

**To Teach or Not to Teach:  
Assessing the Impact of High-Stakes Accountability Policies on College  
Graduates' Decisions to Enter Teaching**

May 2004

Felicia Kong Estrada  
Advisor: Susanna Loeb  
Stanford University, Stanford, CA 94309 USA  
E-mail: [festrada@stanford.edu](mailto:festrada@stanford.edu)

**Abstract**

Many states throughout the 1990s implemented the use of a number of large-scale state assessments as a way to measure student achievement and make teachers and schools directly accountable for student improvement. The research on the adoption of assessment-based accountability on teachers' career decisions is largely unformed. This paper examines the effects of accountability on a group of individuals entering the teaching profession – recent college graduates. The two main questions evaluate whether states that implemented strong accountability policies experience a differential change in the characteristics and work attitudes of new teachers. The study follows the career decisions of college graduates from the High School and Beyond Study of 1980 and the National Education Longitudinal Survey of 1988. Overall, the study finds no consistent, statistical evidence that strong accountability has a detrimental effect on the proportion of high quality, male, or non-white college graduates who decide to become teachers.

Keywords: Accountability, Career Choice, Teaching (Occupation)

## **Acknowledgements**

I would like to thank Professor Susanna Loeb for her invaluable guidance in developing the idea for this paper and her continual interest and support. Also, I would like to thank Professor Martin Carnoy and members of the Licensing Office at the United States Department of Education National Center of Education Statistics for their help in accessing my data. I also thank my friends and family for their support and motivation throughout the past year. Finally, I thank members of the Stanford University Undergraduate Research Programs Office for their interest in my topic and the generous support of my research grant. All errors and opinions expressed herein are my own

## Table of Contents

<b>1. Introduction .....</b>	<b>3</b>
<b>2. Review of the Literature.....</b>	<b>5</b>
2.1 Previous Research on Accountability and Teacher Characteristics.....	5
2.1.1 Effects on Teacher Quality in Low-Performing Schools.....	5
2.1.2. Effects on Teacher Quality in Tested Grades.....	6
2.1.3 National Effects of Accountability on Teacher Quality.....	7
2.2 Models on the Decision to Teach.....	8
<b>3.Data and Methodology.....</b>	<b>10</b>
3.1 Limitations of the Analysis .....	10
3.2 Data.....	10
3.2.1 Inconsistencies With the Data .....	14
3.3 Methodology.....	15
3.3.1 Baseline Effects.....	17
3.3.2 Differential Effects.....	18
3.3.3 Descriptive Analysis.....	19
<b>4. Results.....</b>	<b>20</b>
4.1 Basic Teacher Entry Trends.....	20
4.2 Differential Effects of Strong Accountability.....	21
4.2.1 Effects on Teacher Quality.....	21
4.2.2 Effects on Teacher Gender.....	22
4.2.3 Effects on Teacher Racial and Ethnic Composition.....	22
4.2.4 Effects on Likely Teachers.....	23
4.3 Descriptive Analysis.....	24
4.3.1 Effects on Teacher Job Satisfaction.....	24
<b>5. Conclusion.....</b>	<b>26</b>
<b>6. Reference List.....</b>	<b>29</b>
<b>7. Data Products.....</b>	<b>31</b>
<b>8. Figures and Tables.....</b>	<b>33</b>

## 1. Introduction

Tests have always been used as tools for student assessment. Historically, teachers, parents and students have used tests in the classroom as purely evaluative tools to gauge student understanding of material and identify areas in which the student needs help. In the 1990s, the role of testing changed dramatically through the widespread adoption of assessment-based accountability reforms, when a number of states mandated the use of standardized tests as a way to make school districts, schools, and teachers more accountable for student learning. As a result, schools, teachers, and students received rewards for good student performance, but faced penalties for poor student performance. In some cases, part of teachers' pay was even linked to the gains their students made on annual assessments (Carnoy and Loeb 2003; Clotfelter, Ladd, Vigdor, and Diaz 2003). Because very few studies have looked at accountability reform's effects on teacher career decisions, it is not clear how teachers have responded to increased scrutiny under these recent reforms. Since no studies have attempted to look at accountability's effects on people entering the teaching profession, it is even less clear on how *potential* teachers are reacting to the changes.

Each year almost half of new teachers are recent college graduates (Digest of Education Statistics 2002). Considering that accountability policies may place additional demands on teachers to ensure students meet achievement standards, we are uncertain if teaching will remain as attractive a career option to young people weighing the decision to teach. College graduates who perceive accountability policies negatively might be less inclined to teach for a number of reasons. For example, some individuals may choose not to teach because they dislike the idea of being evaluated by narrow measures like standardized tests. Others may be discouraged from

teaching because they believe the strong focus on student outcomes will force them to teach to a test, and thus teaching will become uninteresting or less challenging. On the other hand, college students may not mind accountability reforms and become teachers regardless of the policy change. Or as an extreme hypothesis, the challenge of meeting high performance standards under strong accountability may serve as a tool to bring more respect to the craft of teaching so more people will want to enter teaching.

More important than changes in the number of college graduates who become teachers due to accountability are changes in types and quality of individuals who become teachers. This paper contributes to the literature on accountability reform by being the first to study how the characteristics and attitudes of new teachers have changed in states that implemented strong accountability policies than states that did not adopt reforms. The two main questions ask:

- ❖ Have states that adopted stronger accountability observed a differential change in the composition of new teachers with respect to high-quality candidates, racial and ethnic groups, gender, and other groups of likely teachers?
- ❖ Do teachers in strong accountability states feel as satisfied in their work lives in the areas of pay and benefits, job importance and challenge, and job autonomy as teachers in states with weak or no accountability systems?

## 2. Review of the Literature

### 2.1 Previous Research on Accountability and Teacher Characteristics

Three studies have attempted to look at changes in teachers' exit, transfer, quit rates and other career decisions due to stronger accountability policies. While I expect that accountability affects the decision to teach differently than it does the decision to leave, I review past studies' results to better grasp how teachers characteristics have responded to changes in accountability.

#### 2.1.1 *Effects on Teacher Quality in Low-Performing Schools*

Clotfelter, Ladd, Vigdor, and Diaz (2003) study the case of North Carolina's ABCs program (which stands for the principles of accountability, basic skills, and local control) on teacher retention in low-performing schools. The ABCs program began in 1996 and was characterized by: annual testing in grades K-8; required achievement of school-wide student test score gains and performance standards; and bonus pay for teachers at schools whose students met the gain and performance criteria. The most relevant question to my research Clotfelter, Ladd, Vigdor, and Diaz (2003) address is how the accountability program influenced changes in teacher characteristics among the state's low-performing schools.

The authors found that the ABCs program negatively affected the quality of teachers at low-performing schools. They choose to measure teacher quality by the proportion of novice teachers and selectivity of teachers' undergraduate institutions. The authors consider novice teachers to be lower-quality since first year teachers are less likely to raise student achievement than experience teachers. They chose to focus on decreased teacher college selectivity at low-performing schools as an indicator of less average academic ability, and by extension, an

indicator of less effective teaching. They recognize that while college selectivity is an imprecise measure of academic ability, it serves as an average indicator of teacher quality.

The results of Clotfelter, Ladd, Vigdor, and Diaz (2003) indicate that throughout the periods before and after the ABCs program, low-performing schools had a higher proportion of low-quality teachers, as indicated by levels of novice teachers and levels of teachers from uncompetitive undergraduate institutions. In spite of that, their results of accountability's relative effects in the post-accountability were mixed. No adverse effect was detected when evaluating whether teacher quality decreased at low-performing schools in comparison to higher-performing schools. On the other hand, when considering the quality trends among only low-performing schools over time, there was some evidence to indicate more novice teachers and teachers from uncompetitive colleges were employed after the policy change. Overall, their results inclined to show damaging effects of accountability to low-performing schools, but in regards to lowered teacher quality, results were statistically insignificant.

### *2.1.2. Effects on Teacher Quality in Tested Grades*

In 1998 the New York State Education Department implemented a revised assessment system for fourth and eighth graders in English languages arts and math. The associated consequence of this reform was that school performance was publicly reported. Boyd, Lankford, Loeb and Wycoff (2003) studied the effects of this accountability policy change on New York fourth grade teachers in comparison to teachers of other elementary grades one through six. The authors used an across-grade comparison to observe the relative effect accountability might have in who schools select to teach and what type of teachers leave a grade level that is associated with a major state-wide assessment. The study's results do not indicate that teacher quality decreased among new teachers to the fourth grade. This may be due to schools purposely

allocating more experienced, highly skilled teachers to the tested grade. Also, teachers who were new to teaching were more likely to have attended highly competitive undergraduate institutions and were less likely to quit the teaching in low-performing schools *after* accountability was implemented. On the whole, Boyd, Lankford, Loeb and Wycoff (2003) did not find evidence of damaging effects of accountability on teachers with high-quality characteristics.

### *2.1.3 National Effects of Accountability on Teacher Quality*

Loeb and Estrada (2003) conducted the only national study on accountability's effects on teachers' career decisions. Part of their analysis focused on how the characteristics of teachers have changed between 1993-94 and 2000-01, before and after many states adopted assessment-based accountability systems. If we concentrate on the observed effects of teacher quality, as defined by college competitiveness, states with stronger accountability had an increase in the proportion of teachers who attended more competitive colleges, but the result is not statistically significant. This beneficial result is surprising since the general trend was there was a decrease in the teachers' college competitiveness. When they restrict their sample to new teachers with 2 or fewer years of experience, they find no evidence that stronger accountability changes the expectation a teacher comes from a more or less competitive college.

