Federal Mandates and Mortgage Supply: Regression Discontinuity Analyses of the Community Reinvestment and GSE Acts

by

Neil Pravin Bhutta

B.A., Economics and Physics, Emory University (2000)

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2008

Signature of Author ....
Department of Economics
29 July 2008

Certified by ....
Michael Greenstone
3M Professor of Environmental Economics
Thesis Supervisor

Certified by ....
David Autor
Professor of Economics
Thesis Supervisor

Accepted by ....
Peter Temin
Elisha Gray II Professor of Economics
Chairperson, Departmental Committee on Graduate Students

Federal Mandates and Mortgage Supply:
Regression Discontinuity Analyses of the Community Reinvestment and GSE Acts

By

Neil Pravin Bhutta

Submitted to the Department of Economics
on July 29, 2008 in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

ABSTRACT

In this dissertation, I provide evidence of the causal impact on mortgage supply of the Community Reinvestment Act (CRA) and the "Government-Sponsored Enterprises (GSE) Act", laws requiring banks and the GSEs (Fannie Mae and Freddie Mac), respectively, to help improve credit access for low-income households and neighborhoods. While financial markets evolved rapidly since the early 1990's, I use discontinuities in the laws' eligibility rules to identify their effects. To implement the analyses, I use a census of mortgage applications collected under the Home Mortgage Disclosure Act. Overall, these programs appear to have had limited impact.

I first analyze CRA's effect on mortgage lending in targeted neighborhoods: census tracts with a median family income (MFI) under 80% of MSA MFI. The regression discontinuity (RD) estimates suggest an overall credit supply shift of at least \$6 billion (\$2007) from 1994 and 2002 in targeted neighborhoods. In addition to CRA's direct effect on bank lending, I also find that unregulated institutions lend more in targeted tracts ("crowd-in"). Further analysis suggests that information spillovers from increased bank lending helps generate crowd-in.

In Chapter 2, I examine CRA's effect on home purchase mortgage lending to *households* with income under 80% of the MSA MFI. In both Chapters 1 and 2, I find CRA's impact is concentrated in the largest MSAs, where enforcement is most intense. The RD estimates indicate that CRA caused a 6% increase in large MSA bank home purchase lending at the cutoff. Unlike in Chapter 1, there is no theoretical basis for crowd-in and none is found. Nor do I find that banks crowd-out unregulated institutions.

Finally, I measure the impact of one of the three goals established under the GSE Act. Under this goal the GSEs target census tracts with MFI under 90% of MSA MFI. The RD estimates suggest this goal led to a 3-4% increase in GSE purchases, and increased GSE-eligible originations by 2-3% at the cutoff. Unlike previous research, I find no evidence that the GSEs crowd-out FHA and subprime loans. The results imply a lower bound of the goal's impact of \$2.4 billion between 1997 and 2002.

Thesis Supervisor: Michael Greenstone
Title: 3M Professor of Environmental Economics

Thesis Supervisor: David Autor
Title: Professor of Economics

Acknowledgements

This dissertation represents the end point of my graduate education at MIT. There are many to whom I owe thanks.

I am thankful for the generous support (financial and otherwise) of the MIT Department of Economics and its staff, the National Science Foundation, the George and Obie Shultz Fund, the Dewey Library and its staff (in particular Katherine McNeill-Harman), the MIT GIS Lab, and the NCAA.

With respect to this research, I am very grateful for guidance from Michael Greenstone, David Autor and Amy Finkelstein. Thanks also to James Berry, Tonja Bowen, Christine Brickman, Raymond Guiteras, Hui Shan, Christopher L. Smith, Bill Wheaton, seminar participants at MIT, the Board of Governors of the Federal Reserve, the Furman Center for Real Estate and Public Policy (NYU) and the Bureau of Labor Statistics for helpful comments.

I would like to express my sincerest gratitude to Michael Greenstone for inviting me to work with him, for his continued support and for pushing me farther than I thought I could go. I am also extremely thankful for the time and expertise provided to me by David Autor and Amy Finkelstein. I feel unbelievably fortunate to have had the opportunity to work with and learn from Michael, David, and Amy, as well as the other faculty and students in the Economics Department at MIT. Thanks, in particular, also to Christopher L. Smith for all his help, advice and insights (and candy!) during the past five years. He is a fantastic friend and colleague.

I am indebted to my undergraduate advisor at Emory University, Christopher Curran. I do not think I would have had the tremendous benefit of attending MIT were it not for his guidance and belief in me.

Special thanks to my parents, Pravin and Rajal Bhutta. Their love and support throughout my life is truly remarkable. Thanks also to my sister, Sonya Borrero, for encouraging and inspiring me.

Finally, I am most grateful for my wife-to-be, Christine Brickman. I hope I find it in myself to always show her the generosity, compassion and patience that she shows me.

Contents

Chapter 1

Giving Credit where Credit is Due? The Community Reinvestment Act and Mortgage Lending in Low Income Neighborhoods

1.1. Introduction

The U.S. government has long intervened in credit markets to improve credit access. Economic and sociological theories suggesting that credit access and homeownership generate important individual and social benefits such as improved self-esteem and child outcomes, reduced crime and increased voter turnout often motivate such policies.[1] Congress passed the Community Reinvestment Act (CRA) in 1977 to ensure that low-income, often predominantly minority, communities have adequate access to mortgage credit. CRA stipulates that federally insured banks have an affirmative obligation to supply credit throughout their local market. In practice, regulators evaluate banks' lending record in low-income neighborhoods and can penalize non-compliant banks.

CRA's impact on credit supply is cast in doubt, especially since it does not subsidize banks and does not set explicit lending goals, leaving the determination of bank compliance to individual regulators' discretion. But changes in CRA enforcement in the early and mid 1990's likely mounted pressure on banks to expand their CRA-qualified activity. Indeed, banks increasingly established flexible lending programs and entered into "CRA lending agreements" since the mid 1990's (see Figure 1; Avery et al 2000). At the same time, homeownership and household mortgage debt rose dramatically, drawing attention to CRA.[2] But CRA's impact has not been clearly differentiated from other major financial market changes, and Federal Reserve Chairman Ben Bernanke (2007) recently declared, "Distinguishing with certainty the effects that the CRA had on 'CRA-type' activity from the effects of simultaneous regulatory and market changes over this period has not been possible."

This paper utilizes a simple yet compelling strategy for identifying CRA's impact. I take advantage of a discontinuity in CRA's eligibility rule to identify its effect on credit flow in targeted neighborhoods. CRA targets census tracts with a median family income less than 80% of its MSA's median family income, hereafter "low and moderate income" (LMI) tracts. This rule provides the basis for a regression discontinuity (RD) design where the essential idea is that tracts just below and above the cutoff are identical except for CRA-eligibility and so a substantive difference in lending across the two groups can be attributed to CRA. Given the considerable transformation of financial markets over

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

[2] See Li (2005) for an overview of these trends and potential causes.

the past two decades, the importance of exploiting this discontinuity to identify CRA's effects cannot be overemphasized.

Policy makers have recently debated CRA fiercely (Golberg 2000, Chen 2004) and it will likely re-emerge as policy makers rethink mortgage market regulations and policies to avoid future crises (e.g. Bowyer 2008). This paper informs this debate not only by providing estimates of CRA's effect on bank (regulated) lending, but also its indirect effect on non-bank (unregulated) lending. While CRA may spark increased bank lending, crowd-out of non-banks could negate these benefits. On the other hand, non-bank lending could increase if targeted neighborhoods improve or information spillovers exist (e.g. Lang and Nakamura 1993), a frequently cited rationale for government intervention in credit markets (Lacker 1995).[3]

More broadly, studying CRA may yield insights into the costs and benefits of expanding credit supply. For example, CRA could provide an experimental framework to study the impact of credit access on crime (e.g. Garmaise and Moskowitz 2006). And measuring the performance and profitability of marginal CRA mortgages could improve our understanding of the sustainability of increased lending and homeownership, and as a result also provide evidence on the competitiveness of the banking industry, a key regulatory issue as this industry consolidates.

To implement the analysis, I use comprehensive mortgage application data collected under the Home Mortgage Disclosure Act (HMDA). For all MSA's between 1994 and 2002, I find that bank mortgage origination volume was 3% higher in LMI tracts at the cutoff. Further analysis reveals that CRA's impact is concentrated entirely in large MSA's, where enforcement is more intense. In this subsample, the discontinuity in bank lending grows from 4% between 1994 and 1996 to 8% between 1997 and 2002, consistent with a reform implemented by 1997 strengthening CRA.[4] Notably, this reform added incentives for banks to target LMI *households* and in a companion paper using a similar RD strategy (Bhutta 2008a) I find that CRA's effect on this margin is also concentrated in large MSA's.

For non-banks, I find evidence of "crowd-in". In large MSAs in post-reform years, I find non-bank lending increased by 3%. Further, the increase in credit by unregulated lenders is concentrated in tracts that have had relatively low previous home sales, while the discontinuity in bank lending occurs in both low and high sales tracts. These results are consistent with a model of information externalities inhibiting credit flow in thin markets (see Section 3).

It is important to note that these results are sensitive to controlling for MSA and tract size. However, conditional on MSA and tract size, other covariates that explain an additional 20-25% of the variation in loan volume are very well-balanced across the cutoff, making a strong case for a causal interpretation of the results. Also important to this interpretation, the RD strategy does not readily identify discontinuities at non-CRA points.

Finally, I take advantage of the *change* in LMI status of some tracts following the release of 2000 Census data by formulating a "two-dimensional" regression discontinuity design – the change in treatment status depends on two variables instead of one – to assess the effects of CRA in 2004 and 2005. While summary statistics reveal

[3] Caplin and Leahy (1998) develop a similar model in the context of retail development in New York City.
[4] Zinman (2002) uses this reform to help identify his study of CRA's effect on small business lending.

extraordinary growth in loan volume in "switching" tracts, indicative of the explosion in credit supply at this time, the RD strategy identifies a modest effect of CRA on bank lending of 4-5% to these tracts. As this finding is for a set of newly targeted census tracts, it strengthens the causal interpretation of the earlier results.

In the next section, I discuss CRA in more detail and review previous research trying to identify CRA's effect on credit supply. Section 3 describes CRA's potential effects, including a short discussion of the information externality model. Section 4 discusses the data and regression discontinuity strategy. Section 5 presents the results and discusses the "two-dimensional" regression discontinuity approach, and Section 6 concludes.

1.2. Background & Related Literature

Congress passed the CRA in response to claims that banks were continuing to irrationally redline low-income, urban neighborhoods.[5] To ensure that banks and thrifts (hereafter banks) supply credit in both high and low income neighborhoods within their operating market, regulators periodically inspect their lending records. Importantly, CRA does not cover non-deposit independent mortgage companies or credit unions. And, CRA does not automatically cover non-deposit mortgage subsidiaries of banks, but banks have the option of including their subsidiaries' lending in their CRA evaluation.

Examiners rate banks in each of their "assessment area(s)" – generally the MSAs and/or counties where they have branches – separately and then combine these ratings into an overall grade (e.g. outstanding, needs to improve, etc.).[6] For multi-MSA institutions, examiners pay close attention to the largest markets they serve (see discussion in Section 5.2). Regulators can then penalize banks earning a poor CRA rating by denying applications for bank mergers, opening and closing new branches and other activities.

Interest in CRA as a tool to help expand homeownership opportunities grew in the late 1980's and early 1990's. Lawmakers pushed regulators to improve enforcement, beginning with an 1989 amendment forcing regulators to publicly disclose CRA evaluations and culminating in a 1995 reform (phased in by mid-1997) that shifted examiners' focus towards banks' actual record of lending rather than procedural efforts to meet CRA, such as spending time meeting with community groups.[7, 8] A strengthened CRA combined with accelerating bank merger activity following bank deregulation may have made CRA more effective during the late 1990's.

Community organizations also play a role in enforcement. Regulators must solicit and weigh public comments on a bank before deciding on its merger application. Community groups can and often do use the HMDA data and other resources to build a

[5] FHA underwriting manuals through the 1950's explicitly warned lenders against lending in minority neighborhoods (see Jackson (1985). Gotham (2000) cites examples of continued discriminatory language in private underwriting and appraisal manuals through the 1970's.

[6] Banks define their assessment area for CRA evaluations.

[7] Banks with an exam scheduled between January 1, 1996 and June 30, 1997 had the option of being evaluated under the old CRA criteria. Most took that option suggesting the new standards were more rigorous (Thomas 1998).

[8] See Barr (2005) and Fishbein (1992) for enforcement and legislative history of CRA.

case against a bank.[9] In response, banks have entered into "CRA Agreements" with community groups where banks pledge resources to targeted neighborhoods. Consistent with CRA becoming more salient for banks during the 1990's, the value of new CRA agreements rose considerably at this time (Figure 1). And bankers interviewed by Harvard's Joint Center for Housing Studies (JCHS) said that many banks have taken considerable steps to increase CRA-qualified loans to avoid CRA-associated difficulties and bad press (Belsky et al 2000).

A major component of the CRA evaluation is a bank's volume of lending to LMI neighborhoods compared to that of its peers as well as to the bank's non-LMI lending volume. Regulators therefore give banks "CRA-credit" as a discontinuous function of tract income. Therefore, to identify CRA's effect, I will measure the jump in loan volume at the point where CRA-credit jumps.

This strategy differs from most other studies of CRA. A few studies test for bank-specific reactions to particular CRA incentives, with mixed results. Dahl et al (2002) show that changes in CRA-type lending is uncorrelated with CRA rating downgrades in the early 1990's. In contrast, Bostic and Robinson (2005) find that banks increase targeted lending during the years in which they have a CRA agreement in effect, and Bostic et al (2005) find that merging banks' CRA-qualified lending increases prior to acquisitions of other banks.

Other studies, as this one, try to estimate CRA's impact more broadly. JCHS researchers (2002) show that banks make a higher fraction of their loans to and are less likely to deny loan applications from CRA-targeted populations inside the bank's assessment area (i.e. where banks get "CRA credit") compared to outside those geographic boundaries. But because bank operations likely differ in areas where they have branches relative to where they do not – for instance, better knowledge of the local market and population which may allow more lending to low-income borrowers – this test lacks a causal interpretation.[10]

A few studies (Schill and Wachter 1994, Evanoff and Segal 1996, Bostic and Surette 2004) find that unregulated LMI loan growth exceeds that of regulated lenders, implying CRA's effect has been limited. But excessive unregulated loan growth suggests that the trend in unregulated lending is not likely to provide a valid counterfactual to test CRA's effect on bank behavior, perhaps because unregulated lending operates under a different business model. Another concern is that reductions in underreporting during the 1990's likely inflate loan growth for independent mortgage companies more so than for banks (JCHS 2002, Bhutta 2008a).

Similar to the current paper's strategy, Berry and Lee (2007) exploit the CRA discontinuity, but they test for a discontinuity in loan rejection rates and find no effect. This test assumes that banks respond to CRA by lowering credit standards, which may not be the case (see Section 3). Even if banks do lower credit standards, testing for a discontinuity in the denial rate will be biased if average applicant credit quality changes in response. For instance, more high-risk types might apply as the probability of

---

[9] For instance, protest of a merger application by WesBanco in 2001 resulted in a year long delay (NCRC 2002).

[10] They also exploit variation across MSA's in CRA Agreements. As the authors acknowledge, this strategy also does not have a causal interpretation since MSA's where banks sign into agreements are likely to be different from other MSA's and CRA agreements locations may be endogenous.

acceptance rises. Indeed, Canner et al (1999) shows that during the 1990's, the expansion of subprime credit led to an increase in credit supply and denial rates. Both in this paper and in Bhutta (2008a) mentioned earlier, I detect a jump in application volume at the CRA cutoff, calling into question Berry and Lee's assumptions. Aside from this concern, the current study is also more comprehensive in that I look at CRA's effect on several loan types and dollars lent, and evaluate CRA's indirect effects on non-banks.

1.3. Neighborhood Credit Supply and the CRA[11]

Two types of lenders generally supply mortgage credit: banks, which take deposits, and mortgage companies (non-banks) that get loanable funds from a secondary market of investors. Roughly, banks serve prime (low-risk) borrowers, while non-banks provide FHA (government insured) and subprime loans in addition to conventional prime loans (Nichols et al 2005). The availability of prime, FHA and subprime mortgages suggests that a variety of risk types can obtain credit (Pennington-Cross et al 2000). Still, a large fraction of applications are denied indicating that many may have difficulty obtaining a mortgage.

The effect of CRA on targeted-neighborhood credit supply depends on how banks respond to CRA. Banks may attract applications away from non-banks, for instance through increased neighborhood advertising or by providing incentives to mortgage brokers and real estate agents to recommend customers to them, leaving credit supply basically unchanged.

Alternatively, banks may offer credit at a lower price in targeted areas. Some crowd-out may occur, but an overall increase in credit flow would result as demand would rise in response and more borrowers would meet payment-to-income ratio requirements. This scenario would raise consumer welfare and potentially generate positive neighborhood externalities as borrowers would be less likely to default (e.g. Immergluck and Smith 2006).

Banks may also lower their credit standards, which could crowd-out non-banks' FHA and subprime lending. Otherwise, such marginal loans might have differing effects over time, at first increasing credit supply and homeownership but subsequently resulting in more defaults and perhaps destabilizing neighborhoods.

Other research indicates that banks often work with community groups and state agencies to help provide credit to marginal borrowers (see Avery et al 2000, Campen and Callahan 2001, Quercia et al 2001). For instance, banks may allow a lower credit score but would require the borrower to attend a third party's financial literacy course. Banks might also conduct more thorough credit checks, on their own or with the help of a third party, allowing them to extend more credit to marginal borrowers.

Increased credit supply due to CRA that generates more home sales and/or increased homeownership could result in observed increases in *non-bank* lending. For instance, as homeownership rises the demand for refinance and home improvement credit should rise – demand that would likely be met by both banks and non-banks.

Lang and Nakamura (L-N 1993) provide a more subtle reason for "crowd-in". They hypothesize that increased home sales generates *public* information about neighborhood home values that can increase appraisal precision and thus lead to greater equilibrium credit supply. More concretely, a lender's net value of a loan, V_t, at time t is

[11] This section draws partially on a discussion in Avery et al (2003)

a function of the loan size, L_t, appraisal uncertainty, σ_t, and a vector of other loan characteristics $\boldsymbol{\theta}_t$: $V_t = V(L_t, \sigma_t, \boldsymbol{\theta}_t)$. Lenders choose a maximum loan size, L_t^*, given σ_t and $\boldsymbol{\theta}_t$ to maximize value. The implicit derivative of the value-maximizing loan size with respect to appraisal uncertainty is $(-\partial V_t / \partial \sigma_t)/(\partial V_t / \partial L_t)$. This derivative is negative since both partial derivatives are negative: an increase in either loan size or appraisal uncertainty, all else equal, reduces lenders' valuation of a loan. Importantly, asset-value uncertainty negatively affects loan value even for risk neutral lenders because of the asymmetry in payoffs: there are no gains to lenders from low appraisals to offset the default losses from high appraisals. Thus, lenders reduce offered loan size to counterbalance higher appraisal uncertainty (or, more generally, restrict credit supply in some way). Finally, L-N argue that appraisal precision depends on information generated from previous home sales, but the non-excludability of this information (e.g. real estate transaction details are made public) diminishes lenders' incentive to lend, leading to suboptimal credit supply. For instance, a negative shock to neighborhood home sales in one period may propagate as lenders wait to profit from the information generated by other lenders' lending activity.

A few studies establish a positive relationship between home sales in one period and subsequent credit access, but none have a quasi-experimental setup.[12] This study may provide evidence of the L-N hypothesis by taking advantage of the shock to home sales induced by CRA, combined with data on historical tract-level sales volume. If the L-N hypothesis is true, then an increase in bank home purchase loans due to CRA should have a greater effect on *non-bank* home purchase lending in previously low-sales areas than in high-sales areas. More precisely, the discontinuity in non-bank lending at the CRA cutoff amongst low-sales tracts should exceed that for high-sales tracts.

1.4. Data & Empirical Strategy

1.4.1. Overview

I take advantage of a sharp discontinuity in the CRA eligibility rule to identify its impact on credit flow. Census tracts with a median family income less than 80% of MSA median family income qualify as "low-and-moderate" income (LMI) under CRA regulatory procedures and are targeted by banks. In the regression discontinuity (RD) analysis that follows, this income ratio (*TM*) is the "assignment" (or "running") variable. The impact of the CRA *at the cutoff* will be identified by measuring the jump in loan volume at $TM = 0.80$.

Importantly, the value of *TM* used in this paper is identical to that used by bank regulators. As such, the key right-hand-side variable in the regressions to follow is measured precisely. 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 the 1990 Census and 1993 MSA definitions (by OMB).[13] In 2003, 2000 Census data was used to calculate

---

[12] See Blackburn and Vermilyea (2006), Harrison (2001), Ling and Wachter (1998) and Calem (1996). Also, Calem and Wachter 1999, LaCour-Little and Malpezzi 2003, Capozza et al 2005 show a link between appraisal quality and subsequent default – the intervening mechanism in the L-N model.

[13] 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.

TM. *TM* changed again in 2004 as MSA definitions changed considerably (see Section 5.2). I perform two separate analyses – one for 1994 to 2002 and another for 2004-2005.

4.2. 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 a large number of independent mortgage companies previously exempt began reporting under HMDA.[14]

HMDA provides a unique lender ID that I use to separate loans by three lender types: banks, mortgage subsidiaries of banks and independent mortgage companies. Credit unions, which extend a very small share of mortgage credit, are included in this last group because they also fall outside the reach of CRA. I also use the census tract of the property, the loan amount, the disposition of the loan (e.g. approved, originated, denied, etc.), the loan purpose (e.g. refinance), and whether the loan was sold in the secondary market and to whom it was sold. Finally, lenders report some borrower characteristics such as race, gender, and income, which I use to measure changes in portfolio risk. See Table 1 for a full list and description of the variables available in HMDA.

I also use census tract-level characteristics from the 1990 and 2000 Censuses, compiled and distributed by Geolytics. I merge these characteristics to the HMDA loan data using the census tract codes to create a tract-level dataset, with yearly mortgage activity variables from HMDA summed up by tract for different periods.

The analysis focuses on census tracts in MSAs as HMDA data are unreliable in rural areas (Avery et al 2007). In terms of MSA coverage, I provide evidence in the follow-up paper using Census and CPS data that HMDA accounts for the vast majority of home purchase lending activity at least by 1999 and that coverage increased considerably since 1994. Further, although independent banks in MSAs with less than $30 million in assets are exempt from reporting, using FDIC Call Report and Summary of Deposit data I find that these small banks held only 1.6% of the value of all 1-4 family residential loans held by all banks with at least one branch in an MSA at the end of 1998.

In addition to rural tracts, I exclude census tracts in MSAs in Hawaii and Alaska. Of the remaining states, I use only census tracts that have been in an MSA since 1994 in order to maintain a constant set of geographies. Of tracts in eligible MSAs around the cutoff ($0.75 \leq TM \leq 0.85$), I use just over 96% in the analysis. I keep census tracts with at least 100 housing units, at least one "specified" owner-occupied unit, less than 30% of the population in group quarters, at least one application of any type in each year, and those with between 0.25 and 20 originations per (1990) owner-occupied unit from 1994 to 2002.

Table 2 provides means of tract-level mortgage activity and tract characteristics for three groups of tracts: those just below the cutoff (column 2), those just above the cutoff (column 3), and those with median family income close to that of the MSA (column 1). Panel A shows average mortgage volume by loan type and the three lender types between 1994 and 2002. Column 4 provides the p-Value for a test of the equality of the means in columns 2 and 3.

[14] See Avery et al (2007) for more details on HMDA coverage and reliability. Also see FFIEC website (www.ffiec.gov) for rules defining which lending institutions are required to report under HMDA.

Table 2 makes clear that loan volume is highly correlated with tract relative income. Summing all loans by all lenders, total loan volume for median income tracts is around 30% higher than for the CRA-ineligible group and more than 60% higher than the CRA-eligible group. The differences in terms of owner-occupied housing units are 18% and 39%, respectively. In fact, substantive differences across the cutoff occur for nearly all the housing and demographic variables in panel B. The regression discontinuity strategy discussed next provides a reliable way in theory to deal with the apparent selection problem when trying to identify the impact of CRA.

1.4.3. Regression Discontinuity

Consider the following tract-level regression of potential outcomes (e.g. log total originations by banks between 1994 and 2002) on a CRA treatment indicator variable, $D_i = \mathbf{1}[TM_i \leq 0.80]$:

$$Y_i = \alpha + \beta D_i + e_i \tag{1.1}$$

The essential premise of the regression discontinuity design is that:

$$\lim_{h \to 0} \{ \mathrm{E}[e_i \mid 0.80 - h < TM_i < 0.80] - \mathrm{E}[e_i \mid 0.80 \leq TM_i \leq 0.80 + h] \} = 0 \tag{1.2}$$

In words, CRA targeted and not-targeted tracts arbitrarily close to the cutoff ($TM = 0.80$) 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 CRA treatment effect. Crucially, under assumption (1.2) the major changes in mortgage markets over the study period and the substantial differences between LMI and non-LMI census tracts do not threaten identification of CRA's effect at the tract-eligibility threshold.

