Public and Private Learning in the Market for Teachers: Evidence from the Adoption of Value-Added Measures

Michael Bates *

December 2, 2016

Abstract

Informational asymmetries between employers may inhibit optimal worker mobility. However, evidence is limited because researchers rarely observe shocks to employers' information. I exploit two school districts' adoptions of value-added (VA) measures of teacher effectiveness—informational shocks to some, but not all, employers—to provide direct tests of asymmetric employer learning. I develop a learning model and test its predictions for teacher mobility. I find that adopting VA increases within-district mobility of high-VA teachers, while low-VA teachers move out-of-district to uninformed principals. These patterns evidence asymmetric employer learning. This sorting from widespread VA adoption exacerbates inequality in access to effective teaching. (*JEL* D83 I24 J63)

*Department of Economics, University of California at Riverside, Riverside, CA 92521, United States of America, (email: michael.bates@ucr.edu). I would like to express my sincere gratitude to Todd Elder for his guidance throughout the process of this project. I would also like to thank Mike Conlin, Scott Imberman, and Jeffrey Wooldridge for their thoughtful advice. Thanks also goes to Soren Anderson, Brad Hershbein, Mindy Marks, Michael Podgurski, Michael Waldman, and participants at the 2014 UM-MSU-UWO Labor Day Conference, the Causal Inference in Education Research at the University of Michigan, the University of California at Riverside, Michigan State University, University of Illinois at Urbana Champaign, and the NBER Fall Education Meetings for their helpful comments and discussion. I also thank Kara Bonneau and the North Carolina Education Research Data Center as well as representatives of Guilford County Schools, Winston Salem/Forsyth Community Schools, Charlotte-Mecklenburg Schools, and Cumberland County Schools. All errors are my own.

1 Introduction

Incomplete information inhibits the market from achieving the optimal allocation of workers across employers (Spence, 1973; Jovanovic, 1979; Gibbons and Katz, 1991; Farber and Gibbons, 1996; Altonji and Pierret, 2001). While a large literature focuses on informational asymmetries between workers and employers, Waldman (1984) and Greenwald (1986) gave rise to another literature focusing on asymmetric information between current and prospective employers. Prior empirical studies uses models of asymmetric employer learning to explain empirical facts, such as wage dynamics with respect to job tenure versus experience, variability of wages after a job loss, and selection of mobile or promoted workers on easy or difficult to observe characteristics (Schönberg, 2007; Pinkston, 2009; DeVaro and Waldman, 2012; Kahn, 2013). If the current employer enjoys an informational advantage over other prospective employers, it becomes a monopsonist of that information, permitting persistent gaps between workers' wages and their marginal products of labor (Milgrom and Oster, 1987). Furthermore, workers may not flow to the employers or positions at which they would be most productive (Waldman, 1984; Greenwald, 1986).

Despite these important implications and the intuitive appeal of the theory, the existing evidence is mixed. Further, it is limited by an absence of direct measures of productivity, and more importantly, a lack of exogenous variation in the informational landscape in which employers operate. This work seeks to fill this gap. I use the release of worker-level performance data to some, but not all, employers as a unique natural experiment, to test the degree to which the information spreads among employers, whether mobility responds in accordance with theory, and the type of learning that had previously prevailed.

I develop models of public and private employer learning in the context of the market for middle and elementary school teachers. I then use statewide, micro-level, administrative data from North Carolina to formulate value-added (VA) measures of teacher productivity.[1] Lastly, I exploit the adoption of teacher VA by two of the largest school districts in the state, which provides an asymmetric shock to employers' information sets, to provide a direct test of asymmetric employer learning. Thus, this setting allows me to disentangle employer learning from other forms of human capital accumulation.

The adoption of VA in North Carolina provides a context with rich informational variation to examine employer learning. Each of the two large districts that adopted VA did so in different ways and separately from the rest of the state. This provides three different

[1] VA measures calculate how much a teachers' students learn in comparison to how much those students are expected to learn. There are several methods for estimating VA. I do not have access to the exact ones issued to teachers and principals. I estimate teacher VA using multiple methods. The primary specification estimates teacher fixed effects in the regression of student test scores on student covariates including past test scores. Results are robust to alternative formulations of VA.

informational landscapes: one in Guilford County Schools (to be referred to as Guilford), where the teacher, the current (or retaining) principal, and any hiring principal within the district were given direct access to the teacher's VA; one in Winston Salem/Forsyth Community Schools (to be referred to as Winston-Salem), in which only teachers and their current principals received value-added reports; and lastly, in the rest of the state, where the information structure remained relatively constant. Examining how the relationship between teacher quality and teacher mobility changes within and across these settings reveals the degree to which VA was informative, and spread throughout the market.

