When Regulations Backfire: The Case of the Community Reinvestment Act

Konstantin Golyaev

University of Minnesota

September 15, 2010

Motivation

- Home mortgage lending industry had grown considerably in the mid-2000s
  - As approval rates increased, more loans went bad starting in 2007
- The Community Reinvestment Act (CRA) had been accused to add to the problem
  - CRA encourages banks to lend more in low- and moderate-income areas (lower income areas)
- Existing empirical evidence on the question is inconclusive

The Community Reinvestment Act: History

- Late 1930s: “Redlining” policy instituted by the FHA
  - Banks are strongly encouraged not to lend in certain neighborhoods
- 1950s: Supreme Court declares redlining unconstitutional
  - Banks de-facto stick to the old policies
- 1977: The Community Reinvestment Act is passed
  - Idea: lending to someone must only be determined by how likely s/he is to pay back, not by where s/he lives
  - Banks are encouraged to seek creditworthy borrowers in lower-income areas

Question

Question

Did the CRA contribute to the mortgage crisis?

Two sub-questions really:

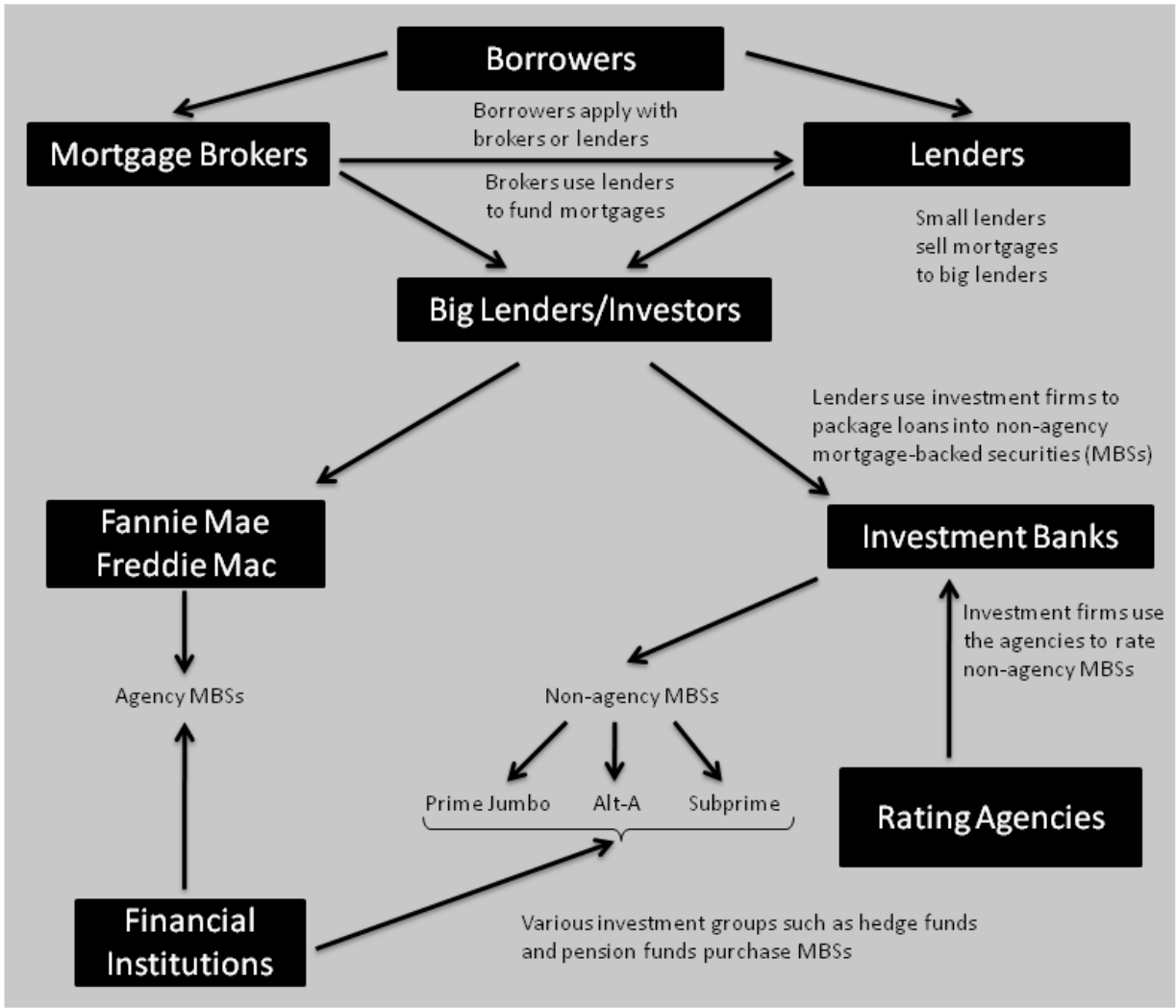
1. Does the CRA cause banks to approve more loans?
2. If yes, how did those extra loans perform?

