

# Performance Information and Personnel Decisions in the Public Sector: The Case of School Principals

Julie Cullen, Eric Hanushek, Gregory Phelan and Steven  
Rivkin

October 2018

# Accountability and principal labor market outcomes

- The passage of No Child Left Behind (NCLB) followed years of state experimentation with test-score accountability
- This emerged from a concern that weak incentives adversely affected school quality
  - There appeared to be little relationship between job performance and labor-market outcomes for many educators including school principals

# Potential sources of a weak association between educator effectiveness and labor-market outcomes

- Little information on student performance including test scores
- There were subjective evaluations of teachers, but very few received unsatisfactory ratings
- Rigid salary schedules and strong job protections

- Weak incentives for administrators to take action in the face of poor performance or even to differentiate among teachers or principals on the basis of effectiveness
- Uninformed school boards could contribute to the problem by lessening the information content of superintendent evaluations
  - Boards of trustees require meaningful information from sources other than the administration

Texas accountability and NCLB require districts to generate and provide performance information

- Test score results by subject, grade and demographic group
- Cruder ratings categories based on test-scores and other factors including attendance
- Panel micro-data on students and schools that can be used to generate achievement growth or value-added measures

We study the effects of performance information on the probability of principal job retention and compensation growth

- Examine the pattern of Regression Discontinuity Design (RDD) estimates of ratings effects to learn more about the use of various types of information
- RDD estimates essentially compare labor-market outcomes for two principals who lead schools with almost identical test-score pass rates except that they fall on opposite sides of the cutoff
  - E.G. 79% versus 80% where the cutoff to receive an acceptable rating is 80%

Why should the principal with a 79 % pass rate have a lower probability of job retention or receive a lower raise?

- The superintendent uses only the rating and does not have knowledge of the underlying pass rate information that determines the rating
- Principals may resign rather than remain due to additional burdens or stigma of a low rating
- Other stakeholders including parents and school-board members only have information on the ratings and pressure the district to take action
  - Media may focus on ratings
  - Superintendents may be reluctant to take job actions or use better performance information

# The use of NCLB and Texas accountability in the evaluation of principals

- Ratings are based largely on the pass rate rather than achievement growth
  - Schools can meet required improvement thresholds in the lower ratings categories
- Pass rates and achievement levels reflect family in addition to school influences
  - Performance ratings based on pass rates unfairly punish principals who work in disadvantaged schools
- It is difficult to infer the performance of the principal from measures of school outcomes



We also study the associations among arguably better measures of principal effectiveness, ratings and labor-market outcomes

- Describe differences in pass rates and a more compelling value added (VA) measure of principal effectiveness by ratings
- Estimate the association between labor-market outcomes and pass rate, VA and rating
  - Examine differences in the association between labor-market success and these variables in the same and in different districts

# UTD Texas Schools Project data

- Stacked panels of students and staff
- Annual testing
- Student demographic characteristics
- Can follow principals and students who switch schools and roles within Texas public schools
- We exclude principals with 25 or more years of experience to limit influence of retirement
- Exclude first year in a school (little impact?)
- 11,428 total principal/year observations

# Texas Accountability System

- Schools are rated unacceptable, acceptable, recognized or exemplary based upon pass rates on standardized tests for all students combined and specific demographic groups with minimum numbers of test-takers
  - Diverse schools with more groups are less likely to realize a higher rating
    - Some exemptions to mitigate this effect
- Schools may meet required improvement thresholds

# Distribution of Campus Accountability Ratings

Unacceptable	0.012
Acceptable	0.381
Recognized	0.438
Exemplary	0.169

# Information Dissemination

- Pass rates for all included tests and demographic groups are reported on line for all schools
- The ratings are widely disseminated and known
  - Extensively discussed in the press and in school-board meetings
- The accountability system does not report information on value added