When considering the limited body of research focused on changes in teacher quality, no single study precisely estimates accountability's. The suggested harmful or beneficial effects on the competitiveness of teachers' undergraduate institutions and the increase of potentially less effective novice teachers tend not to be large either. The state-specific studies enjoy an advantage of having detailed data that enables them to consider very specific questions on teacher labor. On the other hand, the national study allows us to observe the relative impact of stronger versus weaker accountability systems. Although, throughout the studies, I discussion on

the relative impact of accountability on the racial, ethnic, and gender-balanced composition of new teachers lacking. I hope to add to the research by investigating changes in the personal characteristics of teachers in depth.

## **2.2 Models on the Decision to Teach**

The literature on econometrically modeling the decision to teach is also very small. Manski (1985) analyzed the role academic ability and earnings play in the decision to becoming a teacher. He created an elaborate utility maximization model for potential teachers as a function of monetary and non-monetary job characteristics. He controlled for student academic ability by using SAT scores and high school class rank, and used procedures to predict how responsive college graduates would be to entering teaching if policy increased their earnings expectation. I am not wholly convinced that potential teachers intentionally maximize monetary earnings, and since I believe that accountability will play a measurable role in informing a potential teacher about non-monetary job conditions, I do not adopt Masnki's utility or probability of teaching models. Further details about my underlying teacher utility model are found in the methodology section of this paper.

A more applicable study, completed by Murnane, Singer, Willett, Kemple and Olsen's (1991), analyzed the role different demographic characteristics and academic ability had on a college graduate's probability of entering the teaching profession. They used the National Longitudinal Survey (NLS) of Young Men, Young Women, and Youth Databases to follow the individuals who graduated college between 1967 and 1984. In their logistic model the probability of becoming a teacher within five years of graduation is defined as a function of graduation year, race, gender, age at graduation, college major, and standardized test score. They found that overall, the proportion of college graduates who were becoming teachers was declining. This

result is not as alarming because it reflected the falling demand for new teachers because children of the baby boom were beginning to leave schools. Findings that were of concern included estimates that confirmed that higher ability and male college graduates were much less likely to enter the teaching profession. They also found that over the surveyed years, the proportion of blacks entering the teaching profession was declining.

A few elements of their model make it hard to fully incorporate with my study. The first difference is that the Murnane model does not offer a model for evaluating policy effects on teacher career choice. It is also limited to describing the probability of becoming a teacher within one time frame. A drawback of their study is their choice to use intelligence quotient (IQ) and the Armed Forces Qualifying Test scores as proxies for academic ability. In today's political landscape, these are not the most believable measures of ability, so if I were to include a proxy for academic ability I would incorporate more accepted measures like a test of academic skills or college competitiveness. In all, Murnane et al. (1991) did provide help in my ability to form an accessible model for predicting teaching career choice.

### **3. Data and Methodology**

#### **3.1 Limitations of the Analysis**

This study is one of the first to assess the impact of accountability policies on the decision to teach at the national level. A disadvantage in attempting a study with national data is that the quality of the questions is more suitable to question about general employment, not career entry. Other variable information that could be helpful is not available because the surveys were not intentionally designed to model teacher career paths. While I can determine if an individual becomes a teacher, I am limited to a broad definition of whether an individual has ever been employed as a school teacher, and cannot be sure if he or she teaches public or private school<sup>1</sup>. Another disadvantage is that my results reflect only the average effects of accountability by state. I cannot specifically differentiate the effects of strong accountabilities on grade levels or subject areas. Nevertheless, because my sample contains states that have implemented a range of strong and weak accountability systems, I can compare the effects of increasingly strong systems.

#### **3.2 Data**

My measure of a state's accountability strength is an index borrowed from Carnoy and Loeb's (2003) nation-wide accountability study on student outcomes<sup>2</sup>. A zero to five score is assigned to each of the 50 states based on information about a state's accountability policies prior to 2000. The score increases as each state's reform policies apply more external pressures on schools and districts. A score of zero indicates no statewide assessment or standards and a

---

<sup>1</sup> Comparison of the Digest of Education Statistics 1990 and 2002 editions indicate that the proportion of American teachers in private schools has remained at between 13.2 and 14.8 percent. Thus, I do not expect my results to reflect any strong bias due to my inability to limit my study for public school teachers

score of one indicates state assessment but does not issue sanctions or rewards based on the student outcomes. Scores of two and three indicate assessments for elementary and middle grades with sanctions and/or rewards, with the difference that states that score a three also require a high school exit exam. A score of four does not require a high school exit exam, but does indicate that the state has harsh penalties on schools or districts that do not improve scores, like threat of reconstitution. If the state meets both standards of a score of three and four, they receive the highest index score of five. Table 4 contains the index classifications of all 50 states in Carnoy and Loeb's analysis.

I combine my state policy data with college graduate data from two surveys conducted by the National Center for Education Statistics (NCES) as part of the National Education Longitudinal Studies (NELS) program. The general mission of the NELS program is to study “the educational, vocational, and personal development of students at various grade levels, and the personal, familial, social, institutional, and cultural factors that affect that development (Ingels and Taylor 1995, vi). The surveys provide detailed information on a respondent's individual characteristics such as race and gender, family socio-economic status, standardized test performance, and attitudes and expectations about future schooling and careers. Since the NELS studies are longitudinal, they follow the same student over time, beginning when she is a high school sophomore until she reaches her mid-20s. The latter follow-up surveys contain information from the participants' high school and post-secondary institutions, including information about the school selectivity ranking. The follow-ups also contain detailed information about the respondent's labor market participation, occupation history, earnings, and job satisfaction rates.

---

<sup>2</sup> The index was created using a database constructed by the Consortium for Policy Research in Education which is available online at <http://www.cpre.org>.

I use two NELS studies, conducted in different years, to create a pooled cross-section of individuals who entered the work force before and after the widespread introduction of assessment-based accountability policies<sup>3</sup>. The individuals from the High School & Beyond of 1980 (HS&B) serve as my control group of new college graduates entering the workforce prior to the implementation of most assessment-based accountability reforms. Conversely, the individuals from the National Education Longitudinal Study of 1988 (NELS:88) serve as my post-accountability reform comparison group.

Because I am only doing my analysis on college graduates, I limit my sample population to survey respondents who report receiving at least a bachelor degree. Of those who are college graduates I keep only those whom I have information for the in which state they live and work. This identification of the state they work in (WORKSTATE) is necessary because accountability strength is measured at the state level. Typical HS&B respondents who finished college graduated high school in 1982, graduated college in 1986, and are interviewed about their chosen career in 1992. A typical NELS:88 college graduate left high school in 1992, college in 1996, and was interviewed about her career choice in 2001. For a graphical representation of the survey schedule see Figure 1. Additional survey data about respondent's post-secondary institutions like the state location and institution selectivity ranking was extracted from the HS&B and NELS:88 surveys' respective Postsecondary Education Transcript Study (PETS) datasets<sup>4</sup>.

Ultimately, I work with a sample of 2103 college graduates from HS&B and 3996 college graduates from NELS:88. Of these individuals 220 and 421 became school teachers from HS&B and NELS:88. As expected, the majority of teachers were white, but the proportion of

---

<sup>3</sup> Similar procedures to combine HS&B and NELS:88 data have been done previously by other studies: Ingels and Taylor 1995; Rasinski, Ingels, Rock, Pollack, and Wu 1993

<sup>4</sup> All NCES datasets are the restricted-use versions that require a license to use. For information on how to obtain a license contact the NCES Data Security Office at <http://nces.ed.gov/pubsearch/licenses.asp>.

non-whites in the sample of college graduates and teachers was about 25 percent. An interesting fact about the samples is that while the average college graduate had a tenth grade test score and socio-economic score of .61 and .66 standard deviations above all members of the original sample, respectively, whereas average teacher scores of .48 and .45 standard deviations above all sample members<sup>5</sup>. Thus, this implies college graduates have higher ability and come from wealthier, more educated families, but of those graduates, individuals who become teachers are, on average, lower in ability. In my teacher sample, the average new school teacher is a white woman, who attended a suburban high school, and went to a non-selective university. I include an extensive cross-table between college graduate and teacher characteristics and NELS survey membership in Tables 1a and 1b. Tables 2a and 2b list the complete descriptive statistics.

The final source of data I incorporate is from the NCES Digest of Education Statistics 2002, an annual report that summarizes data for all levels of schooling in the public and private institution for teachers, students, and administrators. I assigned to each college graduate's state of residence the corresponding average beginning teacher's salary as found in the Digest. Since salaries change over time I use 1990-91 salaries for college graduates from HS&B and 1998-99 salaries for the NELS:88 graduates. All dollar amounts are listed in 1998-1999 dollars. The wage data serves as a monetary proxy for how attractive an occupation teaching is in the periods before and after accountability policies are implemented. For beginning teacher salary data see Table 3.

---

<sup>5</sup> Tenth grade standardized test scores are derived from a composite score on a reading and math test issued as part of the surveys. The scores in both surveys were separately normalized to allow for comparison between HS&B and NELS:88, inclusive of non-college graduates. The contractors first assigned a respondent's socio-economic status score as a function of family income, parent education and parent occupation, as well as other factors. I normalized SES scores separately for all HS&B and NELS:88 respondents, inclusive of non-graduates to allow for comparison between pre- and post-accountability periods.

