ESSAYS IN LABOR ECONOMICS AND FINANCIAL ECONOMICS

By

Yuqing Zhou

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Economics – Doctor of Philosophy

2017

ABSTRACT

ESSAYS IN LABOR ECONOMICS AND FINANCIAL ECONOMICS

By

Yuqing Zhou

In Chapter 1, I revisit the effects of unilateral divorce laws on female labor supply. I use a variety of models to check the robustness of the results and find that the estimated effects on female labor supply are remarkably robust. The estimates I mainly use in this paper suggest that unilateral divorce laws increase female labor force participation rates by roughly 4-5 percentage points, and that these effects strengthen over time. There are also strong, long-term effects on the weeks and hours of work and on participation in full-time work. In addition, this paper compares the dynamic participation responses of married mothers versus married non-mothers, high education versus low education women, young versus old women and white versus black women.

In Chapter 2, I investigate the racial differences in non-cognitive skills. The disparity in cognitive skills between white and black children has been studied extensively. However, the racial gap in non-cognitive skills has attracted much less attention. In this paper, I use ECLS-K to show that there are significant differences in non-cognitive skills between white and black students, even after controlling for a large set of background variables. I show that the gap is not exaggerated by teacher test-score bias against black students. Because subjective bias, on the contrary, attenuates the racial gap, I keep adding teachers' characteristics and use school fixed effects to measure the gaps of non-cognitive skills less biased. Strikingly, the more accurate gaps are much larger than the original gaps. The huge gap between whites and blacks in non-cognitive skills may help to explain the racial gap in wages, probability of arrest, teenage pregnancy, and other important outcomes in the United States.

In Chapter 3, we show investors are surprised by the lack of a management forecast. Conditional on the same quarterly earnings news, cumulative stock returns prior to the earnings announcement date are higher in firm quarters with no management forecast than those in firm quarters with management forecasts. However, the difference in cumulative stock returns declines significantly after the earnings announcement. One possible explanation for the results is limited strategic thinking: investors underestimate the relation between management's strategic incentive to withhold information and the private information they have, which leads to the initial underreaction and the subsequent reversal. We contribute to the literature by showing that investors are constrained in understanding managers' strategic disclosure decisions, even when such decisions are salient to investors and repeated over time, and analysts and sophisticated institutional investors exist.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

The Effects of Divorce Laws on Labor Supply: A Reconsideration and New Results

1.1. Introduction

Female labor supply increased dramatically following World War II as measured by labor force participation (LFP) and working hours. This phenomenon is important for the women themselves, their families and society at large. But what has caused this? What are the effects? Answering these questions has been the subject of a voluminous literature.

During the 1970's there were significant changes in divorce laws in US. The new unilateral divorce laws allow people to end a marriage without the consent of their spouse. In addition, many states removed fault as a consideration in property division. These law changes might affect the bargaining power within the household and change people's expected value of a marriage. Therefore, they might also change women's returns to housework relative to other options.

The above-mentioned two phenomena prompted researchers to ask whether divorce laws affect female labor supply. If they do, how large and how long are the effects? Theoretically, as Gray (1998) discussed in his paper, according to neoclassical model the household head pools all family resources when determining the family's optimal behavior, therefore this model predicts that unilateral divorce should only affect the time allocations of wives through its effect on divorce probabilities. However, cooperative bargaining models demonstrate that household behavior is sensitive to the earning power within the family. Bargaining model assumes that husbands and wives cooperatively bargain over possible decisions to be made by the household. Changes to divorce laws alter wives leverage in the bargaining process, which in turn is likely to alter married women's demand for leisure and the labor force participation behavior.

Early empirical studies like Peters (1986) and Parkman (1992) only used single year cross-sectional data and showed that unilateral divorce laws are positively correlated with female labor force participation rates. However, Gray (1998) used a difference-in-difference (DID) approach to show these results are problematic and found that (p. 629) “unilateral divorce has no significant impact on married women’s labor-force participation unless the underlying marital-property laws in each state are considered.” In a subsequent article, Stevenson (2008) carefully examined Gray’s argument and showed that these results are not robust to alternative specifications and controls. She concluded that unilateral divorce laws increased female labor force participation, regardless of the pre-existing laws regarding property division. Genadek, Stock and Stoddard (2007) tried to distinguish the responses for married women with and without children. According to their results, new divorce laws increased the labor supply of married mothers relative to married non-mothers.

Though there are several papers investigating the effects of unilateral divorce laws on female labor supply, it is very important to check the robustness of the results by using different estimation methods and functional forms. For example, Wolfers (2006) examined the effects of the divorce law changes on a different, but related outcome: the rate of divorce itself. He investigated the dynamic effects with state-specific time trends. However, Lee and Solon (2011) explored the sensitivity of Wolfers (2006) results to variation in estimation method and functional form. They found that the results are extremely fragile. Lee and Solon then concluded that the impact of unilateral divorce laws remains unclear. Moreover, they suggested that identification in difference-in-differences research becomes weaker in the presence of dynamics, casting doubt on all the estimates of the effects of unilateral divorce laws found by previous research. Because previous studies on female labor supply used the same identification strategy as that in Wolfers (2006) paper, it is necessary to check the results on labor supply and try to assess whether they are also fragile and whether we can successfully measure the effects and what the real effects are.

In summary, my paper has several contributions. First, I use a variety of models to estimates the effects of unilateral divorce laws on female labor supply. I get significant results across all of these models, which implies that the effects on female labor supply are very strong. Specifically, I estimate dynamic effects and robustness of the evidence to the presence of state-specific time trends, which have not been used in previous research on labor supply. I also show that the estimates are quite robust to other estimation methods and functional forms. Second, other than simply considering labor force participation rates, I also estimate the effects on weeks worked last year, usual hours worked per week last year and LFP for full time job that have not been checked before. Moreover, I present sub-sample results that compare the dynamic participation responses of married mothers versus married non-mothers, high education versus low education women, young versus old women and white versus black women.

1.2. The Robustness of Estimates of Divorce Laws on Labor Supply

Previous studies on the effects of unilateral divorce laws on labor supply included state fixed effects to control for unobserved factors varying across states but unchanging within a state over time. Year fixed effects were also included to control for evolving unobserved national factors.

However, there are some other factors that may influence female labor supply differently across states over time. Therefore, we need to check the state-specific time trends that have not been used in previous research. As noted by Friedberg (1998), if we overlook these factors, the estimates will be biased if the divorce reform is endogenous. This means that there may be unobserved attributes that are correlated with the law changes across states and do not change at a national level uniformly (which can be picked up by the year fixed effects). Therefore, we need to check the results when adding state-specific time trends in the regression to test whether these factors do matter. All of the previous research do not add state-specific time trends in their models.

There is another important issue we need to consider. The impact of changes in divorce laws may not be immediate and constant, as individuals may learn about the new policy and then adjust their behavior gradually. Therefore, the state-specific trends may pick up the effects of a policy and not just preexisting trends. In order to solve this problem, we need to add variables that model the dynamic response of divorce explicitly. These variables should identify the entire response function allowing the estimated state-specific time trends to identify preexisting trends.[1] Stevenson (2008) examined the dynamic response of female labor force participation in Table A.5 in her paper, but she did not check the results by adding state-specific time trends at the same time. Therefore, we are still not sure whether the models measure the accurate effects if the trends are different in each state.[2] In order to ensure this, I use specifications that include state-

[1] If the state-specific trends are not linear, adding the state-specific trends and model the dynamic response of divorce would still be not enough to perfectly solve the problem.

[2] Stevenson included all women aged 14 years or greater in her sample. Since the divorce laws change may have little effects on women who are too young (younger than 18), using these observations may attenuate the real effects on young and middle aged adult women. I also checked the effects on women who are older than 50 and find that the divorce law change does not affect them. Therefore, in this paper I use the sample that only includes women between the age of 18 and 49.

specific time trends while also allowing for dynamic responses.[3] Specifically, I focus on the following model:

$$
\begin{aligned}
\text{Labor Force Participation}_{s,t} &= \sum_{k\geq 1}\beta_k \text{Divorce Law has been in effect for } k \text{ periods}_{s,t} \\
&+ \sum_a \beta_a \text{Age}_{s,t} + \sum_r \beta_r \text{Race}_{s,t} + \sum_e \beta_e \text{Education}_{s,t} \\
&+ \sum_s \beta_s \text{State fixed effects}_s + \sum_t \beta_t \text{Time fixed effects}_t \\
&+ \sum_s \beta_{st} \text{State}_s \times \text{Time}_t + \varepsilon_{s,t}
\end{aligned}
\tag{1.1}
$$

In equation 1.1, the dependent variable Labor Force Participation is women's state-level LFP rates in state s in year t. I follow Wolfers (2006) in adding variables meant to model the dynamic response of divorce laws change. These variables are dummy variables for one and two years before the new legal regime, first two years of the new legal regime, for three and four years, for five and six years, and so on. The Age, Race, and Education variables indicate the share of each age, race, and education group, respectively, in each state and year. I control for state fixed effects, time fixed effects and state-specific time trends. In some specifications, I also control for quadratic time trends, and the results are very similar to those from regressions just with time fixed effects.

The data used in this paper comes from Current Population Survey (CPS), March Annual Demographic Files from 1977 to 2012.[4] I restrict the sample to married, spouse present women

[3] In all cases, I also estimate models without state-specific time trends, which yield similar results to models that include such trends. This similarity implies that the factors that influence female labor supply differently across states over time are not correlated with changes in divorce laws.

[4] I do not use CPS data before 1977 because most states were grouped together from 1968 to 1976. In previous research, Parkman (1992) uses 1979 CPS data. Gray (1998) uses three different data. One is 1960, 1970 and 1980 Census data, one is 1968 and 1979 CPS data, another is 1970 and 1980 PSID data. Genadek et al (2007) use 1960 - 1990 Census data. Only Stevenson (2008) uses similar data set that is 1968-1995 CPS data in Table 5 in her paper. In addition, she also uses 1970 and 1980 Census in her paper.

between the age of 18 and 49.[5] Table A.1 shows some basic information about demographic and labor force participation for women in this sample. Specifically, women are on average 36 years old and 88% of them are whites. Half of them attend college for at least 1 year. The number of children in the household is on average 1.68. In this sample, 68% of women are in the labor force and around half of them have full time jobs. I also present the basic information separately for the sample of women before and after the divorce laws change. Before the law change women earn less than women do after law change. This is reasonable since personal income has increased gradually in United States over the past thirty years.

Given the divorce law variation is at the state level, I aggregate all of my data to the state-year level.[6] The state-year level data is constructed from the sample I describe above from the 1977-2012 CPS. I construct the labor force participation rates and share of the observations for each age group, race group and education group by state and year.[7] The unilateral divorce laws specification used in this paper is based on Gruber (2004).[8]

Table A.2 presents estimates of equation (1) without state-specific time trends and quadratic time trends while Table A.3 presents the estimates with state-specific time trends. Column (1) in both of these tables show the results from a basic specification: weighted least squares (WLS), weighting each state and year observation by the state's population. The coefficients on the dynamic responses in column (1), Table A.2 imply that women's labor force participation rates do not significantly increase before the divorce law change and even the first two years after the

---

[5] It is possible that the unilateral divorce law changes may have effects on selection into marriage. For instance, Rasul (2004) showed that the marriage rate declined by about 3 to 4 percent following the adoption of unilateral divorce laws change. In addition, Gray (1988) shows that after taking into account the selection into marriage, the results on divorce rates are similar to those when just using married samples.

[6] Since the unilateral divorce laws change in state level, using individual level data is the same with using aggregate level data and controlling for the average value of each background variables.

[7] When constructing these variables in the regression, I use the CPS sampling weights.

[8] In Table 1 in Gruber (2004), he documents the availability of unilateral divorce in each state from 1910 to the present based on Friedberg (1998) and a careful state-by-state review of the actual divorce laws.

law change. However, after 3 or more years of the law change, women's labor force participation rates increase more than 5 percentages, an effect that is roughly constant across all time periods.

More strikingly, after adding state-specific time trends, we can find from column (1), A.3 that the effects on female labor force participation rates become even stronger in both the short-run as well as the long-run[9]. Specifically, the coefficient of 1-2 years before the divorce law is 0.061, which means that women's labor force participation rates increase a lot even before the law change. As I have controlled state-specific time trends in Table A.3, the pre-law-change effects may not be endogenous trend if the state-specific time trends are linear. This could be the policy lead effects. People may change their LFP decisions even before the unilateral divorce laws has been changed, if they anticipate this law would be passed in a few years.[10] The coefficient of divorce law has been in effect for 0-2 years is 0.074, and this effect increases to nearly 10 percentage points 3-4 or more years after the laws change. This pattern suggests that some people react before the law change but many other people need to take some time to adjust their labor supply based on the new divorce laws. As I have discussed before, the estimates from regressions with state-specific time trends may be less biased than the estimates in Table A.2 if state-specific time trends are linear. If some factors influence female labor supply differently across states over time, excluding state-specific time trends will induce biased estimates. The results of the basic specification imply that divorce laws have robustly positive effects on female labor force participation, even in the long run.[11] However, we need to examine several important issues before reaching a definitive conclusion.

[9] I test the equality of the coefficients in column (1), Table A.2 and those in column (1), Table A.3. They are marginally significantly different in 10% level.

[10] Since I could not check the pre-trends due to lack of data, I could not completely rule out the possibility that the significant pre-trends in Table A.2 and A. 3 are caused by misspecification.

[11] I also use regression with both state-specific time trends and quadratic time trends. The results do not change a lot when including higher-order state-specific time trends. They are quite similar with those in Table A.3.

The basic specification in column (1) is based on the assumption that the error term in each weighted regression is homoskedastic and serially uncorrelated. However, the residuals may have strong serial correlation, and ignoring these autocorrelations could lead to bias in the estimation of standard errors. I use Stata's cluster option to implement Arellano's (1987) method of correcting standard error estimates for both serial correlation and heteroskedasticity. The results in column (3) in Table 2 and 3 cluster at the state level. The standard errors in these specifications are similar to those in column (1), which means that autocorrelation is not a big concern in this setting.

Secondly, only if the error terms for individuals within the state are homoskedastic and independent of each other, weighting by population leads to efficient coefficient estimation. However, error terms are not always homoskedastic. Based on Dickens (1990)'s conclusion, it is likely that individual error terms are positively correlated. Then OLS applied to aggregate data may be more efficient than WLS.[12] To check this issue, I also use OLS and OLS cluster to run regression (1), and the results are shown in columns (2) and (4) of Tables 2 and 3. Now all the coefficients of dynamics response have the same pattern with those in column (1) and column (3). WLS and OLS producing similar results is a finding that is consistent with the models being correctly specified for measuring the effects on female labor supply in the light of DuMouchel and Ducan's (1983) paper. They emphasized that if the estimation model is correctly specified, both WLS and OLS are consistent.

[12] If the individual-level error term is like: $v_{ij} = c_i + u_{ij}$, where c_i is unobserved group-level factors in common. Then the variance of the group-average error term v_i is: $Var(v_i) = \sigma_c^2 + (\sigma_u^2 / J_i)$. If σ_c^2 is substantial and the sample size J_i is sufficiently large, the variance of the group-average error term may be dominated by σ_c^2, which is homoskedastic. In this case, OLS is better than WLS.

Regression models discussed above are all linear. In order to analyze how the results are affected by alternative specifications, I also try nonlinear functional forms for the dependent variable. Specifically, since LFP rate is a fractional variable and always positive, I use the model for the logarithm and the logit[13] of the labor force participation rate. The coefficients are shown in the last two columns in Tables 2 and 3.[14] All the effects on LFP rate are still positive and statistically significant. This is also consistent with the results from other specifications.

Since Stevenson (2008) used 1968-1995 CPS data, which is similar with the data I use in this paper, in Table 5 in her paper, it is important to compare my results with hers. In column (1) and column (2) of Table 4, I copy the results in column (3) and (4) of Table 5 in Stevenson's paper. The sample is restricted to 14 years or greater married women. In column (3) I use the same data, 1968-1995 CPS, to replicate the results in column (1). I only add state fixed effects and year fixed effects, which are the same with the controls used in column (1). The estimation coefficients in these two columns are insignificant and very similar with each other. In column (4), I add control variables that include share of each age group in each state and year, share of each race group and education group in each state and year, state fixed effects and time fixed effects, which are the same with those in Table 2 and 3 in my paper.[15] After adding control variables, the coefficients become larger and strongly significant. Both of the results in column (2) and (4) with control variables are significant and the estimates in column (4) are even bigger. In column (5) I add state-specific time trend. Comparing to column (4) the coefficients become

[13] The dependent variable is log[p/(1-p)] where p is the labor force participation rate.
[14] The coefficients in these two columns are marginal effects and the standard errors come from Bootstrapping. So they are comparable with results in other columns.
[15] According to Stevenson's paper, column (2) also add control variables that include the maximum AFDC rate for a family of four; existence of the AFDC unemployed parent and food stamp programs; the natural log of state personal income per capita, the unemployment rate; age composition variables indicating the share of states' populations aged 14-19; and then ten-year cohorts beginning with age 20 up to a variable for 90+; the Donohue and Levitt Effective access; and the share of the state's population that is black, white and other.

smaller and the standard errors become larger. As a result, the estimates are not significant. In column (6) I restrict the sample to married women between the age of 18 and 49. In column (7) I extend the sample from 1968-1995 CPS data to 1968-2012 CPS data. Both of these two changes make the estimation coefficients bigger. As I discuss above, since the divorce laws change may have little effects on women who are too young or too old, using these observations may attenuate the real effects on young and middle aged adult women.

In summary, based on what I showed in Table 2 and 3, the results are robust to variation in estimation methods and functional forms.[16] According to the results in Lee and Solon (2011)'s paper, the effects on divorce rates are still unclear, however that is not the case here. Based on the results in Table 3 that comes from probably preferable specifications, changes in divorce laws have strong effects on female labor force participation both in the short run as well as in the long run.

1.3. Other results

1.3.1. Working Week, Working Hour and the Full Time Job

Most of the previous research only focuses on the effects on extensive margin of labor supply, i.e. labor force participation[17]. However, people may change their types of jobs or working hour even if they still choose to stay in the job market since completely exiting the labor market is a very big and sharp decision. Therefore, in this paper I try to find some new results and investigate

[16] I also use models that include both linear and quadratic time trends. The results from these models, which I do not put in the paper, are nearly identical to the results from models that only include linear trends.
[17] In Genadek, Stock and Stoddard (2007)'s paper, they use OLS to get the effects on weeks worked last year and hours worked last week. However, they do not check dynamic effects and also do not add state specific time trends.

the effects on a special case of LFP, i.e. full time job. Furthermore, naturally we would like to know the effects on the weeks and hours worked as well. This part of the paper focuses on a model that is similar to equation (1). The dependent variable represents the average weeks worked last year, usual hours worked per week last year or LFP for full time job for married, spouse present women between the age of 18 and 49 in state s in year t. The dummy variables for dynamic response here are different from those in equation (1). Since the coefficients for these dummy variables are quite similar in Table 2 and 3, it is better to use a more concise model. In this section, the dynamic response variables I use are dummy variables for one to two years before the new legal regime, for first two years of the new legal regime, for three to four years and for 5 and more years. The definitions of independent variables Age, Race, Education and other variables are the same with those in equation (1).

Table 5 reports the OLS cluster estimates[18], using weeks worked last year unconditional on participating in the labor force, weeks worked last year conditional on participating in the labor force, usual hours worked per week last year unconditional on participating in the labor force, usual hours worked per week last year conditional on participating in the labor force and LFP rates for full time job as dependent variables.[19] The results presented in column (1) of Table 5 indicate that, unconditional on participating in the labor force, women work several more weeks per year even before the laws change and the effects increase gradually. Specifically, one to two years before law change, women work around 1.48 more weeks. Within two years after the divorce laws reform, women work nearly 2.4 more weeks. After that, the weeks they work

[18] According to the results in Table A.2 and A.3, we can find that the standard errors from WLS are bigger than those from OLS. So Dickens' argument is right in this case; OLS is more efficient. Therefore, in the rest of the paper I use OLS instead of WLS.

[19] I also check the sensitivity of results for *weeks worked last year* and for *usual hours worked per week last year*. Based on these results, I can find that the results of for *weeks worked last year* and for *usual hours worked per week last year* also very robust.

increase. In the long run they work around 3.3 more weeks. This is a very large effect. As noted before, the reason the effects within two years are smaller is probably because people need some time to change the expectation of their marriage and then adjust their labor supply behavior. In column (3), we can find that, unconditional on participating in the labor force, women also increase their working hours per week after the law change. They work roughly 1.8 more hours per week right after the laws changed and keep increasing hours worked gradually. In the long run, women in reformed states work around 2.4 hours more per week than their counterparts in other states. Based on the results in column (2) and (4), it is clear that, conditional on participating in the labor force, there is not significant change for either weeks worked last year or hours worked per week last year. Lastly, in column (5) the similar pattern could be seen. In the short run, the full time job participation rate increases 2.7 percentage points. After 5 or more years of the law change, the participation rate increases more, i.e. up to 4 percentage points. However, the effects on LFP of full time job are not significant. It is also possible that the results I find above are affected by selection. These results may imply that women who are induced to enter the labor force prefer to work longer than the average level of weeks women worked last year and hours worked per week last year unconditional on participating in the labor force.

1.4. Sub-sample Results

1.4.1. Married Mother and Non-Mother

In Genadek, Stock and Stoddard (2007)'s paper, they tried to separate the effects on married women who have no child, who have younger children and who have older children. Using

IPUMS Census data, they found married mothers are more likely to increase participation in the labor force as well as the weeks worked. On the contrary, married non-mothers reduced LFP after the reform. In this section, I use CPS to investigate the heterogeneous responses of these three groups of married women. The specification includes state-specific time trends while also allowing for dynamic responses. In order to estimate the aggregate specification (2), I construct averages of all variables by state, year, and by whether there are children under age 6 or aged 6-18 in the household. The dependent variable represents LFP rates, average weeks worked last year, usual hours worked per week last year or LFP rates for full time job for each group of women. The independent variables Child under6 and Child 6-18 are both dummy variables and they equal to one if in this group women's youngest child is under 6 or 6 to 18 respectively. In addition, the vector X includes other control variables: age square, non-labor income and non-labor income square[20]. The definitions of all other variables are similar with those in section 3.1, except that they are constructed by state and year within the three subgroups.

To compare the results with those in Genadek, Stock and Stoddard (2007)'s paper, I use OLS cluster.[21] According to the results in Table 6, the effects of unilateral divorce laws vary across different types of women. In the short run, the presence of children is associated with a large differential response to divorce laws-reform. The changes in divorce laws have very strong effects on married non-mothers, smaller effect on married mothers with older children, and hardly any significant impact on married mothers with young children in the short run. However, the effects are strong for all groups of women in the long run. In column (1) of Table 6, overall

[20] I add some more control variables here in this part because Genadek, Stock and Stoddard (2007) also added these controls in their regressions. Then the results are more comparable to those in their paper, though adding these control variables has little effects on the coefficients.

[21] I also check the sensitivity of these results by using the specifications I used in Table A.2 and they are still very robust.

we can find that married non-mothers increase their labor force participation a lot, no matter how many years since the unilateral divorce laws have passed. Specifically, the coefficient on 0-2 years later suggests that the probability of LFP for non-mothers in reformed states in two years after the reform is 5.6 percentage point larger than that for non-mothers in states without divorce laws reform. The -0.055 negative coefficient on 0-2 years later* child under6 implies married mothers with children under 6 have nearly zero (0.1 percentage point) net increase of LFP and the standard error of the net increase is 0.020. For married mothers with children of age 6-18, there is a net increase of a 4.9 percentage point probability of LFP, which is also a very large effect and the standard error of the net increase is 0.022. The net increase of LFP of married mothers with children of age 6-18 is significant in 5% level. However, in the long run, the gaps of the effects between married non-mothers, married mothers with older children and married mothers with young children gradually disappear. All of the women in reformed states increase their LFP a lot.

According to the results in column (1), Table 6, unilateral divorce laws have effects on all women, and in the short run the effects are much larger for married non-mothers. The reason why the effects of unilateral divorce laws on married non-mothers are much stronger in the short run is probably because the marriage is much more stable if a couple have already had children. Therefore, divorce laws reform does not change their expectation of marriage as quickly as for other people.

1.4.2. Education, Age and Race

In Table 7 I present sub-sample results by education, age and race. The model used in this section is similar with equation (1) except that the dynamic response variables I use are dummy variables for first two years of the new legal regime, for three to four years and for 5 and more years. The definitions of independent variables Age, Race, Education and other variables are the same as those in equation (1). The dependent variable is LFP rates.

Column (1) and Column (2) report the OLS cluster estimates of the effects on women with high school or lower education and women with college or higher education in. The results indicate that probability of LFP for low education women in reformed states in the two years after the reform is 5.9 percentage points larger than that for their counterparts in states without divorce laws reform. On the contrary, the effects on high education women are not significant. The coefficients are smaller than those in Column (1) and the standard error are larger. All in all, the effects of unilateral divorce laws are stronger and clearer on low education women than on high education women. It is possible that for high education women, their marriages are more stable so that their LFP decisions are not strongly affected by divorce laws change.

Column (3) and Column (4) report the effects of unilateral divorce laws on women that are 35 years old or younger and women that are older than 35 separately. I include all women, not only the married mothers, in calculating LFP rates in columns (3) and (4). It is clear that the estimates of effects on both of these two groups of women are large. However, the standard error of the estimates of the effects on younger women are much larger than those in Column (4). Therefore, the effects on older women are significant but are not on younger women. Maybe middle-aged women generally face a similar condition. Majority of them are likely to be married and their marriage stability is affected by the unilateral divorce laws. As for younger women, some are single, some are just married, some are still in schools. The within sample variation of the effects is larger for younger women and therefore it is hard to get the precise effects of divorce law change on younger women.

Last, Column (5) and Column (6) present the effects on white and black women separately. Unilateral divorce laws have significant and large effects on white women. They have higher LFP rates both in the short run and in the long run. Due to small samples of black women, the estimated coefficients are too imprecise to draw conclusions for this group.

1.5. Summary and Discussion

There are several papers about the effects of unilateral divorce laws on female labor supply. In this paper I expand upon their analyses but testing alternative specifications. In order to accurately measure the results, I first investigate the dynamic effects with state-specific time trends. In addition, I carefully check the sensitivity of the effects on female labor supply with other estimation methods and functional forms. Previous research on the effects of unilateral divorce laws on divorce rates have found extremely fragile results; therefore, the impact of changes in divorce laws on divorce rates remains unclear. In this paper, I find more robust results suggest that there are strong effects on the female labor force participation rate even in the long run. The robustness of effects on female LFP is different from the fragile results of divorce rates in the previous literature. There are also strong and long term effects on weeks worked per year and usual hours worked per week. Moreover, I compare the dynamic response to labor supply among married mothers with that of married non-mothers. The results show that in the long run all married women increase their labor supply, however in the short run married non-mothers have much larger response to unilateral divorce laws change. In addition, I also find that the effects of unilateral divorce laws are stronger on low education women than on high education women, but do not find clear pattern of differences between old and young women or white and black women. Though I find robust and strong results by using different specifications in this paper, people still need to interpret the results in my paper with caution. First, I could not check and account for pre-trends because of the lack of data, therefore the results might be biased if there are pre-trends. Second, the identification is tenuous by using differences in differences when there are complicated dynamics.

As Lee and Solon (2011) discussed in their paper, "the DID research design with unit-specific time trends is essentially a type of regression discontinuity design, with time as the 'running variable'." "When the shift in the dependent variable may vary with the length of time since the policy change, and especially when that complication is accompanied by other differences across states in time trend, the sharpness of the identification strategy suffers." Therefore, when using the same identification strategy by exploiting the divorce laws change, the results of other outcomes may also be sensitive as those of divorce rates. However, on the contrary, the results on female labor supply suggest much more robust effects. After controlling for state-specific time trends, no matter which estimation method and functional form I use, the results show female labor force participation rate increases a lot. Why is there such a big difference between the sensitivity of results on divorce rates and that of female labor supply? It is possible that the unilateral divorce laws have little effects on divorce rates since any decision to divorce is relative to a small portion of people who are around the margin of divorce. On account of the fact that the effects are so small, it is hard to measure them precisely and find any robust answer. On the contrary, unilateral divorce laws may have strong effects on female labor supply through the mechanism of changing the expectation of marriage among all adults, but not through divorce. In other words, since all married women, not just the women around the margin of divorce, need to reconsider their labor force participation decisions, the effects could be very large and easy to measure. As a result, the identification strategy is sharper than that of effects on divorce rates.

