses. * p<0.10, ** p<0.05, *** p<0.01. All regressions include MSA fixed effects, covariates (see Table 1 for list) and two tract-level scale variables measured in 1990: (log) owner-occupied units and (log) total housing units. Dependent variables the (log) number of GSE-eligible refinance and home purchase mortgages originated in each of the three periods specified above at the census tract level. For all regressions, bandwidth (h) is 0.02.

Table 7:
Discontinuity Estimates of the Effect of the GSE Act on GSE-Ineligble Originations in Census Tracts with $TM \leq 0.90$ ($h = 0.02$), 1997-2002

Dependent variable: *(log) # GSE-* ***in****eligible originations, 1997-2002*

	(1)	(2)	(3)
1[$TM \leq 0.90$]	0.0123	0.0134	0.0089
	(0.0198)	(0.0127)	(0.0123)
TM'			
(TM')***1**[$TM \leq 0.90$]			
(log) # GSE-Ineligible		0.5774***	0.5877***
originations, 1994-96		(0.0207)	(0.0215)
Kernel Weights	rect	rect	tri
R-Squared	0.804	0.911	0.915
N	2208	2206	2201

Notes: Standard errors, clustered at MSA-level, shown in parentheses. * p<0.10, ** p<0.05, *** p<0.01. All regressions include MSA fixed effects, covariates (see Table 1 for list) and two tract-level scale variables measured in 1990: (log) owner-occupied units and (log) total housing units. Dependent variable is the (log) number of refinance and home purchase mortgages originated not likely to be eligible for purchase by GSEs (i.e. loans above the conforming limit, FHA or VA loans, or loans originated by subprime lenders) at the census tract level. For all regressions, bandwidth (h) is 0.02.

Table 8:
Summary Statistics: Tracts that Switched to being Targeted in 2004 versus Tracts Remaining Untargeted

	Switchers[1]	Non-Switchers[2]	p-Value[3]
Number of Tracts	1,850	5,759	-
A. Tract Characteristics			
TM_{old}	0.969	1.009	< 0.01
	(0.049)	(0.055)	
TM_{new}	0.827	1.045	< 0.01
	(0.064)	(0.119)	
Owner-Occ Units, 1990	1015.1	1153.4	< 0.01
	(459.5)	(474.9)	
Owner-Occ Units, 2000	1080.9	1340.3	< 0.01
	(524.6)	(631.7)	
Total Housing Units, 1990	1773.5	1744.8	0.293
	(743.8)	(725.7)	
Total Housing Units, 2000	1883.3	1973.3	< 0.01
	(827.8)	(894.9)	
Minority Pop Share, 1990	0.105	0.077	< 0.01
	(0.082)	(0.074)	
Minority Pop Share, 2000	0.212	0.128	< 0.01
	(0.163)	(0.119)	
Med House Value, 1990 ($2007)	136,363.40	160,108.50	< 0.01
	(78,397.57)	(97,344.43)	
Med House Value, 2000 ($2007)	135,432.00	166,516.20	< 0.01
	(71,495.71)	(93,116.70)	
B. Mortgage Originations, 2005-2006			
GSE-Eligible Originations (per year)	105.4	144.3	< 0.01
	(84.6)	(133.4)	
Purchased by GSEs	28.5	40.6	< 0.01
	(22.4)	(37.3)	
GSE-Ineligible Originations[4]	76.7	77.6	0.783
	(98.3)	(88.2)	
Purchased by GSEs	1.9	2.0	0.254
	(2.1)	(1.9)	
C. Mortgage Originations, 2001-2002			
GSE-Eligible Originations	129.4	197.5	< 0.01
	(89.8)	(152.9)	
Purchased by GSEs	58.3	89.6	< 0.01
	(44.6)	(75.9)	
GSE-Ineligible Originations[5]	54.5	61.6	< 0.01
	(42.4)	(65.9)	
Purchased by GSEs	1.5	3.1	< 0.01
	(2.6)	(6.1)	

Notes: Standard deviations in parentheses. Sample tracts comprised of those in an MSA with TM_{old} between 0.90 and 1.10 and minority population share in 1990 less than 0.30. (1) Sample tracts with $TM_{new} \leq 0.90$. (2) Sample tracts with $TM_{new} > 0.90$. (3) p-Value from test of difference between columns 1 and 2; clustered at MSA level. (4) Ineligible mortgages are either above the single family conforming limit, FHA- or VA- insured, or identified as high-price in the HMDA. (5) Ineligible originations are owner-occupied home purchase and refinance loans either above the single family conforming limit, FHA- or VA- insured, or originated by a subprime specialist institution.