If VA measures are informative, they provide teachers with a signal of their own ability. Thus, the model predicts that VA measures increase the likelihood that effective teachers move from one school to another within the districts where the signals are public. If the information spreads easily through the market there should be no difference between the impacts of VA for moves within-district and out of Guilford or Winston-Salem. However, if retaining principals keep teachers' VA measures private, ineffective teachers may become more likely to move out-of-district. Thus, the asymmetric employer learning model predicts adverse selection of teachers out-of-district.

Understanding informational asymmetries in the teacher labor market is also important in its own right, as there are currently an estimated 3.1 million teachers employed in the United States (NCES, 2016). Further, previous findings that effective teachers have large, meaningful impacts on the lives of their students, though there is wide variation in the teachers' ability to do so (Chetty et al., 2011, 2014). While Staiger and Rockoff (2010) and Rivkin et al. (2005) illustrate the difficulty in identifying effective teachers at the point of hire, Jacob and Lefgren (2008); Chingos and West (2011), and Rockoff et al. (2012) each present evidence of principals learning about the quality of their teaching force. However, there is little understanding of how much of that information spreads to principals of other schools nor how widespread changes in available information about teacher quality may change teacher mobility.

In the teacher labor market, wage rigidities force the market to clear on other amenities. A large literature demonstrates that in general teachers prefer to teach in schools that are closer in proximity to their homes, higher performing, and for white teachers, schools with a lower percentage of black students (Boyd et al., 2008; Jackson, 2009; Boyd et al., 2013). Consequently, as VA signals provide good teachers with more choice over where to teach, they may also exacerbate the divide in access to high quality education. This work provides the first examination of whether the release of VA leads to further sorting of teachers to schools. Rising inequity may be an important consequence of the policy that has been previously overlooked.

Using differences-in-differences analysis, I find that by releasing VA measures to teachers and principals, both districts increase the probability that high-VA teachers will move within district, particularly to higher-performing schools. I estimate that the release of VA increases the probability that a teacher with a one standard deviation higher VA moves within-district to higher-performing schools by about 10% suggesting that VA provided new public information into those markets. I find that the selection of mobile teachers due to adopting VA is less positive for teachers moving to schools outside of Guilford and Winston-Salem. The policy leads teachers who are a full standard deviation below average to become roughly 30% more likely to move from Guilford to a higher-performing school in the rest of the state. In Winston-Salem, the effect of the policy on the probability that a high-VA teacher moves to a higher-performing school is 60% smaller for teachers moving out-of-district than it is for teachers moving within-district. The fact that we see positive selection to principals with access to the information and much smaller effects and even negative selection for moves to those without access to the VA measures is consistent with asymmetric employer learning.

This rising mobility of effective teachers to high-performing schools evidences rising inequality in the distribution of teachers in the market. These results are reinforced with similar teacher mobility away from schools with higher shares of black students. Further, I find increased growth in school performance for high-VA teachers, particularly in Winston-Salem. Given that 38 states currently require teacher evaluations to incorporate teachers' impacts on student achievement on standardized exams, this threat to educational equity is an important and perhaps widespread unintended consequence.

2 Setting

Shocks to the information available on workers' productivity are rare. Shocks to the information of some, but not all, employers in a market are rarer still. To my knowledge, this is the first study directly testing a general model of public and private learning by exploiting information shocks to a large, important labor market. Guilford County Schools (Guilford) contracted with SAS (originally called "Statistical Analysis System") to receive teacher EVAAS (Education Value-Added Assessment System) measures of teacher effectiveness in 2000. These measures are based on the model developed by Sanders et al. (1997) under the name "Tennessee Value-Added Assessment System" (TVAAS). In fact, the adoption of VA by Guilford accompanied the transition of TVAAS to EVAAS, as the system came under the management of SAS, which began at North Carolina State University. The district gave teachers, principals, and hiring principals within the district direct access to these teacher VA measures. Consequently for moves within Guilford, the introduction of VA provides a shock to the public information.

The rest of the state of North Carolina adopted EVAAS measures of school effectiveness in 2008. Winston-Salem/Forsyth Community Schools (Winston-Salem) took an additional step, providing SAS with student-teacher matches necessary to receive the same teacher specific measure of effectiveness already present in Guilford. In Winston-Salem, only the teachers and their own principals directly received the VA reports. The VA measures were not directly given to principals at other schools in the district.

However, the introduction of VA in Winston-Salem is theoretically also public. As in Grossman (1981) and Milgrom (1981), each teacher contemplating moving within the district has as incentive to voluntarily disclose his score. Because all principals in the district know that the VA exists, if a teacher chooses not to reveal his score, hiring principals within the district assume that he is as good as the average teacher who chooses not to reveal his score. Consequently, all teachers with scores above that average have an incentive reveal their scores. The average score of those who do not disclose drops until only teachers with scores at the minimum are indifferent between revealing and keeping the information private. If teachers act as predicted, all teachers voluntarily disclose their EVAAS reports, and the VA alter the information available to both current and hiring principals within Winston-Salem, just as they do in Guilford.