Answer(s)

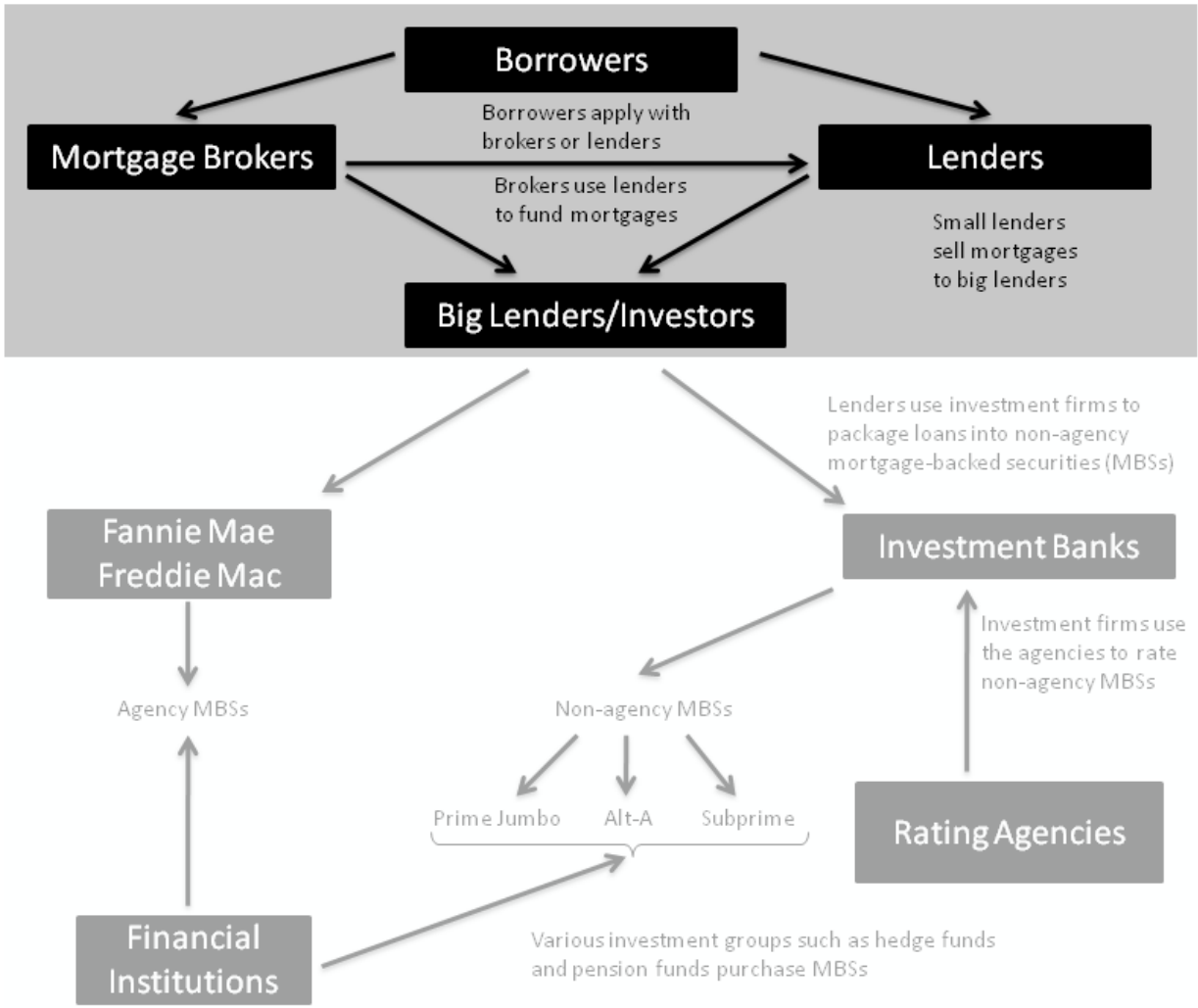
1. Yes, the CRA does have a significant effect on loan approval
   - Average marginal effect of 33% suggests almost $500,000$ extra loans approved
2. Indirect measures suggest poor peformance:
   - Foreclosure rates are 5.43 times higher in CRA-eligible areas
   - This sugggests 1 out of 6 CRA-induced loans had failed to perform
   - Other studies find similar picture, i.e. Demyanyk and van Hemert (2009), Bajari, Chu and Park (2009)

Outline

The Mortgage Origination Industry



The Mortgage Origination Industry



Data

- HMDA 2000-2005: all home mortgage loan applications ($\sim$ 50 mln. obs.)
  - Use 2005 applications for single family owner-occupied home purchase loans in California
  - Use 2000-2004 for credit scores proxies
- CRA 2005 – Census-tract-level definitions of assessment areas
- FDIC Summary of Deposits 2005 – bank branches' locations
- Census 2000 – Census-tract-level socio-economic characteristics
- Crime Rates 1999-2005 – California Attorney General's office
- 2010 Foreclosure Data – The Local Initiatives Support Corporation (LISC) and the New York Fed

Approach

- I use the discontinuities in the CRA rules to identify its causal impact
- CRA makes banks define Assessment Areas (AAs)
  - must roughly correspond to areas of their primary market activities
  - cannot cut across census tracts
  - must be a “connected” area (“holes” or “gaps” discouraged)
  - must do over 50% of their business in AAs
  - regulators look much harder at bank activities within AAs
- Regulators may forbid the bank to expand if its CRA performance is poor
- Use boundaries of assessment areas for identification

