Essays on Regime Change and Education Policy Reform

By

Michael A Roberts

Submitted to the graduate degree program in the Department of Economics and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Chair: David Slusky

Donna Ginther

Dietrich Earnhardt

Kevin Pisciotta

Argun Saatcioglu

Date Defended: 28 August, 2018

The Dissertation Committee for Michael A Roberts certifies that this is the approved version of the following dissertation:

Essays on Regime Change and Education Policy Reform

Chair: David Slusky

Date Approved: 28 August, 2018

Abstract

This dissertation consists of three chapters, each representing a self-contained research paper in health economics.

The first chapter formalizes a model which generalizes several political models of collective action and regime change. It considers the impact that an extremist party can have on the choices made by the population in choosing whether or not to take part in a revolt. This third party is usually a personal interest group that benefits from a revolution, or in some cases benefits from the current regime in power, and thus will try to persuade the general population into pursuing an action that is in the extremist's best interest. The paper presents several applications of the model with political and economic roots. These models add insight to revolutions in the present day as well as throughout history, particularly those aided by outside benefactors.(JEL codes: D5,D72,D91)

In the second chapter, I study the effect that NCLB had on teacher turnover and compare it to the impact from state accountability systems that existed prior to the passage of NCLB. I find that, while state accountability systems have no significant effect on teacher turnover, teachers are 5 percentage points more likely to remain in the field following NCLB. The driving force behind this result is the year that a teacher earned his or her bachelor's degree. Receiving a bachelor's degree after 2002 makes an individual 27 percentage points more likely to stay in the field after NCLB was passed. I believe there is a self-selection process to explain this result, as after NCLB became law only those individuals who felt comfortable teaching under an accountability system earned their education degree and became a teacher. I find further evidence of this result by considering where the teachers earned their degrees. The state where a teacher earned their degree does not have a large effect on the likelihood of continuing to teach, regardless of whether or not that state had some kind of prior accountability system. This suggests that the self-selection process occurs before the individuals earn their degree, and not in response to a change in curriculum from the college or university at which they earned their degree. (JEL codes: I28, J08, J48)

In the third chapter, I consider a teacher's response to earning tenure as it relates to classroom performance. For elementary and secondary school teachers, earning tenure makes it very expensive and time consuming for a school to terminate them. Critics of tenure argue that this creates an incentive for teachers to expend less effort and energy into their teaching as they are able to avoid the penalty of job loss. There is a long research history of teacher characteristics and the effects they have on student achievement. Surprisingly, the impact of earning tenure remains unknown. This chapter uses student level data and pairs it with teacher and school characteristics in order to find the effect earning tenure has on test scores. By taking advantage of the staggered issuance of tenure, we can isolate the impact of earning tenure. I find that immediately after being awarded tenure, student test scores drop. However, they increase and overtake previous scores the following year. This suggests teachers earn extra benefit when their students are successful, and thereby, resist the incentive to provide lower quality teaching. (JEL codes: I20, I28, J08)

Acknowledgements

I would like to thank my advisor David Slusky as well as the many individuals who have offered me comments and advice over the years, including Dietrich Earnhardt, Donna Ginther, Argun Saatcioglu, Tarun Sabarwal, Tsvetan Tsvetanov, and Jianbo Zhang.

In addition, I would like thank seminar participants at the University of Kansas and Kansas State University, as well as conference participants at Southern Economic Conference, Missouri Valley Economic Conference, and Graduate Student Economic Conference.

I am immensely grateful to the staff at the National Center for Education Statistics at the Department of Education for their help in acquiring the data for my research and to Julie Fugett, Carl Taylor, and Rick Evanhoe for providing and facilitating a secure space to work with this data.

This paper explores behavior changes in response to No Child Left Behind and tenure using data obtained from the National Center for Education Statistics. The findings and conclusions in this paper are those of the author and do not necessarily represent the views of NCES, or the U.S. Department of Education.

Finally, thank you to my friends and family for all your support, especially to my wife Becca, my children, Alexander and Savannah, my siblings, Katie, Andrea and Jonathan, and my parents, Paul and Beth.

All remaining errors are my own.

Contents

Acceptance……………………………………………………...……………………………ii

Abstract……………………………………………………………………………………...iii

Acknowledgment……………………………………………………………………………v

1 A Generalization of Models of Collective Action and Regime Change………………………...1

1.1 Introduction…………………………………………………………………………1

1.2 The Model…………………………………………………………………………..2

1.3 Comparative Statistics……………………………………………………………13

1.4 Application: Common Knowledge Cost………………………………………..15

1.5 Application: Uncertain Payoffs…………………………………………………18

1.6 Application: Pro-Government Extremist……………………………………….19

1.7 Conclusion………………………………………………………………………...20

2 The Impact of No Child Left Behind on Teacher Turnover…………………………………21

2.1 Introduction………………………………………………………………………..21

2.2 Theoretical Model………………………………………………………………24

2.3 Data and Methods………………………………………………………………28

2.4 Results……………………………………………………………………………..32

2.5 Discussion and Robustness Checks…………………………………………….45

2.6 Conclusion………………………………………………………………………...52

3 Teacher Reaction to Tenure…………………………………………………………….54

3.1 Introduction………………………………………………………………………..54

3.2 Background on Teacher Tenure………………………………………………...56

3.3 Theoretical Model………………………………………………………………58

3.4 Data and Methods………………………………………………………………60

3.5 Results……………………………………………………………………………..63

3.6 Discussion…………………………………………………………………………71

3.7 Conclusion……………………………………………………………………...72

Appendices………………………………………………………………………………….74

References……………………………………………………………………………..91

1 A Generalization of Models of Collective Action and Regime Change

.Introduction

The purpose of this paper is to formalize a model which generalizes several political models. We will mainly examine models of collective action and regime change. We will also consider the impact that an extremist party can have on the choices made by the population. This third party is usually a personal interest group who benefits from a revolution and thus will try to persuade the general population into revolting. An easy example of this would be arms dealers who receive increased sales if a revolution were to occur. Clearly this type of model can be used to add insight to revolutions in the present day as well as throughout history. However, the uses of this model can expand even beyond these examples.

Take, for example, adoption of a new technology. This is also an example of a regime change; the players are deciding whether or not to adopt this new technological advancement or to keep the current method in place. There is even a third party who is trying to sway the public opinion into adopting the new technology. This “extremist" may be the company in charge of new technology and their payoff increases as more people adopt their advancement. In this example, the “extremist" action could be the amount of advertising they choose to purchase in an attempt to increase the number of citizens who adopt their technology.

The layout of the paper is as follows: The first section contains the model of interest and defines the necessary variables. I then look at existence theorems and examine some applications of this model. Finally, the conclusion wraps up the paper and gives ideas for future research.

The Model

Imagine there is a current regime governing a country. As with all regimes a revolt is a very real concern. Represent the current population being governed as a continuum of players each deciding whether to participate in the revolt or not. We will say that if a player plays action a_1 then they participate in the revolt and if that player plays action a_0 then they do not participate in the revolt.

The population has some anti-government sentiment, Θ, which is normally distributed with mean θ and variance σ_{Θ}^2. That is Θ~N(θ, σ_{Θ}^2). This variable represents how much the population as a whole values a change of regime. In addition, each citizen has his or her own personal anti-government sentiment, θ_i, given by the equation: $\theta_i = \Theta + \varepsilon_i$ with ε_i normally distributed with mean 0 and variance σ_{ε}^2, ε~N(0, σ_{ε}^2). It is important to note that Θ and ε_i are independent from each other for each *i*. So, each player has type θ_i which measures the benefit that player *i* earns if the current regime is overthrown, and that benefit is some independent deviation away from the population mean.

Additionally, each citizen has a type associated with his or her cost of revolting, c_i, which is uniformly distributed over positive values in the interval $[\underline{c}, \overline{c}]$. So c_i~U($\underline{c}, \overline{c}$) where $0 < \underline{c} < \overline{c}$. It would not make sense for the typical player to have a negative cost of revolting, which is why we have the positive restrictions on the values of $\underline{c}$ and $\overline{c}$.

The payoff for player *i* is given by the following matrix:

		Result of Revolt: Success	Result of Revolt: Failure
P_i	a_1= revolt	$\theta_i + \rho - c_i$	$s - \mu - c_i$
	a_0= don't revolt	θ_i	s

Let N measure the proportion of the population that chooses to revolt. To formulate a realistic uncertainty of the result of the revolt, we will let the success of the revolt be dependent on N. That is *Pr(success) = N*. The variables θ_i and c_i are player's individual types drawn as described above. The status quo payoff, *s*, is the payoff received by each player under the current regime. The value *s* is common knowledge for all players. Thus if a revolution fails everyone earns a payoff of *s*. Finally, notice that in the event of a successful revolution, there is a reward reserved for those citizens who participated in the revolution.

This “privilege” is equal to ρ with (0 ≤ ρ ≤ 1) this means that in a successful revolution everyone receives θ_i, but the additional ρ is reserved for those who helped to overthrow the previous regime. This means there is an incentive for a player to revolt (play a_1) if he or she believes that a revolution will be successful.

After learning his or her types, player *i* formulates beliefs about the anti-government sentiment of the whole population, Θ. To do this I will use Bayes' Rule in the case of normal priors and normal signals as seen in Degroot (1970) and I use Morris and Shin (2003) in the case of uniform priors and posteriors. This gives us a posterior belief for player *i*. He or she now believes that Θ is normally distributed with mean $\overline{\theta_i}$ and variance σ_1^2 where $\overline{\theta_i} = \lambda\theta_i + (1-\lambda)\bar{\theta}$ and $\sigma_1^2 = \lambda\sigma_\varepsilon^2$ with $\lambda = \frac{(\sigma_\varepsilon^2)}{(\sigma_\Theta^2 \sigma_\varepsilon^2)}$. Notice that $\overline{\theta_i}$ is increasing in θ_i, this means that the more anti-government an individual is the more anti-government he or she believes the rest of the population is, and thereby increases the likelihood that he or she will choose to revolt (play a_1).

As in Baliga and Sjostrom (2012), we will add an “extremist" party to this model whose goal is to manipulate the actions chosen by the population in order to increase its own payoff. The extremists choose to expend some level of energy on producing a signal which will be observed by the population. The players then update their beliefs again before playing the revolution game. The extremist chooses some level of effort, e in [0,1), which they will use to send a message which will be seen by the entire population. The message is a public signal which will send the population some information about the true value of Θ similar to the policy effects in Angeletos, Hellwig, and Pavan (2006). Notice that the extremist does not know the true value of Θ but rather is trying to convince the general population that Θ is large and therefore the population's anti-government sentiment is high. The message, $m()$, is a function of population anti-government sentiment, Θ, and effort, e. In particular, $m(\Theta, e) = \Theta + e + \eta$, with η normally distributed with mean 0 and variance σ_η^2, $\eta \sim N(0, \sigma_\eta^2)$. Notice that the message the extremist sends increases in both Θ and e, this implies that the more anti-government sentiment present in a society the easier it is to persuade that population to revolt. Edmond (2013) finds that the number of information manipulations that a population witness makes the regime easier to overthrow. This can be represented in my model by aggregating all the messages received into one value of m.

As mentioned above, after the extremist issues their message, the population updates their beliefs for a second time. Following the information gained from the extremist's message, player i now has beliefs that Θ is normally distributed with mean $\bar{\bar{\bar{\theta}}}$ and variance of σ_2^2.

$\bar{\bar{\bar{\theta}}} = \psi(m - e*) + (1 - \psi)\bar{\bar{\theta}}$ and $\sigma_2^2 = \psi\sigma_\eta^2$ with $\psi = \frac{(\sigma_\eta^2)}{(\sigma_2^2 + \sigma_\eta^2)}$. In addition, each member of the population has some belief about the level of effort expended by the extremists. I denote this value as e* and acknowledge that this value is common knowledge between all players.

At this point, player i has some belief about the value of Θ, as well as having received his or her cost signal. So, player i can use the available information to generate one type that uniquely describes himself or herself, $y_i = \theta_i - c_i$. Then, he or she can use y_i to formulate some beliefs about player j's signal $y_j = \theta_j - c_j$. From the information received up to this point, player

i believes that the opposing player's signal is normally distributed with mean $\bar{\bar{\bar{\theta}}} - c_i$ and variance σ_2^2.

The sequence of the game is as follows:

1. Players receive their types and formulate beliefs about other players
2. The extremist sends their message
3. Players update their beliefs
4. The Revolution Game is played

The extremist party receives their payoff after the revolution game is played as it is dependent on the number of citizens who revolt. The payoff for the extremists is given by $u(N) - f(e)$ where $u(N)$ describes the benefit received when a proportion N of the population chooses to revolt. Furthermore, $f(e)$ measures the cost of exerting e level of effort. Notice that, $\frac{\partial}{\partial N} u(N) \geq 0$ and $\frac{\partial}{\partial e} f(e) \geq 0$.

We now consider the previously introduced variable, $y_i = \theta_i - c_i$. It can be seen that y_i is the benefit of revolting minus the cost of revolting for player i. A player will revolt if y_i is large enough. This is similar to saying that if θ_i, the benefit of revolting, is large enough or if c_i, the cost of revolting, is small enough then player i benefits more from revolting (playing a_1) than if he or she chose to not revolt (play a_0). I will assume that the players will play a cutoff strategy. That is, given some cutoff point x, player i will choose to revolt if and only if $y_i \geq x$, otherwise he or she will not revolt.

We begin by finding the expected payoff for player i if he chooses to revolt.

$$Pr(success)(\theta_i + \rho - c_i) \; + \; (1 - Pr(success))(s - \mu - c_i) \tag{1.1}$$

Similarly, we find the expected payoff for player i if he chooses to not revolt.

$$Pr(success)(\,\theta_i) + \; \big(1 - Pr(success)\big)(s) \tag{1.2}$$

As mentioned above, player *i* will choose to revolt if equation (1.1) is greater than equation (1.2). That is, if

$$Pr(success)(\theta_i + \rho - c_i) + (1 - Pr(success))(s - \mu - c_i) \geq Pr(success)(\theta_i) + \big(1 - Pr(success)\big)(s).$$

Rearranging this inequality we obtain expected net payoff.

$$Pr(success)\rho - \mu + Pr(success)\mu - ci \geq 0 \tag{1.3}$$

If equation (1.3) is true then player *i* will choose to revolt. Note if the left hand side of equation (1.3) is equal to 0 then player *i* is indifferent between revolting and not revolting, but for the purposes of this paper we will assume that he or she will revolt.

Now we consider player *j*'s best response if player *i* follows the strategy defined using the expected net payoff function above. To achieve this best response function I use the strategy found in Baliga and Sjostrom (2011) and subtract y_j from player *i*'s expected net payoff and multiply by -1. This results in the inequality:

$$y_j \geq \theta_j - c_j + c_i - Pr(success)\rho + \mu - Pr(success)\mu \tag{1.4}$$

Notice that equation (1.4) is a cutoff rule for player *j*. Thus, given the probability of a successful revolution, $Pr(success)$, player *j* will revolt and play a_1 if and only if the inequality in equation (1.4) is true. It is also important to note that this can be expanded a further. First, assume player *i* is playing the strategy defined by cutoff point x, then

$$Pr(success) = Pr(\theta_i - c_i \geq x).$$

In fact, by using Player *j's* beliefs about Player *i*'s type we find

$$Pr(success) = (1 - \Phi(\frac{x - \bar{\bar{\bar{\theta}}} + c_i}{\sqrt{\sigma_2^2}}))$$

where Φ is the CDF of a normal distribution. Thus, equation (1.4) becomes

$$y_j \geq \theta_j - c_j + c_i - (1 - \Phi(\frac{x - \bar{\bar{\bar{\theta}}} + c_i}{\sqrt{\sigma_2^2}}))\,\rho + (\Phi(\frac{x - \bar{\bar{\bar{\theta}}} + c_i}{\sqrt{\sigma_2^2}})\,\mu \tag{1.5}$$

Particular attention should be drawn to the right hand side of equation (1.5) as that can be used to define a function Γ. In the following equation, let *x* be a cutoff point for player -*i*.

$$\Gamma(x) = \theta_j - c_j + c_i - (1 - \Phi(\frac{x - \bar{\bar{\bar{\theta}}} + c_i}{\sqrt{\sigma_2^2}}))\,\rho + (\Phi(\frac{x - \bar{\bar{\bar{\theta}}} + c_i}{\sqrt{\sigma_2^2}})\,\mu \tag{1.6}$$

It is also important to note that I am only interested in symmetric equilibria, as discussed by Shadmehr and Bernhardt (2011). This implies that $\theta_j - c_j = y_j$, which is equivalent to the cutoff point for player *j*. Thus the equation becomes

$$\Gamma(x) = x + c_i - (1 - \Phi(\frac{x - \bar{\bar{\bar{\theta}}} + c_i}{\sqrt{\sigma_2^2}}))\,\rho + (\Phi(\frac{x - \bar{\bar{\bar{\theta}}} + c_i}{\sqrt{\sigma_2^2}})\,\mu \tag{1.7}$$

Notice that $\Gamma(x)$ is the best response function for player *j*. Given cutoff points for all the other players, this function will determine the cutoff point that player *j* should use to maximize his or her utility. More importantly, any fixed point of this function defines an equilibrium for the model. That is, cutoff point *x* is an equilibrium if $\Gamma(x) = x$.

I declare that at least one equilibrium for this game exists. The outcome in which each player chooses to never revolt is an equilibrium. That is, each player will choose to play a_0 regardless of the signal that he or she receives. This outcome is represented by having each player's cutoff point be equal to infinity. Thus $y_j < \infty$ for all i, which implies that no player will ever revolt. Proving that this outcome is an equilibrium is easily verified by setting $x = \infty$ in the right hand side of equation (1.5), the best response function for player *j*.

Now I will show that an equilibrium in which there is a nonzero chance of participation exists. Since I am assuming that equilibrium is symmetric we can say that each player will play the same cutoff point, *x*. If a player has type greater than this value he or she will choose to revolt, but if his or her type is less than that value he or she will not revolt. However, if a player has type equal to that value, $\theta_i - c_i = x$, then he or she is indifferent between revolting and not revolting. For those players who are indifferent, the best response function becomes

$$\Gamma(x) = x + c_i - (1 - \Phi(\frac{x - \bar{\bar{\bar{\theta}}} + c_i}{\sqrt{\sigma_2^2}}))\,\rho + (\Phi(\frac{x - \bar{\bar{\bar{\theta}}} + c_i}{\sqrt{\sigma_2^2}})\,\mu \tag{1.8}$$

In equilibrium $\Gamma(x) = x$, thus equation (1.8) simplifies to

$$0 = c_i - (1 - \Phi(\frac{x - \bar{\bar{\bar{\theta}}} + c_i}{\sqrt{\sigma_2^2}}))\,\rho + (\Phi(\frac{x - \bar{\bar{\bar{\theta}}} + c_i}{\sqrt{\sigma_2^2}})\,\mu$$

Therefore an equilibrium is defined by

$$c_i = (1 - \Phi(\frac{x - \bar{\bar{\bar{\theta}}} + c_i}{\sqrt{\sigma_2^2}}))\,\rho + (\Phi(\frac{x - \bar{\bar{\bar{\theta}}} + c_i}{\sqrt{\sigma_2^2}})\,\mu \tag{1.9}$$

Notice this equation states that in equilibrium the marginal player has equal cost and benefit of revolting. I will now use this equality to show that an equilibrium with positive participation exists. The second term in equation (1.7) is the probability that a revolution is successful, $1 - \Phi$, multiplied by the payoff that is reserved for those who chose to revolt, ρ and Φ multiplied by the punishment incurred by revolters in a failed attempt, μ. Thus that value is the expected benefit received for revolting which I will now refer to as $EB()$.