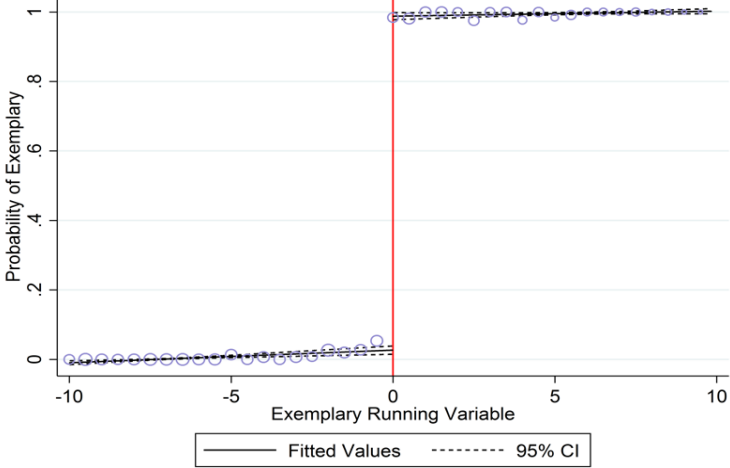
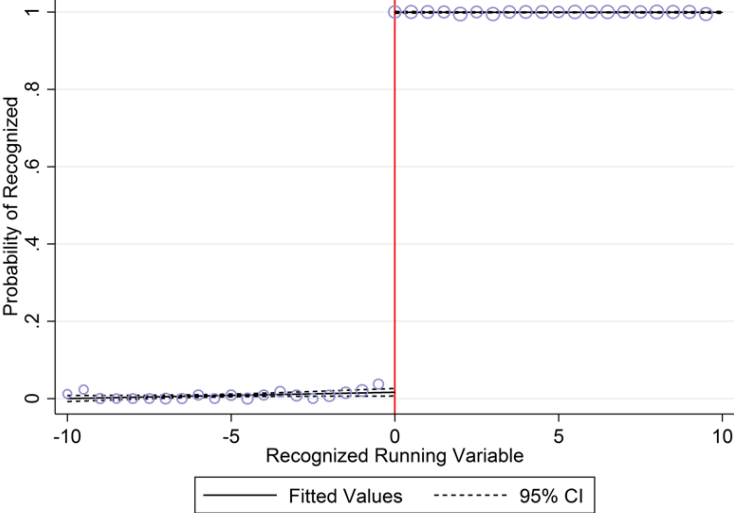
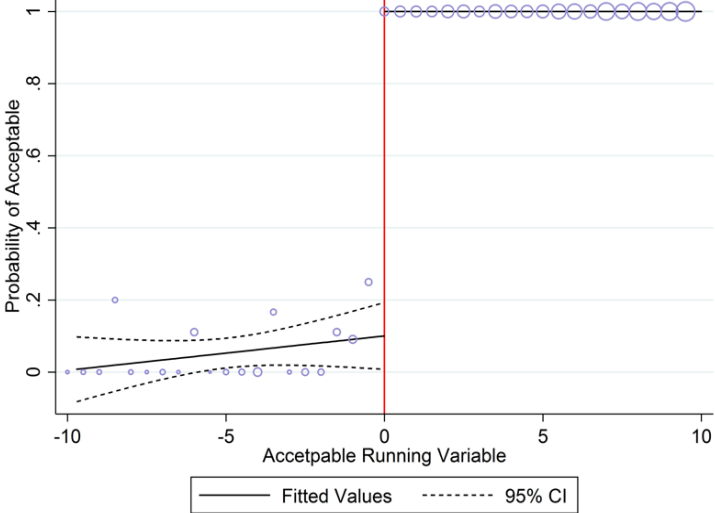
# Accountability system structure enables use of RDD methods

- The running variable is the pass-rate of the lowest-scoring demographic group relative to its cutoff for a given rating that counts for accountability purposes
  - Must be at least 30 and possibly up to 50 students to count
  - System allows groups to meet a standard even if they do not meet the minimum specified pass rate through exceptions
  - Running variable is based on marginal group and measure