### 3.2.1 *Inconsistencies with the data*

When I combined the NELS:88 and HS&B data sets, I encountered some problems due to inconsistencies in the way variables were coded. The main difference that may have affected my results was a difference in how the surveys identified where the participant lives and works. Even though the NELS:88 survey includes a variable that indicates where the respondent live as an adult, HS&B does not specify a state of residence. I resolved this problem by finding the mode university attended by students for each HS&B high school. I found the corresponding state of each mode university in the PETS dataset to define the HS&B WORKSTATE variables. This procedure is based under the assumption that the majority of individuals go to college and work in the state where they grew up (Boyd, Lankford, Loeb and Wycoff 2003; Murnane, Singer, Willett, Kemple and Olsen 1991)<sup>6</sup>. However, given that not all participants in the original survey go on to post-secondary institutions and not all continue to participate the follow-ups, I was unable to create a WORKSTATE variable for every respondent in the pre accountability group. Respondents without an identified WORKSTATE were dropped from the sample. Nonetheless, my sample of individuals with bachelor degrees is still representative of 48 of the 50 states: only Alaska and Delaware were excluded.

Regrettably, I realize that when considering the geographic distribution of teachers within the HS&B and NELS:88 surveys, a number of low-population, low-accountability states were not represented. The HS&B sample of teachers did not include people from nine states: Arkansas, Maine, Nevada, North Carolina, North Dakota, Rhode Island, South Dakota, Vermont and Virginia. Additionally, the NELS:88 teacher sample did not contain any people from Montana and West Virginia. Since my logits capture the state effects of accountability, I

---

<sup>6</sup> I owe many thanks to Professor Susanna Loeb for her insight and suggestion to use this strategy for defining the HS&B state of residence.

recognize that the disproportionate representation of people from certain states might create larger standard errors, and skew the significance and/or direction of my results. I account for the standard errors later by clustering my logits by state.

### **3.3 Methodology**

When approaching the question of how to model the probability of an individual becoming a teacher, I first consider how individuals choose a profession. Utility maximization in the area of occupational choice theory asserts that people choose the profession that gives them the most happiness or offers them the most utility. For many professions, one might think that a college graduate simply chooses a profession that maximizes salary given her set of skills and next best employment alternatives in the labor market. However, the decision to teach is much more complex than a decision to maximizing earnings. For example, in one survey of high school students' attitudes toward becoming a teacher, the primary reason for the desire to teach was the responsibility of preparing youth for the future (Valadez 2003). This is not to say salaries do not make a difference for who goes into teaching. While I cannot control for all the intrinsic and altruistic preferences that are maximized when choosing to teach, I cannot exclude the possibility that money does not affect the decision to teach. Therefore, I assume average beginning teachers' salary in a state at least serves as a benchmark for the wage in which teaching becomes a viable career choice.

When choosing a career or job, individuals maximize non-monetary benefits as well, like working conditions. For a teacher, working conditions might include quality of the facilities or peer support. If an individual feels she will have strong administrative and parental support or if she is convinced she can make a difference in the classroom, she might be more likely to go into or continue teaching (Farkas, Johnson, and Faleno 2000). In the same way, strength of

accountability policies serve as one type of tool college graduates can use to infer working conditions. If a potential teacher thinks the school work environment is more hostile and restrictive because of increased teacher scrutiny and narrow focus on student test performance in strong accountability environments, she may be less inclined to believe she could make a difference as a teacher and be influenced not to teach. Conversely, if a potential teacher views accountability as a challenge that the administration and parents are equally willing to embrace, she may be more inclined to teach in such a supportive accountability environment.

Other issues I include in a college graduate's decision to teach are her other employment opportunities in careers other than in teaching. Typically, education literature accepts standardized test score performance or selectivity of an individual's undergraduate institution as proxies for ability. I use these variables under the assumption that the higher one's ability, the more likely she has additional, potentially higher paying career opportunities outside of teaching. While standardized test performance and undergraduate institution selectivity cannot perfectly describe what determines a high-ability or high-quality teaching candidate, they help describe what next best career alternatives exist for an individual.

I also make a few assumptions about the pathway into teaching and what considerations recent college graduates take into account when choosing a career. For simplicity, I assume that there are no barriers for individuals with a bachelor degree to become a teacher. In reality, each state has a certification process that may require additional testing, education-related course hours, in order to be licensed to become eligible to be a school teacher. These practical barriers are undoubtedly associated with additional cost of becoming a school teacher, but since the data is limited and opportunity costs vary widely across states, I cannot consider the effects of external barriers to the teaching profession. In my model, the only policy that affects a college

graduate's decision to teach is the strength of accountability reform in her state. In sum, I believe a college graduate's decision to teach is based on the threshold wage where teaching becomes a considered option, as well as the quality of working conditions as related to the strength of accountability policies in her state, and her ability to find work outside of teaching.

### 3.3.1 *Baseline Effects*

My main analysis tries to find any causal, differential effect of stronger accountability. Before I look at the differential effects of strong accountability on specific sub-groups, I use a baseline regression to estimate the general trends of becoming a teacher due to variation in the policy environment. I use the same variables I believe a college graduate considers when deciding to teach as the underlying model for determining the probability of becoming a teacher. Controls are added for the demographic characteristics of teachers including gender, race and ethnicity, and urbanicity of their own high schools. The logit takes the following form:

$$\text{Probability}(\text{TEACH}=1|x) = e^{x\beta} / (1 + e^{x\beta})^7$$

$$\text{Where, } x\beta = \beta_0 + \beta_1\text{GENDER} + \beta_2\text{RACE} + \beta_3\text{URBANICITY} + \beta_4\text{SALARY} + \beta_5\text{TEST} + \beta_6\text{SELECTIVE} + \beta_7\text{INDEX} + \beta_8\text{NELS88} + \beta_9\text{INDEX*NELS88} \quad (1)$$

Here the probability that a college graduate will become a school teacher is a function of: the average beginning teacher's salary (SALARY) for the state of residence during the fourth follow-up survey; the corresponding accountability index (ACCOUNT) score for the state; and interaction dummy (NELSACCT) on accountability index and NELS:88 membership; where I control for whether he or she is in the NELS:88 (NELS88) study; academic ability as

---

<sup>7</sup> Long 1997.

demonstrated by the standardized composite math and reading score of a survey assessment administered during their sophomore year of high school (TEST) and dummy the selectivity of their undergraduate institution (SELECTIVE); a control for the characteristics of the urbanicity of their high schools (dummies for RURAL and URBAN); a control dummy for the gender, namely if they are male (MALE); and dummies for his or her ethnicity (dummies for ASIAN, BLACK, and HISPANIC). In general, the coefficients of teacher-choice variables and personal characteristics will reflect the general odds of becoming a teacher in both time periods, while the coefficient of the interaction term will indicate a change in the odds of teaching specific to accountably strength. In other words, the coefficient of INDEX\*NELS88 indicates the relative change in proportion of college graduates becoming teachers for states that have adopted stronger accountability policies. The logit is weighted for survey design and clustered by state to adjust standard errors since accountability is measured at the state-level.

### 3.3.2 *Differential Effects*

I use a similar logit model to find differential effects on various teacher characteristics in states that adopt strong accountability systems on the probability of becoming a school teacher using the following general equation.

$$\text{Probability}(\text{TEACH}=1|x) = e^{x\gamma}/(1+ e^{x\gamma})$$

$$\begin{aligned} \text{Where, } x\gamma = & \gamma_0 + \gamma_1 \text{GENDER} + \gamma_2 \text{RACE} + \gamma_3 \text{SALARY} + \gamma_4 \text{TEST} + \gamma_5 \text{SELECTIVE} + \\ & \gamma_6 \text{INDEX} + \gamma_7 \text{NELS88} + \gamma_8 \text{INDEX*NELS88} + \gamma_9 \text{INDEX*CHAR} + \gamma_{10} \text{NELS88*CHAR} + \\ & \gamma_{12} \text{CHAR*INDEX*NELS88} \end{aligned} \quad (2)$$

In these models I retain the same controls used in the baseline equation, but I also chose a specific teacher characteristic, such as teachers from selective colleges, to evaluate and interact its effect with other variables. The characteristic of interest (CHAR) is separately interacted with index and finally the characteristic is interacted with both index and NELS88. The final three-way interaction's coefficient specifically indicates the differential effect of stronger accountability. For example, if I am evaluating the relative changes in college graduates from selective colleges becoming teachers in higher accountability states, I will interpret the coefficient of SELECTIVE\*INDEX\*NELS88. I also conduct procedures to weight my observations for survey design and cluster by states so that the standard errors reflect that accountability strength is measure at the state level.

### 3.3.3 *Descriptive Analysis*

The last analysis on teacher job satisfaction rates is purely descriptive. For questions that appear in both the pre- and post-accountability periods, I consider if teachers' responses about teacher pay or job challenges and importance differ for states that adopted stronger accountability systems than teachers in states with weak or no accountability. For the one survey item that is only available during the post-accountability period, I determine whether teachers in strong accountability states express the same levels of job autonomy as teachers in lower accountability states.