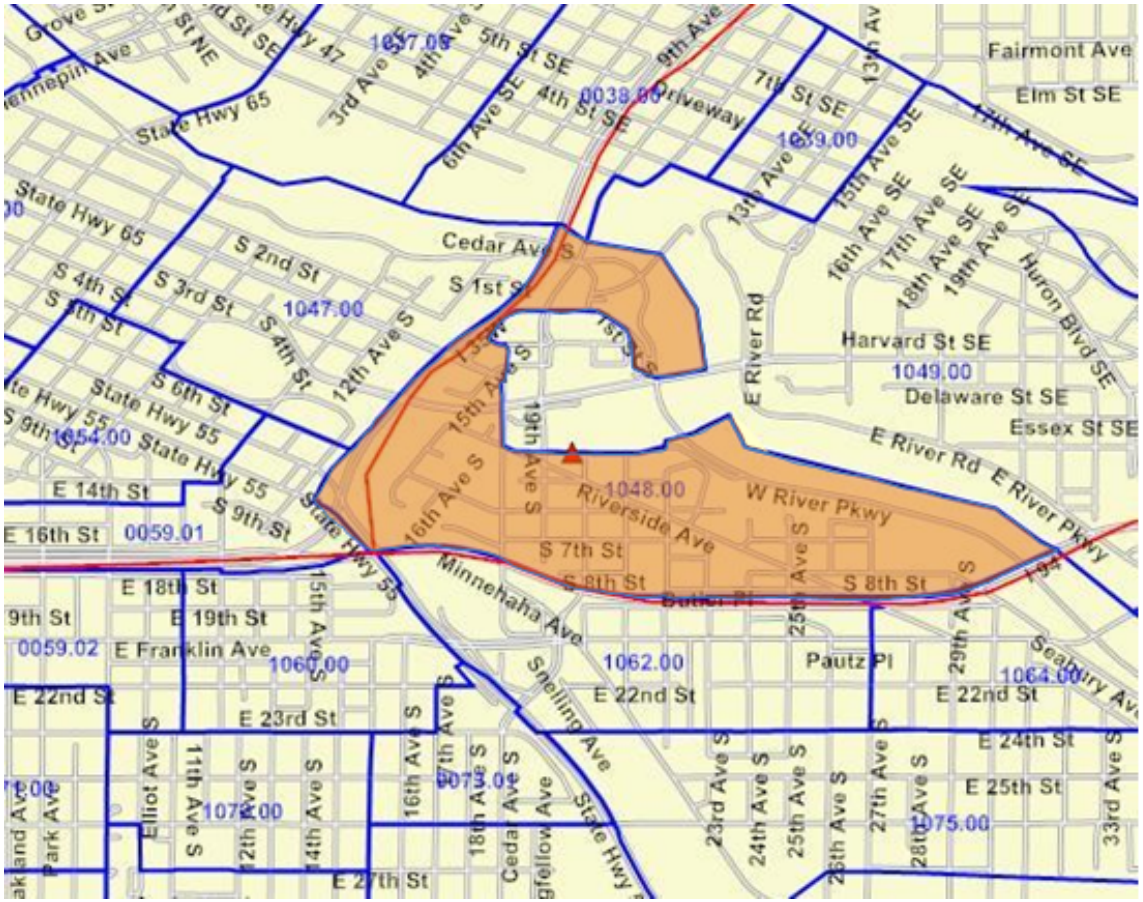
Identification

- Tract eligibility criterion:

$$\frac{\text{Tract Median Income}}{\text{MSA Median Income}} \leq 0.8$$

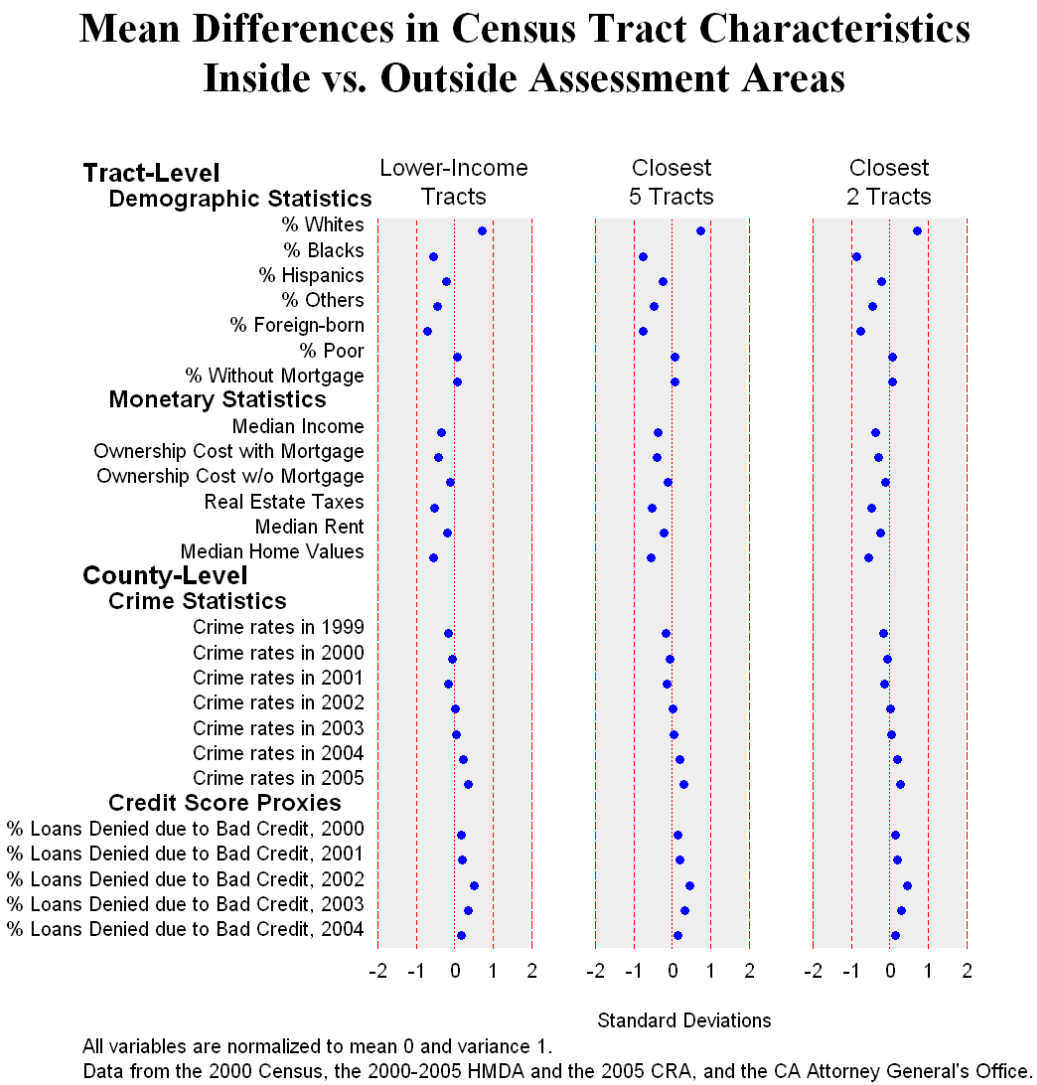
- Look at CRA-eligible census tracts along the boundaries of assessment areas
- Pick a collection of tracts that are close to each other
  - and are very similar in all observable characteristics
- Compare loan approval rates in tracts inside and outside assessment areas
- Interpret difference as the CRA causal impact