# Additional Issues

- Absence of information on salary and job offers limits inferences that can be drawn about personnel practices
  - A principal may respond to a rating by voluntarily leaving a position
- Ratings are released in August, so we focus on labor market outcomes in the second year after the release

# Figure 2: Regression Discontinuity First Stages by Rating





# RDD specification tests

- Little evidence of manipulation or bunching that would invalidate the RDD design
  - Demographic characteristics move smoothly through discontinuity
  - Densities show little or no evidence of bunching
- This is not a surprise given complexity of accountability system and uncertainty of rating at time of testing

# Appendix Figure A2. Running Variable Densities Around Accountability Rating Cutoffs

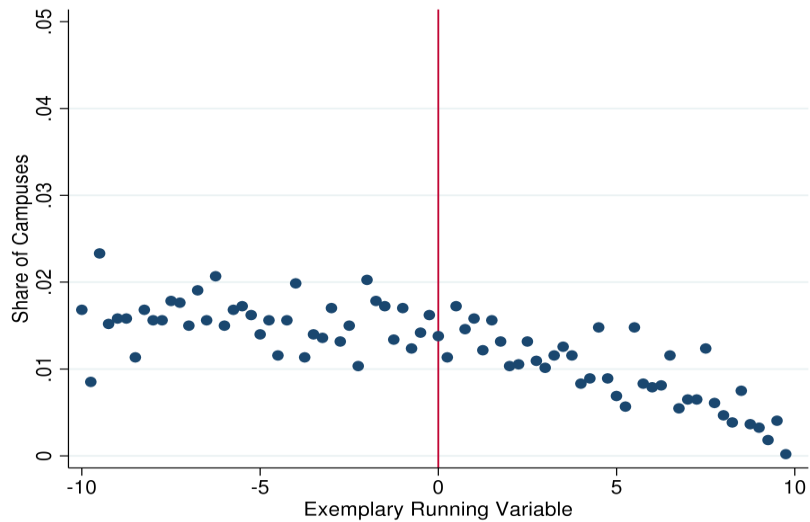
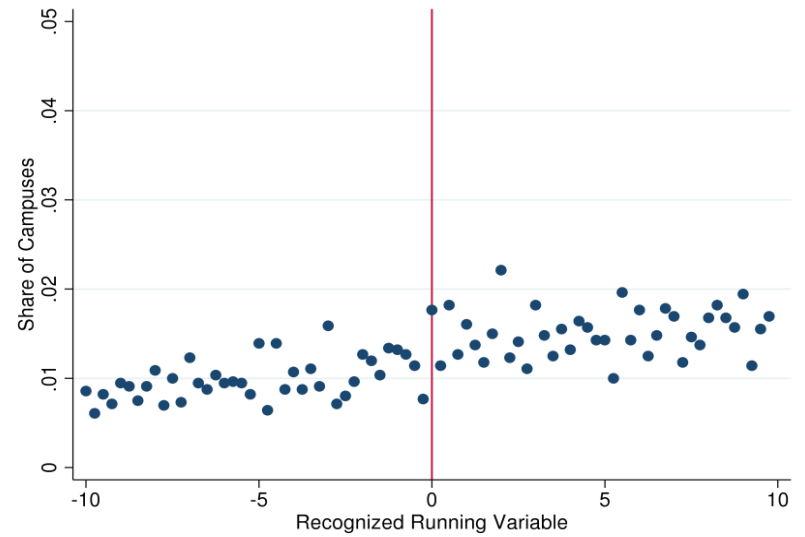
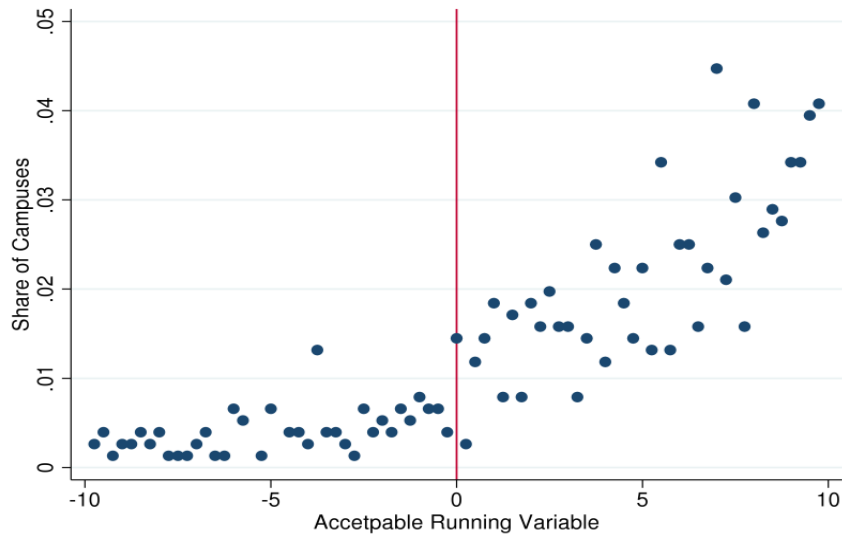