One approach to estimating β is to compare the outcome mean for tracts "just below" the cutoff to that for tracts "just above" the cutoff. But because loan volume is highly correlated with the assignment variable, as illustrated in Table 2, a more attractive strategy is to model the underlying relationship between loan volume and TM in the vicinity of the cutoff, and β will be measured as the difference in the right and left limits of this function at the cutoff (Imbens and Lemieux 2007).

To see this, rewrite (1.1) as

$$\mathrm{E}[Y_i \mid D_i, TM_i] = \beta D_i + \mathrm{E}[e_i \mid D_i, TM_i] \tag{1.3}$$

And since D is determined entirely by TM (1.3) becomes

$$\mathrm{E}[Y_i \mid TM_i] = \beta D_i + \mathrm{E}[e_i \mid TM_i] \tag{1.4}$$

And finally, the RD estimating equation is given by

$$Y_i = \beta D_i + \mathrm{E}[e_i \mid TM_i] + \mu_i \tag{1.5}$$

where $\mu_i = Y_i - \mathrm{E}[Y_i \mid TM_i]$. $\mathrm{E}[e_i \mid TM_i]$ is the "control function" – it controls for the expected value of all excluded variables that affect y and are correlated with D. It will be approximated by linear and polynomial functions of TM.

1.5. Results

1.5.1. Testing the Identification Assumption

If observable, "pre-treatment" tract characteristics change smoothly across the cutoff, that would lend credence to the identification assumption that all tract characteristics affecting lending, except for CRA-eligibility status, change smoothly across the cutoff. Figure 2 displays a test of the identification assumption. It plots the predicted values from a regression of the (log) total number of originations (all lenders and loan types) between 1994 and 2002 for tracts with *TM* between 0.75 and 0.85 on the set of 1990 tract characteristics listed in Panel B of Table 2.[15] Each data point shown in Figure 2 represents the mean of the predicted values for tracts in a half-percentage point interval of *TM*. Also shown is an estimated regression line fit to the underlying tract-level predicted values that allows for an intercept shift at the cutoff. Importantly, the set of tract characteristics used explains more than 75% of the variation in tract loan volume and so this test is quite informative about loan volume around the cutoff in the absence of CRA.

A few features stand out. One is the steep slope of the data, reflecting the quick change in tract characteristics over *TM* around the cutoff. Second, the pattern of the data suggests that the *ex-ante* relationship between loan volume and *TM* around the cutoff is approximately linear. Finally, the data do not indicate any sharp change in tract characteristics at the cutoff.[16] There does appear to be considerable variation in tract characteristics just to the left of the cutoff and near $TM = 0.75$ that may make estimating CRA's effect more difficult.

1.5.2. RD Estimates of CRA's Effect on Bank Lending

Figure 3, which is analogous to Figure 2 except that the Y-axis variable is (log) total bank originations, mirrors the patterns observed in Figure 2. However, the data to the left of the cutoff have shifted up relative to the data for the control tracts when compared with Figure 2, suggesting a positive CRA effect. For instance, the first point to the left of the cutoff is well above the points just to the right, unlike in Figure 2. And the next two points, while being considerably lower, as they are in Figure 2, appear to lie along the dashed trend line whereas they fall below the dashed line in Figure 2.

The top left corner of Table 3 shows the regression result corresponding to Figure 3, and indicates a discontinuity estimate (standard error) in bank loan volume of nearly 7% (5%). The variance estimates throughout the paper allow for within-MSA spatial correlation. Column 2 shows estimates after including MSA fixed effects. The R^2 increases substantially and the standard error falls by about 25%, but the point estimate is also cut nearly in half, implying that the MSA composition of tracts is not well-balanced across the cutoff. Columns 3 and 4 add in tract-level controls, helping bring down the

---

[15] I log-transform the number of owner-occupied units, total number of housing units and median home value before entering them into the regression.

[16] The point estimate of the discontinuity in predicted loan volume is nearly 2% and is not statistically significant.

standard error substantially. At the same time, the point estimates across these columns are stable suggesting that *within MSA* tracts are comparable across the cutoff.

Columns 5 thru 7 use larger bandwidths to help estimate the control function more precisely. The power gained by using more data will be offset to some degree by using a higher order polynomial control function. In addition, to account for a potential discontinuity in loan volume due to the GSE Act, I allow for an intercept shift at $TM = 0.90$.[17] The estimates in Columns 5 thru 7 are relatively precise – all are significant at the 5% level. The point estimates range from 2.7% (Column 6) to 3.8% (Column 3). This stability suggests that the conditional relationship between loan volume and *TM* around the cutoff is well-behaved.

Panels B, C and D show RD estimates separately for census tracts grouped by three MSA size classes. Although CRA applies to all MSAs, enforcement is likely to be greater in large cities for several reasons. First, while large, multi-MSA banks dominate the credit and banking markets in cities of all sizes[18], regulators tend to focus on these banks' activity in the largest cities to help conserve regulatory resources.[19]

In addition, community groups that may challenge bank applications and negotiate CRA agreements are active mainly in large cities. Indeed, an historical list of CRA agreements provided by the National Community Reinvestment Coalition (NCRC 2005) suggests that large cities such as Boston, Washington DC, New York City, Detroit, and Philadelphia have commanded a large proportion of bank commitments.[20]

Finally, the higher probability of public scrutiny in large cities may affect the response to CRA of relatively small banks operating in large cities. First, because smaller banks generally operate in a single market, negative publicity in their sole market could be quite harmful. And secondly, smaller banks may hope to position themselves as acquisition targets[21], and they might be relatively more attractive if a potential buyer can be sure regulators will not place costly delays or conditions on the application because of negative public comments.

Panels B and C show results for census tracts in small (1990 population less than 500,000) and medium (1990 population between 500,000 and 2 million) MSAs, respectively. After controlling for both MSA and tract size (log number of housing units in 1990) the point estimates in Column 3 show no discontinuity in loan volume, although the standard errors are large (5.6% and 4.1%). Including all the other covariates does not affect the point estimates substantively, and subsequently increasing the bandwidth again

[17] The GSE Act of 1992 mandates that Fannie Mae and Freddie Mac make a certain proportion of their mortgage purchases of loans in neighborhoods with $TM \leq 0.90$. See Bhutta (2008b) for an evaluation of its impact.

[18] For instance, I find that in 1997, banks with more than $5 billion in assets (less than 3% of banks) originated nearly 40% of all HMDA-reported single family loans and controlled close to 45% of deposits in MSAs with a (1990) population of less than 1 million. At the same time, the median bank in this size class had branches in 7 different MSAs and had more than 70% of its total deposits coming from MSAs with more than 1 million people.

[19] See Goldberg (2000). The guidelines for evaluating multi-MSA banks instruct regulators to consider the size of an MSA when choosing which of a bank's set of markets to do a "full-scope" review (www.ffiec.gov).

[20] Also see Shwartz (1998)

[21] One reason banks may hope to be acquired is because of possible increases in shareholder wealth due to the acquisition (Cornett and De 1991, Carow and Kane 2001)

increases precision. Overall, there is little evidence of a CRA effect on bank lending in both small and medium size MSAs.

Panel D indicates that CRA's effect is concentrated in the largest MSAs – those with at least 2 million people (in 1990), of which there are 23 (out of 321 MSAs in the sample) accounting for nearly 40% of census tracts. Columns 3 thru 7 indicates a discontinuity in bank lending of more than 7% with last three estimates have p-Values under 0.01. However, the estimates are again sensitive to MSA and tract size controls, which is troubling in terms of satisfying the RD identifying assumption. But the robustness of the point estimate to additional controls that explain a considerable amount loan volume variation (e.g. the R^2 increases by more than 25% from Columns 3 to 4 in Panel D) makes a strong case for a causal interpretation of the estimates in Columns 3 thru 7.

Figure 4 provides results of a falsification exercise where I test for discontinuities in large MSA bank lending at non-CRA cutoffs – i.e. points where there should not be a discontinuity.[22] Each point in Figure 4 represents the estimated discontinuity at the given value of *TM* using the Column 5 specification of Table 3. Nearly all of the estimates away from the CRA cutoff fall close to the zero line, indicating that discontinuities in the data are not easily generated. These results also reinforce the causal interpretation of the discontinuity at 0.80.

A discontinuity does appear at *TM* = 0.75. In testing more than twenty points, it is not surprising to find one statistically significant discontinuity. Probing deeper, I find that pre-existing tract characteristics change rapidly around 0.75. Performing an analysis similar to that in Figure 2, I find a discontinuity in *predicted* loan volume at *TM* = 0.75, conditional on tract size and MSA, of -8.4% (standard error of 4.3%). In contrast, this same exercise yields no discontinuity at the CRA cutoff (point estimate of -0.7%).

Table 4 shows separate estimates (using the Column 6 specification in Table 3) of CRA's effect on different loan types – home purchase loans and refinance/home improvement loans – and for two different periods that represent years before and after full implementation of the 1995 CRA reform. As in Table 3, Panel A shows estimates for loan originations. The point estimate for all loans combined (Columns 1 and 2) is about twice as large in the second period, consistent with the reform strengthening CRA's effect.[23] The results also show a relatively large effect after 1997 in both home purchase lending (8.5%) and refinance/home improvement lending (8.1%).

Panel B shows estimates for CRA's effect on the number of loan applications to banks. As argued earlier, applications may react to CRA if banks, for instance, increase their advertising in targeted neighborhoods or borrowers perceive their chances of obtaining a loan to be greater. Indeed, Table 4 indicates that applications for both loan types do rise in response to CRA. Finally, Panel C shows estimates for the discontinuity in dollars lent. While the point estimates in Panels A and C suggest that home purchase lending rises by a similar proportion in both number and value, the point estimate at the bottom of Column 4 is somewhat lower than the rise in refinance/home improvement originations (0.081 versus 0.057). This result coincides with anecdotal evidence that

[22] One exception is that the GSE Act may induce a discontinuity in lending at *TM* = 90.
[23] The standard error of the difference is about 0.027 – not quite statistically significant.

community groups push banks to increase the supply of small loans to help provide liquidity to homeowners (NCRC 2005).[24]

1.5.3. Risk Characteristics of Marginal Loans

An ideal way to discern the risk of marginal loans would be to look at how the distribution of borrower credit scores changes across the CRA cutoff. Unfortunately, borrower credit score is not available in HMDA. Instead, I create a risk measure by using the available loan and borrower characteristics along with application outcomes (e.g. application denied by lender) to predict each application's probability of being denied given those characteristics. Of course, predicted denial probabilities will reflect underlying risk as well as other factors. For instance, high denial rates for minorities could reflect both lower average credit scores as well as lender discrimination.

To estimate the risk of an application, I first run the following regression:

$$deny_{jMt} = \alpha + \mathbf{x}_{jMt}\boldsymbol{\beta}_{Mt} + \lambda_{Mt} + \varepsilon_{jMt} \tag{5.1}$$

where $deny_{jMt}$ is an indicator variable equal to one if application j in MSA M and year t was denied, $\mathbf{x}$ is a vector of loan and borrower characteristics, λ is a set of MSA by year fixed effects and the coefficients on $\mathbf{x}$ are allowed to vary by MSA and year. (5.1) is run using home purchase applications at banks between 1997 and 2002 with TM between 0.50 and 1.10 and for which a credit decision was made.[25] In this sample, about 18% of applications were denied.

Using the estimated coefficients I generate predicted values and then I identify each application's quintile within its MSA-by-year (predicted) distribution. The predicted values are informative: applications predicted to be in the lowest risk (bottom quintile) have a 15% denial rate while applications predicted to be the highest risk (top quintile) have a 50% denial rate.

The gray bars in Figure 5 show the share of home purchase originations to each of the five risk groups as well as to investors (non-owner-occupiers) in tracts just above the CRA cutoff ($0.80 \leq TM < 0.82$). The modal owner-occupying borrower in these tracts is in the middle risk quintile.[26] The black bars provide estimates of the effect of CRA on the risk distribution from discontinuity regressions as in Table 4, except the outcome variable is the share of home purchase originations to a particular risk group (i.e. I run six regressions). The difference between the black and grey bars represents the estimated discontinuity in loan share.

[24] Small home improvement loans may relatively expensive because of fixed costs associated with any loan contract (see Cleary and Zimmerman 2006). Martin-Guerrero (2002) finds that a large proportion home improvement projects are funded by savings, tax refunds and/or family gifts, consistent with home improvement credit being costly.

[25] I don't include applications withdrawn before the lender's decision was made. I also drop farm loans and loans with missing or zero loan amounts. I use only home purchase applications because borrower characteristics are more reliably reported compared to refinance and home improvement loans. The variables in $\mathbf{x}$ are a set of dummy variables for borrower income (12 groups), amount-to-income ratio (10 groups), race (5 groups) and applicant/co-applicant gender (5 groups; e.g. male/female, male only, etc.).

[26] The distribution in relatively high-income tracts is skewed toward low-risk borrowers and vice versa (not shown).

Figure 5 provides very little evidence for a increase in portfolio risk at the cutoff. Without a better set of risk indicators (e.g. credit score, employment history, etc.) it may be difficult to detect more subtle changes in risk across the cutoff. Another point to keep in mind is that while marginal CRA loans may be going to typical borrowers, the typical bank borrower in tracts near the cutoff is relatively low-income – the median borrower's income is about 85% of the MSA median family income – and has a relatively high likelihood of denial – the denial rate around the cutoff is just over 20%.

1.5.3. RD Estimates of CRA's Effect on Non-Bank Lending

The results thus far show that banks have expanded their credit supply – both for home purchase and refinance/home improvement – in CRA-targeted areas. Also, they are getting more applications but do not appear to take on more risk. As mentioned earlier, these results may reflect banks attracting applications away from non-banks. Crowd-out seems especially likely in the case of increased home purchase lending since an underlying good (i.e. a house) must be supplied. I next test for whether bank lending substitutes for or compliments non-bank lending.

Table 5 provides estimates of CRA's effect for independent mortgage companies (IMC) and mortgage bank subsidiaries (BMC) separately. The positive discontinuity in BMC lending (Columns 3 and 4) is difficult to interpret because CRA may affect these lenders directly (see Sections 2). Nevertheless, this result rejects the concern some have raised of banks obtaining CRA loans from their subsidiaries to improve their CRA grade (e.g. Marsico 2006).

The results for IMCs (Columns 1 and 2) suggest that CRA generated an equilibrium with greater lending. In post-reform years, IMC lending is 3% higher at the cutoff. Importantly, there is no pre-existing discontinuity in non-bank lending – the IMC point estimate for 1994-1996 is 0.006. Also, the modest discontinuity in originations after 1997 does not appear to be driven by any single year. When I divide the 1997-2002 period into two sub-periods (1997-1999 and 2000-2002), I find a discontinuity in originations of similar magnitude in both sub-periods (the point estimates are 0.032 and 0.033, respectively; both are significant at the 10% level).

As discussed in Section 3, if the L-N hypothesis is true, then the increase in bank home purchase loans due to CRA would have a greater subsequent effect on *non-bank* lending in areas with relatively few past home sales. To test this hypothesis, I divide large MSA census tracts into two groups based on tract home purchase loan volume by all lenders in 1993 and 1994 per (1990) owner-occupied unit. I define a "low-sales" (high-sales) group of tracts as those with a home purchase loan rate below (above) the median in their MSA-by-income cell, where tract income takes three values: low ($50<TM<70$), medium ($70\leq TM<90$) or high ($90\leq TM<110$).

Columns 1 and 2 of Table 6 show estimates of the discontinuity in bank home purchase lending in low-sales and high-sales tracts. For both groups of tracts there is a similar (proportional) effect. Analogous estimates for IMCs (Columns 3 and 4) show a large discontinuity in home purchase lending in low-sales tracts, but no effect in high-sales tracts. These results are consistent with predictions of the L-N model.

1.5.4. Exploring Alternative Explanations for Crowd-In

Low-sales and high-sales tracts are different, which may lead to differential effects of CRA for reasons other than an information externality. For instance, the population of low-sales tracts around the cutoff ($0.78 \leq TM \leq 0.82$) is 24% African-American, compared to 13% for high-sales tracts. If minority neighborhoods are better represented by community groups, CRA's direct effect on banks may be stronger in the low-sales group. But Table 6 shows that the discontinuity in bank home purchase lending is similar for both groups, and I also find a similar effect across both groups for home-improvement/refinance lending.[27]

Another possibility is a differential effect in bank purchases of loans in the secondary market, for which banks get "CRA-credit" if the purchase is for a loan originated in an LMI tract. Interestingly, I find a large discontinuity in bank secondary market purchases – for 1997 and 2002 in large MSAs, the point estimate is 0.125 with a standard error of 0.020.[28] However, I do not find a difference in that effect across high and low sales tracts. The difference in the point estimates across the two groups is 0.005.

Along these same lines, I test whether non-banks are more likely to sell their loans to banks in low-sales LMI neighborhoods.[29] Specifically, I test for a discontinuity in non-banks' share of home purchase loans in low-sales tracts sold to banks. While the point estimate is positive, it is very small – 0.0035 with a standard error of 0.0024. Likewise, the discontinuity in home purchase loans sold to banks as a share of all such loans sold is similar (0.0053).

A final reason that I explore for increased non-bank lending in CRA-targeted tracts stems from the GSE Act, which establishes funding goals for secondary market institutions Freddie Mac and Fannie Mae (the GSEs). The "Special Affordable Goal" (SAG) in particular targets census tracts with $TM < 0.80$, similar to the CRA. However, the SAG also requires that the purchased loan be to a borrower with income less than 80% of the MSA median family income to count towards the GSEs' goal.

A few considerations suggest that the SAG will not explain the results thus far. First, I know of no reason that the GSE Act should have a differential effect in large MSAs. And second, direct tests suggest that the GSE Act's effect on credit supply is small (Bhutta 2008b). Nevertheless, I test for a confounding GSE effect by measuring the discontinuity in non-bank home purchase lending in low-sales tracts to borrowers with income *above* the SAG limit. The point estimate (standard error) of 0.103 (0.039) is nearly the same as that in Table 6 (Column 3), suggesting that the SAG is not driving the earlier results.[30]

1.5.5. The Effect of CRA on Bank and non-Bank Lending, 2004-2005

In addition to cross-sectional variation in treatment status, *TM* also varies over time. After each Census, CRA regulators update each tract's assignment variable value (I will refer to the new value as TM_{new} and the old value as TM_{old}). This change generated a

[27] Point estimates (standard errors) for the discontinuity in bank home-improvement/refinance lending are 0.077 (0.039) for low-sales tracts and 0.078 (0.021) for high-sales tracts.

[28] Again, I use the Column 6 specification from Table 3 and use the log number of bank loan purchases as the outcome variable.

[29] Lenders do not report information on the lender from whom they purchased a loan in HMDA, but lenders do report information on the type of institution to whom they sell a loan (see Table 1).

[30] Similar analyses suggest the GSE Act does not affect bank lending either.

set of tracts targeted by CRA beginning in 2004[31] that were not targeted in prior years (i.e. $TM_{old} \geq 0.80$ and $TM_{new} < 0.80$) that can be used to estimate CRA's effect more recently.[32] For a sample of tracts not targeted prior to 2004, consider the following regression model:

$$(5.2) \qquad \Delta Y_i = \alpha + \beta \Delta D_i + \Delta e_i$$

where Δ represents the change across 2004. Equation (5.2) is not identified if unobserved factors 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, this situation is similar to the previous case, except that the regressor of interest (ΔD) is now a deterministic function of two variables instead of just one: TM_{new} and TM_{old}. This observation leads to an estimating equation analogous to (1.5) above:

$$(5.3) \qquad \Delta Y_i = \alpha + \beta \Delta D_i + \mathrm{E}[\Delta e_i \mid TM_{i,new}, TM_{i,old}] + \eta_i$$

where $\Delta D_i = \mathbf{1}[TM_{i,new} < 80]$. To implement (5.3) I substitute for the third term on the right-hand-side a flexible function of TM_{old} and TM_{new}:

$$(5.4) \qquad TM_{old} + \sum_k (TM_{new}^k + TM_{old} * TM_{new}^k) * \mathbf{1}(k > 0)$$

where k indexes the polynomial order. The idea of the identification strategy is to compare the change in lending for tracts that *just switched* to LMI to those that *almost switched* into the treatment group – in other words, a difference-in-difference estimate at the cutoff. In order to use more data away from the cutoff, I control for the relationship between the assignment variable*s* and loan growth using (5.4).

Table 7 provides group means of various housing and credit flow variables for tracts that switched status in 2004 from non-LMI to LMI ("switchers") versus tracts that remained non-LMI ("non-switchers"). I limit the sample to tracts with TM_{old} between 80 and 90 because the vast majority of switchers come from this group. While TM_{old} is similar across the two groups (first row, panel A), their TM_{new} values indicate two groups are trending in opposite directions. Other characteristics in panels A and B provide further evidence of these divergent trends.[33]

The difference between the two numbers in the first row of Panel C provides a basic difference-in-difference (DD) estimate of CRA's effect – a remarkable 15% despite

---

[31] Tracts' median family income based on the 2000 Census was first used in 2003 for CRA enforcement. 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 largely temporary, while *TM* values assigned in 2004 were essentially permanent.

[32] A basic RD test as described earlier will not work after 2004 because of potentially confounding effects of earlier CRA enforcement. For instance, control tracts in 2004 may have been treated between 1994 and 2002.

[33] I link 2000 census tracts to 1990 census tracts using the Census Tract Relationship File and keep the set of tracts with minimal boundary changes.

the fact that switching tracts are getting poorer. IMC mortgage growth in switching tracts also exceeds that in non-switching tracts, again implying crowd-in, assuming the DD estimate is well-identified.

Table 8 shows estimates of (5.3). Column 1 provides a DD estimate controlling for MSA and TM_{old}, which lowers the coefficients relative to the baseline DD estimates calculated from Table 7. Columns 2 thru 6 adopt progressively more flexible control function specifications and include covariates in columns 4 and 6 (see table notes for list of covariates). Controlling for TM_{new} and its interaction with TM_{old} in column 2 reduces the point estimates substantially.

The estimates in columns 2-6 are stable over the various specifications and suggest CRA had a marginal effect of nearly 5% on bank lending. As this result is for a new set of targeted tracts, it bolsters the causal interpretation of the earlier results.

In contrast, Panel B shows that CRA did not affect IMC lending this period. The absence of crowd-in might be explained by the fact that these tracts have just begun being targeted by CRA. Earlier, it was found that CRA did not have an effect on non-bank lending until after 1996. Also, the magnitude of the CRA effect for banks is similar to what it was in the 1994-1996 period (Table 4), and in these years there is no evidence of positive spillovers to IMCs (Table 5).

1.6. Summary & Discussion

The CRA is a longstanding and large scale government credit market intervention to expand credit market access in low and moderate income (LMI) communities. Much controversy surrounds this regulation because its impact is not well understood. This paper helps inform this debate by exploiting the fact that CRA targets census tracts below a known income cutoff to estimate CRA's causal effect on mortgage credit flow to LMI neighborhoods.

The results indicate that CRA's marginal effect on bank lending is about 3% on average across all MSAs between 1994 and 2002, but the effect appears to be entirely concentrated in large MSAs where enforcement is likely to be strongest. Separate discontinuity estimates for small and medium MSAs suggest no impact of CRA, in contrast to a 7% discontinuity in large MSAs. I also find, using a modified RD design, a 4-5% increase in bank lending between 2004 and 2005 in large MSA tracts newly targeted due to the release of 2000 Census data.

Focusing on large MSAs, there is no evidence that increased bank lending crowds-out unregulated or partially regulated lenders. On the contrary, I find small increases at the cutoff in lending by banks' mortgage subsidiaries, which may be due directly to CRA, and by unregulated independent mortgage companies. Consistent with a theory of information externalities hampering credit flow in thin markets, the increase in unregulated lending is found only in a subset of census tracts that have had a historically low rate of home sales.

The estimated discontinuity (standard error) for all loan types by all lenders in large MSAs between 1994 and 2002 is 5% (1%), translating into a little more than 65 extra loans and about $8 million ($2007) of additional credit per tract *at the cutoff.*[34] A

[34] This estimate is generated using the Column 6 specification in Table 3. The point estimate for dollars lent is similar, but less precise with a standard error of 1.7%. Large MSA tracts just below the cutoff (0.78

notable limitation of the RD design is that it only identifies CRA's impact at the eligibility cutoff. Without making additional assumptions, the RD results say little about CRA's effect away from the cutoff. A reasonable assumption may be that the proportional effect holds for tracts within *at least* 0.05 of the cutoff. Deflating total credit flow for the 824 sample large MSA tracts with *TM* between 0.75 and 0.80 establishes a lower bound total impact of 53,000 loans at a value of $6.3 billion.