Chapter 2

Racial Differences in Non-Cognitive Skills

2.1. Introduction

How the lives of different racial groups in United States are influenced on account of their race is an extremely important issue and also has become a very popular topic in recent decades. Previous research shows that there are large differences among racial groups in the United States in nearly every aspect of life, such as education, family structure, labor market outcomes, incarceration, health, and so on.

Previous research tries to investigate what affects the racial gaps among adults and find that the test score gap is an important factor. Neal and Johnson (1996) and O'Neill (1990) find that most of the observed black-white wage gaps among adults disappears when lower eighth grade test scores among blacks are taken into account. Therefore, eliminating the test score gap in high school may be a critical component of reducing racial wage inequality. In order to find ways to lessen or even eliminate serious racial inequalities, researchers started to pay attention to the emergence of racial gaps among children, especially the gap in cognitive skills. Surprisingly, even though there is no difference between white and black infants at eight months of age, large test score and mental ability gaps have been found in children even as young as two years old Scott and Sinclair (1997); Fryer and Levitt (2013). The evolution of racial gaps in test scores is still uncertain, based on many different results. Some of them show that a large black-white test score gap emerges in early childhood; others show that the gap in kindergarten is small and can be explained by socioeconomic status (SES) but widens sharply in the early years of schooling Jencks and Phillips (1998); Fryer and Levitt (2004, 2006).

A wide range of potential explanations for the racial gaps in test scores have been found. Some argue that test score gaps are affected by differences in genetic make-up Hernstein & Murray, (1984); Jensen (1973, 1998). Socioeconomic status and the effects of poverty are also important factors in explaining racial differences in educational achievement Brooks-Gunn & Duncan (1997); Mayer (1997); Brooks-Gunn et al., (1994, 1995, 2000). In addition, differences in school quality Cook & Evans, (2000), racial bias in testing or teachers' perceptions Delpit, (1995); Ferguson (1998); Rodgers & Spriggs (1996), and differences in culture, socialization, or behavior Cook & Ludwig (1998; Fordham & Ogbu (1986); Fryer (2002); Steele & Aronson, (1998) are also possible explanations. However, it is still puzzling that even after controlling for many background factors, such as birth weight, SES, family structure, other aspects of home environments or even school environments, a substantial black-white test score gap still remains Campbell *et al.* (1966); Burkett *et al.* (1995); Rushton (1995); Fryer and Levitt (2013).

For a long period of time people paid more attention to the importance of cognitive skills. In recent years, the importance of non-cognitive skills gradually has attracted more and more attention. Heckman and Rubinstein (2001) demonstrate the quantitative importance of non-cognitive skills in determining earnings and educational attainment by using evidence from the GED testing program in the United States. Heckman et al. also identify and estimate joint evolution of cognitive and non-cognitive skills over the life cycle of children Heckman *et al.* (2008); Cunha *et al.* (2010). A lot of other research documents that non-cognitive skills play a significant role in determining educational achievement, wages, probability of engaging in criminal activities, and other outcomes Heckman *et. al* (2006); Flossmann *et al.* (2006); Agan

(2011); Segal 2(013). Specifically, Heckman, Stixrud, and Urzua (2006) establish that cognitive and non-cognitive skills are equally important in explaining a variety of labor market outcomes, such as wages, employment, and work experience, and behavioral outcomes, such as teenage pregnancy and marriage, smoking, marijuana use, and participation in illegal activities. All in all, for many dimensions of social performance cognitive and non-cognitive skills are equally important.

Since non-cognitive skills are very important, it is natural for people to further question the heterogeneity of non-cognitive skills in different groups and the factors that underlie these differences. Gender has been identified as an important correlate of non-cognitive skills. Bertrand and Pan (2013) explore the importance of home and school environments in explaining the gender gap in disruptive behavior. They find that non-cognitive returns to parental inputs differ markedly by gender. Other than gender, the differences in non-cognitive skills among racial groups are another important problem that can help us understand non-cognitive skills themselves and various aspects of racial gaps. Though there are large numbers of papers that study racial differences and non-cognitive skills separately, racial gaps in non-cognitive skills attract little attention. In chapter 4 of the book *Steady Gains and Stalled Progress*, Magnuson and Waldfogel use the Early Childhood Longitudinal Study: Kindergarten Cohort (ECLS-K) to assess whether non-cognitive skills can help explain persisting black-white achievement gaps. Based on the results of some basic regressions, they find that the non-cognitive skills may contribute to the racial reading gap at kindergarten entrance. Some papers in other fields also suggest that certain non-cognitive skills may play an implicit role in the development and performance of cognitive skills Diamond (2000); Raver *et al.* (2007); Blair *et al.* (2007). Other than research about the role non-cognitive skills play in explaining the racial gaps in cognitive skills, there is little research about racial gaps in non-cognitive skills themselves. Does a racial gap in non-cognitive skills exist? If it does, how large it is? Are measurements of non-cognitive skills objective or subjective? If they are subjective, how reliable they are? More specifically, does subjectivity affect the racial gap? If it does, what can we do to adjust the measures to get the real racial gap in non-cognitive skills? Since non-cognitive skills play a significant role in explaining nearly every aspect of adult performance, it is very important to find the answers to these questions.

In this paper, I use the non-cognitive skills measures in ECLS-K to derive the racial gap in non-cognitive skills. To my knowledge, this is the first paper that focuses on the racial difference in non-cognitive skills. In the analyses presented here, I provide estimates of the unadjusted black-white gaps in externalizing behavior and approaches to learning. The raw differences are very large and significant. After controlling for many important background covariates, the racial gap in externalizing behaviors in grade 5 is still strongly significant, while the racial gap in approaches to learning in grade 5 becomes insignificant. Since all of the measurements of non-cognitive skills are teacher-reported ratings, it is possible that the existence of racial gaps is the result of teachers' test-score bias. In order to investigate the veracity of this speculation, I use the information of students' item response theory (IRT) scores and teachers' subjective assessments of cognitive skills to infer teachers' behavior. I assume that teachers' biases are consistent, so that the direction of the bias of subjective assessments on cognitive skills is the same as that on non-cognitive skills. Surprisingly, teachers do not favor white students. On the contrary, they tend to give higher grades to black students in grade 3 conditional on background variables. Furthermore, I find that teachers are more generous to students in minority-dominated schools, in public schools, and coming from lower SES families. Based on the assumption I made for the consistency of teachers' behavior, black students also tend to get better assessments of non-cognitive skills. Then I use school fixed effects and add teachers' characteristics to get less statistically biased racial gaps in non-cognitive skills. The findings in this paper demonstrate that the adjusted racial gaps in non-cognitive skills are much larger than the unadjusted gaps. The huge gaps between whites and blacks in non-cognitive skills may help to explain the racial gap in wages, probability of arrest, teenage pregnancy and many other important outcomes in the United States.

2.2. Data

My analysis in this paper is based on data from the ECLS-K. The ECLS-K is a nationally representative longitudinal survey that followed the same children who entered kindergarten in the 1998–1999 school years. Information was collected in the fall and spring of kindergarten and the spring of first, third, fifth, and eighth grades. Over twenty thousand children are included in the sample. Information on the children's home environment, home educational activities, school environment, curriculum, and teacher qualifications was collected in each survey year. In addition, there is detailed information about children's cognitive skills and non-cognitive skills. Specifically, ECLS-K has reading, math, and science IRT scores. Other than the objective assessments, academic achievement was also measured with subjective assessments. Teacher-reported grades measure students' mastery of specific skills in reading, math, and science. The grades are reflected on the continuous 0 to 4 point Academic Rating Scale (ARS), where 0 indicates no understanding of the content and 4 indicates complete mastery. It is important to note that when teachers reported their subjective assessments they did not know any students' IRT test scores. Other than measures of cognitive skills, there are also teacher-reported measures of non-cognitive skills, which include externalizing behavior, approaches to learning, self-control, interpersonal skills, and internalizing problems. Each of the five non-cognitive measures averages answers to several questions that are rated on a scale from 1 (never) to 4 (very often). Specifically, the measurement of externalizing problem behaviors is based on information about the rate of the frequency with which a child acts impulsively, interrupts ongoing activities, fights with other children, gets angry, and argues. The measurement of approaches to learning is based on the information about the rate of a child's attentiveness, task persistence, eagerness to learn, learning independence, flexibility, and organization. The third scale, self-control, includes four items that measure a child's ability to control his or her behavior. These items are respecting the property rights of others, controlling his or her temper, accepting peer ideas for group activities, and responding appropriately to pressure from peers. Interpersonal skills measure a child's ability to interact with others on the basis of five items: forming and maintaining friendships; getting along with people who are different; comforting or helping other children; expressing feelings, ideas, and opinions in positive ways; and showing sensitivity to the feelings of others. Finally, the measure for internalizing problem behaviors includes four items that rate the presence of anxiety, sadness, loneliness, and low self-esteem. The National Center for Education Statistics does not release data on all of these questions individually but instead aggregates the data to the five composite scales mentioned earlier, known as Social Rating Scales. This is widely used survey technique for detecting social and behavioral problems Gresham and Elliott (1990). As Bertrand and Pan (2013) mention, these non-cognitive skill measures are highly reliable.[22]

[22] In their paper, Bertrand and Pan (2013) cite what Neidell and Waldfogel (2011) note—that the ECLS-K non-cognitive measures appear to have relatively high "validity based on test-retest reliability, internal consistency, interrater reliability, and correlations with other, more advanced behavioral constructs (Elliott et al., 1988) and are considered the most comprehensive assessment that can be widely administered in large surveys such as the ECLS-K (Demaray et al., 1995)."

Although ECLS-K allows tracking the racial differences for many different types of non-cognitive skills, I focus on externalizing behavior and approaches to learning in this paper, since they are in the set of non-cognitive skills that map into future educational and labor market outcomes. The results in Bertrand and Pan (2013) show that "externalizing behavior is a crucial determinant of school suspension, which itself has been shown to directly matter for long-term educational outcomes." Therefore, their analysis focus on externalizing behavior. Moreover, Cornwell, Mustard, and Parys (2013) find that approaches to learning have the greatest explanatory power in terms of students' overall performance in school. Therefore, my analysis mainly focuses on the racial differences in these two measurements. In order to not overlook the information contained in other measurements of non-cognitive skills, such as self-control, interpersonal skills, and internalizing problems, I also use principal component analysis to investigate the racial gaps in students' overall non-cognitive skills. I restrict the sample to children who have non-missing data on race and gender. In all of my specifications, I use the eighth grade panel weights provided in ECLS-K to weight.[23] Therefore I also restrict the sample to children who have non-missing and non-zero value for the weights. Although I control for many other background characteristics, such as age, region, urbanicity, family structure, mother's age at first birth, and family SES in the fall of kindergarten,[24] I retain all of the observations that have missing values for these background variables. This is because I would lose a large number

[23] There are many weights that can be chosen in ECLS-K. We can choose the weight based on the variables used. In this paper, I use the eighth grade parent panel weight (C1_7FP0). According to *ECLS-K Combined Eighth Grade and K-8 User's Manual* (2009), this weight is suggested for the analysis of parent interview data from six rounds of data collection (fall–kindergarten, spring–kindergarten, spring–first grade, spring–third grade, spring–fifth grade, and spring–eighth grade), alone or in combination with (a) child-assessment data from any of these six rounds; (b) data from any fall–kindergarten, spring–kindergarten, spring–first grade, spring–third grade, spring–fifth grade, or spring–eighth grade teacher questionnaire (teacher-level or child-level); (c) data from any spring–kindergarten, spring–first grade, spring–third grade, spring–fifth grade, or spring–eighth grade school-administrator questionnaire; or (d) data from any spring–kindergarten, spring–first grade, spring–third grade, or spring–fifth grade school-facilities checklist.
[24] In ECLS-K, there are multiple observations for some social and economic variables, such as SES and family structure. For all specifications in this paper, I include only the measures recorded in the fall kindergarten survey, in order to ensure consistency. Including all the values of these variables from each survey does not change the results.

of observations if I drop all of the observations that have missing values, which would result in a serious sample selection problem.[25][26] Finally, for each grade, I further restrict my sample to children who have valid teacher-reported non-cognitive skills.[27]

Table C.1 presents descriptive statistics for children's background characteristics. Column 1 lists full sample means and standard errors of each of the listed variables. Summary statistics by race for all background variables are presented in columns 2 and 3. Based on the results in Table 1, white and black children grow up in different environments. For historical reasons a large proportion of blacks live in the southern United States. In addition, more blacks live in the center of cities. They also tend to spank children more and have lower SES. Another important issue in US society is that only a small proportion of black children grow up in families with two biological parents. This finding is supported by the ECLS-K. We find that 44.7 percent of black children live with a single mother, whereas only 11.1 percent of white children live with a single mother. In addition, nearly half of black women give birth to children when they are younger than twenty, whereas only 18.2 percent of white women do. Last, we find no systematic differences across whites and blacks in gender ratio, age in the fall of kindergarten, birth weight,[28] and families' warmth index.[29] Since the majority of these home-environment variables and children's background characteristics are different across race, they must play an important

[25] In order to keep observations that have missing values, I change the missing value "." to "-99" and generate a dummy variable for each background control. They are equal to one if the observation has missing values and equal to zero otherwise. Then I add all these dummy variables into regressions.

[26] There are not many missing values for family background control variables. However, the missing data problem is more serious for school environment variables and parents input variables. If I drop all of observations with missing values for school environment variables and parents input variables, the number of observations used by the main regressions will drop from around 6000 to around 2000. After dropping these observations with missing values, the racial gaps in non-cognitive skills I find become even larger.

[27] Since there are problems of attrition and missing values, the observations I use for different grades and different dependent variables (measures of non-cognitive skills) are different.

[28] The birth weights of black children are slightly lower than those of white children, but the difference is not significant.

[29] Please refer to the Appendix that contains the definition of warmth index.

role in explaining racial differences, including the difference in non-cognitive skills. I add them to the analysis to assess how they affect the racial gap and how important they are.

Summary statistics for non-cognitive skills in each grade by race are presented in Table C.2. I standardize all of the teacher-reported ratings to have a mean of zero and a standard deviation of one in the weighted sample, using all of the observations that have non-missing values for these ratings.[30,31] The first two columns show the raw means for whites and blacks separately, and the last column presents the raw mean differences. It is obvious that whites get better scores in all of the five non-cognitive skills measures in all grades. The differences are strongly significant, except for the internalizing problems. On average, externalizing behavior of blacks is 0.37 standard deviations higher than for whites in the fall of kindergarten. The gap widens gradually and grows to around 0.51 standard deviations in fifth grade. Similar patterns can be found for approaches to learning, self-control, and interpersonal skills. The only exception is the internalizing problems. The average scores for black children are 0.16 and 0.20 standard deviations worse than for white children in the first and third grades, respectively. However, we find no systematic differences in internalizing problems in kindergarten and fifth grade. The remaining rows in Table 2 show information about whether children have ever been retained from the fall of kindergarten to eighth grade and school suspension in grade 8. Blacks are 0.13 standard deviations more likely to repeat a grade and 0.22 standard deviations more likely to be

[30] I do not restrict the sample here, to avoid the sample selection issue. Though there are some observations that have valid non-cognitive skills measures but do not have valid information for gender, race, or eighth grade panel weights, it is better to include them when standardizing, to get a more accurate distribution of non-cognitive skills.

[31] I also use the same method to standardize IRT scores and teacher-reported cognitive skills in this paper. Appendix F Figure F.1 presents the distribution of teacher-reported ratings of externalizing behavior in grade 5 before standardization. Appendix F Figure F.2 presents the distribution of math IRT scores, teacher-reported math grades, and teacher-reported ratings of externalizing behavior after standardization. Because there are a lot of missing values for math grades in grade 5, I use all of these three measures in grade 3 instead in Appendix F Figure F.2.

suspended. Overall, blacks perform much worse in each grade than whites in terms of non-cognitive skills.

2.3. The Racial Gap in Non-Cognitive Skills

In order to understand the racial gap in non-cognitive skills and the forces that drive this gap, I use regressions to estimate the racial gap. My regression approach involves the weighted least-squares estimation of equations, that have the following form:

$$\text{Non-Cognitive Skills}_i = \sum_r \beta_r \text{Race}_i + \Gamma X_i + \varepsilon_i, \tag{2.1}$$

where i indexes children and r stands for race. A full set of race dummies is included in all of the specifications, with white as the omitted group. The vector X captures a wide range of control variables, which varies across columns in Table 3, and ε is an error term. For all specifications, the estimation is done by using weights corresponding to the eighth grade panel weights.

In Table C.3, columns (1) to (3) show the racial gap of externalizing behavior in grade 5 and columns (4) to (6) present the racial gap of approaches to learning in grade 5. The last three columns show the racial gap of suspension in grade 8. The first, fourth, and seventh columns of Table 3 present the differences in means, including the race, female dummy variables, age at assessment at fall kindergarten, and age-squared. Gender and age variables show the physical differences among students and are not affected by their performance in school. The results are nearly the same with the raw teacher-reported non-cognitive skills gaps in Table C.2. Other than black-white gaps in non-cognitive skills, it is also interesting to find that Asian children's performance are much better than those of white children. They have much less externalizing behavior, much higher grades on approaches to learning and fewer suspensions in grade 8. In addition, Hispanic children do as well as white children except for approaches to learning, for which they have lower scores. The second specification shown in columns (2), (5), and (8) add background controls, which include family structure, mother's age at first birth, family SES, birth weight, and geographic location variables, which include dummies for region and urbanicity.[32] All of these variables are related to home environment, which has lots of variation among different racial groups. Controlling for these variables substantially reduces the estimated black-white gaps in non-cognitive skills. The black-white gap in externalizing behavior in fifth grade falls to about one half of that in column (1); the gap in suspension in grade 8 also falls from 0.22 to 0.12. More strikingly, after controlling for family backgrounds, the racial gap in approaches to learning in grade 5 becomes insignificant. These unadjusted gaps show that a large part of the racial differences in non-cognitive skills between blacks and whites can be explained by the variation in home environment. However, the racial gaps still remain, and the remaining gaps cannot be explained by all of these background and geographical variables. What's more, home environment has no explanatory power for why Asian children have better non-cognitive skills than white children. After controlling for home environment variables, the gaps between whites and Asians remain the same. In order to determine what other factors affect the racial gaps in non-cognitive skills, I add school environments[33] in columns (3), (6), and (9). We find that

[32] The components used in the SES measure in ECLS-K are parental education, parental occupational status, and household income.

[33] Controls for school environments include average age at kindergarten entry, kindergarten type, emphasis on reading and math, emphasis on homework, emphasis on achievement/behavior/cooperation/following directions, time spent on physical education, time spent on recess, whether the school has a formal retention policy, overall kindergarten environment index (emphasis on reading, homework, retention policy), kindergarten peers, teachers'

controlling for all these variables has few effects on the racial gaps, and the gaps, no matter white-black gaps or Asian-white gaps, are nearly the same as those in columns (2), (5), and (8). Though I do not present it in Table C.3, I also add into the regressions parents' time inputs from kindergarten to grade 5[34] and two other inputs, including parental warmth index and whether the child was spanked in the last week. Just like school environments, parents' time inputs have almost no impact on the results. From Table 3 we can determine that, generally speaking, home environment plays an important role in forming the racial differences in non-cognitive skills. However, the racial gaps cannot be fully explained by it. Moreover, though people may think that the qualities of the schools that different racial groups attend are different, they actually have few effects. This is also true for parental time inputs.

The evolution of these non-cognitive skills' gaps is also very important. Table C.4 presents a series of estimates of the racial gap in non-cognitive skills from fall kindergarten to fifth grade. Since the racial gaps are not sensitive to parents' time inputs, I control for race, female dummy variables, age at assessment at Fall kindergarten and age-squared, background controls, geographic location variables and school environment[35] for all specifications in this table. In other words, I use the same specification as the one used in columns (3), (6), and (9) in Table C.3. The first five columns show the evolution of racial gaps in externalizing behavior. At the time that children are just entering kindergarten, the gap already exists, and the externalizing behavior of black children is 0.18 standard deviations higher than that of their white counterparts. This gap becomes 0.36 in grade 3. The result in column (5) shows that the gap in the fifth grade is still

gender.

[34] Controls for parents' time inputs include reading to child, telling stories, singing songs, helping child create art, helping children do chores, playing games, teaching nature or science, building something with child, engaging in sports, visiting the library, going to a concert, visiting a museum, visiting a zoo, attending a sporting event, helping with homework, helping children practice number.

[35] Although the racial gaps are also not sensitive to school environments, I still add them in the regressions. This is because we are always concerned that blacks and whites attend schools of different quality.

significant.[36] The last five columns in Table C.4 present the racial gaps in approaches to learning. There are some differences across grades, but this may be due to scaling.

2.4. The Factors of Teachers' Subjective Bias and Their Effects on Racial Gaps in Non-Cognitive Skills

2.4.1. Does test-score bias play a role in explaining the racial gaps in non-cognitive skills?

Based on the results in section 2.3, I find that the raw differences in non-cognitive skills between black and white children are very large. Moreover, the racial gap in externalizing behavior still exists and is strongly significant after controlling for a large set of background variables. The estimates of the racial gap in approaches to learning are on the margin of being significant. However, people may suspect that this gap is caused by teachers' test-score bias toward black students. Since the measures of non-cognitive skills are subjective and there are not any other objective measures of non-cognitive skills, it is difficult to rule out this concern directly.

To deal with this issue, I use information on cognitive skills. Fortunately, there are both objective and subjective measures of cognitive skills in ECLS-K. For each student from kindergarten to grade 5, we know not only their IRT math and reading scores but also the teachers' subjective assessment (ARS) of math and reading grades. Since the ARS measures the same skills as those found on the objective reading and math IRT scores, this gives us a chance to see how teachers assess students. This information can be used to judge the teacher-reported measures of non-cognitive skills. Therefore, I compare students' IRT scores and teachers'

[36] This may be because of the missing value problem. When I further restrict the sample to observations that have valid information for the background characteristics that racial gaps are sensitive to (background controls, parental inputs, and family quality), the racial gap in grade 5 is larger than that in grade 3.

subjective assessments of reading and math.[37] I use the difference between these two measurements as an index for the difference between the objective and subjective assessment of non-cognitive skills. The assumption here is that the direction of the bias of teachers' subjective assessment is the same for cognitive skills and non-cognitive skills. If teachers perform consistently, this assumption should make sense.

Specifically, I calculate the difference between IRT score and teachers' subjective assessment separately for math and reading for each student. Then I take the mean of these two differences and use it as an index of bias of teachers' assessment. It is possible that teachers see something else that the IRT tests cannot observe, therefore the difference between IRT scores and teachers' subjective grades of cognitive skills might contain both of bias of teachers' assessment and measurements of something that tests cannot observe.

Index of Bias=0.5(Teachers' Assessment of Reading - Reading IRT Score)+0.5*(Teacher's Assessment of Math - Math IRT Score).*

For Grade 5, since there are too many missing values for teachers' assessment of math, I just use the information of reading scores to get the index of bias:[38]

Index of Bias (Grade 5)=Teachers' Assessment of Reading - Reading IRT Score

In order to have a general idea about what brings about the difference between students' IRT scores and teachers' subjective assessment, I regress the index of bias on the race and female

[37] At first, I use the regression to get the racial gap in IRT scores and teachers' subjective assessments. The results show that the racial gap of IRT scores is large and significant, especially math scores. However, the racial gap of teachers' assessments of reading and math is much smaller and just marginally significant, even in grade 5. This is evidence that teachers' subjective assessments are biased. Therefore, it is natural to suspect that teachers' subjective assessments are also biased.

[38] Since the difference between IRT score and teachers' subjective assessment for math and reading are always similar, this will not affect the index a lot.

dummy variables. From Table C.5, it is clear that teachers do not discriminate against black students at all. On the contrary, they even give higher grades to black students.

2.4.2. Understanding of subjective bias

From these results I find that the racial gap in non-cognitive skills is unlikely to be caused by teachers' test-score bias. After ruling out the concern about test-score bias, another important question comes up: Is this the real racial gap in non-cognitive skills? According to the coefficients of black dummy variables in Table C.5, teachers treat white and black students differently and are more generous to black students. If my assumption about the consistency of teachers' assessments is right, the racial gaps in non-cognitve skills would be attenuated by teachers' subjective bias.

In order to assess this more clearly, I use the index of bias as the dependent variable and add all of the control variables in equation (1) into the regression to see whether these variables can help explain the teachers' possible subjective reverse advantage. The results are shown in Table 6.[39] It is clear that after controlling for home and school background variables, teachers' possible reverse advantage greatly decreases. Moreover, teachers give higher grades to students who come from families with lower SES and those from public schools. Based on these results, teachers might give more generous assessments to disadvantaged students. Since a much larger portion of black students come from families with lower SES and from public schools, most of the students who received lower grades from their teachers are white students. This partially causes the strongly significant results in Table C.5. Therefore, after controlling for these important

[39] Since there are too many control variables in this specification, I just present some important and significant coefficients in Table A.5.

background variables, the coefficients for the black student dummy variable decrease greatly. Another issue of possible concern is that students' non-cognitive skills may affect teachers' subjective assessments for cognitive skills. In order to check this, in some specifications that I do not show, I also add into the regression the measurements of externalizing behavior and approaches to learning for different grades, when using index of bias for different grades as an independent variable. The results show that adding these controls affects only the coefficients of female dummy variables which decrease to nearly zero. Other important variables discussed earlier are nearly unchanged. This evidence tells us that non-cognitive skills affect teachers' subjective assessments for cognitive skills only through gender.[40]

2.4.3. Unobserved variables affect the subjective bias

Based on results in Table C.6, it is obvious that though adding all of these family and school background variables can partially solve the problem caused by subjective bias, the bias does not disappear. Teachers still give higher grades to black students, especially in grade 3 and grade 5. This implies that the racial gaps in non-cognitive skills shown in Table C.3 and Table C.4 are still attenuated by teachers' subjective bias.

To show the estimation bias more formally and clearly, it is useful to present a simple econometric model. Suppose Y represents the unbiased measures of non-cognitive skills, and ε is the measurement error. Since there is no objective and unbiased measure of non-cognitive skills, I need to use other information. $Y^* = Y + \varepsilon$ are the biased measures of non-cognitive skills

[40] This mechanism is fully discussed by Cornwell, Mustard, and Parys (2013). In their paper, the results show that boys who perform as well as girls on reading, math, and science tests are graded less favorably by their teachers, but this less favorable treatment essentially vanishes when non-cognitive skills are taken into account.

contained in ECLS-K. X_{11} is all of the control variables in equation (2.1) except for the race dummy variables. Then regression (1) can be written as $Y^* = \alpha_1 RaceDummies + X_{11}\beta_1 + \nu$, where ν is the error term. If ε is i.i.d, the estimation of α_1 would be unbiased, and the racial gaps presented in Table C.3 and Table C.4 are accurate.[41] However, as discussed earlier, I find that the subjective measurements errors are not random. From Table C.5 it is obvious that ε is correlated with many home background variables and school environment variables. Since Y^* represents teacher-reported non-cognitive skills that contain a teacher's subjective opinion and judgment, it is reasonable to expect that ε is also correlated with teachers' characteristics. Some of these variables correlated with ε have already been controlled in equation (2.1), but some others have not. I define X_{12} as the variables that cause biases but are not contained in equation (2.1); then $\varepsilon = X_{11}\delta_1 + X_{12}\delta_2 + \eta$, where η is the i.i.d error term. Here X_{12} is also correlated with X_{11}. If I run a regression like equation (2.1): $Y^* = \alpha_1 RaceDummies + X_{11}\beta_1 + \nu$ [42] and define $M_{11} = I - X_{11}(X_{11}'X_{11})^{-1}X_{11}'$, then:

$$\begin{aligned} E(\hat{\alpha}_1) &= E[(R_1'M_{11}R_1)^{-1}R_1'M_{11}Y^*] \\ &= E[(R_1'M_{11}R_1)^{-1}R_1'M_{11}Y] + E[(R_1'M_{11}R_1)^{-1}R_1'M_{11}\varepsilon] \\ &= \alpha_1 + E[(R_1'M_{11}R_1)^{-1}R_1'M_{11}(X_{11}\delta_1 + X_{12}\delta_2 + \eta)] \\ &= \alpha_1 + 0 + E[(R_1'M_{11}R_1)^{-1}R_1'M_{11}X_{12}\delta_2] + 0 \\ &= \alpha_1 + E[(R_1'M_{11}R_1)^{-1}R_1'M_{11}X_{12}\delta_2] \end{aligned} \tag{2.2}$$

[41] If ε is i.i.d, then regression $Y^* = \alpha_1 RaceDummies + X_{11}\beta_1 + \nu$ can be rearranged as $Y = \alpha_1 RaceDummies + X_{11}\beta_1 + \nu + \varepsilon$. It is obvious that the coefficients are the same in both of these regressions.