Figure 3: Probability of Retention Around the Accountability Cutoffs

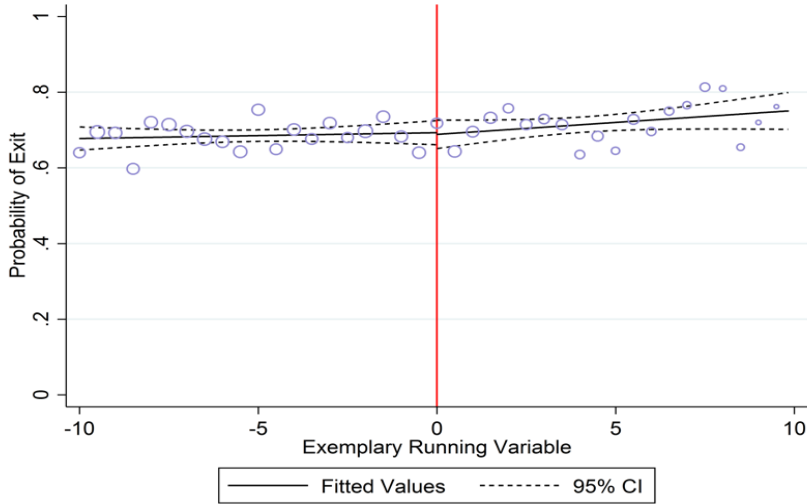
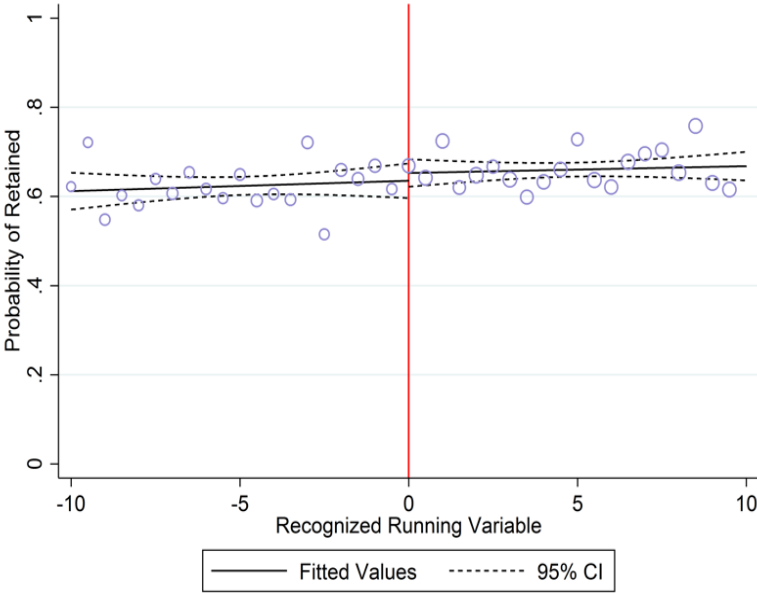
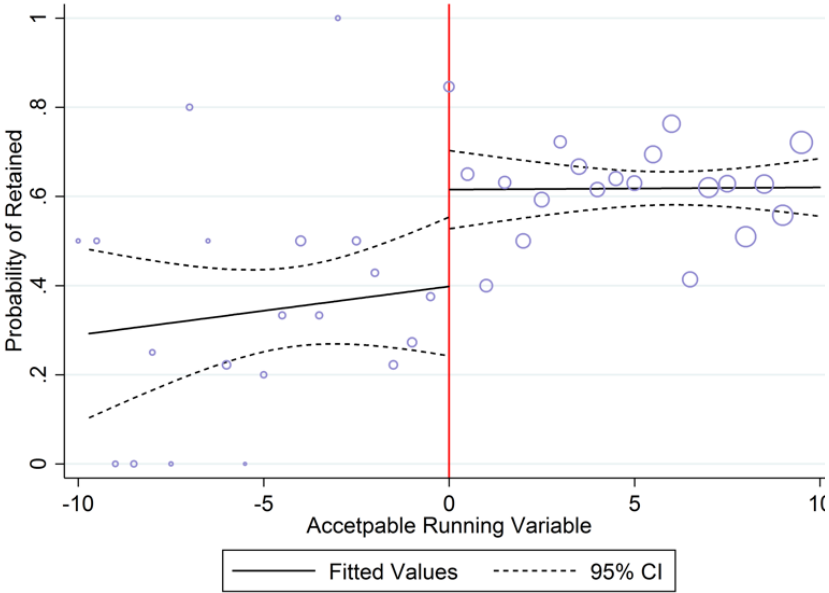