Using this new terminology, it can be see that the best response function, equation (1.6) is a function of the newly defined $EB()$. In fact, Γ(x)=$\Gamma(x) = y_j + c_j - EB(\theta, c)$. Therefore, in order to achieve an equilibrium, we want$EB(\theta, c) = c_i$.

I will first show that if $EB = c$ then θ is positive.

Theorem 1. Any value of θ, such that $EB = c$, is positive.

Proof. Assume that $EB = c$. Since $c > 0$ by definition we know $EB > 0$. Now, $EB() = (1 - \Phi(\frac{x-\bar{\bar{\theta}}+c_i}{\sqrt{\sigma_2^2}}))\, \rho + (\Phi(\frac{x-\bar{\bar{\theta}}+c_i}{\sqrt{\sigma_2^2}})\, \mu$ and $1 - \Phi\left(\frac{x-\bar{\bar{\theta}}+c_i}{\sqrt{\sigma_2^2}}\right) > 0$ by definition. This implies that ρ and μ must both be positive, which is true by definition. ■

Theorem 2. A strategy profile *x* which defines an equilibrium with positive participation exists if and only if $EB = c$

Proof. For the purposes of simplifying the above equations, let $A = \frac{(\psi-1)\lambda}{\sqrt{\sigma_2^2}}$, $B = \frac{1}{\sqrt{\sigma_2^2}}$, and $C = \frac{x-\psi(m-e*)-(1-\lambda)(1-\psi)\bar{\theta}}{\sqrt{\sigma_2^2}}$. We then create a function $f(\hat{\theta}, \hat{c}, m) = A\hat{\theta} + B\hat{c} - C$. This equality is achieved by substituting in the values of $\bar{\bar{\bar{\theta}}}$ and $\bar{\bar{c}}$ into equation (4) and then making the necessary substitutions for A, B, and C. Thus the expected net payoff function, equation (1.3), becomes

$$c_i - \left(1 - \Phi\left(f\left(\hat{\theta}, \hat{c}, m\right)\right)\right)\rho + \left(\Phi\left(f\left(\hat{\theta}, \hat{c}, m\right)\right)\right)\mu \tag{1.10}$$

So, as long as $EB(\theta, c) = c_i$ we can declare that an equilibrium exists. To do the remaining proof I will write $EB(\theta, c)$ as $\left(1 - \Phi\left(f\left(\hat{\theta}, \hat{c}, m\right)\right)\right)\rho + \left(\Phi\left(f\left(\hat{\theta}, \hat{c}, m\right)\right)\right)\mu$.

Now to find the fixed point of Γ(x), it is enough to find the point(s) where $EB() = c$. To do this, let us consider the shape of $EB()$. I will first show that $EB()$ is increasing in θ. By differentiating and rearranging we find that equation (4) is increasing in θ if and only if

$$A \leq \frac{1 - \Phi\left(f\left(\hat{\theta}, \hat{c}, m\right)\right)}{\hat{\theta}\phi\left(f\left(\hat{\theta}, \hat{c}, m\right)\right)} \tag{1.11}$$

and decreasing otherwise. Thus, we can prove that $EB()$ is single peaked by showing that there exists a θ for which equation (1.11) is true thereby making $EB()$ increasing, and by showing there exists a θ for which equation (1.11) is false which would make $EB()$ decreasing. First, let us consider $\lim_{\theta\to 0} EB()$, notice that the numerator will be some positive number whereas the denominator will approach 0. Thus, $\lim_{\theta\to 0} EB() = \infty 1$, which implies that there is a value of θ for which $EB()$ is increasing.

Now we consider $\lim_{\theta\to\infty} EB()$.

$$\lim_{\theta\to\infty} EB() = \lim_{\theta\to\infty} \frac{1-\Phi\left(f\left(\hat{\theta},\hat{c},m\right)\right)}{\hat{\theta}\phi\left(f\left(\hat{\theta},\hat{c},m\right)\right)}$$

$$\lim_{\theta\to\infty} \frac{1-\Phi\left(f\left(\hat{\theta},\hat{c},m\right)\right)}{\hat{\theta}\phi\left(f\left(\hat{\theta},\hat{c},m\right)\right)} = \lim_{\theta\to\infty} \frac{-\phi\left(f\left(\hat{\theta},\hat{c},m\right)\right)f_\theta(\hat{\theta},\hat{c},m)}{\phi\left(f\left(\hat{\theta},\hat{c},m\right)\right)+\hat{\theta}\phi'\left(f\left(\hat{\theta},\hat{c},m\right)\right)f_\theta(\hat{\theta},\hat{c},m)}$$

$$\lim_{\theta\to\infty} \frac{-\phi\left(f\left(\hat{\theta},\hat{c},m\right)\right)f_\theta(\hat{\theta},\hat{c},m)}{\phi\left(f\left(\hat{\theta},\hat{c},m\right)\right)+\hat{\theta}\phi'\left(f\left(\hat{\theta},\hat{c},m\right)\right)f_\theta(\hat{\theta},\hat{c},m)} = \lim_{\theta\to\infty} \frac{-\phi\left(f\left(\hat{\theta},\hat{c},m\right)\right)f_\theta(\hat{\theta},\hat{c},m)}{\phi\left(f\left(\hat{\theta},\hat{c},m\right)\right)+\hat{\theta}\left(f\left(\hat{\theta},\hat{c},m\right)\right)\phi'\left(f\left(\hat{\theta},\hat{c},m\right)\right)f_\theta(\hat{\theta},\hat{c},m)}$$

$$\lim_{\theta\to\infty} \frac{-\phi\left(f\left(\hat{\theta},\hat{c},m\right)\right)f_\theta(\hat{\theta},\hat{c},m)}{\phi\left(f\left(\hat{\theta},\hat{c},m\right)\right)+\hat{\theta}\left(f\left(\hat{\theta},\hat{c},m\right)\right)\phi'\left(f\left(\hat{\theta},\hat{c},m\right)\right)f_\theta(\hat{\theta},\hat{c},m)} = \lim_{\theta\to\infty} \frac{f_\theta(\hat{\theta},\hat{c},m)}{\hat{\theta}\left(f\left(\hat{\theta},\hat{c},m\right)\right)f_\theta(\hat{\theta},\hat{c},m)}$$

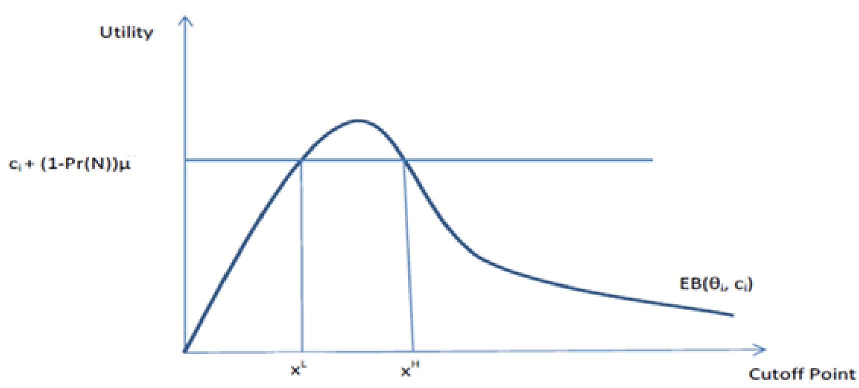
Finally, notice that this value is 0 since the numerator is a constant but that the denominator increases as θ increases. Thus, there exists a θ for which $EB()$ is decreasing. Therefore, $EB()$ is single peaked.

Similarly, by differentiating and rearranging, we find that $EB()$ is decreasing in *c* if and only if B ≥ 0, which is always true. This result comes from the definition of B above. Therefore, we see that $EB()$ is single peaked and continuous. From this we determine that an equilibrium exists as long as $EB()$ equals c_i. This implies that the height of the peak will alter the number of equilibria that exist.

If the maximum value of $EB()$ falls below c_i, then an equilibrium in which a revolution occurs does not exist. If the maximum value is exactly c_i then there is only one equilibrium with successful revolution. Finally, if the maximum value is greater than c_i then there are two equilibria in which there is a nonzero probability of revolting. Thus, an equilibrium exists, but the exact number depends on the maximum value of $EB()$. ■

Figure 1.1 below shows a graphical representation of $EB()$, as in Bueno de Mesquita (2014) the curve is single peaked.

Figure 1.1



I will now focus on the situation in which there are three equilibria, one with zero participation and two with nonzero participation. Let x^{∞} define the equilibrium with no participation. Furthermore, let the equilibria with nonzero participation be defined by x^L and x^H where $x^L < x^H$. It can now be shown that x^{∞} and x^L are stable equilibria, but x^H is not stable.

Theorem 3. The equilibrium defined by x^L is stable, but the one defined by x^H is unstable.

Proof. Since the equilibria of the game are symmetric we can say that Γ(x) intersects the 45 degree line at x^L and x^H . In particular Γ(x) crosses the 45 degree line from below at x^L and from above at x^H . Therefore $|\frac{d}{dx}\Gamma(x^L)| < 1$ and $\left|\frac{d}{dx}\Gamma(x^H)\right| > 1$ ■

For the rest of the paper I will assume that players will not end up at x^H because of the previous result.

I will now determine the effort that the extremist party will choose in equilibria. I begin by considering the equilibrium defined by an infinite cutoff point. Recall that x = ∞ implies that there is no participation from the population. Because of this, the extremist will choose to use minimum effort, that is, $e = 0$. This can easily be seen by considering the extremists payoff, $u(N) - f(e)$. If $x = \infty$, then nobody chooses to revolt which implies that $N = 0$. Therefore, the extremists are now maximizing $u(0) - f(e)$ which is achieved when $f(e)$ is at its minimum. By definition, $f(e)$ is smallest when *e* is at its lowest possible value.

Now we consider the stable equilibrium with positive probability of revolting. Again we consider the extremist's problem: $max\ u(N) - f(e)$ choosing the level of effort, *e*, to produce. In order to examine this problem further we must determine the value of *N*.

Recall, that *N* is the number of citizens who are choosing to revolt from our continuum of players. By definition, this means that *N* is also the proportion of players that revolt as the total population is normalized to 1. To determine this value, we consider an arbitrary player *j* . I declare that Player *j* revolts if his or her type is greater than or equal to the cutoff point *x* that defines the equilibrium. That is, Player *j* revolts if $\theta_j - c_j \geq x$. Using the definitions of θ_j and c_j we know Player *j* will revolt if

$$\Theta + \varepsilon_j - c_j \geq x \tag{1.12}$$

Now if all players use cutoff point x, which is true in equilibrium, then N is simply the proportion of players for which equation (1.12) is true.

Thus,

$$N = 1 - \Phi \frac{x - \bar{\theta} + c_i}{\sqrt{\sigma_\Theta^2 + \sigma_\varepsilon^2}} \tag{1.13}$$

So the extremist's problem becomes $\max u\left(1 - \Phi \frac{x - \bar{\theta} + c_i}{\sqrt{\sigma_\Theta^2 + \sigma_\varepsilon^2}}\right) - f(e)$. Since I assumed that players will not reach x^H, I can now simplify the extremists problem. Since we are only considering equilibria at this moment, it is important to note that in equilibrium the players' belief about the effort exerted by the extremist is correct. That is, $e = e*$. I previously showed that the extremist chooses $e = 0$ in the equilibrium defined by x^∞. Similarly, in the equilibrium defined by x^L the extremist chooses to exert level of effort $e = e*$, with $e* > 0$.

Comparative Statics

I first consider the equilibrium x^L from the above model. I will now show that the cutoff point x which defines the equilibrium is increasing in ρ, the payoff reserved for players who help in the regime overthrow. I will also show later, that in the event of a fixed c_i, the equilibrium is also decreasing in cost, c.

Theorem 4: The equilibrium cutoff point is increasing in ρ.

Proof. Consider the best response function, equation (1.7), at $x = x^L$ which defines an equilibrium. The comparative statics can be found by differentiating with respect to the parameter of interest.

$$\frac{\delta}{\delta\rho}\Gamma(x^L) = \theta_i(1 - \Phi\left(\frac{x-\bar{\bar{\theta}}+c_i}{\sqrt{\sigma_2^2}}\right)) > 0$$

This inequality is true because 0 ≤ 1 – Φ ≤ 1 by definition and $\theta_i \geq 0$ in equilibrium as was previously shown. ■

Let us now consider the probability that a revolution will actually occur. We know this will happen as long as $N \geq 0$. Essentially, we want to know what is $\Pr(N \geq 0)$. First, I will define *N*, the number of people who revolt.

$$N = \Pr(\Theta + \varepsilon_i - c_i \geq x)$$

Which can be rewritten as

$$N = \left(1 - \Phi\left(\frac{x - \bar{\theta} + c_i}{\sqrt{\sigma_\varepsilon^2 + \sigma_\Theta^2}}\right)\right)$$

Therefore the probability that a revolution occurs is equivalent to the proportion of the population for which $N \geq 0$. So

$$Pr\left(1 - \Phi\left(\frac{x - \bar{\theta} + c_i}{\sqrt{\sigma_\varepsilon^2 + \sigma_\Theta^2}}\right)\right) \geq 0$$

Since *N* is normally distributed we can rewrite the equation

$$1 - \Phi\left(\frac{-1 + \Phi\left(\frac{x - \bar{\theta} + c_i}{\sqrt{\sigma_\varepsilon^2 + \sigma_\Theta^2}}\right)}{\sigma^2}\right) \tag{1.14}$$

Where σ^2 is the variance of N. Equation (1.14) is represents the probability that a revolution will be successful. As long as (1.14) is not equal to 0 a revolt has the potential to be successful regardless of the number of participants, this supports the findings of Schelling (1960) and Hardin (1996) on spontaneous revolution.

An interesting implication of this model, that seems to contradict most research on the topic of regime change, is that the extremist in this model does not convince a population to revolt. In fact, if we are in equilibrium, the effort that the extremist spends on their message, e, is equal to the level of effort that the population believes they have exerted, $e*$. This means that, on the average, the players correctly guess the value of e. One might argue that an extremist is only present in those countries which experience an overthrown regime. However, according to this model, the presence of an extremist solely represents that the country is ripe to revolution. This can be seen because an extremist will only choose a nonzero amount of effort in the equilibrium x^L. If, instead, the country is in the equilibrium represented by x^∞ then the extremist will choose to produce effort $e = 0$. This means that only those countries that have a nonzero probability of revolution will have an active extremist. Essentially, an extremist will only choose to send a message to the population if they can influence the decision and since this only happens in the equilibrium defined by x^L , that is the only equilibrium where the extremist puts forth nonzero effort.

Application: Common Knowledge Cost

Consider a situation in which the cost of revolting is a fixed value and that it is common knowledge for all players. I will represent this by setting $c_i = c$ for all i, where $c > 0$ is an arbitrary constant. By doing this, I create a model related to the one considered above, in which the players only have one type, their anti-government sentiment, similar to the model presented by Bueno de Mesquita (2010). As before, we have a continuum of players choosing between actions a_1 and a_0. We also set cost equal to a constant, $c_i = c$. The status quo payoff remains equal to value *s* as before. This application considers the scenario in which only anti-government sentiment is player dependent and cost is the same for all players. This results in the following model:

		P_{-i}	
		$N \geq T$	$N < T$
P_i	a_1	$\theta_i - c$	$s - c$
	a_0	$(1-\gamma)\theta_i$	s

In this application, player's types are only the anti-government sentiment as cost is the same for all players. Again, I assume that players will play cutoff strategies in this game.

As before, players receive their types and then create beliefs about the other players' types. In this model, that means player *i* sees his or her anti-government sentiment, θ_i, which is still normally distributed as before. After receiving his or her type, player *i* formulates a belief about player *j*'s type. Since, in the introductory model, the formulated beliefs of θ_j are independent of a player's cost type, I declare that player *i*'s belief about player *j*'s type remains the same as before. The only difference in this application is that player *i* does not formulate a belief about player *j*'s cost type, as that is common knowledge.

Next, the extremist chooses the level of violence to display which then leads the players to re-update beliefs further, just as we saw above. Assume for a moment, that player *j* is using cutoff strategy *x*, in our previous model this meant that he would revolt (play a_1) if $\theta_j - c_j \geq x$, and would not revolt (play a_0) otherwise. However, in this application, c_j is a fixed value. Thus, we can simplify this strategy by saying that player *j* will revolt if $\theta_j \geq x$. This means, from Player *i'*s point of view, if arbitrary Player *j* plays with cutoff point *x*, the probability that Player *j* revolts is $\Pr(\theta_j \geq x)$. Given Player *i*'s beliefs about Player *j*'s type, $\Pr(\theta_j \geq x) = 1 - \phi(x - \bar{\bar{\theta}})$. As before, we can use this probability to define the proportion of the population that will revolt. This further implies that, given a cutoff point *x* for Player -*i*, Player *i's* best response is $\Gamma(x) = \theta_i + c - \gamma\theta_i\left(1 - \phi\left(x - \bar{\bar{\theta}}\right)\right)$. The fixed point of Γ(x) is the equilibrium in this example which was proven to exist earlier.

The extremist present in this model behaves strictly as a player who wants to see the revolution occur, and tries to persuade the other players into revolting, by sending a message about the value of the population's anti-government sentiment. In this example, the extremist's payoff depends on whether or not a revolution occurs. That is, $u(N \geq T) - 1$ and $u(N < T) = 0$ while still experiencing the costs associated with more effort spent on the message. So, the Extremist's payoffs are: $1 - f(e)$ if $N \geq T$ and -f(e) if N < T where $f(e)$ remains the cost function of providing effort level e.

The comparative statics presented previously remain true in this application. The equilibrium cutoff point is increasing in γ the reserve benefit for choosing to revolt if the revolt is successful (similar to ρ). However, now there is an additional parameter which can impact outcomes. It is easy to show that the equilibrium is also decreasing in c. This tells us that a government could potentially stifle attempts at a revolution by increasing punishment for players who participated in a failed revolution. This implies an increase in the cost of revolting and therefore a reduction in the number of players who choose to revolt.

Application: Uncertain Payoffs

In this next application, we consider a situation in which the exact value of the benefit of a successful revolution is unknown. This represents a more realistic situation than many of the other examples, as the precise payoff of a person's actions may not be entirely realized until after the game has finished. Players may not know the exact payoffs that they will receive from different outcomes, but they will use all the information available to them to formulate a belief about what they will be.

As before, we have a continuum of players. We also declare that a revolution is successful if a proportion of the population greater than or equal to T decides to revolt. We then have cost of revolting c_i equal to a constant c. We also choose a value γ in such a way that $(1-\gamma)\theta_i = \theta_i - c$. This results in a model which is similar to the one presented by Shadmehr and Bernhardt (2011).