Table 9:
Changing Tract Status in 2004 and Growth in Credit Supply

	(1)	(2)	(3)	(4)	(5)
A. GSE-Purchases[1]					
1[*Switcher*]	0.0963***	0.0836***	0.0364**	0.0160	0.0188
	(0.0161)	(0.0141)	(0.0167)	(0.0215)	(0.0159)
TM'_{new}			-0.0030***	-0.0038	
			(0.0011)	(0.0024)	
TM'_{new} * **1**[*Switcher*]			-0.0002	-0.0036	
			(0.0019)	(0.0038)	
R-Squared	0.544	0.557	0.558	0.579	0.625
N	7597	6113	6113	3750	1980
B. GSE-Eligible Originations[2]					
1[*Switcher*]	0.0895***	0.0744***	0.0277**	0.0229	0.0139
	(0.0139)	(0.0119)	(0.0124)	(0.0168)	(0.0126)
TM'_{new}			-0.0026***	-0.0021	
			(0.0009)	(0.0020)	
TM'_{new} * **1**[*Switcher*]			-0.0011	-0.0028	
			(0.0017)	(0.0032)	
R-Squared	0.565	0.593	0.594	0.625	0.682
N	7603	6117	6117	3752	1981
B. GSE-Ineligible Originations[3]					
1[*Switcher*]	0.0582***	0.0433***	0.0222	0.0214	0.0113
	(0.0146)	(0.0133)	(0.0138)	(0.0175)	(0.0124)
TM'_{new}			-0.0014	-0.0034	
			(0.0010)	(0.0021)	
TM'_{new} * **1**[*Switcher*]			0.0001	0.0049	
			(0.0020)	(0.0036)	
R-Squared	0.356	0.387	0.387	0.415	0.489
N	7604	6118	6118	3753	1981
Bandwidth (h)[4]	-	0.20	0.20	0.10	0.05
Covariates[5]	no	no	no	no	yes

Notes: Standard errors, clustered at MSA-level, shown in parentheses. * p<0.10, ** p<0.05, *** p<0.01. All regressions include MSA-level fixed effect. Mortgages include those for home purchase and refinance of owner-occupied sites. (1) Dependent variable is the change in (log) number of GSE-eligible & purchased loans between 2001/02 and 2005/06. (2) Dependent variable is the change in (log) number of GSE-eligible loans between 2001/02 and 2005/06. (3) Dependent variable is the change in (log) number of GSE-ineligible loans between 2001/02 and 2005/06. (4) The bandwidth around TM_{new} = 0.90. (5) Includes (log) median home value in 2000, (log) number of owner-occupied units in 2000, (log) total number of housing units in 2000, minority population share in 2000 and TM_{old} .

Table 10:
Testing for Heterogeneous Effects Due to Safety and Soundness Considerations

	(1)	(2)	(3)
Outcome Variable:	Δ (log) GSE Purchases	Δ (log) Eligible Originations	Δ (log) Ineligible Originations
1[*Switcher*]	0.0617**	0.0270	0.0040
	(0.0295)	(0.0251)	(0.0262)
TM'_{old} * **1**[*Switcher*]	-0.0054*	-0.0016	0.0009
	(0.0031)	(0.0027)	(0.0030)
R-Squared	0.625	0.682	0.489
N	1980	1981	1981
Bandwidth (h)[1]	0.05	0.05	0.05
Covariates[2]	yes	yes	yes

Notes: Standard errors, clustered at MSA-level, shown in parentheses. * p<0.10, ** p<0.05, *** p<0.01. All regressions include MSA-level fixed effect. Mortgages include those for home purchase and refinance of owner-occupied sites. (1) The bandwidth around TM_{new} = 0.90. (2) Includes (log) median home value in 2000, (log) number of owner-occupied units in 2000, (log) total number of housing units in 2000, minority population share in 2000 and TM'_{old}.