It is unclear whether increased lending by banks in targeted tracks is efficient. While crowd-in by non-banks points to information externalities generating suboptimal credit supply, the CRA seems too blunt to be motivated primarily as a response to this problem. Credit may be undersupplied more generally because of externalities, inadequate competition in mortgage markets or discrimination. Additional data from other sources on longer term, tract-level outcomes such as loan performance, crime and home values could shed light on these issues.

CRA may also be motivated simply along equity lines. In this regard its success may be clearer, although the possibility exists that CRA encourages banks to engage in deceptive ("predatory") lending practices that harm CRA's intended beneficiaries or excessively risky lending leads to neighborhood decline. Again, analysis of other outcomes as they become available might help us understand more clearly the broader consequences of CRA. This paper establishes motivation for further analysis and a strategy for conducting it.

References

Avery, Robert B., Raphael W. Bostic, and Glenn B. Canner. 2000. "CRA Special Lending Programs" *Federal Reserve Bulletin*, 86(11): 711-731.

Avery, Robert B., Paul S. Calem, and Glenn B. Canner. 2003. "The Effects of the Community Reinvestment Act on Local Communities" *Proceedings*.

Avery, Robert B., Kenneth P. Brevoort, and Glenn B. Canner. 2007. "Opportunities and Issues in using HMDA Data" *Journal of Real Estate Research*, 29(4): 351-380.

Barr, Michael S. 2005. "Credit Where it Counts: The Community Reinvestment Act and its Critics" *New York University Law Review*, 80(2): 513-652.

Belsky, Eric S., Matthew Lambert, Alexander Von Hoffman, and Nicolas P. Retsinas. 2000. "Insights into the Practice of Community Reinvestment Act Lending: A Synthesis of CRA Discussion Groups" , Joint Center for Housing Studies Working Paper CRA00-1.

$\leq TM < 0.80$) averaged about 1400 originations during this period at a total value of $167.6 million. Deflating these values by 5% yields the level impacts.

Bernanke, Ben S. 2007. "The Community Reinvestment Act: Its Evolution and New Challenges." Speech at The Community Affairs Research Conference, Washington, DC.

Berry, Christopher R. and Sarah L. Lee. 2007. "The Community Reinvestment Act: A Regression Discontinuity Analysis" , Harris School Working Paper Series 07.04.

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

------- 2008b. "Regression Discontinuity Estimates of the Effects of the GSE Act of 1992" Ph.D. MIT.

Bostic, Raphael W. and Breck L. Robinson. 2005. "What Makes Community Reinvestment Act Agreements Work? A Study of Lender Responses" *Housing Policy Debate*, 16(3/4): 513-545.

Bostic, Raphael, Anna Paulson, Hamid Mehran, and Marc Saidenberg. 2005. "Regulatory Incentives and Consolidation: The Case of Commercial Bank Mergers and the Community Reinvestment Act" *Advances in Economic Analysis & Policy*, 5(1): 1392-1392.

Bostic, Raphael and Brian J. Surette. 2004. "Market Forces Or CRA-Induced Externalities: What Accounts for the Increase in Mortgage Lending to Lower-Income Communities?" *Lusk Center for Real Estate Working Paper No. 2004-1013.*

Bowyer, Jerry. 2008. "Don't Blame the Markets" *The New York Sun.* April 18.

Calem, Paul S. and Susan M. Wachter. 1999. "Community Reinvestment and Credit Risk: Evidence from an Affordable-Home-Loan Program" *Real Estate Economics*, 27(1).

Calem, Paul S. 1996. "Mortgage Credit Availability in Low- and Moderate-Income Minority Neighborhoods: Are Information Externalities Critical?" *The Journal of Real Estate Finance and Economics*, 13(1): 71-89.

Campen, James T. and Thomas M. Callahan. 2001. "Boston's Soft Second Program: Reaching Low-Income and Minority Home Buyers in a Changing Financial-Services Environment" *Proceedings*: 363-399.

Canner, Glenn B., Wayne Passmore, and Elizabeth Laderman. 1999. "The Role of Specialized Lenders in Extending Mortgages to Lower-Income and Minority Homebuyers" *Federal Reserve Bulletin*: 709-726.

Caplin, Andrew and John Leahy. 1998. "Miracle on Sixth Avenue: Information Externalities and Search" *Economic Journal*, 108(446): 60-74.

Capozza, Dennis R., Ryan D. Israelsen, and Thomas A. Thomson. 2005. "Appraisal, Agency and Atypicality: Evidence from Manufactured Homes" *Real Estate Economics*, 33(3): 509-537.

Carow, Kenneth A. and Edward J. Kane. 2001. "Event-Study Evidence of the Value of Relaxing Longstanding Regulatory Restraints on Banks, 1970-2000" , NBER 8594.

Chen, David W. 2004. "U.S. Set to Alter Rules for Banks Lending to Poor" *New York Times.* October 20.

Cleary, Yahonnes and Ken Zimmerman. 2006. "House Rich, Pocket Poor and Under Threat: Home Repair Financing and Homeownership Preservation in New Jersey" , New Jersey Institute for Social Justice Report.

Dahl, Drew, Douglas D. Evanoff, and Michael F. Spivey. 2002. "Community Reinvestment Act Enforcement and Changes in Targeted Lending" *International Regional Science Review*, 25(3).

Evanoff, Douglas D. and Lewis M. Segal. 1996. "CRA and Fair Lending Regulations: Resulting Trends in Mortgage Lending" *Economic Perspectives*: 19-46.

Fishbein, Allen J. 1992. "The Ongoing Experiment with 'Regulation from Below': Expanded Reporting Requirements for HMDA and CRA" *Housing Policy Debate*, 3(2): 601-636.

Garmaise, Mark J. and Tobias J. Moskowitz. 2006. "Bank Mergers and Crime: The Real and Social Effects of Credit Market Competition" *Journal of Finance*, 61(2): 495-538.

Goldberg, Deborah B. 2000. "The Community Reinvestment Act and the Modernized Financial Services World" *ABA Bank Compliance*, 21(1): 13.

Gotham, K. F. 2000. "Separate and Unequal: The Housing Act of 1968 and the Section 235 Program" *Sociological Forum*, 15: 13-37.

Guttentag, Jack M. and Susan M. Wachter. 1980. *Redlining and Public Policy*. New York: New York University, Graduate School of Business Administration, Salomon Brothers Center for the Study of Financial Institutions.

Harrison, David M. 2001. "The Importance of Lender Heterogeneity in Mortgage Lending" *Journal of Urban Economics*, 49(2): 285-309.

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*, 13(2).

Imbens, Guido and Thomas Lemieux. 2007. "Regression Discontinuity Designs: A Guide to Practice" , NBER 0337.

Immergluck, Dan and Geoff Smith. 2006. "The External Costs of Foreclosure: The Impact of Single-Family Mortgage Foreclosures on Property Values" *Housing Policy Debate*, 17(1): 57-79.

Jackson, Kenneth T. 1985. *Crabgrass Frontier: The Suburbanization of the United States*. New York: Oxford University Press.

Joint Center for Housing Studies. 2002. "The 25th Anniversary of the Community Reinvestment Act: Access to Capital in an Evolving Financial Services System." Ford Foundation Sponsored Report.

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

Lacker, Jeffrey M. 1995. "Neighborhoods and Banking" *Economic Quarterly*: 13-38.

LaCour-Little, Michael and Stephen Malpezzi. 2003. "Appraisal Quality and Residential Mortgage Default: Evidence from Alaska" *The Journal of Real Estate Finance and Economics*, 27(2): 211-233.

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

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

Ling, David C. and Susan M. Wachter. 1998. "Information Externalities and Home Mortgage Underwriting" *Journal of Urban Economics*, 44(3): 317-332.

Litan, Robert E., Noclas P. Retsinas, Eric S. Belsky, and Susan W. Hagg. 2000. "The Community Reinvestment Act After Financial Modernization: A Baseline Report" , U.S. Department of the Treasury Report.

Marsico, Richard D. 2006. "The 2004-2005 Amendments to the Community Reinvestment Act Regulations: For Communities One Step Forward and Three Steps Back" *Clearinghouse Review*, 39.

Martin-Guerrero, Alvaro. 2002. "Financing Home Improvement Projects: The use of Home-Secured Credit" , Joint Center for Housing Studies Working Paper N02-1.

Meeker, Larry and Forest Myers. 1995. "Community Reinvestment Act Lending: Is it Profitable?" *Financial Industry Perspectives*: 13-35.

Millon Cornett, Marcia and Sankar De. 1991. "Medium of Payment in Corporate Acquisitions: Evidence from Interstate Bank Mergers" *Journal of Money, Credit and Banking*, 23(4): 767-76.

National Reinvestment Coalition. 2005. "CRA Commitments: 1977-2005"

------. 2002. "CRA Manual" .

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*, 30(2): 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.

Quercia, Roberto G., Michael A. Stegman, Walter R. Davis, and Eric Stein. 2001. "Community Reinvestment Lending: A Description and Contrast of Loan Products and their Performance" , Joint Center for Housing Studies Working Paper LIHO-01.11.

Schill, Michael H. and Susan M. Wachter. 1994. "Borrower and Neighborhood Racial and Income Characteristics and Financial Institution Mortgage Application Screening" *The Journal of Real Estate Finance and Economics*, 9(3): 223-239.

Thomas, Kenneth H. 1998. *The CRA Handbook.* New York: McGraw-Hill.

Zinman, Jonathan. 2002. "The Efficacy and Efficiency of Credit Market Interventions Evidence from the Community Reinvestment Act" , Joint Center for Housing Studies CRA02-2.

Figure 1

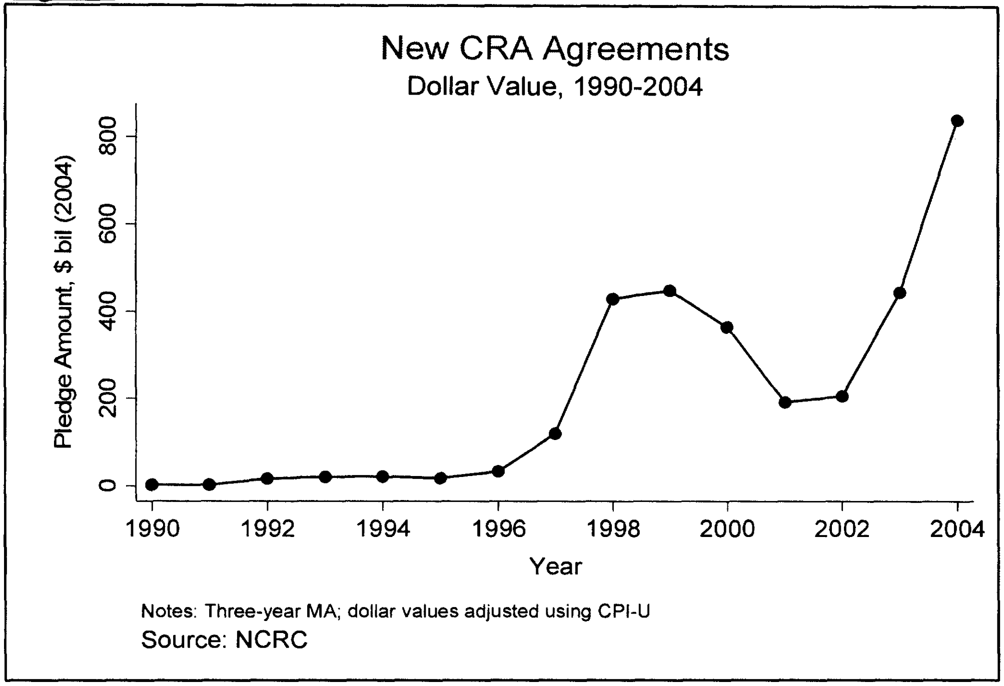
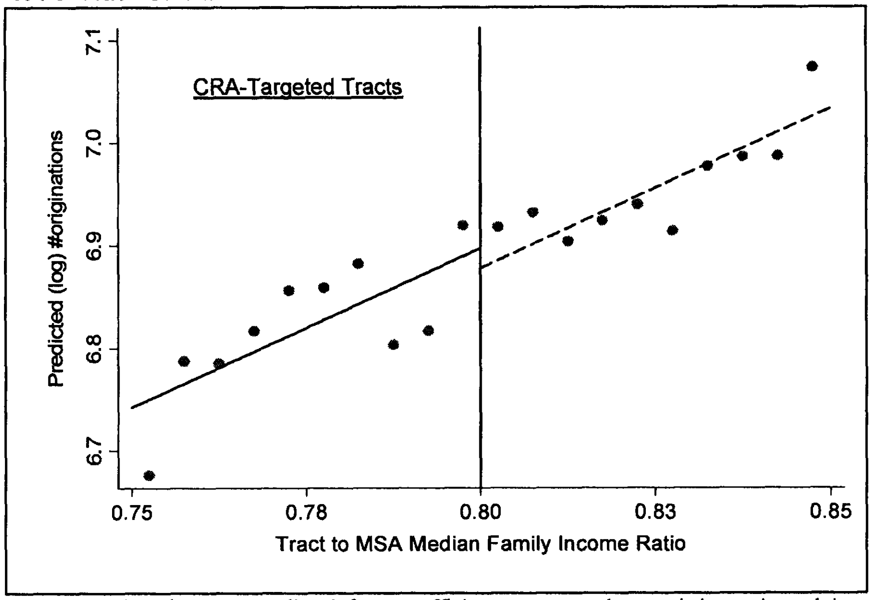

New CRA Agreements
Dollar Value, 1990-2004
Pledge Amount, $ bil (2004)
0
200
400
600
800
1990
1992
1994
1996
1998
2000
2002
2004
Year
Notes: Three-year MA; dollar values adjusted using CPI-U
Source: NCRC

Figure 2

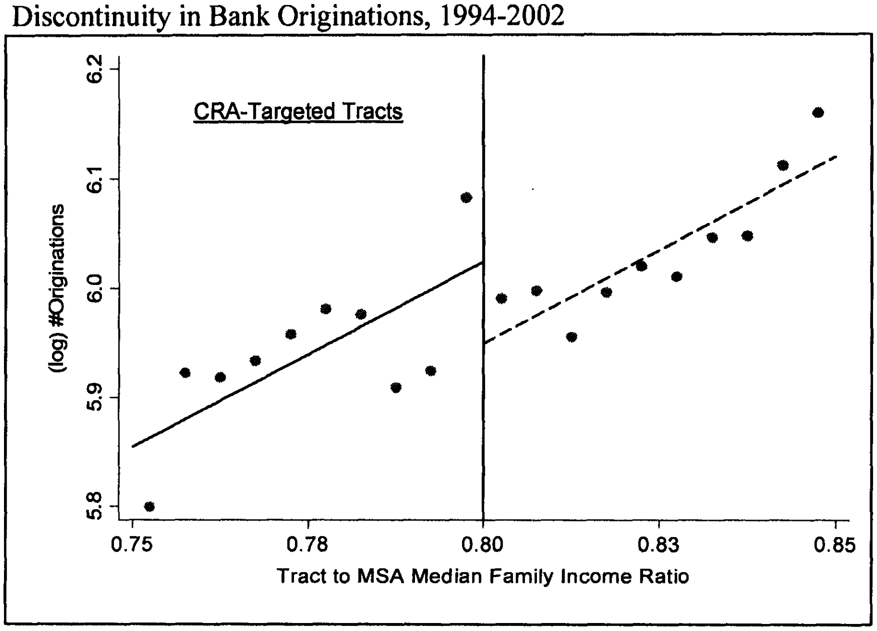
Test of Identification Assumption - Mortgage Originations Predicted by 1990 Tract Characteristics



Notes: Y-axis values are predicted from coefficients on tract characteristics estimated in a regression of (log) originations between 1994 and 2002 on tract characteristics and MSA fixed effects. Data points represent mean of predicted values for tracts within 0.5 percentage point intervals of *TM*. Fitted lines generated from regression of predicted values on TM and dummy for CRA cutoff.

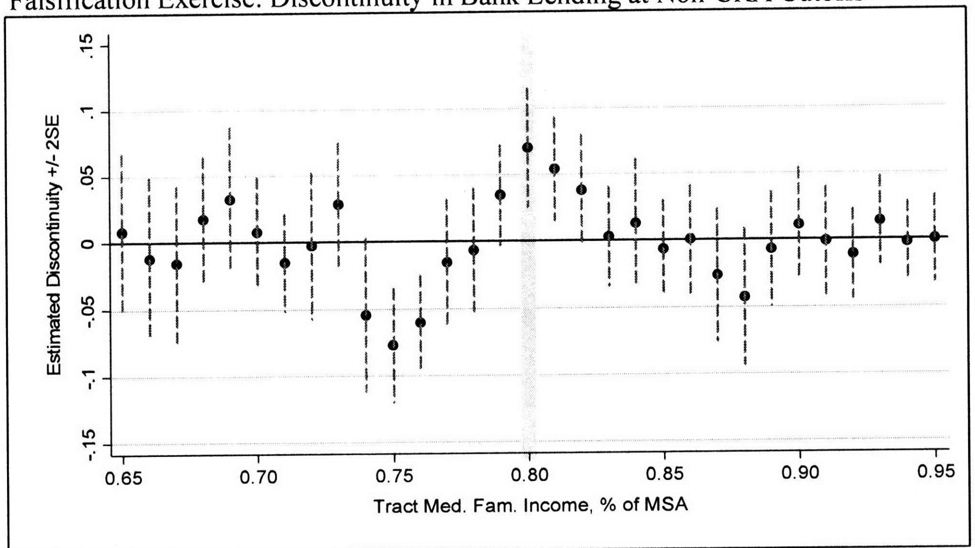
Figure 3:

Discontinuity in Bank Originations, 1994-2002



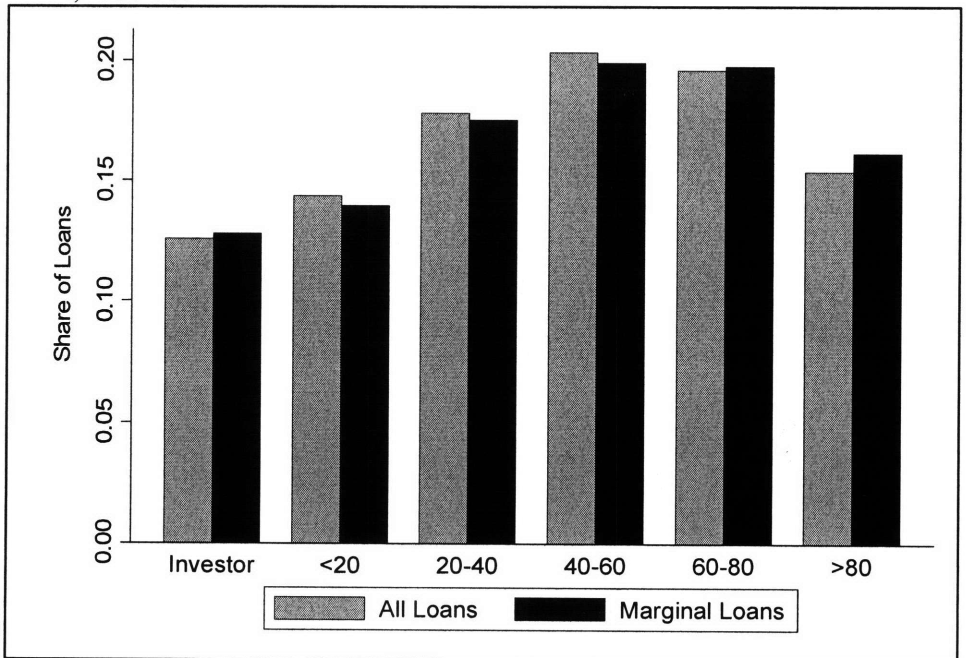
Notes: Each data point represents the mean of the Y-axis variable for tracts in 0.5 percentage point bins of *TM*. Lines are fit from a tract-level regression of the Y-axis variable on *TM* and a treatment dummy variable.

Figure 4:
Falsification Exercise: Discontinuity in Bank Lending at Non-CRA Cutoffs



Notes: Each point represents an estimated discontinuity from a separate regression using a bandwidth of 0.15 around each point and a cubic control function.

Figure 5:
Estimated Risk Distribution of Marginal Home Purchase Loans by Banks in Large MSAs, 1997-2002



Notes: Labels along X-axis define borrower risk groups. The first group refers to loans not for owner-occupancy. The other five groups refer to quintiles of the estimated risk distribution. See text for estimation strategy and full description of figure.

Table 1:
Primary Variables Available in HMDA Dataset, 1994-2005

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
HOEPA Status	2004-2005	Indicator for high-cost loan: APR at consummation exceeds yield for comparable Treasury by more than 8 percentage points.
Lien Status	2004-2005	Loan secured by first or subordinate lien
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	2004-2005	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)	(2)	(3)	(4)
	$0.95<TM<1.05$	$0.70<TM<0.80$	$0.80\leq TM<0.90$	p-Value
Number of Tracts	5826	3932	5297	
		A. Loans per Tract per Year, 1994-2002		
Banks & Thrifts				
Home Purchase Originations	31.2	19.0	23.9	<0.01
	(29.1)	(17.4)	(21.9)	
Refinance & Home Improvment Originations	54.5	32.3	40.9	<0.01
	(40.0)	(24.8)	(31.3)	
Mortgage Company Subsidiaries				
Home Purchase Originations	27.4	16.4	20.9	<0.01
	(31.3)	(17.8)	(23.3)	
Refinance & Home Improvment Originations	27.3	15.1	19.5	<0.01
	(26.6)	(14.1)	(18.9)	
Independent Mortgage Companies				
Home Purchase Originations	30.5	20.0	25.0	<0.01
	(36.5)	(20.6)	(28.1)	
Refinance & Home Improvment Originations	34.6	22.8	27.5	<0.01
	(32.2)	(19.2)	(24.2)	
		B. Census Tract Characteristics (1990)		
Total Housing Units	1917.2	1793.2	1860.9	0.01
	(1020.67)	(999.50)	(960.93)	
Owner-Occupied Units	1197.5	863.9	1016.6	<0.01
	(647.16)	(582.07)	(578.82)	
Med Value Own-Occ Units ($2007)	160,326.86	127,788.33	140,115.35	<0.01
	(98,084.81)	(92,812.21)	(95,221.16)	
Prop Units Detached	0.629	0.496	0.549	<0.01
	(0.235)	(0.269)	(0.256)	
Prop Units Multifamily	0.149	0.197	0.175	<0.01
	(0.181)	(0.217)	(0.202)	
Prop Units Mobile/Trailer	0.067	0.069	0.076	0.04
	(0.109)	(0.133)	(0.126)	
Prop Units Built 1980-1989	0.162	0.121	0.140	<0.01
	(0.157)	(0.130)	(0.142)	
Prop Units Built 1940-1969	0.435	0.438	0.444	0.41
	(0.225)	(0.206)	(0.214)	
Prop Units Built <1940	0.178	0.264	0.221	<0.01
	(0.196)	(0.248)	(0.225)	
Prop non-Hisp Black	0.072	0.177	0.119	<0.01
	(0.152)	(0.265)	(0.211)	
Prop Hispanic	0.060	0.124	0.087	<0.01
	(0.113)	(0.197)	(0.156)	
Prop of Pop 65+ Years	0.133	0.137	0.140	0.17
	(0.070)	(0.079)	(0.074)	
Prop Living in Group Qtrs	0.014	0.016	0.015	0.28
	(0.031)	(0.036)	(0.034)	

Notes: Standard deviations in parenthesis. p-Value is from test of difference between means in columns 2 and 3, clustered at MSA-level.