		Outcome	
		Success	*Failure*
P_i	a_1	$\theta_i - c$	$s - c$
	a_0	$\theta_i - c$	s

In this model, the exact value of θ_i is unknown, which implies that the exact value of θ_i −c is also unknown. Thus, a player's type is the signal that they receive about the value of θ_i. The players observe this value and then formulate beliefs about the other player's types as described above. Using these beliefs, they then play a cutoff strategy as was previously defined. The uncertainty of the payoffs causes this game to exhibits both games of strategic complements and games of strategic substitutes. This can be seen from the shape of the expected net payoff function. The function is single peaked which implies that for cutoff values less than the maximum point, players are more likely to revolt if the other player is more likely to revolt. However, if the cutoff strategy is to the right of the maximum then players are less likely to revolt if the other player is more likely to revolt. It can be explained intuitively as follows: If Player *i* is willing to revolt even when receiving a poor signal, then player *j* is less likely to revolt since he or she begins to doubt the actual value of a successful revolution.

Application: Pro-government Extremist

Now we consider the situation in which the extremist instead wants to see the current regime stay in power. This is represented by a simple change in which the extremist's message becomes $m(\Theta, e) = \Theta - e + \eta$. Instead of adding the value of the effort, it is subtracted from the populations' anti-government sentiment, Θ, along with some noise. This implies that the value of *m* is now lower. This impacts $\bar{\bar{\bar{\theta}}}$ by lowering its value relative to the previous model. Thus, each player now views society as having a lower anti-government sentiment. This, in response will decrease the value of $\bar{\bar{\bar{\theta}}}$ in equation (9), which defines an equilibrium cutoff point. The lower value of $\bar{\bar{\bar{\theta}}}$ increases the numerator thereby increasing the argument inside the normal CDF. Thus the ρ component of equation (9) is now smaller, but the μ component is bigger.

Taking the derivative with respect to $\bar{\bar{\bar{\theta}}}$, we get

$$\frac{\partial}{\partial \bar{\bar{\bar{\theta}}}} = -\rho\phi\left(\frac{x - \bar{\bar{\bar{\theta}}} + c_i}{\sqrt{\sigma_2^2}}\right) - \mu\phi\left(\frac{x - \bar{\bar{\bar{\theta}}} + c_i}{\sqrt{\sigma_2^2}}\right) < 0$$

Thus, when $\bar{\bar{\bar{\theta}}}$ decreases, the cutoff point also decreases, thereby causing more of the population to revolt and leading to a higher probability of a successful revolution.

Conclusion

As seen above, the model presented in this paper has several useful applications to real world scenarios. Some possible additions to the literature would be adding the government as a player and having them try to quell the revolutionaries by giving additional signals to the population. In this situation, one could investigate the possibility that the government can successfully threaten the citizen; would a threat be enough or would they have to follow through with these threats? One could also generalize the model further by having arbitrary distributions and verifying that equilibria still exist.

2 The Impact of No Child Left Behind on Teacher Turnover

Introduction

No Child Left Behind (NCLB) was signed into law by George W. Bush on January 8th 2002. The law's purpose was to increase student achievement. A particular emphasis was made on reducing the achievement gap between high-performing and low-performing schools. The law attempted to accomplish this goal by striving to make every student proficient in math and reading as measured by a standardized test. Any school that did not show the necessary improvements in proficiency between years would have sanctions placed on them. These sanctions ranged from a warning for the first year a school did not reach the required level of proficiency, all the way to complete government takeover of the school if it failed to reach the required level for six consecutive years. Although state-level accountability systems did exist prior to NCLB, this type of federal accountability system had never been seen before in elementary and secondary schools in the U.S. This was a fundamental change in the field of education and as such could have a direct impact on teachers. The way that teachers respond to this type of accountability system directly impacts the achievement of students since any kind of turnover negatively affects a student's test scores (Hanushek, Rivkin, and Schiman 2016).

The academic gains made by students under NCLB have been largely positive. Initial reports including Rouse et al (2007) and Rockoff and Turner (2008) found that NCLB achieved its goal by increasing test scores in target schools and target areas by anywhere between 0.04 and 0.2 standard deviations. This is a modest effect in absolute terms, but does represent about 7% of the effect witnessed between students having a mother who is a high school graduate versus a high school dropout as seen in Dahl and Lochner (2012)[1]. Ballou and Springer (2011) and Dee and Jacobs (2011) also find that NCLB does have a positive impact on student achievement, particularly in math. While most research involving NCLB has been rightly focused on academic achievement, there is surprisingly very little inquiry into the impact it has had on school personnel. It is intuitive that an accountability system as extreme and pervasive as NCLB would have some sort of effect on teachers as well as students. The papers that do study teachers focus on how teachers responded to NCLB and how they approach teaching under an accountability system. Manna (2011) finds that teachers changed the way they teach following NCLB. Several papers also consider how NCLB changes the way teachers approach the field, such as spending more time on tested subjects (math and reading) as found by Dee, Jacobs, and Schwartz (2013), searching for better or more efficient teaching styles as in Murnane and Papay (2010), or even to teach to the test (Jacob 2005; Figlio and Rouse 2006).[2]

There is minimal research on how teachers are responding to accountability systems, such as NCLB. One, possibly unintentional, impact of this type of reform, on teachers is a change in teachers' job satisfaction. Barksdale-Ladd (2000) and Hoffman, Assaf, and Paris (2001) find that, under an accountability system, teachers feel more pressure to deliver high test scores. Cavanaugh (2012), Gerson (2007), and Toppo (2007) all find that accountability systems increase pressure which leads to teacher stress. Others also find that teachers feel a reduction in job security (Finnigan and Gross 2007, Luna and Turner 2001, Grissom, Nicholson-Crotty, and Harrington 2014). This pressure felt by teachers is amplified when considering teachers at schools that are at or below Adequate Yearly Progress (AYP), (Reback et al. 2014). If teachers

[1] They find that students whose mothers are high school graduates had test scores that were 17% of a standard deviation higher than students whose mothers were high school dropouts.
[2] This happens when the curriculum being taught is heavily focused on only preparing the students for what they will see on a standardized test. This type of teaching is thought to lack passion and meaning, and can even be considered unethical by some as it is not an accurate representation of the abilities of the teacher.

are feeling more pressure, less job security, or generally less job satisfaction, then it is likely that more teachers will leave the field. This is valuable since most research shows that teacher turnover has a negative effect on students' learning (Ronfeldt, Loeb, and Wyckoff (2013); and Hanushek, Rivkin, and Schiman (2016)).

Sun, Saultz, and Ye (2014) use pre-NCLB and post-NCLB data to study teacher attrition caused by NCLB, but ignore the staggered introduction of accountability systems in states with a prior policy. Several states had a statewide accountability system in place before NCLB. The assumption used in the literature is that states with a prior accountability system are not affected or treated by NCLB because the federal government used the existing systems as guides when they drafted NCLB. However, the 26 states that had some prior form of an accountability system before 2002 experienced different implementation years of these programs. This means that the states with a prior system were "treated" during different years. I account for this staggered introduction of accountability systems.

This paper will fill the gap in the literature by answering three questions: 1) How are teachers responding to NCLB? In particular, why are teachers reacting this way to NCLB and which subset of the teaching population is the driving force behind these changes? 2) Can we generalize these changes observed in NCLB and compare them to the effects observed from state level accountability systems that were introduced prior to 2002? Is there a difference between NCLB and these previous systems? 3) Is NCLB changing who chooses to become a teacher? What effect does NCLB have on the demographic information of college graduates with degrees in education?

The rest of the paper is as follows. In the next section I introduce a theoretical model that is representative of an individual teacher who has been affected by NCLB. I then describe the data sources and explain the methodology used in the paper. Following that I look at the results obtained from my analysis and discuss the results. Finally, I conclude.

Theoretical Model

The following section introduces the theoretical model I will use to explain a teacher's earnings. I will then use this model to determine how changes in environment or experience can influence a teacher's likelihood to remain teaching. To do this, I introduce an accountability system into the model by adding a wage shock to the teacher. Using this shock I am able to compare how the likelihood of staying in the field changed after the shock was introduced.

Consider a model in which workers are trying to maximize their utility. Since utility is a monotonic transformation of wage, I consider the natural logarithm of the individual's wage that they are currently earning which is given by the following equation:

$$w_i = \alpha t_{ij} + \beta S_j + \tau X_i + \gamma h_{ij} t_{ij} + \varepsilon \qquad (2.1)$$

where t_{ij} is worker *i's* tenure at job *j*, S_j represents the environmental effects of job *j*, X_i is a vector of worker characteristics. Thus wage is a function of an individual's tenure at their current job, as well as where they work and any individual characteristics that are important (Topel and Ward 1992). Finally, h_{ij} is a measure of worker-job match. That is, h_{ij} measures worker *i's* aptitude for job *j*. Note that h_{ij} is not a simple measure of ability, but rather a measure of the quality of the match between the worker and the job (Jovanovic 1979). Finally, ε is the error term with $E[\varepsilon] = 0$. A worker's wage increases in tenure, experience as measured in X_i because the workers gain general skills as well as specific skills through on the job training (Neal 1995, Parent 2000). Note that a higher aptitude for a particular job leads to higher wages. This represents the fact that the worker is more productive and is compensated for that higher production. Additionally, the product $h_{ij}t_{ij}$ shows that a more productive worker benefits more from more experience than less productive workers.

Assume that a teacher with characteristics X is employed at a school with firm effect S'. The teacher has a worker job match of h' and therefore is earning a wage of

$$w = \alpha t_{teach} + \beta S' + \tau X' + \gamma h' t_{teach} + \varepsilon \tag{2.2}$$

We will compare this to an industry job with firm effect $\hat{S}$ and a worker-job match of $\hat{h}$ which will have wage

$$w = \alpha t_{other} + \beta \hat{S} + \tau X' + \gamma \hat{h} t_{other} + \varepsilon \tag{2.3}$$

Additionally assume that all workers are employed for N years before retiring so in the model $t \epsilon [0,N]$. Now assume that at time $t=T$ there is wage shock experienced by the teacher, such as the introduction of an accountability system, which affects the worker-job match. We will represent this by shifting the value of the current worker- job match, h, by the value of the shock, v.

So the teacher's new wage becomes

$$w = \alpha t_{teach} + \beta S' + \tau X' + \gamma (h + v) t_{teach} + \varepsilon \tag{2.4}$$

Notice that the shock has a greater magnitude on longer tenured teachers.

Following this shock, a teacher can choose to remain in the field and earn their post-shock wage or they can leave the field and become employed at a different job to avoid the effect of the shock. So we can now compare all the possible wages available to this teacher.

$$teaching\ wage = \alpha t_{teach} + \beta S' + \tau X' + \gamma h' t_{teach} + \varepsilon \tag{2.5}$$

$$post\_shock\ teaching\ wage = \alpha t_{teach} + \beta S' + \tau X' + \gamma (h' + v) t_{teach} + \varepsilon \tag{2.6}$$

$$other\ industry\ wage = \alpha t_{other} + \beta \hat{S} + \tau X' + \gamma \hat{h} t_{other} + \varepsilon \tag{2.7}$$

So, if a wage shock occurs, a teacher's decision can be simplified to choosing whether to stay in the field or to leave the field. The graphs below show what happens to the teacher's lifetime utility.

Figure 2.1: Lifetime Utility Following a Wage Shock

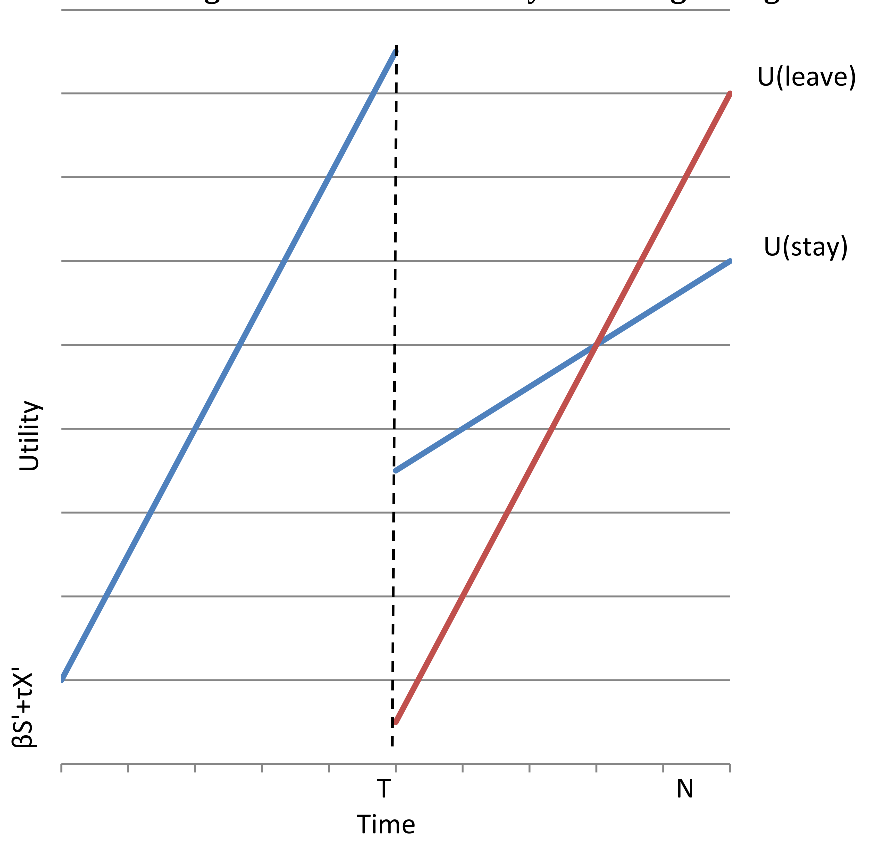


Figure 1: Left of T shows log wage of a teacher before the wage shock takes place. Right of T shows their decision after it occurs, continue teaching and earn log wage measured by U(stay) or leave teaching and earn log wage U(leave).

We can say that lifetime utility for a teacher who remains in the field is given by:

$$U(stay) = \sum_{t=0}^{T} \delta^t(\alpha t_{teach} + \beta S' + \tau X' + \gamma h' t_{teach} + \varepsilon) + \sum_{t=T}^{N} \delta^t(\alpha t_{teach} + \beta S' + \tau X' + \gamma(h' + v) t_{teach} + \varepsilon) \quad (2.8)$$

Whereas the lifetime utility for a teacher who leaves and becomes employed elsewhere is given by:

$$U(leave) = \sum_{t=0}^{T} \delta^t(\alpha t_{teach} + \beta S' + \tau X' + \gamma h' t_{teach} + \varepsilon) + \sum_{t=T}^{N} \delta^t(\alpha t_{other} + \beta S' + \tau X' + \gamma h' t_{other} + \varepsilon) \quad (2.9)$$

Which means a teacher only needs to consider the net change in lifetime utility by leaving the field, which is given by:

$$U(change) = U(leave) - U(stay)$$

$$= \sum_{t=T}^{N} \delta^t(\alpha t_{other} + \beta \hat{S} + \tau X' + \gamma \hat{h} t_{other} + \varepsilon - \sum_{t=T}^{N} \delta^t(\alpha t_{teach} + \beta S' + \tau X' + \gamma(h' + v)t_{other} + \varepsilon) \quad (2.10)$$

If this equation is greater than 0, then the teacher gets more utility from leaving the profession, but if it is negative then the teacher will remain in the field. If it is equal to 0 then we will assume that the teacher will stay.

Solving the sum, the above equation becomes:

$$U(change) = (\tfrac{N-T}{2})\delta^{N-T}(\alpha(N-T) + 2\beta\hat{S} + 2\tau X' + \gamma\hat{h}(N-T) - \alpha(N+T) - 2\beta S' - s\tau X' - \gamma(h' + v)(N+T)) \quad (2.11)$$

Which simplifies to

$$U(change) = (\tfrac{N-T}{2})(-\delta^N \alpha T - \delta^T \alpha T + \delta^N \beta(\hat{S} - S') + \delta^T \beta(\hat{S} - S') + \delta^N \gamma(\hat{h}(N-T) - (h' + v)(\delta^N N + \delta^T T)) \quad (2.12)$$

Now we can look at the comparative statics for this utility function.

$$\frac{\partial U}{\partial S'} = \left(\frac{N-T}{2}\right)(-\delta^N \beta - \delta^T \beta) \leq 0$$

$$\frac{\partial U}{\partial h'} = \left(\frac{N-T}{2}\right)(-\delta^N \gamma N - \delta^T \gamma T) \leq 0$$

$$\frac{\partial U}{\partial \hat{S}} = \left(\frac{N-T}{2}\right)(\delta^N \beta + \delta^T \beta) \geq 0$$

$$\frac{\partial U}{\partial \hat{h}} = \left(\frac{N-T}{2}\right)\left(\delta^N \gamma(N-T)\right) \geq 0$$

$$\frac{\partial U}{\partial v} = \left(\frac{N-T}{2}\right)(-\delta^N \gamma N - \delta^T \gamma T) \leq 0$$

Here we see that as S', the schools environment, increases then the net utility from leaving decreases. This is also true if h', the worker-job match for teaching, increases. However, if either $\hat{S}$, the environment for the other profession, or $\hat{h}$, the worker-job match for the non-teaching profession, increases then the net utility from leaving also increases which makes it more likely that the individual will leave. Finally, if the value of ν, the utility shock, increases then the individual is more likely to remain in teaching.

Using this model, I will test to determine what effect an accountability system has on the likelihood of leaving. Additionally, we can test to see if an increase in teacher aptitude makes an individual more likely to stay in the field. Since, job-worker match is hard to measure, we can instead use a measure of happiness as a proxy. We can test the effect work environment has on a teacher's decision to stay. To do this, I can look at a teacher's satisfaction with their current school and stratify my results along those findings.

Data and Methods

I use three waves of the Schools and Staffing Survey (SASS) to combine cross-sectional time series data on teachers and schools. SASS is a nationally representative survey of schools and school personnel issued by the National Center for Education Statistics (NCES). I pair these three waves with the corresponding waves of the Teacher Follow-Up Survey (TFS) which is also issued by NCES. The SASS waves that I use contain information during the 1993-1994, 1999-2000, and 2007-2008 school years. While the TFS waves I use are issued sometime during the following year, 1994-1995, 2000-2001, and 2008-2009. Since NCLB was signed into law in January of 2002, I have two pre-NCLB waves, the 1993-94 and 1999-2000 school years, and one post-NCLB wave, the 2007-08 school year. This differs from the current literature since I am only using one post-NCLB wave.

As Reback, Rockoff, and Schwartz (2014) showed, states had some control over how to implement NCLB. A state had to submit its own guidelines to the government which then had to be approved under the NCLB guidelines. This meant that a state's personal plan had to be approved by the federal government. This approval process would have caused a delay in when NCLB was implemented among states. Therefore, while NCLB was passed in 2002 and theoretically would have gone into effect during the 2002-2003 school year, it is ambiguous as to whether or not a state would have experienced any effects from NCLB until much later. Thus, I am omitting the SASS data for the 2003-2004 school year. Many other papers such as Dee and Jacobs (2011), Sun, Saultz, and Ye (2014), and Grissom, Nicholson-Crotty, and Harrington (2014) use all four waves. I believe this is a mistake because any change in turnover during this time period would not be caused by the accountability system itself, but instead the future expectations of the system being in place. More recent waves of the SASS are available, such as the 2011-2012 wave, but I chose not to use them just as Sun, Saultz, and Ye (2014) chose not to use them, as any teacher turnover post-2008 could be strongly influenced by the Great Recession making it difficult to separate any effects of NCLB from those caused by the recession.