## 4. Results

### 4.1 Basic Teacher Entry Trends

The coefficients of Equation 1 are listed on Table 5. It gives us a glimpse of how individuals are making the decision to teach in a world not simply influenced by changes other than in response to accountability policy strength. As we know, there were many simultaneous policies affecting the teacher work life and teacher demand during the 1980s and 1990s, which I have not specifically controlled for. One way to interpret the coefficients of this analysis is to observe what types of individuals were more or less likely to teach for both HS&B and NELS:88 groups. In general, the odds of becoming teachers are 67 percent lower for men, 65 percent lower for Asians, and 44 percent lower for Blacks, irrespective of accountability strength. These odds are statistically significant at the .01 levels. Also, as one might expect, higher ability individuals who achieved better scores on a standardized test and who went to more selective universities are less likely to become teachers. Scoring a standard deviation higher than the average college graduate decreases one's odds of becoming a teacher by 11 percent, and going to a selective college decreases one's odds of teaching by 27 percent. The decreased odds of teaching for high-ability individuals are significant at the .05 levels. This suggests that many highly qualified college graduates are attracted to professions outside of teaching. The coefficients on these specific teacher characteristics can be attributed to the general policy environment, not limited to accountability. From these preliminary results, the consequence of whether stronger accountability will further discourage male, minority, and high-quality candidates from teaching is an immediately relevant policy question.<sup>8</sup>

---

<sup>8</sup> In general, the regression on coefficient on NELS:88 indicates that the odds of becoming a teacher for college graduates in the NELS88 survey are 23 percent lower than their HS&B counterparts. This result is surprising

The interaction between NELS:88 and the accountability index introduces a base relative change of the likelihood to teach for states that adopted stronger accountability systems. At first glance, the odds ratio above one can be interpreted that the odds of becoming a teacher in a higher accountability state increase by 8 percent for each one-unit increase in accountability index score. While this effect might seem large, it is important to note that its effect is not statistically relevant, even at the .10 level. To better understand how sub-populations of college graduates were affected by strong accountability systems, I turn to my second group of analyses.

#### **4.2. Differential Effects of Strong Accountability**

I consider the differential causal effects of strong accountability on high-quality candidates, men, non-whites, and other groups of people I consider likely teachers. When evaluating these effects I describe the three-term interaction variable bolded at the end of each table.

##### *4.2.1. Effects on Teacher Quality*

Tables 6a and 6b describe the relative change states with high accountability saw in graduates with higher test scores and college graduates who attended more selective undergraduate institutions becoming teachers. Since this population of high-quality teacher candidates is already less likely to teach without the effects of accountability, a strong and statistically significant decrease in the odds of becoming a teacher would provide evidence for a detrimental effect of accountability. Table 6a indicates that in states that implement higher accountability, a college graduate who scores a standard deviation above her peers decreases her probability of teaching by 3 percent. Even so, this suggested negative effect of accountability is

---

because it suggests that the NELS88 sample of recent graduates entering the workforce did not pick up any increased demand of teachers in the post-accountability period due to the increased demand because of class-size reduction policies in the 1990s. However, this coefficient is not significant on its own and its large standard error indicates that it is not being precisely measured. Also, I choose not to comment on the urbanicity effects, because

small and statistically insignificant. Table 6b shows in states that implemented stronger accountability systems, college graduates from more selective colleges observe a 6 percent decrease in the odds of becoming a teacher, but this effect is not statistically significant either. One interesting finding is that in each of the regressions, when evaluating the differential effect of one quality measure the alternative quality measure becomes significant. This phenomenon suggests that high test scores and attending a more selective institution are highly correlated with each other. To summarize, the analysis provides no conclusive evidence accountability discourages high-quality individuals from teaching.

#### *4.2.2. Effects on Teacher Gender*

The differential effect of men becoming teachers is available in Table 6c. In this analysis, stronger accountability decreases the relative probability of a man becoming a teacher by a strong and statistically significant amount. The odds of a man becoming a teacher decrease by 27 percent in states that have adopted a higher accountability policy. This discouraging effect of accountability is significant at the .05 level.

#### *4.2.3. Effects on Teacher Racial and Ethnic Composition*

In the case of non-whites, the evidence for a strong accountability effect is not as clear as the case with men (See Table 6d). The interacted effect of becoming a teacher for a non-white minority living in a higher accountability state, during the relevant policy period seems to increase the odds of becoming a teacher. Since this effect is not statistically significant, it is hard to be certain of accountability's positive influence on minorities choosing to teach. This may reflect better minority recruitment strategies in high-accountability states or it may reflect that non-whites are more responsive to higher wages that are correlated with strong accountability

---

the standard errors for graduates from urban high schools are so large, while the effects of rural seem to be very large.

states. Even so, the case may be that the data is not rich enough to estimate any effect of accountability on the probability of non-whites teaching to any degree of statistical certainty. Overall, there is no conclusive evidence to suggest non-white teachers are negatively influenced by strong accountability policies, and there is weak evidence that indicates non-white may be more attracted to teaching in high accountability states.

#### *4.2.4. Effects on Likely Teachers*

Tables 6e and 6f offer analysis on the differential change states with stronger accountability systems have experience with respect to groups of college graduates I consider “likely teachers.” The first type of likely teachers I evaluate are undergraduate majors in either elementary or secondary education, special education, physical education, or other education major. While not all education majors become teachers, and not all teachers major in education, I believe it is important to evaluate whether stronger accountability systems have disproportionately affected the career choice of potential teachers. The second group of likely teachers I analyze are individuals who have expressed the desire to teach. Both surveys asked respondents in the sophomore year of high school and the sophomore year of college, “What job or occupation do you plan to have when you are age 30?” If the respondent answered “school teacher” in either high school or college, I create a positive response for a “WNT2TCH” dummy. I interact the accountability index score and NELS88 participation dummy separately for education majors and college graduates who wanted to become school teachers. From the regression analysis, it is likely teachers from strong accountability states are not discouraged from teaching. In fact, the odds of an education major becoming a teacher in a high accountability state are 14 percent higher than education majors in low accountability states, and the odds of a college graduate who expressed the desire to become a school teacher increases by

4 percent. However, the strength of beneficial effects of accountability on the increasing the likelihood to teach is not precisely estimated, and is not statistically significant at the .10 level. Once again, the results provide no conclusive evidence to support the case for accountability's negative effects on the supply of new teachers.

### **4.3 Descriptive Analysis**

#### *4.3.1 Effects on Teacher Job Satisfaction*

The descriptive analysis begins with Figure 2a. The first work life question I consider asks teachers if they are satisfied with their job pay. Teachers in the post-accountability period are universally less satisfied with job pay than teachers in the pre-accountability period. However, no accountability group experiences a significant change in the decrease of satisfied teachers. The second question issued in both survey asks teachers, "Would you say that you (are/were) satisfied or dissatisfied with each of the following elements of your job" in terms of job challenge/importance. Unlike their responses to job pay, post-accountability teachers are universally more satisfied with the level of their job challenge and importance at all levels of accountability strength. The increase in teachers' job challenge/importance satisfaction is statistically significant for states that implemented the strongest types of reforms (states with accountability index scores of 4 or greater). In particular, the increase of 86.09 percent of teachers claiming to be satisfied during the pre-accountability period to 92.08 percent satisfied in the post-reform years is statistically significant at the .05 level.

The final teacher work life question is only available for NELS:88. School teachers were asked, "Which one of the following four statements best describes your job?" Respondents could choose one of four responses: 1) someone else decides what you do and how you do it; 2) someone else decides what you do, but you decide how to do it; 3) you have some freedom in

deciding what you do and how to do it; or 4) you are basically your own boss. Because the responses are so few in the extreme responses, I combine responses 1 and 2 as unsatisfied and response 3 and 4 as experiencing more job autonomy. As Figure 3 indicates, on average, schools with weaker accountability systems have a greater proportion of teachers that feel they have more autonomy over their work. Nevertheless, the difference in satisfaction rates between weak and high accountability states is not significant at the .10 level. In general, teachers' responses about job satisfaction in regards to pay, importance and challenge, and autonomy provide no strong evidence that stronger accountability weakens the perceptions of how teachers view their jobs.

## 5. Conclusion

I find no conclusive evidence that adopting strong accountability policies decreases a state's ability to attract high-quality, non-white, or other likely teaching candidates to teaching. In fact, states with stronger accountability may be observing are *more* likely to observe an increase in the proportion of non-white college graduates and education majors achieving careers as teachers. Additionally, I find no conclusive evidence that strong accountability significantly decreases teachers' job satisfaction or decreases their sense of job autonomy. The only damaging effect my results confirm is the decreased likelihood to teach for male college graduates in states that have implemented strong accountability policies. Taken as a whole, my research indicates that accountability is not markedly detrimental to recent college graduates' decisions to become teachers.

One reason why my results are inconclusive is due to the quality of my data. My sample size of teachers across states may not have been diverse enough to capture the true effects varying accountability strength nationally. As data sources improve and become readily available, it is feasible that similar studies about accountability's effects on the decision to teach could continue at the state or national level. Additional studies about the effects of accountability on the teacher labor supply might one day extend to sub-groups like mid-career-changing professionals or specifically skilled college graduates in math and science.