[42] I use R_1 as the simplified form of all of race dummy variables.

It is obvious that the estimation of α_1 is biased, and the estimation error is $(R_1' M_{11} R_1)^{-1} R_1' M_{11} X_{12} \delta_2$.

In order to show how X_{12} would bias the estimator more intuitively; I present an example to look at the IRT scores and teachers' subjective assessments by racial composition of schools. Since I do not control for the racial composition of schools in equation (2.1), it can be regarded as capturing some effects of X_{12} no matter how racial composition itself affects the teachers' subjective bias or how it captures some unobserved variables in X_{12}. Figure B.1 shows the mean of IRT math scores and teachers' assessments of math in grade 3[43] for all students by racial composition[44]. Students' IRT scores decrease with the increasing of minorities in schools, which corresponds with the well-known phenomenon that, on average, white-dominated schools are higher achieving and whites have higher IRT scores. What is surprising is that teachers' subjective assessments do not match the pattern of IRT scores. The average teachers' grades are even higher in the schools with 50–75 percent of minorities than in those in white-dominated schools. More surprisingly, teachers give around 0.3 standard deviations lower grades than IRT scores in white-dominated schools (percentage of minorities less than 10 percent) and, on the contrary, give 0.3 standard deviations higher grades in minority-dominated schools (percentage of minorities larger than 75 percent). Generally speaking, it is very clear that teachers give much more generous grades to their students in schools with a higher percentage of minorities.

For the purpose of showing teachers' different treatment of white and black students more accurately, I then look at the IRT scores and teachers' subjective assessments for whites and

[43] I use IRT score and teachers' assessments in grade 3 instead of grade 5 here. The reason is that the trend of the racial gaps in grade 5 is smaller than in grade 3, which is different from the trend from kindergarten to grade 3.
[44] Racial composition is a categorical variable in ECLS-K: percentage of minorities is less than 10 percent, 10–25 percent, 25–50 percent, 50–75 percent, and larger than 75 percent.

blacks separately by racial composition of schools, in Figure B.2 and Figure B.3. The most important information we can get from Figure B.2 is that teachers do not favor white students and give them higher grades. We can also note that teachers in white-dominated schools are stricter on average. The grades they give to white students are much lower than the students' IRT math scores. In minority-dominated schools, teachers give grades similar to IRT scores to white students.

Figure B.3 also shows the mean of IRT math scores and teachers' assessments of math in grade 3 by racial composition, but for black students. At first we see a well-known phenomenon—that there are huge gaps in cognitive skills between blacks and whites. Second, the pattern of grades is similar to that in Figure B.2. Teachers in minority-dominated schools are more generous to students. In addition, there is another important pattern worth mentioning. After comparing Figure B.3 with Figure B.2, it is clear that no matter the racial composition, teachers are more generous to black students. Specifically, in minority-dominated schools, the grades black students get are much higher than their IRT scores; even in white-dominated schools, black students still get nearly the same or even slightly higher grades than IRT scores.

All in all, we can see that teachers in schools with different racial compositions perform differently. Students get higher grades for cognitive skills in schools with more minority students, conditional on their IRT scores. This clearly shows that if there are variables that cause biases but are not contained in equation (1), the estimation would be biased, and the racial gaps would be attenuated because of this problem.

2.5. The Racial Gap in Non-Cognitive Skills

2.5.1. School fixed effects and teachers' characteristics

According to the earlier discussion, the measurements of non-cognitive skills are biased. In order to ascertain the real racial gaps in non-cognitive skills, I need to find a way to reduce the bias. From equation (2.2) we know that I can add X_2 variables into the regression if I know them, in order to solve the estimation-bias problem. As a result, the estimation error will be reduced, and I will have more accurate measures of racial gaps in non-cognitive skills. Specifically, if I run $Y^* = \alpha_1 RaceDummies + X_{11}\beta_1 + X_{12}\beta_2 + \nu$ and stipulate that X_1 includes X_{11} and X_{12}, $M_1 = I - X_1(X_1'X_1)^{-1}X_1'$, then:

$$\begin{aligned} E(\hat{\alpha}_1) &= E[(R_1'M_1R_1)^{-1}R_1'M_1Y^*] \\ &= E[(R_1'M_1R_1)^{-1}R_1'M_1Y] + E[(R_1'M_1R_1)^{-1}R_1'M_1\varepsilon] \\ &= \alpha_1 + E[(R_1'M_1R_1)^{-1}R_1'M_1(X_{11}\delta_1 + X_{12}\delta_2 + \eta)] \\ &= \alpha_1 \end{aligned} \tag{2.3}$$

From these equations we find that after adding X_{12} into the regression, the estimation bias is successfully eliminated.

Based on this discussion, I try to improve my estimation by finding and adding part of X_{12} in the regression (1). Since the measurements of non-cognitive skills in ECLS-K are teacher-reported assessments, some variables that belong to X_{12} can be related to a teacher's characteristics. As a result, the best way is to use teacher fixed effects. However, owing to the limit of ECLS-K, the number of students taught by the same teacher is not large enough to use teacher fixed effects. Here I use school fixed effects instead and control for some important teachers' characteristics. Specifically, I control for teachers' birth year, education, teaching experience, and race in the regression.

In Table C.7, I present the evolution of non-cognitive skills gaps after using school fixed effects and adding teachers' characteristics. The results in the first five columns are racial gaps in externalizing behavior. We find that the gap is nearly 0.40 standard deviations, even when children were just in the fall of kindergarten. After that, it becomes more than 0.48 standard deviations in grade 1. In grades 3 and 5, the racial gaps in externalizing behavior stay around 0.40 standard deviations. In addition, columns (6) to (10) present the racial gaps in approaches to learning. Comparing the results in Table C.4, the gaps not only become much larger but also have a clear evolutionary pattern. In kindergarten the gaps are around 0.20 standard deviations. The gaps reach 0.38 standard deviations in grade 3. In summary, these results are consistent with the phenomenon I find in previous sections and this section. Since many more black students attend schools where teachers tend to give higher grades to students, the original racial gaps are attenuated. When controlling for the factors that cause bias of the measurements, the racial gaps are larger.

Though the results in Table C.7 are likely less biased than those in Table C.4, it is still not clear how good the measures of racial gaps in non-cognitive skills are. The method I use is not perfect, because in reality it is difficult to determine all of the variables that cause bias in non-cognitive skills measures. Therefore, the method can only remove part of the second part of the bias $(R_1'M_{11}R_1)^{-1}R_1'M_{11}X_{12}\delta_2$.

So as to check whether this method, which is adding school fixed effects and teachers' characteristics, can reduce a large proportion of the bias, I still use information concerning IRT scores and subjective measurements of cognitive skills. When using only control variables that are the same as those in Table C.4, the racial gaps in teachers' assessments of cognitive skills are much smaller than those in IRT scores. This is consistent with the discussion in the previous chapter. Since teacher-reported measures of cognitive skills are subjective, they also suffer the same type of bias with teachers' assessment of non-cognitive skills. In Table C.8, I present the racial gaps in IRT scores and teacher-reported cognitive skills, after using school fixed effects and controlling for teachers' characteristics. It is obvious that the racial gaps in IRT reading scores and teachers' subjective assessments of reading are very similar. This implies that using school fixed effects and adding teachers' characteristics can help eliminate teachers' subjective bias for reading scores. However, in terms of math scores, the difference in racial gaps is still large.[45] All in all, the results in Table C.8 imply that using school fixed effects and adding teachers' characteristics can help to decrease the subjective bias but cannot fully eliminate it. In Appendix Table D.1, I present additional evidence to support this conclusion. I use the index of bias as the dependent variable and the same control variables used in Table C.6 but add school fixed effects and teachers' characteristics. In Table C.6 we discern that teachers' subjective bias is still significant, especially in grade 3 and grade 5. In Appendix Table D.1 the coefficients on the black dummy variables are close to zero in the spring of kindergarten, grade 1, and grade 5. However, coefficients in the fall of kindergarten and grade 3 are still positively significant. Therefore, it is likely that the racial gaps in non-cognitive skills shown in Table C.7 lie between

[45] Why using school fixed effects and adding teachers' characteristics can help eliminate teachers' subjective bias for reading scores but not for math scores is not clear. It might be because teachers measure different skills from those measured by math IRT scores when giving math scores.

the unbiased racial gaps and those gaps in Table C.4. In other words, the unbiased racial gaps are even larger than those in Table C.7.

2.5.2. Principal component analysis

Though I focus on externalizing behavior and approaches to learning in this paper, self-control, interpersonal skills, and internalizing problems are also important measurements of non-cognitive skills. Here I use principal component analysis to investigate the racial gaps in students' overall non-cognitive skills. Specifically, I use information concerning all these five measurements to do the principal component analysis. The first principal component of these five non-cognitive skills measures explains, on average, 65 percent of the variation of all children's performance.[46] Furthermore, the first component is the only one with an eigenvalue significantly greater than 1.[47] The Kaiser rule recommends retaining only factors with eigenvalues exceeding unity. Therefore, I use the first component as the measure of overall non-cognitive skills. This is the linear combination of the five original non-cognitive skills variables and accounts for maximum possible variance. The results in Table 9 show that the racial gaps in overall non-cognitive skills are even larger than those in externalizing behavior and approaches to learning. In the fall of kindergarten, the gap is 0.59 standard deviations. In grade 3 it reaches 0.86 standard deviations.

[46] Please see Appendix E Table E.2, which presents the component loadings of measures of noncognitive skills in grade 5.
[47] Please see Appendix F Figure F.3, which presents eigenvalues of first five components of non-cognitive skills in grade 5.

2.5.3. Distributional differences

In the previous sections I have shown average racial gaps in non-cognitive skills. However, the differences between the distributions of white students and black students are also very important. In Figure B.4 and Figure B.5, I present the distributions of externalizing behavior and approaches to learning in grade 5, respectively, for white and black children separately. From Figure B.4 it is clear that the distribution of black students is much more flat, which means that the variance of black students' performance is much smaller than that of white students. On the contrary, the distribution of white students' externalizing behavior is strongly skewed to the right. Nearly half of white students have very little externalizing behavior, but only around 30 percent of black students get similarly good scores. All in all, the racial gaps in externalizing behavior are mainly caused by the performance of the top students. Though a small proportion of both white and black students have much more bad behaviors, a much larger proportion of white students perform very well. On the contrary, many more black students get middle-level scores in externalizing behavior. Figure B.5 shows that the distributions of approaches to learning have a pattern similar to that of externalizing behavior.

2.6. Conclusion

Racial inequality has long existed in the United States. Even now, relative to whites, blacks earn less, get less education, have worse health, live fewer years, and have a higher probability of being in jail and being unemployed. What causes these phenomena and how to eliminate these differences are extremely important for US society. In recent years, the racial difference in cognitive skills has attracted a lot of attention. Much research has shown that the black-white gap in cognitive skills emerges at a young age and increases gradually. At the same time, people are gradually becoming aware of the importance of non-cognitive skills. They play a significant role in determining educational achievement, wages, probability of engaging in criminal activities, and other outcomes. This paper is the first paper that relates these two important issues and offers a new perspective on the serious racial gap that exists in the United States.

Using a nationally representative dataset, ECLS-K, I find that there are significant differences in non-cognitive skills between white and black students, even after controlling for a large set of background variables. Specifically, blacks have around 0.2 standard deviations more externalizing behaviors than whites in grade 5 and are 0.12 standard deviations more likely to be suspended in grade 8. People might be concerned that this gap has been caused by the teachers' test-score bias against black students. However, the evidence from ECLS-K shows that teachers do not discriminate against black students. In addition, teachers give higher grades to students from families of lower SES, those from public schools, and students in schools that have a higher percentage of minorities. Since reverse advantages for minorities exist and attenuate the racial gap, I try to use school fixed effects and control for teachers' characteristics to deal with the bias. The more accurate racial gaps in non-cognitive skills are very large. They are more than twice as large as the original gaps in third grade and fifth grade.

This large racial gap in non-cognitive skills gives us a new angle to understand racial inequality in the United States. Since non-cognitive skills have a large impact on many important outcomes, they can help to explain the racial gap in wages, probability of arrest, teenage pregnancy, and other important outcomes in the United States. This research also gives us a new potential method to reduce and eliminate the racial inequality between blacks and whites. If we can successfully lessen the black-white differences in non-cognitive skills, the racial gap in wages, education, and other outcomes will substantially decrease.

Chapter 3

The tale of silent dogs:

Do stock prices fully reflect the implication of news withholding?

3.1. Introduction

Many prior studies on voluntary disclosure assume that investors understand the implication of nondisclosure. In this view, investors correctly infer the range of news that management withholds. Studies in non-capital market settings, however, show that economic agents do not correctly infer information solely based on other agents' strategic actions (Eyster and Rabin, 2005). We test whether investors are also subject to bias when interpreting nondisclosure. Examining quarters in which management does not provide a guidance, we find that investors underestimate the magnitude of bad news associated with nonguidance. The finding extends our understanding of the way investors respond to strategic disclosures, and has implications on management's voluntary disclosure decisions.

We use the setting of management forecast to test whether investors understand the implication of nondisclosure. Nondisclosure refers to the absence of a management forecast in a quarter. First, nondisclosure is prevalent. For firms that had at least one management forecast between 2003 and 2014, management does not provide guidance for about half of the time. The nontrivial proportion of nondisclosure makes it more likely that investors are able to learn the implication of nondisclosure. Second, management forecast decisions are well scrutinized by investors Ball and Shivakumar (2008) and repeated over time.

If investors correctly interpret the implication of nondisclosure for stock price, conditional on the same underlying news, investors' reactions to disclosure and nondisclosure *on average* should not be systematically different. For example, if management withholds all news below \$5 and investors are rational, rational investors' reaction to nondisclosure should be equal to their average reaction to the disclosure of news below \$5[48]. Contrary to this prediction, we demonstrate that nondisclosure is on average associated with better market reactions than disclosures *conditional on the same underlying* news, proxied by quarterly earnings surprise. The result is robust to controlling for some determinants of disclosure choice, calendar year-quarter fixed effects, the exact timing of disclosure and nondisclosure, and other possible behavioral biases such as post earnings announcement drifts.

Importantly, we further show that the difference in market reaction to disclosure and nondisclosure declines significantly around the earnings announcement. The result suggests that investors did not interpret nondisclosure correctly prior to the earnings announcement and subsequently adjust their interpretation after observing the announced earnings. One possible explanation of the results we find is that investors are subject to limited strategic thinking (LST thereafter) when interpreting nondisclosure. That is, although investors might understand nondisclosure implies bad news, as demonstrated by prior studies (e.g., Chen *et al.* (2011)), LST predicts that they underestimate how bad the news really is.

It is possible that omitted variables drive the results. These omitted variables should explain the systematic better market reaction prior to the earnings announcement and the subsequent adjustments. In particular, they must change around the earnings announcement and be observed by investors in such a way that relative to firm quarters with disclosures, they increase firm value

[48] We emphasize the on average nature of the better market reaction to nondisclosure than disclosure, because nondisclosure is associated a range of news and for individual cases it can happen that nondisclosure is better or worse than a disclosure of the same news.

before the earnings announcement and decrease firm value following the earnings announcement. We argue that such factors are more likely to be investor related than firm related, because we do not expect the latter to systematically change directions around the earnings announcement.

We next explore factors that affect the magnitude of the bias. We find that the results are mostly driven by negative earnings surprises, that is, bad news firm quarters. The decline in the nondisclosure premium is larger both economically and statistically for nondisclosure firm quarters with negative earnings surprises. The result suggests that investors suffer from the bias when managers attempt to withhold bad news, which is consistent with prior literature that shows withholding information is more beneficial than disclosure when news is bad. We also find the results are more significant for firm quarters with high analyst following. If more analyst following is associated with high investor attention, the evidence suggests that the bias is distinctive from limited attention. Third, we find weak evidence that the results are stronger for high investor uncertainty, using return volatility as a proxy. The decline in nondisclosure premium is larger for the high investor uncertainty group. However, the nondisclosure premium prior to the earnings announcement is higher for low investor uncertainty group, suggesting the bias is not driven by disagreement among investors but something systematic.

We then do some robustness tests to provide more evidence. We show our results are robust to using size adjusted returns. Our results are consistent with LST when we include bundled forecasts in our estimation.

The paper contributes to two streams of literature. First, we show investors are subject to limited strategic thinking; that is, they do not fully understand the implications of managers being strategic about information disclosure. Limited strategic thinking is a type of behavioral bias in the financial market and is different from other types of cognitive constraints. It has important implications because investors often need to form inferences from firms' strategic actions, for example, the lack of a disclosure decision. Our findings suggest that they have limited ability to do so. This finding is significant, because market forces should eliminate limited strategic thinking, which include but are not limited to the existence of analysts and sophisticated institutional investors, the repeated game nature of information disclosure, and the disclosure of true earning each fiscal period. Second, we contribute to the voluntary disclosure literature. Prior studies show investors interpret withholding information as bad news. However, whether investors' response to withholding is consistent with rational strategic thinking has not been explored. The answer to this question is important for our understanding of why managers do not always voluntarily provide information before earnings announcements. Our results suggest withholding information is not as costly as a rational model predicts. Managers can even benefit from withholding information by making use of investors' cognitive constrains and achieve better pricing. The finding is also policy relevant. It suggests that increased disclosure can benefit investors if the disclosure choice is expected to be strategic and investors do not understand it correctly.

3.2. Literature Review

The paper is related to three streams of literature, investor response to nondisclosure, managers' voluntary disclosure decisions, and economic agents' cognitive constraints in games of strategic interactions.

First, although a large literature examines how investors respond to information disclosed by management Ball and Brown (1968); Beaver (1968), fewer studies investigate how investors respond to information that management strategically withholds. Consistent with analytical research Verrecchia (1983); Dye (1985), prior studies find that nondisclosure is associated with bad news. For example, Chen *et al.* (2011) investigate firms that announced their decisions to stop providing earnings guidance and document negative three-day return around the announcement day. Although Chen *et al.* (2011) show that investors perceive nondisclosure as bad news, they do not examine the nature of investors' response to nondisclosure. Unlike their study, we examine all firms that stop providing guidance in a quarter, and primarily focus on whether investors fully understand the strategic decision of information withholding. Giglio and Shue (2014) also examine whether investors fully understand the absence of news; in their setting, no news is related to the passage of time after announcing a merger. Our paper is different. We focus on the disclosure decision of management instead of the passage of time.

Second, this paper is related to the voluntary disclosure literature. Most models of voluntary disclosure assumes that investors are rational when managers strategically withhold information Verrecchia (1983); Dye (1985). Examining investor reactions to strategically withheld news provides insights about the cost and benefit trade-off of managers' disclosure decisions. Specifically, such under-reaction makes nondisclosure less costly and thus provides another reason for withholding information.

Third, this paper complements the findings in nonfinancial settings that economic agents are subject to limited strategic thinking (e.g., Eyster and Rabin (2005); Brown *et al.* (2012)), that is, economic agents cannot correctly infer other agents' private information from observing their actions. We provide evidence on investor limited strategic thinking in the financial market.

3.3. Hypotheses

In this section we discuss the hypothesis. According to Grossman (1981) and Milgrom (1981), in equilibrium managers will only withhold if earnings surprise St is the worst, denoted by SA. In other words, if the manager withholds St, rational investors will expect:

$$E(St|ND)= SA.$$

Investors defined above are rational. Limited strategic thinking is one possible explanation, we define it in the context of the strategic disclosure game described above.

Definition 1. Limited strategic thinking (LST)

We call investors to be λ (λ ≥ 0) cognitive constrained and denote them as λ LST investors, if their expectation of S_t conditional on observing nondisclosure takes the form:

$$(1 - \lambda)E(St|ND) + \lambda\mu, \tag{3.1}$$

where μ = E(St), the prevailing expectation of St before observing the manager's disclosure decision, and E(St|ND) is rational investors' expectation accounting for the manager's nondisclosure incentive.

From Definition 1, LST investors' expectation of firm value following nondisclosure is higher than that of rational investors. The extent of limited strategic thinking varies with λ. When λ = 1, investors completely ignore the strategic nature of nondisclosure and perceive nondisclosure as uninformative. When λ = 0, investors are rational and completely understand the strategic nature of nondisclosure.

3.3.1 Empirical predictions

We derive two testable empirical predictions about the bias. First, investors' expectation of firm value following nondisclosure is different from the rational investors if they are subject to bias. If this bias is caused by LST, according to Definition 1, investors' expectation of firm value following nondisclosure is higher than that of rational investors.

To test this prediction, we examine the change in investors' expectation of firm value following a disclosure decision. Specifically, we measure cumulative stock returns from six days after the previous quarter earnings announcement to various dates around the earnings announcement date of that quarter. The beginning of the return measurement window, combined with our sample selection procedure (dropping firm quarters with disclosures bundled with the earnings announcement), ensures that management has not yet made a disclosure decision at that point and that investors have not yet reacted to a disclosure decision. Vary the ending dates of the return measurement window has two benefits. First, it allows us to robustly capture investors' reaction to nondisclosure despite not knowing the exact date on which they realize nondisclosure. Second, we would like to observe how information regarding nondisclosure impounds into the stock price over time.

The test also requires measuring rational investors' reaction to nondisclosure. We argue that 5 in cases where management provides a disclosure, investors' reaction to the disclosure of news is not subject to the bias, because by definition investors do not need to form expectations about the manager's nondisclosure incentive. LST predicts higher *average* cumulative stock returns prior to earnings announcements for nondisclosure firm-quarters than for disclosure firm-quarters *conditional on the same earnings news*. One immediate concern with this approach is the underlying reason why the same earnings news is disclosed or withheld might lead to the differential market reactions. We leave the relevant discussions until the research design section and assume for now that these reasons do not change firm value. We summarize the prediction below.

Hypothesis 1: The cumulative stock returns prior to the earnings announcements for nondisclosure firm-quarters are on average higher than those for disclosure firm-quarters, conditional on the same earnings news.

We emphasize the “on average” nature of investors’ reaction to nondisclosure. Rational investors only guess the range of news that management withholds. That is, they are correct on average. It follows that LST should imply an average better reaction to nondisclosure than to disclosure conditional on the same underlying news. A second aspect of the hypothesis worth explaining is that conditioning on the same earnings news does not imply that investors know the earnings news before its release. The right interpretation is that given the same ex-post news content, whether investors on average can infer the news correctly based on firm disclosure decisions.

A second implication is that investors will be surprised by the earnings announcement following nondisclosure. When earnings are announced, investors can correct their expectations based on the information from the earnings announcement. As a result, after the earnings announcement, the higher cumulative stock return following nondisclosure predicted in Hypothesis 1 would become smaller or even disappear, as investors correct any bias caused by the bias. Crucially, if investors are not subject to the bias or more specifically, LST, we will not observe such a correction, because in that case investors should on average correctly incorporate relevant information into the stock price prior to the earnings announcement.

Hypothesis 2: The cumulative return difference predicted in Hypothesis 1 declines after the [6] earnings announcement.

3.4. Sample Selection

Our empirical tests examine quarterly management forecast decisions. Testing the bias requires that nondisclosure be a strategic decision made by management, a feature well documented by prior research (e.g., Chen *et al.* (2011)). Another requirement is that investors expect a disclosure and understand the lack of a disclosure. The prevalence of management quarterly earnings forecasts makes it more likely that investors expect a forecast issuance and, importantly, understand the absence of a forecast. This feature is crucial, because it is less likely our results are driven by limited attention. Prior research shows that saliency of signals can affect investors' judgement, (e.g., Giglio and Shue (2014;) So and Weber (2015)). Finally, quarterly management forecasts often directly map into the forthcoming quarterly earnings. This feature allows us to use the quarterly earnings surprise as our proxy for the underlying news behind the management forecast decision, and to have a proxy for news even when management withholds information.

To construct our sample, we first collect management forecasts issued between 2003 and 2014 from I/B/E/S Guidance. We choose 2003 as the starting year, because data coverage after 2003 is more comprehensive, which reduces the chances of classifying missing disclosures as nondisclosure. Management forecast decisions are generated by matching forecast announcement dates with earn- ings announcement dates. For example, if a firm announces its first quarter earnings on April 15 and its second quarter earnings on July 15, we will call the second firm quarter (that announces earnings on July 15) a “disclosure quarter,” if at least one forecast is issued between these two dates, and a ”nondisclosure quarter,” if no forecast is issued between these two dates. In doing so, we require all firms to have valid PERMNOs and have non-missing earnings announcement dates throughout the sample period.

Data cleaning involves three steps. To minimize the presence of irregular firms, we delete firms that have at least one case of larger than 120-day gap between adjacent fiscal quarter end dates, and firms that have at least one case of larger than 180-day gap or smaller than 28-day gap between adjacent earnings announcement dates. We then keep only earnings related forecasts. The final step is to restrict the timing of management forecasts. First, we drop forecasts issued after the fiscal quarter end date, because these forecasts are typically pre-announced earnings, the lack of which are less likely to be the type of strategic decision we are interested in, that is, withholding bad news. Second, for our main analysis, we drop forecasts bundled with an earnings announcement to avoid confounding investors’ reactions to forecasts and actual announced earnings.

Finally, we merge the data with stock return from CRSP and realized earnings and other financial information from Compustat-CRSP merged data. Table 1 presents our sample selection criteria. We have 58,469 unique firm quarters with data on stock returns and control variables.

3.5. Research Design

Our research design focuses on quarterly cumulative stock returns, measured from 6 days subsequent to the previous quarter earnings announcement date to various days around the current quarter earnings announcement date. The ending date of the return measurement window varies from 30 days prior to the current quarter earnings announcement date to 5 days subsequent to it. We measure returns from 6 days subsequent to the previous quarter earnings announcement date to reduce the impact of previous quarter earnings announcement and allow sufficient time for investors to adjust their expectation. By removing bundled forecasts, we ensure that management has not yet made a disclosure. Varying the ending dates allows us to capture investors' response to nondisclosure, the timing of which is unobservable, as well as to observe how information is impounded into the stock price around the earnings announcement. We later examine the robustness of our results to bundled forecasts.

The regression model is:

$$CR_{iqt} = \beta_0 + \beta_1 ND_{iq} + \beta_2 After_{iq} * ND_{iq} + \beta_3 EarnSurp_{iq} + \beta_4 After_{iq} * EarnSurp_{iq} + \beta_5 After_{iq} + \varepsilon_{iqt}, \tag{3.2}$$

where CR_{iqt} is firm *i*'s raw cumulative stock returns from 6 days subsequent to the previous quarter earnings announcement date to *t* days prior to the current quarter earnings announcement; *t* varies from 30 to -5, which means from 30 days *prior* to the current quarter earnings announcement to 5 days subsequent to the current quarter earnings announcement; ND_{iq} is a dummy variable that equals to one when management of firm *i* does not provide a forecast in quarter q; $EarnSurp_{iq}$ is the earnings news of quarter *q* for firm *i*, measured as the difference between current quarter earnings and the same quarter earnings of last year, scaled by the closing stock price five days subsequent to the previous quarter earnings announcement; $After_{iq}$ is a dummy variable that equals to one if the cumulative return ending window is on or after the current quarter earnings announcement date, and zero otherwise; ε_{iqt} is the error term. β_1 and β_2 are what we are most interested in.