However, this concern does not apply to the 2007-2009 SASS and TFS waves since the recession began in 2008 after the SASS survey was completed and any decision to stay or leave the profession was already made at that point and likely not influenced by the Great Recession (Hyatt and McEntarfer 2012).

I remove private school teachers from the dataset as they were not affected by NCLB. While it is true that some private schools chose to follow NCLB guidelines (Christensen et. al.

2007), it is impossible to know which ones from the SASS datasets. Furthermore, it is difficult to know whether or not any sanctions were placed on any private schools that did not meet Adequate Yearly Progress (AYP) or if the schools just wanted to see where they stood in relation to other schools.[3] Finally, I removed observations from my dataset if their contract was not renewed by the school as I want to view how NCLB affected voluntary attrition and so I only want to consider the individuals who chose to leave the field of teaching during this timeframe. Figure 2.2 shows the timing of the SASS and TFS and how each individual teacher is labeled once they respond to the TFS.

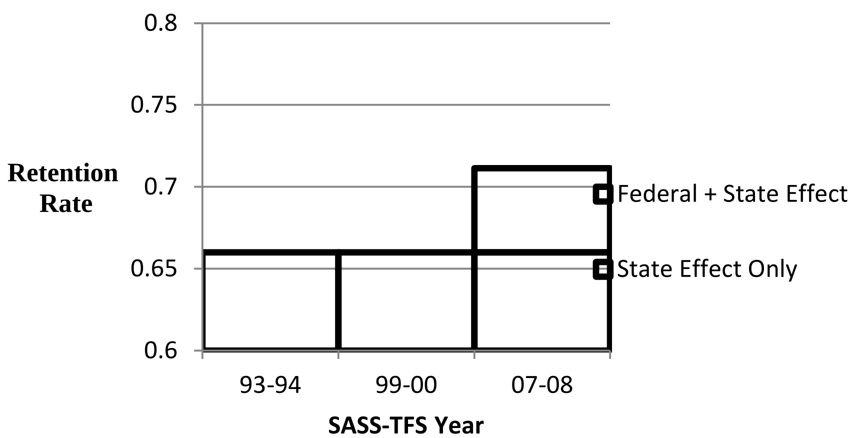
Figure 2.2: SASS-TFS Wave for school year *t-t+1*	
Year *t* SASS	Year *t+1* TFS
All teachers in sample.	**Stayers**: continuing to teach in the TFS **Leavers**: Anybody not teaching in the TFS. (Contains retirees, terminated, and voluntary leavers)

The method I use in this paper is similar to the method found in Dee and Jacobs (2011) and Sun, Saultz, and Ye (2014), a difference-in-differences model to determine the effect of not only NCLB, but state-level teacher accountability systems as well. As in Dee and Jacobs (2011) as well as Sun, Saultz, and Ye (2014), I consider a model in which a comparison is made between a "treatment group," consisting of states that had no accountability system prior to NCLB, and a "comparison group" made up of states that had some form of an accountability policy before NCLB. As Sun, Saultz, and Ye (2014) explain, the premise of this strategy is that NCLB was

[3] I do include the private teachers as a robustness check later in the paper.

modeled after the accountability systems that existed in states such as Texas, Tennessee, and North Carolina. From 1990 to 2000, 26 states formed their own accountability systems for teachers. The federal government then used these programs as a basis for NCLB (Murnane and Papay 2010). I use the dating method used by Dee and Jacobs (2011) which is presented in Appendix Figure 2.1. The basic idea in the model is that states that transitioned from a previous accountability system to NCLB would not experience any effects as they are already "treated" by the previous program. However, the 24 states that did not have any prior system will be truly "treated" by NCLB. This difference in effects is illustrated in Figure 2.3.

Figure 2.3: The Effect of Accountability Systems on Teacher Retention



I use a linear probability model as follows:

$$Y_{ijst} = \beta_0 + \beta_1 T_s + \beta_2 YEAR + \beta_3 (T_s * YEAR) + \beta_4 X_{ijst} + \beta_5 S_{jst} + \beta_6 STATE_{st} + \varepsilon_{ijst} \quad (2.13)$$

The model looks at the likelihood of teacher i in school j in state s during year t remaining in the field of teaching following that year. The dependent variable is a dummy which equals 1 if teacher i is still in the field in the Teacher Follow-up Survey. T_s is a treatment indicator which is equal to 1 if the year is post-introduction of an accountability system in that state. I also control for year fixed effects, *YEAR*, as well as individual teacher characteristics, X_{ijst}, school characteristics, S_{jst}, and state fixed effects, $STATE_{st}$.

I also use this model to consider the effect that any previous accountability systems had on teacher turnover. In the 1993-1994 wave of SASS, only 2 states had some kind of accountability system in place (Illinois and Wisconsin). However, by the next SASS wave in 1999-2000, that number had risen to 26. Therefore, we can use a difference-in-differences model to find the effect that these accountability programs had on these 24 states.

Results

Table 2.1 shows descriptive statistics for teachers during all years of the survey. I further separate teachers between leavers and stayers. Notice that the teachers that choose to stay during these 3 waves are on average 6 years younger than those that leave and have nearly 6 fewer years of experience. The percentage of private school teachers, Asian teachers, black teachers, Hispanic teachers, and Native American teachers largely does not vary between stayers and leavers. Finally notice that leavers are less likely to have a bachelor's degree, but more likely to have a master's degree. This seems to imply that teachers from both ends of the education spectrum are leaving the field.

Table 2.1. Summary Statistics			
	(1) All Teachers	(2) Stayers	(3) Leavers
Age	40.5	38.4	44.6
Experience	11.7	9.7	15.6
Private	0.25	0.25	0.25
Native American	0.018	0.017	0.021
Asian	0.026	0.026	0.026
Black	0.066	0.068	0.062
Hispanic	0.05	0.054	0.042
Female	0.72	0.74	0.70
Union	0.53	0.53	0.53
Salary	33,470	33,226	33,941
Bachelors	0.97	0.98	0.95
Masters	0.38	0.35	0.44
Stay	0.66	1	0
N	25,210	16,610	8,600

Table 2.2 shows how selected variables differ among teachers across years and across the teacher’s leave status. The average age of stayers steadily decreases over time while the age of leavers doesn’t follow any noticeable pattern. Average total experience of stayers remains relatively constant prior to NCLB but drops significantly afterward. By pairing this decrease in experience with the increase in retirement rates among leavers, it seems as if NCLB could be pushing some teachers out of the field. The percentage of non-white stayers also rises over time. Finally, the percentage of teachers that have a bachelors and the percentage that have a masters rise over time for both stayers and leavers.

Table 2.2. Teacher by Year by Status						
	(1)	(2)	(3)	(4)	(5)	(6)
	93-94	99-00	07-08	93-94	99-00	07-08
	Stayers	Stayers	Stayers	Leavers	Leavers	Leavers
Age	38.5	38.2	36.7	45.1	44.1	44.9
Experience	9.7	9.8	7.7	15.9	15.8	15.5
Private	0.30	0.27	0.15	0.25	0.20	0.24
Native American	0.013	0.014	0.016	0.021	0.021	0.019
Asian	0.019	0.021	0.026	0.020	0.025	0.023
Black	0.052	0.061	0.070	0.050	0.054	0.071
Hispanic	0.050	0.058	0.063	0.036	0.042	0.049
Female	0.73	0.73	0.73	0.66	0.68	0.75
Union	0.56	0.48	0.57	0.56	0.52	0.54
Salary	26,099	31,506	39,401	28,195	33,618	41,914
Bachelors	0.97	0.98	0.98	0.96	0.96	0.96
Masters	0.32	0.33	0.35	0.44	0.45	0.47
N	3,970	4,380	3,480	2,310	2,370	1,260

Table 2.3 shows the results from the simple model in which we do not control for any individual or school characteristics. This table shows the impact of NCLB on turnover between the 1999-2000 and 2007-2008 schools years. The reform effect for the 2007-2008 school year is positive and significant. It appears that any prior accountability system does not affect teacher turnover, but NCLB increases the likelihood of staying by over 5 percentage points.

Table 2.3: Effect of Federal Accountability System on Staying	
	(2)
	’99-‘07
Type of reform	Federal
Reform	0.0513***
	(0.0191)
Observations	8,050
Number of state	50
Year FE	YES
State FE	YES

OLS model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Below, I include a graphical representation of the average age and experience of teachers over time. These results can be found in Figure 2.5 and Figure 2.6 respectively. In both age and total experience there is a sharp decline in stayers after 2003-04. However, among leavers, the average age and experience is slightly increased following the signing of NCLB. The effect on age and experience after a state-level accountability system appears to be nonexistent. These descriptive statistics seem to suggest that teachers are reacting to NCLB by leaving the field which allows for new teachers to take their place thereby lowering the average age and experience significantly, which contradicts the findings of the model in Table 2.3. This type of result is possible because the sheer number of teachers in field increased following NCLB. The average student to teacher ratio for the United States fell during my time frame, 17.3 in 1995 to 15.3 in 2008[4], which suggests that there are more teachers overall after NCLB was passed. If

[4] National Center for Education Statistics: Fast Facts on Teacher Trends (https://nces.ed.gov/fastfacts/display.asp?id=28)

most of these new teachers were younger individuals with little experience it allows the previous figures and the effects found in Table 2.3 to be true.

Figure 2.5: Average age over time

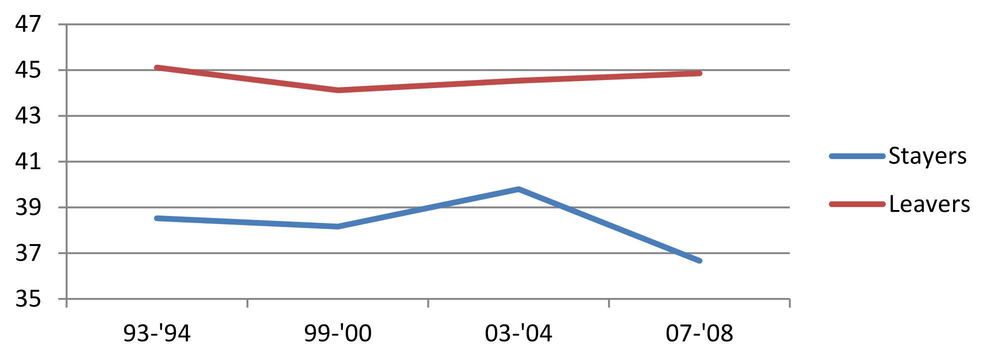
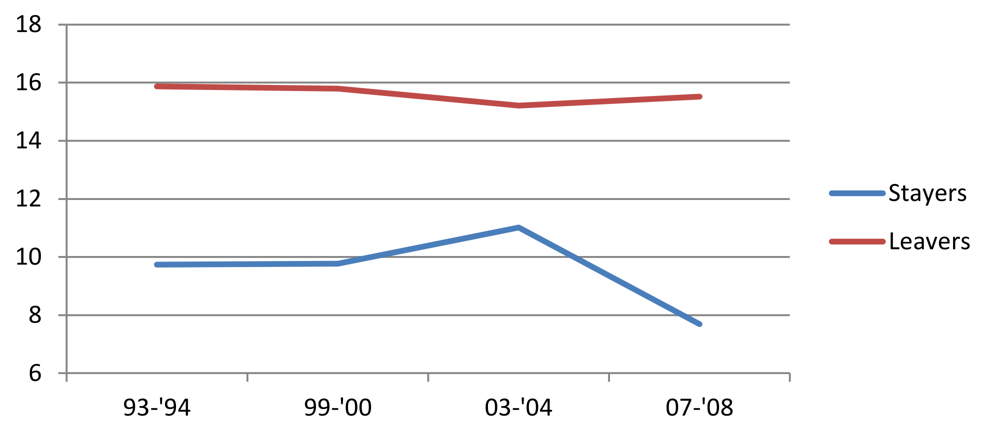


Figure 2.6: Average experience over time



Accepting the results found in the tables and figures above, it appears teachers are more likely to remain in the field following a national accountability system such as NCLB, even though the media suggests that NCLB is largely viewed as unpopular. Following the typology suggested by Perrow (1972) this would lead to more variety and higher analyzability among tasks. One possible explanation is that following implementation of an accountability system, the field of education becomes less craft-like and more engineering-like which causes a shift in what is expected of teachers. This shift makes teaching more structured and mechanical and thereby will chase few teachers away even if they dislike the changes.

The next question to answer is whether or not different subgroups are affected differently. I first look at whether NCLB affects male teachers and female teachers differently. The results are in Table 2.4.

Table 2.4: Effect of NCLB on Staying Stratified by Sex

	(1) '99-'07 Females	(2) '99-'07 Males	(3) Diff
Type of reform	Federal	Federal	Federal
Reform	0.0468**	0.0353	0.0115
	(0.0226)	(0.0360)	(0.0425)
Observations	5,780	2,270	
Number of state	50	50	
Year FE	YES	YES	
State FE	YES	YES	

OLS with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. The results are stratified by gender. Column 3 shows a difference in mean test for the 2 columns it follows. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Columns 1 and 2 show the models where I use the 1999-2000 and 2007-2008 school years. These results show the effect of NCLB. Female teachers experience a slightly smaller but more significant effect of NCLB whereas males experience a larger but less significant effect. However, column 3 shows that the male and female teachers are not significantly different from each other.

Table 2.5: Effect of NCLB on Staying by Education Level			
	(1) ’99-‘07 Bachelors	(2) ’99-‘07 Masters	(3) Diff
Type of reform	Federal	Federal	
Reform	0.0864***	0.00236	0.0840**
	(0.0235)	(0.0332)	(0.0407)
Observations	4,820	3,120	
Number of state	50	50	
Year FE	YES	YES	
State FE	YES	YES	

OLS with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher was still teaching in the Teacher Follow-up Survey. The results are stratified by education level. I dropped the teachers with no Bachelor’s. Column 3 shows the results of a difference in mean test for the 2 columns it follows. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

I then test how teachers react to accountability systems based on the level of education they have. I am dropping the 5% of teachers who did not have either a Bachelor’s or a Master’s degree. Interestingly, the effect of NCLB is large, positive, and significant on teachers who solely have a bachelor’s degree as seen in column 1. This effect disappears once a teacher earns a master’s degree represented in column 2. This is most likely explained by the fact that after receiving a master’s degree, the teacher has more job opportunities available to them and can pursue a different career.

Table 2.6: Effect of NCLB on Staying by Age			
	(1)	(2)	(3)
	Less than 25	25-35	35-60
Type of Reform	Federal	Federal	Federal
Reform	0.0743**	0.0522	0.0410
	(0.0375)	(0.0354)	(0.0293)
Observations	1,370	2,370	3,890
Number of state	50	50	50
Year FE	YES	YES	YES
State FE	YES	YES	YES

OLS model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. The federal reform is NCLB. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

I then consider whether or not teachers respond to accountability systems differently depending on their age. Columns 1 through 3 show the impact that a federal level accountability system, such as NCLB has on the likelihood of various age groups remaining in the field. NCLB has a significant effect on those teachers who are less than 25 years old, but no significant effect for any other age group.

Table 2.7: Effect of NCLB on Staying by Experience

	(1) NCLB 0-3 years experience	(2) NCLB 3-15 years experience	(3) NCLB 15-30 years experience
Type of Reform	Federal	Federal	Federal
Reform	0.0562**	0.0310	0.0454
	(0.0255)	(0.0381)	(0.0499)
Observations	3,180	2,500	1,570
Number of state	50	50	50
Year FE	YES	YES	YES
State FE	YES	YES	YES

OLS model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. The federal reform is NCLB. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

In Table 2.7 I show how the effect of an accountability system is impacted by a teacher's experience. Columns 1 through 3 show the impact that a federal reform, or NCLB, has on a teacher's likelihood of staying stratified along experience. Column 1 shows new teachers with less than 3 years of experience, column 2 shows the effect on teachers with 3-15 years of experience, and column 6 shows the effect on teachers with more than 15 years of experience. NCLB has a positive and significant effect for new teachers. This impact disappears as the individual gains more experience. This suggests that teachers are more likely to stay in the field following NCLB if they began teaching after it was enacted, since fewer than 3 years of experience during the 2007-2008 school year would imply that the teacher, more than likely, earned their degree after 2002.

Table 2.8: Effect of NCLB on Staying by Year of Bachelor's Degree

	(1)	(2)	(3)
	Pre reform	Post reform	Diff
Type of reform	Federal	Federal	
Reform	0.0291	0.299***	0.270***
	(0.0251)	(0.0838)	(0.0875)
Observations	6,420	1,640	
Number of state	50	50	
State FE	YES	YES	
Year FE	YES	YES	

OLS model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. The federal reform is NCLB. Column 3 shows the results of a difference mean test for the 2 column it follows. Standard errors are clustered at the state level.

*** p<0.01, ** p<0.05, * p<0.1

I test this result by stratifying according to the year that a teacher earned their bachelor's degree. I split them into 2 groups: teachers who earned their degree before the accountability system was passed and those who earned their degree after. The results are in Table 2.8. Column 1 shows the effect of NCLB on those teachers who earned their degree before 2002, when NCLB was signed into law. Column 2 shows the effect for those who earned a Bachelor's degree after 2002. A teacher who earned their degree before NCLB took effect was not significantly impacted by the reform. However, teachers who worked in states that were first affected by NCLB were more likely to stay in the field if they earned their degree after 2002 when NCLB was signed into law. Column 3 shows that the difference between the two groups is significant at the 0.01% level.

This can be explained by a self-selection process. Anyone who became a teacher after 2002 knew that NCLB was signed and that they would have to teach under its laws. Any prospective teachers that felt they would not be able to do so, chose to pursue another discipline.

This is further emphasized in the survey results as every single teacher surveyed majored in education for their Bachelor's degree. Therefore, there were no individuals who majored in some non-education discipline that became teachers.

Table 2.9: Effect of NCLB on Staying by View on Continuing Teaching

	(1) '99-'07 would teach again	(2) '99-'07 would not teach again	(3) Diff	(1) '93-'99 would teach again	(2) '93-'99 would not teach again	(3) Diff
Type of reform	Federal	Federal	Federal	State	State	State
Reform	0.0564*** (0.0203)	-0.0369 (0.0699)	0.0933 (0.0728)	-0.00129 (0.0218)	-0.156*** (0.0540)	0.155*** (0.0582)
Observations	6,410	910		1,320	6,350	
Number of state	50	50		50	50	
Year FE	YES	YES		YES	YES	
State FE	YES	YES		YES	YES	

OLS model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. A state reform is any state-level accountability system enacted before 2002. The federal reform is NCLB. Columns 3 and 6 show the results for a difference in mean test for the 2 columns it follows. A teacher's willingness to teach again is from the survey question "Would you be a teacher again?" Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Now I test all testable hypotheses from my model in section II. As was found at the end of that section, the likelihood of an individual teacher staying in the field increases as h, the variable measuring job-worker match, rises. Since it is difficult to measure job-worker match, I will use satisfaction with the career as a proxy. On the SASS, there is a question that asks whether or not the individual would be a teacher if they could go back to college and start over. I create a new dummy variable equal to 1 if the teacher responded "yes" to this question. I then stratify along this satisfaction variable. The results are presented in Table 2.9. Notice that for a prior accountability system, an individual is more likely to leave the field if they are unhappy with the field of teaching as seen in column 2. However, the difference between those teachers that are happy and those that are not is not significant for the federal reform, with those results found in column 6. This is because, according to the model in section II, we are measuring the effect of $h+v$, job-worker match plus the wage shock. Additionally, the sign and magnitude of v is unknown and therefore makes the sign and magnitude of $h+v$ unknown. This uncertainty can explain why the difference is no longer significant for the federal reform.