Table 3:
RD Estimates of CRA's Effect on Bank Lending

Dependent Variable: *(log) # originations in tract by banks, 1994-2002*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				A. Sample: All MSAs *(mean) 57.8*			
$1[TM < 0.80]$	0.0680	0.0312	0.0381	0.0337*	0.0348**	0.0267**	0.0301**
	(0.0529)	(0.0388)	(0.0273)	(0.0187)	(0.0151)	(0.0113)	(0.0131)
R-Squared	0.007	0.404	0.673	0.868	0.864	0.877	0.877
# Tracts	4708	4708	4708	4708	13633	25445	25445
				B. Sample: Small MSAs *61.6*			
$1[TM < 0.80]$	-0.0323	-0.0637	0.0120	-0.0045	0.0016	-0.0048	-0.0040
	(0.0798)	(0.0779)	(0.0562)	(0.0348)	(0.0245)	(0.0189)	(0.0222)
R-Squared	0.011	0.473	0.748	0.889	0.873	0.883	0.883
# Tracts	1266	1266	1266	1266	3716	7143	7143
				C. Sample: Medium MSAs *55.6*			
$1[TM < 0.80]$	0.0837	0.0698	0.0061	0.0114	0.0146	0.0015	0.0030
	(0.0717)	(0.0642)	(0.0413)	(0.0299)	(0.0231)	(0.0173)	(0.0203)
R-Squared	0.009	0.289	0.609	0.860	0.853	0.868	0.869
# Tracts	1642	1642	1642	1642	4796	8751	8751
				D. Sample: Large MSAs *57.1*			
$1[TM < 0.80]$	0.1257	0.0506	0.0782*	0.0764**	0.0716***	0.0729***	0.0778***
	(0.1111)	(0.0630)	(0.0431)	(0.0274)	(0.0222)	(0.0158)	(0.0182)
R-Squared	0.005	0.449	0.684	0.869	0.871	0.880	0.880
# Tracts	1800	1800	1800	1800	5121	9551	9551
Bandwidth	0.05	0.05	0.05	0.05	0.15	0.30	0.30
Control Function	linear	linear	linear	linear	cubic	cubic	quintic
MSA Fixed Effects		Y	Y	Y	Y	Y	Y
Tract Size Control[1]			Y	Y	Y	Y	Y
Other Tract Covariates[2]				Y	Y	Y	Y

Notes: Standard errors clustered at MSA-level shown in parentheses. * p<0.10, ** p<0.05, *** p<0.01. Includes home purchase, refinance and home improvement originations. (1) log total number of housing units in 1990. (2) See Table 2 for list of covariates and description in text. Regressions in columns 5-7 include variable indicating for $TM \leq 0.90$ to control for potential effects of the GSE Act. Small MSAs are those with popualation <500k, medium MSAs have population between 500k and 2 million, and large MSAs are those with population >2 million according to 1990 Census. Mean of outcome variable (per year) for selected sample in italics and measured using tracts just above cutoff ($0.80 \leq TM < 0.82$)

Table 4:
RD Estimates of CRA's Effect on Bank Lending in Large MSAs, by Loan Type and Period

	(1)	(2)	(3)	(4)	(5)	(6)
	All Loan Types		Home Purchase Loans		Refinance & Home Improvement Loans	
	1994-1996	1997-2002	1994-1996	1997-2002	1994-1996	1997-2002
			A. Originations			
mean [1]	*41.4*	*64.9*	*17.3*	*22.1*	*24.1*	*42.8*
$\mathbf{1}[TM < 0.80]$	0.0424**	0.0821***	0.0430**	0.0845***	0.0303	0.0812***
	(0.0198)	(0.0195)	(0.0183)	(0.0207)	(0.0248)	(0.0247)
R-Squared	0.856	0.865	0.770	0.785	0.851	0.875
N	9548	9551	9518	9547	9539	9551
			B. Applications			
mean	*67.9*	*118.0*	*24.9*	*33.4*	*43.1*	*84.6*
$\mathbf{1}[TM < 0.80]$	0.0319*	0.0633***	0.0213	0.0713***	0.0327*	0.0585***
	(0.0164)	(0.0150)	(0.0174)	(0.0172)	(0.0174)	(0.0165)
R-Squared	0.860	0.870	0.773	0.782	0.866	0.883
N	9551	9551	9540	9551	9550	9551
			C. Amount Originated			
mean ($000's) [2]	*3,620.77*	*7,744.20*	*2,056.80*	*3,202.10*	*1,563.97*	*4,542.10*
$\mathbf{1}[TM < 0.80]$	0.0436*	0.0699***	0.0427*	0.0872***	0.0277	0.0571**
	(0.0235)	(0.0202)	(0.0240)	(0.0229)	(0.0278)	(0.0220)
R-Squared	0.811	0.837	0.763	0.780	0.784	0.853
N	9548	9551	9518	9547	9539	9551

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, tract scale variables and covariates (see Table 2 and text for description), and are run using a bandwidth of 0.30 and cubic control function. Regressions also include variable indicating for TM $\leq$ 0.90 to control for potential effects of the GSE Act. (1) Means of outcome variables in levels (per year), calculated using tracts just above cutoff (0.80 $\leq$ TM < 0.82). (2) Adjusted to year 2007 dollars using CPI-U.

Table 5:
Net Effect of CRA on Non-Banks in Large MSAs

	(1)	(2)	(3)	(4)
	Independent Mort. Co.		Mortg. Co. Subsidiary	
	1994-1996	1997-2002	1994-1996	1997-2002
		A. Originations		
mean [1]	*37.9*	*63.3*	*18.8*	*53.9*
$1[TM < 0.80]$	0.0060	0.0335**	0.0496**	0.0376**
	(0.0196)	(0.0143)	(0.0177)	(0.0181)
R-Squared	0.782	0.825	0.804	0.856
N	9548	9551	9514	9551
		B. Applications		
mean	*71.5*	*142.5*	*28.1*	*87.0*
$1[TM < 0.80]$	-0.0043	0.0229	0.0500**	0.0363**
	(0.0202)	(0.0159)	(0.0185)	(0.0164)
R-Squared	0.780	0.821	0.811	0.850
N	9551	9551	9536	9551
		C. Loan Amount		
mean ($000's) [2]	*41,694.33*	*82,478.17*	*22,533.33*	*78,939.33*
$1[TM < 0.80]$	-0.0043	0.0376*	0.0556**	0.0381
	(0.0257)	(0.0182)	(0.0229)	(0.0238)
R-Squared	0.761	0.791	0.792	0.838
N	9548	9551	9514	9551

Notes: Standard errors clustered at MSA-level shown in parentheses. * p<0.10, ** p<0.05, *** p<0.01. Includes home purchase, refinance and home improvement originations. All regressions include MSA fixed effects, tract scale variables and covariates (see Table 2 and text for description), and are run using a bandwidth of 0.30 and cubic control function. Regressions also include variable indicating for TM ≤ 0.90 to control for potential effects of the GSE Act. (1) Means of outcome variables in levels, calculated using tracts just above cutoff (0.80 ≤ TM < 0.82). (2) Adjusted to year 2007 dollars using CPI-U.

Table 6:
Test of the Lang-Nakamura Information Externality Hypothesis

Dependent Variable: *(log) # home purchase originations in tract by lender, 1997-2002*

	(1)	(2)	(3)	(4)
	Banks		Indep. Mort. Co.	
	Low-sales Census Tracts	High-sales Census Tracts	Low-sales Census Tracts	High-sales Census Tracts
mean	*19.5*	*24.8*	*24.6*	*32.6*
$1[TM < 0.80]$	0.0876***	0.0922***	0.0882**	0.0037
	(0.0293)	(0.0278)	(0.0317)	(0.0320)
R-Squared	0.801	0.792	0.770	0.778
N	4759	4788	4762	4785

Notes: Standard errors clustered at MSA-level in parentheses. * p<0.10, ** p<0.05, *** p<0.01. All regressions include MSA fixed effects, tract scale variables and covariates (see Table 2 and text for description), and are run using a bandwidth of 0.30 and cubic control function. Regressions also include variable indicating for TM ≤ 0.90 to control for potential effects of the GSE Act. (1) Means of outcome variables in levels, calculated using tracts just above cutoff (0.80 ≤ TM < 0.82). Sample includes only large MSA census tracts. See text (Section 5.3) for details on how low-sales and high-sales samples are created.

Table 7:
Group Means: Tracts that Switched to LMI versus Tracts that Remain non-LMI

	Switchers	non-Switchers	p-Value
A. Tract Income Characteristics			
TM, 1990	84.47	85.49	< 0.01
	(0.11)	(0.12)	
TM, 2000	71.36	91.96	< 0.01
	(0.63)	(0.66)	
Med. Fam Inc, 1990	$35,028.21	$35,961.18	0.17
	(1053.85)	(1254.41)	
Med. Fam Inc, 2000	$41,313.65	$53,035.66	< 0.01
	(1723.38)	(1839.21)	
B. Tract Housing Characteristics			
Owner-Occ Units, 1990	813.55	955.72	0.02
	(104.37)	(59.32)	
Owner-Occ Units, 2000	833.33	1097.99	< 0.01
	(103.46)	(83.06)	
Total Housing Units, 1990	1756.06	1708.48	0.34
	(82.89)	(66.91)	
Total Housing Units, 2000	1812.25	1881.11	0.29
	(80.80)	(91.60)	
Med House Value, 1990	$111,486.50	$115,514.20	0.67
	(16,380.98)	(13,126.07)	
Med House Value, 2000	$135,655.60	$153,274.10	0.11
	(17,466.56)	(12,803.15)	
C. Mortgage Lending by Banks			
Change (log) Bank Originations, 2001/02 - 2004/05	0.3057	0.1555	< 0.01
	(0.060)	(0.057)	
Bank Originations, 2001/02	118.66	193.71	< 0.01
	(18.51)	(20.34)	
Change (log) Bank Originations, 1998/99 - 2001/02	0.2373	0.317	0.03
	0.0564	0.0471	
D. Mortgage Lending by Ind Mort Co			
Change (log) IMC Originations, 2001/02 - 2004/05	0.486	0.381	< 0.01
	(0.062)	(0.046)	
IMC Originations, 2001/02	115.30	162.54	< 0.01
	18.83	20.50	
Change (log) IMC Originations, 1998/99 - 2001/02	0.079	0.180	< 0.01
	(0.053)	(0.036)	

Notes: Standard errors (clustered at MSA level) in parentheses. Sample comprised of tracts from MSA's with more than 2 million people in 1990 with TM_{old} between 80 and 90 that experienced at most minor boundary changes between 1990 and 2000.

Table 8:
Effect of Changing CRA Status on Loan Origination Growth

	(1)	(2)	(3)	(4)	(5)	(6)
A. Banks						
1[switch=1]	0.1086***	0.0588*	0.0385	0.0473**	0.0476*	0.0450*
	(0.0216)	(0.0338)	(0.0251)	(0.0205)	(0.0251)	(0.0223)
R-Squared	0.462	0.466	0.471	0.570	0.472	0.570
N	1662	1662	1662	1660	1662	1660
B. Independent Mort Co						
1[switch=1]	0.0664**	-0.0208	0.0087	0.0031	0.0205	0.0018
	(0.0282)	(0.0281)	(0.0291)	(0.0223)	(0.0237)	(0.0193)
R-Squared	0.292	0.302	0.305	0.411	0.306	0.411
N	1651	1651	1651	1645	1651	1645
Control Function Specification[1]	k=0	k=1	k=3	k=3	k=4	k=4
Covariates[2]	no	no	no	yes	no	yes

Notes: Robust standard errors (clustered at MSA-level) in parentheses; * p<0.10, ** p<0.05, *** p<0.01. All regressions include MSA fixed effects. Sample includes tracts with TM_{old} between 80 and 90 and in MSA's with >2 million population. (1) See text (Section 5.5) for details on control function specification. (2) Covariates include (log) number owner-occ units in 2000, (log) med house value in 2000, (log) total housing units in 1990, change in (log) median family income from 1990 to 2000 and growth in originations between 1998/99 and 2001/02.

Chapter 2

The Effect of the Community Reinvestment Act on Mortgage Lending to Low-Income Borrowers

2.1. Introduction

Between 1994 and 2005 the U.S. homeownership rate grew by about 8% (see Figure 1), representing an expansion in homeownership not seen since the 1950's. This sharp change in the rate of growth in homeownership coincided with distinct changes in the operation of financial markets. Innovation is one aspect of these changes. Information technology improved sharply in the mid 1990's (e.g. automated underwriting and use of credit scoring), reducing the cost of mortgage lending and improving the ability of mortgage lenders to price risk (see Straka 2000). Perhaps as an offshoot of such advancements lenders also expanded the menu of mortgage contracts, providing potential borrowers with a greater range of options to finance the purchase of a home (see Garriga et al 2006). Indeed, the subprime market grew tremendously during this period, moving the mortgage market closer to one based on price rationing where more of those with sub-par credit histories can still obtain mortgage credit (see Chomsisengphet and Pennington-Cross 2006).

At the same time, a number of regulatory changes and public policies aimed at ensuring equal access to credit and increasing homeownership, especially among minority and low-income households, were initiated or strengthened in the 1990's. For instance, the "Government-Sponsored Enterprise (GSE) Act" of 1992 established mortgage purchasing goals for Fannie Mae and Freddie Mac targeted at low income households and communities (see Bhutta 2008b). Regulators and policy makers also increased enforcement of and strengthened fair lending laws between the late 1980's and mid 1990's. Congress expanded The Home Mortgage Disclosure Act (HMDA 1975), which requires lenders to disclose details on mortgage applications they receive, in 1989 and 1993 to cover more lenders and to force lenders to disclose more information that could help regulators identify discriminatory behavior. And the Community Reinvestment Act (CRA 1977), which requires federally-insured deposit institutions to help meet the credit needs of low-income communities and households, became more strictly enforced in the early 1990's and then was expanded under the Clinton Administration in 1995.

All of these changes may have helped expand access to mortgage credit and push up the homeownership rate among lower income households. Indeed, Figure 2 shows that since 1994 homeownership rates of relatively low-income households have grown considerably faster than high-income households. At the same time, a "fixed-coefficient" analysis shown in Figure 3 suggests that neither demographic changes nor income growth during the 1990's explain the trend in homeownership.[1] One alternative story is that sharp increases in credit access for historically marginal groups may have helped raise homeownership rates.[2]

In this paper, I test whether the CRA has had a causal impact on mortgage lending to low-income households. CRA has recently been at the forefront of contentious policy debates (Goldberg 2000, Chen 2004), and may very well come up again as policy makers try to understand the current mortgage market problems (e.g. Bowyer 2008). However, CRA's effects are not well-understood, confounded by the coincident changes in the mortgage market discussed. Indeed, Federal Reserve Chairman Ben Bernanke (2007) recently declared, "Distinguishing with certainty the effects that the CRA had on 'CRA-type' activity from the effects of simultaneous regulatory and market changes over this period has not been possible."

I make progress on this question by exploiting a discontinuity in CRA's selection rule to gain traction. In particular, since mid-1997 regulators have assessed banks' compliance with CRA based in part on their volume of lending to borrowers with income below 80% of their MSA's median family income (referred to as "low and moderate income", or LMI). This classification of loans as a sharp, discontinuous function of borrower income serves as the basis of a regression discontinuity design to identify CRA's impact where I essentially compare the volume of originations and applications on either side of this cutoff. More precisely, I employ local nonparametric and semi-parametric regression discontinuity methods similar to those described in Imbens and Lemieux (2008) and implemented in Ludwig and Miller's study of Head Start (2007).

The primary result of this analysis is that CRA's impact on bank lending to LMI borrowers on average across all MSA's between 1998 and 2005 is no larger than 3%. I do find, however, a substantive discontinuity of about 6% in bank loan activity (originations, applications and dollars lent) at the cutoff in the largest MSA's, where enforcement is most intense, during the five years from 1998 to 2002. For a number of reasons described below, including regulatory changes implemented in 2004 and 2005, CRA's effect could have declined in recent years. Notably, these results are similar to the findings in Bhutta (2008a) that also uses a regression discontinuity strategy to estimate CRA's impact on lending to low-income *neighborhoods*. In that paper, I also find that CRA's impact is heavily concentrated in large MSAs.

Interpreting the discontinuity for large MSA's as the causal effect of CRA is bolstered by the lack of a discontinuity for lenders that are not regulated under CRA (i.e. independent mortgage companies). The finding is also robust across alternative bandwidths and specifications of the control function. Further, I find that the relationship between loan volume and the assignment variable is otherwise smooth – that is, I do not find discontinuities at points away from the cutoff, suggesting that the one found at the cutoff is not spurious. Finally, one concern from an interpretive standpoint is whether banks might simply round *down* the income of borrowers just above the threshold in order to obtain CRA credit. But if banks are actually extending credit to new borrowers, then other loan characteristics might be different across the cutoff. One result that suggests loan characteristics below the cutoff are changing is that the proportion of loans

[1] This finding is similar to Doms and Krainer (2007).

[2] Quercia et al 2003, Gabriel and Rosenthal 2003, Doms and Krainer 2007, Chambers et al 2005 and Bostic and Surette 2004 look into the causal factors of increased homeownership.

sold into the secondary market is discontinuously lower below the cutoff, which is consistent with marginal loans not conforming to the standards of secondary market purchasers. Other results that look directly at loan and borrower characteristics available in the HDMA do not yield strong evidence that loan composition is changing across the cutoff.

In the next section, I discuss CRA in more detail and some of the previous literature trying to identify its impact. Section 3 discusses the HMDA data in more detail and also compares summary statistics from these data to comparable statistics drawn from the Census and CPS in order to gauge HMDA's reliability. Section 4 discusses the regression discontinuity strategy and Section 5 presents the results. Finally, in Section 6 I summarize and draw some conclusions about the impact CRA may have had on homeownership over the period in question based on the empirical results from Section 5 combined with other available evidence.

2.2. Background & Related Literature

Congress passed the CRA in 1977 in response to claims that banks were helping to uphold a historical practice of irrationally redlining low-income, urban neighborhoods.[3] CRA originally mandated that federally insured deposit institutions must help meet the banking and credit needs of the entire community where they earn deposits. In practice, this means providing evidence to regulators that they are supplying credit in both high and low income neighborhoods within their operating market. A 1995 reform, which went into full effect by mid-1997, added non-geographic criteria to CRA. In particular, banks also have to demonstrate that they are providing credit to low-income borrowers (regardless of the borrower's neighborhood income) and to small businesses.[4] To be clear, this paper focuses on estimating CRA's effect on lending to low-income borrowers.

Regulators periodically inspect banks' lending records (largely using HMDA) to judge their compliance with CRA. CRA applies to bank and thrift institutions covered by federal deposit insurance. It does not apply to non-deposit independent mortgage companies or credit unions. Non-deposit mortgage subsidiaries of banks are not covered either, but banks can include lending by their subsidiaries in their CRA evaluations. Examiners rate banks in each of their "assessment area(s)" – generally the MSA's and/or counties where they have branches – separately and then combine these ratings into an overall grade (outstanding, satisfactory, needs to improve and substantial noncompliance).

Regulators can penalize banks with poor CRA evaluations. Applications that banks submit to merge with another bank or open a bank branch can be denied by regulators on CRA grounds. CRA also provides a formal outlet for community organizations to voice dissatisfaction with banks operating in their area. Regulators must weigh evidence brought by the public regarding bank compliance with CRA before deciding on a bank's merger application. Community groups, which are particularly active in large cities, use the detailed mortgage lending data available through HMDA

---

[3] For example, FHA underwriting manuals through the 1950's explicitly warned lenders against lending in minority neighborhoods, and Gotham (2000) cites examples of continued discriminatory language in private underwriting and appraisal manuals through the 1970's. Also see Jackson (1985) for an account of FHA's redlining policies.

[4] See Barr (2005) and Fishbein (1992) for more on CRA's history.

and small business lending data disclosed by banks under CRA to challenge bank applications.[5] One response by banks has been to enter into "CRA Agreements" with community groups where banks pledge a certain amount of resources to targeted neighborhoods (see Bhutta 2008a). Bostic and Robinson (2005) find that banks do increase targeted lending during the years in which an agreement is in effect. And bankers interviewed by Harvard's Joint Center for Housing Studies (JCHS) said that many banks have taken considerable steps toward increasing CRA-qualified loans to avoid CRA-associated difficulties and bad press (Belsky et al 2000).

In order to quantify banks' efforts toward meeting CRA objectives, regulators define low and moderate income (LMI) borrowers as those with income below 80% of their MSA's median family income. And then as part of the CRA evaluation, regulators measure a bank's volume of lending to LMI borrowers and compare this volume to that of its peers as well as to the bank's non-LMI lending volume. Regulators therefore give banks "CRA-credit" as a discontinuous function of borrower income. So to measure CRA's effect, I will measure the jump in loan volume at the point where CRA-credit jumps.

This strategy differs from most other studies of CRA. A few studies test for bank-specific reactions to particular CRA incentives, with mixed results. Dahl et al (2002) show that changes in CRA-type lending is uncorrelated with CRA rating downgrades in the early 1990's. In contrast, Bostic and Robinson (2005) find that banks increase targeted lending during the years in which they have a CRA agreement in effect, and Bostic et al (2005) find that merging banks' CRA-qualified lending increases prior to acquisitions of other banks.

Other studies, as this one, try to estimate CRA's impact more broadly. JCHS researchers (2002) show that banks make a higher fraction of their loans to and are less likely to deny loan applications from CRA-targeted populations inside the bank's assessment area (i.e. where banks get "CRA credit") compared to outside those geographic boundaries. But because bank operations likely differ in areas where they have branches relative to where they do not – for instance, better knowledge of the local market and population which may allow more lending to low-income borrowers – this test lacks a causal interpretation.[6]

A few studies (Schill and Wachter 1994, Evanoff and Segal 1996, Bostic and Surette 2004) find that unregulated LMI loan growth exceeds that of regulated lenders, implying CRA's effect has been limited. But *excessive* unregulated loan growth suggests that the trend in unregulated lending is not likely to provide a valid counterfactual to test CRA's effect on bank behavior, perhaps because unregulated lending operates under a different business model. Another concern is that reductions in underreporting during the 1990's likely inflate loan growth for independent mortgage companies more so than for banks (JCHS 2002).

Finally, Berry and Lee (2006 working paper) conduct a regression discontinuity analysis and conclude CRA has no impact on banks, but their study differs from the one

---

[5] For instance, protest of a merger application by WesBanco in 2001 resulted in a year long delay. See NCRC (2002).

[6] They also exploit variation across MSA's in CRA Agreements. As the authors acknowledge, this strategy also does not have a causal interpretation since MSA's where banks sign into agreements are likely to be different from other MSA's and CRA agreements locations may be endogenous.

here in a number of ways. First, they test for a discontinuity in loan rejection rates at the cutoff, which assumes that application volume and quality is exogenous. But banks may advertise in order to generate more applications and (high-risk) borrowers are more likely to apply when the likelihood of acceptance rises, so it is important to analyze loan volume and application volume separately. Indeed, Bhutta (2008a) finds that both originations and applications rise in CRA-targeted census tracts, calling into question Berry and Lee's assumptions.

Second, I allow for differential effects of CRA over time and across MSA's of different size, guided by information about CRA enforcement practices. Since CRA's impact is likely to be concentrated in large MSA's (see Bhutta 2008a), focusing on these areas can help to identify, quantify and describe CRA's impact. Also importantly, Berry and Lee pool data from 1993 to 2003, but CRA did not award banks for lending to LMI *borrowers* until after the 1995 reform. In other words, they combine years when CRA should not have had an effect with years when it may have, which reduces the ability to identify small effects. Finally, I analyze separately banks and their mortgage company subsidiaries. Banks have the option to include lending by their subsidiaries in their CRA exam. Some have argued (e.g. Marsico 2006) that banks may game CRA, for instance by referring non-LMI borrowers to their subsidiaries to help increase the bank's proportion of LMI loans. Separately analyzing banks and their subsidiaries will allow me to test this "crowd-out" scenario. Similarly, I will test for a discontinuity in lending by independent mortgage companies that will indicate whether these lenders are crowded out by increased bank lending (as suggested in the JCHS study).

2.3. Data

Data on individual mortgage applications 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 expanded coverage to include the majority of independent mortgage companies, which extend about a third of home purchase loans.[7]

Importantly, HMDA includes lender identification information that allows one to match each application to its class of lender. I divide institutions into the three groups: "banks", which fall under CRA authority; "mortgage subsidiaries of banks" which may be directly affected by CRA because of their bank affiliation; and "independent mortgage companies", which fall entirely outside the purview of CRA. Credit unions, which extend a very small share of mortgage credit, are included in this last group because they also fall outside the reach of CRA.

Applicants' gross family income is another key piece of information in the HMDA for this analysis as it allows determination of which applications qualify as CRA loans. Some of the other variables available in the HMDA that will be used in this analysis are the census tract of the property, the type of loan being applied for (e.g. home purchase), the loan amount, the disposition of the application (e.g. denied), whether the loan was sold by the lender into the secondary market, and a few demographic

[7] See FFIEC website (www.ffiec.gov) for details on which institutions are still exempt from reporting.

characteristics such as race and gender. Table 1 provides a more complete description of the HMDA variables.[8]

For each year of HMDA data used in the analysis, I use owner-occupied home purchase applications for which income is reported (more than 95% of reported applications) and can be matched to an MSA (i.e. loans that have a valid census tract code, which I then match to that tract's MSA). I exclude Hawaii and Alaska MSA loans from the analysis.

The top panel of Table 2 provides some descriptive statistics from the HMDA using all home purchase loans originated in MSA's. The left half of the table shows levels statistics for 1999. I chose this year so that I could compare the statistics of borrowers in the HMDA to those of recent home buyers from the 2000 Census (Panel B). The first row indicates that nearly 3.7 million loans were reported with a just under 1.2 million of those being to LMI borrowers. This volume compares to about 4.7 million new home buyers in the Census. These are household heads that indicate having moved in 2000 or 1999, own their current residence and have a mortgage. Much of the difference in volume between HMDA and the Census is likely explained by the fact that the Census includes movers during the first quarter of 2000 in addition to movers during 1999. As such, these results suggest that the vast majority of home purchase mortgage activity in MSA's by 1999 is covered by HMDA. This is an important observation that has not, to my knowledge, been documented elsewhere.