For our first hypothesis, if investors are subject to LST, we expect β_1 to be positive, that is, conditional on the same underlying news, nondisclosure is associated with higher stock return prior to the earnings announcement. We do not expect a positive β_1, if, on average, investors correctly understand the relation between managers' strategic disclosure incentive and stock price.

For our second hypothesis, we examine whether the cumulative stock return difference between disclosure and nondisclosure firm quarters significantly declines around the earnings announcement date. A negative β_2 is consistent with investors adjusting their originally incorrect expectation after observing the information from the earnings announcement. We do not expect this to happen if investors were on average correct about the underlying news.

The threat to internal validity of our regression model is that omitted factors affect disclosure choice as well as cumulative stock returns, leading to spurious effects. These omitted factors need to satisfy three conditions to explain our empirical results. First, they have to affect the disclosure decision, that is, the independent variable, prior to the earnings announcement. Second, they have to lead to a better market reaction to nondisclosure than to disclosure conditional on the same underlying news prior to the earnings announcement. Third, they have to change around the earnings announcement to explain the decline in the cumulative stock return difference between nondisclosure and disclosure firm quarters around the earnings announcement. Importantly, the change cannot be anticipated by investors, because otherwise investors would incorporate it into the stock price prior to the earnings announcement, which would not explain the decline in the cumulative return difference.

We believe that factors satisfying the criteria above are hard to find, because firms are not likely to systematically make decisions to improve market reaction in nondisclosure firm quarters but only to reverse it subsequent to the earnings announcement. Because we cannot directly rule our the existence of these factors, to alleviate any residual endogeneity concerns, we add control variables to the main specification presented in equation (2) following prior studies:

1. Log(Market Value of Equity)$_{i,q-1}$: previous quarter log total market value of equity, which controls for size risk and to some extent firm information environment;

2. Return Volatility$_{i,q-1}$: previous quarter daily return volatility, because uncertainty affects disclosure decisions Waymire (1985) and return volatility also controls for idiosyncratic risk from the previous quarter;

3. Number of Analist Following$_{i,q-1}$: previous quarter number of analyst following, which controls for investors' attention;

4. Earnings per Share$_{i,q-1}$: previous quarter earnings per share defined as the primary earnings per share excluding extraordinary items, scaled by stock price five days subsequent to the earnings announcement two quarters ago.

5. Earnings Surprise$_{i,q-1}$: previous quarter earnings surprise, which controls for potential post earnings announcement drift from the previous quarter and overall firm performance.

6. Market to Book Ratio$_{i,q-1}$: previous quarter market to book ratio, which controls for firm growth and growth risk.

7. Cumulative Return$_{i,q-1,-5}$: previous quarter cumulative stock return (from 6 days subsequent to the earnings announcement two quarters ago to 5 days subsequent to the previous quarter earnings announcement date), which controls for price momentum.

In Table F.2, we present the summary statistics of these control variables.

3.6. Limited Strategic Thinking in the Stock Market

3.6.1. Descriptive statistics

This section describes how investors react to management quarterly forecast decisions conditional on actual earnings surprise, which proxies the news of that quarter. We first provide descriptive evidence on to what extent investors correctly price the information content of forthcoming earnings based on disclosure decisions prior to the earnings announcement. Following Hypothesis 1, on average investors' reaction to nondisclosure should not be better than their reaction to disclosure conditional on the same underlying news. Following Hypothesis 2, if investors are right on average, market reactions should not be systematically different before and after the earnings announcement.

To illustrate the two hypotheses, in Figure H.2, we divide our sample into four parts based on the quartiles of quarterly earnings surprise, measured as the difference between quarterly earnings and the same quarter earnings last year divided by stock price five days after the previous quarter earnings announcement date. For each earnings surprise quartile, we plot the cumulative stock return from 6 days subsequent to the previous quarter earnings announcement to various dates around the current earnings announcement.

Figure H.2a shows that for bad news quarters (the bottom quartiles of earnings surprise), cumulative stock return difference between disclosure and non-disclosure firm quarters are very small after the earnings announcement, suggesting that their actual information contents are similar. However, when management did not provide management forecast, cumulative stock returns prior to the earnings announcement are higher for non-disclosure firm quarters than nondisclosure firm quarters, suggesting that investors over-price non-disclosure firm quarters prior to the earnings announcement. This is consistent with our hypothesis. Investors are subject to LST when management withholds news: they underreact relative to the implication of nondisclosure for stock price and subsequently adjust their underreaction after the earnings announcement. In contrast, we observe no strong adjustment for disclosure firm quarters, which suggests that investors correctly interpreted news from management forecast. Figure H.2b shows a similar result, although non-disclosure firm quarters have better reaction than disclosure firm quarters even after the earnings announcement.

The over-pricing and subsequent adjustment for nondisclosure firm quarters are observed only in the first two quartiles of earnings surprise distribution. For good news firm quarters (Figure H.2c and Figure H.2d), the first observation is that the cumulative stock return trends within a quarter are almost parallel for disclosure and nondisclosure firm quarters. The difference in cumulative stock returns between disclosure and non-disclosure firm quarters are very small both prior and subsequent to the earnings announcement for the third quartile earnings surprise (Figure

H.2c). Disclosure firm quarters outperform nondisclosure firm quarters both prior and subsequent to the earnings announcement for the fourth quartile earnings surprise (Figure H.2d). Finally, both Figure H.2c and Figure H.2d show an upward adjustment prior to the earnings announcement, consistent with information leakage.

Overall, Figure H.2a shows investors react to good news and bad news differently. Because our model does not predict management to withhold good news, we perform our empirical analysis separately for good news and bad news samples and expect our results to manifest in bad news quarters. Another pattern is that investors incorporate part of firm earnings news prior to the earnings announcement. This feature provides support to our assumption that investors are likely to realize the lack of a disclosure.

Table G.2 provide the distribution of our main variables and control variables. Including bundled forecasts, the average nondisclosure frequency is 97%. After dropping bundled forecasts, the disclosure frequency is 54%. We winsorize all continuous control variables at 1% and 99% of their distribution.

In addition to the decision to issue a forecast, managers have the discretion to provide forecasts at any time within a quarter, which makes it difficult for researchers to capture the timing at which investors realize nondisclosure. We next examine the timing of management forecasts. Figure H.3 presents the probability density function of voluntary disclosure date relative to current period earnings announcement date. We find that managers provide the first management forecast at different points in time within a quarter. The large variation in the timing of forecast provision and the possibility that investors gradually realize the lack of a forecast are consistent with our choice of cumulative stock returns as our main dependent variables of interest.

3.6.2. Main results

This section presents multivariate results on investors' limited strategic thinking. Table G.3 presents the results for our main specification of equation (3.2).

In Column (1) and (2), the dependent variable is the cumulative stock return from 6 days after the earnings announcement date of the previous quarter to a window around the current quarter earnings announcement date. The window is 30 days prior to the current quarter earnings announcement and 5 days after the current quarter earnings announcement respectively. There are 36 observations for each firm quarter. The independent variable of interest is ND, a dummy variable that indicates the absence of management forecast in a quarter, and ND*After, where After is one if the ending date of cumulative return is on or after the earnings announcement.

In Column (1), we can find that conditional on the same earnings surprise, the cumulative stock return is 2% higher if the manager chooses not to provide a forecast than otherwise. The magnitude is both economically and statistically significant. The finding suggests that investors pay a premium for nondisclosure even holding the underlying news constant, which is consistent with our first hypothesis that investors do not fully respond to the information content of nondisclosure. In addition, the third row of Column (1) presents that the cumulative return premium for nondisclosure drops 0.5% after the earnings announcement. This provides the evidence supporting the second hypothesis, which is the cumulative return difference predicted in hypothesis 1 declines after the earnings announcement. Another point deserves to be mentioned is that the explanation power of earnings surprise on cumulative return sharply increases after the earnings announcement. This result once again implies that the information that earnings surprise contains are not fully incorporated in the stock price before the earnings announcement. In

Column (2), we add control variables to alleviate the endogeneity concern. The results are very similar with those in Column (1).

Because investors might not realize the lack of disclosure on 30 days prior to the current quarter earnings announcement, in Column (3) and (4) we use the same regression but only use cumulative returns that end 5 days around the current quarter earnings announcement date. Therefore, there are only 11 observations for each firm quarter for the short window analysis in Table G.3. The advantage of short window analysis is that investors are likely to realize the lack of disclosure and also fully react to disclosed managers' forecast a few days before the earnings announcement. The downside is that investors might also realize actual earnings and correct their bias right before the earnings announcement, which may attenuate the results we find in Column (1) and (2). The results in Column (3) show that conditional on the same earnings surprise, the cumulative stock return is 2.3% higher if the manager does not provide a forecast than otherwise. After earnings announcement, the nondisclosure premium declines 0.7%. In Column (4), the coefficients are nearly the same with those in Column (3). In summary, when using short window analysis, the pattern is the same with that when using long window analysis.

3.6.3. Sub-sample results

Although we include control variables and argue that unobserved heterogeneity is unlikely to affect our results, our empirical tests are still not perfect to show the bias. To make the discussion more convincing, we provide evidence on factors that affect the level of the bias. The cross sectional results serve two purposes. First, we provide patterns that vary with the extent of investors' subjective bias. Second, we show that these patterns are consistent with the bias but alternative channels cannot explain some of the results.

3.6.4. Bad news versus good news

Recall that our model predicts the nondisclosure occurs when earnings surprise of firm quarters are lower than the equilibrium threshold S^*. In other words, we expect that management withholds bad news. However, management might have incentives to withhold good news.

If investors are subject to the bias and management withholds good news, investors' underreaction to good news withholding will lead to lower market reaction to nondisclosure prior to the earnings announcement and subsequent increase in cumulative stock return following the earnings announcement. Because the two represent distinctive non-disclosure incentives and potentially opposite predictions, we examine them separately in this subsection.

Before discussing our subsample results, the subsample analysis assumes that although investors might not interpret the implication of nondisclosure correctly, they should have the ability to distinguish the sign of the news. Figure H.2 is consistent with this argument. At least on average, investors predict the sign of earnings news correctly prior to the earnings announcement. Cumulative stock returns tend to decrease for bad news firm quarter and increase for good news firm quarter prior to the earnings announcement.

Table F.4 present the results for good news sample and bad news sample by using regression model 2 with and without controls and year*quarter fixed effects. For firm quarters with non-positive earnings surprise, conditional on the same earnings surprise, the cumulative stock return is around 3.5% higher if management does not provide a forecast. After the earnings announcement, the nondisclosure premium drops by 1.5%. The evidence of the bad news sample supports LST and features larger economic magnitude. On the contrary, we find small nondisclosure premium of the sample with good news, especially after adding controls. Moreover, there is also no different reaction to earnings announcement between nondisclosure and disclosure firm quarters when the underlying news is good.

Figure H.6 plots the coefficients of NG separately for positive earnings surprises sample and non- positive earnings surprises sample by using regression model 2 with controls and year*quarter fixed effects. We again find support for LST for the bad news sample but not the good news sample.

Ideally, we want to estimate our empirical model using only observations below the threshold of disclosure to test LST. Because we cannot pin down the threshold, our main empirical tests in Table F.3 rely on the full sample. Focusing on the full sample is more conservative in that we include all good news cases from which we may even find opposite premium. However, the estimation results from the sample, which are restricted to non-positive firm quarters, are also essential. To further test whether the subsample results in this section is consistent with LST, in Figure H.7 we plot the raw stock return 5 days around earnings announcement date separately for disclosure and nondisclosure firm quarters. To figure out how does the news content affect the stock return differently for both of the two samples, we further divide our sample into four parts based on the quartiles of quarterly earnings surprise.

Figure H.7 suggests that for bad news quarters (Figure H.7a and Figure H.7b), investors are surprised by the earnings announcement, only when management provided no earnings forecasts prior to the earnings announcement. In contrast, we observe no such pattern for disclosure firm quarters, which suggests that investors correctly interpreted news from management forecast and earnings announcement provides little extra information. For good news quarters, investors are also surprised by the earnings announcement of good news when there is no management earnings forecasts prior to the earnings announcement, evident from the increase in cumulative stock return around the earnings announcement. Somewhat surprisingly, we also observe increase in cumulative stock returns around the earnings announcement for disclosure firm quarters. This result is consistent with investors mistrusting the good news disclosure management voluntarily provided and only react to good news disclosure during the actual earnings announcement. One explanation for the mistrust might be good news voluntary disclosure is not audited. While the pattern that the stock return of nondisclosure firm quarters increases after the earnings announcement is consistent with the prediction of the LST model, we need to interpret results from the good news sample with caution. First, investors do not seem to understand management forecast with good news correctly, making it not be a good benchmark. In addition, our model does not predict management to withhold good news. Some other factors may affect managers' decisions and investors' reactions during good news quarters, which makes it difficult to tease out LST in good news quarters.

3.6.5. The number of analysts following

Analysts are one of the most important information intermediaries. We want to assess how analyst following affects the extent of LST. We divide our sample into two sub-samples, one with the number of analyst following of the previous quarter being larger than the median and the other with the number of analyst following of the previous quarter being smaller than the median. Then we compare the nondisclosure premium between the two subsamples.

Table F.5 present the sub-sample results by using regression model 2 with controls and year*quarter fixed effects. For the sample with the number of analyst following of the previous quarter being larger than the median, both of the nondisclosure premium and the subsequent drop are significant.

In contrast, we find no evidence for LST for the sample with the number of analyst following of the previous quarter being smaller than the median.

Figure H.8 plots the coefficients of NoGuide for the sub-sample results by using regression model 2 with controls and year*quarter fixed effects. Red dots show the results for firm quarters with high number of analyst following. For firms quarters with the same earnings surprise, on average the cumulative returns from the last period announcement date to 30 days prior to the current period announcement date is 1.8% higher if the manager choose not to disclose the earnings guidance than otherwise. Again, the nondisclosure premium declines after earnings announcements. The results are consistent with limited strategic thinking. Blue dots show the results for firm quarters with low number of analyst following. There is little nondisclosure premium and no change after earnings announcement.

The results found in this section suggest that limited attention is unlikely to be the reason that the nondisclosure premium exists. If the nondisclosure premium is mainly driven by limited attention, we should find much smaller premium for the sample with high number of analyst following than that for the sample with low number of analyst following, as analysts pay more attention to firms' information than unsophisticated investors. At this stage, we are not sure about the reason we could not find nondisclosure premium for the sample with low number of analyst following. We plan to provide further evidence on this issue.

3.6.6. Uncertainty

We now evaluate whether the bias changes with investor uncertainty. Our measure of uncertainty, return volatility in the prior quarter, captures the general uncertainty investors have about a firm.

Table G.6 present our subsample analysis by using regression model 2 with controls and year*quarter fixed effects. Figure H.9 plots the coefficients of NoGuide for the sub-sample results by using regression model 2 with controls and year*quarter fixed effects. We divides our sample into two sub-samples, one with return volatility being larger than the median of the reporting month and the other with return volatility being smaller than the median of the reporting month[49].

Table G.6 and Figure H.9 document that in the subsample of small return volatility, nondisclosure firm quarters have higher cumulative stock returns prior to the earnings announcement than dis- closure firm quarters. The difference significantly declines after the earnings announcement. The results suggest that investors under-react to nondisclosure in this subsample. However, we have weak evidence that the bias exists for the high return volatility group. While nondisclosure is associated with higher cumulative stock returns prior to the earnings announcement, the premium only slightly declines after the earnings announcement, suggesting investors think their beliefs prior to the earnings announcement were correct. We believe that if nondisclosure premium doesn't change even after the earnings announcement, it is likely to represent other types of biases or uncontrolled factors that affect our results.

[49] Specifically, we compare whether the return volatility of a firm is larger than the median return volatility of all firms that share the same fiscal quarter end dates. The idea is to compare whether the return volatility of a given firm is larger than those of other contemporaneous firms. We also use the overall distribution of return volatility for our subsample analysis, which also employs time series variation. The idea is then to compare whether the return volatility of a given firm is larger than those of all firms across all time.

Overall, we do not find clear pattern for the level of the subjective bias varying with investors' overall uncertainty. While nondisclosure is associated with higher cumulative stock returns in both subsamples, the lack of decline in the high return volatility group following the earnings announcement cautions against drawing strong inferences.

3.7. Robustness Tests

3.7.1 Size adjusted cumulative return

It is possible that changes in firm risk may affect stock returns as well as managers' disclosure choices. For the purpose of our study, we are interested in how investors perceive the underlying information that is withheld. If the information is related to change in firm risk, then we also want to include it as part of our analysis, because investors might not understand the change in risk correctly prior to the earnings announcement. Nevertheless, following prior research that focuses on unexpected returns, we use size adjusted cumulative stock return as the dependent variable to test the robustness of our results.

We construct size adjusted cumulative return as follows. First, we find the market value of equity at the end of June of each year[50]. Second, we merge the firm year level data with value weighted 3-size portfolio returns[51]. The portfolio return data also has size cut-off information, based on which we divide firms into three groups each year, size below 30 percentile, size

[50] We use June, because size portfolio returns are computed using size cut-offs formed in June each year. The sorting will be valid until the next June.

[51] We download the size portfolio return data from Professor French's data library on his home page: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data Library/det port form sz.html.

between 30 and 70 percentile and size above 70 percentile. Then, we subtract size portfolio return from the raw stock return and calculate the cumulative size adjusted return.

Table G.7 and Figure H.10 present the results by using regression model 2 and model 2, with controls and year*quarter fixed effects, respectively. From both of the table and the figure, we find very similar patterns with those using cumulative return as the dependent variable. Conditional on the same underlying news, nondisclosure firm quarters have higher size adjusted cumulative return and the nondisclosure premium drops significantly after the earnings surprise. According to what we find in this section, size risk is unlikely to drive our results.

3.7.2. Bundled management forecasts

In our main sample, we drop management forecasts issued from one day prior to the last period earnings announcement to 5 days subsequent to the last period earnings announcement. Such bundled forecasts create two problems. First, in cases of bundled forecasts in the *prior* quarter earnings announcement, capturing investors' reaction to disclosure (that is, bundled forecasts) requires disentangling (1) investors' reaction to prior quarter earnings surprise and (2) price adjustment to previous quarter bias if it exists. Second, in cases of bundled forecasts in the current quarter earnings announcement, capturing the bias related price adjustment following the earnings announcement requires disentangling the effects of the bundled forecasts. Both require additional assumptions on their relations with market returns.

Despite problems with bundled forecasts, to ensure the robustness of our results, we examine bundled forecasts. Our test utilizes the timing of market reaction to disclosure. We argue that the bias should mainly manifest around the time investors realize the existence of a disclosure decision. We predict that for bundled forecasts, the effect should concentrate early in the quarter, whereas for non-bundled forecasts, the effect should concentrate later in the quarter.

To capture the timing of the bias, we measure cumulative stock returns in two ways: (1) cumulative stock returns from the prior quarter earnings announcement to five days after the prior quarter earnings announcement, (2) cumulative stock returns from six days after the prior quarter earnings announcement to one day prior to the current quarter earnings announcement. We expect for bundled forecasts, the coefficient on NG to be positive for bundled forecast when using the cumulative stock return measured early in the quarter, and insignificant or negative when using the cumulative stock return measured late in the quarter. The coefficient can be negative, if investors have completed reacting to the bundled disclosure but are slow in realizing or responding to nondisclosure. For non-bundled forecasts, results from previous sections have demonstrated that coefficient on NG is positive when cumulative stock returns are measured from six days after the prior quarter earnings announcement. We do not have any predictions for non-bundled forecasts early in the quarter, because it depends on the relative timing of investors realizing the existence of a forecast in that quarter. Investors only observe nondisclosure for the first five days of the quarter for non-bundled forecasts.

Table G.8 shows the results, using observations from bad news firm quarters. From column (1), the coefficient on NG is positive and significant at 5% level. Relative to bundled forecasts, nondisclosure is on average associated with a 0.2% higher cumulative stock return for the five days subsequent to the previous quarter earnings announcement, conditional on the same quarterly earnings news. In contrast, column (2) shows that the coefficient on NG is negative when we measure cumulative stock returns from six days subsequent to the prior earnings announcement to one day prior to the current earnings announcement but the coefficient is insignificant at 10% level. Although the evidence is consistent with our predictions, we caution against over-interpreting the results. The results hold only to the extent that we sufficiently control for all confounding factors, which is unlikely. In addition, investors process a great amount information from earnings announcements, potentially adding noise to our results. Finally, the timing of reaction to nondisclosure can be ill captured. It is possible that investors' reaction to nondisclosure is more gradual than our ad-hoc five-day cut off.

Column 3 and 4 of Table G.8 suggests that the bias only manifests after five days subsequent to the previous quarter earnings announcement. The result is consistent with the idea that investors haven't observed a disclosure decision and do not differentiate disclosure and nondisclosure.

3.8. Predictable Return Under Limited Strategic Thinking

This section examines whether nondisclosure premium documented in the full sample can generate predictable returns. The basic idea is that if investors overprice nondisclosure prior to the earnings announcement and such overpricing is adjusted on and/or subsequent to the earnings announcement, a short position in stocks that did not issue management forecasts will on average earn a positive return.

We implement this idea by constructing calendar-quarter portfolios based on investors' knowledge of nondisclosure. In the first step, for each firm that announced earnings in a given calendar quarter, we sort firms into three buckets based on the size of its earnings surprise. We choose 30% and 70% as the cut-off points, which are computed based on the distribution of earnings surprise in the previous calendar quarter. This step aims to hedge out some post earnings announcement drifts. In the second step, conditional on each earnings surprise bucket, we short firms that did not issue management forecasts during the quarter and long firms that issued earnings forecasts. In cases where we cannot find any match for a nondisclosure firm, we short the nondisclosure firm for five days and long the next available disclosure firm for five days when it becomes available. We exclude firms that pre-announce earnings and firms that bundle forecasts with earnings announcement from the portfolio construction. Doing so means that our portfolios face very significant idiosyncratic risk, because we are left with fewer firms to select after dropping bundled forecasts. The advantage is that the stock price movement following earnings announcement is not confounded by the additional earnings forecasts from the earnings announcement.

Table G.9 shows the portfolio results. Consistent with LST, on average, the portfolio strategy earns a positive return of 1.2% five days subsequent to the portfolio formation. The portfolio is not without risk, because LST does not rule out the possibility that the portfolio strategy can earn negative returns for individual cases. For about 25% of the time, the strategy earns returns lower than -3%. But for about 25% of the time, the strategy earns returns higher than 5%.

3.9. Conclusion

In this paper, we examine whether investors in the capital market are surprised by the nondisclosure. Some economics studies find economic agents have cognitive constraints, which present them from correctly infer managers' private information from their strategic disclosure decisions. However, in capital market it is a repeated game over time, there are analysts and sophisticated institutional investors and realized earnings are disclosed each period. Therefore, it is an empirical question that whether investors in capital market subject to limited strategic thinking.

We test the existence of subjective bias by examining whether stock prices fully reflect the news that management strategically withholds. We features a nondisclosure premium that firm quarters with no voluntary earnings guidance have higher cumulative stock return than other firm quarters with guidance for the same news content. The positive correlation between disclosure and stock return exists among firm quarters with non-positive earnings surprise but not among firm quarters with positive earnings surprise. Besides, after realized earnings announcement date, the nondisclosure premium becomes insignificant.

There are a lot of factors affect managers disclosure choices and the stock return. If any factor affect both of managers disclosure choices and the stock return, the estimation would be biased. In order to address the omitted-variable problem, we control for some potential factors which may affect disclosure and stock return at the same time. We find that the nondisclosure premium still exists and even a little bit bigger conditional on these factors.

We contribute to the literature by showing that investors do not fully understand managers strategic actions even when such actions are repeated over time, when analysts and sophisticated institutional investors exist and when realized earnings are disclosed each period. Presenting this phenomenon also helps investors to adjust their expectation in the stock market and is useful for policy makers to set the best disclosure policy.

APPENDICES

Appendix A

Tables for Chapter 1

Table A.1: Summary Statistics

Variable Name	Variable Description	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
		Full Sample		Before Law Change		After Law Change	
Demographic							
Age	In years	35.84	7.85	32.83	8.45	35.86	7.84
White	= 1 if woman is white	0.88	0.32	0.96	0.19	0.88	0.33
Black	= 1 if woman is black	0.06	0.24	0.002	0.04	0.06	0.24
College	= 1 if woman attended college for at	0.50	0.50	0.39	0.49	0.50	0.50
Number of Child	Number of own children in household	1.68	1.26	2.13	1.63	1.68	1.26
Labor Force Participation							
Work	= 1 if woman is in the labor force	0.69	0.46	0.58	0.49	0.69	0.46
Full Time Job	= 1 if woman has full time job	0.52	0.50	0.41	0.49	0.53	0.50
Working Weeks[52]	Weeks worked last year	32.68	22.73	26.93	22.52	32.72	22.73
Working Hours	Hours worked last week	26.27	18.52	22.55	18.80	26.30	18.51
Salary Income	Wage and salary income	14606.95	23839.9	3812.53	5432.26	14698.22	23914.5
Other Income	Non-wage and salary income	1893.06	8330.93	553.48	2371.49	1904.39	8362.32
Household Income	Total household income	58821.93	58808.56	23071.09	14698.03	59124.22	58948.83

Notes: Sample is restricted to married, spouse present women between the age of 18 and 49 in CPS 1977-2012. The number of observation is 837726. Aggregate data used in this paper is constructed from this individual sample.

[52] The summary statistics of working hours, working weeks and salary income in Table A.1 are conditional on working.

Table A.2: Dynamic Effects of Unilateral Divorce Laws on Labor Force Participation Rate (Without state-specific time trends)

Specification:	(1) WLS	(2) OLS	(3) WLS, Cluster	(4) OLS, Cluster	(5) WLS, Log(LFP)	(6) WLS, Logit
1-2 years before	0.024	0.032**	0.024**	0.032**	0.040*	0.012
	(0.021)	(0.014)	(0.010)	(0.013)	(0.011)	(0.016)
0-2 years later	0.029	0.035**	0.029**	0.035**	0.041*	0.021
	(0.021)	(0.014)	(0.012)	(0.014)	(0.014)	(0.016)
3-4 years later	0.058***	0.064***	0.058***	0.064***	0.072***	0.052***
	(0.018)	(0.012)	(0.016)	(0.016)	(0.012)	(0.015)
5-6 years later	0.052***	0.061***	0.052***	0.061***	0.060***	0.052***
	(0.017)	(0.011)	(0.015)	(0.017)	(0.011)	(0.013)
7-8 years later	0.059***	0.055***	0.059***	0.055***	0.067***	0.060***
	(0.017)	(0.011)	(0.015)	(0.017)	(0.010)	(0.013)
9-10 years later	0.061***	0.058***	0.061***	0.058***	0.067***	0.063***
	(0.017)	(0.011)	(0.016)	(0.018)	(0.011)	(0.013)
11-12 years later	0.056***	0.053***	0.056***	0.053**	0.059***	0.060***
	(0.017)	(0.011)	(0.017)	(0.021)	(0.010)	(0.013)
13-14 years later	0.054***	0.053***	0.054***	0.053**	0.056***	0.060***
	(0.017)	(0.011)	(0.018)	(0.021)	(0.010)	(0.013)
> 15 years later	0.045***	0.060***	0.045**	0.060***	0.045**	0.054***
	(0.017)	(0.010)	(0.020)	(0.022)	(0.010)	(0.013)
Observations	1.836	1.836	1.836	1.836	1.836	1.836
R-squared	0.876	0.860	0.876	0.860	0.876	0.865

Notes: Standard errors in parentheses. All regressions based on aggregate level data constructed from CPS 1977-2012. Sample restricted to married women age 18 to 49. Dependent variable is LFP rates. Control variables include: share of each age group in each state and year, share of each race group and education group in each state and year, state fixed effects and time fixed effects. The results in column (5) and column (6) are marginal effects and the standard errors come from Bootstrapping. ***significant at 1% **5% *10%.