Table 2.10: Effect of Accountability System on Satisfaction Stratified by Year of Bachelor's Degree

	(1) Pre	(2) Post	(3) Diff	(4) Pre	(5) Post	(6) Diff
Type of reform	State	State	State	Federal	Federal	Federal
Reform	-0.1246* (0.0685)	0.0176 (0.0566)	0.1422 (0.0889)	0.0136 (0.0320)	0.105 (0.135)	0.0914 (0.1387)
Observations	5,120	5,540		6,300	1,600	
Number of state	50	26		50	50	
Year FE	YES	YES		YES	YES	
State FE	YES	YES		YES	YES	

OLS model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. A state reform is any state-level accountability system enacted before 2002. The federal reform is NCLB. Columns 3 and 6 each show the results for a difference in mean test for the two columns it follows. A teacher's satisfaction with teaching is determined by their response to the survey question "Would you teach again?" Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

I am interested on the effect that an accountability system has on a teacher's view on teaching. I continue to use the survey question "Would you be a teacher again?" as a proxy for their satisfaction with the field. I run an OLS model with this satisfaction variable as the dependent variable. The results are presented in Table 2.10. While there is no significant effect of either a state-level or federal-level accountability system on teacher satisfaction, the size of the difference in the effect accountability systems have on teacher happiness is notable.

Table 2.11: Effect by where the teacher earned their bachelor's degree

	(1) Treated degree	(2) Control degree	(3) Diff
Type of reform	Federal	Federal	
Reform	0.0469	0.0325	0.0144
	(0.0455)	(0.0388)	(0.0598)
Observations	3,110	4,320	
Number of state	50	50	
Year FE	YES	YES	
State FE	YES	YES	

OLS model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. A treated degree is when the teacher earned their bachelor's degree in state that was affected by NCLB. A control degree is if the teacher earned their degree in a state that had a previous accountability system. The federal reform is NCLB. Standard errors are clustered at the state level.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

For one final test, I split the sample into two categories based on where they earned their Bachelor's degree. If the degree was earned in one of the 24 states that did not have a state-wide accountability system prior to NCLB, then that teacher is said to have earned their degree in a "treated" state. The remaining teachers that earned their degree in a state with some previous system in place is said to have earned it in a "control" state. I stratify the sample into these two groups and present the results in Table 2.11. These results show if there was a difference between colleges and universities after they became "treated." Column 1 shows the effect of NCLB on the likelihood of staying for teachers who earned their degree in a "treated" state while column 2 shows the results for those who graduated from a school in the "control" group.

Neither of these results is statistically significant. Additionally, the groups are not statistically different from each other.

Discussion and Robustness Checks

One of the concerns with using the SASS-TFS questionnaires is that a smaller subset of individuals replies to the TFS than those that fill out the SASS. This could negatively impact my results if the reason for replying to the TFS is endogenous with some characteristics. However, by testing for the differences in means between the SASS sample and the TFS sample, I find the two are statistically equivalent. This can be found in Appendix Table 2.1.

NCLB, as well as any state accountability system, solely impacts public schools. Private schools are not required to obey the guidelines found in these systems. Therefore, I can use the private teachers in the sample as a robustness check to determine that only the public teachers were impacted by NCLB while private teachers remained largely unaffected. I split my sample into two groups, those that teach in a public school and those that teach in a private school. I further stratify the sample by considering the effect of the accountability systems relative to whether or not the teacher earned their bachelor's degree before the system was implemented or after. The results are presented in Appendix Table 2.2. It can be seen that the effect on private school teachers is significantly different from the effect on public school teachers for those that earned their degree after 2002, but relatively similar for teachers who earned their Bachelor's before 2002, in that neither private school teachers nor public school teachers are significantly impacted by the reform. This result is also found in statewide accountability systems at larger magnitudes. So earning a degree before the state implemented their system does not have a significant impact on the likelihood of staying regardless of whether or not the individual teaches in a private school or a public school. However, if he or she earned their degree after the accountability system was implemented in their state, then they are much more likely to stay in the field if they are working in a public school relative to their private school counterparts.

It is also likely that school characteristics for the school at which the individual teaches, largely impacts a teacher's decision to stay in the field following implementation of an accountability system. I consider many of these in the appendix. Of particular interest is the composition of students which can be seen in Appendix table 2.3 through Appendix table 2.7. Appendix Table 2.3 shows the effect stratified along the percent of students who are not native English speakers or limited English proficient (lep). I rank all the schools according to the percent of lep students that attend that school. I then split these schools into quintiles and consider the effect of an accountability system stratified by which quintile the school falls in. I consider the extreme cases in Appendix Table 2.3 of the first quintile with the lowest proportion of lep students and the fifth quintile of schools with the largest proportion. There is no significant effect on a teacher's likelihood of staying for a state-level system. Yet, for a federal-level system, teachers are 8 percentage points more likely to stay if they teach in a school with a low number of lep students. However, this is not significantly different from those that teach in a fifth quintile school.

Appendix Table 2.4 shows the effect stratified by the percent of students who are eligible for a free or reduced lunch (frl) again considering the first quintile and the fifth quintile. There is not significant effect for either state-level or federal-level accountability systems. Appendix Table 2.5 and 2.6 show the effect of an accountability system, NCLB in Appendix Table 2.5 and state accountability system in Appendix Table 2.6, stratified by racial quintiles, again comparing first and fifth quintiles for Asian students, black students, Hispanic students, Native American students, and white students. After NCLB, teachers that work at schools with the lowest proportion of black students were 23 percent more likely to stay in the profession. This effect was significantly different from those that taught in schools with the largest proportion of black students (or the fifth quintile). This effect and the difference between the quintiles is not present for the state-level system. Appendix Table 2.7 stratifies along student proficiency as measured by the schools making AYP or missing AYP in consecutive years. This table only considers the effect of a federal-level system as AYP is a component of NCLB. The effect on staying is not significant for schools that made or missed AYP for a single year.

I also include interaction terms in my model. I interact the treatment with age and experience to see if the effect of age or experience changes after the accountability system is present. The results are in Appendix Table 2.8. Column 1 shows the effect of the age interaction and experience interaction from a state-level accountability system. I then stratify according to when a Bachelor's degree was earned. Column 2 shows those teachers who earned their degree before their respective state passed an accountability system. Column 3 shows the results for those who earned their degree after it was passed. Columns 4 to 6 show the effect of interaction terms on the likelihood of teachers staying in the field for federal-level systems. Column 4 shows the effect on all teachers in the sample. Column 5 shows the effect on teachers who earned their Bachelor's degree prior to NCLB and column 6 shows the effect on those teachers who earned their degree after NCLB.

In Appendix Tables 2.10 through 2.17, I show the results from a probit model to consider the effect of accountability systems on the likelihood of staying. The logit results mirror what was found from the probit models and therefore I only include the probit results in the appendix. The interpretations of the effects do not change, although the magnitude of the effects is somewhat diminished.

As a final investigative test I determine how the general makeup of teachers changed after No Child Left Behind was signed. I focus on three main changes: the probability of an individual teacher specializing in math while earning their degree, the probability of an individual teacher specializing in reading while earning their degree, and the probability that a teacher failed at least one Praxis examination. I focus on those who specialized in math and reading because those are the two subjects tested by NCLB. This would suggest that any teacher who taught one of those two subjects would be under more scrutiny than another subject since math and reading scores directly influence whether a school makes Adequate Yearly Progress. This additional scrutiny could lead to education majors leaving those subjects and choosing some other specialization. The results from these models can be found in Appendix Table 2.18.

Additionally, I determine the likelihood of an individual teacher failing at least one Praxis test as a proxy for teacher quality. The Praxis test measures test taker's knowledge and skills. It is a crucial component of licensing and certifications for teachers. There are several content specific Praxis tests that the potential educators must take in order to be a certified teacher. I run a model with a binary dependent variable which is equal to 1 if the teacher never failed a Praxis test. The results are also presented in Appendix Table 2.18.

In column 1 of Appendix Table 2.18, I have the change in likelihood of a teacher specializing in math while earning their degree after NCLB was signed. Teachers were 12 percentage points less likely to choose this specialization after the federal reform took effect. This effect is not present for teachers specializing in reading, which can be seen in column 2. In column 3 I show how the likelihood of a teacher failing at least one Praxis test changes after NCLB takes effect. Teachers are 1.5 percentage points more likely to fail at least one Praxis test after No Child Left Behind took effect.

The introduction of NCLB seems to have a positive effect on teachers choosing to stay in the field. This seems to contradict the media's interpretation of NCLB's effect on teachers. Toppo (2007), Hefling (2012) all find that teachers describe the policy as "unfair," and provide generally unfavorable views. However, teachers' actions say otherwise as several studies considering NCLB's impact on teaching environment have shown the opposite result. Grissom, Nicholson-Crotty, and Harrington (2014) find that NCLB has a positive effect on teachers' job satisfaction[5]. This higher job satisfaction should lead to lower turnover. Similarly, Sun, Saultz, and Ye (2014) find that teachers are less likely to leave the profession after NCLB. They find an initial increase in turnover in 2003-2004, but as I discussed above, it is ambiguous whether or not that effect is caused by NCLB. Additionally, they find that teachers are more likely to stay in the year 2007-2008. Finally, Loeb and Cunha (2007) also find no evidence that NCLB has resulted in an increase in turnover rate. While there is evidence that teachers dislike NCLB (Center on Education Policy (CEP) 2006; Deniston and Gerrity 2010; Sunderman, Tracey, Kim, and Orfield 2004), it also seems that this displeasure is not large enough to leave the field.

One might think that, after NCLB or a similar accountability system was passed, the institutes of higher learning where the teachers earned their degrees, would change their curriculum to best prepare their students in how to handle the new landscape of teaching. By better preparing their students, these institutes could directly assist in lowering teacher turnover. In a somewhat surprising result, I find that the colleges and universities where the teachers are studying do not appear to have a strong influence on whether or not the teacher continues in the field. Instead, it appears that what state the teacher teaches in has a much larger effect on the

---

[5] It should be noted that they do find teachers feel they have less job security following No Child Left Behind, but overall satisfaction is up.

likelihood of an individual remaining in the field. It seems that schools in states with an accountability system before NCLB, can better prepare their new teachers for the stress that comes from operating under these types of systems. When it comes to teacher response to NCLB, it is not where you earned your degree that matters, but instead where you are currently teaching.

Since teachers who earned a degree after NCLB are more likely to stay in the field, perhaps there is a self-selection process that is occurring. I already showed that this process is not happening in the curriculum that is being presented by the colleges and universities that are preparing these teachers. Therefore, it must be happening prior to this curriculum being presented. Perhaps the general makeup of teachers is changing. That is, who chooses to become a teacher was affected by the passing of NCLB. The argument can be made that teachers may choose to avoid specializing in math or reading while earning their degree since those are the two tested subjects and therefore would put added stress on the teacher. My results do indicate that teachers avoided math education after NCLB. This suggests that potential teachers avoided specializing in what they might consider a “difficult” subject to teach. This result does not hold true for reading education, as there is not drop-off in the probability of specializing in reading after federal reform. Perhaps this is because reading is viewed as an “easier” subject to teach relative to math.

Finally, I looked at the quality of students who were becoming teachers by considering the probability of failing a Praxis test. After NCLB, potential educators were more likely to fail at least one of their Praxis examinations. This suggests that the pool of potential teachers are of a lesser quality following implementation of NCLB. It could be that a country-wide accountability system is scaring away high quality teachers because they do not wish to teach in a high stress environment.

I now consider the assumption made in this paper as well as previous literature that states with a prior accountability system are not treated by NCLB. I consider the effect of these state reforms on turnover.

Table 2.12: Effect of State Accountability Systems on Staying

	(1) ’93-‘99
Type of reform	State
Reform	-0.0147 (0.0196)
Observations	8,330
Number of state	50
Year FE	YES
State FE	YES

OLS model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

There is no significant impact of a statewide accountability system on turnover. Furthermore, stratifying these results according to age, experience, education level, and year of Bachelor’s Degree does not change this outcome. In each of these stratifications the effect of NCLB on turnover is not significant. These results do seem to highlight a problem with the model that is used as teachers do not respond the same to a prior accountability system as they do to NCLB. However, the results are still telling, since states without a prior accountability system did experience a much stronger reaction to the federal system as compared to their counterparts.

Future research can expand on and correct this issue by pairing control states and treated states that have similar guidelines and requirements.

Conclusion

I find that NCLB does not decrease the likelihood that a teacher remains in the field, but rather makes them 5 percentage points more likely to stay. Furthermore, any state accountability system that was enacted before NCLB had no measurable impact on teacher turnover. In addition, it seems that, even though NCLB was built on the pre-existing accountability polies, teachers are responding differently to NCLB than they did to prior systems.

Most subgroups of teachers also follow this trend, whether the teacher is male, female, new, experienced, young, or old, teachers are more likely to stay in the field following the passage of NCLB. The largest driving force behind these results is when the individual earned their teaching degree. If he or she graduated after NCLB was signed, then they are more likely to stay in the field relative to their colleagues who earned a degree before 2002. This seems to suggest a self-selection process among potential teachers.

Since there is no difference between whether the teacher's alma mater was located in a "treated" state or a "control" state, it seems that it is not the colleges and universities preparing their students for a career under NCLB that is causing this self-selection process to occur. Therefore, this process must be happening among the potential teachers, perhaps even before admission into an education program. Further research would include using admission data from departments and schools of education to determine the viability of this claim.

These findings can be used as an aid in understanding how accountability systems will impact student achievement as well as teacher behavior, which can further aid in the creation of educational policies that advance student learning and improve teacher quality. Given the passage of waiver system signed by President Barack Obama in 2015, states are now able to remove themselves from NCLB guidelines. The next step in determining how teachers react to accountability systems is to observe if there is any change in turnover after states have excused themselves from the federal accountability system and return to what was previously considered normal.

3 Teacher Reaction to Tenure

Introduction

For primary and secondary educators, earning tenure is an expected part of the job. In this paper I look at the effect that earning tenure has on teachers. In particular, how do teachers change their behavior in response to tenure and what impact does this change in behavior have on their students? Teacher tenure is a policy that prevents elementary and secondary teachers from being fired without just cause. It is a contract that guarantees employment except in cases of severe misconduct or incompetence. Even in instances of irresponsibility, misbehavior, or lack of performance, the process of firing a tenured teacher is time consuming and extremely expensive. Before tenure was standard, a teacher could be fired because of age, race, religion, or even favoritism. Women could even be fired if they became pregnant.

The biggest benefit that teachers receive from tenure is job security. Once they earn tenure a teacher can try new techniques or different methods that they were afraid to try before. It becomes incredibly difficult to fire a tenured teacher, as the cost to do so rises dramatically.

Critics of tenure argue that this increased cost could lead to less accountability for the teacher. Historically, experience in the school district is the primary factor (and in some cases the only factor) in making tenure decisions. Before 2009 not a single state based its tenure decisions on any measure of teacher effectiveness (2013 NCTQ Report).

Lazear (2003) illustrates the potential problems that can arise from this type of tenure process. After a teacher earns tenure it becomes very costly to fire that individual. In an 18 year span there were only 39 tenured teachers who were fired in the entire state of Illinois (Kersten 2006). Similarly, teachers are aware that the likelihood of getting fired decreases after earning tenure, this could cause the quality of education to diminish for some teachers since there is no real punishment or rather the threat of punishment is reduced (Brill 2009 and Medina 2010).

Han (2015) suggests tenure does not protect poor teachers since the districts know the quality of education issued from an individual instructor and can choose not to renew his or her contract. However, the incentive conflict occurs after the teacher has earned tenure. My question is not “does tenure protect bad teachers?” but rather “does tenure create bad teachers?” There is also no research that has been done on the effect of tenure on student performance. Jones (2015) finds evidence that teachers change their behaviors the year before they are up for tenure. This includes spending more money on their classroom, spending more time outside of school working on lesson plans and teaching strategies, and even spending more time meeting with parents. He also finds that a teacher’s behaviors return to normal within two years after earning tenure. However, tenure’s impact on student achievement is nonexistent.

Several papers have looked at the effects that teacher behavior has on student performance. Ronfeldt, Loeb, and Wyckoff (2013) find that teacher turnover negatively impacts student test scores, a finding reinforced by Hanushek, Rivkin, and Schiman (2016). Hanushek (1971), Murnane and Phillips (1981), Rivkin, Hanushek, and Kane (2005), as well as Kane and Staiger (2008) all look at various teacher characteristics and the affect they have on student achievement. Sass, Hannaway, Xu, Figlio, and Feng (2012), Taylor and Tyler (2012), and Rothstein (2014) all look at the impact schools and school policy can have on student achievement. Yet, in all of these papers, tenure is absent. I aim to remedy that by determining the effect of earning tenure on student test scores.

I use a pooled cross sectional data set of student level data and pair it with teacher and school data during the same time frame. I find that immediately after earning tenure, student test scores drop. However, for each year of teaching after earning tenure those test scores begin to climb, quickly overtaking pre-tenure levels.

Background on Teacher Tenure

The history of tenure dates back to the 19th century when teachers had little to no protection from being fired. In 1885, the National Education Association (NEA) asked for political assistance in protecting teachers. Massachusetts became the first state to pass a pre-college tenure law for teachers in 1886. New Jersey was the first state to pass a comprehensive tenure law that protected all elementary and secondary teachers in 1909. Currently, in most states, a teacher is awarded tenure after he or she has taught for a certain period of time, somewhere between 1 and 5 years. Following the Great Depression, teacher unions gained a lot of power and fought for job protection and benefits, and by the 1950's, 80% of all K-12 teachers were tenured (McGuinn 2010). As of 2008, 2.3 million teachers in America are tenured and this does not include those in post-secondary education (Stephey 2008). Each state has its own regulations for awarding tenure to a teacher, but generally the probationary period ranges from 1 to 5 years.

Several states have begun to modify their tenure laws (Christie and Zinth 2011). New York, for example, now holds their teachers accountable for the quality of their instruction that students are receiving as well as putting an emphasis on merit and performance. Colorado has made it possible for tenured teachers to lose their tenure. Ohio and California have increased the number of probationary years that a teacher must teach before earning tenure.

Many of these changes came as a response to President Obama's 2009 Race to the Top program which offered over $4 billion in grants to states who require schools to consider student achievement when making tenure decisions (US Department of Education 2009). Yet, as a whole, states continue to ignore teacher performance in tenure reviews, only about half of all states use any kind of teaching evaluation while making tenure decisions (2013 NCTQ Report).