As for future policy recommendations, the consistently negative correlation between men, Asians, Blacks, and high-ability college graduates and the event of choosing to teach suggests that policy makers must develop and execute recruitment strategies to ensure the quality and diversity of tomorrow's teachers. The same applies to recruitment strategies for all areas of

hard-to-staff schools and subject areas. Because my data could not describe what grade level, subjects, or school types respondents taught at, there may well be negative differential effects to college graduates choosing to teach tested grades, tested subjects, or teach in schools that serve low-performing students – all of which were research questions addressed by the previous studies of accountability and existing teachers’ career decisions. Also, since there are slightly higher correlations between high accountability states and beginning teacher wages, additional research would be helpful in determining what populations of high-quality college graduates are most responsive to increased wages. Specifically, in my analysis I observed that non-whites were more inclined to teach in high accountability schools. Because my research is limited in scope, I was not able to determine if non-white college graduates were choosing to teach because of the higher wages offered by high accountability states or because high accountability states tend to have larger non-white populations to begin with. In all, policy makers should seek out detailed analysis about how they can recruit highly skilled college graduates of all backgrounds and what wage policies may be likely to compliment accountability reforms.

The importance of understanding how strong accountability policies differentially affect who decides to teach increases daily following the more recent reforms mandated by the No Child Left Behind (NCLB) Act of 2001. During the 1990s states freely chose the strength of accountability systems they adopted. In contrast, NCLB mandates that *all* states adopt statewide assessments in multiple subject areas for elementary and secondary students in an effort “to help all students meet high academics standards”(US Department of Education 2002, p. 9). NCLB explicitly requires that schools ensure students achieve at certain standards to meet “adequate yearly progress.” If schools meet or exceed these standards they are eligible for “academic achievement awards,” but if they fail, districts and schools will “be targeted for

assistance...subject to corrective action, and ultimately restructuring” (US Department of Education 2002, p. 10). The reforms require that all states adopt strong assessment-based accountability systems, so the future effects of accountability policies will reflect more uniform policy strategies. In order to understand how accountability will affect potential teachers differently in the future, continual emphasis should be placed on making good data available for future research. While introduction of NCLB will make it impossible to distinguish the effects of the 1990s, the universal accountability effects due to NCLB requirements will be of even greater significance. If policy makers recognize the importance of improving the recruitment and retention of highly qualified teacher candidates during this period of reform, I speculate that the effects of accountability on people entering teaching can remain negligible, and perhaps with greater cooperation, even beneficial.

## 6. Reference List

- Boyd, Donald, Hamilton Lankford, Susanna Loeb, and James Wyckoff. 2003. "Do Mandatory Tests Affect Teachers' Exit and Transfer Decisions? The Case of the 4<sup>th</sup> Grade Test in New York State." Unpublished Manuscript.
- Boyd, Donald, Hamilton Lankford, Susanna Loeb, and James Wyckoff. 2003. "The Draw of Home: How Teachers' Preferences for Proximity Disadvantage Urban Schools." National Bureau of Economic Research Working Paper 9935.
- Carnoy, Martin, and Susanna Loeb. 2003. Does External Accountability Affect Student Outcomes? A Cross-state Analysis. *Educational Evaluation and Policy Analysis* 24(4): 305-331.
- Clotfelter, Charles T., Helen F. Ladd, Jacob L. Vigdor, and Roger A. Diaz. 2003. "Do School Accountability Systems Make It More Difficult for Low Performing Schools to Attract and Retain High Quality Teachers?" Unpublished manuscript.
- Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander Mood, Fredrick Weinfeld, & Robert L. York. 1988. *Equality of Educational Opportunity*. Washington, DC: Office of Education, National Center for Education Statistics, 1966. Reprint, Salem, NH: Ayer Company Publishers.
- Darling-Hammond, Linda, Mary E. Dilworth, and Marcy Bullmaster. 1996. *Educators of Color: The recruitment, preparation, and retention of persons of color in the teaching profession*. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.
- Ehrenberg, Ronald G., Daniel D. Goldhaber, and Dominic J. Brewer. 1995. Do Teachers' Race, Gender, and Ethnicity Matter? Evidence from the National Educational Longitudinal Study of 1988. *Industrial and Labor Relations Review* 48(3): 547-561.
- Farkas, Steve, Jean Johnson, and Tony Faleno. 2000. "A Sense of Calling: Who Teaches and Why." New York: Public Agenda Foundation. Available from <http://www.publicagenda.org>. Accessed 26 May 2004.
- Hanushek, Eric A., and Richard R. Pace. 1995. Who Chooses To Teach (and Why)? *Economics of Education Review* 14(2): 101-117.
- Ingels, Steven, and John Taylor. 1995. National Education Longitudinal Study of 1988: Conducting Cross-Cohort Comparisons Using HS&B, NAEP, and NELS:88 Academic Transcript Data. Database on-line. Available from ERIC, ED418110.

- Loeb, Susanna, and Felicia Estrada. 2003. "Have Assessment-Based School Accountability Reforms Affected the Career Decisions of Teachers?" in Proceedings of the Educational Testing Service Invitational Conference, Measurement and Research Issues in a New Accountability Era. Forthcoming.
- Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Advanced Quantitative Techniques in the Social Sciences Series. Thousand Oaks: SAGE Publications.
- Manski, Charles F. 1987. Academic Ability, Earnings, and the Decision to Become a Teacher: Evidence from the National Longitudinal Study of the High School Class of 1972" in *Public Sector Payrolls*, ed. David A. Wise. Chicago: University of Chicago Press.
- Murnane, Richard J., Judith D. Singer, John B. Willet, James J. Kemple, and Randall J. Olsen. 1991. *Who Will Teach? Policies That Matter*. Cambridge: Harvard University Press.
- National Center for Education Statistics. *Digest of Education Statistics 2002*. March 2003.
- National Center for Education Statistics. *Digest of Education Statistics 2000*. March 2001.
- National Center for Education Statistics. *Digest of Education Statistics 1990*. March 1991.
- Rasinski, Kenneth, Steven Ingels, Donald Rock, Judith Pollack, and Shi-Chiang Wu. America's High School Sophomores a Ten Year Comparison. Database on-line. Available from ERIC, ED360318.
- U.S. Department of Education. Office of the Under Secretary. 2002. *No Child Left Behind: A Desktop Reference*. Washington, D.C.
- Valadez, Jacqueline P. 2003. High School Students' Perceptions of Becoming an Educator: A Case Study in South Texas. Ph.D. diss., Saint Mary's University.
- Wise, Arthur E. 2003. Your Child's Learning Depends on Teacher Knowledge and Skill. *Our Children* 28(6): 9-10.

## 7. Data Products

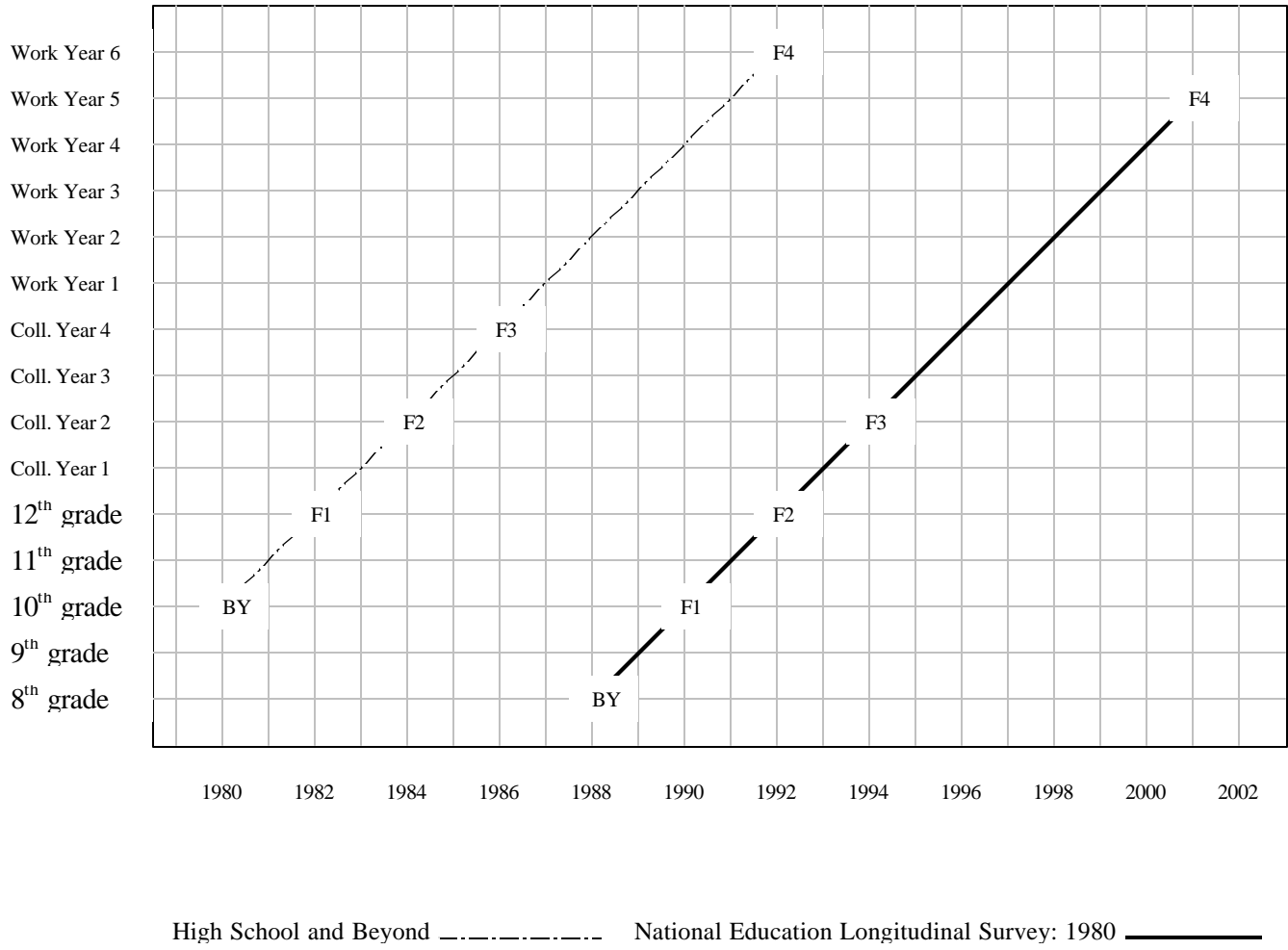
National Center for Education Statistics. High School & Beyond Sophomore Cohort, Restricted Data. CD-ROM (NCES 98-135).