HMDA and Census statistics on race, gender and marital status are similar although in all cases HMDA proportions for such groups are lower than in the Census. This is because a sizeable number of loans are submitted via the internet or mail and in many instances borrower characteristics, especially race, are not reported for these loans. Overall, LMI borrowers/buyers are younger, less educated, more likely to be of a minority group, less likely to be married, and are more likely to purchase a condominium or mobile home.

The right half of Panel A shows annualized growth in the variables between 1994 and 2002. The right half of Panel B provides benchmarks using the CPS over a similar period, again drawing respondents that recently moved and own their home (mortgage status is not identified). One notable result is that loan growth in the HMDA greatly exceeds that indicated by the CPS because of underreporting in earlier years. Further, the difference in growth rates across the HMDA and CPS is greater for the non-LMI group. These results suggest HMDA is not likely to provide an accurate description of changes in the mortgage market. Near the bottom of Panel A one can see that banks' (i.e. CRA regulated) share of LMI loans (i.e. CRA-eligible) are falling. This finding echoes that in Bostic and Surette (2004). But as argued earlier, this result tells us little about CRA's impact because of, among other reasons, the unreliability of HMDA in measuring changes over time.

2.4. Empirical Design

2.4.1 Regression Discontinuity

I take advantage of a sharp discontinuity in CRA's eligibility rule to identify the impact of CRA on home purchase loan volume. Since mid-1997 banks' CRA performance is based in part on their volume of lending to low and moderate income

[8] Also see Avery et al (2007) for an in depth review of the HMDA data.

(LMI) individuals. LMI borrowers are defined as those with gross family income less than 80% of median family income of the MSA of residence[9]. In the regression discontinuity (RD) analysis that follows, this ratio of applicant income to MSA median family income is the "assignment" (or "running") variable and is referred to as R. Therefore, families with $R<0.80$ form CRA's target population. MSA median family income is measured each year and reported by the Department of Housing and Urban Development (HUD). Importantly, these MSA-level income estimates are also used both by lenders and regulators to determine which loans qualify for "CRA credit". Thus, for each loan I am able to calculate the true value of R.

For each of the three types of institutions, I count the number of home purchase loans within MSA-year-income bins where income is in $1000 increments corresponding to the fact that lenders report income rounded to the nearest thousand.[10] I focus on CRA's effect in MSA's since CRA enforcement has historically emphasized urban areas, and because HMDA coverage of MSA's is more consistent and reliable than for non-metro areas (Avery et al 2007).

The strategy for quantifying CRA's causal impact is to measure the discontinuity in bank-originated loan volume at the 0.80 cutoff. The key advantage of the RD design is that characteristics of the underlying population are nearly identical just above and below this cutoff. In other words, in the absence of CRA, one would not expect a jump in loan volume at $R = 0.80$. As such, this strategy is not confounded by the major concurrent changes in financial markets. I implement local RD methods described in Imbens and Lemieux (2008). The causal effect of CRA on loan volume, β, will be estimated by the following regression.

$$(4.1) \qquad Y_{IMt} = \beta D_{IMt} + C(R_{IMt}) + e_{IMt}$$

where Y_{IMt} is the (log) number of loans to borrowers of income I in MSA M in year t, $D_{IMt}=\mathbf{1}[R_{IMt}<0.80]$, and $C(R_{IMt})$ is a (unknown) smooth function controlling for the relationship between loan volume and relative income. Observations will be weighted using kernel weights $W_{IMt} = K([R_{IMt} - 0.80]/h)$, where h is some chosen bandwidth – i.e. observations more than h away from the cutoff will not be used to estimate the discontinuity.

The primary obstacles in estimating (4.1) are modeling the control function (i.e. $C(R_{IM})$) and choosing h. I pursue a strategy similar in spirit to that of Ludwig and Miller (2007), presenting several estimates under different bandwidths and choices of the control function. I also present results using both a simple rectangular kernel – $K(a) = \mathbf{1}[0\leq |a|\leq 1]$ – as well as a triangular kernel – $K(a) = \mathbf{1}[0\leq |a|\leq 1]*(1-|a|)$. A credible discontinuity should be robust to the choice of weighting scheme (Imbens and Lemieux 2008). In all regressions, I include MSA by year fixed effects so that β is identified only from variation in loan volume across the cutoff within MSA by year cells.

[9] For borrowers outside of an MSA, LMI is defined using the median family income of the state's non-metropolitan counties.

[10] Although income is measured discretely, the running variable is essentially continuous since regulators do not round R and thus I am not concerned about specification error due to discreteness as described in Lee and Card (2008).

As a starting point, I show nonparametric estimates for a small bandwidth (h = 0.02). These estimates simply compare mean loan volume just to the left and to that just right of the cutoff. In terms of (4.1), this strategy sets $C(R_{IM})$ to zero. As shown by Porter (2003), bias under this method increases in the slope of the relationship between Y and R.

If the outcome and assignment variables are highly correlated around the cutoff, a preferred RD strategy is local linear regression. Under this method two (kernel weighted) linear functions are estimated (one on either side of the cutoff) and the effect of CRA is calculated as the difference in the limits of these functions at the cutoff. In terms of (4.1), $C(R_{IM})$ takes the form:

$$C(R_{IMt}) = R'_{IMt} + R'_{IMt} * \mathbf{1}[R_{IMt} < 0.80], \quad R'_{IMt} = R_{IMt} - 0.80 \tag{4.2}$$

Local linear regression based estimates will be presented for bandwidths of 0.04 and 0.06. Finally, for the largest bandwidth used (h = 0.06), I also estimate quadratic functions on either side of the cutoff as opposed to linear functions. OLS robust standard errors are reported.[11] It is expected that a substantive discontinuity will be robust to these alternative specifications.

Figure 4 plots loan volume from 1998 to 2005 against reported borrower (nominal) income. Reported income does not appear to be smoothly distributed as there are spikes at different income levels, particularly those incomes divisible by \$5000. The most pronounced spike occurs at \$60,000. These spikes do not necessarily reflect poor reporting by banks – family income in the CPS generates a similar pattern (not shown). Regardless, in the regressions that follow I will include two dummy variables – one indicating for income divisible by \$5000 and another indicating for income of \$60,000 – to help ensure that the discontinuous distribution of income does not drive the results.

2.4.2. Gaming Behavior by Banks

If a discontinuity in loan volume is detected, an important question from an interpretive standpoint is whether lenders misreport income in order to help satisfy their CRA obligations. For instance, perhaps banks deliberately round *down* the income of an applicant that would have been just above the cutoff in order to get "CRA credit". One reason this may not actually occur in practice is that regulators test banks' reporting quality by cross-checking HMDA records with actual loan records and can punish banks for misreporting.[12]

Nevertheless, I can run two tests that might help resolve this issue. One test utilizes the micro loan data to estimate "who" the marginal borrowers are. If banks are actually extending credit to new borrowers, then borrower composition just below the cutoff should be of lower quality than above the cutoff.

A second test looks at loan sales into the secondary market around the cutoff. If borrower quality just below the cutoff is in fact deteriorating, this may show up in the

[11] Clustering the standard errors at the MSA level may be theoretically justified, but clustering in this way generally yielded smaller standard errors and so I present only robust standard errors.
[12] See Interagency Examination Procedures at http://www.fdic.gov/news/news/financial/2006/fil06033.html

form of fewer sales of loans into the secondary market since these loans may not conform to the usual guidelines of secondary market purchasers (e.g. Quercia et al 2001).

These tests are conducted within the same RD framework as described above, except Y_{IMt} will represent variables such as the proportion of loans in bin *IMt* that were sold into the secondary market.

2.5. Results

2.5.1. Aggregate Results

The top panel of Figure 5 plots bank home purchase loan volume against the assignment variable, *R*, aggregating data from all MSA's and years 1998 to 2005. Each data point represents loan volume in a two percentage point bin of *R*. A line is drawn at the CRA cutoff and loan volume appears to change smoothly across the cutoff (i.e. there is no indication of a CRA effect).

The relationship is highly nonlinear, especially near the cutoff (the rise and then fall of loan volume across income reflects the total number of households in the underlying population in each bin), suggesting that a "global" RD strategy would be difficult to implement reliably. The bottom panel zooms in on data in a relatively small window around the cutoff ($0.60 \leq R \leq 1.00$), and also displays a nonparametric fit of the data on either side of the cutoff.[13] Again, there is no indication of a CRA effect.

Table 3 provides regression results corresponding to Figure 5 using the methodology described in Section 4. Columns 1 and 2 use rectangular and triangular kernel weights, respectively, with h=0.02 and does not include a control function. These specifications yield similar point estimates indicating an effect of CRA on bank home purchase loan volume of between 2% and 3%. Columns 3 and 4 set *h* to 0.04 and again exclude a control function, and the point estimates rise slightly but are not substantively different from those using $h = 0.02$.

Columns 5 and 6 include a linear control function as specified in (1.2) above. The slope coefficients (third and fourth rows) suggest a negative relationship between loan volume and *R* to the right of the cutoff, and including the control function pushes the discontinuity point estimates close to zero. In fact, the point estimates in columns 3 and 4 lie further than two standard errors away from the point estimates in columns 5 and 6, respectively. As discussed previously, the bias of the nonparametric RD estimates in the first four columns will rise in the (absolute value of the) slope of the control function.

Columns 7 thru 12 set *h* to 0.06. A larger bandwidth utilizes more data, which helps in estimating the control function. At the same time, a larger bandwidth is more likely to generate bias due to an improperly specified control function. The standard errors in columns 9 and 10 are smaller than those in columns 5 and 6, while the point estimates are about three times larger and the discontinuity estimate in column 9 is statistically significant. The last two columns of Table 3 include quadratic terms in the control function. These terms turn out to be significant and both discontinuity point estimates become statistically indistinguishable from zero. Overall, Table 3 suggests CRA has <u>on average</u> had little effect, if any, on loan volume.

---

[13] For this exercise I simply use the Stata "lowess" command to fit the data using the default options which employs a bandwidth that uses 80% of the data and tri-cube kernel weights.

2.5.2. Disaggregated Results

Now I allow for the effect of CRA to vary over time and across MSA's of different size. As argued in Bhutta (2008a) CRA enforcement is sharpest in the largest MSAs, and RD estimates CRA's effect on lending to LMI census tracts found in Bhutta (2008a) reflect such an enforcement pattern.

CRA's effect may also vary over time for several reasons. One is regulatory changes to CRA implemented in 2004 and 2005 that may have weakened its impact (Marsico 2006). Another is the surge in subprime credit supply during the first half of this decade (see Chomsisengphet and Pennington-Cross 2006) that may have made CRA superfluous.[14] And merger activity, which makes CRA more salient an issue for banks, peaked in the late 1990's and then dropped to considerably lower levels by 2003.

The top, middle and bottom panels of Figure 6 show RD results for small, medium and large MSA's, respectively. Each point represents the estimated discontinuity using specification 6 of Table 3 (i.e. triangular kernel with $h = 4$ and linear control function) for the sample of MSA's listed at the top of the graph and four years of data around the year listed on the x-axis (see Figure 6 notes for more details). A dashed vertical line was placed at 1997.5, indicating the point in time at which the 1995 changes to CRA (i.e. including banks' lending to LMI borrowers in their CRA exam) were fully implemented.

Mirroring the pattern of results in Bhutta (2008), the top two panels provide no evidence of a CRA impact in small or medium size MSA's while the bottom panel reveals evidence of a CRA impact in the largest MSA's in the 5-6 year period following full implementation of the 1995 reforms. The estimates labeled as being for 1999, 2000 and 2001 utilize data from 1997 to 2002 and show a statistically significant impact of just over 5%. The first and second points in this graph use data largely prior to the reform and so are expected to be zero. The point estimates do not fall right on zero, but they cannot be distinguished from zero. The last three points suggest CRA's impact was only temporary consistent with market forces eliminating the "need" for CRA and/or the 2004-2005 changes reducing CRA's impact.

Panel A of Table 4 tests the robustness of the finding of a positive CRA impact for large MSA's. Table 4 also provides similar estimates for the other two types of lending institutions: somewhat regulated mortgage company subsidiaries (Panel B) and unregulated independent mortgage companies (Panel C). I do not report the slope coefficients in Table 4 to save space.

The discontinuity estimates in Panel A for banks are quite robust across the different specifications. Unlike in Table 3, the point estimates are very similar in moving from columns 3 and 4 to columns 5 and 6 which include a control function. Indeed, the slope coefficients (not shown) for this subsample are comparatively small and not statistically significant as they were in Table 3.[15] In other words, the relationship

[14] In other words, borrowers both above and below the cutoff could more easily obtain credit in the new environment of relaxed credit standards. For instance, in a push to expand loan volume Countrywide eased credit standards considerably by 2003 (Hagerty and Simpson 2008). Indeed, I find that home purchase loan volume (all lenders) in the HMDA for MSA borrowers with R around the CRA cutoff (those between 70 and 90) rose by nearly 30% between 2002 and 2005 after zero growth from 1999 to 2002.

[15] For instance, the coefficients (standard error) on the terms on R' and $(R')*1[R \leq 0.80]$ for the regression in column 5 are -0.0025 (0.0079) and 0.0094 (0.011), respectively.

between bank home purchase loan volume and R seems relatively flat in the neighborhood of the cutoff and so adding a control function does not matter greatly. As such, the estimates in columns 3 and 4 which exclude a control function provide both reliable and relatively efficient estimates of the discontinuity.

Panel B provides some evidence of a CRA impact on home purchase lending volume by banks' mortgage subsidiaries of about 2-4%. The point estimates are quite similar across the specifications (the largest point estimate is 0.037 in columns 2 and 9 and the smallest point estimate 0.022 in column 3), but not all are statistically significant. Importantly, these results bear no sign of banks referring non-LMI loans to their mortgage affiliates while servicing LMI loans themselves as some have suggested they might do (e.g. Marsico 2006). Perhaps manipulation of this sort is costly relative to the benefits of achieving a higher CRA grade.

Panel C provides no evidence that CRA affects unregulated independent mortgage companies. Eight of the ten point estimates are no larger than 0.02 and none are statistically significant. Again, there is no evidence that increased lending by banks due to CRA crowds out non-bank lenders. At the same time, the lack of a *positive* discontinuity for independent mortgage companies reinforces the interpretation of the results in Panel A that the measured discontinuity for banks represents the causal effect of CRA.

Figure 7 provides results of another falsification exercise where I test for discontinuities at points in the neighborhood around the CRA cutoff. An important identification assumption is that the relationship between loan volume and R would be smooth at the CRA cutoff in the absence of CRA. While not directly testable, confidence in this assumption would be bolstered if the empirical relationship were found to be smooth (i.e. no discontinuities) at points around the cutoff.

The top, middle and bottom panels show results of this falsification exercise using specifications 1, 3 and 5, respectively, from Table 4. In the top panel the only statistically significant discontinuity other than at the CRA cutoff is at $R = 0.81$. But this discontinuity actually reflects the CRA-induced discontinuity since the RD estimate at 0.81 compares a control mean that uses data between 0.81 and 0.83 against a "treatment" mean that uses data between 0.79 and 0.81 (i.e. some observations are part of the true treatment group). Elsewhere, the estimated discontinuities lie around the zero line and none are statistically significant. Similarly, the middle panel indicates that the largest discontinuity under the 3rd specification occurs at the true CRA cutoff and then the estimates trail off smoothly on either side of this cutoff and are zero almost everywhere else. Approaching $R = 0.70$ small negative discontinuities appear, which reflects the downward sloping portion of the relationship between loan volume and R (see top panel of Figure 5) combined with a larger bandwidth relative to estimates in the top panel (i.e. increased bias).

The bottom panel is somewhat less well-behaved, generating three negative discontinuities within 3-percentage points of the CRA cutoff. Again, these non-CRA discontinuities may simply reflect the CRA-induced discontinuity at 0.80 since these non-CRA RD estimates combine data above and below the actual CRA cutoff. Outside of the points immediately around the CRA cutoff, the RD estimates lie much closer to the zero line and none are statistically significant. Overall, the results in Figure 7 are encouraging in that they indicate the empirical relationship between loan volume and R is generally smooth except for at the CRA cutoff.

2.5.3. Changes in Portfolio Characteristics

If banks did in fact extend more credit to marginal applicants because of CRA then the risk characteristics of the loan portfolio just below the CRA cutoff should be different than the characteristics of the portfolio just above. In this section I perform two tests to gauge whether such a change has occurred.

In the first test I use individual loan and borrower characteristics along with application outcomes (e.g. application denied by lender) to predict each application's probability of being denied given those characteristics. I will then test for sharp differences in the share of loans to the (predicted) highest and lowest "risk" borrowers across the CRA cutoff. Of course, predicted denial probabilities reflect underlying risk as well as other factors. For instance, I find that minorities are more likely to be denied, which reflects both the fact that minorities typically have lower credit scores (not observed in HMDA) as well as discrimination they may encounter.

The second test uses information in HMDA on whether a loan is sold to secondary market purchasers, specifically testing whether there is a sharp change in the share of loans sold across the CRA cutoff. Generally, for a loan to be sold into the secondary market its risk characteristics (e.g. borrower income and credit score, loan amount to house value ratio, etc.) relative to its interest rate must conform to certain specifications. "Affordable" loans designed for marginal borrowers are less likely to be "cookie-cutter" loans typically sold into secondary and instead will be held in portfolio by the originating institution (e.g. see Avery et al 2000 and Quercia et al 2001). As such, if banks are making specialized loans to reach out to new borrowers then the share of loans just below the cutoff that are sold should be lower than that share just above the cutoff.

To estimate the risk of an application, I first run the following regression:

$$(5.1) \qquad deny_{jMt} = \alpha + \mathbf{x}_{jMt}\boldsymbol{\beta}_{Mt} + \lambda_{Mt} + \varepsilon_{jMt}$$

where $deny_{jMt}$ is an indicator variable equal to one if application j in MSA M and year t was denied by the lender, $\mathbf{x}$ is a vector of loan and borrower characteristics, λ is a set of MSA by year fixed effects and the coefficients on $\mathbf{x}$ are allowed to vary by MSA and year. (5.1) is run using all home purchase applications at banks with R between 0.70 and 0.90 and for which a credit decision was made. I don't include applications that were withdrawn by the lender due to incompleteness or were withdrawn by the applicant before the lender's decision was made.[16] The variables in $\mathbf{x}$ are as follows:

- Amount-to-income ratio; I create five groups: (0, 0.5], (0.5, 1.5], (1.5, 3], (3, 4], (4, ∞)
- (log) Median home value of census tract where property is being purchased
- Race/Ethnicity; I create five groups: White, Asian, Black, Hispanic and Other, which is comprised mostly of mail/internet applications where race is unreported, but also includes Native American

[16] I also drop loans recorded as being farm loans.

- Primary applicant sex; there are three groups: male, female and unreported on mail/internet application[17]
- Co-applicant sex; there are four groups: male, female unreported on mail/internet application and no co-applicant
- Loan type: there are three groups: conventional, FHA-insured, VA-insured

Using the estimated coefficients I generate predicted values which yield predicted risk distributions by MSA and year:

$$\widehat{deny}_{jMt} = \hat{\alpha} + \mathbf{x}_{jMt}\hat{\boldsymbol{\beta}}_{Mt} + \hat{\lambda}_{Mt} \quad (5.2)$$
$$\widehat{deny}_{jMt} \rightarrow \widehat{f_{Mt}}$$

Finally, I assign each application a number 1-5 corresponding to its quintile in its MSA by year distribution.

Table 5 shows summary statistics that result from this process for the year 2000. The first number in the top row shows that there were almost 93,000 home purchase applications in the sample for year 2000. The next two columns show statistics for those loans that fall into the last and first quintiles, respectively, of their predicted risk distribution as described above. As such, about 20% of all applications fall into each category. The second row shows that about 19% of all applications are denied. About 37% classified as high risk are denied while only 9% of applications categorized as low risk are denied.

Table 5 also indicates that high risk loans tend to have a relatively low loan amount-to-income ratio. Since, all else equal, risk rises with the amount-to-income ratio, the negative correlation in the data indicates that low ratios must reflect other deficiencies of the application.

Further comparison of columns 2 and 3 shows that higher denial rates are correlated with neighborhoods that have relatively low home values, with being Black or Hispanic, as well as with being a mail/internet application. A substantial share (about 23%) of low risk applications are FHA-insured. Applicants that meet the requirements to apply for an FHA loan represent a low risk to the bank since these loans are insured and therefore are unlikely to be denied. Finally, the last row indicates that the probability of a loan being sold into the secondary market is negatively correlated with risk.

Panels A and B of Table 6 show RD results where the outcome variables are the (log) high and low risk proportion of bank originations, respectively. Panel A does not provide strong evidence that the share of highest risk applications has increased. Eight of ten coefficients are positive, but only one specification generates a statistically significant coefficient. The lack of a sharp rise in Panel A suggests the rise in lending by banks is not concentrated at the upper tail of the risk distribution.

Panel B provides some evidence of a fall in the proportion the lowest risk applications. Six of the ten specifications yield statistically significant estimates around 0.04 or 4%. Such a drop indicates that the increase in bank lending of about 6% as

[17] A very small fraction (less than 0.1%) of applications in HMDA are recorded under a fourth group "NA". I drop these applications.

measured earlier occurs disproportionately outside of the lower tail of the risk distribution. This conclusion is tempered by the finding in columns 5, 6, 9 and 10. The somewhat more flexible specifications in columns 5, 6 are not significant, and the estimates in columns 9 and 10 which use a quadratic control function yield estimates just below zero.

The results of the second test, given in Panel C, are more robust and indicate that the risk characteristics of marginal loans differ from the average. All of the specifications in Panel C yield similar and statistically significant results suggesting a decline in the proportion of loans sold by about 4%. This result implies that a majority of the "extra" loans made by banks because of CRA are not sold into the secondary market, which is consistent with these loans being made under flexible terms aimed at marginal borrowers such that they do not conform to the requirements of secondary market purchasers.

2.5.4. The GSE Act as a Potential Confounder

The GSE Act (1992) created mortgage funding goals for Fannie Mae and Freddie Mac. One of these, the "Special Affordable" goal, targets borrowers with income less than 80% of the MSA median family income (i.e. identical to CRA) that are purchasing a home in a census tract that has median family income less than 80% of the MSA median family income. The GSE Act may therefore confound the interpretation that discontinuity in loan volume at $R = 0.80$ is due to CRA. However, a number of factors suggest that the observed impact is due to entirely to CRA.

For one, the only discontinuity in loan volume found thus far occurs in large MSAs, consistent with patterns of CRA enforcement. Any GSE effect is not expected to follow this pattern. Second, this discontinuity is found only for banking institutions and somewhat for banks' subsidiaries, while no discontinuity is found for independent mortgage companies. Again, if the GSE program were effective it would be expected to affect all types of lenders. Finally, of the three goals of the GSE Act, the 'Special Affordable' is allotted the fewest resources.

Table 7 provides more concrete evidence that the GSE program does not help explain the discontinuity in loan volume at $R = 0.80$. In Table 7 I estimate the discontinuity in loan volume at $R = 0.80$ separately for two groups of loans: those loans that are for purchase of a property in a GSE-qualified census tract (Panel A) and those loans that are for purchase of a property outside of GSE-qualified census tracts (Panel B). If the GSE program has a substantive effect, then the estimates in Panel A should exceed those in Panel B.

First, notice that only about 16% of loan volume near the CRA cutoff (4,106 loans out of 25,485) occurs in GSE-qualified tracts, implying that the GSE effect would have to be extremely large in order to explain the discontinuity in loan volume. The results in columns 1-4 do suggest that the discontinuity may be larger in GSE-qualified tracts, but in the remainder of the columns where a control function is included in the regressions the estimates in Panel A are smaller and insignificant relative to those in Panel B. These results along with the reasoning above suggest the GSE Act has very little, if anything, to do with the discontinuity in bank lending.

2.5.4. CRA's Effect on Applications and Dollars Lent

The final empirical exercise explores the effect of CRA on loan applications and dollars lent at the cutoff. Again, I will focus on the largest MSA's between 1998 and 2002 where a substantive discontinuity in originations is found.

Panels A and B in Table 8 display estimates of the discontinuity in applications received by banks and mortgage company subsidiaries of banks, respectively. I include all applications a lender receives and records in HMDA including those that are withdrawn prior to a credit decision being made. The point estimates are very similar to those found for Table 4 for total originations. Since the outcome variables have been log-transformed, the results imply that applications are increasing at the about the same rate as originations. This finding is important because it indicates that applications are endogenous. Researchers may be tempted to use the origination *rate* as the outcome variable (e.g. Berry and Lee 2008), but such a strategy will typically generate misleading results if banks react to CRA by soliciting more applications or if application volume and quality responds to banks' increased willingness to lend in response to CRA.