There is a lot of discussion as to the benefits and consequences that can arise from tenure. Under current tenure laws, teachers are protected from being fired for a variety of reasons that were previously concerns. Teachers can no longer be fired for discriminatory purposes nor personal grudges or favoritism. It is also illegal to fire a tenured teacher in order to hire a less expensive replacement. In addition, teachers cannot be fired for teaching new or controversial ideas.

There is, however, a negative side to tenure. By only firing a tenured teacher in the event of severe misconduct, it is impossible to impose accountability for student achievement. Since most states give tenure to all teachers after a certain period of time, there is no guarantee that all of these protected teachers are great or even sufficient educators. In addition, even if the teacher is guilty of misconduct, it is incredibly costly to fire them. For example, in 2010 the Los Angeles Unified School District spent $3.5 million in order to fire 7 teachers and the process took 5 years (Barrett 2012). The sheer magnitude of time and money spent to fire a tenured teacher will deter any serious disciplinary action in all but the most extreme cases.

Critics of tenure argue that the policy fosters mediocrity among students and lowers the academic achievement in schools. A recent study by the New Teacher Project finds that 81% of schools admit to having low-performing tenured teachers, but over half of those schools will not act because of the tenure laws (Weisberg et al. 2009). However, those in favor of teacher tenure fear that abolishing these laws will once again make teachers susceptible to unjust firings.

Theoretical Model

Consider a general utility formula of

$$\delta u = w + f(e) - e + (b + q)(v - u) \tag{3.1}$$

Where δ is a discount factor and $0 < \delta < 1$, u is the utility obtained from working, *w>0* is the wage received from the current job. The amount of effort an individual exerts in their profession is represented by e, and *f(e)* is function of effort that measures the psychological payoffs a teacher receives from working (Lortie 1975). Let $f(0) = 0$, and $f'(e) > 0$. The variable b is the probability of the individual being fired for reasons not related to performance (downsizing, budget cuts, etc), q is the probability of being fired for shirking, with both $0 < b < 1$, and $0 < q < 1$. Finally, *v* is the utility earned from being unemployed.

My model closely follows Stiglitz and Shapiro (1984) model of shirking and efficiency wages with a measure of additional utility not related to pay included. Given equation (3.1), a teacher has one single decision to make, whether to shirk or to put forth effort. If an individual chooses to shirk then $e = 0$. Additionally, $f(e) = 0$ by definition and $q > 0$. This gives us a specific utility function of

$$u_{shirk} = \frac{w + (b + q)v}{\delta + b + q} \tag{3.2}$$

If instead, a teacher chooses to put forth nonzero effort into their teaching, then $e > 0$ which implies $f(e) > 0$. Also, by not choosing to shirk, the probability of being fired for shirking is 0, $q = 0$. This gives us a utility function for working.

$$u_{work} = \frac{w + f(e) - e + bv}{\delta + b} \tag{3.3}$$

A teacher will choose to shirk if and only if $u_{shirk} > u_{work}$. This inequality simplifies to

$$q(\delta v - w) - (\delta + b + q)(f(e) - e) > 0 \tag{3.4}$$

As long as equation (3.4) is true, the individual will shirk. The first term measures the expected difference between unemployment benefits and current wage. The second term measures the discounted expected difference between benefit and cost of effort. Thus, if equation (3.4) is true, then the worker would get a higher payoff from being unemployed then the payoff received by spending positive effort.

If we assume that teachers get no psychological payoff from working, that is $f(e) = 0$, then we obtain the findings found in Stiglitz and Shapiro (1984). Unless there is a penalty associated with unemployment, everyone will shirk. However, by introducing *f(e)*, it is now possible for an individual to obtain a nonwage payoff from increased effort. Thus, by choosing not to shirk, the second term in equation (3.4) increases thereby introducing an implicit cost to unemployment benefits, the opportunity cost of teaching.

If *w* is high enough, a teacher will choose not to shirk. This cutoff wage, *x*, is obtained by solving equation (3.4) for *w*.

$$w < \frac{(\delta + b + q)\big(e - f(e)\big)}{q} + \delta v = x \tag{3.5}$$

From here, we can find comparative statics for the cutoff wage.

$$\frac{\partial x}{\partial e} = \frac{\delta - \delta f'(e) + b - bf'(e) + q - qf'(e)}{q} \tag{3.6}$$

$$\frac{\partial x}{\partial \delta} = \frac{e - f(e)}{q} \tag{3.7}$$

$$\frac{\partial x}{\partial b} = \frac{e - f(e)}{q} \tag{3.8}$$

$$\frac{\partial x}{\partial q} = \frac{(\delta + b)(f(e) - e)}{q^2} \tag{3.9}$$

$$\frac{\partial x}{\partial v} = \delta \tag{3.10}$$

If an employer pays a high enough wage the employee will choose not to shirk. For teachers who receive some psychological payoffs, this wage decreases in δ, and b if and only if $f(e) > e$. Similarly, the cutoff wage decreases in e if $f'(e) > 1$. This means that a school can pay a lower wage to its teachers and still be able to prevent them from shirking as long as the psychological payoffs received from teaching overcome the effort spent. I want to put extra emphasis on equation (3.9) as it shows that the cutoff wage increases in q, the probability of being fired for shirking, as long as $f(e) > e$. This suggests that after a teacher earns tenure, and q drops, the cutoff wage will similarly drop as long as the benefit from effort outweighs the cost.

Data and Methods

In order to examine the impact of tenure on test scores, I acquired student level data including standardized test scores. In addition to student level data, I also collected data on teacher and school characteristics which includes whether or not a teacher earns tenure in a given year. The National Assessment of Educational Progress (NAEP) is the largest nationally representative and continuing assessment of what America's students know and can do in various subject areas. Assessments are conducted periodically in mathematics, reading, science, writing, art, civics, economics, geography, history, and even technology and engineering.

Since NAEP assessments are administered uniformly using the same sets of test booklets across the nation, NAEP results serve as a common metric for all states. The assessment stays essentially the same from year to year, with only carefully documented changes. This permits NAEP to provide a clear picture of student academic progress over time. The number of questions in a given booklet is very large and each student only answers a small subset of all available questions. NAEP uses the answers received from students to predict what the score would be if these students took the entire test using Item Response Theory (IRT).

IRT determines a method that can accurately predict the probability that a given student will answer a question correctly. The probability for a student of ability θ_k to select answer i for question j is $Pr_{ij}(\theta_k)$. Once outcome x is observed, the likelihood function can be found:

$$\Pr(x|\theta_k, item\ parameters) = \prod_{j=1}^{n} \prod_{i=0}^{m_j-1} Pr_{ij}(\theta_k)^{u_{ij}}$$

Where m_j is the number of categories for answers to question j for the given question, and u_{ij} is an indicator variable equal to

$$u_{ij} = \begin{cases} 1\ if\ response\ x\ is\ in\ category\ i \\ 0\ otherwise \end{cases}$$

The likelihood function develops a relationship between as student's answers and the parameters of the question as well as to the student's ability. The question's parameters are found using marginal maximum likelihood methods which are iterative procedures used to estimate these parameters from an initial distribution of scale scores. Based on this initial distribution, interim estimates of parameters are calculated which are then used to recalculate new parameters. This process is continued until the values converge on estimates for the IRT model. These estimates can be used to find a likelihood function for a student by using the items answered by that student. This likelihood function can then be used to make inferences about score distributions from samples where no student answers all of the items. For NAEP calculations, estimates of questions parameters are obtained using a NAEP BILOG/PARSCALE program, which combines the findings of Mislevy and Bock (1982) and Muraki and Bock (1991).

The NAEP then normalizes these results so that a value of 0 is the mean and a value of 1.0 is one standard deviation away from the mean. I will use this normalized value as the measure of student achievement from 2003 to 2007, which gives me 3 years of data since the NAEP is issued every other year. I use the NAEP results based on representative samples of students in grades 4 and 8. These grades and ages are chosen by the NCES because they represent critical junctures in academic achievement. These assessments follow the framework developed by the National Assessment Governing Board, and use the latest advances in assessment methodology.

I will follow Jones (2015) and take advantage of the staggered issuance of tenure. Some states such as Texas and New York assign tenure after 3 years, whereas some states such as Indiana and Missouri award teachers tenure after 5 years. Some states even award tenure after one year of experience (Mississippi for example). Existing literature finds that the vast majority of teachers teach close to where they grew up. Reininger (2011) finds that 85% of teachers teach within 40 miles of their hometown. Using these results, I make the assumption that the time before tenure is exogenous to the teacher and they are not using it to determine where they teach. The list of tenure requirements for each state is found in Appendix Table 3.1 and graphically in Appendix Figure 3.1. This allows me to operate with the assumption that all teachers with *t* years of experience are the same, but some of them are awarded tenure and others are not. I then use a difference in difference approach to examine the impact that tenure has on teachers. The treated group is the teachers who were awarded tenure and the control group is the other teachers with the same amount of experience that have not been given tenure yet due to their state's longer probationary period.

My empirical specification is as follows

$$V_{itsy} = earn_{itsy} + \sum_{k=1}^{80} k\_years\_post + X_{its} + T_{its} + S_{sy} + Y_{tsy} + \mu_{itsy}$$

The dependent variable (*V*) *is* normalized NAEP score for student *i*, paired with teacher *t*, at school *s*, during year *y*. The two variables of interest are *earn* which is a dummy equal to 1 if the teacher earned tenure this year. We also draw attention to *k_years_post* which is a dummy variable equal to one if the teacher is in their kth year after earning tenure. *X*, *T*, *S*, and *Y* contain characteristics for the student, the teacher, the school, and the year respectively. The basic structure of using a dependent variable at the student level and comparing effects of characteristics at the teacher level is found throughout the literature, specifically Hanushek (1971).

Results

A brief overview of the descriptive statistics that describe the students and teachers in my sample follow. Table 3.1 shows information on the students and Table 3.2 shows means for selected teacher characteristics.

Table 3.1: Teacher Statistics

	(1) Sample	(2) Non-tenured	(3) Tenured
Age	40.52	33.853	45.277
Experience	11.698	4.687	16.701
Percent Female	0.723	0.721	0.724
Percent Asian	0.016	0.025	0.014
Percent Black	0.065	0.083	0.061
Percent Hispanic	0.033	0.055	0.029
Percent Native American	0.015	0.018	0.014
Percent White	0.624	0.687	0.610
Percent with Bachelors	0.998	0.997	0.998
Percent with Masters	0.693	0.507	0.734
Observations	127,460	23,100	104,360

The average teacher in the sample is 40.5 years old and has almost 12 years of experience. When isolating only teachers who have not earned tenure yet, those numbers drop to roughly 34 years old and almost 5 years of experience. For currently tenured teachers, the average age is 45 with nearly 17 years of experience. Women make up nearly three fourths of the sample regardless of subset. Over 60 percent of teachers are white, a number that approaches 70 percent when considering untenured teachers. While the vast majority of all teachers have a

Bachelor’s Degree, a Master’s Degree is much more common among tenured teachers with 73 percent owning one compared to the 50 percent of untenured individuals.

Table 3.2: Student Statistics

	(1) Sample	(2) Non-tenured	(3) Tenured
Age	10.842	10.817	10.848
Percent Female	0.490	0.497	0.489
Percent Asian	0.046	0.048	0.045
Percent Black	0.172	0.206	0.165
Percent Hispanic	0.146	0.175	0.139
Percent Native American	0.024	0.022	0.024
Percent White	0.601	0.537	0.615
Percent IEP	0.137	0.098	0.145
Percent LEP	0.074	0.081	0.073
Percent FRL	0.448	0.484	0.439
Observations	2,166,860	392,720	1,774,140

Table 3.2 shows descriptive statistics for students in the sample as a whole and then those same statistics stratified by whether their teacher has tenure or not. The numbers are largely the same for all three columns. The average age of students is 10.8, since the NAEP data tests 4th graders and 8th graders. So the majority of students in the sample are 9, 10, 13, or 14 years old. Almost half of the students are female, 5 percent are of Asian descent, 20 percent are Black,

roughly 15 percent are Hispanic, and a little over 2 percent are Native American. White students make up 60 percent of the sample. Almost 14 percent of students have an Individualized Education Plan (IEP), which is generally assigned to students with disabilities or learning disorders. 7 percent have Limited English Proficiency (LEP). Almost half of all students come from households where the total income is low enough to qualify that student for a Free or Reduced Lunch (FRL).

Table 3.3: Teacher Statistics Stratified by Tenure

	(1) Year Before	(2) Year After	(3) Diff
Percent Asian	0.023	0.022	0.001
	(0.1504)	(0.1467)	
Percent Black	0.093	0.097	-0.004
	(0.290)	(0.2961)	
Percent Hispanic	0.057	0.053	0.004
	(0.2320)	(0.2245)	
Percent Native American	0.017	0.019	-0.002
	(0.1294)	(0.1354)	
Percent White	0.666	0.674	-0.008
	(0.4716)	(0.4689)	
Percent with Bachelors	0.998	0.999	-0.001
	(0.0448)	(0.0373)	
Percent with Masters	0.569	0.595	-0.026
	(0.4951)	(0.4909)	
Observations	5,900	5,970	

Table 3.3 compares teachers the year before they are eligible for tenure and the year after they earn tenure in order to show that both sides are functionally similar.

Figure 3.1: Normalized Reading Scores by When Teacher Earns Tenure

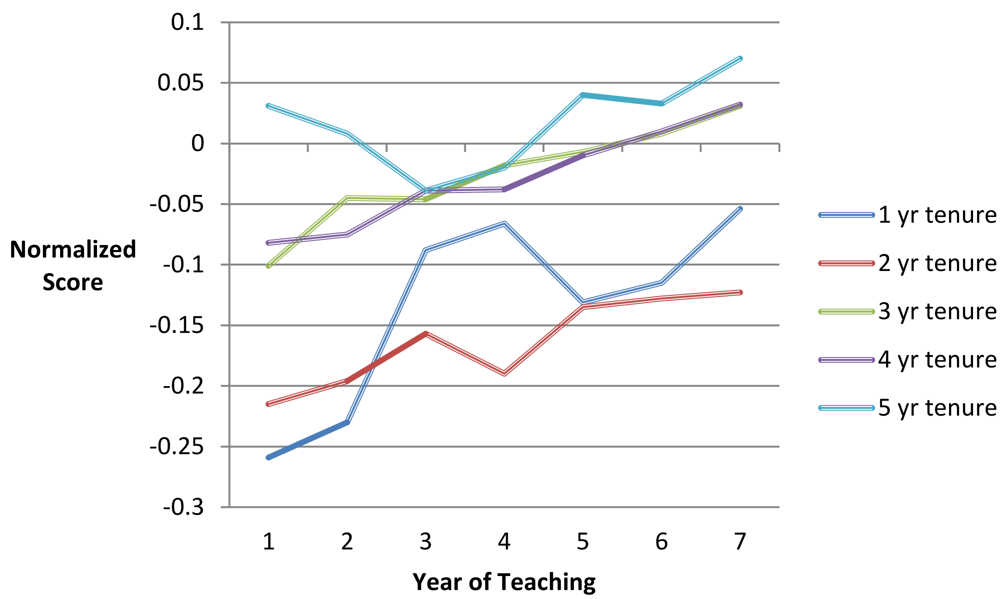
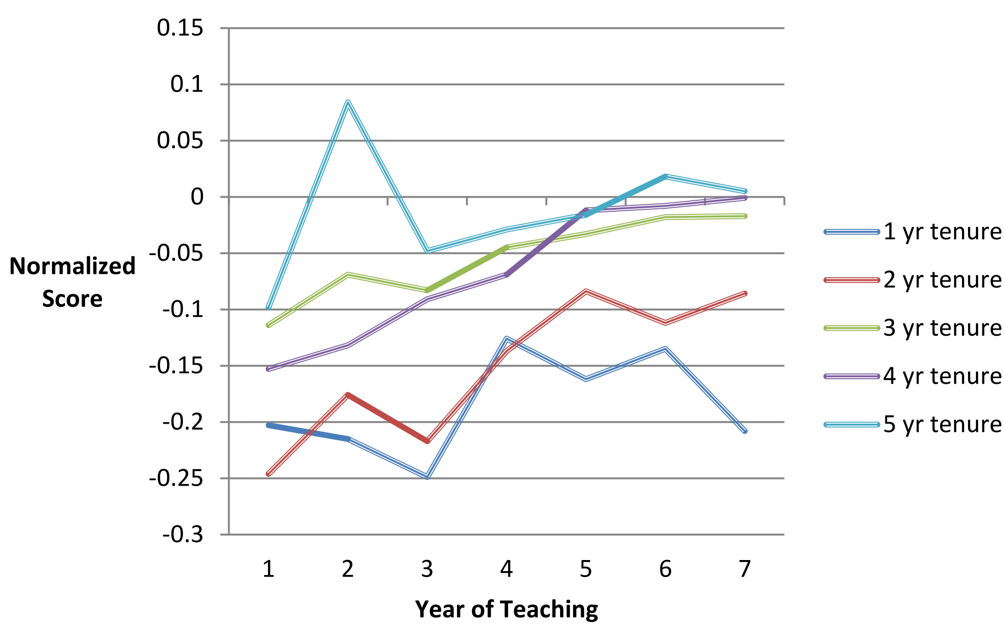
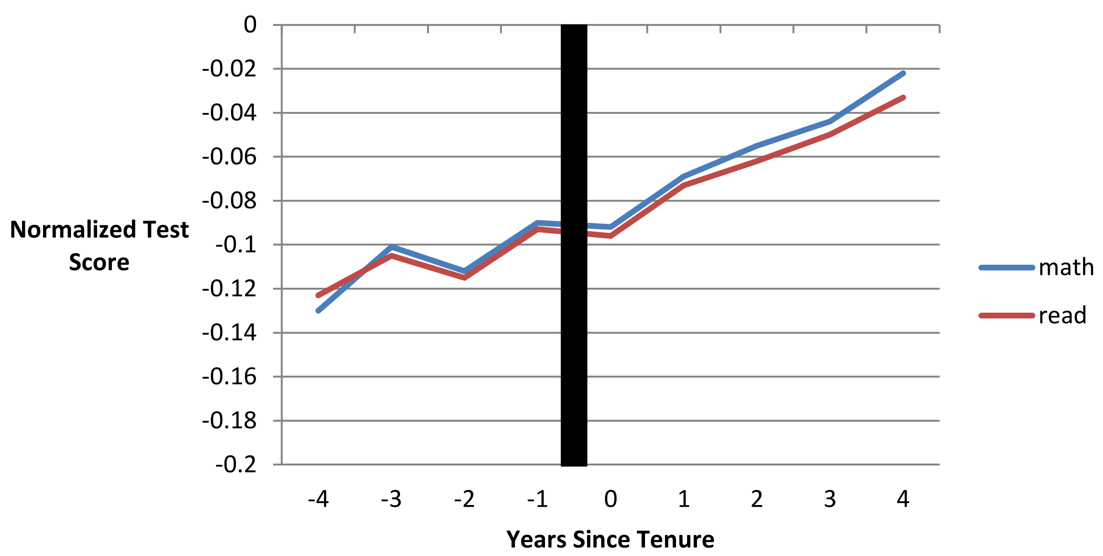


Figure 3.2: Normalized Math Scores by When Teacher Earns Tenure



In both Figure 3.1 and Figure 3.2, we do not observe many reductions in test scores following tenure-ship (The change is negative only for Math scores when the teacher earns tenure after 1 or 2 years). However, the increase in test scores for teachers the year after they earn tenure is lower than for their counterparts not earning tenure that year. These figures also illustrate that a teacher earning tenure before they have gained 3 years of experience seems to stunt their ability as an educator. The scores for teachers who earn tenure after 1 or 2 years is lower than for those who earn tenure after 3 to 5 years. The linear trend for all groups follow the same trajectory, but there is a noticeable gap between the top 3 groups and the bottom 2 groups.