National Center for Education Statistics. High School & Beyond Sophomore Cohort: 1980-92 Postsecondary Education Transcripts and Supplement, Restricted Data. CD-ROM (NCES 2000-194).

National Center for Education Statistics. National Education Longitudinal Study of 1988: Base Year through Fourth Follow-up, Restricted Data. CD-ROM (NCES 2003-348).

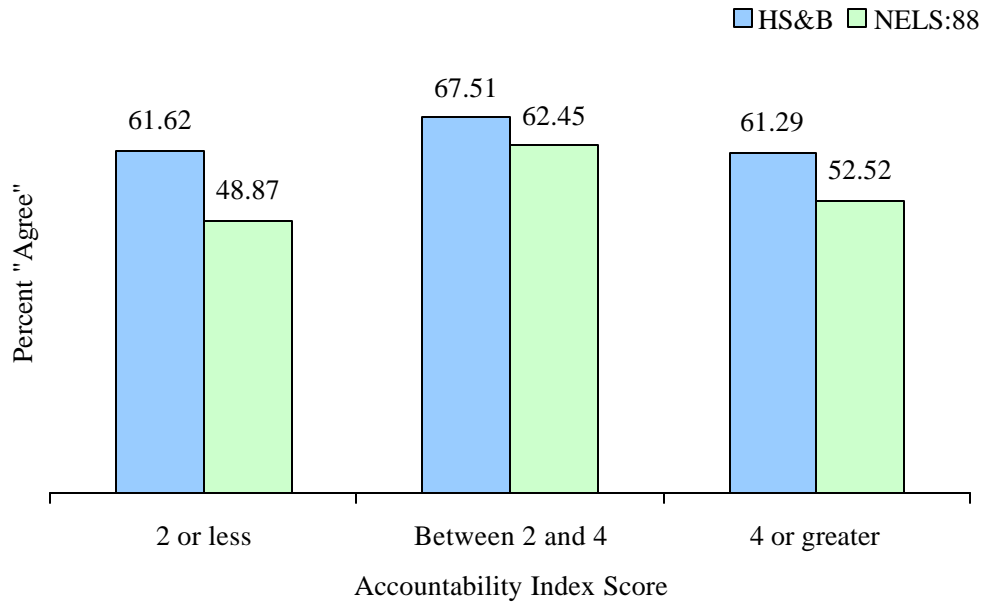
National Center for Education Statistics. National Education Longitudinal Study of 1988/2000: Postsecondary Education Transcript Study, Restricted Data. CD-ROM (NCES 2003-402).

**Figure 1. NELS Survey Schedules**

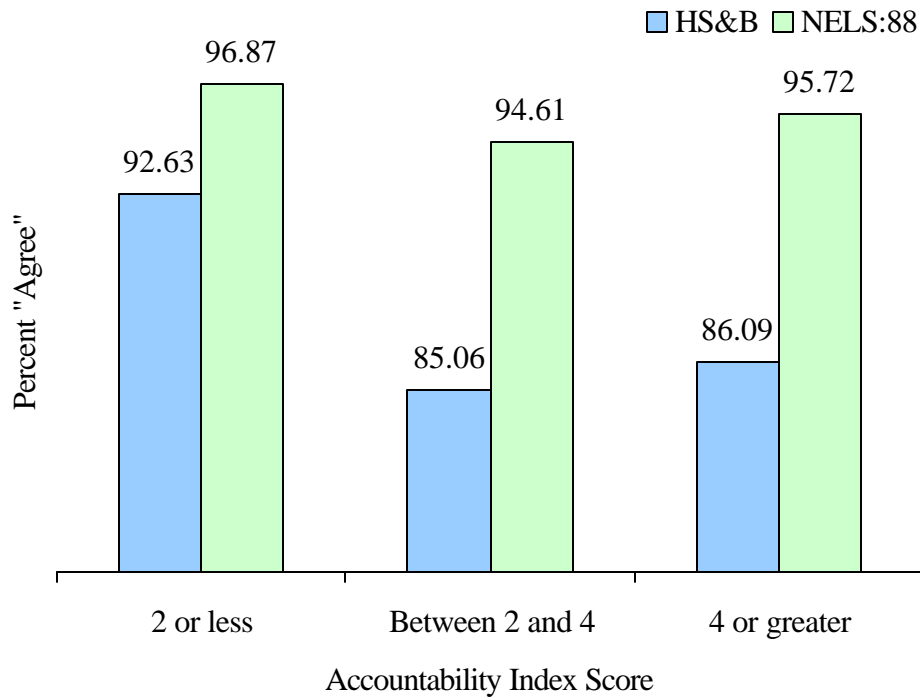


BY=Base Year F1=1<sup>st</sup> Follow-Up F2=2<sup>nd</sup> Follow-Up F3=3<sup>rd</sup> Follow-up F4=4<sup>th</sup> Follow-up

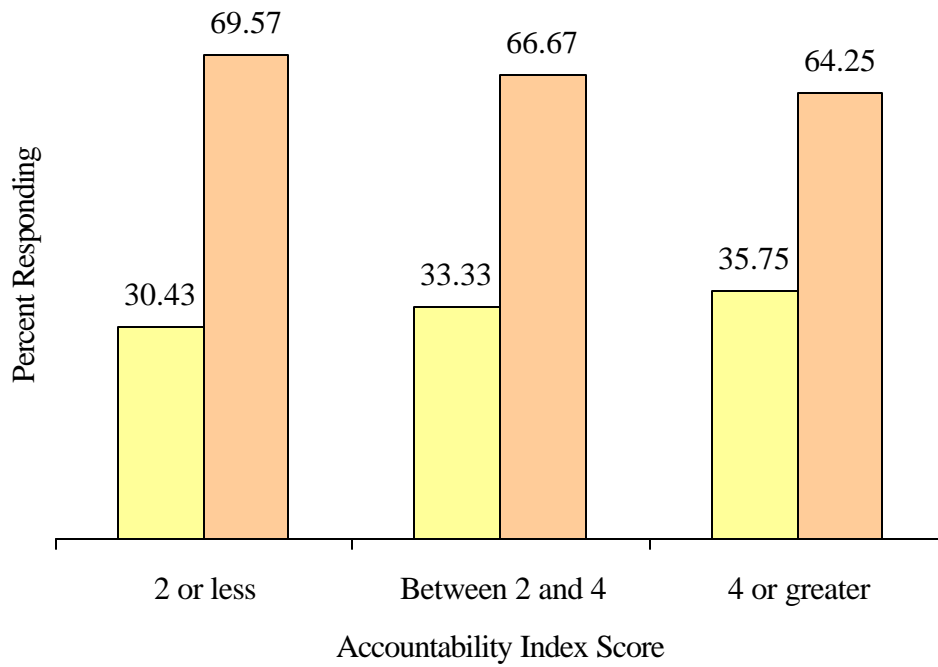
**Figure 2a. Percent of Teachers That Are Satisfied with Job Pay. HS&B and NELS88 Fourth Follow-ups**



**Figure 2b. Percent of Teachers That Are Satisfied with Job Importance/Challenge. HS&B and NELS88 Fourth Follow-ups**



**Figure 3. Teachers' Perceived Job Autonomy in 2001 from the NELS88 Fourth Follow-up**



- "Someone else decides what you do and how you do it" or "Someone else decides what to do, but you decide how to do it"
- "You have some freedom in deciding what you do and how you do it" or "You are basically your own boss"

**Table 1a. College Graduate-Only Data by Survey**

	HS&B	NELS 88	Total
BACHELOR	2103	3996	6099
TEACH			
NO	1883	3575	5458
YES	220	421	641
FEMALE	1093	2234	3327
MALE	1008	1762	2770
WHITE	1556	2997	4553
NON-WHITE	540	999	1539
EDMAJOR			
NO	1893	3613	5506
YES	210	383	593
WNT2TCH			
NO	1713	3220	4933
YES	223	305	528
SELECTIVE			
NO	1299	2438	3737
YES	262	1298	1560
URBAN HIGH	402	1285	1687
SURBAN HIGH	1291	1713	3004
RURAL HIGH	410	951	1361
WORKSTATE	AK		
NOT	DE		
REPRESENTED			

**Table 1b. Teacher-only Data Table by Survey**

	HS&B	NELS	Total
TEACH	303	471	774
FEMALE	232	369	601
MALE	70	102	172
	302	471	773
EDMAJOR			
NO	176	215	391
YES	127	256	383
WNT2TCH			
NO	152	332	484
YES	131	91	222
WHITE	199	370	569
NON-WHITE	102	101	203
SELECTIVE			
NO	193	336	529
YES	34	100	134
URBAN HIGH	61	126	187
SURBAN HIGH	172	193	365
RURAL HIGH	70	148	218
WORKSTATE	AR	MT	
NOT	ME	WV	
REPRESENTED	NV		
	NC		
	ND		
	RI		
	SD		
	VT		
	VA		