Table 6. Regression discontinuity estimates of the impact of attaining the higher rating on the probability of job retention

bandwidth	7.5	5	2.5	optimal
Acceptable	0.270**	0.365***	0.467***	0.425***
n	497	299	140	222
Recognized	0.019	0.024	0.021	0.021
n	4,252	2,879	1,458	1,457
Exemplary	0.014	0.018	0.038	0.024
n	3,925	2,690	1,419	1,767

# Figure 4: Change in Log Salary Around the Accountability Cutoffs

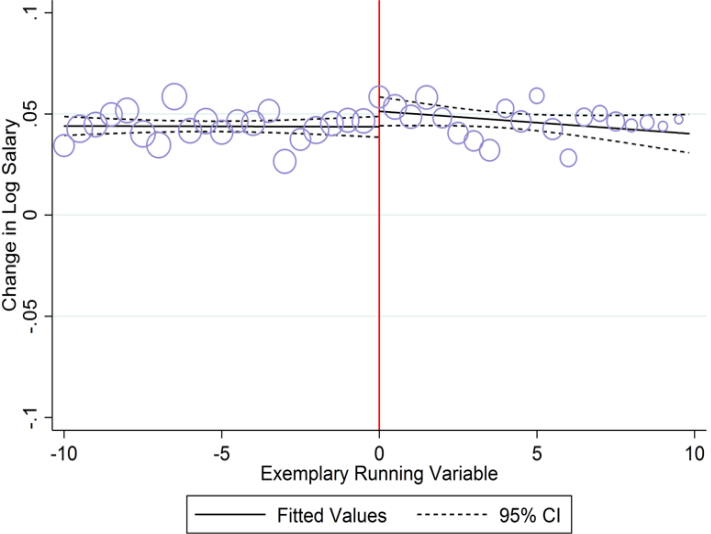
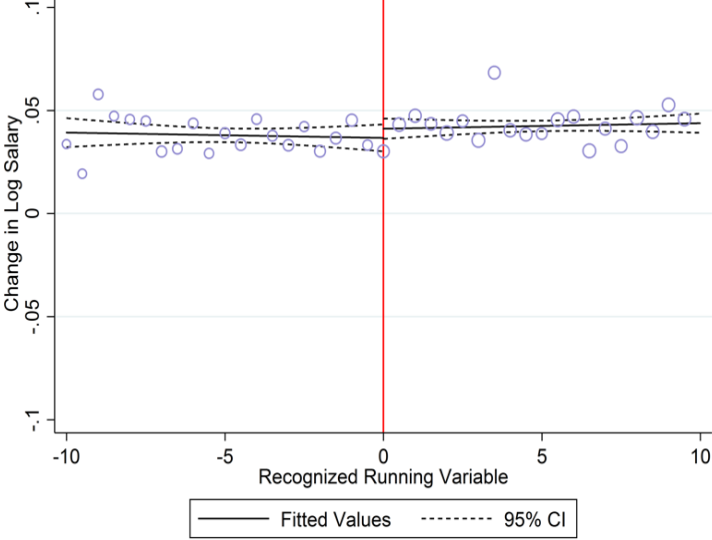
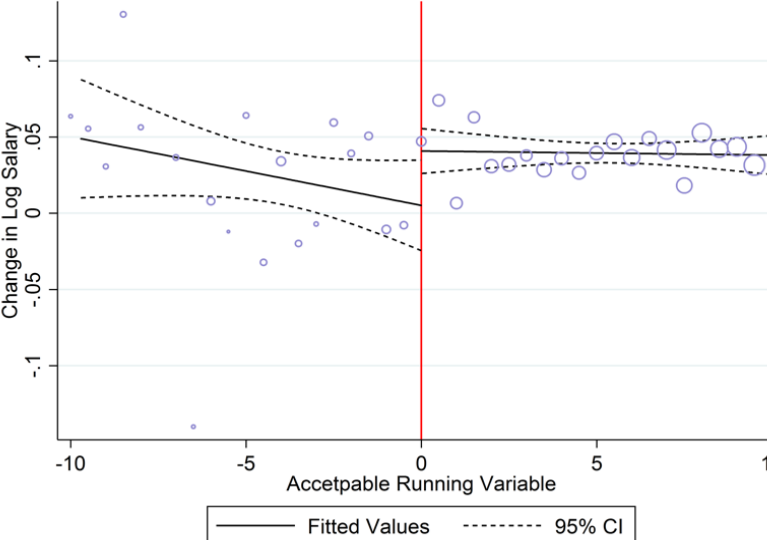


Table 8. Regression discontinuity estimates of the impact of attaining the higher rating on salary growth (change in log salary)

bandwidth	7.5	5	2.5	optimal
Acceptable	0.035*	0.052**	0.070**	0.060**
n	413	238	112	187
Recognized	0.002	0.001	-0.002	-0.002
n	3,762	2,546	1,285	1,284
Exemplary	0.011*	0.010	0.008	0.006
n	3,573	2,443	1,290	1,604

# Potential channels that contribute to the effect of an unacceptable rating

- Lack of information on underlying pass rates?
  - Absence of discontinuities at recognized and exemplary thresholds suggests that districts have information on underlying pass rates
- Voluntary principal departures to avoid oversight and restrictions that follow multiple years of unacceptable rating
  - Following table reveals little or no evidence of a stronger effect in schools that were classified as unacceptable in the previous year

Table 7. Regression discontinuity estimates of the impact of attaining an acceptable rating on the probability of job retention, by prior year rating

bandwidth	7.5	5	2.5	optimal
Previously rated				
unacceptable	-0.478	-0.247	-0.799	-0.305
n	35	24	12	19
Not previously rated				
unacceptable	0.340***	0.424***	0.559***	0.488***
n	462	275	128	203



Other potential channels that contribute to the effect of an unacceptable rating

- School board members or parents demand action in response to an unacceptable rating
  - Consistent with anecdotal evidence including newspaper articles
- An unacceptable rating raises the probability of voluntary quits