Table A.3: Dynamic Effects of Unilateral Divorce Laws on Labor Force Participation Rates (With state-specific time trends)

	(1)	(2)	(3)	(4)	(5)	(6)
Specification:	WLS	OLS	WLS, Cluster	OLS, Cluster	WLS, Log(LFP)	WLS, Logit
1-2 years before	0.061***	0.055***	0.061***	0.055***	0.070***	0.056**
	(0.021)	(0.013)	(0.013)	(0.016)	(0.016)	(0.014)
0-2 years later	0.074***	0.063***	0.074***	0.063***	0.079***	0.070***
	(0.022)	(0.014)	(0.018)	(0.022)	(0.017)	(0.015)
3-4 years later	0.101***	0.095***	0.101***	0.095***	0.109***	0.095***
	(0.021)	(0.015)	(0.018)	(0.020)	(0.018)	(0.016)
5-6 years later	0.099***	0.098***	0.099***	0.098***	0.103***	0.097***
	(0.021)	(0.015)	(0.020)	(0.022)	(0.017)	(0.016)
7-8 years later	0.104***	0.095***	0.104***	0.095***	0.107***	0.103***
	(0.022)	(0.015)	(0.022)	(0.025)	(0.018)	(0.017)
9-10 years later	0.108***	0.098***	0.108***	0.098***	0.109***	0.106***
	(0.022)	(0.016)	(0.023)	(0.024)	(0.018)	(0.017)
11-12 years later	0.104***	0.097***	0.104***	0.097***	0.103***	0.105***
	(0.022)	(0.016)	(0.022)	(0.023)	(0.019)	(0.017)
13-14 years later	0.106***	0.099***	0.106***	0.099***	0.103***	0.108***
	(0.022)	(0.016)	(0.023)	(0.023)	(0.019)	(0.017)
> 15 years later	0.107***	0.111***	0.107***	0.111***	0.103***	0.111***
	(0.023)	(0.017)	(0.025)	(0.025)	(0.020)	(0.018)
Observations	1.836	1.836	1.836	1.836	1.836	1.836
R-squared	0.911	0.900	0.911	0.900	0.916	0.901

Notes: See notes of Table A.2. Control variables also include state-specific time trends. ***significant at 1% **5% *10%.

Table A.4: The Effects of Unilateral Divorce Laws on Labor Force Participation Rates: Different Specifications

Specification:	(1) Stevenson, 2008 column (3), Table 5	(2) Stevenson, 2008 column (4), Table 5	(3) Stevenson, column (3), Table 5: Replication	(4) Add Controls	(5) Add Controls & State Time Trend	(6) Add Controls & State Time Trend, 18-49 Women	(7) Add Controls & State Time Trend, 18-49 Women, 1968-2012
1-3 years prior to change	-0.003	0.000	0.003	0.003	-0.003	0.009	0.026*
	(0.010)	(0.009)	(0.010)	(0.008)	(0.011)	(0.013)	(0.014)
Year of change	-0.002	0.017*	0.001	0.022**	0.015	0.031*	0.055**
	(0.013)	(0.010)	(0.011)	(0.010)	(0.015)	(0.017)	(0.023)
1-3 years later	0.001	0.017**	0.009	0.030***	0.023	0.041**	0.064***
	(0.010)	(0.008)	(0.009)	(0.008)	(0.015)	(0.019)	(0.018)
4-6 years later	0.010	0.027***	0.017*	0.041***	0.026	0.057**	0.086***
	(0.009)	(0.008)	(0.009)	(0.008)	(0.019)	(0.024)	(0.023)
7-9 years later	0.004	0.026***	0.009	0.038***	0.017	0.048*	0.082***
	(0.010)	(0.008)	(0.009)	(0.008)	(0.020)	(0.025)	(0.025)
10 years or more later	0.016	0.027***	0.012	0.043***	0.015	0.040	0.083***
	(0.008)	(0.009)	(0.010)	(0.009)	(0.021)	(0.026)	(0.025)
Observations	1116	1116	1.116	969	969	969	1.836
R-squared	/	/	0.903	0.896	0.922	0.912	0.898

Notes: Standard errors in parentheses. All regressions based on aggregate level data. Sample restricted to married women age 14 or older. Dependent variable is LFP rates for all columns. Standard errors are robust as those in Steven (2008)'s paper. The results in column (1) are in column (3), Table 5 of Stevenson, 2008 paper. In column (2) I use the same data that are 1968-1995 CPS data to replicate the results. In column (3), I add control variables that include: share of each age group in each state and year, share of each race group and education group in each state and year, state fixed effects and time fixed effects. In column (4) I add state-specific time trends. In column (5) I restrict the sample to 18-49 years old women and in column (5) I use observations from 1995 to 2012. ***significant at 1% **5% *10%.

Table A.5: Dynamic Effects of Unilateral Divorce Laws on Weeks and Hours of Work and LFP of Full Time Job

Specification:	(1) OLS, Cluster Weeks Worked, Unconditional	(2) OLS, Cluster Weeks Worked, Conditional	(3) OLS, Cluster Hours Worked, Unconditional	(4) OLS, Cluster Hours Worked, Conditional	(5) OLS, Cluster LFP of Full Time Job
1-2 years before	1.482**	0.090	0.802	-0.445	0.006
	(0.659)	(0.216)	(0.630)	(0.298)	(0.020)
0-2 years later	2.400***	0.726	1.825***	0.598	0.027
	(0.820)	(0.696)	(0.538)	(0.493)	(0.020)
3-4 years later	2.938***	0.379	2.009***	0.003	0.037
	(1.080)	(0.701)	(0.659)	(0.367)	(0.026)
> 5 years later	3.336***	0.352	2.433***	0.259	0.043
	(1.010)	(0.646)	(0.690)	(0.473)	(0.028)
Observations	1.836	1.836	1.836	1.836	1.836
R-squared	0.922	0.908	0.884	0.822	0.853

Notes: See notes of Table 2. Dependent variables are weeks worked last year unconditional on participation in the labor force in column (1), weeks worked last year conditional on participation in the labor force in column (2), usual hours worked per week last year unconditional on participation in the labor force in column in column (3), usual hours worked per week last year conditional on participation in the labor force in column in column (4) and LFP of full time job in column (5). Control variables also include state-specific time trends. ***significant at 1% **5% *10%.

Table A.6: Dynamic Effects of Unilateral Divorce Laws on Married Mothers and Non-Mothers

Specification:	(1) OLS, Cluster LFP
0-2 years later	0.056***
	(0.017)
3-4 years later	0.052***
	(0.015)
> 5 years later	0.045***
	(0.012)
0-2 years later* child under 6	-0.055
	(0.034)
3-4 years later* child under 6	-0.023
	(0.029)
> 5 years later* child under 6	0.004
	(0.012)
0-2 years later* child 6-18	-0.007
	(0.031)
3-4 years later* child 6-18	0.003
	(0.018)
> 5 years later* child 6-18	0.010
	(0.006)
Observations	5.508
R-squared	0.839

Notes: Standard errors in parentheses. All regressions based on aggregate level data from CPS 1977-2012. Sample restricted to married women age 18 to 49. Dependent variables are LFP rates in Control variables include: age, age square, race, education, non-labor income, non-labor income squared, state fixed effects, time fixed effects and state-specific time trends. ***significant at 1% **5% *10%.

Table A.7: Dynamic Effects of Unilateral Divorce Laws on Women with High and Low Education

	(1)	(2)	(3)	(4)	(5)	(6)
	Women with Low Education	Women with High Education	Younger Women	Older Women	White Women	Black Women
0-2 years later	0.059***	0.030	0.052**	0.033***	0.041***	0.002
	(0.010)	(0.024)	(0.022)	(0.011)	(0.015)	(0.098)
3-4 years later	0.085***	0.069**	0.078***	0.058***	0.067***	0.111
	(0.013)	(0.027)	(0.024)	(0.011)	(0.018)	(0.200)
> 5 years later	0.093***	0.050	0.072**	0.062***	0.064***	0.113
	(0.015)	(0.031)	(0.027)	(0.014)	(0.021)	(0.157)
Observations	1.836	1.836	1.836	1.836	1.836	1.687
R-squared	0.809	0.752	0.796	0.848	0.893	0.225

Notes: See notes of Table 2. Dependent variables are LFP rates. Control variables also include state-specific time trends. ***significant at 1% **5% *10%.

Appendix B

Figures for Chapter 2

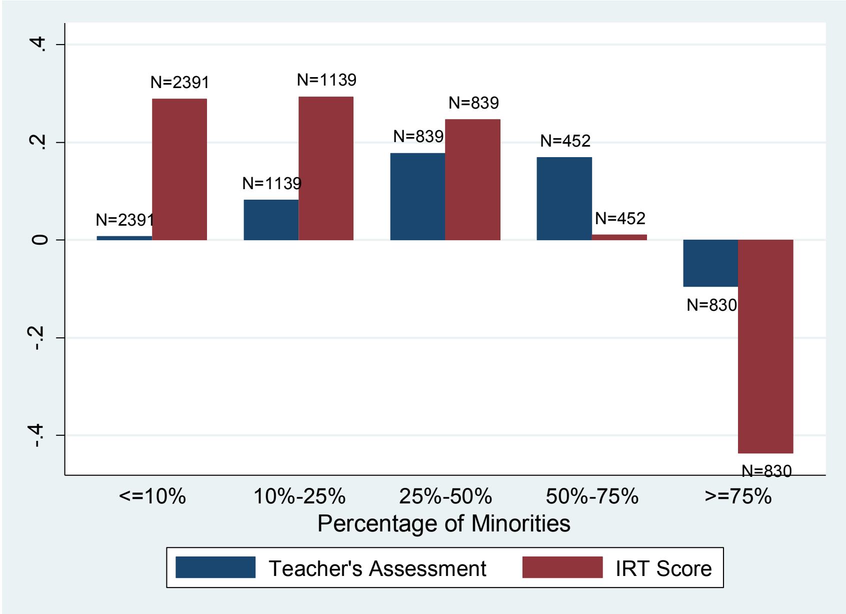


Figure B.1: Grade 3 Math IRT Scores and Teachers' Subjective Assessments by Racial Composition of Schools, All Students.

Notes: At first, I restrict my sample to be the same as I use when running the regression for externalizing behavior in grade 3. Then I further restrict the sample to the students who have valid information for racial composition of schools they attend. The racial composition variable is a categorical variable in ECLS-K: the percentage of minorities is less than 10 percent, 10–25 percent, 25–50 percent, 50–75 percent, and larger than 75 percent. The IRT scores and teachers' subjective assessments are all standardized. The number of observations are shown in the graph.

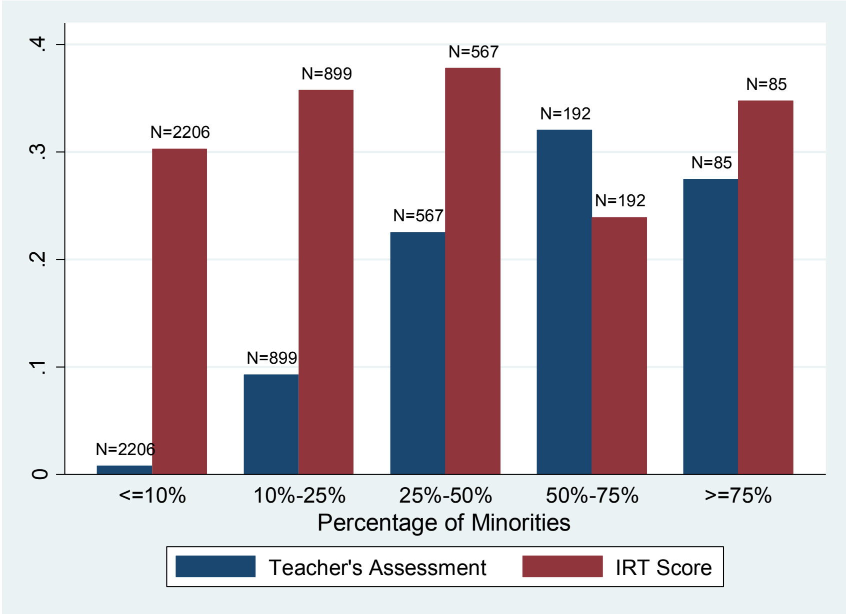


Figure B.2: Grade 3 Math IRT Scores and Teachers' Subjective Assessments by Racial Composition of Schools, White Students.

Notes: The sample restriction is the same as that in Figure 1, except that the observations here are only for white students. The IRT scores and teachers' subjective assessments are all standardized. The number of observations are shown in the graph.

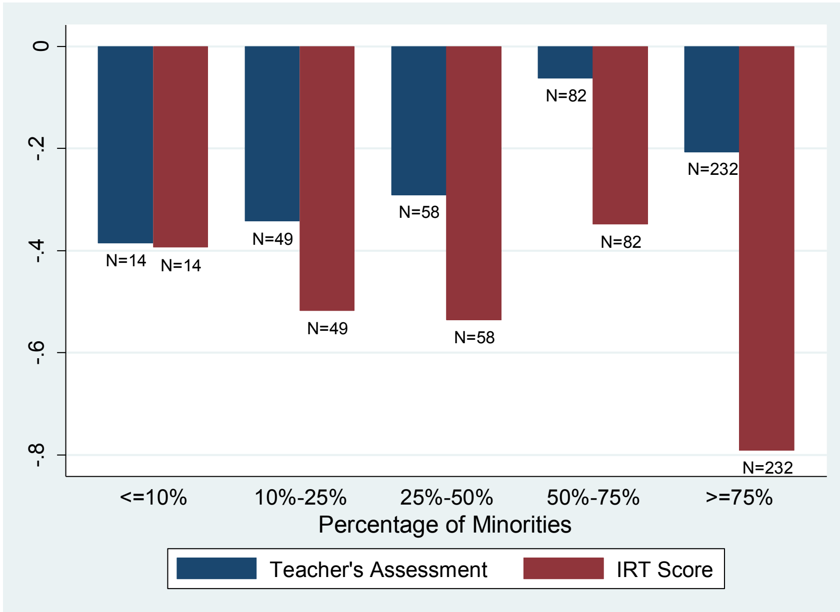


Figure B.3: Grade 3 Math IRT Scores and Teachers' Subjective Assessments by Racial Composition of Schools, Black Students.

Notes: The sample restriction is the same as that in Figure 1, except that the observations here are only for black students. The IRT scores and teachers' subjective assessments are all standardized. The number of observations is shown in the graph.

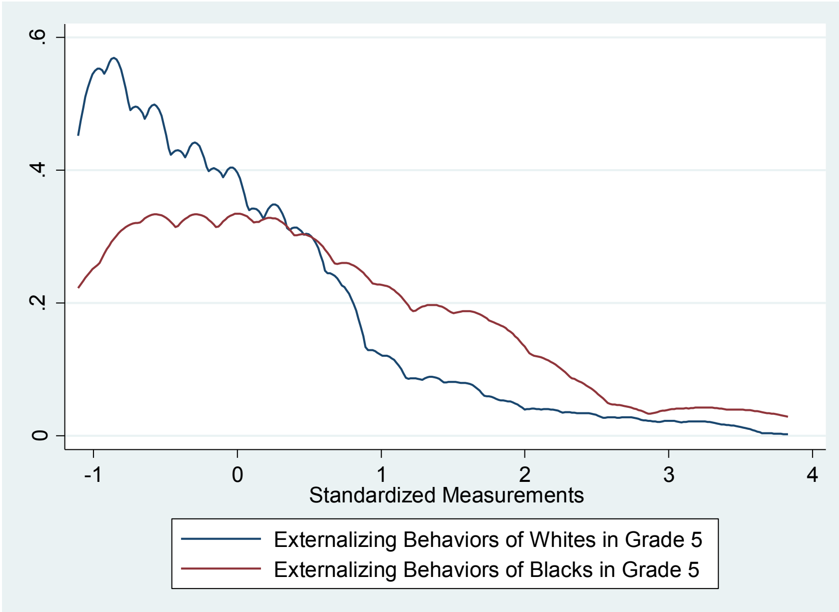


Figure B.4: Distribution of Externalizing Behavior for White and Black Students in Grade 5.
Notes: The sample is restricted to all white and black students who have valid information for externalizing behaviors in grade 5. The measurements of externalizing behaviors are all standardized.

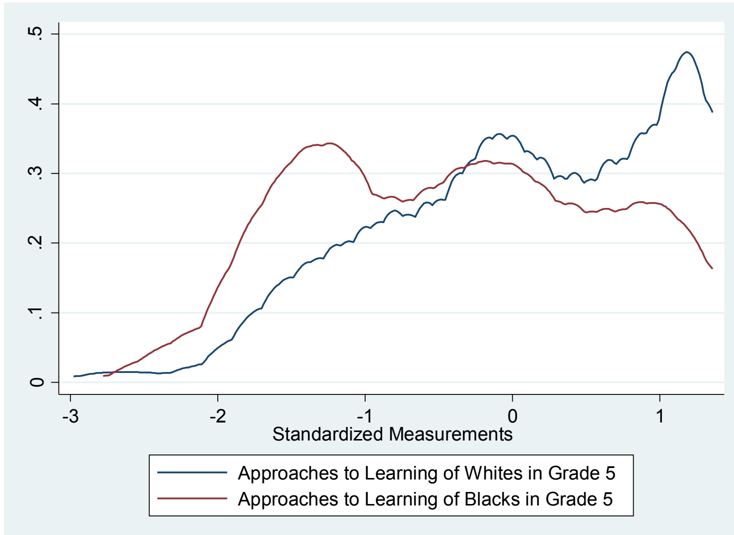


Figure B.5: Distribution of Approaches to Learning for White and Black Students in Grade 5.
Notes: The sample is restricted to all white and black students who have valid information for approaches to learning in grade 5. The measurements of externalizing behaviors are all standardized.

Appendix C

Tables for Chapter 2

Table C.1: Descriptive Statistics

Variables	Full Sample (N=6,859)	Whites (N=4,553)	Blacks (N=572)
Female	0.481	0.476	0.468
	(0.500)	(0.499)	(0.499)
Age (fall kindergarten): month	68.483	68.779	68.361
	(4.230)	(4.247)	(4.209)
Birth weight: pounds	7.376	7.473	7.069
	(1.326)	(1.347)	(1.342)
Region			
Northeast	0.180	0.221	0.122
	(0.384)	(0.415)	(0.327)
Midwest	0.231	0.290	0.149
	(0.422)	(0.454)	(0.357)
South	0.389	0.343	0.675
	(0.488)	(0.475)	(0.469)
West	0.200	0.147	0.054
	(0.400)	(0.354)	(0.226)
Location			
Central city	0.367	0.264	0.505
	(0.482)	(0.441)	(0.500)
Urban fringe and large town	0.421	0.488	0.321
	(0.494)	(0.500)	(0.467)
Small town and rural	0.212	0.248	0.174
	(0.409)	(0.432)	(0.380)
Families' warmth index<median	0.525	0.518	0.514
	(0.499)	(0.500)	(0.500)
Spanked child last week	0.273	0.240	0.357
	(0.445)	(0.427)	(0.480)
Mother's age at first birth less	0.256	0.182	0.455
	(0.437)	(0.386)	(0.498)
Family Structure:			
Two biological parents	0.687	0.775	0.349
	(0.464)	(0.417)	(0.477)
Single mother	0.176	0.111	0.447
	(0.381)	(0.314)	(0.498)
SES:			
1st Quartile (lowest)	0.174	0.068	0.304
	(0.379)	(0.251)	(0.460)
2nd Quartile	0.188	0.161	0.277
	(0.390)	(0.368)	(0.448)
3rd Quartile	0.189	0.190	0.214
	(0.392)	(0.392)	(0.411)
4th Quartile	0.220	0.260	0.152
	(0.414)	(0.439)	(0.360)
5th Quartile (Highest)	0.230	0.322	0.052
	(0.421)	(0.467)	(0.223)

Notes: Sample is restricted to children with non-missing values for female dummy variables, race, weights, and each outcome cited.

Table C.2: Teacher Ratings of Non-Cognitive Skills, Retention, and Suspension

	Whites	Blacks	Difference
Externalizing Behaviors:			
Fall Kindergarten	-0.056	0.314	0.370***
	(0.970)	(1.104)	[0.078]
Spring Kindergarten	-0.083	0.392	0.475***
	(0.977)	(1.094)	[0.074]
Grade 1	-0.061	0.359	0.420***
	(0.967)	(1.108)	[0.074]
Grade 3	-0.079	0.469	0.548***
	(0.956)	(1.123)	[0.082]
Grade 5	-0.066	0.445	0.511***
	(0.942)	(1.171)	[0.083]
Approaches to Learning:			
Fall Kindergarten	0.090	-0.319	-0.409***
	(0.989)	(0.993)	[0.064]
Spring Kindergarten	0.095	-0.350	-0.445***
	(0.976)	(1.042)	[0.070]
Grade 1	0.069	-0.340	-0.409***
	(0.965)	(1.057)	[0.075]
Grade 3	0.059	-0.329	-0.388***
	(0.995)	(1.019)	[0.074]
Grade 5	0.085	-0.364	-0.449***
	(0.977)	(1.012)	[0.070]
Self-Control:			
Fall Kindergarten	0.089	-0.354	-0.444***
	(1.000)	(0.982)	[0.065]
Spring Kindergarten	0.106	-0.400	-0.506***
	(0.967)	(1.078)	[0.073]
Grade 1	0.074	-0.312	-0.385***
	(0.985)	(1.039)	[0.073]
Grade 3	0.073	-0.392	-0.465***
	(0.970)	(1.072)	[0.078]
Grade 5	0.075	-0.406	-0.481***
	(0.951)	(1.122)	[0.081]
Interpersonal Skills:			
Fall Kindergarten	0.077	-0.262	-0.339***
	(1.020)	(0.957)	[0.064]
Spring Kindergarten	0.094	-0.276	-0.370***
	(0.980)	(1.052)	[0.072]
Grade 1	0.048	-0.277	-0.325***
	(0.999)	(1.021)	[0.074]
Grade 3	0.056	-0.292	-0.348***
	(0.997)	(1.045)	[0.077]
Grade 5	0.053	-0.333	-0.385***
	(1.001)	(1.047)	[0.075]
Internalizing Problems:			
Fall Kindergarten	-0.012	-0.007	0.005
	(1.005)	(0.953)	[0.062]
Spring Kindergarten	-0.023	0.065	0.088
	(0.977)	(1.089)	[0.074]
Grade 1	-0.021	0.137	0.158**

Table C.2 (cont'd)

	(0.976)	(1.150)	[0.080]
Grade 3	-0.007	0.189	0.196**
	(0.969)	(1.129)	[0.085]
Grade 5	0.022	0.018	-0.003
	(1.032)	(0.946)	[0.066]
Ever been retained (from fall kindergarten to Grade			
	0.104	0.230	0.126***
	(0.305)	(0.421)	[0.029]
In/Out of School Suspension in Grade 8			
	0.130	0.346	0.216***
	(0.336)	(0.476)	[0.031]

Notes: Summary statistics are based on children with non-missing values for female dummy variables, race, weights, and each outcome cited. Measurements of non-cognitive skills are standardized to have a mean of zero and standard deviation of one in the weighted sample. Please refer to the text for sample restrictions. Observations are weighted using eighth grade parent panel weights. Robust standard errors are reported for differences in the means across genders. ***Significant at 1 percent, **5 percent, *10 percent.

Table C.3: Racial Gaps in Non-Cognitive Skills:

Externalizing Behavior in Grade 5, Approaches to Learning in Grade 5, and Suspension in Grade 8

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Externalizing Behavior in Grade 5			Approaches to Learning in Grade 5			Suspension in Grade 8		
Black	0.510***	0.213**	0.193**	-0.441***	-0.106	-0.111	0.215***	0.119***	0.118***
	[0.081]	[0.085]	[0.082]	[0.064]	[0.072]	[0.072]	[0.029]	[0.032]	[0.031]
Hispanic	-0.053	-0.212***	-0.229***	-0.111**	0.091	0.090	0.014	-0.031	-0.040*
	[0.042]	[0.056]	[0.057]	[0.046]	[0.058]	[0.057]	[0.016]	[0.021]	[0.021]
Asian	-0.329***	-0.319***	-0.331***	0.445***	0.462***	0.464***	-0.076***	-0.072***	-0.067***
	[0.069]	[0.071]	[0.069]	[0.055]	[0.057]	[0.059]	[0.015]	[0.017]	[0.017]
Other	0.112	-0.021	-0.037	-0.182**	-0.036	-0.030	0.076**	0.045	0.048
	[0.076]	[0.074]	[0.075]	[0.083]	[0.078]	[0.078]	[0.034]	[0.032]	[0.032]
Female	-0.507***	-0.490***	-0.497***	0.610***	0.587***	0.593***	-0.155***	-0.155***	-0.155***
	[0.041]	[0.039]	[0.038]	[0.039]	[0.036]	[0.036]	[0.015]	[0.014]	[0.014]
Age at Assessment at fall kindergarten	-0.340***	-0.341***	-0.323***	0.415***	0.381***	0.421***	-0.010	-0.008	-0.016
	[0.125]	[0.115]	[0.117]	[0.127]	[0.115]	[0.121]	[0.052]	[0.049]	[0.050]
Age-squared	0.003***	0.003***	0.002***	-0.003***	-0.003***	-0.003***	0.000	0.000	0.000
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.000]	[0.000]	[0.000]
Single Mom		0.223***	0.225***		-0.171***	-0.171***		0.075***	0.080***
		[0.064]	[0.064]		[0.054]	[0.054]		[0.025]	[0.025]
Other Family Structure		0.342***	0.332***		-0.264***	-0.259***		0.092***	0.089***
		[0.073]	[0.071]		[0.069]	[0.068]		[0.029]	[0.029]
Age at First Birth < 20		0.277***	0.265***		-0.184***	-0.181***		0.104***	0.101***
		[0.060]	[0.059]		[0.052]	[0.052]		[0.022]	[0.022]
1st SES Quintile		0.217***	0.213***		-0.519***	-0.512***		0.055*	0.047
		[0.076]	[0.076]		[0.068]	[0.069]		[0.029]	[0.029]
2nd SES Quintile		0.181***	0.193***		-0.330***	-0.336***		0.029	0.025
		[0.063]	[0.062]		[0.062]	[0.063]		[0.023]	[0.022]
3rd SES Quintile		0.143***	0.141***		-0.223***	-0.223***		0.059***	0.056***

Table C. 3 (cont'd)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		[0.053]	[0.052]		[0.053]	[0.053]		[0.022]	[0.021]
4th SES Quintile		0.083*	0.080*		-0.128**	-0.126**		0.009	0.009
		[0.048]	[0.048]		[0.051]	[0.052]		[0.018]	[0.017]
Birth Weight		0.019	0.017		-0.008	-0.007		0.005	0.005
		[0.014]	[0.014]		[0.015]	[0.015]		[0.006]	[0.005]
School Environment	No	No	Yes	No	No	Yes	No	No	Yes
Geographic Location	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Observations	6.261	6.261	6.261	6.300	6.300	6.300	6.756	6.756	6.756
R-squared	0.117	0.171	0.185	0.136	0.201	0.207	0.093	0.138	0.150

Notes: Each column corresponds to a different regression that controls for various subsets of variables. Columns (1) to (3) show the racial gap of externalizing behavior in grade 5, and columns (4) to (6) show the racial gap of approaches to learning in grade 5. Columns (7) to (9) show the racial gap of suspension in grade 8. Columns (1), (4), and (7) control for race, female dummy variables, and age at assessment at fall kindergarten, age-squared. Columns (2), (5), and (8) add family-quality variables that include mother's age at first birth and family SES, birth weight, and geographic location variables, including dummies for region and urbanicity. Columns (3), (6), and (9) include school environments. Sample is restricted to those with non-missing observations on racial dummies, gender, weights, and specific noncognitive skill measures. Observations are weighted using eighth grade parent panel weights. Robust standard errors are reported. ***Significant at 1 percent level, **5 percent level, and *10 percent level.