**Table 2a. HS&B and NELS88 Combined College Graduate Sample Descriptives**

Variable	Obs	Mean	Std. Dev.
Male	6097	45.43%	
White	6092	74.74%	
Black	6092	7.30%	
Hispanic	6092	7.65%	
Asian	6092	9.37%	
Nonwhite	6092	25.26%	
SES	5809	0.6391	0.90
TEST	5682	0.6705	0.83
SELECTIVE	5297	29.45%	
URBAN	6052	27.88%	
SUBURBAN	6052	49.64%	
RURAL	6052	22.49%	
WNT2TCH	5461	9.67%	
EDMAJOR	6099	9.72%	
TEACH	6099	10.51%	
SJOBPAY	6009	72.99%	
SJOBIMP	6001	83.10%	
SJOBSEC	5956	84.13%	
SJOBAUT	3976	65.87%	
WEIGHT	6099	223.03	218.54

**Table 2b. HS&B and NELS88 Combined Teacher Sample Descriptives**

Variable	Obs	Mean	Std. Dev.
Male	773	22.25%	
White	772	73.70%	
Black	772	8.03%	
Hispanic	772	12.18%	
Asian	772	4.27%	
Nonwhite	772	26.30%	
SES	740	.5091	0.92
TEST	715	0.4615	0.86
SELECTIVE	663	20.21%	
URBAN	770	24.29%	
SUBURBAN	770	47.40%	
RURAL	770	28.31%	
WNT2TCH	706	31.44%	
EDMAJOR	774	49.48%	
SJOBPAY	766	59.27%	
SJOBIMP	767	92.70%	
SJOBSEC	768	85.94%	
SJOBAUT	471	66.67%	
WEIGHT	774	224.87	218.18

**Table 3. Minimum and average teacher salaries by state 1990-91 and 1998-99  
( in 1998-99 dollars)**

State	1990-91	1998-99		Percent change, 1990-91 to 1998-99 (constant dollars)	
	Minimum (Beginning) Salary	Minimum (Beginning) Salary	Minimum (Beginning) salary as a percent of average salary	Minimum (beginning) salary	Average salary
<b>United States</b>	<b>\$26,468</b>	<b>\$26,639</b>	<b>65.7</b>	<b>0.6</b>	<b>0.4</b>
Alabama	27,171	29,092	81.2	7.1	8.6
Alaska	36,799	32,884	68.1	-10.6	-9.5
Arizona	26,263	26,163	75.7	-0.4	-8.5
Arkansas	21,450	21,273	64.9	-0.8	12.3
California	30,189	29,105	62.8	-3.6	-3.6
<b>Colorado</b>	<b>24,311</b>	<b>25,489</b>	<b>66.8</b>	<b>4.8</b>	<b>-2.4</b>
<b>Connecticut</b>	<b>31,101</b>	<b>31,391</b>	<b>62.4</b>	<b>0.9</b>	<b>-5.7</b>
<b>Delaware</b>	<b>25,940</b>	<b>29,981</b>	<b>69.4</b>	<b>15.6</b>	<b>-0.2</b>
<b>Florida</b>	<b>26,255</b>	<b>24,402</b>	<b>67.9</b>	<b>-7.1</b>	<b>-4.3</b>
<b>Georgia</b>	<b>25,152</b>	<b>27,908</b>	<b>71.6</b>	<b>11</b>	<b>9.6</b>
Hawaii	29,233	28,315	70.1	-3.1	-2.0
Idaho	19,272	20,814	61.1	8	8.7
Illinois	26,975	28,954	63.9	7.3	6.4
Indiana	24,877	26,171	63.6	5.2	1.7
Iowa	23,841	24,333	69.5	2.1	1.9
<b>Kansas</b>	<b>23,289</b>	<b>23,006</b>	<b>66.4</b>	<b>-1.2</b>	<b>0.0</b>
<b>Kentucky</b>	<b>23,727</b>	<b>24,387</b>	<b>68.9</b>	<b>2.8</b>	<b>-1.1</b>
<b>Louisiana</b>	<b>21,485</b>	<b>23,500</b>	<b>73.4</b>	<b>9.4</b>	<b>-0.5</b>
<b>Maine</b>	<b>23,195</b>	<b>24,962</b>	<b>71.5</b>	<b>7.6</b>	<b>-0.4</b>
<b>Maryland</b>	<b>28,933</b>	<b>27,605</b>	<b>64.9</b>	<b>-4.6</b>	<b>-9.6</b>
Massachusetts	26,785	28,005	63.6	4.6	-0.7
Michigan	27,523	27,822	57.1	1.1	4.9
Minnesota	25,838	24,462	61.4	-5.3	-2.2
Mississippi	23,284	21,346	72.2	-8.3	-2.3
Missouri	24,934	25,164	75.2	0.9	-1.5
<b>Montana</b>	<b>22,608</b>	<b>21,676</b>	<b>68.7</b>	<b>-4.1</b>	<b>-3.9</b>
<b>Nebraska</b>	<b>22,539</b>	<b>22,611</b>	<b>68.8</b>	<b>0.3</b>	<b>0.6</b>
<b>Nevada</b>	<b>29,928</b>	<b>28,842</b>	<b>67.8</b>	<b>-3.6</b>	<b>-1.9</b>
<b>New Hampshire</b>	<b>25,354</b>	<b>24,406</b>	<b>65.2</b>	<b>-3.7</b>	<b>-2.7</b>
<b>New Jersey</b>	<b>30,103</b>	<b>29,112</b>	<b>56.3</b>	<b>-3.3</b>	<b>9.5</b>

New Mexico	23,497	24,393	75.8	3.8	1.5
New York	32,407	30,808	62.0	-4.9	-3.9
North Carolina	24,340	25,338	68.7	4.1	2.9
North Dakota	19,996	19,136	66.0	-4.3	0.1
Ohio	22,672	23,087	56.7	1.8	3.7
<b>Oklahoma</b>	<b>22,823</b>	<b>25,258</b>	<b>81.2</b>	<b>10.7</b>	<b>3.9</b>
<b>Oregon</b>	<b>25,012</b>	<b>28,589</b>	<b>65.3</b>	<b>14.3</b>	<b>10.4</b>
<b>Pennsylvania</b>	<b>28,567</b>	<b>29,793</b>	<b>61.5</b>	<b>4.3</b>	<b>9.4</b>
<b>Rhode Island</b>	<b>25,664</b>	<b>26,237</b>	<b>56.7</b>	<b>2.2</b>	<b>-1.4</b>
<b>South Carolina</b>	<b>24,275</b>	<b>23,827</b>	<b>69.1</b>	<b>-1.8</b>	<b>-0.3</b>
South Dakota	20,490	21,376	75.3	4.3	3.3
Tennessee	24,758	22,645	63.8	-8.5	2.3
Texas	24,758	26,261	76.2	6.1	-0.2
Utah	21,175	22,957	67.5	8.4	8.9
Vermont	22,742	25,435	69.3	11.8	0.5
<b>Virginia</b>	<b>27,284</b>	<b>25,777</b>	<b>68.4</b>	<b>-5.5</b>	<b>-6.1</b>
<b>Washington</b>	<b>25,326</b>	<b>23,645</b>	<b>61.4</b>	<b>-6.6</b>	<b>-4.9</b>
<b>West Virginia</b>	<b>23,011</b>	<b>23,316</b>	<b>68.1</b>	<b>1.3</b>	<b>7.3</b>
<b>Wisconsin</b>	<b>25,420</b>	<b>24,839</b>	<b>63.1</b>	<b>-2.3</b>	<b>-3.1</b>
<b>Wyoming</b>	<b>23,637</b>	<b>22,836</b>	<b>68.2</b>	<b>-3.4</b>	<b>-6.0</b>

Source: National Center for Education Statistics. Digest of Education Statistics 2000.