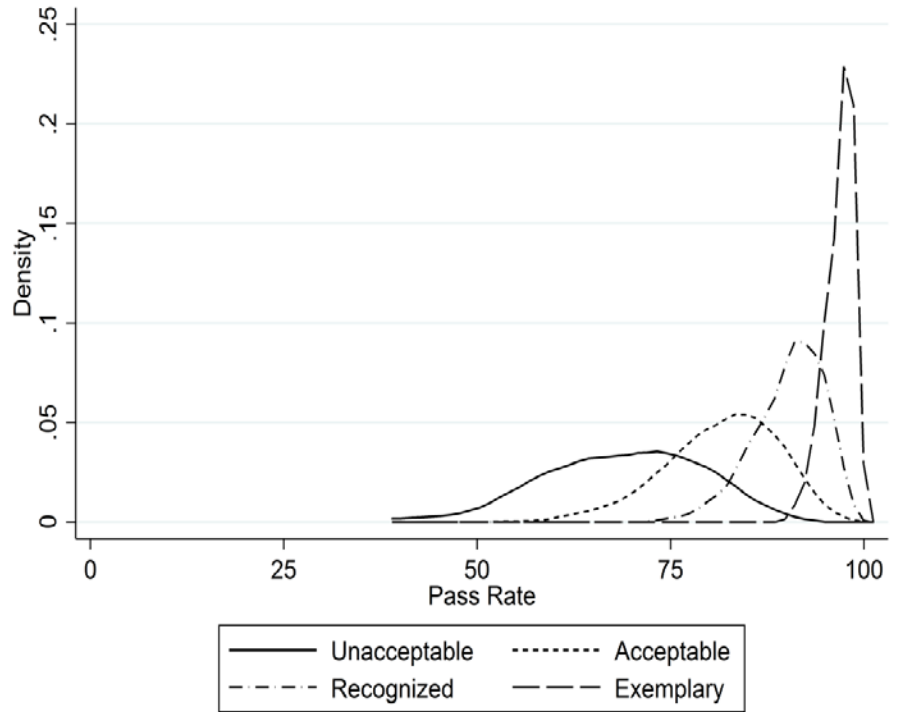
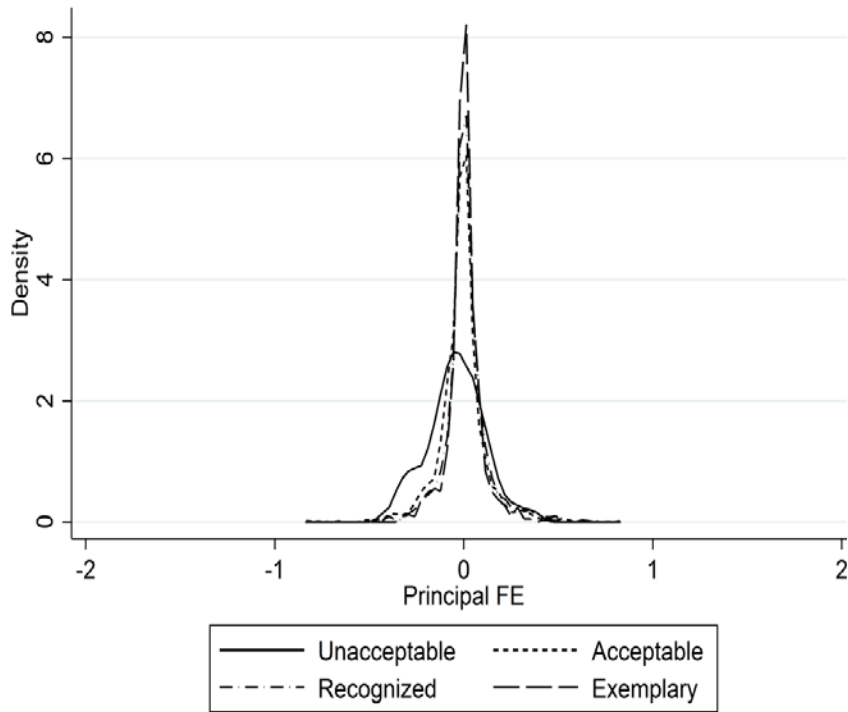
Does a higher probability of departures in schools rated unacceptable lead to disproportionate exits of less effective principals?