Figure 3.3: Math and Reading Scores in Relation to Tenure



In Figure 3.3 we see that test scores largely follow an increasing trend over time. However, when a teacher earns tenure and moves from -1 year since tenure (year before) to 0 years since tenure (year of), test scores largely hold constant for both math and reading.

Table 3.4: Effect of Tenure

	(1) Reading	(2) Math
Earn	-0.0125***	-0.01089***
	(0.00410)	(0.00409)
1 Year Post	0.0258***	0.0362***
	(0.00419)	(0.00411)
2 Years Post	0.0375***	0.0452***
	(0.00424)	(0.00427)
3 Years Post	0.0366***	0.0495***
	(0.00434)	(0.00432)
4 Years Post	0.0338***	0.0542***
	(0.00472)	(0.00471)
Observations	863,980	878,410
Number of state	50	50
Year FE	YES	YES
State FE	YES	YES

Test scores are negatively impacted the year a teacher earns tenure. The effect is rather small, falling less than 0.02 standard deviations for both Math and Reading. Furthermore, for each year after earning tenure, in both Math scores and Reading scores, test scores rise. This pattern holds through the first 20 years post tenure-ship.

Table 3.5: Effect of Tenure on Reading by Race

	(1) Asian	(2) Black	(3) Hispanic	(4) Native American	(5) White
Earn	0.00772	-0.0455***	0.0230	0.0488	-0.01024***
	(0.0321)	(0.0155)	(0.0206)	(0.0395)	(0.00508)
1 Year Post	0.0220	0.0665***	0.0558***	0.00533	0.0496***
	(0.0368)	(0.0151)	(0.0208)	(0.0388)	(0.00519)
2 Years Post	0.0211	0.0764***	0.0725***	0.0775**	0.0823***
	(0.0351)	(0.0152)	(0.0209)	(0.0333)	(0.00527)
3 Years Post	0.126***	0.0633***	0.0977***	-0.00746	0.0871***
	(0.0351)	(0.0161)	(0.0229)	(0.0366)	(0.00538)
4 Years Post	0.0545	0.0599**	0.0821***	0.152***	0.0808***
	(0.0377)	(0.0175)	(0.0241)	(0.0396)	(0.00585)
Observations	1,030	4,890	2,290	990	45,290
Number of state	50	50	50	50	50
Year FE	YES	YES	YES	YES	YES
State FE	YES	YES	YES	YES	YES

Table 3.6: Effect of Tenure on Math by Race

	(1) Asian	(2) Black	(3) Hispanic	(4) Native American	(5) White
Earn	0.0262	-0.0190	-0.0595***	0.0538	-0.0178***
	(0.0343)	(0.0187)	(0.0223)	(0.0442)	(0.00639)
1 Year Post	0.0309	0.0550***	0.00906	0.0594	0.0703***
	(0.0328)	(0.0175)	(0.0221)	(0.0424)	(0.00630)
2 Years Post	0.117***	0.0846***	0.0582**	-0.00955	0.0786***
	(0.0376)	(0.0183)	(0.0231)	(0.0438)	(0.00652)
3 Years Post	0.165***	0.0371**	0.128***	-0.0271	0.125***
	(0.0374)	(0.0187)	(0.0239)	(0.0411)	(0.00654)
4 Years Post	0.247***	0.0502**	0.0898***	0.114**	0.132***
	(0.0409)	(0.0210)	(0.0269)	(0.0539)	(0.00715)
Observations	950	3,180	1,780	660	30,770
Number of state	50	50	50	50	50
Year FE	YES	YES	YES	YES	YES
State FE	YES	YES	YES	YES	YES

Tables 3.5 and 3.6 show the effect of tenure on reading and math scores respectively while stratifying the teacher by race. The positive influence of further experience following tenure is largely universal among both subjects and all races. However, the statistically significant negative impact of tenure is not. For reading test scores only black and white teachers exhibit a negative change, whereas for math scores it is Hispanic teachers as well as white teachers that show a significant decrease once earning tenure.

Discussion

Tenure does seem to lead to a decrease in test scores. Regardless of subject, students exhibit between a 0.01 to 0.02 standard deviation decline during the year following when their teacher earns tenure. This decline appears to be temporary since test scores to start to rise again for each year taught following earning tenure.

This seems to contradict what we would expect to see once an employee is awarded with “guaranteed job security.” It is surprising that test scores do increase after the slight drop following tenure being awarded. This could speak to the type of people that generally choose to become teachers. These individuals are likely to be more altruistic and less rational in an economic sense. Teachers gain utility not only from financial value, but also from personal, emotional, and “psychic” rewards (Lortie 1975). This could suggest that teachers are not rationally choosing to put forth less effort, but instead are recovering from the previous year (their final year before earning tenure). As Jones 2015 finds, teachers spend more time and energy during this year, and it is likely that these same teachers use their first tenured year as recuperation.

My theoretical model also supports this claim as long as we assume that during the year the teacher earns tenure they are exerting enough effort to cause $f(e) < e$, thereby increasing the likelihood of shirking. However, after the individual has recuperated from the effort spent to earn tenure, they return to the point where $f(e) > e$ thereby earning the psychological payoff associated with teaching.

The test scores for a teacher who just earned tenure is -0.1101 and the linear effect of experience on test scores is 0.0073. So a back of the envelope calculation tells us that earning tenure is associated with lower test scores of the magnitude 0.1174. One of the biggest arguments against tenure for elementary and secondary school teachers is that it provides an incentive to stop quality instruction. While this is true in the short run, the slight reduction in test scores is short lived and is overcome the following year. There are many other discussions that can be had about tenure pros and cons, but the effect on student performance should not be one of them.

Conclusion

One perennial fear with awarding tenure to elementary and secondary teachers is that tenure incentivizes teachers to put forth less effort into their teaching since they will not be punished in the form of job loss. In this paper I considered the effect that earning tenure can have on student test scores by taking advantage of the staggered issuance of tenure among school districts. While it does appear that students are less productive the year after their teacher earns tenure, this effect is short lived and scores return to, and overtake, pre-tenure levels the very next year.

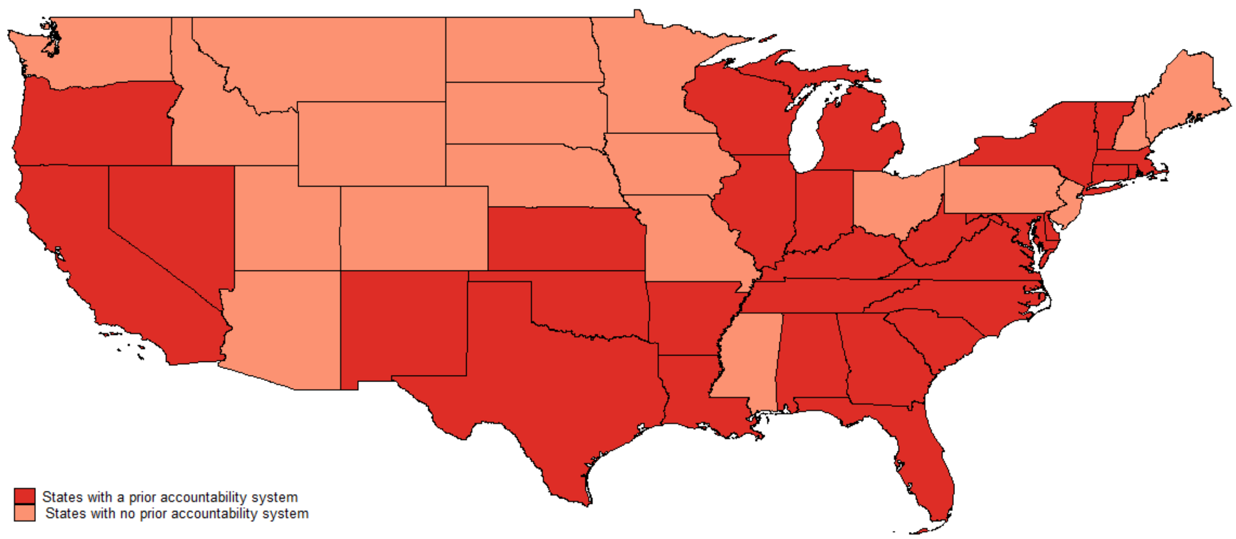
In addition, the effect that tenure has on student achievement is minor compared to other student characteristics such as whether the student has an IEP, or whether or not the student is LEP. Additionally, socio-economic status is incredibly important in determining a student's achievement.

While the result from my analysis is statistically significant, the magnitude of the decrease (slightly greater than 0.01 standard deviations) is a very small practical effect. Dahl and Lochner (2012) find that this is equivalent to roughly $160-$170 reduction in income. There are other potential problems and concerns associated with tenure policy, but the negative impact on students due to poor incentives is minor and should not be used as the sole reason to discontinue tenure policy.

Appendix

The following figure shows the first year that an accountability system was introduced in each state. These years are obtained from Dee and Jacobs (2011).

Appendix Figure 2.1: Prior Accountability System by State



Alaska introduced an accountability system in 1997. Hawaii had no prior accountability system and was treated by NCLB.

Appendix Table 2.1. SASS and TFS robustness

	SASS values	TFS values	Diff
Age	42.5	40.1	2.4
	(10.734)	(12.525)	
Experience	14.0	11.4	2.6
	(9.831)	(11.092)	
Asian	0.026	0.022	0.004
	(0.159)	(0.147)	
Black	0.058	0.059	0.001
	(0.233)	(0.236)	
Hispanic	0.042	0.052	-0.010
	(0.201)	(0.221)	
Native	0.024	0.016	0.008
American	(0.152)	(0.126)	
Female	0.682	0.716	-0.034
	(0.466)	(0.451)	
Bachelors	0.980	0.972	0.008
	(0.141)	(0.165)	
Masters	0.432	0.372	0.060
	(0.495)	(0.483)	

Appendix Table 2.2: Private Teachers Robustness Check

	(1) Private	(2) Public	(3) Diff
Type of reform	Federal	Federal	
Degree earned	0.0433	0.0073	0.036
before	(0.0474)	(0.0243)	(0.0533)
Type of reform	Federal	Federal	
Degree earned	-0.181	0.298***	0.479***
after	(0.157)	(0.0808)	(0.1766)
Type of reform	State	State	
Degree earned	-0.364	0.732***	1.096***
before	(0.243)	(0.102)	(0.2635)
Type of reform	State	State	
Degree earned	0.0614	-0.0485	0.110
after	(0.0843)	(0.0602)	(0.1036)

OLS model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. A state reform is any state-level accountability system enacted before 2002. The federal reform is NCLB. The sample is split into those teachers who earned their Bachelor's degree before the reform occurred and those who earned their degree after. Column 3 shows the difference in mean between private teachers and public teachers for each category. Standard errors are clustered at the state level.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Appendix Table 2.3: Effect by lep quintile

	(1)	(2)	(3)	(4)	(5)	(6)
	First quintile limited English proficiency	Fifth quintile limited English proficiency	Diff	First quintile limited English proficiency	Fifth quintile limited English proficiency	Diff
Type of reform	State	State	State	Federal	Federal	Federal
Reform	0.00482	0.0213	0.0165	0.104***	0.0452	0.0588
	(0.0316)	(0.0556)	(0.0640)	(0.0345)	(0.0426)	(0.0548)
Observations	2,550	1,650		2,960	1,730	
Number of state	50	49		50	49	
Year FE	YES	YES		YES	YES	
State FE	YES	YES		YES	YES	

OLS model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. A state reform is any state-level accountability system enacted before 2002. The federal reform is NCLB. The sample is divided into those teachers who teach at a school with a low number of students that are limited English proficient as opposed to those that teach at schools with a high proportion of students that are limited English proficient. Columns 3 and 6 each show the result of a difference in mean test for the two columns before it. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 2.4: Stratified along district salary

	(1) Low percent free reduced lunch	(2) High percent free reduced lunch	(3) Diff	(4) Low percent free reduced lunch	(5) High percent free reduced lunch	(6) Diff
Type of reform	State	State		Federal	Federal	
Reform	-0.190	0.133	0.323	0.0515	0.249	0.198
	(0.333)	(0.141)	(0.3616)	(0.197)	(0.232)	(0.304)
Observations	670	1,690		1,380	1,420	
Number of state	49	49		50	50	
State FE	YES	YES		YES	YES	
Year FE	YES	YES		YES	YES	

OLS model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. A state reform is any state-level accountability system enacted before 2002. The federal reform is NCLB. To represent district income I look at the proportion of students that are eligible for free and reduced lunch. The sample is divided into teachers at schools with a high proportion eligible for frl and those that teach at a school with low proportion of frl. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 2.5: NCLB Effect Stratified by Race Quintile

	(1) First quintile	(2) Fifth quintile	(3) Diff	Number of state	State FE	Year FE
Asian	0.0458	0.0518	0.006	49	YES	YES
	(0.0400)	(0.0466)	(0.0614)	49	YES	YES
Black	0.230***	0.0324	0.198**	48	YES	YES
	(0.0467)	(0.0453)	(0.0651)	38	YES	YES
Hispanic	0.00134	0.0543	0.0530	49	YES	YES
	(0.0466)	(0.0465)	(0.0658)	42	YES	YES
Native	0.0253	0.0580	0.0327	50	YES	YES
	(0.0310)	(0.0435)	(0.0534)	49	YES	YES
White	0.0939**	0.0394	0.0545	45	YES	YES
	(0.0426)	(0.0461)	(0.0628)	48	YES	YES

OLS model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. The federal reform is NCLB. I divide the sample into quintiles based on the proportion of students that are the given races. I then look at the results for the first and fifth quintiles. Column 3 shows the results of a difference in mean test for each race. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 2.6: Prior System Effect Stratified by Race Quintile

	(1) First quintile	(2) Fifth quintile	(3) Diff	Number of state	State FE	Year FE
Asian	-0.0045	0.0029	0.0074	49	YES	YES
	(0.0345)	(0.0430)	(0.0551)	49	YES	YES
Black	0.0953*	0.0154	0.0799	50	YES	YES
	(0.0500)	(0.0434)	(0.0662)	39	YES	YES
Hispanic	-0.0869*	-0.0056	0.0813	49	YES	YES
	(0.0496)	(0.0469)	(0.0683)	40	YES	YES
Native	-0.0143	3.11e-05	0.0143	50	YES	YES
	(0.0265)	(0.0454)	(0.0526)	49	YES	YES
White	0.0074	-0.0089	0.0163	46	YES	YES
	(0.0426)	(0.0481)	(0.0643)	48	YES	YES

OLS model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. The federal reform is NCLB. I divide the sample into quintiles based on the proportion of students that are the given races. I then look at the results for the first and fifth quintiles. Column 3 shows the results of a difference in mean test for each race. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 2.7: Effect of NCLB stratified by Adequate Yearly Progress (AYP)

	(1) Made AYP	(2) Missed AYP	(3) Diff	(4) Made AYP	(5) Missed AYP twice	(6) Diff
Reform	-0.00648	0.339	0.345	0.683*	-0.0575	0.741
	(0.101)	(0.324)	(0.339)	(0.376)	(0.372)	(0.529)
Observations	1,640	1,350		860	490	
Number of state	49	49		46	41	
State FE	YES	YES		YES	YES	
Year FE	YES	YES		YES	YES	

OLS model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. A state reform is any state-level accountability system enacted before 2002. The federal reform is NCLB. A school is determined as having made Adequate Yearly Progress (AYP) if enough student subgroups are considered proficient according to the guidelines set by that state. If a school fails to reach AYP two consecutive years then sanctions are placed on that school. Those schools make up the sample for column 5. Diff shows the results of a difference in mean test for the two columns immediately before. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 2.8: Effect of interaction terms on staying

	(1) All	(2) Pre	(3) Post	(4) All	(5) Pre	(6) Post
Type of reform	State	State	State	Federal	Federal	Federal
Reform*age	0.000908	0.00332*	-0.0120***	0.00285**	0.00321**	-0.00953
	(0.00140)	(0.00182)	(0.00413)	(0.00144)	(0.00160)	(0.00702)
Reform*exp	8.39e-05	-0.00223	0.00893*	-0.00261	-0.00173	-0.0220
	(0.00162)	(0.00214)	(0.00465)	(0.00172)	(0.00185)	(0.0184)
Observations	7,660	4,040	3,630	7,320	5,670	1,650
Number of state	50	50	50	50	50	50
State FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

OLS model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. A state reform is any state-level accountability system enacted before 2002. The federal reform is NCLB. Pre contains all teachers who earned their Bachelor's degree before the listed reform took effect. Post inludes those who earned their degree after the reform occurred. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 2.9: Effect of Accountability System on attitude toward teaching stratified by year of bachelor's degree

	(1) All	(2) Post reforem	(3) Pre reform	(4) Diff	(5) All	(6) Post reform	(7) Pre reform	(8) Diff
Type of reform	State	State	State		Federal	Federal	Federal	
Reform	0.0058	0.1709**	0.1963*	0.0254	0.0331	0.129	0.0339	0.0951
	(0.0268)	(0.0841)	(0.1037)	(0.1335)	(0.0273)	(0.137)	(0.0349)	(0.1414)
Observations	7,660	2,540	5,120		7,130	1,590	5,540	
Number of state	50	26	50		50	50	50	
State FE	YES	YES	YES		YES	YES	YES	
Year FE	YES	YES	YES		YES	YES	YES	

OLS model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. A state reform is any state-level accountability system enacted before 2002. The federal reform is NCLB. Pre contains all teachers who earned their Bachelor's degree before the listed reform took effect. Post includes those who earned their degree after the reform occurred. Diff shows the results of a difference in mean test for the two columns immediately before. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 2.10: Effect on NCLB on staying, probit model

	(1) All	(2) Pre	(3) Post	(4) Diff
Reform	0.0935	0.0118	0.987***	0.9752**
	(0.0660)	(0.0768)	(0.372)	(0.3798)
Observations	8,100	6,450	1,610	

Probit model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. The federal reform is NCLB. Pre contains all teachers who earned their Bachelor's degree before the listed reform took effect. Post includes those who earned their degree after the reform occurred. Diff shows the results of a difference in mean test for the two columns immediately before. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 2.11: Effect of NCLB on staying stratified by age, probit model

	(1) Less than 25	(2) 25-35	(3) 35+
Reform	0.261	0.174	0.0165
	(0.197)	(0.127)	(0.0896)
Observations	1,340	2,390	4,320

Probit model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. The federal reform is NCLB. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 2.12: Effect of NCLB on staying stratified by degree, probit model

	(1) Bachelors Only	(2) Masters	(3) Diff
reform	0.208**	-0.0465	0.2545*
	(0.0896)	(0.102)	(0.1358)
Observations	4,850	3,140	

Probit model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. The federal reform is NCLB. Diff shows the result of a difference in mean test for the two columns prior. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 2.13: Effect of NCLB on staying stratified by experience, probit model

	(1) Less than 3	(2) 3-15	(3) 15+
Reform	0.157	0.0314	0.00359
	(0.120)	(0.122)	(0.120)
Observations	3,190	2,490	2,370

Probit model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. The federal reform is NCLB. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 2.14: Effect of prior system on staying, probit model

	(1) All	(2) Pre	(3) Post	(4) Diff
Reform	-0.0920	0.500*	0.0360	0.464
	(0.0644)	(0.266)	(0.223)	(0.3471)
Observations	7,660	5,360	5,200	

Probit model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. A state reform is any state-wide accountability system that took effect prior to 2002. Pre contains all teachers who earned their Bachelor's degree before the listed reform took effect. Post includes those who earned their degree after the reform occurred. Diff shows the results of a difference in mean test for the two columns immediately before. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 2.15: Effect of prior system on staying stratified by age, probit model

	< 25	25-35	35+
Prior	0.0760	-0.0724	-0.0929
	(0.231)	(0.116)	(0.0858)
Observations	890	2,450	4,290

Probit model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. A state reform is any state-wide accountability system that took effect prior to 2002. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 2.16: Effect of prior system on staying stratified by experience, probit model

	< 3	3-15	15+
Reform	0.0282	-0.207*	-0.0994
	(0.120)	(0.107)	(0.121)
Observations	2,750	2,780	2,130

Probit model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. A state reform is any state-wide accountability system that took effect prior to 2002. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 2.17: Effect of prior system on staying stratified by degree, probit model

	(1) Bachelors Only	(2) Masters	Diff
Reform	-0.0512	-0.192*	0.1408
	(0.0846)	(0.104)	(0.1341)
Observations	4,750	2,810	

Probit model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. A state reform is any state-wide accountability system that took effect prior to 2002. Diff shows the results of a difference in mean test for the previous two columns. Standard errors are clustered at the state level.