**Table 4. Carnoy and Loeb's Accountability Index, by State, 1999-2000**

State	Grades with State Testing in 1999-2000	School Accountability 1999-2000	Repercussion for Schools 1999-2000	Strength of Repercussion for Schools 1999-2000	HS Exit Test in 2000	Grade HS Test First Given	First Grad Class	Index
Alabama	3-11	School report cards	Ratings, intervention	Strong	Yes	10	2001	4
Alaska	4-7	None	None	None	Yes	10	2002	1
Arizona	3,5,8,10	Report cards	'Public shame'	Weak	Yes	10	2002	2
Arkansas	4, 6	None	None	None	No			1
California	2-11	Report cards	Ratings, awards, intervention	Strong	No	10	2004	4
Colorado	3, literacy	None	None	None	No			1
Connecticut	4,6,8,10	Reporting scores to state	Identify schools with needs	Weak	No			1
Delaware	3,5,8,10,11	None	None	None	No	10	2004	1
Florida	4,5,8,10	Report cards	Ratings, subject to vouchers	Strong	Yes	10	1988	5
Georgia	3,4,5,8,11	School reports	None	None	Yes	11	1995	2
Hawaii	3,5,8,10	None	None	None	No			1
Idaho	ITBS, 3-8	None	None	None	No			1
Illinois	3,4,5,8,10	Academic improvement	Watch lists, warnings, intervention	Moderate	No			2.5
Indiana	3,6,8,10	Performance Assessment	Accreditation	Moderate	Yes	10	1999	3
Iowa	None	None	None	None	No			0
Kansas	3,4,5,8,10	School reports	Accreditation	Weak	No			1
Kentucky	4,5,7,8,10-12	Meeting state improvement goals	Monetary rewards, intervention	Strong	No			4
Louisiana	LEAP,4,8	Report cards, growth targets	Intervention	Moderate	Yes	10	1991	3
Maine	4,8,11	None	None	None	No			1
Maryland	3,5,8	School Performance Index	Monetary rewards, reconstitution	Strong	Yes	10,11,12	2001	4
Massachusetts	4,8,10	Students only	Student promotions	Implicit only	Yes	10	2003	2
Michigan	4,5,7,8	School rating	Accreditation	Weak	No			1
Minnesota	3,5,8,10	School reports	None	None	Yes	8,10		2
Mississippi	2-8	Only districts accountable, based on test scores	Public recognition, loss of accreditation	Moderate to strong at district level	Yes	11	1994	3
Missouri	3-11	School can be deemed academically deficient	Possible audit	Weak	No			1.5
Montana	4,8,11	None	None	None	No			1
Nebraska	None	None	None	None	No			0

Nevada	4,8,10	School reports	None	Weak	Yes	11	1999	1.5
New Hampshire	3,6,10	None	None	None	No			1
New Jersey	4,5,11	Mostly district level, 75% pass rate	Audits, possible state takeover	Strong	Yes	11		5
New Mexico	1-9	School ratings and district rankings	Some money rewards, probation	Moderate to strong	Yes	10	1990	4
New York	4,5,8,11	State review of school performance	Freeze on pupil registration	Strong	Yes	10	1998	5
North Carolina	3-8	School ratings	Money rewards, intervention	Strong	Yes	9	1994	5
North Dakota	4,8,12	Improve student learning	Accreditation	Weak	No			1
Ohio	4,6,9,12	Report cards, but mainly district level	Money for schools, sanctions for districts	Moderate	Yes	9		3
Oklahoma	5,8	Reports to state	Accreditation	Weak	No			1
Oregon	3,5,8,10	School performance ratings	Write school improvement plans	Weak to moderate	Yes	10	1991	2.5
Pennsylvania	5,6,8,9,11	High schools have ratings	Money for HS improvement	Weak	No			1
Rhode Island	3,4,7,8,10	Yearly progress on test results	Reconstitution	Weak implementation	No			1
South Carolina	3-8, 10	District only	District defined as impaired	Moderate	Yes	10	1990	3
South Dakota	2,4,5,8,9,11	Test reports	None	None	No			1
Tennessee	3-8, 9	Test reports	Accreditation	Weak	Yes	9		1.5
Texas	3-8,10	Report cards	School ratings, interventions	Strong	Yes	10	1991	5
Utah	3,5,8,11	None	Accreditation	Weak	No	10	2007	1
Vermont	2,4,8,10	School reports	Identify schools for assistance	Weak	No			1
Virginia	3,4,5,6,8,9	Report tests, other data	Standards of Accreditation	Weak to moderate	No			2
Washington	2-10	School reports	Accreditation	Weak	No	10	2008	1
West Virginia	3-8	Performance audits	Intervention	Strong	No			3.5
Wisconsin	3,4,8,10	Continuous Progress Indicator	Ratings of schools	Weak to moderate	No	11	2004	2
Wyoming	4,8,11	Only district	Accreditation	Weak	No		2001	1

**Table 5. Logit Estimates of a Recent College Graduate Becoming a Teacher: Odds Ratios  
& Z-Statistics**

Sample Size = 4907 Pseudo R<sup>2</sup> = 0.0510

**Table 6a. Logit Estimates of Differential Change in New Teachers with Higher Test Scores**

Variable	Odds Ratio	Z-Statistic
Male	.33	-7.64
Asian	.35	-3.39
Black	.56	-2.88
Hispanic	1.16	.56
Urban	1.03	.14
Rural	1.41	2.55
State Average Beginning Teacher Salary	1.00	.47
Standardized Tenth Grade Test Score	.89	1.66
Selective Undergraduate Institution	.73	-2.27
State Accountability Index Score	1.00	.06
NELS88	.77	.19
NELS88 * Accountability Index	<b>1.08</b>	<b>.08</b>

**in States Implementing Stronger Accountability: Odds Ratios & Z-Statistics**

Variable	Odds Ratio	Z-Statistic
Male	0.33	-7.59
Asian	0.36	-3.35
Black	0.56	-2.89
Hispanic	1.15	0.56
State Average Beginning Teacher Salary	1.01	0.48
Selective Undergraduate Institution	0.73	-2.24
Standardized Tenth Grade Test Score	0.90	-0.52
NELS88	0.74	-0.81
State Accountability Index Score	1.01	0.08
NELS88 * Accountability Index	1.10	0.7
Test Score * Accountability Index	1.00	-0.06
NELS88 * Test Score	1.07	0.21
Test Score * NELS88 * Accountability Index	<b>0.97</b>	<b>-0.22</b>

**Table 6b. Logit Estimates of Differential Change in New Teachers with From Selective Universities in States Implementing Stronger Accountability: Odds Ratios (Z-Statistics)**

Variable	Odds Ratio	Z-Statistic
Male	0.33	-7.63
Asian	0.36	-3.42
Black	0.56	-2.89
Hispanic	1.16	0.57

State Average Beginning Teacher Salary	1.01	0.31
Selective Undergraduate Institution	0.94	-0.12
Standardized Tenth Grade Test Score	0.88	-1.74
NELS88	0.79	-0.76
State Accountability Index Score	0.99	-0.14
NELS88 * Accountability Index	1.10	1.13
Selective Institution * Accountability Index	1.06	0.37
NELS88 * Selective Institution	0.71	-0.46
Selective Institution * NELS88 * Accountability Index	<b>0.94</b>	<b>-0.29</b>

**Table 6c. Logit Estimates of Differential Change in New Male Teachers in States Implementing Stronger Accountability: Odds Ratios (Z-Statistics)**

Variable	Odds Ratio	Z-Statistic
Male	0.16	-5.01
Asian	0.35	-3.41
Black	0.55	-3.12
Hispanic	1.16	0.54
State Average Beginning Teacher Salary	1.01	0.41
Selective Undergraduate Institution	0.73	-2.27
Standardized Tenth Grade Test Score	0.88	-1.72
NELS88	0.62	-1.81
State Accountability Index Score	0.93	-1.38
NELS88 * Accountability Index	1.18	2.2
Male * Accountability Index	1.33	2.28
NELS88 * Male	2.23	1.47
Male * NELS88 * Accountability Index	<b>0.73</b>	<b>-2.12</b>

**Table 6d. Logit Estimates of Differential Change in New Non-White Teachers in States Implementing Stronger Accountability: Odds Ratios (Z-Statistics)**

Variable	Odds Ratio	Z-Statistic
Male	0.33	-7.9
Non-White	0.87	-0.2
State Average Beginning Teacher Salary	1.01	0.48
Selective Undergraduate Institution	0.73	-2.34
Standardized Tenth Grade Test Score	0.88	-1.72
NELS88	0.78	-0.99
State Accountability Index Score	1.00	-0.07
NELS88 * Accountability Index	1.09	1.21
Non-White * Accountability Index	1.03	0.17
NELS88 * Non-White	0.48	-1.13
Non-White * NELS88 * Accountability Index	<b>1.08</b>	<b>0.38</b>

**Table 6e. Logit Estimates of Differential Change in New Teachers Who Majored in Education in States Implementing Stronger Accountability: Odds Ratios (Z-Statistics)**

Variable	Odds Ratio	Z-Statistic
Male	0.46	-5.91
Asian	0.46	-3.2
Black	0.71	-1.63
Hispanic	1.24	0.63
State Average Beginning Teacher Salary	1.02	0.83
Selective Undergraduate Institution	0.97	-0.19
Standardized Tenth Grade Test Score	1.08	0.81
Education Major	15.57	4.51
NELS88	0.69	-0.86
State Accountability Index Score	1.06	0.55
NELS88 * Accountability Index	1.04	0.27
Education Major * Accountability Index	0.91	-0.61
NELS88 * Education Major	1.67	0.78
Education Major * NELS88 * Accountability Index	<b>1.14</b>	<b>0.71</b>

Sample Size = 4907 Pseudo R2 = .2694

**Table 6f. Logit Estimates of Differential Change in New Teachers that Intended to Teach in States Implementing Stronger Accountability: Odds Ratios (Z-Statistics)**

Variable	Odds Ratio	Z-Statistic
Male	0.39	-6.09
Asian	0.35	-3.37
Black	0.68	-1.79
Hispanic	1.25	0.71
State Average Beginning Teacher Salary	1.02	0.82
Selective Undergraduate Institution	0.77	-1.64
Standardized Tenth Grade Test Score	0.97	-0.24
Wanted to Teach in High School and/or College	13.96	5.22
NELS88	1.76	1.64
State Accountability Index Score	1.08	0.99
NELS88 * Accountability Index	0.98	-0.19
Wanted to Teach * Accountability Index	0.97	-0.21
NELS88 * Wanted to Teach	0.20	-2.84
Wanted to Teach * NELS88 * Accountability Index	<b>1.04</b>	<b>0.21</b>

Sample Size = 4608 Pseudo R2 = .1106