Census Tract Containing UMN Economics

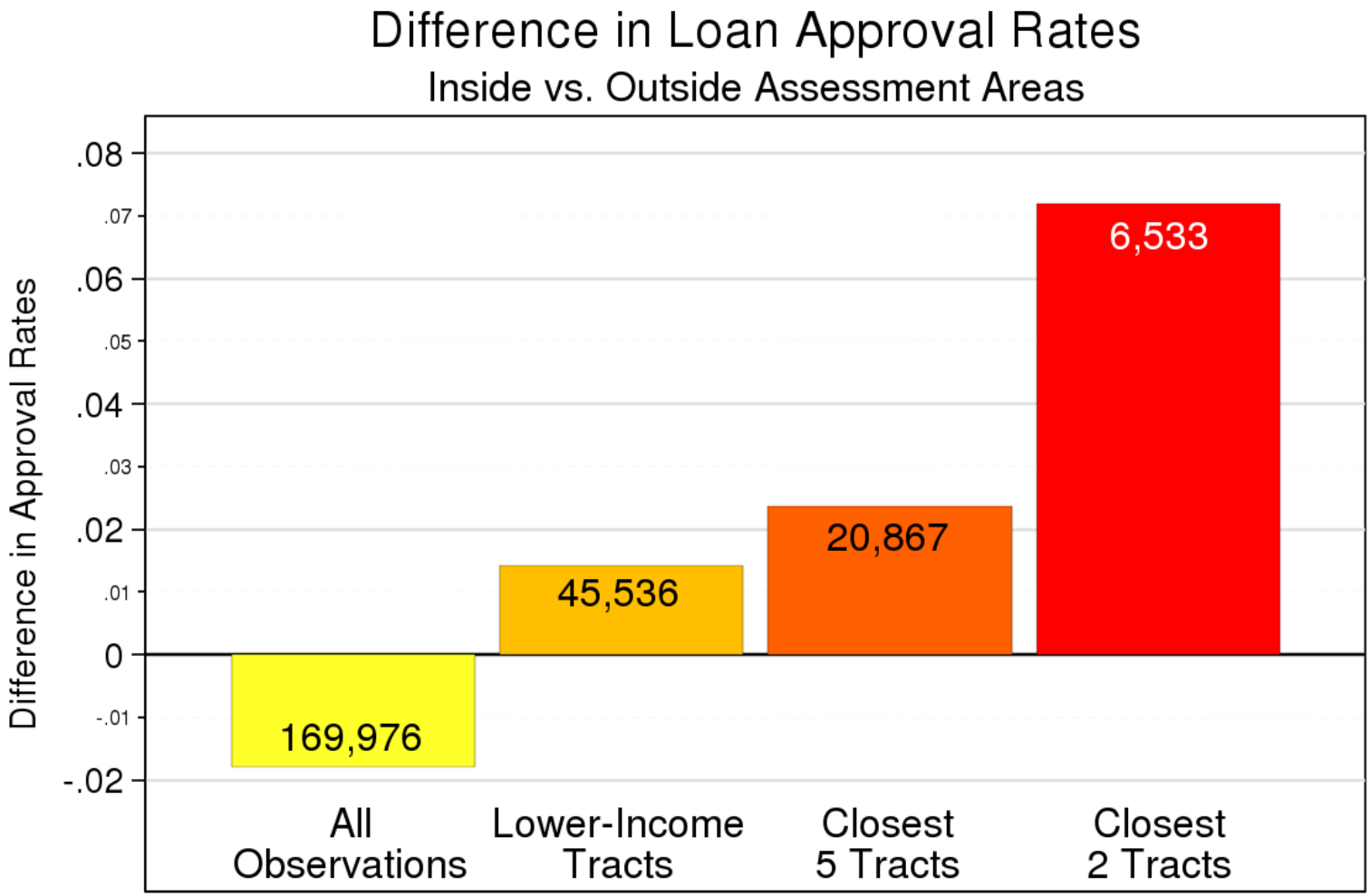


Selected Census Tract Statistics	
Population	7534
% Minority in Population	59.41%
2009 MSA Median Family Income	$83,900
2009 Tract Median Family Income	$25,254
Tract/MSA Income Ratio	0.301
CRA Eligible?	Yes

Matching Results



Preliminary Evidence



Data from the 2005 HMDA and CRA.
Numbers on the bars represent the number of loan applications used.

Regression Results

Table 1: Linear Probability Model for Loan Approval

		Lower-Income Tracts			High-Income Tracts	
	All Data	All	Closest 5	Closest 2	All	Closest 5
Loan in Assessment Area	−0.015***	0.013**	0.030***	0.068***	−0.024***	−0.005
	(0.002)	(0.004)	(0.007)	(0.013)	(0.004)	(0.005)
Loan size, 100k	−0.001*	−0.003*	−0.002	0.002	−0.003***	−0.004***
	(0.000)	(0.002)	(0.002)	(0.004)	(0.001)	(0.001)
Annual Income, 100k	0.001	0.010***	0.007	0.011**	0.002**	0.004*
	(0.001)	(0.003)	(0.004)	(0.004)	(0.001)	(0.001)
Applicant Female	−0.006*	0.004	0.004	−0.000	−0.008*	−0.009
	(0.002)	(0.004)	(0.006)	(0.011)	(0.004)	(0.006)
Applicant Not White	−0.024***	−0.023***	−0.021***	−0.017	−0.032***	−0.037***
	(0.002)	(0.004)	(0.006)	(0.011)	(0.004)	(0.005)
Applicant Hispanic	−0.088***	−0.092***	−0.088***	−0.085***	−0.054***	−0.058***
	(0.002)	(0.004)	(0.006)	(0.011)	(0.004)	(0.006)
Has a Co-Applicant	0.052***	0.048***	0.044***	0.035**	0.053***	0.051***
	(0.002)	(0.004)	(0.006)	(0.011)	(0.004)	(0.005)
Constant	0.799***	0.785***	0.769***	0.724***	0.864***	0.845***
	(0.003)	(0.006)	(0.010)	(0.017)	(0.005)	(0.008)
Number of obs.	169,964	44,546	20,867	6,533	47,039	24,393

Dependent variable: loan approval indicator. Data from the 2005 HMDA and CRA. Standard errors in parentheses.

* p< 0.05, ** p< 0.01, *** p< 0.001

Instrumental Variables

- Banks may draw assessment area boundaries strategically and nonrandomly
  - The matching procedure might fail to solve this problem completely
- CRA effect on loan approval unlikely to be constant
  - Use distance from nearest bank branch to AA boundary as instrument
- Measurement error interpretation applies here

IV Model

Main equation: linear probability model for loan approval:

$$y_i = AA_i \cdot \beta + x_i' \gamma + \varepsilon_{i,2},$$

- y_i – loan approval indicator
- i – indexes loan applications
- x_i – observable covariates
- AA_i – indicator for loan being inside the CRA assessment area

Model CRA impact via auxiliary equation:

$$AA_i = dist_i \cdot \delta_1 + x_i' \delta_2 + \varepsilon_{i,1},$$

- $dist_i$ – distance from assessment area boundary to nearest branch, bank-specific
- (β, γ, δ) – parameters for estimation