*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 2.18: How Teacher Makeup Changed

	(1) Teach math	(2) Teach reading	(3) Failed test
Type of reform	Federal	Federal	Federal
Mean	0.053	0.109	0.022
Reform	-0.12**	0.0012	0.015***
	(0.0059)	(0.0084)	(0.0044)
Observations	10,260	10,260	4,240
Number of state	50	50	50
State FE	YES	YES	YES
Year FE	YES	YES	YES

OLS model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. The federal reform is NCLB. Teach math is a dummy variable equal to one if the teacher majored in math education for their Bachelor's degree. Teach reading is a dummy variable equal to one if the teacher majored in reading education for their Bachelor's degree. Failed test is a dummy variable equal to 1 if the teacher failed at least one Praxis test. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 2.19: Effect of Accountability System on Staying by Sex

	(1) '93-'99 Females	(2) '93-'99 Males	(3) Diff
Type of reform	State	State	State
Reform	-0.0238	-0.00456	0.0192
	(0.0230)	(0.0379)	(0.0443)
Observations	5,940	2,390	
Number of state	50	50	
Year FE	YES	YES	
State FE	YES	YES	

OLS with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. The results are stratified by gender. Column 3 shows a difference in mean test for the 2 columns it follows. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 2.20: Effect of Accountability System on Staying by Education Level			
	(1) ’93-‘99 Bachelors	(2) ’93-‘99 Masters	(3) Diff
Type of reform	State	State	
Reform	-0.00498	-0.0611*	0.0561
	(0.0248)	(0.0358)	(0.0436)
Observations	4,680	2,780	
Number of state	50	50	
Year FE	YES	YES	
State FE	YES	YES	

OLS with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher was still teaching in the Teacher Follow-up Survey. The results are stratified by education level. I dropped the teachers with no Bachelor’s. Column 3 shows the results of a difference in mean test for the 2 columns it follows. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 2.21: Effect of Accountability System on Staying by Age			
	(1) Less than 25	(2) 25-35	(3) 35-60
Type of Reform	State	State	State
Reform	0.0107	-0.0122	-0.0239
	(0.0532)	(0.0370)	(0.0277)
Observations	910	2,430	4,100
Number of state	50	50	50
Year FE	YES	YES	YES
State FE	YES	YES	YES

OLS model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. A state reform is any state-level accountability system enacted before 2002. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 2.22: Effect of Accountability System on Staying by Experience

	(1) Prior 0-3 years experience	(2) Prior 3-15 years experience	(3) Prior 15-30 years experience
Type of Reform	State	State	State
Reform	0.0143	-0.0675	-0.0572
	(0.0305)	(0.0355)	(0.0421)
Observations	2,720	2,750	1,820
Number of state	50	50	50
Year FE	YES	YES	YES
State FE	YES	YES	YES

OLS model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. A state reform is any state-level accountability system enacted before 2002. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 2.23: Effect of Accountability System on Staying by Year of Bachelor's Degree

	(1) Pre reform	(2) Post reform	(3) Diff
Type of reform	State	State	
Reform	0.127*	0.124	0.003
	(0.0731)	(0.0820)	(0.1100)
Observations	6,960	5,100	
Number of state	50	26	
State FE	YES	YES	
Year FE	YES	YES	

OLS model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. A state reform is any state-level accountability system enacted before 2002. Column 3 shows the results of a difference mean test for the 2 column it follows. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 2.24: Effect of Accountability System on Staying by View on Continuing Teaching

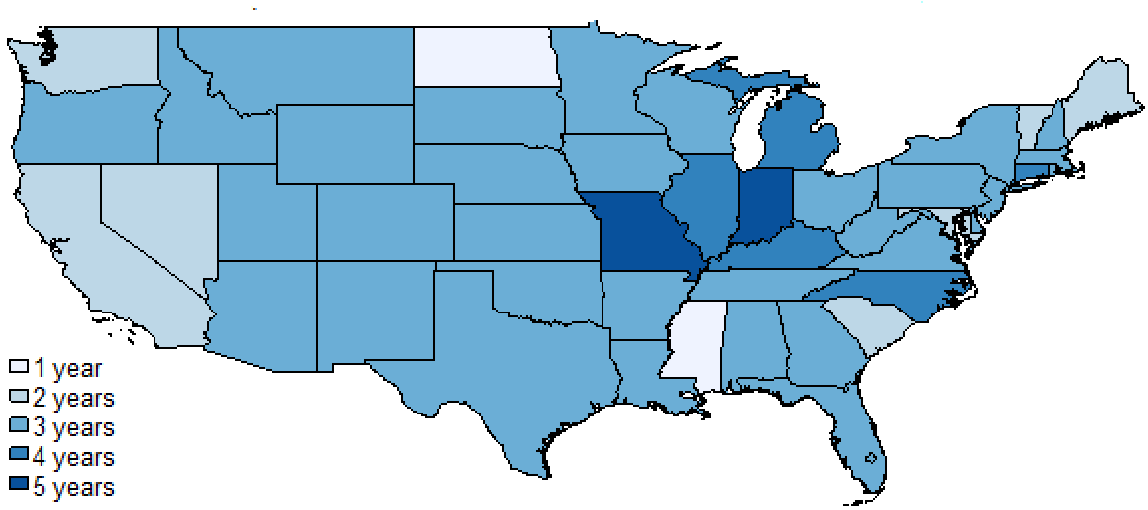
	(1) '93-'99 would teach again	(2) '93-'99 would not teach again	(3) Diff
Type of reform	State	State	State
Reform	-0.00129	-0.156***	0.155***
	(0.0218)	(0.0540)	(0.0582)
Observations	1,320	6,350	
Number of state	50	50	
Year FE	YES	YES	
State FE	YES	YES	

OLS model with stay as the dependent variable. Stay is a dummy variable equal to 1 if the teacher is still teaching in the Teacher Follow-up Survey. A state reform is any state-level accountability system enacted before 2002. Column 3 shows the results for a difference in mean test for the 2 columns it follows. A teacher's willingness to teach again is from the survey question "Would you be a teacher again?" Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 3.1: Experience Required for Tenure by State

1 year	2 years	3 years		4 years	5 years
Hawaii	California	Alabama	New Jersey	Connecticut	Indiana
Mississippi	Maine	Alaska	New Mexico	Illinois	Missouri
North Dakota	Maryland	Arizona	New York	Kentucky	
	Nevada	Arkansas	Ohio	Michigan	
	South Carolina	Colorado	Oklahoma	North Carolina	
	Vermont	Delaware	Oregon		
	Washington	Florida	Pennsylvania		
		Georgia	Rhode Island		
		Idaho	South Dakota		
		Iowa	Tennessee		
		Kansas	Texas		
		Louisiana	Utah		
		Massachusetts	Virginia		
		Minnesota	West Virginia		
		Montana	Wisconsin		
		Nebraska	Wyoming		
		New Hampshire			

Appendix Figure 3.1: Experience Required for Tenure by State



References

Angeletos, George-Marios, Christian Hellwig, and Alessandro Pavan. "Signaling in a global game: Coordination and policy traps." *Journal of Political economy* 114.3 (2006): 452-484.

Baliga, Sandeep, and Tomas Sjöström. "The strategy of manipulating conflict." *American Economic Review* 102.6 (2012): 2897-2922.

Ballou, Dale, and Matthew G. Springer. "Achievement trade-offs and no child left behind." *Manuscript. Peabody College of Vanderbilt University. En: www. caldercenter. org* (2008).

Barksdale-Ladd, Mary Alice, and Karen F. Thomas. "What's at stake in high-stakes testing: Teachers and parents speak out." *Journal of Teacher Education* 51.5 (2000): 384-397.

Barrett, B. "LAUSD's dance of the lemons." *LA Weekly* 11 (2012).
Brill, Steven. "The rubber room: The battle over New York City's worst teachers." *The New Yorker* 29 (2009).

Bueno De Mesquita, Ethan. "Regime change and revolutionary entrepreneurs." *American Political Science Review* 104.3 (2010): 446-466.

Bueno de Mesquita, Ethan. "Regime Change and Equilibrium Multiplicity." *Unpublished manuscript, Harris School, University of Chicago* (2014).

Cavanaugh, S. "Survey: N.C. teachers say high-stakes tests dominate classes." *Education Week* (2012).

Christensen, Gayle S., et al. "Private School Participants in Federal Programs under the" No Child Left Behind Act" and the" Individuals with Disabilities Education Act": Private School and Public School District Perspectives." *US Department of Education* (2007).

Christie, K., and K. Zinth. "Ensuring successful student transitions from the middle grades to high school." *Denver, CO: Education Commission of the States* (2008).

Dahl, Gordon B., and Lance Lochner. "The impact of family income on child achievement: Evidence from the earned income tax credit." *American Economic Review* 102.5 (2012): 1927-56.

Dee, Thomas S., and Brian Jacob. "The impact of No Child Left Behind on student achievement." *Journal of Policy Analysis and management* 30.3 (2011): 418-446.

Dee, Thomas S., Brian Jacob, and Nathaniel L. Schwartz. "The effects of NCLB on school resources and practices." *Educational Evaluation and Policy Analysis* 35.2 (2013): 252-279.

DeGroot, Morris. "H., 1970, Optimal Statistical Decisions."

Deniston, Richard D., and Kevin W. Gerrity. "Elementary school teachers' perceptions of No Child Left Behind and its effect on morale." *scholarlypartnershipsedu* 5.2 (2011): 4.

Doherty, Kathryn M, et al. "Connect the Dots: Using Evaluations of Teacher Effectiveness to Inform Policy and Practice." *State of the States 2013,* National Council on Teacher Quality, Oct. 2013,www.nctq.org/dmsView/State_of_the_States_2013_Using_Teacher_Evaluations_NCTQ_Report

Edmond, Chris. "Information manipulation, coordination, and regime change." *Review of Economic Studies* 80.4 (2013): 1422-1458.

Figlio, David N., and Cecilia Elena Rouse. "Do accountability and voucher threats improve low-performing schools?." *Journal of Public Economics* 90.1 (2006): 239-255.

Finnigan, Kara S., and Betheny Gross. "Do accountability policy sanctions influence teacher motivation? Lessons from Chicago's low-performing schools." *American Educational Research Journal* 44.3 (2007): 594-630.

Gerson, Jack. "NCLB: Bad for teachers, bad for kids." *United Teacher* 37.2 (2007): 14.

Grissom, Jason A., Sean Nicholson-Crotty, and James R. Harrington. "Estimating the effects of No Child Left Behind on teachers' work environments and job attitudes." *Educational Evaluation and Policy Analysis* 36.4 (2014): 417-436.

Han, Eunice S. "The Myth of Unions' Overprotection of Bad Teachers: Evidence from the District-Teacher Matched Panel Data on Teacher Turnover." *Job Market Paper, Wellesley College and NBER* (2015).

Hanushek, Eric. "Teacher characteristics and gains in student achievement: Estimation using micro data." *The American Economic Review* 61.2 (1971): 280-288.

Hanushek, Eric A., Steven G. Rivkin, and Jeffrey C. Schiman. "Dynamic effects of teacher turnover on the quality of instruction." *Economics of Education Review* 55 (2016): 132- .

Hardin, Russell. *One for all: The logic of group conflict*. Princeton University Press, 1997.

Hefling, K. "Obama administration's challenge to schools: Embrace digital textbooks within 5 years." *Retrieved Dec. 4th* (2012).

Hoffman, James V., Lori Czop Assaf, and Scott G. Paris. "High-stakes testing in reading: Today in Texas, tomorrow?." *The reading teacher* 54.5 (2001): 482-492.

Hyatt, Henry R., and Erika McEntarfer. "Job-to-job Flows and the Business Cycle." (2012). 148.

Jacob, Brian A. "Accountability, incentives and behavior: The impact of high-stakes testing in the Chicago Public Schools." *Journal of public Economics* 89.5 (2005): 761-796.

Jones, Michael D. "How do teachers respond to tenure?." *IZA Journal of Labor Economics* 4.1 (2015): 8.

Jovanovic, Boyan. "Job matching and the theory of turnover." *Journal of political economy* 87.5, Part 1 (1979): 972-990.

Kane, Thomas J., and Douglas O. Staiger. *Estimating teacher impacts on student achievement: An experimental evaluation*. No. w14607. National Bureau of Economic Research, 2008.

Kersten, Thomas A. "Teacher Tenure: Illinois School Board Presidents' Perspectives and Suggestions for Improvement." *Planning and Changing* 37.3/4 (2006): 234.

Lazear, Edward P. "Teacher incentives." *Swedish Economic Policy Review* 10.2 (2003): 179-214.

Loeb, Susanna, and Jesse Cunha. "Have Assessment-Based Accountability Reforms Influenced the Career Decisions of Teachers and Principals?." (2007).

Lortie, Dan. "Schoolteacher: A sociological analysis." *Chicago: University of* (1975).

Luna, Catherine, and Cara Livingstone Turner. "The impact of the MCAS: Teachers talk about high-stakes testing." *The English Journal* 91.1 (2001): 79-87.

Manna, P. "Education: Federal government programs and issues." (2011).

McGuinn, Patrick. "Ringing the bell for K-12 teacher tenure reform." *Washington, DC: Center for American Progress*(2010).

Medina, Jennifer. "Last day of "rubber rooms" for teachers." *The New York Times* (2010): 331-59.

Mislevy, Robert J., and R. Darrell Bock. "Biweight estimates of latent ability." *Educational and psychological measurement*42.3 (1982): 725-737.

Morris, Stephen, and Hyun Song Shin. "Global games: Theory and applications." (2001).

Muraki, E., and R. D. Bock. "PARSCALE: Parameter scaling of rating data." *Computer program]. Chicago: Scientific Software, Inc* (1991).

Murnane, Richard J., and Barbara R. Phillips. "What do effective teachers of inner-city children have in common?." *Social Science Research* 10.1 (1981): 83-100.

Murnane, Richard J., and John P. Papay. "Teachers' views on No Child Left Behind: Support for the principles, concerns about the practices." *The Journal of Economic Perspectives* 24.3 (2010): 151-166.

Neal, Derek. "Industry-specific human capital: Evidence from displaced workers." *Journal of labor Economics* 13.4 (1995): 653-677.

Parent, Daniel. "Industry-specific capital and the wage profile: Evidence from the national longitudinal survey of youth and the panel study of income dynamics." *Journal of Labor Economics* 18.2 (2000): 306-323.

Reback, Randall, Jonah Rockoff, and Heather L. Schwartz. "Under pressure: Job security, resource allocation, and productivity in schools under No Child Left Behind." *American Economic Journal: Economic Policy* 6.3 (2014): 207-241.

Reininger, Michelle. "Hometown disadvantage? It depends on where you're from: Teachers' location preferences and the implications for staffing schools." *Educational Evaluation and Policy Analysis* 34.2 (2012): 127-145.

Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. "Teachers, schools, and academic achievement." *Econometrica* 73.2 (2005): 417-458.

Rockoff, Jonah, and Lesley J. Turner. "Short-run impacts of accountability on school quality." *American Economic Journal: Economic Policy* 2.4 (2010): 119-147.

Ronfeldt, Matthew, Susanna Loeb, and James Wyckoff. "How teacher turnover harms student achievement." *American Educational Research Journal* 50.1 (2013): 4-36.

Rothstein, Jesse. "Revisiting the impacts of teachers." *UC-Berkeley Working Paper* (2014).

Rouse, Cecilia Elena, et al. *Feeling the Florida heat? How low-performing schools respond to voucher and accountability pressure*. No. w13681. National Bureau of Economic Research, 2007.

Sass, Tim R., et al. "Value added of teachers in high-poverty schools and lower poverty schools." *Journal of Urban Economics* 72.2-3 (2012): 104-122.

Schelling, Thomas C. "The strategy of conflict." *Cambridge, Mass* (1960).

Shadmehr, Mehdi, and Dan Bernhardt. "Collective action with uncertain payoffs: coordination, public signals, and punishment dilemmas." *American Political Science Review*105.4 (2011): 829-851.

Shadmehr, Mehdi, and Dan Bernhardt. *Coordination Games with Information Aggregation*. Mimeo—University of Miami, 2012.

Stephey, M. J. "A brief history of tenure." *Time Magazine* 17 (2008).

Sun, Min, Andrew Saultz, and Yincheng Ye. "Federal policy and the teacher labor market: Exploring the effects of NCLB on teacher turnover." *annual meeting of The Association for Education Finance and Policy, San Antonio, TX*. 2014.

Sunderman, Gail L., et al. "Listening to teachers: Classroom realities and no child left behind." (2004).

Taylor, Eric S., and John H. Tyler. "The effect of evaluation on teacher performance." *American Economic Review* 102.7 (2012): 3628-51.

Teacher Retirement System of Texas (History of Benefit Changes). *www.trs.texas.gov*, *2016* *www.trs.texas.gov/TRS%20Documents/history_benefit_changes.pdf*,

Topel, Robert H., and Michael P. Ward. "Job mobility and the careers of young men." *The Quarterly Journal of Economics* 107.2 (1992): 439-479.

Toppo, Greg. "How Bush education law has changed our schools." *USA Today* 8 (2007).

United States Department of Education. "Race to the Top program: Executive summary." (2009).

Weisberg, Daniel, et al. "The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness." *New Teacher Project* (2009).