Table C.4: The Evolution of Racial Gaps in Non-Cognitive Skills

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Externalizing Behavior					Approaches to Learning				
	Fall-K	Spring-K	Grade 1	Grade 3	Grade 5	Fall-K	Spring-K	Grade 1	Grade 3	Grade 5
Black	0.175**	0.232***	0.311***	0.362***	0.193**	-0.130*	-0.103	-0.190**	-0.139*	-0.111
	[0.079]	[0.081]	[0.079]	[0.083]	[0.082]	[0.069]	[0.070]	[0.078]	[0.080]	[0.072]
Hispanic	-0.134**	-0.114*	-0.082	-0.152**	-0.229***	0.104*	0.095*	0.104*	0.095*	0.090
	[0.062]	[0.059]	[0.063]	[0.069]	[0.057]	[0.055]	[0.055]	[0.058]	[0.056]	[0.057]
Asian	-0.247***	-0.189**	-0.177**	-0.270***	-0.331***	0.167**	0.201***	0.213**	0.292***	0.464***
	[0.079]	[0.087]	[0.083]	[0.072]	[0.069]	[0.077]	[0.071]	[0.089]	[0.079]	[0.059]
School Environment	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Location	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Background Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age and Gender	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6.678	6.,606	6.246	5.741	6.261	6.759	6.629	6.275	5.760	6.300
R-squared	0.159	0.162	0.130	0.151	0.185	0.180	0.202	0.166	0.177	0.207

Notes: Each column corresponds to a different regression, which has different independent variables. The control variables in all of these regressions are the same as those in column (3), Table 3. They are race, female dummy variables, age at assessment at fall kindergarten, age-squared, and background controls, which include family-quality variables, birth weight, geographic location variables, and school environments. Please refer to Table 3 for detailed information about these control variables. Sample is restricted to those with non-missing observations on racial dummies, gender, weights, and specific non-cognitive skill measures. Observations are weighted using eighth grade parent panel weights. Robust standard errors reported. ***Significant at 1 percent level, **5 percent level, and *10 percent level.

Table C.5: The Factors of the Difference between IRT Scores and Teachers' Subjective Assessments

	(1)	(2	(3)	(4)	(5)
			Index of Bias		
	Fall-K	Spring-K	Grade 1	Grade 3	Grade 5
Black	0.139**	0.178***	0.181***	0.521***	0.365***
	[0.054]	[0.052]	[0.056]	[0.065]	[0.063]
Hispanic	0.040	0.085**	0.212***	0.326***	0.305***
	[0.052]	[0.035]	[0.037]	[0.041]	[0.042]
Asian	-0.347***	-0.179**	-0.055	0.259***	0.267***
	[0.103]	[0.075]	[0.070]	[0.064]	[0.059]
Other	0.092	0.038	0.065	0.185**	0.173***
	[0.081]	[0.062]	[0.061]	[0.079]	[0.063]
Female	0.126***	0.105***	0.091***	0.131***	0.167***
	[0.039]	[0.031]	[0.031]	[0.033]	[0.035]
Observations	4.639	6.269	5.956	5.331	6.203
R-squared	0.014	0.014	0.021	0.073	0.044

Notes: The control variables are race and female dummy variables. The sample is restricted to observations that have valid information for female, race, weights and index of bias. Observations are weighted using eighth grade parent panel weights. Robust standard errors reported. ***Significant at 1 percent level, **5 percent level, and *10 percent level.

Table C.6: The Factors of the Difference between IRT Scores and Teachers' Subjective Assessments: with Control Variables

	(1)	(2)	(3)	(4)	(5)
			Index of Bias		
	Fall-K	Spring-K	Grade 1	Grade 3	Grade 5
Black	0.079	0.101*	0.036	0.359***	0.204***
	[0.065]	[0.061]	[0.057]	[0.066]	[0.067]
Hispanic	-0.012	0.013	0.130***	0.179***	0.087*
	[0.053]	[0.040]	[0.042]	[0.050]	[0.049]
Asian	-0.397***	-0.207***	-0.097	0.199***	0.180***
	[0.092]	[0.073]	[0.066]	[0.066]	[0.064]
Other	0.040	-0.010	0.021	0.125	0.074
	[0.080]	[0.060]	[0.062]	[0.078]	[0.062]
Female	0.111***	0.103***	0.085***	0.131***	0.177***
	[0.038]	[0.030]	[0.030]	[0.032]	[0.034]
Single Mom	0.130**	0.046	0.064	0.008	-0.045
	[0.053]	[0.047]	[0.048]	[0.049]	[0.049]
Other Family	0.013	0.065	0.021	0.077	-0.069
	[0.055]	[0.055]	[0.053]	[0.058]	[0.063]
Age at First Birth < 20	-0.025	0.019	0.058	0.013	-0.043
	[0.049]	[0.040]	[0.043]	[0.045]	[0.048]
1st SES Quintile	0.122	0.073	0.078	0.244***	0.372***
	[0.078]	[0.060]	[0.054]	[0.060]	[0.065]
2nd SES Quintile	0.232***	0.094*	0.158***	0.172***	0.189***
	[0.065]	[0.052]	[0.050]	[0.054]	[0.055]
3rd SES Quintile	0.239***	0.168***	0.185***	0.094**	0.109**
	[0.059]	[0.047]	[0.045]	[0.047]	[0.050]
4th SES Quintile	0.180***	0.144***	0.097**	0.035	0.132***
	[0.058]	[0.047]	[0.039]	[0.045]	[0.047]
	[0.060]	[0.044]	[0.039]	[0.044]	[0.047]
Observations	4.639	6.269	5.956	5.331	6.203
R-squared	0.065	0.059	0.063	0.113	0.093

Table C.6 (con'td)

Notes: The control variables in all of these regressions are the same as those in column (3), Table 3. Other control variables are race, female dummy variables, age at assessment at fall-kindergarten, age-squared, and background controls, which include family-quality variables, birth weight, geographic location variables, and school environments. Please refer to Table 3 for detailed information about these control variables. The sample is restricted to observations that have valid information for female, race, weights, and index of bias. Observations are weighted using eighth grade parent panel weights. Robust standard errors reported. ***Significant at 1 percent level, **5 percent level, and *10 percent level.

Table C.7: The Evolution of Racial Gaps in Non-Cognitive Skills: School Fixed Effects and Teachers' Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Externalizing Behavior					Approaches to Learning				
	Fall-K	Spring-K	Grade 1	Grade 3	Grade 5	Fall-K	Spring-K	Grade 1	Grade 3	Grade 5
Black	0.399***	0.462***	0.476***	0.412***	0.384***	-0.214**	-0.198**	-0.248***	-0.380***	-0.247**
	[0.124]	[0.088]	[0.099]	[0.099]	[0.102]	[0.096]	[0.088]	[0.096]	[0.091]	[0.096]
Hispanic	-0.061	-0.121*	-0.037	-0.111	-0.210***	0.024	0.123**	0.050	0.119*	0.148**
	[0.071]	[0.069]	[0.067]	[0.071]	[0.071]	[0.064]	[0.059]	[0.063]	[0.063]	[0.071]
Asian	-0.275***	-0.280***	-0.065	-0.270***	-0.354***	0.215**	0.218**	0.253***	0.417***	0.515***
	[0.081]	[0.083]	[0.094]	[0.078]	[0.096]	[0.090]	[0.089]	[0.093]	[0.100]	[0.091]
Geographic Location	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Background Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age and Gender	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Teacher Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5.813	5.708	5.690	5.342	5.924	5.880	5.731	5.715	5.350	5.957
R-squared	0.404	0.431	0.450	0.556	0.623	0.457	0.449	0.475	0.529	0.597

Table C.8: The Evolution of Racial Gaps in Cognitive Skills: School Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Reading					Math				
	Fall-K	Spring-K	Grade 1	Grade 3	Grade 5	Fall-K	Spring-K	Grade 1	Grade 3	Grade 5
Panel A: IRT Scores										
Black	-0.117*	-0.056	-0.115	-0.286***	-0.434***	-0.180***	-0.273***	-0.418***	-0.529***	-0.505***
	[0.062]	[0.068]	[0.071]	[0.098]	[0.087]	[0.068]	[0.071]	[0.068]	[0.098]	[0.101]
Hispanic	-0.174***	-0.104*	-0.053	-0.115*	-0.013	-0.246***	-0.189***	-0.205***	-0.196***	-0.071
	[0.057]	[0.055]	[0.057]	[0.064]	[0.061]	[0.053]	[0.055]	[0.052]	[0.065]	[0.067]
Asian	0.233*	0.336**	0.447***	0.088	0.063	0.213*	0.198*	0.103	0.065	0.172*
	[0.138]	[0.151]	[0.110]	[0.094]	[0.092]	[0.123]	[0.107]	[0.106]	[0.110]	[0.096]
Observations	5,598	5,666	5,682	5,340	5,963	5,799	5,789	5,732	5,355	5,968
R-squared	0.463	0.447	0.527	0.594	0.649	0.510	0.503	0.512	0.574	0.642
Panel B: Teachers' Subjective Assessments										
Black	-0.002	-0.140*	-0.288***	-0.330***	-0.430***	0.095	-0.102	-0.309***	-0.304***	-0.214
	[0.080]	[0.082]	[0.090]	[0.101]	[0.081]	[0.086]	[0.084]	[0.089]	[0.089]	[0.139]
Hispanic	-0.238***	-0.183***	-0.073	-0.094	-0.060	-0.102	-0.135**	-0.110*	-0.077	0.114
	[0.062]	[0.062]	[0.066]	[0.072]	[0.066]	[0.065]	[0.060]	[0.058]	[0.069]	[0.108]
Asian	-0.049	0.021	0.167	0.129	0.220**	0.233**	0.151*	0.180*	0.300***	0.354***
	[0.102]	[0.091]	[0.109]	[0.113]	[0.093]	[0.101]	[0.090]	[0.098]	[0.116]	[0.108]
Observations	5,463	5,743	5,717	5,216	5,921	4,441	5,678	5,663	5,106	2,936
R-squared	0.552	0.507	0.467	0.518	0.603	0.623	0.534	0.522	0.570	0.681
School Environment	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Location	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Background Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age and Gender	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Teacher Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table C.8 (cont'd)

Notes: Each panel corresponds to a different adjusted method. Each column corresponds to a different regression that has different independent variables. All of these regressions have the same control variables as those in column (3), Table 3. They are race, female dummy variables, age at assessment at fall-kindergarten, age-squared, and background controls, which include family-quality variables, birth weight, geographic location variables, and school environments. Please refer to Table 3 for detailed information about these control variables. Sample is restricted to those with non-missing observations on racial dummies, gender, weights, and specific adjusted non-cognitive skill measures. Observations are weighted using eighth grade parent panel weights. Robust standard errors reported. ***Significant at 1 percent level, **5 percent level, and *10 percent level.

Table C.9: The Evolution of Racial Gaps in Non-Cognitive Skills: School Fixed Effects and Teachers' Characteristics

	(1)	(2)	(3)	(4)	(5)
	Principal component analysis				
	Fall-K	Spring-K	Grade 1	Grade 3	Grade 5
Black	-0.593***	-0.546***	-0.681***	-0.863***	-0.481***
	[0.185]	[0.157]	[0.171]	[0.186]	[0.170]
Hispanic	0.043	0.197*	0.120	0.322**	0.425***
	[0.129]	[0.107]	[0.110]	[0.125]	[0.128]
Asian	0.403***	0.430***	0.209	0.580***	0.888***
	[0.134]	[0.124]	[0.186]	[0.180]	[0.180]
Geographic Location	Yes	Yes	Yes	Yes	Yes
Background Controls	Yes	Yes	Yes	Yes	Yes
Age and Gender	Yes	Yes	Yes	Yes	Yes
School Fixed Effects	Yes	Yes	Yes	Yes	Yes
Teacher Characteristics	Yes	Yes	Yes	Yes	Yes
Observations	5.404	5.614	5.576	5.209	5.737
R-squared	0.460	0.466	0.501	0.551	0.623

Notes: Each column corresponds to a different regression with different independent variables. All of these regressions have the same control variables as those in column (3), Table 3. They are race, female dummy variables, age at assessment at fall-kindergarten, age-squared, and background controls, which include family-quality variables, birth weight, and geographic location variables. Please refer to Table 3 for detailed information about these control variables. I also add school fixed effects and teachers' characteristics in the regression. Sample is restricted to those with non-missing observations on racial dummies, gender, weights, specific non-cognitive skill measures, and school ID. Observations are weighted using eighth grade parent panel weights. Robust standard errors reported. ***Significant at 1 percent level, **5 percent level, and *10 percent level.

Table C.10: The Evolution of Racial Gaps in Non-Cognitive Skills (Class of 2010–2011): School Fixed Effects and Teachers' Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Externalizing Behavior			Approaches to Learning			PCA		
	Fall-K	Spring-K	Grade 1	Fall-K	Spring-K	Grade 1	Fall-K	Spring-K	Grade 1
Black	0.092	0.166*	0.245***	-0.128**	-0.131*	-0.244***	-0.315***	-0.374***	-0.501***
	[0.074]	[0.087]	[0.074]	[0.065]	[0.068]	[0.062]	[0.121]	[0.139]	[0.119]
Hispanic	-0.207***	-0.137***	-0.178***	0.068	0.035	0.116**	0.166*	0.127	0.302***
	[0.053]	[0.053]	[0.050]	[0.047]	[0.049]	[0.048]	[0.090]	[0.091]	[0.089]
Asian	-0.170**	-0.221***	-0.289***	0.102	0.190***	0.422***	0.084	0.290**	0.645***
	[0.070]	[0.067]	[0.068]	[0.070]	[0.066]	[0.071]	[0.138]	[0.125]	[0.133]
Geographic Location	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Background Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age and Gender	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Teacher Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5.983	6.126	6.915	6.121	6.140	6.942	5.316	5.996	6.732
R-squared	0.260	0.285	0.334	0.309	0.323	0.361	0.325	0.326	0.371

Notes: Each column corresponds to a different regression with different independent variables. All of these regressions have the same control variables as those in column (3), Table 3. They are race, female dummy variables, age at assessment at fall-kindergarten, age-squared, and background controls, which include family-quality variables, birth weight, and geographic location variables. Please refer to Table 3 for detailed information about these control variables. I also add school fixed effects and teachers' characteristics in the regression. Sample is restricted to those with non-missing observations on racial dummies, gender, weights, specific non-cognitive skill measures, and school ID. Observations are weighted using eighth grade parent panel weights. Robust standard errors reported. ***Significant at 1 percent level, **5 percent level, and *10 percent level.

Appendix D

Additions for Chapter 2

Warmth Index

The Warmth Index is based on the sum of parental responses to the following questions in Spring-K. Each question was recoded such that 0 indicates the most negative response and 3 indicates the warmest response. The scale had a total score of 39. Individuals with missing components were deleted case-wise. Cronbach's alpha: 0.70.

Is the statement (1) completely true, (2) mostly true, (3) somewhat true, (4) not at all true?

a) Child and I often have warm, close times together.
b) Most of the time I feel that child likes me and wants to be near me.
c) I am usually too busy to joke and play around with child.
d) Even when I'm in a bad mood, I show child a lot of love.
e) By the end of a long day, I find it hard to be warm and loving toward child.
f) I express affection by hugging, kissing, and holding child.
g) Being a parent is harder than I thought it would be.
h) Child does things that really bother me.
i) I find myself giving up more of my life to meet child's needs than I ever expected.
j) I feel trapped by my responsibilities as a parent.
k) I often feel angry with child.
l) Child seems harder to care for than most.
m) I find taking care of a young child more work than pleasure.

Appendix E

Supplemental Tables for Chapter 2

Table E.1: The Factors of the Difference between IRT Scores and Teachers' Subjective Assessments: School Fixed Effects and Teachers' Characteristics

	(1)	(2)	(3)	(4)	(5)
			Index of Bias		
	Fall-K	Spring-K	Grade 1	Grade 3	Grade 5
Black	0.176**	0.060	0.004	0.156**	0.013
	[0.068]	[0.059]	[0.062]	[0.070]	[0.080]
Hispanic	0.044	0.033	0.076*	0.043	-0.029
	[0.059]	[0.045]	[0.043]	[0.046]	[0.053]
Asian	-0.136	-0.076	-0.063	0.148*	0.164*
	[0.102]	[0.071]	[0.065]	[0.084]	[0.089]
Other	0.021	-0.112	-0.023	0.143*	-0.062
	[0.080]	[0.073]	[0.053]	[0.074]	[0.072]
Female	0.071**	0.099***	0.079***	0.110***	0.172***
	[0.032]	[0.026]	[0.023]	[0.027]	[0.027]
Observations	4.639	6.266	5.956	5.31	6.203
R-squared	0.492	0.432	0.498	0.598	0.607

Notes: The control variables in all of these regressions are the same as those in Table 6. Please refer to Table 6 for detailed information about these control variables. The sample is restricted to observations with valid information for female, race, weights, and index of bias. Observations are weighted using eighth grade parent panel weights. Robust standard errors reported. ***Significant at 1 percent level, **5 percent level, and *10 percent level.

Table E.2: The Component Loadings of Measures of Non-Cognitive Skills in Grade 5

	(1) Component 1	(2) Component 2	(3) Component 3	(4) Component 4	(5) Component 5
Externalizing Behavior	0.4502	-0.2178	0.7983	0.2278	0.2463
Approaches to Learning	0.4721	-0.0199	-0.4571	0.7481	-0.0905
Self-control	0.4958	-0.2274	0.0006	-0.4072	-0.7326
Interpersonal Skills	0.4903	-0.1215	-0.3707	-0.4633	0.6267
Internalizing Problems	0.2973	0.9411	0.1274	-0.0897	-0.0411

Notes: The sample is restricted to all of the observations with valid information for all five measures of non-cognitive skills in grade 5.

Appendix F

Supplemental Figures for Chapter 2

Figure F.1: Distribution of Externalizing Behavior in Grade 5 (Before Standardized)
Notes: The sample is restricted to all of the observations with valid information for externalizing behaviors in grade 5.

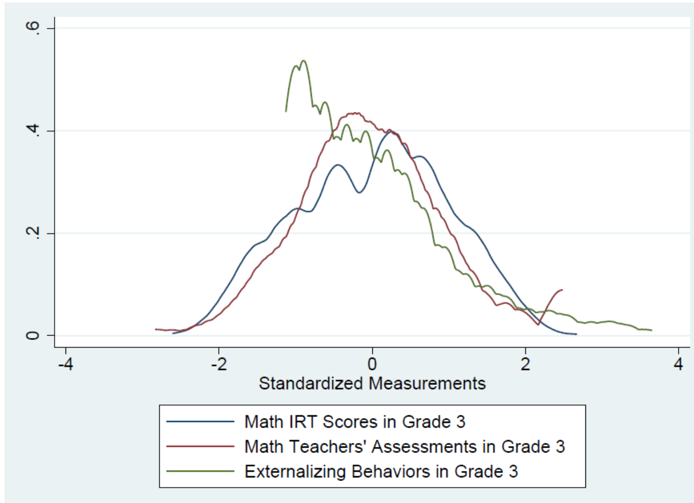


Figure F.2: Distribution of IRT Scores, Teachers' Assessments of Cognitive Skills and Externalizing Behavior in Grade 3 (After Standardization).
Notes: The sample is restricted to all of the observations with valid information for math IRT scores, teachers' assessments, and externalizing behaviors in grade 3. The IRT scores, teachers' subjective assessments, and measurements of externalizing behaviors are all standardized.

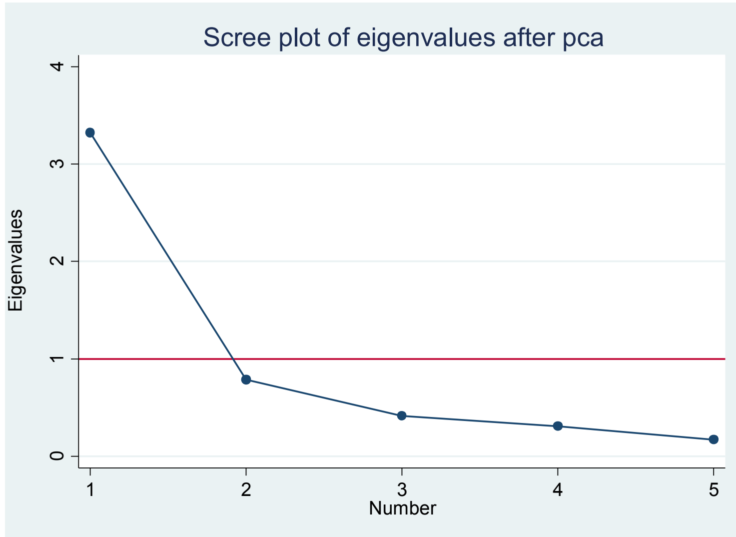


Figure F.3: The Eigenvalues of First Five Components of Non-Cognitive Skills in Grade 5.
Notes: The sample is restricted to all of the observations with valid information for all five measures of non-cognitive skills in grade 5.

Appendix G

Tables for Chapter 3

Table G.1: Sample Selection

Sample	Observations	
Panel A: CRSP-Cumpustat Merged unique firm quarters 2002-2014		336.801
Keep firms that have IBES ticker	(20.426)	
Delete firms with missing earnings announcement date	(33.359)	
Delete firms with missing quarter information (>120 days between the two adjacent fiscal quarter end dates)	(9.252)	
After deleting firms with the earnings announcement date > 120 days away from the fiscal quarter end date	(20.398)	
After deleting firms with more than 120 days or smaller than 28 days between adjacent earnings announcement dates	(8.948)	
CRSP-Cumpustat Merged final sample		244.418
Panel B:		
IBES Guidance 2002-2014; US firms with valid PERMNO		322.791
Merge with CRSP-Cumpustat Merged from Panel A		425.728=
From IBES Guidance 2002-2014		271.364+
From CRSP-Cumpustat Merged from Panel A alone		154.364
Keep observations after 2002	(32.250)	
Delete firms that never issue a forecast between 2003 and 2014	(59.494)	
Keep earnings related disclosures	(145.180)	
Drop pre-announcements	(7.299)	
Keep the earliest forecast of each quarter	(59.427)	
Drop forecasts issued within 5 days of the prior quarter earnings announcement	(54.357)	
Keep non-missing stock return and control variables	(9.278)	
Final sample		

Table G.2: Summary Statistics

Variable	N	Mean	SD	Percentile 5th	25th	50th	75th	95th
$CRr_{i,q,30}$	58,443	0.015	0.138	-0.227	-0.056	0.018	0.093	0.243
$CRr_{i,q,-5}$	58,443	0.013	0.195	-0.338	-0.085	0.024	0.127	0.318
$NG_{i,q}$	58,443	0.970	0.171	1.000	1.000	1.000	1.000	1.000
$NG^{b}_{i,q}$	122,078	0.538	0.499	0.000	0.000	1.000	1.000	1.000
Earnings $surprise_{i,q}$	58,443	0.001	0.083	-0.080	-0.007	0.002	0.010	0.080
$Log(MVE)_{i,q-1}$	58,443	13.087	1.903	10.055	11.757	12.979	14.331	16.541
EP $ratio_{i,q-1}$	58,443	-0.005	0.078	-0.108	-0.006	0.011	0.020	0.054
Return $volatility_{i,q-1}$	58,443	0.029	0.017	0.010	0.017	0.024	0.036	0.064
Number of $analysts_{i,q-1}$	58,443	6.592	6.953	0.000	1.000	4.000	9.000	22.000
$MTB_{i,q-1}$	58,443	2.708	3.731	0.460	1.165	1.841	3.120	8.583

Notes: This table describes the distribution of the variables used in the empirical analysis. $CRr_{i,q,t}$ is log cumulative stock return from 6 days subsequent to the earnings announcement for quarter q-1 to *t* days *prior* to the earnings announcement quarter q (negative t means after the earnings announcement). $NG_{i,q}$ is a dummy that is equal to one if management of firm i does not issue a forecast in quarter *q*. $NG^{b}_{i,q}$ is defined the same way as $NG_{i,q}$ but the distribution is computed with the sample including bundled forecasts. Earnings $surprise_{i,q}$ is the difference between current quarter earnings and the same quarter earnings of last year, scaled by the closing stock price five days subsequent to the previous quarter earnings announcement. $Log(MVE)_{i,q}$ is the average natural logarithm of firm *i*'s market value of equity for quarter q. EP $ratio_{i,q}$ is primary earnings per share excluding extraordinary items of firm in quarter *q*, scaled by stock price five days subsequent to the earnings announcement of firm *i* for quarter *q-1*. Return $volatility_{i,q}$ is the standard deviation of daily stock return of firm *i* in quarter *q*. Number of $analysts_{i,q}$ is the number of analysts issuing a forecast for firm *i* in quarter *q*. $MTB_{i,q}$ is the market value of equity divided by book value of equity of firm *i* in quarter *q*. All variables are winsorized at 1% and 99%.

Table G.3: The Differences in the Cumulative Return between Firms With and Without Guidance

VARIABLES	(1) Long Window	(2)	(3) Short Window	(4)
$ND_{i,q}$	0.020***	0.019***	0.023***	0.023***
	[0.003]	[0.003]	[0.004]	[0.004]
$After_{i,q}$	0.004**	0.002	0.004**	0.003*
	[0.002]	[0.002]	[0.001]	[0.002]
$ND_{i,q}$* $After_{i,q}$	-0.005**	-0.004*	-0.007***	-0.007***
	[0.002]	[0.002]	[0.002]	[0.002]
Earnings $surprise_{i,q}$	0.085***	0.084***	0.111***	0.112***
	[0.009]	[0.010]	[0.011]	[0.012]
Earnings $surprise_{i,q}$* $After_{i,q}$	0.135***	0.138***	0.105***	0.108***
	[0.008]	[0.008]	[0.006]	[0.006]
$Log(MVE)_{i,q-1}$		0.001		0.001*
		[0.001]		[0.001]
EP $ration_{i,q-1}$		0.094***		0.110***
		[0.012]		[0.014]
Return $volatility_{i,q-1}$		-0.531***		-0.704***
		[0.059]		[0.068]
Number of $analysts_{i,q-1}$		-0.000		-0.000
		[0.000]		[0.000]
No analyst $following_{i,q-1}$		-0.006***		-0.003
		[0.002]		[0.002]
Earnings $surprise_{i,q-1}$		-0.011		-0.019
		[0.011]		[0.012]
$MTB_{i,q-1}$		-0.001***		-0.001***
		[0.000]		[0.000]
$CR_{i,q-1,-5}$		-0.005		-0.001
		[0.004]		[0.004]
Year * Quarter Fixed Effects	Yes	Yes	Yes	Yes
Observations	2.247.696	2.103.948	686.796	642.873
R^2	0.139	0.150	0.159	0.172

Notes: This table presents regression outputs using our main specification. $CR_{iqt} = \beta_0 + \beta_1 ND_{iq} + \beta_2 After_{iq}*ND_{iq} + \beta_3 EarnSurp_{iq} + \beta_4 After_{iq}*EarnSurp_{iq} + \beta_5 After_{iq} + \varepsilon_{iqt}$, where CR_{iqt} is firm *i*'s raw cumulative stock returns from 6 days subsequent to the previous quarter earnings announcement date to t days prior to the current quarter earnings announcement; in Column 1 and 2, *t* is from -5 to 30, which means that from 30 days prior to the current quarter earnings announcement to 5 days subsequent to the current quarter earnings announcement; in Column 3 and 4, *t* is from -5 to 5; $ND_{i,q}$ is a dummy variable that equals to one when management of firm *i* does not provide a forecast in quarter *q*; $EarnSurp_{i,q}$ is the earnings news of quarter *q* for firm *i*,

Table G.3 (cond't)

measured as the difference between current quarter earnings and the same quarter earnings of last year, scaled by the closing stock price five days subsequent to the previous quarter earnings announcement; $After_{i,q}$ is a dummy variable that equals to one if the cumulative return ending window is on or after the current quarter earnings announcement date, and zero otherwise; ε_{iqt} is the error term. X_{iq} are defined in Table 2. We also control for No Analyst $Following_{i,q-1}$, which is a dummy variable that equals to one if firm i in quarter $q-1$ is not followed by any analyst, and year*quarter fixed effects. Standard errors are clustered at the firm level.