- The higher probability of principal turnover below the acceptable threshold is evidence of an information failure
- However, it could nevertheless raise the quality of leadership and instruction given the information deficiencies if principals rated unacceptable are significantly less effective than those rated acceptable

# Empirical specification used to estimate principal effectiveness

- Student-level value added model controlling for cubic in prior achievement
  - Demographic and program characteristics as controls
  - campus fixed effects
  - Principal fixed effects
- Sample excludes the first year as principal in a school and schools with only a single principal
- Demean estimates by connected network fixed effects
  - Quality is relative to other principals in the network
  - Campus is the network in many cases

# Figure 7. Principal Fixed Effect and School Pass Rate Densities, by Accountability Rating



Describe the association between labor market outcomes and multiple measures of student outcomes

- OLS regression with district by year fixed effects, cubics in prior achievement and demographic and program characteristics as controls
- Measures of student outcomes
  - Rating indicators
  - Pass rate
  - Principal fixed effect on achievement

# OLS estimates of relationships between outcomes and ratings, pass rate, and principal fixed effect

	Probability of retention	Change in log(salary)
Principal fixed effect	0.003	0.02
Pass rate	0.001	0.0017***
unacceptable	-0.17**	-0.02
recognized	0.03	0.00
exemplary	0.02	0.012**

# Investigate differences in information use between current and outside districts

- Current district has additional information on job performance (e.g. evaluations) which is related to productivity and effectiveness
- Given very small number of out of district moves we create a single measure of labor market success
- Indicator of success equal to 1 if
  - Retained in job; or
  - Salary growth exceeds median salary growth
  - Change in student demographic exceeds the median
    - Equal to 0 for principals who continue in current position

# Estimation of differences by destination

- RDD analysis run separately by destination
  - Note that other destination is included with non-successes
- Multinomial logit estimates
  - Not success is omitted category
  - Those who exit the Texas public schools are classified as not success
  - No evidence based on RDD that a higher rating raises the probability of exit from Texas public schools



Table 7. Regression discontinuity estimates of the impact of attaining an acceptable rating on the probability labor-market success, by job location

bandwidth	7.5	5	2.5	optimal
Same district	0.222*	0.289**	0.224	0.306*
n	497	299	140	222
Different district	-0.120	-0.125	-0.055	-0.111
n	497	299	140	222

# Multinomial logit estimates of relationships between school performance metrics and composite labor market success within district and out of district

	Success in same district (80.2%)	Success in other district (4.8%)
Principal fixed effect	0.600*	-0.095
Average pass rate	0.035***	0.039***
unacceptable	-1.011***	0.261
recognized	0.267***	-0.068
exemplary	0.335**	0.160

# Summary

- Little or no discontinuity in labor market outcomes at the recognized and exemplary boundaries
- Crossing the acceptable boundary raises retention and salary growth
- Pattern of estimates suggests that pressure for the district to respond to an unacceptable rating or the principal decision to voluntarily quit account for the discontinuity

# Summary

- The ratings based on pass rates fail to sort principals by effectiveness as measured by VA
  - However, many effective principals with high VA receive an unacceptable rating
- Current district appears to make use of information related to VA; destination districts do not appear to use it

# Policy implications

- An unacceptable rating induces a response and raises the probability of job separations
- Labor-market outcomes are significantly related to ratings, pass rate, and VA in the case of the same district
- Basing accountability system on metrics more closely aligned with learning could
  - elevate the quality of instruction
  - treat schools serving disadvantaged students more favorably