Panels C and D display the estimates of CRA's effect on total dollars originated by banks and their subsidiaries. For banks, all of the specifications generate statistically significant results and with the exception of specifications 3 and 4 all of the estimates lie around 0.06. This magnitude is similar to the effect on loan volume, which suggests that the loan amount of marginal loans is similar to the average loan. This result may seem somewhat surprising – if banks are making more loans to marginal borrowers, one might expect them to provide smaller loans to mitigate this risk. However, a larger loan amount than might otherwise be approved given the risk characteristics of the borrower is also a way banks may increase their lending. In other words, there is no clear expectation for whether CRA's effect (in proportional terms) on dollars lent should be bigger or smaller than its effect on originations.

That the average loan amount is not changing at the cutoff implies that the loan-to-income ratio is also not changing (since income is virtually the same on either side of the cutoff). Yet, although this risk factor is not changing, marginal loans are less likely to be sold to secondary market purchasers as indicated by the results in Table 6 (Panel C), suggesting that some other risk factor may be changing, such as the loan to property value ratio, or the credit score of marginal borrowers.

2.6. Summary and Discussion

In this paper, I have identified CRA's impact on home purchase lending to low-income borrowers by exploiting a discontinuity in the rules regulators use to judge banks' compliance with CRA. Using local regression discontinuity techniques I find that CRA's impact at the cutoff is on average not statistically different from zero and no larger than 3%. Exploring the data more finely, I do find a robust jump in loan volume of about 6% at the cutoff in the largest MSA's during a 5-6 year period after CRA was strengthened and merger activity was at a peak. These results mirror those in Bhutta (2008a), which finds that CRA's effect on lending to low and moderate income *neighborhoods* is concentrated in the largest MSA's between 1997 and 2002.

Interpreting this discontinuity as a causal effect of CRA is bolstered by the absence of a discontinuity for unregulated independent mortgage companies in large MSAs during the same period. Further, I find that the relationship between loan volume and the assignment variable is otherwise smooth – that is, I do not find discontinuities at points away from the cutoff, suggesting that the one found at the cutoff is not spurious.

Importantly, I do not find evidence of crowd-out or gaming behavior that has been suggested elsewhere. In particular, mortgage company subsidiaries of banks do not show a jump in *non*-LMI loan volume at the cutoff, nor is there evidence that independent mortgage companies are crowded out by the increased lending by banks in the large MSA's.

These results may underestimate CRA's true impact if banks have reacted to CRA in a "smooth" way. In other words, CRA may have changed the slope of the relationship between loan volume and the assignment variable over time rather than inducing a large "jump" in lending at the cutoff. Another shortcoming of the regression discontinuity design is that it says little about CRA's effect away from the cutoff.

The 6% impact for banks in large MSA's implies a level impact at the cutoff of about 1,400 home purchase loans at a value of approximately $200 million. If one assumes this effect applies up to five percentage points away from the cutoff, the level impact would measure approximately 4,600 loans at a value of $630 million. And if one assumes a constant effect of 6% for all loan volume below the cutoff (i.e. observed loan volume below the cutoff in large MSAs is 6% higher than it would be in the absence of CRA), this would amount to nearly 35,000 loans at a value of $3.9 billion. Importantly, even under this generous assumption and assuming that all of these marginal loans are for first time homeownership, CRA's impact on the homeownership rate would be negligible. Consistent with this conclusion that CRA has likely had little to do with the rise in homeownership rates, I find essentially no difference in the growth rate of homeownership between large and small MSAs between 1994 and 2004 (see Figure 8).

References

Avery, Robert B., Raphael W. Bostic, and Glenn B. Canner. 2000. "CRA Special Lending Programs" *Federal Reserve Bulletin*, 86(11): 711-731.

Avery, Robert B., Kenneth P. Brevoort, and Glenn B. Canner. 2007. "Opportunities and Issues in using HMDA Data" *Journal of Real Estate Research*, 29(4): 351-380.

Barr, Michael S. 2005. "Credit Where it Counts: The Community Reinvestment Act and its Critics" *New York University Law Review*, 80(2): 513-652.

Belsky, Eric S., Matthew Lambert, Alexander Von Hoffman, and Nicolas P. Retsinas. 2000. "Insights into the Practice of Community Reinvestment Act Lending: A Synthesis of CRA Discussion Groups" , Joint Center for Housing Studies Working Paper CRA00-1.

Bernanke, Ben S. 2007. "The Community Reinvestment Act: Its Evolution and New Challenges." Speech at The Community Affairs Research Conference, Washington, DC.

Berry, Christopher R. and Sarah L. Lee. 2007. "The Community Reinvestment Act: A Regression Discontinuity Analysis" , Harris School Working Paper Series 07.04.

Bhutta, Neil. 2008a. "Giving Credit Where Credit is due? the Community Reinvestment Act and Mortgage Lending in Low Income Neighborhoods." .

------. 2008b. "Regression Discontinuity Estimates of the Effects of the GSE Act of 1992" Ph.D. MIT.

Bostic, Raphael W. and Breck L. Robinson. 2005. "What Makes Community Reinvestment Act Agreements Work? A Study of Lender Responses" *Housing Policy Debate*, 16(3/4): 513-545.

Bostic, Raphael W. and Brian J. Surette. 2001. "Have the Doors Opened Wider? Trends in Homeownership Rates by Race and Income" *The Journal of Real Estate Finance and Economics*, 23(3): 411-34.

Bostic, Raphael and Brian J. Surette. 2004. "Market Forces Or CRA-Induced Externalities: What Accounts for the Increase in Mortgage Lending to Lower-Income Communities?" *Lusk Center for Real Estate Working Paper No. 2004-1013.*

Bowyer, Jerry. 2008. "Don't Blame the Markets" *The New York Sun*(April 18).

Chambers, Matthew, Carlos Garriga, and Don E. Schlagenhauf. 2007. "Accounting for Changes in the Homeownership Rate" (2007-21).

Chen, David W. 2004. "U.S. Set to Alter Rules for Banks Lending to Poor" *New York Times*(October 20).

Chomsisengphet, Souphala and Anthony Pennington-Cross. 2006. "The Evolution of the Subprime Mortgage Market" *Federal Reserve Bank of St.Louis Review*, 88(1): 31-56.

Doms, Mark and John Krainer. 2007. "Innovations in Mortgage Markets and Increased Spending on Housing" (2007-05).

Fishbein, Allen J. 1992. "The Ongoing Experiment with 'Regulation from Below': Expanded Reporting Requirements for HMDA and CRA" *Housing Policy Debate*, 3(2): 601-636.

Gabriel, Stuart A. and Stuart Rosenthal. October, 2003. "The Causes of Increased Homeownership in the 1990's" , Working Paper No. 03-01.

Garriga, Carlos, William T. Gavin, and Don Schlagenhauf. 2006. "Recent Trends in Homeownership" *Review*: 397-412.

Goldberg, Deborah B. 2000. "The Community Reinvestment Act and the Modernized Financial Services World" *ABA Bank Compliance*, 21(1): 13.

Hagerty, James R. and Glenn R. Simpson. April 30, 2008. "Countrywide Loss Focuses Attention on Underwriting" *The Wall Street Journal*, B.

Imbens, Guido W. and Thomas Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice" *Journal of Econometrics,* 142(2): 615-635.

Jackson, Kenneth T. 1985. *Crabgrass Frontier: The Suburbanization of the United States*. New York: Oxford University Press.

Joint Center for Housing Studies. 2002. "The 25th Anniversary of the Community Reinvestment Act: Access to Capital in an Evolving Financial Services System" , Ford Foundation Sponsored Report.

Lee, David S. and David Card. 2008. "Regression Discontinuity Inference with Specification Error" *Journal of Econometrics,* 142(2): 655-674.

Ludwig, Jens and Douglas L. Miller. 2007. "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design*" *Quarterly Journal of Economics*, 122(1): 159-208.

Marsico, Richard D. 2006. "The 2004-2005 Amendments to the Community Reinvestment Act Regulations: For Communities One Step Forward and Three Steps Back
" *Clearinghouse Review*, 39.

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*, 12(1): 29-59.

Quercia, Roberto G., Michael A. Stegman, Walter R. Davis, and Eric Stein. 2001. "Community Reinvestment Lending: A Description and Contrast of Loan Products and their Performance" , Joint Center for Housing Studies Working Paper LIHO-01.11.

Straka, John W. 2000. "A Shift in the Mortgage Landscape: The 1990's Move to Automated Credit Evaluations" *Journal of Housing Research*, 11(2).

Figure 1

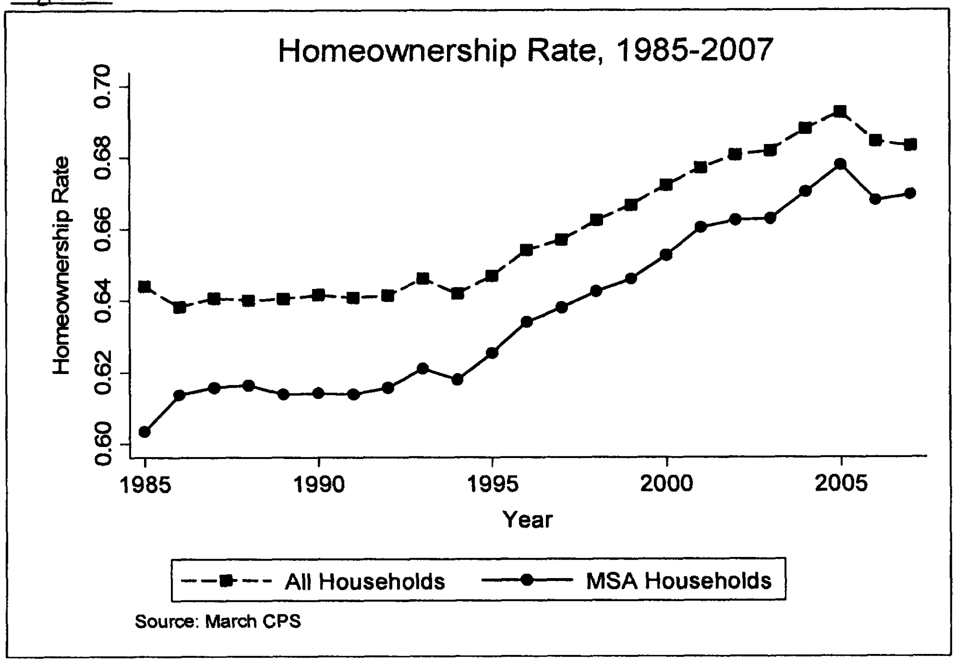
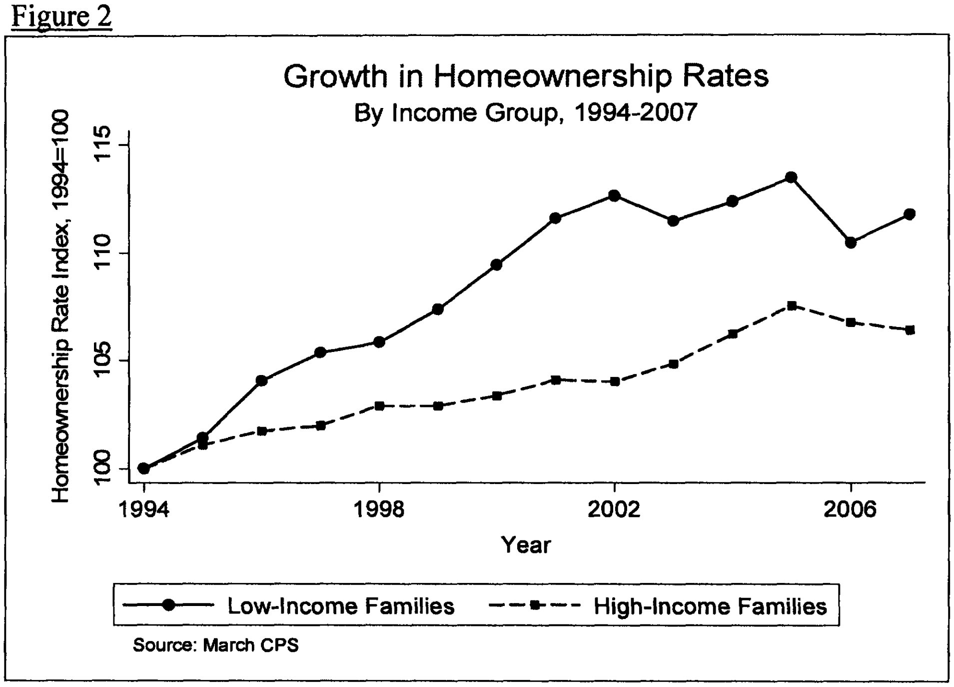
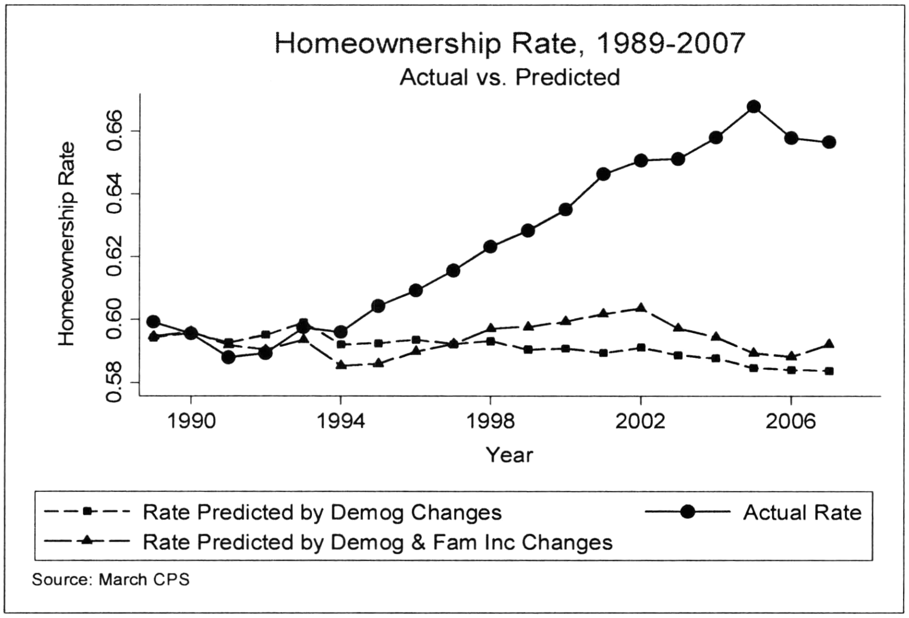


Figure 2



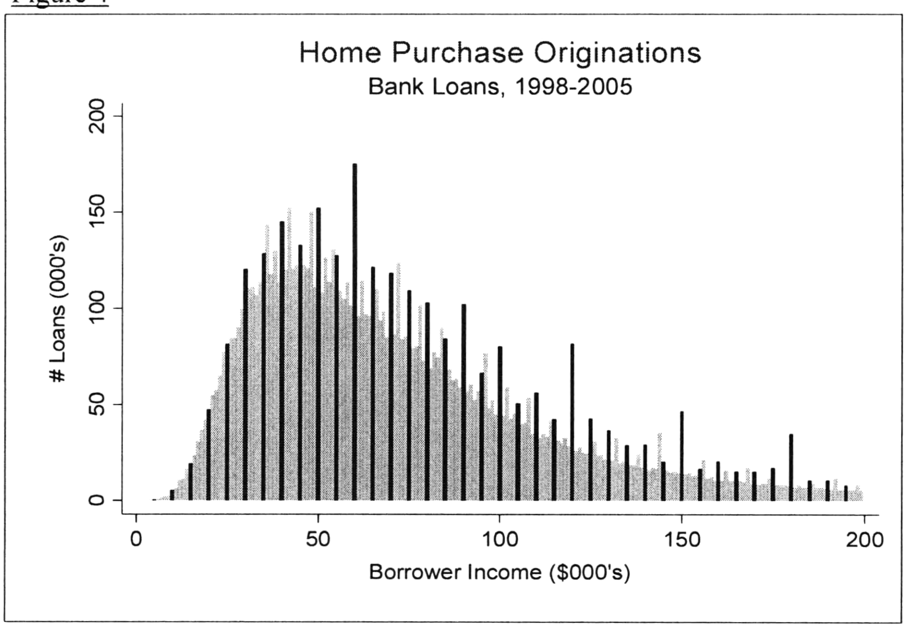
Notes: Sample includes households in an MSA. Low (high) income families defined as those with family income (income of household head plus that of spouse) below (above) the median family income in the family's state and year of survey; state-by-year median income calculated by author from CPS

Figure 3



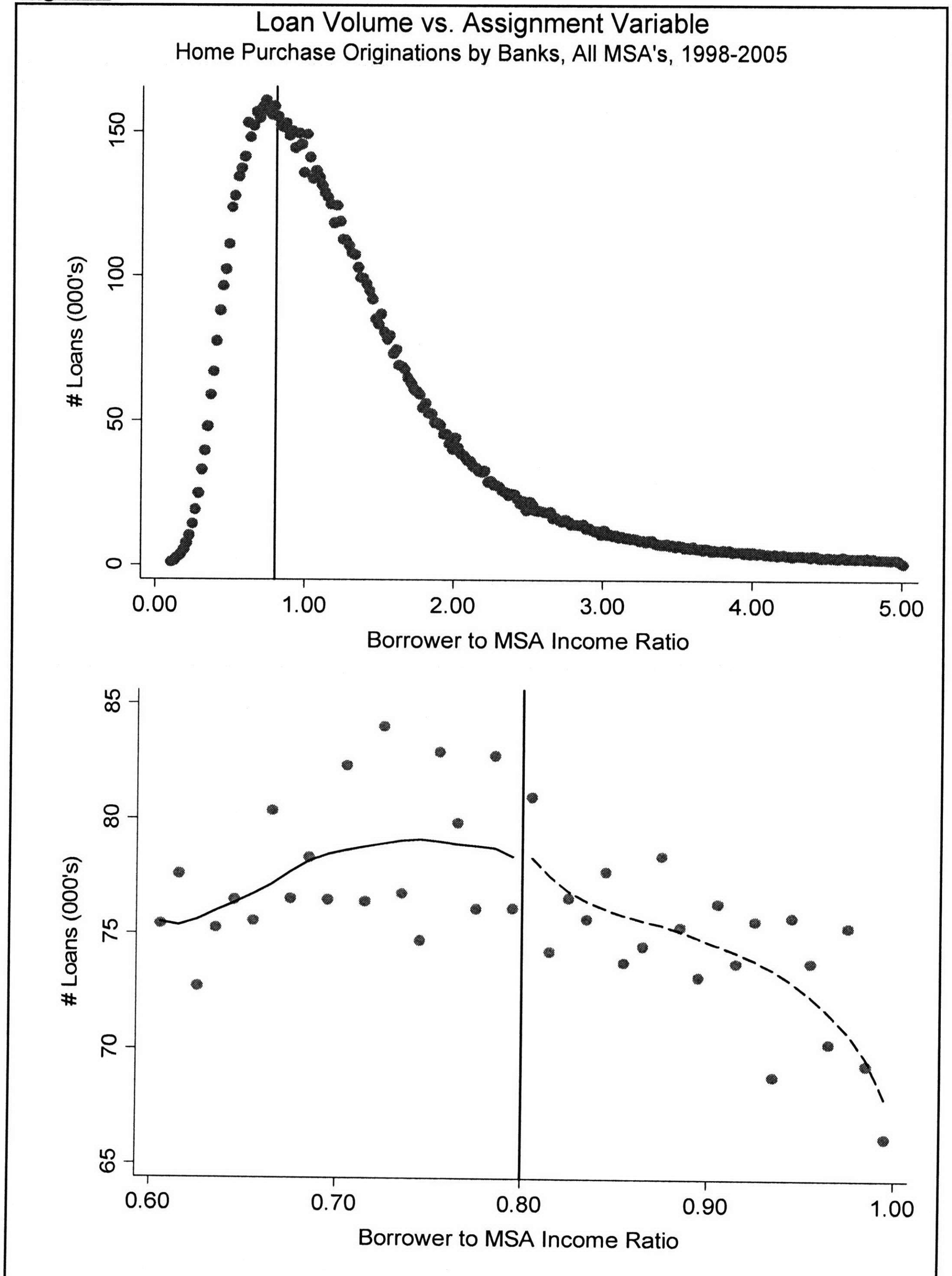
Notes: Sample includes households in an MSA with household head aged between 24 and 60. Predicted homeownership rates calculated by applying coefficients estimated from a linear regression of homeownership dummy on household head characteristics using a pooled sample from 1989 to 1991 to household head characteristics in subsequent years. Characteristics used are dummy variables for college-educated, married, black race, hispanic ethnicity, female and eight age groups. These variables are fully interacted in one regression and interacted with family income (defined as income of household head plus spouse) in the other. Family income is adjusted for inflation using CPI-U.

Figure 4



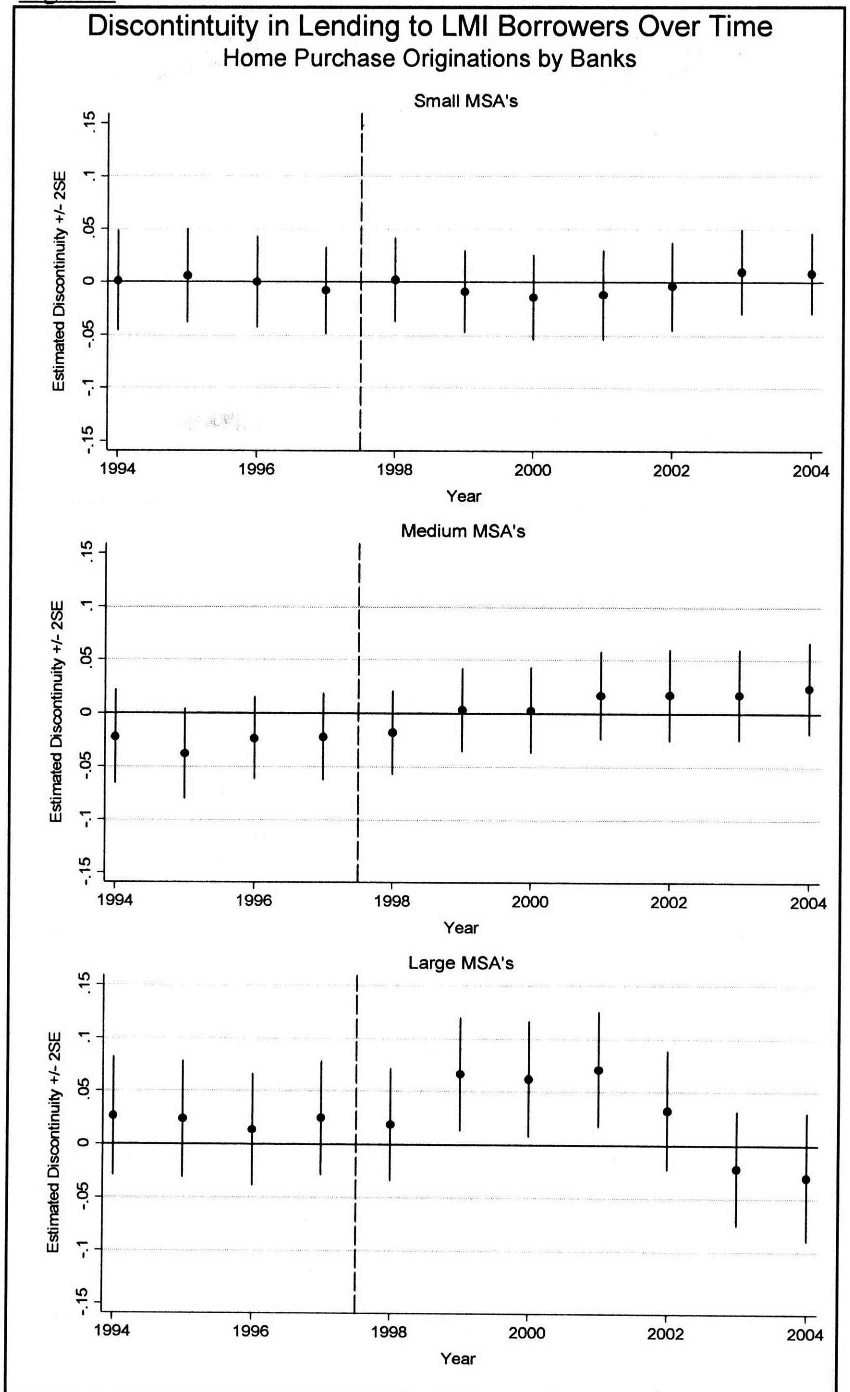
Notes: displays total number of home purchase loans originated by banks between 1998 and 2005 in MSA's by income in $1,000 increments. Black bars indicate incomes divisible by $5000.