Table G.4: The Differences in the Cumulative Return between Firms With and Without Guidance: By Earnings Surprise

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			Earnings Surprise ≤			Earnings Surprise > 0		
VARIABLES	Long Window			Short Window	Long Window			Short Window
$ND_{i,q}$	0.034***	0.038***	0.035***	0.041***	0.008*	0.005	0.010**	0.007
	[0.005]	[0.006]	[0.005]	[0.006]	[0.004]	[0.004]	[0.004]	[0.005]
$After_{i,q}$	-0.004	-0.001	-0.007	-0.002	0.010***	0.008***	0.007***	0.006***
	[0.003]	[0.002]	[0.003]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]
$ND_{i,q}$* $After_{i,q}$	-0.015***	-0.017***	-0.012***	-0.016***	0.002	0.002	0.000	0.000
	[0.003]	[0.002]	[0.003]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]
Earnings $surprise_{i,q}$	0.118***	0.141***	0.070***	0.084***	-0.025*	0.035**	-0.021	0.044***
	[0.014]	[0.016]	[0.016]	[0.018]	[0.013]	[0.015]	[0.015]	[0.017]
Earnings $surprise_{i,q}$* $After_{i,q}$	0.067***	0.050***	0.075***	0.054***	0.055***	0.053***	0.039***	0.038***
	[0.012]	[0.009]	[0.012]	[0.009]	[0.010]	[0.011]	[0.008]	[0.009]
$Log(MVE)_{i,q-1}$			0.002***	0.004***		-0.001		-0.002**
			[0.001]	[0.001]		[0.001]		[0.001]
EP $ration_{i,q-1}$			1.878***	1.677***		1.966***		2.358***
			[0.322]	[0.375]		[0.273]		[0.326]
Return $volatility_{i,q-1}$			-0.432***	-0.619***		-0.639***		-0.809***
			[0.086]	[0.099]		[0.080]		[0.092]
Number of $analysts_{i,q-1}$			0.000	0.000		-0.000		0.000
			[0.000]	[0.000]		[0.000]		[0.000]
No analyst $following_{i,q-1}$			-0.006**	-0.005		-0.003		0.002
			[0.003]	[0.003]		[0.003]		[0.003]
Earnings $surprise_{i,q-1}$			-0.029*	-0.036**		0.017		0.013
			[0.015]			[0.013]		[0.015]
$MTB_{i,q-1}$			-0.001**	-0.001***		-0.001**		-0.001**
			[0.000]			[0.000]		[0.000]
$CR_{i,q-1,-5}$			-0.019***	-0.018***		-0.004		-0.003
			[0.005]	[0.006]		[0.005]		[0.006]
Year * Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	944,316	888,084	288,541	271,359	1,302,380	1,215,864	398,255	371,514
R^2	0.144	0.157	0.171	0.186	0.136	0.146	0.148	0.158

Table G.4 (cond't)

Notes: This table presents regression outputs using our main specification. $CR_{iqt} = \beta_0 + \beta_1 ND_{iq} + \beta_2 After_{iq}*ND_{iq} + \beta_3 EarnSurp_{iq} + \beta_4 After_{iq}*EarnSurp_{iq} + \beta_5 After_{iq} + X_{iq} + \varepsilon_{iqt}$, where CR_{iqt} is firm *i*'s raw cumulative stock returns from 6 days subsequent to the previous quarter earnings announcement date to t days *prior* to the current quarter earnings announcement; in Column 1 and 2, t is from -5 to 30, which means that from 30 days prior to the current quarter earnings announcement to 5 days *subsequent* to the current quarter earnings announcement; in Column 3 and 4, t is from -5 to 5; ND_{iq} is a dummy variable that equals to one when management of firm *i* does not provide a forecast in quarter *q*; $EarnSurp_{iq}$ is the earnings news of quarter *q* for firm *i*, measured as the difference between current quarter earnings and the same quarter earnings of last year, scaled by the closing stock price five days subsequent to the previous quarter earnings announcement; $After_{iq}$ is a dummy variable that equals to one if the cumulative return ending window is on or after the current quarter earnings announcement date, and zero otherwise; ε_{iqt} is the error term. X_{iq} are control variables defined in Table 3. From Column (1) to Column (4) we use firm quarters with non-positive earnings surprise. From Column (5) to Column (8) we use firm quarters with positive earnings surprise.

Table G.5: The Differences in the Cumulative Return between Firms With and Without Guidance: By Number of Analyst Following

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Number of Analysts Following > 50 Percentile				Number of Analysts Following ≤ 50 Percentile			
VARIABLES	Long Window		Short Window		Long Window		Short Window	
$ND_{i,q}$	0.027***	0.031***	0.023***	0.027***	0.003	0.007	0.007	0.011
	[0.004]	[0.004]	[0.004]	[0.004]	[0.008]	[0.009]	[0.008]	[0.010]
$After_{i,q}$	0.006***	0.006***	0.003	0.004***	-0.005	-0.003	-0.005	-0.002
	[0.002]	[0.002]	[0.002]	[0.002]	[0.005]	[0.004]	[0.005]	[0.004]
$ND_{i,q}$* $After_{i,q}$	-0.004*	-0.007***	-0.003	-0.006***	0.001	-0.003	0.001	-0.004
	[0.002]	[0.002]	[0.002]	[0.002]	[0.005]	[0.004]	[0.005]	[0.004]
Earnings $surprise_{i,q}$	0.110***	0.139***	0.109***	0.139***	0.070***	0.095***	0.073***	0.099***
	[0.015]	[0.017]	[0.016]	[0.019]	[0.012]	[0.014]	[0.013]	[0.015]
Earnings $surprise_{i,q}$* $After_{i,q}$	0.130***	0.094***	0.127***	0.091***	0.142***	0.117***	0.148***	0.121***
	[0.012]	[0.010]	[0.013]	[0.010]	[0.010]	[0.008]	[0.011]	[0.008]
$Log(MVE)_{i,q-1}$			0.001	0.002**			0.002**	0.002
			[0.001]	[0.001]			[0.001]	[0.001]
EP $ration_{i,q-1}$			0.115***	0.124***			0.084***	0.106***
			[0.023]	[0.025]			[0.015]	[0.017]
Return $volatility_{i,q-1}$			-0.067	-0.189			-0.741***	-0.941***
			[0.102]	[0.116]			[0.074]	[0.085]
Number of $analysts_{i,q-1}$			0.000	-0.000			-0.000	-0.001
			[0.000]	[0.000]			[0.001]	[0.002]
No analyst $following_{i,q-1}$							-0.003	-0.004
							[0.003]	[0.004]
Earnings $surprise_{i,q-1}$			-0.043**	-0.048**			0.003	-0.007
			[0.018]	[0.021]			[0.013]	[0.015]
$MTB_{i,q-1}$			-0.000	-0.000			-0.001***	-0.001***
			[0.000]	[0.000]			[0.000]	[0.000]
$CR_{i,q-1,-5}$			-0.008	-0.005			-0.002	0.002
			[0.006]	[0.007]			[0.005]	[0.006]
Year * Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	929.340	908.353	283.965	277.552	1.162.296	1.041.912	355.146	318.362
R^2	0.150	0.170	0.163	0.185	0.133	0.153	0.144	0.166

Table G.5 (cond't)

Notes: This table presents regression outputs using our main specification. $CR_{iqt} = \beta_0 + \beta_1 ND_{iq} + \beta_2 After_{iq}*ND_{iq} + \beta_3 EarnSurp_{iq} + \beta_4 After_{iq}*EarnSurp_{iq} + \beta_5 After_{iq} + X_{iq} + \varepsilon_{iqt}$, where CR_{iqt} where CR_{iqt} is firm *i*'s raw cumulative stock returns from 6 days subsequent to the previous quarter earnings announcement date to *t* days prior to the current quarter earnings announcement; in Column 1 and 2, t is from -5 to 30, which means that from 30 days prior to the current quarter earnings announcement to 5 days subsequent to the current quarter earnings announcement; in Column 3 and 4, *t* is from -5 to 5; ND_{iq} is a dummy variable that equals to one when management of firm *i* does not provide a forecast in quarter *q*; $EarnSurp_{iq}$ is the earnings news of quarter *q* for firm *i*, measured as the difference between current quarter earnings and the same quarter earnings of last year, scaled by the closing stock price five days subsequent to the previous quarter earnings announcement; $After_{iq}$ is a dummy variable that equals to one if the cumulative return ending window is on or after the current quarter earnings announcement date, and zero otherwise; ε_{iqt} is the error term. X_{iq} are control variables defined in Table 3. From Column (1) to Column (4) we use firm quarters with number of analysts following larger than 50 percentile. From Column (5) to Column (8) we use firm quarters with number of analysts following smaller than or equal to 50 percentile.

Table G.6: The Differences in the Cumulative Return between Firms With and Without Guidance: By Return Volatility

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Return Volatility ≤ 50 Percentile				Return Volatility > 50 Percentile			
VARIABLES	Long Window		Short Window		Long Window		Short Window	
$ND_{i,q}$	0.006	0.007	0.013***	0.016***	0.022***	0.025***	0.022***	0.026
	[0.007]	[0.007]	[0.007]	[0.007]	[0.004]	[0.004]	[0.004]	[0.004]
$After_{i,q}$	-0.003	-0.001	-0.002	-0.001	0.004**	0.005***	0.004*	0.005***
	[0.004]	[0.003]	[0.004]	[0.003]	[0.002]	[0.001]	[0.002]	[0.001]
$ND_{i,q}$* $After_{i,q}$	-0.006	-0.007***	-0.006	-0.008***	-0.001	-0.004**	-0.001	-0.004
	[0.004]	[0.003]	[0.004]	[0.003]	[0.002]	[0.002]	[0.002]	[0.002]
Earnings $surprise_{i,q}$	0.074	0.099***	0.074***	0.100***	0.118***	0.153***	0.121***	0.155***
	[0.011]	[0.012]	[0.011]	[0.013]	[0.018]	[0.020]	[0.019]	[0.022]
Earnings $surprise_{i,q}$* $After_{i,q}$	0.132***	0.106***	0.137***	0.109***	0.139***	0.099***	0.142***	0.102***
	[0.009]	[0.007]	[0.009]	[0.007]	[0.015]	[0.012]	[0.016]	[0.012]
$Log(MVE)_{i,q-1}$			0.004***	0.004**			0.000	0.000
			[0.001]	[0.001]			[0.001]	[0.001]
EP $ration_{i,q-1}$			0.079***	0.097***			0.106***	0.121***
			[0.014]	[0.015]			[0.023]	[0.026]
Return $volatility_{i,q-1}$			-0.881***	-1.118			0.648***	0.611***
			[0.88]	[0.101]			[0.147]	[0.169]
Number of $analysts_{i,q-1}$			-0.000	-0.000			-0.000	-0.000
			[0.000]	[0.000]			[0.000]	[0.000]
No analyst $following_{i,q-1}$			-0.003	0.002			-0.001	0.000
			[0.003]	[0.003]			[0.003]	[0.003]
Earnings $surprise_{i,q-1}$			-0.009	-0.015			0.008	-0.002
			[0.013]	[0.014]			[0.017]	[0.019]
$MTB_{i,q-1}$			-0.001**	-0.001***			-0.000	-0.000
			[0.000]	[0.000]			[0.000]	[0.000]
$CR_{i,q-1,-5}$			-0.002	0.002			-0.0018***	-0.012
			[0.004]	[0.005]			[0.006]	[0.007]
Year * Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1.073.160	327.910	1.057.932	323.257	1.070.604	1.046.016	327.129	319.616
R^2	0.132	0.154	0.141	0.165	0.198	0.217	0.200	0.218

Table G.6 (cond't)

Notes: This table presents regression outputs using our main specification. $CR_{iqt} = \beta_0 + \beta_1 ND_{iq} + \beta_2 After_{iq} * ND_{iq} + \beta_3 EarnSurp_{iq} + \beta_4 After_{iq} * EarnSurp_{iq} + \beta_5 After_{iq} + X_{iq} + \varepsilon_{iqt}$, where CR_{iqt} is firm i's raw cumulative stock returns from 6 days subsequent to the previous quarter earnings announcement date to *t* days prior to the current quarter earnings announcement; in Column 1 and 2, *t* is from -5 to 30, which means that from 30 days prior to the current quarter earnings announcement to 5 days subsequent to the current quarter earnings announcement; in Column 3 and 4, *t* is from -5 to 5; ND_{iq} is a dummy variable that equals to one when management of firm *i* does not provide a forecast in quarter *q*; $EarnSurp_{iq}$ is the earnings news of quarter q for firm i, measured as the difference between current quarter earnings and the same quarter earnings of last year, scaled by the closing stock price five days subsequent to the previous quarter earnings announcement; $After_{iq}$ is a dummy variable that equals to one if the cumulative return ending window is on or after the current quarter earnings announcement date, and zero otherwise; ε_{iqt} is the error term. X_{iq} are control variables defined in Table 3. From Column (1) to Column (4) we use firm quarters with return volatility smaller or equal to 50 percentile. From Column (5) to Column (8) we use firm quarters with return volatility larger than 50 percentile.

Table G.7: The Differences in the Size Adjusted Cumulative Return between Firms With and Without Guidance

	(1)	(2)	(3)	(4)
VARIABLES	Long Window		Short Window	
$ND_{i,q}$	0.012***	0.018***	0.014***	0.018***
	[0.003]	[0.003]	[0.003]	[0.003]
$After_{i,q}$	-0.001	-0.001	0.001	-0.001
	[0.002]	[0.002]	[0.001]	[0.002]
$ND_{i,q}$* $After_{i,q}$	-0.003	-0.003*	-0.005***	-0.003*
	[0.002]	[0.002]	[0.001]	[0.002]
Earnings $surprise_{i,q}$	0.060***	0.062***	0.084***	0.062***
	[0.008]	[0.009]	[0.009]	[0.009]
Earnings $surprise_{i,q}$* $After_{i,q}$	0.117***	0.127***	0.091***	0.127***
	[0.007]	[0.007]	[0.005]	[0.007]
$Log(MVE)_{i,q-1}$		0.002***		0.002**
		[0.001]		[0.001]
EP $ration_{i,q-1}$		0.058***		0.058***
		[0.001]		[0.010]
Return $volatility_{i,q-1}$		-0.463***		-0.463***
		[0.051]		[0.051]
Number of $analysts_{i,q-1}$		-0.000		-0.000
		[0.000]		[0.000]
No analyst $following_{i,q-1}$		-0.002		-0.002
		[0.002]		[0.002]
Earnings $surprise_{i,q-1}$		-0.005		-0.005
		[0.009]		[0.019]
$MTB_{i,q-1}$		-0.000***		-0.000***
		[0.000]		[0.000]
$CR_{i,q-1,-5}$		0.007**		0.007**
		[0.004]		[0.004]
Year * Quarter Fixed Effects	Yes	Yes	Yes	Yes
Observations	2.246.544	2.095.704	686.444	640.354
R^2	0.009	0.017	0.014	0.022

Notes: This table presents regression outputs using our main specification. $CR_{iqt} = \beta_0 + \beta_1 ND_{iq} + \beta_2 After_{iq} * ND_{iq} + \beta_3 EarnSurp_{iq} + \beta_4 After_{iq} * EarnSurp_{iq} + \beta_5 After_{iq} + X_{iq} + \varepsilon_{iqt}$, where $CRS_{iq}t$ is firm *i*'s size adjusted cumulative stock returns from 6 days subsequent to the previous quarter earnings announcement date to *t* days prior to the current quarter earnings an- nouncement; in Column 1 and 2, *t* is from -5 to 30, which means that from 30 days prior to the current quarter earnings announcement to 5 days subsequent to the current quarter earnings an- nouncement; in Column 3 and 4, t is from -5 to 5; ND_{iq} is a dummy variable that equals to one when management of firm i does not provide a forecast in quarter *q*; $EarnSurp_{iq}$ is the earnings news of quarter *q* for firm *i*, measured as the difference between current quarter earnings and the same quarter earnings of last year, scaled by the closing stock price five days subsequent to the previous quarter earnings announcement; $After_{iq}$ is a dummy variable that equals to one if the cumulative return ending window is on or after the current quarter earnings announcement date, and zero otherwise; ε_{iqt} is the error term. X_{iq} are control variables defined in Table 3 except that all stock returns are size adjusted.

Table G.8: The Differences in the Cumulative Return between Firms With and Without Guidance: Bundled Guidance

VARIABLES	(1) Long Window	(2) Long Window	(3) Short Window	(4) Short Window
$ND_{i,q}$	0.002***	-0.002	-0.000	0.031***
	[0.001]	[0.002]	[0.002]	[0.006]
Earnings $surprise_{i,q}$	0.026***	0.121***	0.0013	0.097***
	[0.007]	[0.017]	[0.009]	[0.021]
$Log(MVE)_{i,q-1}$	0.002***	0.004***	0.002*	0.003**
	[0.000]	[0.001]	[0.000]	[0.001]
EP $ration_{i,q-1}$	0.047***	-0.012	0.050***	0.024
	[0.009]	[0.020]	[0.010]	[0.024]
Return $volatility_{i,q-1}$	-0.290***	-0.428***	-0.331***	-0.443
	[0.040]	[0.090]	[0.047]	[0.110]
Number of $analysts_{i,q-1}$	-0.000	-0.000	0.000	0.000
	[0.000]	[0.000]	[0.000]	[0.000]
No analyst $following_{i,q-1}$	-0.003***	0.002	-0.002	0.002
	[0.002]	[0.003]	[0.002]	[0.004]
Earnings $surprise_{i,q-1}$	0.034***	0.011	0.033**	0.013
	[0.008]	[0.018]	[0.009]	[0.021]
$MTB_{i,q-1}$	-0.000***	-0.000	-0.001***	-0.000
	[0.000]	[0.000]	[0.000]	[0.000]
$CR_{i,q-1}^{[0,5]}$	-0.020***	-0.010	-0.017**	-0.008
	[0.006]	[0.011]	[0.007]	[0.015]
	-0.011***	-0.014**	-0.013	-0.012
$CR_{i,q-1}^{[6,T-1]}$	[0.003]	[0.006]	[0.004]	[0.008]
Year * Quarter Fixed Effects	Yes	Yes	Yes	Yes
Observations	39.985	39.985	22.191	22.191
R^2	0.021	0.017	0.028	0.022

Notes: This table presents regression outputs using our main specification. $CRS_{iq}^{[a,T-b]} = \beta_0 + \beta_1 ND_{iq} + \beta_2 After_{iq} * ND_{iq} + \beta_3 EarnSurp_{iq} + \beta_4 After_{iq} * EarnSurp_{iq} + \beta_5 After_{iq} + X_{iq} + \varepsilon_{iqt}$, where $CRS_{iq}^{[a,T-b]}$ is firm *i*'s size adjusted cumulative stock returns from a days subsequent to quarter *q-1* earnings announcement date to *b* days prior to quarter *q* earnings announcement date; ND_{iq} is a dummy variable that equals to one when management of firm *i* does not provide a forecast in quarter *q*; $EarnSurp_{iq}$ is the earnings news of quarter *q* for firm *i*, measured as the difference between current quarter earnings and the same quarter earnings of last year, scaled by the closing stock price five days subsequent to the previous quarter earnings announcement; $After_{iq}$ is a dummy variable that equals to one if the cumulative return ending window is on or after the current quarter earnings announcement date, and zero otherwise; ε_{iqt} is the error term. X_{iq} are control variables defined in Table 3 except that all stock returns are size adjusted.

Table G.9: Portfolio returns

Days after earnings announcement	N	Mean	SD	5th	25th	50th	75th	95th
0	2723	0.004***	0.049	-0.061	-0.016	0.003	0/024	0.071
1	2811	0.006***	0.075	-0.103	-0.026	0.006	0.041	0.118
2	2830	0.009***	0.076	-0.109	-0.027	0.007	0.046	0.126
3	2841	0.011***	0.085	-0.114	-0.028	0.010	0.049	0.133
4	2841	0.012***	0.089	-0.116	-0.029	0.010	0.051	0.139
5	2841	0.012***	0.090	-0.119	-0.030	0.010	0.053	0.141

Notes: This table constructs calendar-quarter portfolios on earnings announcement dates based on the existence of management forecasts. In the first step, for each firm that announced earnings in a given calendar quarter, we sort firms into three buckets based on the size of its earnings surprise. We choose 30% and 70% as the cut-off points, which are computed based on the distribution of earnings surprise in the previous calendar quarter. This step aims to hedge out post earnings announcement drifts. In the second step, conditional on each earnings surprise bucket, we short firms that did not issue management forecasts during the quarter and long firms that issued earnings forecasts, and hold the portfolio for five days. In cases where we cannot find any match for a nondisclosure firm, we short the nondisclosure firm for five days and long the next available disclosure firm for five days when it becomes available. We exclude firms that pre-announce earnings and firms that bundle forecasts with earnings announcement from the portfolio construction. The table reports the distribution of portfolio stock returns over the next five days after the portfolio construction. N is the number of portfolio.

Appendix H

Figures for Chapter 3

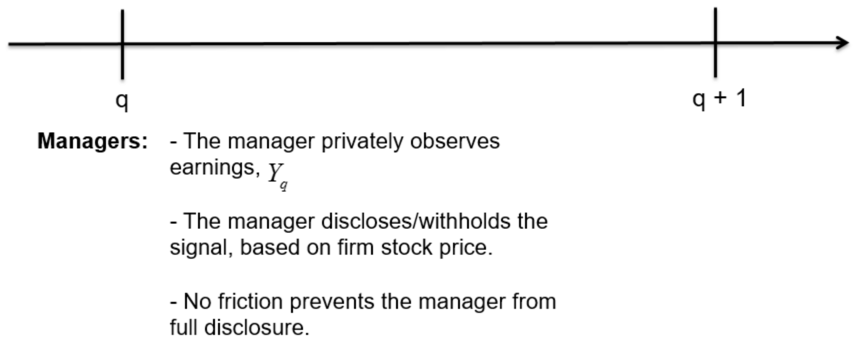
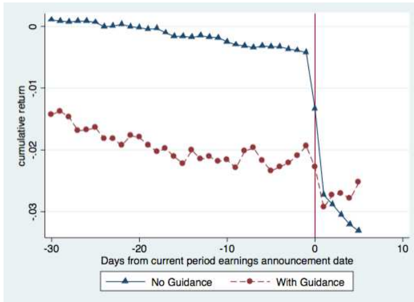
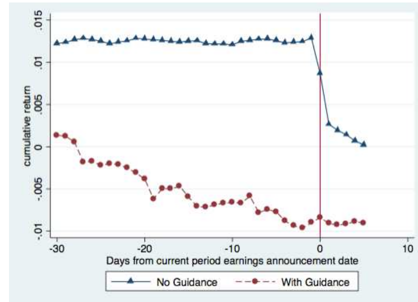


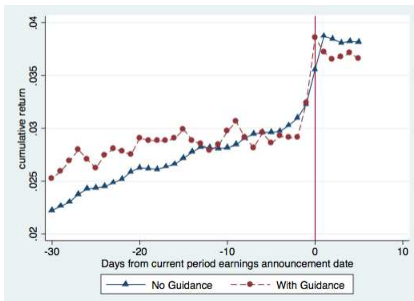
Figure H.1: Time Line.
Notes: This figure shows the time line of the voluntary disclosure.



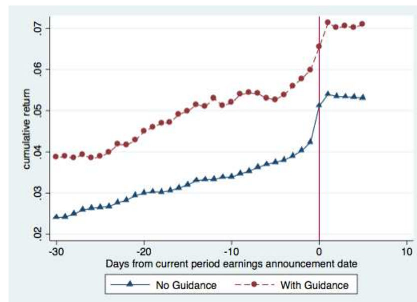
(a) The first quartile of earnings surprise



(b) The second quartile of earnings surprise



(c) The third quartile of earnings surprise



(d) The fourth quartile of earnings surprise

Figure H.2: Cumulative Return of Firm Quarters With and Without Guidance: by Earnings Surprise.

Notes: This figure shows cumulative stock returns from six days after the previous quarter earnings announcement date until five days after the current quarter earnings announcement date separately for firm quarters with and without guidance. The x-axis is days relative to the current quarter earnings announcement. For expositional purpose, the figure only shows 30 days prior to the cur- rent quarter earnings announcement until 5 days subsequent to it. The y-axis is cumulative stock returns. The sample is divided into four parts based on the magnitude of earnings surprise. Fig- ure (a) plots firm quarters with the lowest quartile of the earnings surprise. Figure (b) plots firm quarters with earnings surprise between 25% and 50% of the whole earnings surprise distribution. Figure (c) plots firm quarters with earnings surprise between 50% and 75% of the whole earnings surprise distribution. Lastly, Figure (d) plots firm quarters with the highest quartile of the earnings surprise.

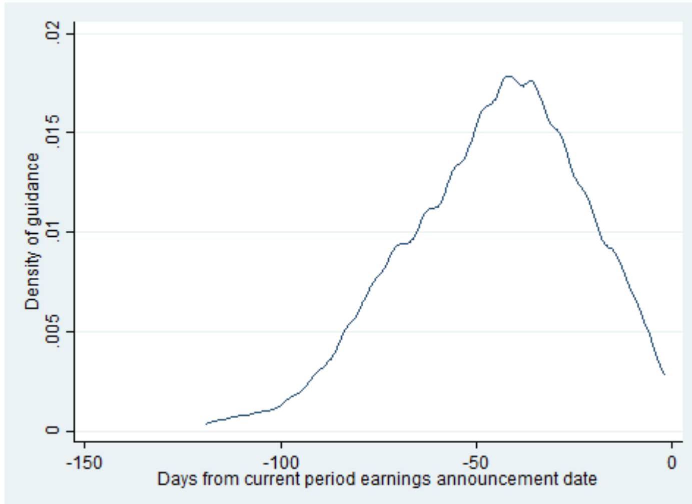


Figure H.3: Distribution of disclosure date
Notes: This figure shows the probability density function of voluntary disclosure dates relative to the forthcoming earnings announcement date. We only show the density function within 120 days of the current quarter earnings announcement date.

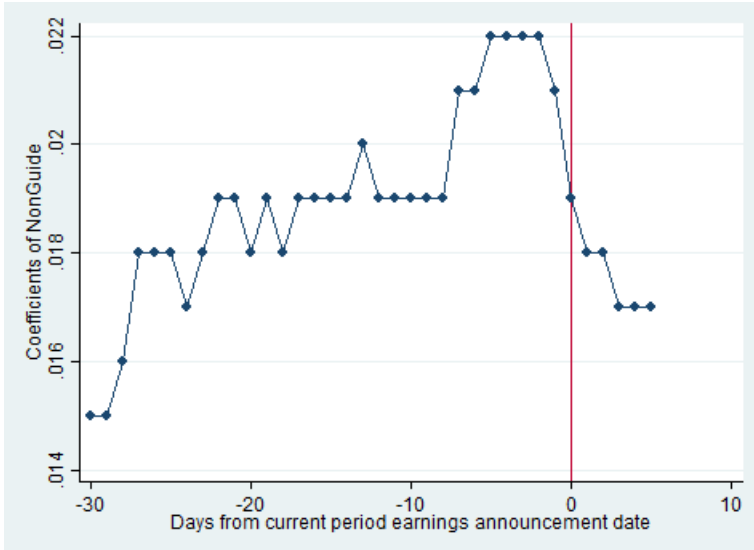


Figure H.4: Coefficients of Non-guidance.
Notes: This figure shows the coefficients of non-guidance for the following regression: $CR_{iqt} = \beta_{0t} + \beta_{1t}ND_{iq} + \beta_{2t}EarnSurp_{iq} + X_{iq} + \varepsilon_{iqt}$, where CR_{iqt} is firm *i*'s raw cumulative stock returns from 6 days subsequent to the previous quarter earnings announcement date to *t* days prior to the current quarter earnings announcement; ND_{iq} is a dummy variable that equals to one when management of firm *i* does not provide a forecast in quarter q; $EarnSurp_{iq}$ is the earnings news of quarter *q* for firm *i*, measured as the difference between current quarter earnings and the same quarter earnings of last year, scaled by the closing stock price five days subsequent to the previous quarter earnings announcement; X_{iq} are the same with those defined in Table G.3. ε_{iqt} is the error term.