2SLS Results

Table 2: 2SLS Linear Probability Model for Loan Approval

	Lower-Income Tracts	Closest 5 Tracts	Closest 2 Tracts
Loan in Assessment Area	0.557***	0.343***	0.247***
	(0.036)	(0.031)	(0.038)
Loan size, 100k	0.031***	0.011***	0.003
	(0.003)	(0.003)	(0.004)
Annual Income, 100k	−0.002	0.007	0.011**
	(0.003)	(0.004)	(0.004)
Applicant Not White	0.028***	−0.010	−0.009
	(0.006)	(0.007)	(0.012)
Applicant Female	0.003	0.004	−0.003
	(0.005)	(0.007)	(0.012)
Has a Co-Applicant	0.010	0.030***	0.024*
	(0.005)	(0.006)	(0.011)
Applicant Hispanic	−0.090***	−0.100***	−0.082***
	(0.005)	(0.006)	(0.011)
Constant	0.377***	0.505***	0.582***
	(0.028)	(0.027)	(0.034)
Number of obs.	44, 546	20, 867	6, 533

Dependent variable: loan approval indicator. Data from the 2005 HMDA and CRA. Distance to nearest bank branch used as an instrument for loan being in assessment area. Standard errors in parentheses. * $p< 0.05$, ** $p< 0.01$, *** $p< 0.001$

Nonlinear Bayesian IV Model

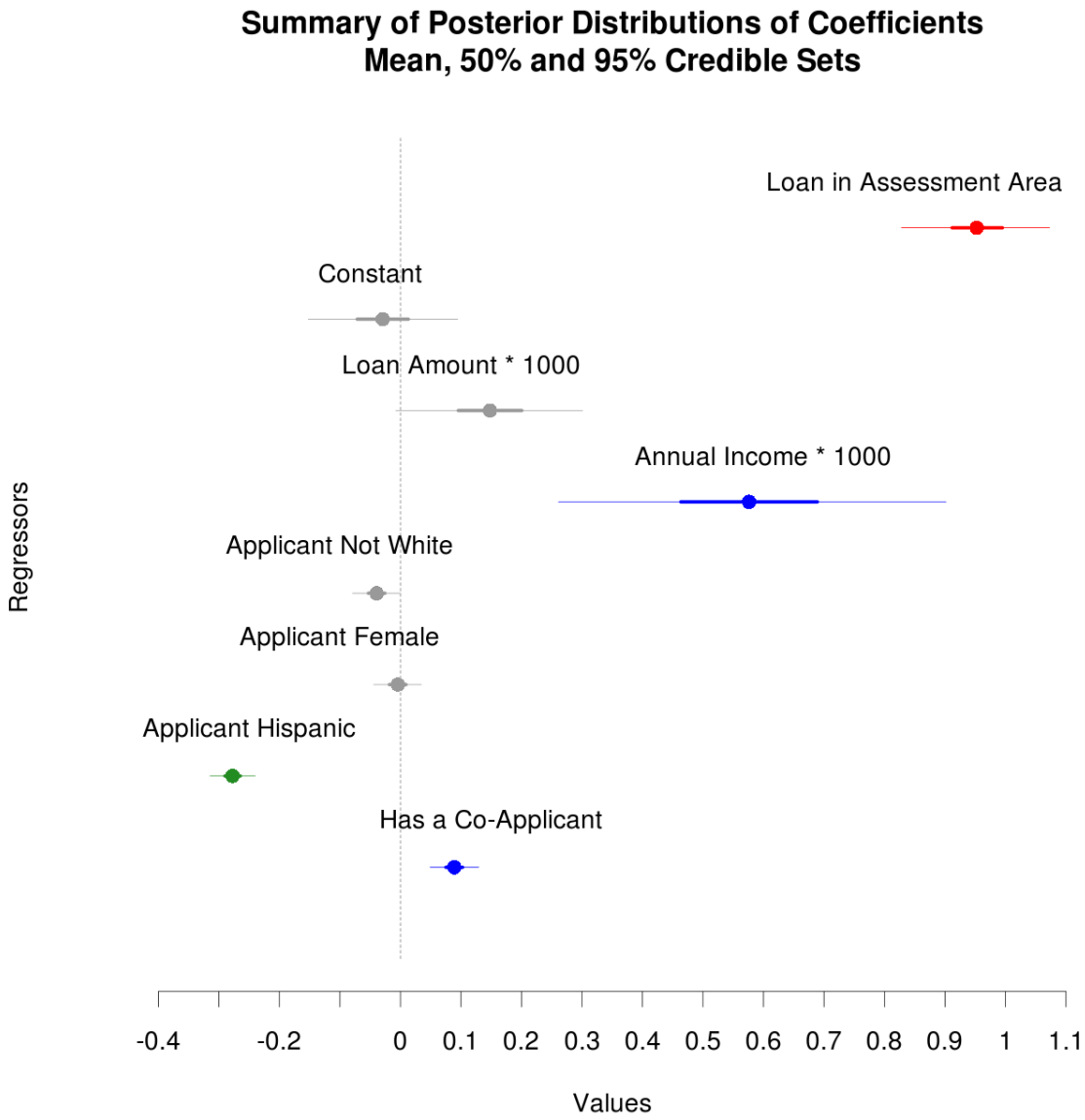
- Proper model of a binary outcome involves nonlinearities (probit)
  - Linear probability model is an approximation
  - Blundell and Powell (2004) show it can be really poor
- Want to allow for unobserved heterogeneity via random coefficients
- Rewrite main equation as loan approval probit:

$$y_i^* = AA_i \cdot \beta + x_i'\gamma + \varepsilon_{i,2}, \quad y_i = I\{y_i^* \geq 0\},$$