Figure 5



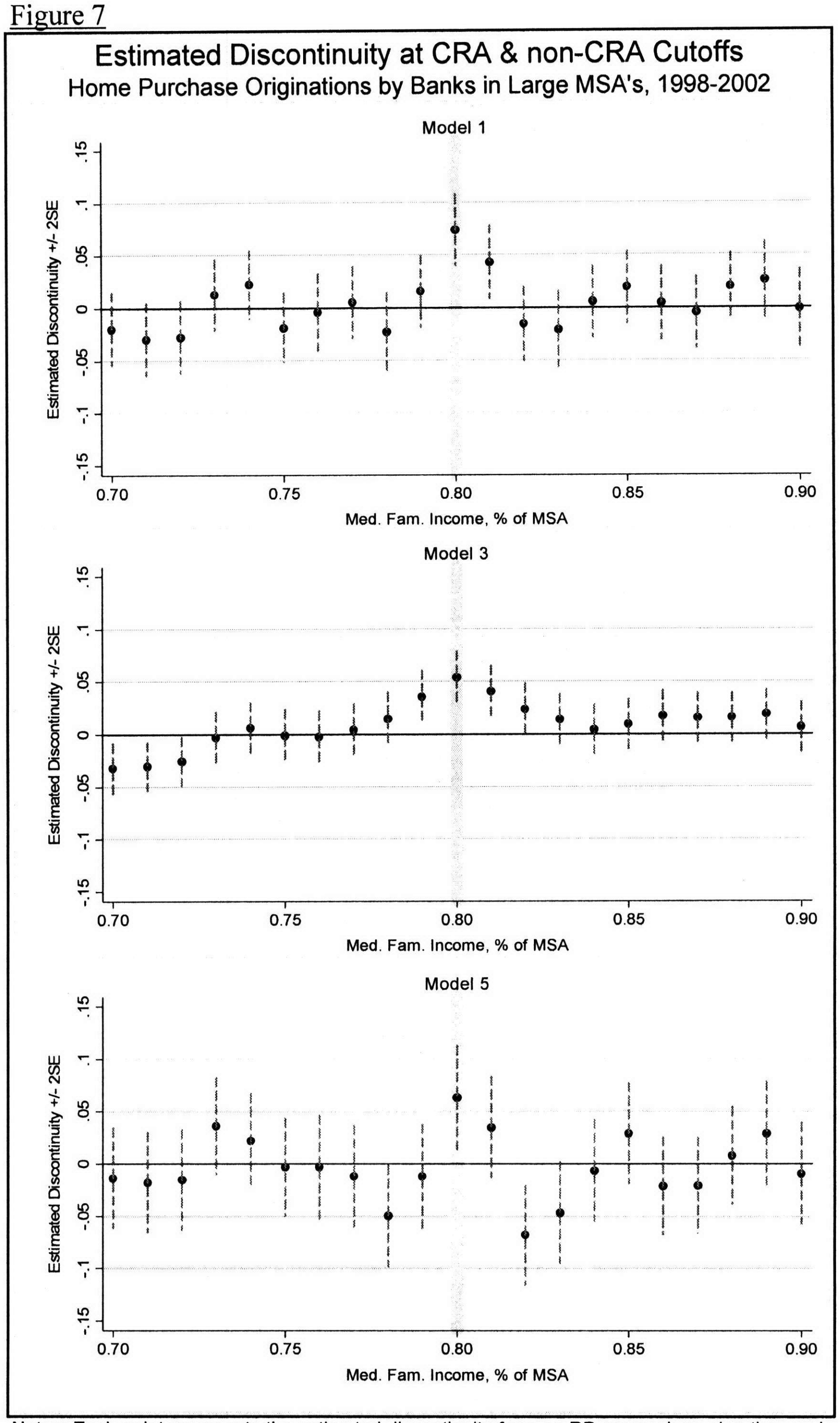
Notes: Top panel shows home purchase loan volume by RD running variable with a line drawn at the CRA cutoff (0.80). Bottom panel zooms in on area around cutoff and displays local linear estimates of relationship between loan volume and running variable on either side of the cutoff. Data points in the top (bottom) panel represents loan volume in a two (one) percentage point bin of the running variable

Figure 6



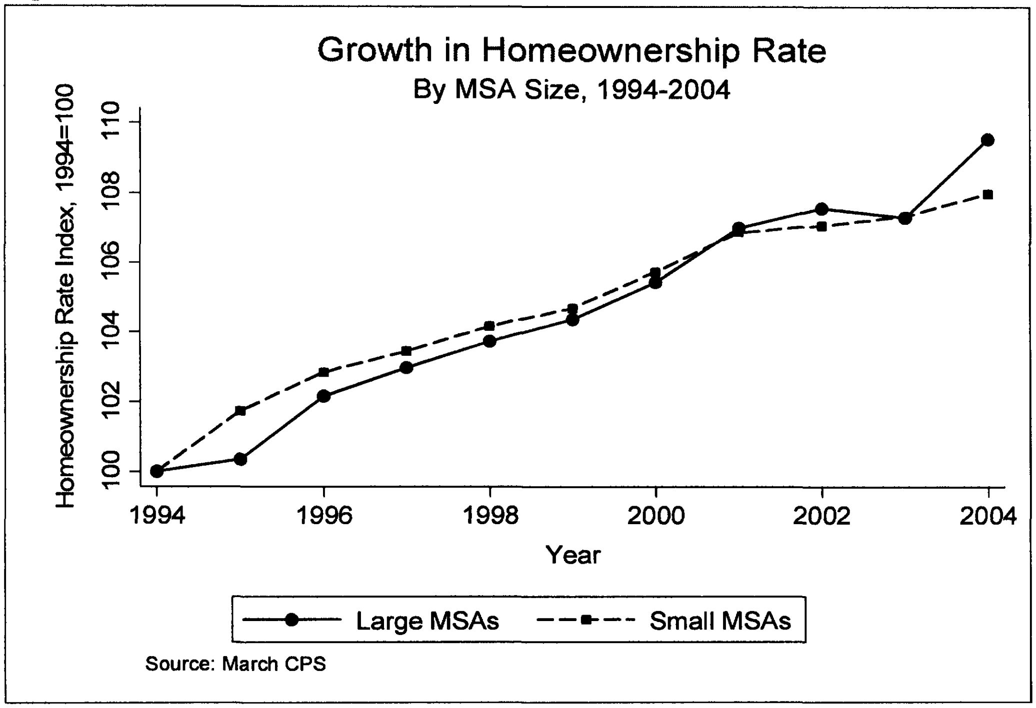
Notes: Each point represents the estimated discontinuity from an RD regression corresponding to specification 6 in Table 2 for the sample of MSA's described above each graph. Each regression uses four years of loan data: the year on the axis plus 2 years before and one year after. MSA size groups based on population in 1990 Census: small MSA's are those with population less than 500k, medium MSA's are those with population between 500k and 2mil, and large MSA's are those with population over 2mil. Groups are not a perfectly consistent set of geographic areas because of major changes in MSA definitions used in the 2004 and 2005 data.

Figure 7



Notes: Each point represents the estimated discontinuity from an RD regression using the x-axis point as the cutoff. The top, middle and bottom panels use specifications 1, 3 and 5, respectively, from Table 4 for each regression.

Figure 8



Notes: Large MSAs comprised of the largest 20 MSAs in terms of total households (as measured in the CPS) in 1994 while the remainder of MSAs make up the set of small MSAs.

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
HOEPA Status	2004-2005	Indicator for high-cost loan: APR at consummation exceeds yield for comparable Treasury by more than 8 percentage points.
Lien Status	2004-2005	Loan secured by first or subordinate lien
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	2004-2005	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:
Summary Statistics: HMDA and Comparable Estimates from Census and CPS

A. HMDA Data Summary Statistics[1]

	Levels, 1999			Annual %Change, 1994-2002		
	All Borrowers	LMI Borrowers	non-LMI Borrowers	All Borrowers	LMI Borrowers	non-LMI Borrowers
No. of Loans Reported (000's)	3,664	1,175	2,489	4.38	5.80	3.77
Med. Family Income ($)	69,681	39,818	88,345	1.05	1.36	1.57
Med. Loan Amount ($)	136,873	94,567	164,247	1.69	3.10	1.83
Prop. Borrowers Black	0.070	0.113	0.050	-0.53	-0.36	-1.26
	(0.006)	(0.010)	(0.004)	(0.33)	(0.49)	(0.43)
Prop. Borrowers Hispanic	0.081	0.118	0.064	3.15	3.55	2.51
	(0.009)	(0.013)	(0.009)	(0.38)	(0.61)	(0.47)
Prop. Borrowers Female	0.233	0.372	0.167	2.49	1.03	3.16
	(0.003)	(0.004)	(0.004)	(0.20)	(0.23)	(0.29)
Prop. Borrowers w/ Co-Applicant	0.517	0.285	0.627	-3.44	-5.02	-2.81
	(0.005)	(0.005)	(0.007)	(0.07)	(0.19)	(0.07)
Prop. Loans FHA	0.197	0.308	0.145	-0.46	0.19	-1.90
	(0.007)	(0.010)	(0.006)	(0.20)	(0.27)	(0.27)
Prop. Loans by Banks	0.328	0.298	0.342	-2.57	-3.01	-2.35
	(0.010)	(0.011)	(0.010)	(0.11)	(0.21)	(0.13)
1990 Med. Home Value in Census	196,067	149,705	217,952	0.00	-0.14	0.22
Tract of Property ($)	(11,218)	(6,861)	(13,155)	(0.00)	(0.10)	(0.07)

B. Comparable Statistics from Census & CPS Data

	Levels, 1999 (Census)			Annual %Change, 1992-2002 (CPS)[2]		
	All Buyers	LMI Buyers	non-LMI Buyers	All Buyers	LMI Buyers	non-LMI Buyers
No. Purchased (000's)[3]	4,669	1,569	3,100	1.36	2.70	0.96
Med. Family Income ($)	70,552	35,089	93,322	0.62	1.60	1.55
Med. House Value ($)	165,527	114,364	195,623	na	na	na
Prop. HH Head Black	0.086	0.121	0.069	0.38	-0.19	0.69
	(0.007)	(0.010)	(0.005)	(0.63)	(1.01)	(1.19)
Prop. HH Head Hispanic	0.103	0.141	0.084	4.44	4.90	5.56
	(0.011)	(0.014)	(0.010)	(0.76)	(1.23)	(1.38)
Prop. HH Head Female	0.271	0.415	0.198	4.72	2.60	9.41
	(0.003)	(0.004)	(0.004)	(0.35)	(0.44)	(0.84)
Prop. Married HH	0.637	0.369	0.772	-0.55	-1.33	-0.09
	(0.004)	(0.006)	(0.004)	(0.14)	(0.43)	(0.15)
Prop. HH Head Age<35	0.386	0.45	0.354	0.06	0.31	-0.16
	(0.005)	(0.006)	(0.005)	(0.24)	(0.45)	(0.39)
Prop. HH Head College+	0.397	0.234	0.479	1.61	4.04	1.81
	(0.009)	(0.009)	(0.010)	(0.27)	(0.89)	(0.33)
Prop. Units Condominium	0.089	0.122	0.073	0.55	-0.41	1.28
	(0.007)	(0.010)	(0.006)	(0.51)	(0.81)	(1.01)
Prop. Units Mobile Home	0.066	0.124	0.037	-4.69	-5.91	-8.64
	(0.005)	(0.009)	(0.003)	(0.66)	(0.87)	(1.46)

Notes: Standard errors in parenthesis. All dollar figures adjusted to 2007 dollars using the CPI-U. (1) HMDA levels summary statisitics based on 5% random sample of 1999 owner-occupied home purchase originations in MSA's identified in the 2000 Census; HMDA %change statistics based on 1% random samples from 1994 and 2002 owner-occupied home purchase originations in MSA's identified in the CPS. (2) Combines data from 1992, 1993, 1994, 2002, 2003 & 2004 rounds of March CPS. (3) Census columns use only buyers with a mortgage; CPS does not seperate out 'free & clear' owners.

Table 3:

RD Estimates of CRA's Effect on Bank Home Purchase Loans to LMI Borrowers

All MSA's, 1998-2005

Dependent variable: (log) number of originations in (MSA)x(Year)x(Income) cells

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	h = 0.02		h = 0.04				h = 0.06					
1[R<80]	0.0255***	0.0223***	0.0318***	0.0250***	0.0061	0.0033	0.0314***	0.0286***	0.0183**	0.0111	-0.0102	-0.0132
	(0.0077)	(0.0085)	(0.0052)	(0.0055)	(0.0106)	(0.0101)	(0.0042)	(0.0046)	(0.0086)	(0.0084)	(0.0131)	(0.0125)
R'					-0.0112***	-0.0128***			-0.0067***	-0.0090***	-0.0223***	-0.0254***
					(0.0032)	(0.0037)			(0.0018)	(0.0020)	(0.0069)	(0.0074)
(R')***1**[R<80]					0.0093**	0.0103**			0.0089***	0.0093***	0.0121	0.0119
					(0.0045)	(0.0051)			(0.0024)	(0.0029)	(0.0097)	(0.0107)
$(R')^2$											0.0027**	0.0034**
											(0.0011)	(0.0014)
$(R')^2$***1**[R<80]											-0.0047***	-0.0062***
											(0.0016)	(0.0020)
Kernel Weights	rect	tri	rect	tri	rect	tri	rect	tri	rect	tri	rect	tri
R-Squared	0.968	0.978	0.951	0.956	0.951	0.956	0.945	0.949	0.946	0.949	0.946	0.949
N	5592	5476	11257	11115	11257	11115	16903	16721	16903	16721	16903	16721

Notes: Robust standard errors in parentheses; * p<0.10, ** p<0.05, *** p<0.01. All regressions include MSA x Year fixed effects and dummy variables indicating for borrower income at multiples of $5,000 and borrower income of $60,000. Loan volume has been log-transformed so coefficients should be interpretted as semi-elasticities.

Table 4:
RD Estimates of CRA's Effect on Home Purchase Loans to LMI Borrowers, by Lender Type
Large MSA's, 1998-2002

Dependent variable: (log) number of originations in (MSA)x(Year)x(Income) cells

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	h = 0.02		*h* = 0.04				*h* = 0.06			
					A. Banks (Regulated Lenders) (loan volume at cutoff[1] = 24,082)					
1[R<80]	0.0740***	0.0680***	0.0541***	0.0581***	0.0632**	0.0679***	0.0644***	0.0655***	0.0633**	0.0608**
	(0.0169)	(0.0192)	(0.0122)	(0.0126)	(0.0249)	(0.0239)	(0.0200)	(0.0198)	(0.0313)	(0.0297)
N	281	275	561	551	561	551	842	834	842	834
					B. Mortgage Company Subsidiaries (Voluntarily Regulated Lenders) (loan volume at cutoff[1] = 34,086)					
1[R<80]	0.0345**	0.0369**	0.0216**	0.0252**	0.0322*	0.0332	0.0241	0.0297*	0.0369	0.0321
	(0.0147)	(0.0165)	(0.0093)	(0.0105)	(0.0195)	(0.0203)	(0.0159)	(0.0162)	(0.0247)	(0.0246)
N	281	275	561	551	561	551	842	834	842	834
					C. Independent Mortgage Companies (Unregulated Lenders) (loan volume at cutoff[1] = 35,767)					
1[R<80]	0.0093	0.0197	-0.0012	0.0037	0.0155	0.0204	0.0058	0.0112	0.0250	0.0294
	(0.0160)	(0.0176)	(0.0105)	(0.0116)	(0.0216)	(0.0220)	(0.0177)	(0.0179)	(0.0279)	(0.0271)
N	281	275	561	551	561	551	842	834	842	834
Control Function	none	none	none	none	linear	linear	linear	linear	quadratic	quadratic
Kernal Weights	rect	tri	rect	tri	rect	tri	rect	tri	rect	tri

Notes: Robust standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01. All regressions include MSA x Year fixed effects, dummy variable indicating for borrower income at multiples of $5,000 and another dummy variable indicating for borrower income of $60,000. Loan volume has been log-transformed so coefficients should be interpretted as semi-elasticities. (1) loan volume at cutoff estimated by calculating mean loan volume in MSA by year cells with *R* between 0.80 & 0.82 and then multiplying this mean by the number of MSA's (23) and years (5) in the sample.

Table 5:
Characteristics of Applications Around CRA Cutoff w/ High & Low Predicted Denial Probability
Large MSA's, 2000

	All Applications[1]	(Predicted) Highest Risk Applications[2]	(Predicted) Lowest Risk Applications[2]
No. Applications	92,779	18,397	18,389
Proportion Applications Denied	0.187	0.371	0.085
	(0.014)	(0.021)	(0.010)
Characteristics of Originated Applications			
Number[3]	67,506	9,321	15,708
Borrower Income ($ 000's)	60.64	61.30	60.50
	(2.17)	(2.33)	(2.15)
Amount-to-Income Ratio	2.20	1.75	2.49
	(0.07)	(0.14)	(0.06)
1990 Med. House Value in Census Tract of Purchased Property ($ 000's)	204.06	184.18	249.27
	(15.52)	(15.74)	(20.17)
Race			
Asian (Proportion)	0.074	0.025	0.119
	(0.011)	(0.006)	(0.029)
Black	0.075	0.279	0.014
	(0.009)	(0.046)	(0.007)
Hispanic	0.117	0.13	0.037
	(0.024)	(0.018)	(0.014)
Other[4]	0.098	0.312	0.016
	(0.010)	(0.036)	(0.005)
Applicant Sex			
Female	0.298	0.267	0.415
	(0.011)	(0.013)	(0.039)
NA - Internet/Mail Application	0.054	0.242	0.004
	(0.008)	(0.032)	(0.002)
Co-Applicant Sex			
Male	0.058	0.112	0.012
	(0.002)	(0.013)	(0.005)
Female	0.362	0.232	0.376
	(0.014)	(0.018)	(0.049)
NA - Mail/Internet Application	0.028	0.16	0.001
	(0.004)	(0.023)	(0.000)
No Co-Applicant	0.552	0.496	0.611
	(0.015)	(0.035)	(0.049)
FHA-Insured Loan	0.097	0.049	0.228
	(0.008)	(0.014)	(0.037)
Loan Sold into Secondary Market[5]	0.518	0.435	0.62
	(0.017)	(0.020)	(0.021)

Notes: Standard errors clustered at MSA-level in parentheses; dollar values adjusted to 2007 dollars using CPI-U. (1) Sample includes home purchase applications from 2000 HMDA with *BM* between 70 & 90 in 23 largest MSA's. (2) Highest (lowest) risk applications are those with predicted probability of denial in the top (bottom) 20%; see text for details on how denial probability is predicted. (3) Number originated is less than the number not denied because some borrowers do not accept the lender's offer. (4) 'Other' comprised mostly of mail/internet applicants not reporting race (5) 'Loans Sold' not used to predict denial probability as it is observed only for originated loans.

Table 6:
RD Estimates of CRA's Effect on Loan and Borrower Characteristics
Large MSA's, 1998-2002

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	h = 0.02		h = 0.04				h = 0.06			
					A. outcome: (log) Proportion of 'High Risk' Loans[1] (mean[2] just above cutoff = 0.14)					
1[R<80]	0.0170	0.0189	0.0268*	0.0203	0.0040	-0.0023	0.0333	0.0239	0.0091	-0.0144
	(0.0215)	(0.0227)	(0.0161)	(0.0163)	(0.0321)	(0.0307)	(0.0255)	(0.0252)	(0.0398)	(0.0389)
N	281	275	561	551	561	551	842	834	842	834
					B. outcome: (log) Proportion of 'Low Risk' Loans[1] (mean[2] just above cutoff = 0.23)					
1[R<80]	-0.0403**	-0.0441**	-0.0357***	-0.0366***	-0.0306	-0.0268	-0.0414**	-0.0337*	-0.0153	-0.0051
	(0.0180)	(0.0189)	(0.0119)	(0.0132)	(0.0257)	(0.0252)	(0.0199)	(0.0204)	(0.0313)	(0.0313)
N	281	275	561	551	561	551	842	834	842	834
					C. outcome: (log) Proportion of Loans Sold into Secondary Market (mean[2] just above cutoff = 0.53)					
1[R<80]	-0.0417***	-0.0430***	-0.0387***	-0.0383***	-0.0416***	-0.0428***	-0.0396***	-0.0382***	-0.0301*	-0.0426***
	(0.0095)	(0.0111)	(0.0068)	(0.0072)	(0.0136)	(0.0137)	(0.0109)	(0.0110)	(0.0161)	(0.0164)
N	281	275	561	551	561	551	842	834	842	834
Control Function	none	none	none	none	linear	linear	linear	linear	quadratic	quadratic
Kernal Weights	rect	tri	rect	tri	rect	tri	rect	tri	rect	tri

Notes: Robust standard errors in parentheses; * p<0.10, ** p<0.05, *** p<0.01. All regressions include MSA x Year fixed effects, dummy variable indicating for borrower income at multiples of \$5,000 and another dummy variable indicating for borrower income of \$60,000. (1) Loan risk estimated as the predicted liklihood of application denial given borrower and loan characteristics. High (low) risk loans represent those in the upper (lower) 20% of the estimated 'risk' distribution. See text for more details. (2) Means calculated using MSA by year cells with R between 0.80 & 0.82.

Table 7:
RD Estimates of CRA's Effect on Home Purchase Loans to LMI Borrowers, Inside and Outside GSE-Qualified Census Tracts
Large MSA's, 1998-2002

Dependent variable: (log) number of originations in (MSA)x(Year)x(Income) cells

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	h = 0.02		h = 0.04				h = 0.06			
A. Sample: Loans Inside GSE-qualified Census Tracts (loan volume at cutoff[1] = 4,106)										
1[R<80]	0.0763***	0.0692**	0.1009***	0.0782***	0.0161	0.0221	0.0472	0.0304	0.0098	0.0058
	(0.0269)	(0.0301)	(0.0191)	(0.0199)	(0.0387)	(0.0379)	(0.0309)	(0.0307)	(0.0483)	(0.0471)
N	281	275	561	551	561	551	842	834	842	834
B. Sample: Loans Outside of GSE-qualified Census Tracts (loan volume at cutoff[1] = 21,379)										
1[R<80]	0.0667***	0.0613***	0.0405***	0.0488***	0.0655**	0.0701***	0.0640***	0.0666***	0.0658**	0.0653**
	(0.0173)	(0.0190)	(0.0126)	(0.0129)	(0.0255)	(0.0242)	(0.0208)	(0.0202)	(0.0319)	(0.0300)
N	281	275	561	551	561	551	842	834	842	834
Control Function	none	none	none	none	linear	linear	linear	linear	quadratic	quadratic
Kernal Weights	rect	tri	rect	tri	rect	tri	rect	tri	rect	tri

Notes: Robust standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01. All regressions include MSA x Year fixed effects, dummy variable indicating for borrower income at multiples of $5,000 and another dummy variable indicating for borrower income of $60,000. Loan volume has been log-transformed so coefficients should be interpretted as semi-elasticities. (1) Loan volume at cutoff estimated by calculating mean loan volume in MSA by year cells with R between 0.80 & 0.82 and then multiplying this mean by the number of MSA's (23) and years (5) in the sample.

Table 8:
RD Estimates of CRA's Effect on Application Volume and Dollars Originated
Large MSA's, 1998-2002

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	h = 0.02		h = 0.04				h = 0.06			
				A. outcome: (log) # Applications at Banks (loan volume at cutoff[1] = 34,177)						
1[R<80]	0.0795***	0.0753***	0.0580***	0.0616***	0.0630**	0.0713***	0.0589***	0.0648***	0.0681**	0.0647**
	(0.0173)	(0.0194)	(0.0125)	(0.0129)	(0.0251)	(0.0242)	(0.0206)	(0.0202)	(0.0321)	(0.0300)
N	281	275	561	551	561	551	842	834	842	834
				B. outcome: (log) # Applications at Mortgage Company Subsidiaries (loan volume at cutoff[1] = 44,953)						
1[R<80]	0.0395**	0.0418**	0.0256**	0.0283**	0.0332	0.0382*	0.0206	0.0290	0.0395	0.0388
	(0.0161)	(0.0186)	(0.0102)	(0.0116)	(0.0217)	(0.0224)	(0.0175)	(0.0181)	(0.0275)	(0.0272)
N	281	275	561	551	561	551	842	834	842	834
				C. outcome: (log) Dollars Originated by Banks (dollars lent at cutoff[1] = $3.3 billion)						
1[R<80]	0.0604***	0.0557***	0.0253**	0.0369***	0.0597**	0.0664***	0.0613***	0.0611***	0.0583*	0.0592*
	(0.0178)	(0.0205)	(0.0127)	(0.0132)	(0.0259)	(0.0250)	(0.0207)	(0.0205)	(0.0322)	(0.0306)
N	281	275	561	551	561	551	842	834	842	834
				D. outcome: (log) Dollors Originated by Mortgage Companies Subsidiaries (dollars lent just above cutoff[1] = $5.1 billion)						
1[R<80]	0.0245*	0.0291*	-0.0089	0.0075	0.0421**	0.0375*	0.0280*	0.0353**	0.0439*	0.0341
	(0.0147)	(0.0164)	(0.0096)	(0.0106)	(0.0194)	(0.0202)	(0.0160)	(0.0161)	(0.0243)	(0.0241)
N	281	275	561	551	561	551	842	834	842	834
Control Function	none	none	none	none	linear	linear	linear	linear	quadratic	quadratic
Kernal Weights	rect	tri	rect	tri	rect	tri	rect	tri	rect	tri

Notes: Robust standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01. All regressions include MSA x Year fixed effects, dummy variable indicating for borrower income at multiples of $5,000 and another dummy variable indicating for borrower income of $60,000. (1) Volumes estimated by calculating mean loan volume in MSA by year cells with R between 0.80 & 0.82 and then multiplying this mean by the number of MSA's (23) and years (5) in the sample. Dollar amounts adjusted to $2007 using CPI-U.