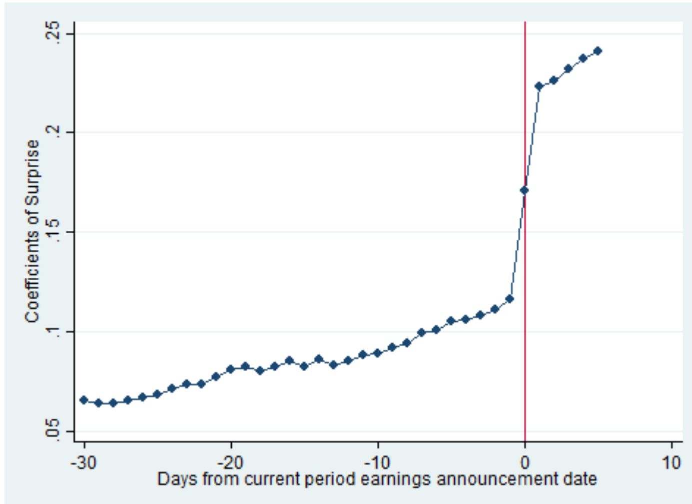


Figure H.5: Coefficients of Earnings Surprise.
Notes: This figure shows the coefficients for earnings surprise for the following regression: $CR_{iqt} = \beta_{0t} + \beta_{1t}ND_{iq} + \beta_{2t}EarnSurp_{iq} + X_{iq} + \varepsilon_{iqt}$, where CR_{iqt} is firm *i*'s raw cumulative stock returns from 6 days subsequent to the previous quarter earnings announcement date to t days prior to the current quarter earnings announcement; ND_{iq} is a dummy variable that equals to one when management of firm *i* does not provide a forecast in quarter *q*; $EarnSurp_{iq}$ is the earnings news of quarter q for firm *i*, measured as the difference between current quarter earnings and the same quarter earnings of last year, scaled by the closing stock price five days subsequent to the previous quarter earnings announcement; X_{iq} are the same with those defined in Table G.3. ε_{iqt} is the error term.

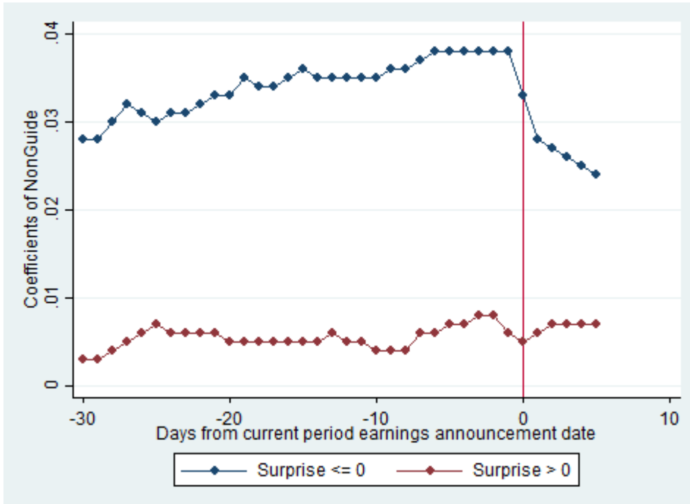
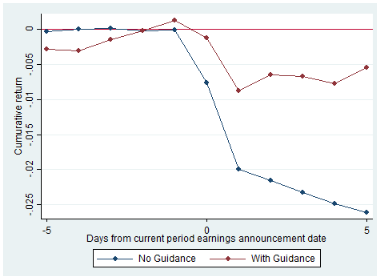
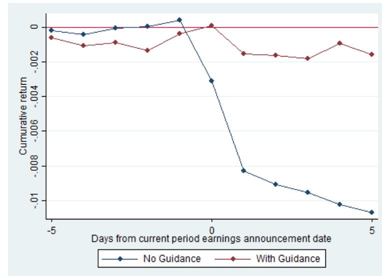


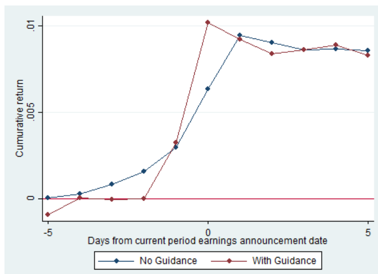
Figure H.6: Coefficients of Non-guidance: by Earnings Surprise.
Notes: This figure shows the coefficients of non-guidance for the following regression: $CR_{iqt} = \beta_{0t} + \beta_{1t}ND_{iq} + \beta_{2t}EarnSurp_{iq} + X_{iq} + \varepsilon_{iqt}$, from sub-sample results (by earnings surprise). The blue dots (the higher line) are for negative earnings surprise sample. The red dots (the lower line) are for positive earnings surprise sample. CR_{iqt} is firm *i*'s raw cumulative stock returns from 6 days subsequent to the previous quarter earnings announcement date to *t* days prior to the current quarter earnings announcement; ND_{iq} is a dummy variable that equals to one when management of firm *i* does not provide a forecast in quarter *q*; $EarnSurp_{iq}$ is the earnings news of quarter *q* for firm *i*, measured as the difference between current quarter earnings and the same quarter earnings of last year, scaled by the closing stock price five days subsequent to the previous quarter earnings announcement; X_{iq} are the same with those defined in Table G.3. ε_{iqt} is the error term.



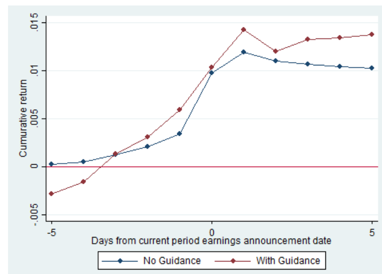
(a) The first quartile of earnings surprise



(b) The second quartile of earnings surprise



(c) The third quartile of earnings surprise



(d) The fourth quartile of earnings surprise

Figure H.7: Raw Return of Firm Quarters With and Without Guidance: by Earnings Surprise.
Notes: This figure shows the raw cumulative stock return 5 days around the earnings announcement date for firm quarters with and without guidance. The sample is divided into four parts based on the earnings surprise. Figure (a) plots the raw stock return for firm quarters with the lowest quartile of the earnings surprise. Figure (b) plots the raw stock return for firm quarters with earnings surprise between 25% and 50% of the whole earnings surprise distribution. Figure (c) plots the raw stock return for firm quarters with earnings surprise between 50% and 75% of the whole earnings surprise distribution. Lastly, Figure (d) plots the raw stock return for firm quarters with the highest quartile of the earnings surprise.

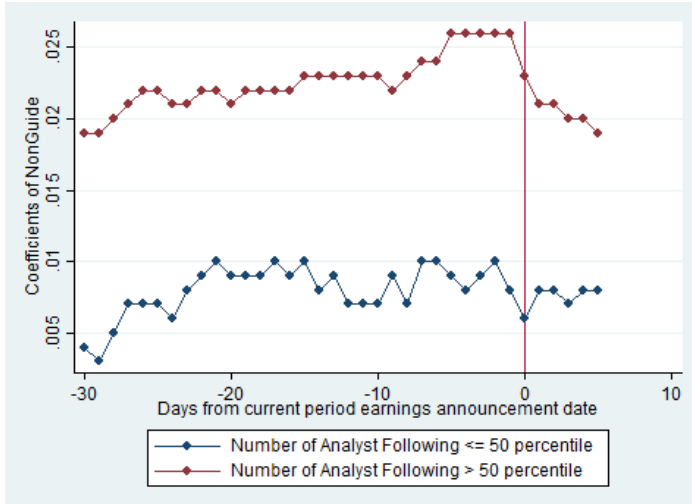


Figure H.8: Coefficients of Non-guidance: by Number of Analysts Following.

Notes: This figure shows the coefficients of non-guidance for regression (1) with control variables from sub-sample results (by number of analysts following). The blue dots (the lower line) are small number of analysts following. The red dots (the higher line) are large number of analysts following. The cutoff is the median number of analysts following for a given reporting month. This figure shows the coefficients of non-guidance for the following regression: $CR_{iqt} = \beta_{0t} + \beta_{1t}ND_{iq} + \beta_{2t}EarnSurp_{iq} + X_{iq} + \varepsilon_{iqt}$, from sub-sample results. CR_{iqt} is firm *i*'s raw cumulative stock returns from 6 days subsequent to the previous quarter earnings announcement date to t days prior to the current quarter earnings announcement; ND_{iq} is a dummy variable that equals to one when management of firm *i* does not provide a forecast in quarter *q*; $EarnSurp_{iq}$ is the earnings news of quarter q for firm i, measured as the difference between current quarter earnings and the same quarter earnings of last year, scaled by the closing stock price five days subsequent to the previous quarter earnings announcement; X_{iq} are the same with those defined in Table G.3. ε_{iqt} is the error term.

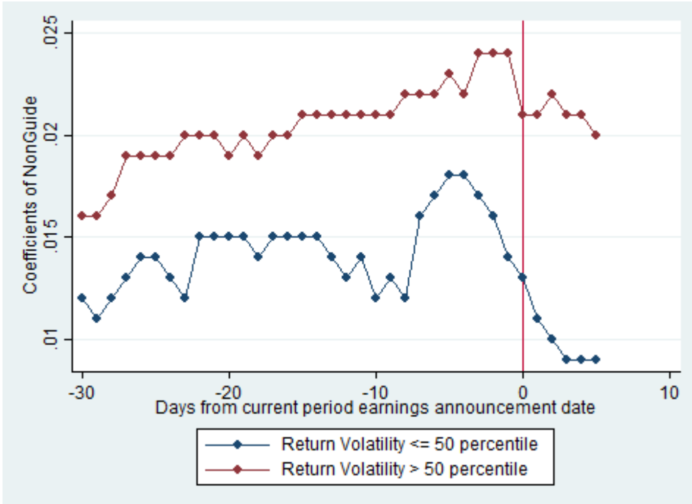


Figure H.9: Coefficients of Non-guidance: by Return Volatility.
Notes: This figure shows the coefficients of non-guidance for regression (1) with control variables from sub-sample results (by return volatility). The blue dots are small return volatility. The red dots are large return volatility. Return volatility is measured as standard deviation of daily stock return in the previous quarter. The cutoff is the median return volatility for a given reporting month. This figure shows the coefficients of non-guidance for the following regression: $CR_{iqt} = \beta_0 + \beta_1 ND_{iq} + \beta_2 EarnSurp_{iq} + X_{iq} + \varepsilon_{iqt}$, from sub-sample results (by return volatility). The blue dots are small return volatility. The red dots are large return volatility. Return volatility is measured as standard deviation of daily stock return in the previous quarter. The cutoff is the median return volatility for a given reporting month. CR_{iqt} is firm *i*'s raw cumulative stock returns from 6 days subsequent to the previous quarter earnings announcement date to *t* days prior to the current quarter earnings announcement; ND_{iq} is a dummy variable that equals to one when management of firm i does not provide a forecast in quarter *q*; $EarnSurp_{iq}$ is the earnings news of quarter *q* for firm *i*, measured as the difference between current quarter earnings and the same quarter earnings of last year, scaled by the closing stock price five days subsequent to the previous quarter earnings announcement; X_{iq} are the same with those defined in Table G.3. ε_{iqt} is the error term.

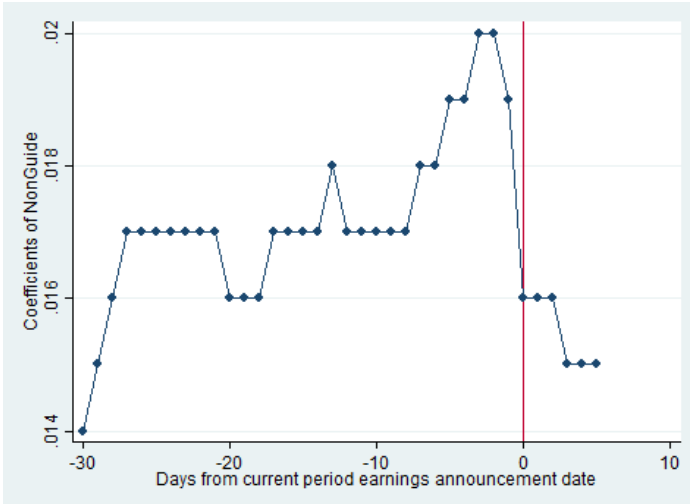


Figure H.10: Coefficients of Non-guidance: Size Adjusted Cumulative Return.
Notes: This figure shows the coefficients of non-guidance for the following regression: $CRS_{iqt} = \beta_{0t} + \beta_{1t}ND_{iq} + \beta_{2t}EarnSurp_{iq} + X_{iq} + \varepsilon_{iqt}$, where CRS_{iqt} is firm *i*'s size adjusted cumulative stock returns from 6 days subsequent to the previous quarter earnings announcement date to t days prior to the current quarter earnings announcement; ND_{iq} is a dummy variable that equals to one when management of firm *i* does not provide a forecast in quarter q; $EarnSurp_{iq}$ is the earnings news of quarter *q* for firm *i*, measured as the difference between current quarter earnings and the same quarter earnings of last year, scaled by the closing stock price five days subsequent to the previous quarter earnings announcement; X_{iq} are the same with those defined in Table G.3. ε_{iqt} is the error term.

BIBLIOGRAPHY

BIBLIOGRAPHY

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. How Much Should We Trust Differences-in-Differences Estimates? Quarterly Journal of Economics 119(1): 249 - 76, 2012.

Dickens, William T. Error Components in Grouped Data: Is It Ever Worth Weighting? Review of Economics and Statistics, 72(2): 328 - 33, 1990.

Dickens, William T. Error Components in Grouped Data: Is It Ever Worth Weighting? Review of Economics and Statistics, 72(2): 328 - 33, 1990.

DuMouchel, William H., and Greg J. Duncan. Using Sample Survey Weights in Multiple Regression Analyses of Stratified Samples. *Journal of the American Statistical Association*, 78(383): 535 - 43, 1983.

Friedberg, Leora. Did Unilateral Divorce Raise Divorce Rates? Evidence From Panel Data, *American Economic Review*, 88(3): 608-627, 1998.

Genadek, Katie R.; Stock, Wendy A. and Stoddard, Christiana. No-Fault Divorce Laws and the Labor Supply of Women with and without Children. *The Journal of Human Resources* Vol. 42, No. 1 pp. 247 - 274, 2007.

Gray, Jeffrey S. Divorce-Law Changes, Household Bargaining, and Married Women's Labor Supply, *American Economic Review*, 88(3): 628 - 642, 1998.

Gruber, Jonathan. Is Making Divorce Easier Bad for Children? The Long-Run Implications of Unilateral Divorce. *Journal of Labor Economics*, 22(4): 799 - 833, 2004.

Haider, Steven; Solon, Gary and Wooldridge, Jeffrey. What Are We Weighting For". *Working Paper No. 18859, National Bureau of Economic Research*, 2013.

Lee, Jin Young and Solon, Gary. The Fragility of Estimated Effects of Unilateral Divorce Laws on Divorce Rates," *B.E. Journal of Economic Analysis and Policy (Contributions)* 11:(1), 2011.

Parkman, Allen. Unilateral Divorce and the Labor-Force Participation rate of Married Women, Revisited" *American Economic Review*, 82(3): 671 - 78, 1992.

Peters, H. Elizabeth. Marriage and Divorce: Informational Constraints and Private Contracting", *American Economic Review*, 76 (3): 437 - 54, 1986.

Rasul, Imran. The Impact of Divorce Laws on Marriage. Unpublished, 2004.

Stevenson, B. Divorce Law and Women's Labor Supply. *Journal of Empirical Legal Studies*, 5: 853 - 873, 2008.

Wolfers, Justin. Did Unilateral Divorce Raise Divorce Rates? A Reconciliation and New Results, *American Economic Review*, 96(5) 1802 - 1820, 2006.

Agan, Amanda. 2011 "Non-Cognitive Skills and Crime," IZA Conference Paper.

Bertrand, Marianne, and Jessica Pan. The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior." *American Economic Journal: Applied Economics*, 5(1): 32 - 64, 2013.

Blair, C., & Razza, R. P. Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*, 78, 647 - 663, 2007.

Brooks-Gunn, J., and G. J. Duncan (Eds.), *The Consequence s of Growing Up Poor* (New York: Russell Sage, 1997).

Brooks-Gunn, J., and G. J. Duncan, Family Poverty, Welfare Reform and Child Development, *Child Development* 71, 188 -196, 2000.

Brooks-Gunn, J., G. J. Duncan, and P. Klebanov, Economic Deprivation and Early-Childhood Development, *Child Development* 65:2, 296 - 318, 1994.

Brooks-Gunn, J., G. J. Duncan, and P. Klebanov, Ethnic Differences in Children's Intelligence Test Scores: Role of Economic Deprivation, Home Environment and Maternal Characteristics," *Child Development* 67, 396 - 408, 1996.

Brooks-Gunn, J., P. K. Klebanov, and G. K. Duncan, "Ethnic Differences in Children's. Intelligence Test Scores: Role of Economic Deprivation, Home Environment, and Maternal Characteristics," *Child Development* 67:2, 396 - 408, 1995.

Burkett, John; McMillen, Marilyn; Owings, Jeffery and Pinkerton, Daniel B. "Statistics in Brief: Making the Cut: Who Meets Highly Selective College Entrance Criteria?" Washington, DC: National Education Longitudinal Study (NELS),1995.

Campbell, Ernest Q.; Coleman, James S.; Hobson, Carol J.; McPartland, James; Mood, Alexander M.; Weinfeld, Frederic D. and York, Robert L. *Equality of Educational Opportunity*. Washington D.C.: U.S. Government Printing Office (GOP), 1966.

Cook, M., and W. Evans, "Families or Schools? Explaining the Convergence in White and Black Academic Performance," *Journal of Labor Economics* 18:4, 729 - 754, 2000.

Cook, P., and J. Ludwig, The Burden of Acting White: Do Black Adolescents Disparage Academic Achievement? (pp. 375–400), in C. Jencks and M. Phillips (Eds.), *The Black-White. Test Score Gap* (Washington, DC: The Brookings Institute, 1998).

Cornwell C, Mustard D, Van Parys J, Noncognitive skills and the gender disparities in test scores and teacher assessments: evidence from primary school, *Journal of Human Resources* 48, 236- 64, 2013.

Cunha, Flavio, James J. Heckman, and Susanne M. Schennach. Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Econometrica*, 78(3): 883 - 931, 2010.

Delpit, L., *Other People's Children: Cultural Conflict in the Classroom* (New York: The New Press, 1995).

Diamond, A. Close interrelation of motor development and cognitive development and of the cerebellum and prefrontal cortex. *Child Development*, 71, 44 - 56, 2000.

Ferguson, R. F., Teachers' Perceptions and Expectations and the Black White Test Score Gap (pp. 273–317), in C. Jencks and M. Phillips (Eds.), *The Black-White Test Score Gap* (Washington, DC: The Brookings Institute, 1998).

Flossman, Anton L., Remi Piatek and Laura Wichert. Going Beyond Returns to Education: The Role of Noncognitive Skills on Wages in Germany. Working Paper, University of Konstanz, ZEW, 2006.

Fordham, S., and J. Ogbu, "Black Students' School Successes: Coping with the Burden of Acting White," *The Urban Review* 18:3, 176 - 206, 1986.

Fryer, R., "An Economic Approach to Cultural Capital," University of Chicago working paper, 2002.

Fryer, Roland G., Jr. and Steven D. Levitt, "Understanding the Black-White Test Score Gap in the First Two Years of School," *Review of Economics and Statistics* 86, 447 - 64, 2004.

Fryer, Roland G., Jr. and Steven D. Levitt, The Black-White Test Score Gap Through Third Grade, *American Law and Economics Review* 8, 249 - 81, 2006.

Fryer, Roland G. Jr., and Steven D. Levitt. Testing for Racial Differences in the Mental Ability of Young Children. *American Economic Review*, 103(2): 981 - 1005, 2013.

Heckman, J. J. Schools, Skills and Synapses, *Economic Inquiry*, 46, 289 - 324, 2008.

Heckman, James J., Yona Rubinstein. The Importance of Noncognitive skills: Lessons from the GED testing program. *American Economic Review*, 91(2): 145 - 149, 2001.

Heckman, James J., Jora Stixrud and Sergio Urzua. The Effects of Cognitive and Non-Cognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, 24(3): 411 - 482, 2006.

Hernstein, R. J., and C. Murray, The Bell Curve: Intelligence and Class Structure in American Life (*The Free Press*, 1994).

Jencks, Christopher and Meredith Phillips, The Black-White Test Score Gap: An Introduction in Christopher Jencks and Meredith Phillips (Eds.), *The Black-White Test Score Gap* (Washington, DC: Brookings Institution Press, 1998).

Jensen, A., How Much Can We Boost IQ and Scholastic Achievement? *Harvard Educational Review* 39 (1969), 1–123. *Educability and Group Differences* (New York: Free Press, 1973). *The G Factor: The Science of Mental Ability* (Greenwood Publishing Group, 1998).

Mayer, S. E., *What Money Can't Buy: Family Income and Children's Life Chances* (Harvard University Press, 1997).

Neal, D., and W. R. Johnson, The Role of Pre-market Factors in Black-White Wage Differences. *Journal of Political Economy* 104, 869 – 895, 1996.

O'Neill, J., The Role of Human Capital in Earnings Differences between Black and White Men." *Journal of Economic Perspectives* 4:4, 25 - 46, 1990.

Raver, C. C., Gershoff, E. T., & Aber, J. L. Testing equivalence of mediating models of income, parenting, and school readiness for White, Black, and Hispanic children in a national sample. *Child Development*, 78, 96 - 115, 2007.

Rodgers, W., and W. Spriggs, What Does AFQT Really Measure: Race, Wages, Schooling and the AFQT Score, *The Review of Black Political Economy* 24:4, 13 - 46, 1996.

Rushton, Philippe J. Race and crime: international data for 1989-1990. *Psychological Reports*, 76(1), 307 - 12, 1995.

Scott, R. and Sinclair, D. "Ethnic-related cognitive profiles of black and white preschool children." *Homo- Journal of Comparative Human Biology*, 28, 116-20, 1996.

Segal, Carmit. Misbehavior, Education and Labor Market Outcomes. The Journal of the European Economic Association, Forthcoming, 2013

Steele, C., and J. Aronson, Stereotype Threat and the Test Performance of Academically Successful African Americans (pp. 401–430), in C. Jencks and M. Phillips (Eds.), *The Black-White Test Score Gap* (Washington, DC: The Brookings Institute, 1998).

Akerlof, G. The market for lemons: Quality uncertainty and the market mechanism. Springer, 1995.

Amihud, Y. Illiquidity and stock returns: cross-section and time-series effects. Journal of Financial Markets 5, 31 - 56, 2002.

Anilowski, C., Feng, M., Skinner, D. J. Does earnings guidance affect market returns? The nature and information content of aggregate earnings guidance. Journal of Accounting and Economics 44, 36 - 63, 2007.

Ball, R., Brown, P. An empirical evaluation of accounting income numbers. Journal of Accounting Research 6, 159 - 178, 1968

Ball, R., Shivakumar, L. How much new information is there in earnings? Journal of Accounting Research 46, 975 - 1016, 2008.

Ball, R. T., Easton, P. Dissecting earnings recognition timeliness. Journal of Accounting Research 51, 1099 - 1132, 2013.

Barberis, N. Investing for the long run when returns are predictable. The Journa l of Finance 55, 225 - 264, 2000.

Barberis, N., Shleifer, A., Vishny, R. A model of investor sentiment. Journal of Financial Economics 49, 307 - 343, 1998.

Beaver, W. H. The information content of annual earnings announcements. Journal of Accounting Research 6, 67 – 92, 1968.

Bernard, V. L., Thomas, J. K. Post-earnings-announcement drift: delayed price response or risk premium? Journal of Accounting Research 27, 1 – 36, 1989.

Brown, A. L., Camerer, C. F., Lovallo, D. To review or not to review? Limited strategic thinking at the movie box office. American Economic Journal: Microeconomics 4, 1 - 26, 2012.

Brown, S., Hillegeist, S. A., Lo, K. The effect of earnings surprises on information asymmetry.

Journal of Accounting and Economics 47, 208 – 225, 2009. Carhart, M. M. On persistence in mutual fund performance. The Journal of finance 52, 57 - 82, 1997.

Chen, S., Matsumoto, D., Rajgopal, S. Is silence golden? An empirical analysis of firms that stop giving quarterly earnings guidance. Journal of Accounting and Economics 51, 134 - 150, 2011.

Choi, J.-H., Myers, L. A., Zang, Y., Ziebart, D. A. Do management EPS forecasts allow returns to reflect future earnings? Implications for the continuation of management's quarterly earnings guidance. Review of Accounting Studies 16, 143 - 182, 2011.

Chuk, E., Matsumoto, D., Miller, G. S. Assessing methods of identifying management fore- casts: CIG vs. researcher collected. Journal of Accounting and Economics 55, 23 - 42, 2013.

Dye, R. A. Disclosure of nonproprietary information. Journal of Accounting Research 23, 123 - 145, 1985.

Einhorn, E. Voluntary disclosure under uncertainty about the reporting objective. Journal of Accounting and Economics 43, 245 – 274, 2007.

Eyster, E., Rabin, M. Cursed equilibrium. Econometrica 73, 1623 – 1672, 2005.

Giglio, S., Shue, K. No news is news: Do markets underreact to nothing? Review of Financial Studies 27, 3389 - 3440, 2014.

Grossman, S. J. The informational role of warranties and private disclosure about product quality. Journal of Law and Economics 24, 461 - 483, 1981.

Hirst, D. E., Koonce, L., Venkataraman, S. Management earnings forecasts: A review and framework. Accounting Horizons 22, 315 - 338, 2008

Houston, J. F., Lev, B., Tucker, J. W. To guide or not to guide? Causes and consequences of stopping quarterly earnings guidance. Contemporary Accounting Research 27, 143 - 185, 2010.

Hutton, A. P., Miller, G. S., Skinner, D. J. The role of supplementary statements with management earnings forecasts. Journal of Accounting Research pp. 867 - 890, 2003.

Hutton, A. P., Stocken, P. C. Prior forecasting accuracy and investor reaction to management earnings forecasts. Working paper, available at SSRN 817108, 2009.

Jin, G. Z., Luca, M., Martin, D. Is no news (perceived as) bad news? An experimental investigation of information disclosure. National Bureau of Economic Research Working Paper, 2016.

Kasznik, R., Lev, B. To warn or not to warn: Management disclosures in the face of an earnings surprise. The Accounting Review 70, 113 - 134, 1995.

Kim, O., Verrecchia, R. E. Trading volume and price reactions to public announcements. Journal of Accounting Research 29, 302 - 321, 1991.

Kothari, S. P., Shu, S., Wysocki, P. D. Do managers withhold bad news? Journal of Accounting Research 47, 241 - 276, 2009.

Kyle, A. S. Continuous auctions and insider trading. Econometrica: Journal of the Econometric Society 53, 1315 - 1335, 1985.

Lennox, C. S., Park, C. W. The informativeness of earnings and management's issuance of earnings forecasts. Journal of Accounting and Economics 42, 439 - 458, 2006.

Milgrom, P. R. Good news and bad news: Representation theorems and applications. The Bell Journal of Economics 12, 380 - 391, 1981.

Ng, J., Tuna, I., Verdi, R. Management forecast credibility and underreaction to news. Review of Accounting Studies 18, 956 - 986, 2013.

Patell, J. M. Corporate forecasts of earnings per share and stock price behavior: Empirical test. Journal of Accounting Research 14, 246 - 276, 1976.

Payzan-LeNestour, E., Bossaerts, P. Learning about unstable, publicly unobservable payoffs. Review of Financial Studies 28, 1874 - 1913, 2015.

Penman, S. H. An empirical investigation of the voluntary disclosure of corporate earnings forecasts. Journal of accounting research 18, 132 – 160, 1980.

Rogers, J. L., Skinner, D. J., Van Buskirk, A. Earnings guidance and market uncertainty. Journal of Accounting and Economics 48, 90 – 109, 2009.

Serrano-Padial, R. Naive traders and mispricing in prediction markets. Journal of Economic Theory 147, 1882 – 1912, 2012.

Skinner, D. J. Why firms voluntarily disclose bad news. Journal of Accounting Research 32, 38 - 60, 1994

Sletten, E. The effect of stock price on discretionary disclosure. Review of Accounting Studies 17, 96 – 133, 2012.

So, E. C., Weber, J. Time will tell: Information in the timing of scheduled earnings news. Available at SSRN 2480662, 2015.

Tucker, J. W. Is openness penalized? Stock returns around earnings warnings. The Accounting Review 82, 1055 - 1087, 2007.

Verrecchia, R. E. Discretionary disclosure. Journal of Accounting and Economics 5, 179 - 194, 1983.

Waymire, G. Earnings volatility and voluntary management forecast disclosure. Journal of Accounting Research 268 -295, 1985.

Zhang, F. X. Information uncertainty and stock returns. The Journal of Finance 61, 105 - 137, 2006.

Zhou, F. S. Disclosure dynamics and investor Learning. Working paper, 201 URL http://home.uchicago.edu/ frankzhou8674/index.html.