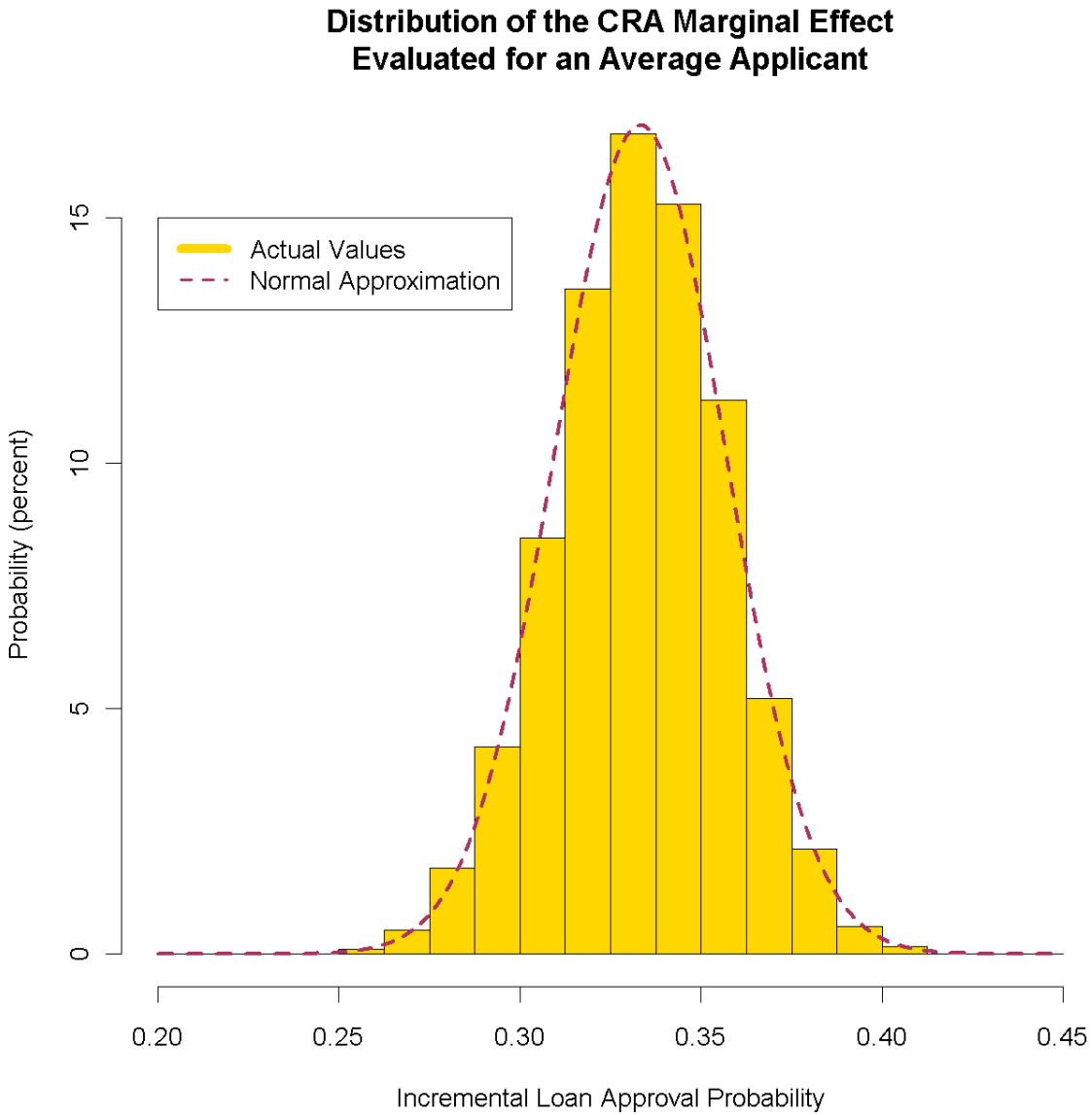
  - y_i^* – latent loan application "score"
- The CRA auxiliary equation is unchanged.