Chapter 3

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

3.1. Introduction

Congress created Fannie Mae and Freddie Mac to facilitate the primary mortgage market by connecting distant borrowers and investors. Although the two institutions have been reconstituted into private "government-sponsored enterprises" (GSEs), they retain implicit and explicit financial benefits from the government (CBO 2001). One of the justifications for these benefits is that they will be passed on to consumers in the form of lower mortgage interest rates. Published estimates typically imply the GSEs lower mortgage rates by 30 to 50 basis points[1]. However, a recent working paper by Sherlund (2008) suggests the GSE impact has fallen considerably since 1996. And work by Passmore et al (2005) suggests that only a small fraction of the financial benefits accruing to the GSEs are passed through to borrowers.

The GSEs are also expected to help promote homeownership opportunities for low-income households and neighborhoods, as well as minority households and communities. The 1992 Federal Housing Enterprise Financial Safety and Soundness Act (hereafter the GSE Act) formalizes this responsibility. The GSE Act mandates that the US Department of Housing and Urban Development (HUD) create "affordable housing goals" that the GSEs must meet. These goals, adjusted periodically, specify that a certain fraction of GSE mortgage purchases must be of loans made to targeted households and neighborhoods.

The GSE Act is one of many 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[2].

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 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. In that paper, I show that the Community Reinvestment Act not only pushes regulated lenders (banks) to increase their lending in targeted low-income neighborhoods, but also

---

[1] See Quigley (2006) for a review of this literature

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

leads to an increase in lending by *unregulated* lenders (independent mortgage companies) in some targeted neighborhoods – specifically, those areas 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.[3] The initial goals (thick black line) were finalized in December 1995, then were increased in 1997 and then again in 2001 and 2005. Numbers generated from analyses by Manchester (2002, 2007) 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 (and much of the literature regarding the GSE Act and GSE activity) 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 a causal impact on GSE purchase activity and subsequent mortgage credit flow in targeted neighborhoods.[4] 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.[5] 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 semi-parametric regression discontinuity methods similar to those described in Porter (2003) and Imbens and Lemieux (2008) and implemented in Ludwig and Miller's (2007) study of Head Start.

Discerning the causal effect of the GSE goals on GSE behavior and overall credit flow is important for several reasons. For one, knowing the degree to which the goals bind can help policy makers as they reset goals in the future. Policy makers should also be interested in whether increased GSE purchasing activity ultimately affects the mortgage supply. One reason the affordable housing goals would have 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

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

[4] Ongoing work analyzes the effects of the other goals.

[5] 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.

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 Community Reinvestment Act (CRA), which forces banks to provide evidence that they are supplying credit in low-income neighborhoods and low-income households. 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.[6]

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.[7] 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.[8] 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.

Although they find no "first-stage" effect, they still find statistically significant negative second-stage relationships between GSE share and the two outcome variables. This result may occur because they do not execute a standard instrumental variable

---

[6] 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.

[7] 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.

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

regression and therefore may actually be capturing the cross-sectional relationship one would expect (i.e. tracts with more prime loans purchased by the GSEs will tend to have fewer subprime loans).[9]

Regardless of these potential shortcomings, looking at 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):

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

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 described by Imbens and Lemieux (2008) to help mitigate identification concerns.

To preview, I find evidence of a direct effect of the UAG on GSE purchasing activity of 3-4% and evidence of a increase in overall GSE-eligible lending of 2-3%. 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.

It is important to note that the results for GSE purchases and GSE-eligible lending 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 and not statistically significant.

In the next section, I discuss the data used in this paper and the regression discontinuity empirical strategy, and then present summary statistics. Section 3 describes the results and Section 4 summarizes and discusses directions for future research.

[9] To be specific, An and Bostic 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.

3.2. Data & Empirical Strategy

3.2.1. Overview

I take advantage of a sharp discontinuity in the GSE Underserved Areas Housing Goal (UAG) eligibility rule to identify the goal's impact on GSE purchase volume and overall credit flow. Since 1996, the GSEs have been required to make a specified share of their mortgage purchases for loans originated in 'underserved neighborhoods' (see top panel of Figure 1). Census tracts with a median family income no greater than 90% of MSA median family income qualify as 'underserved neighborhoods'. 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 is referred to as *TM*. Therefore, 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 the 1990 Census and 1993 MSA definitions (by OMB).[10] Importantly, the value of *TM* used in this paper is identical to that used by the GSEs and HUD. As such, the key right-hand-side variable in the regressions to follow is measured without error.

3.2.2. 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.[11]

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.[12] 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.[13] 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.

[10] 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.

[11] 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.

[12] 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.

[13] Each year, HUD releases a "subprime lender list". See www.huduser.org .

Census tract-level characteristics from the 1990 Census, 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.

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. Census tract and MSA definitions changed 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 high-priced, high-risk FHA and subprime loans, while the rest are VA and "jumbo" loans. 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.[14] 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.

3.2.3. Empirical Strategy: Regression Discontinuity

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

(2.1)
$$Y_i = \alpha + \beta D_i + e_i$$

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

(2.2)
$$\lim_{h \to 0} \{ \mathrm{E}[e_i \mid 0.90 - h \leq TM_i \leq 0.90] - \mathrm{E}[e_i \mid 0.90 < TM_i \leq 0.90 + h] \} = 0$$

(2.2) 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 as in (2.1), 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 ($K(a) = \mathbf{1}[0 \leq |a| \leq 1]$) as well as a triangular kernel ($K(a) = \mathbf{1}[0 \leq |a| \leq 1] * [1-|a|]$) 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.

14 Overall, though, mortgage lenders tend to specialize in either prime or subprime lending. (see Nichols et al 2005)

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:

$$Y_i = \alpha + \beta D_i + TM'_i + TM'_i * 1[TM_i < 0.90] + \mu_i \tag{2.3}$$

where $TM'_i = TM_i - 0.90$. The term $TM'_i + TM'_i * 1[TM_i < 0.90]$ in (2.3) is often referred to as the "control function". For all regressions, robust standard errors clustered at the MSA level are reported.

3.3. Results

3.3.1. Preliminaries: 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 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).

3.3.2. 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 (2.2) 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 (2.2) 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. On the other hand, 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 linear controls for *TM*) and the bandwidth (h) is 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.[15]

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.

Reporting error by HMDA respondents is another reason non-GSE purchases might *appear* to increase. In particular, the *purchaser* variable may be reported with considerable error (Scheesele 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.

3.3.3. 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 between 1997 and 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

[15] The point estimates in columns 9 and 11 rise slightly with the inclusion a linear control function.

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.

3.3.4. 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.[16]

3.4. Summary & Discussion

In this paper, I identify the impact of the Underserved Areas Goal 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%. Unlike previous research, I find no evidence of an offsetting reduction in "GSE-ineligible" lending, which is comprised mainly of FHA and subprime loans. At the same time, that FHA and subprime lending does not increase bolsters the view that the discontinuity in GSE-eligible lending is not spurious. Additionally, the lack of discontinuities at points away from the cutoff (Figure 4) also supports the causal interpretation of the cutoff. One possible weakness of the

[16] 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).

results, however, is that they depend on including covariates and thus deviate from a "true" RD design.

An interesting remaining question is if increased GSE activity is not coming at the expense of FHA and subprime lenders, to whom (i.e. where in the risk distribution) are these loans going? Ongoing research using the limited borrower characteristics from the HMDA and more detailed data from the GSEs will try to answer this question, although the small effects measured here suggests that finding sharp changes in borrower risk across the cutoff may be difficult.

Also in terms of future research, one could exploit the discontinuities and temporal changes in the other GSE goals to evaluate the impact they have had and get a broader sense of the impact the GSE Act has had.

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. This limitation is countered by the credibility of the regression discontinuity design and the ability to easily 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*.[17] 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. This value appears to be considerably smaller than the subsidy the GSEs maintain, which CBO (2001) estimated at over $12 billion (in 2007 dollars) in 2000 alone.

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*, 34(3): 263-273.

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" .

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.*

[17] 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.

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*, 36(2): 207-231.

Avery, Robert B., Kenneth P. Brevoort, and Glenn B. Canner. 2007. "Opportunities and Issues in using HMDA Data" *Journal of Real Estate Research*, 29(4): 351-380.

Bhutta, Neil. 2008. "Giving Credit Where Credit is due? the Community Reinvestment Act and Mortgage Lending in Low Income Neighborhoods." .

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*, 59(3): 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*, LXI(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*, 13(2).

Imbens, Guido W. and Thomas Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice" *Journal of Econometrics,* 142(2): 615-635.

Kubrin, Charis E. and Gregory D. Squires. "The Impact of Capital on Crime: Does Access to Home Mortgage Money Reduce Crime Rates?" Paper presented at 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*, 33: 223-234.

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

Ludwig, Jens and Douglas L. Miller. 2007. "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design*" *Quarterly Journal of Economics*, 122(1): 159-208.

Manchester, Paul B. May, 2002. "Goal Performance and Characteristics of Mortgages Purchased by Fannie Mae and Freddie Mac, 1998-2000" , HF-015.

------. May, 2007. "Goal Performance and Characteristics of Mortgages Purchased by Fannie Mae and Freddie Mac, 2001-2005" , HF-017.

Munnell, Alicia H., M. B. T. Geoffrey, Lynn E. Browne, and James McEneaney. 1996. "Mortgage Lending in Boston: Interpreting HMDA Data" *American Economic Review*, 86(1): 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*, 30(2): 197-219.

Passmore, Wayne, Shane M. Sherlund, and Gillian Burgess. 2005. "The Effect of Housing Government-Sponsored Enterprises on Mortgage Rates" *Real Estate Economics*, 33(3): 427-463.

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*, 12(1): 29-59.

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

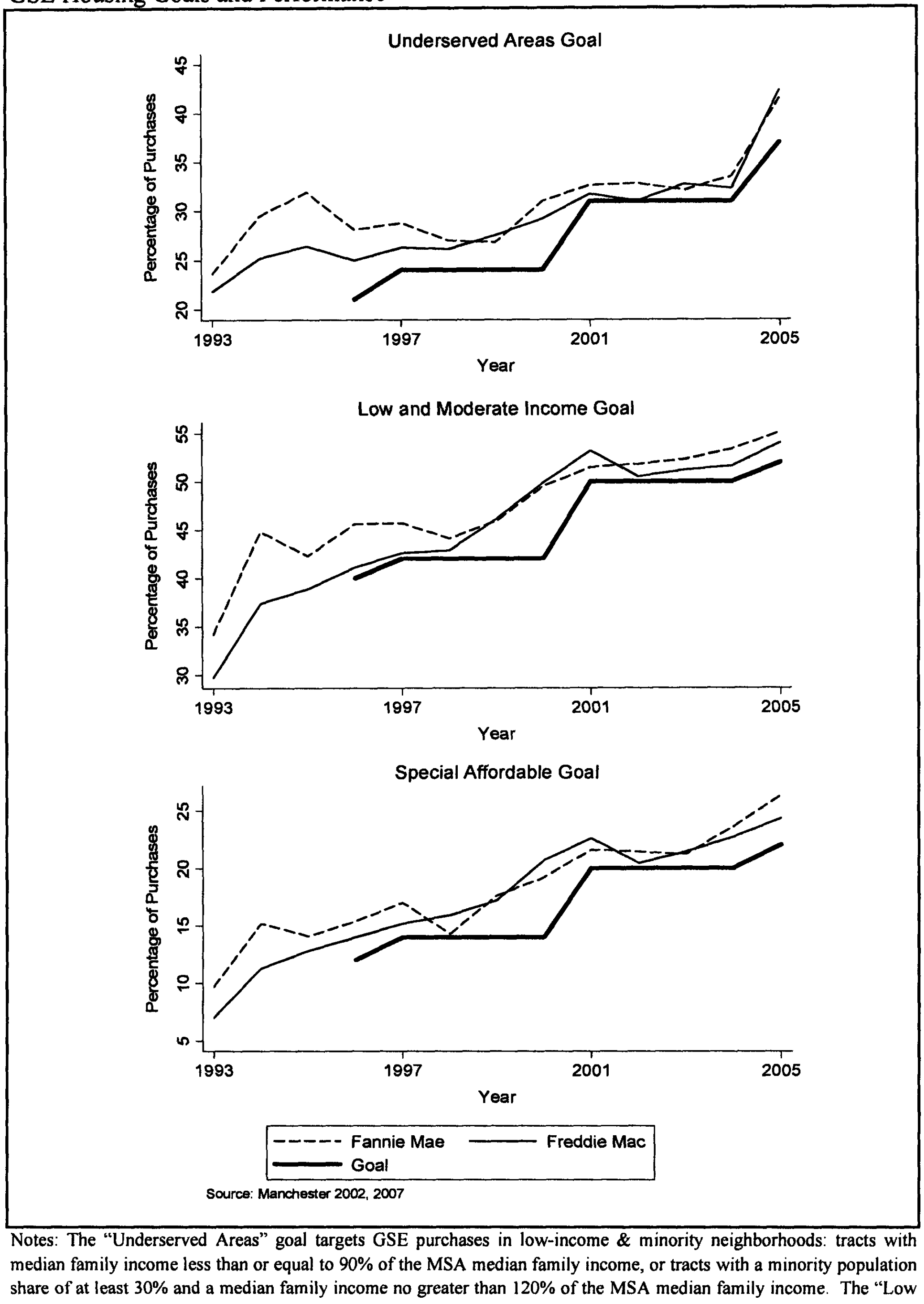
Scheessele, Randall M. 1998. "HMDA Coverage of the Mortgage Market" , HF-007.

Sherlund, Shane M. 2008. "The Jumbo-Conforming Spread: A Semiparametric Approach" (2008-01).

Straka, John W. 2000. "A Shift in the Mortgage Landscape: The 1990's Move to Automated Credit Evaluations" *Journal of Housing Research*, 11(2).

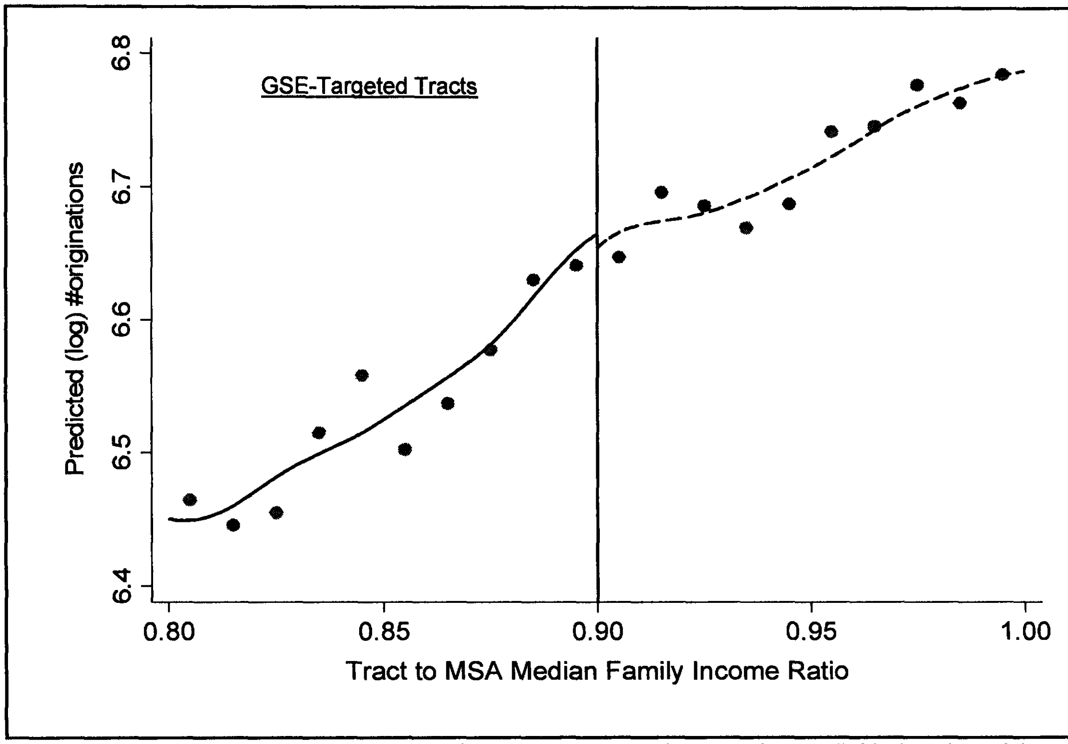
Temkin, Kenneth, Jennifer E. H. Johnson, and Diane Levy. March, 2002. "Subprime Markets, the Role of GSEs, and Risk-Based Pricing" .

Figure 1:
GSE Housing Goals and Performance



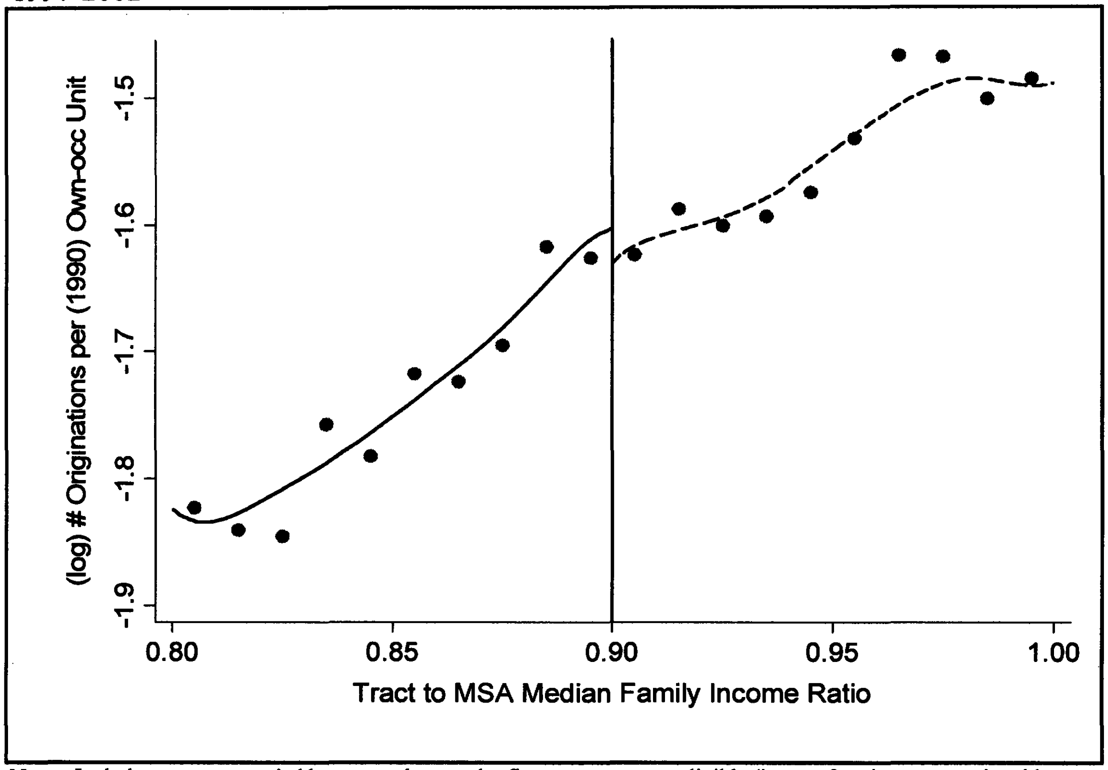
Notes: The "Underserved Areas" goal targets GSE purchases in low-income & minority neighborhoods: tracts with median family income less than or equal to 90% of the MSA median family income, or tracts with a minority population share of at least 30% and a median family income no greater than 120% of the 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



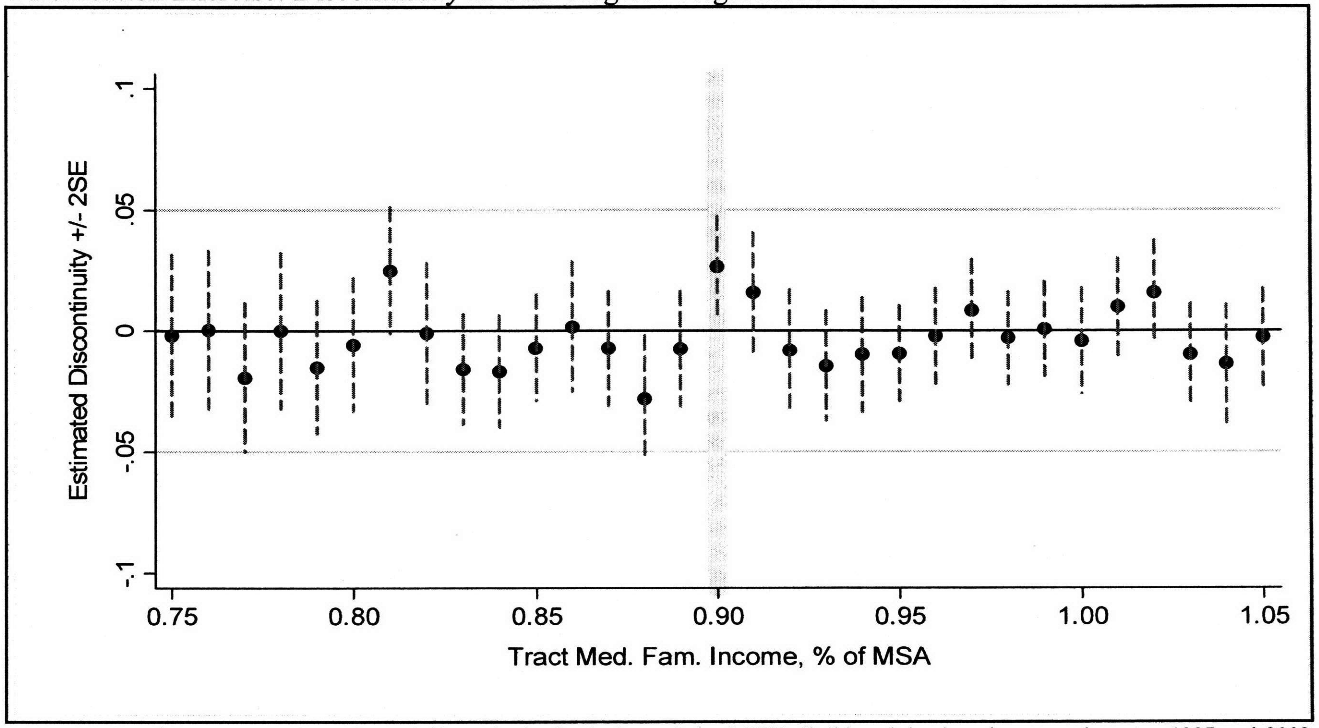
Notes: Y-axis values are predicted using coefficients on census tract characteristics (see Table 1) estimated from a tract-level regression of (log) number of refinance and home purchase originations on tract characteristics and MSA fixed-effects (estimated fixed-effects not used to generate predicted 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, 1992-2002

Variable	Description
Year	Year of mortgage application or purchase
Institution ID	10 Character Lender Identifier
Regulatory Agency ID	Code indicating OCC, Fed, FDIC, OTS, NCUA (credit unions) or HUD as supervisory agency
Loan Type	Conventional or government insured (e.g. FHA, VA)
Loan Purpose	Home purchase, refinance, home improvement or multifamily (i.e. 5+ family property)
Occupancy	Owner-occupied or investment property/second home
Loan Amount	Dollar amount of loan
Action Taken	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)	Institution can provide primary reason(s) for denial (e.g. credit history, insufficient collateral, debt load, etc)
Geography	State, county and census tract of property
Income	Gross annual family income, rounded to the nearest thousand dollar
Applicant(s) Race	Race of primary applicant; race of co-applicant if applicable. May not be provided if telephone/internet application
Applicant(s) Sex	Sex of primary applicant; sex of co-applicant if applicable. May not be provided if telephone/internet application
Purchaser	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) $85 \leq TM \leq 90$	(3) $90 < TM \leq 95$	(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 only 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		0.4019***		0.4050***		0.3918***		0.4036***
purchased by Non-GSEs, 1994-96		(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)
$\mathbf{1}[TM \leq 0.90]$	0.0211	0.0266**	0.0272**
	(0.0167)	(0.0104)	(0.0112)
TM'			
$(TM')*\mathbf{1}[TM \leq 0.90]$			
(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-Eligble 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)
$\mathbf{1}[TM \leq 0.90]$	0.0123	0.0134	0.0089
	(0.0198)	(0.0127)	(0.0123)
TM'			
$(TM')*\mathbf{1}[TM \leq 0.90]$			
(log) # GSE-Ineligible originations, 1994-96		0.5774***	0.5877***
		(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.