⏩ Bayesian IV detailed

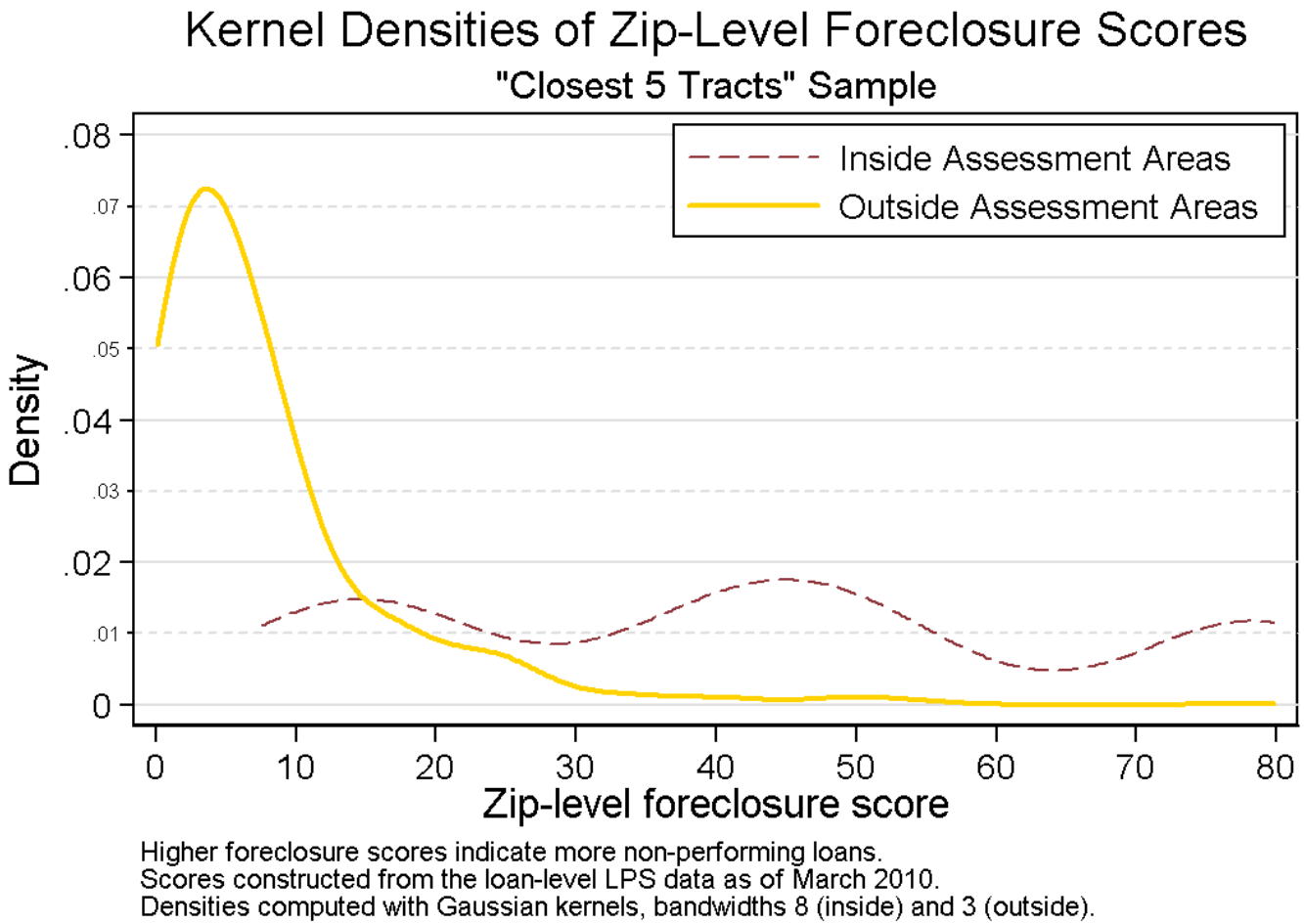
MCMC Results: Main Equation Posteriors



MCMC Results: CRA Marginal Effect



How Did The Extra Loans Perform?



Higher foreclosure scores indicate more non-performing loans.
Scores constructed from the loan-level LPS data as of March 2010.
Densities computed with Gaussian kernels, bandwidths 8 (inside) and 3 (outside).

Mean score outside AA: 7.93; inside AA: 42.52 (5.36 times larger).

Conclusion

- CRA does induce banks to approve more mortgage loans
  - About $500,000$ extra loans had been approved in CA in 2005
- This likely to have exacerbated problems with mortgage defaults
  - By 2010, 1 out of 6 CRA-induced loans had failed to perform

Nonlinear Bayesian IV Model

- Set up estimation as Bayesian IV with Data Augmentation

$$\begin{cases} AA_i = dist_i \cdot \delta_1 + x_i'\delta_2 + \varepsilon_{i,1} \\ y_i^* = AA_i \cdot \beta + x_i'\gamma + \varepsilon_{i,2} \end{cases},$$

$$\begin{pmatrix} \varepsilon_{i,1} \\ \varepsilon_{i,2} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} \right)$$

- Priors:

$$\begin{aligned} \delta &\sim N\left(\mu_\delta, A_\delta^{-1}\right) \\ (\beta, \gamma) &\sim N\left(\mu_{\beta\gamma}, A_{\beta\gamma}^{-1}\right) \\ \Sigma &\sim IW\left(\upsilon_0, V_0\right) \end{aligned}$$

⏩ Back to the Presentation

Nonlinear Bayesian IV Model

- Data augmentation step:

$$\varepsilon_2 \mid \varepsilon_1 = \bar{\varepsilon}_1 \sim N\left(\frac{\sigma_{12}}{\sigma_2^2}\bar{\varepsilon}_1, \sigma_2^2 - \frac{\sigma_{12}^2}{\sigma_1^2}\right),$$

- Treat y_i^* as extra set of parameters, draw them from truncated normal
- Caveats:
  - Model not identified: cannot recover σ_2^2.
  - So do MCMC in non-identified space, then "margin out" the identified parameters
  - Model takes many iterations to converge $(100,000)$

Back to the Presentation