The setting and incentives teachers face differs when moving out of Guilford and Winston-Salem. It is possible that hiring principals in the rest of the state are unaware of the existence of an applying teacher's EVAAS report. Consequently, a teacher may withhold his signal and leave the principal's expectation of his ability unchanged. This informational asymmetry may be avoided by principals thoroughly researching from where their applicants are coming. In which case, the same predictions as were formulated for within-district moves would apply. However, such acquisition of information is costly, and principals may forgo it. Thus, the test between symmetric and asymmetric learning hinges on whether the adoption of VA leads the selection of out-of-district mobile teachers to be significantly more negative than its effects on the selection of within-district movers.

Since principals in both Guilford and Winston-Salem received training about the measures, VA likely served as a more salient signal for principals within the district than for those in the rest of the state. Out-of-district hiring principals may have placed particularly low weight in the measure early in Guilford's adoption of VA. Guilford contracted with SAS just two years after the creation of the EVAAS system, and two years before before the passage of No Child Left Behind. VA were largely absent from education policy discussions. The salience of the signal was likely less of an issue for teachers moving from Winston-Salem, considering school-level EVAAS measures were implemented across the entire state the same year. This may lead the learning results for out-of-district moves to be more pronounced for Guilford than they are for teachers leaving Winston-Salem.

3 Model

This section describes a simple model to illustrate the basic intuition and provide primary predictions for which workers move, and where they go–and how each may change in response to an information shock. Please see Appendix 8.1 for proofs of the predictions of this model. I provide a more comprehensive model in Appendix 8.2, which allows for public learning (nesting symmetric learning as a special case) and allows more realistic dynamics between competing principals and teachers. It also covers additional extensions, which the data allow me test. The primary predictions hold under both versions of the model. These models build on the model of asymmetric employer learning presented in Pinkston (2009), primarily by endogenizing worker mobility, and incorporating discreet information shocks into the continuous learning process. Additional changes to the model allow it to more closely fit this particular labor market. I will highlight peculiarities in the market for primary school teachers and the model structures that accompany them.

3.1 Model Structure

There are two broad classifications of principals: those who are hiring (denoted by the superscript h); and those who are retaining teachers (denoted by the superscript r). Each period, teachers receive two offers, and move to schools that maximize their utility. In the first period both principals are hiring principals. Each subsequent period, teachers receive an offer from their retaining principal and an outside offer from either a principal within or outside of the current district with a given probability.[2] These offers reflect principals' expectations about the effectiveness of the teacher, which is based upon the information available. I itemize the information structure below:

1. True effectiveness is not observable to employers, but is given by, $\mu = m + \epsilon$, where m is observable and is the mean productivity among a worker's reference group and ϵ is mean 0 with variance of σ_ϵ).
2. Private signal:
    (a) For hiring principals (denoted by the superscript h), the private signal is given by$P^h = \mu + \tau^h$ where τ^h is mean 0 and variance $\sigma_\tau(0)$). $\sigma_\tau(0)$ is fixed over time.

[2] Principals face rigid budget constraints, which translate to a fixed number of positions.

(b) For a retaining principal (denoted by the superscript r), the private signal is given by $P_t^r = \mu + \tau_t^r$ where τ_t^r is mean 0 and variance $\sigma_\tau(t))$ and $\frac{\partial \sigma_\tau(t)}{\partial t} < 0$.

3. The VA serves as an additional signal with the form $V = \mu + \nu$, where ν has a mean of 0 and a variance of σ_ν).

4. The noise of each signal is orthogonal to the noise of the other signals.[3]

I assume that teachers know their effectiveness (μ), but cannot credibly reveal it. As a teacher begins her career, all principals begin with the prior belief that she is as good as the average teacher with her same characteristics (m). The teacher encounters two principals to whom he may privately (but noisily) signal his ability akin to an interview, (denoted by P_0^h where 0 indicates no additional private information).

Through interactions, observations, and/or attention to outcomes, retaining principals may obtain private information unavailable to rival employers (P_t^r) the longer a teacher teaches within the school (t). If such private learning occurs, the precision of the current principal's signal ($\sigma_\tau(t)$) increases the longer a teacher works in the school, while hiring principals' private signals from interviewing the teacher have a constantly high variance ($\sigma_\tau(0)$). Thus, the accumulation of private information leads to $\sigma_\tau(t) < \sigma_\tau(0)$ for all $t > 0$. In order to nest symmetric learning within the more flexible model, I maintain that that even in this special case, employers receive a private signal each period, but the variance of the signal is constant over years of tenure ($\sigma_\tau(t) = \sigma_\tau(0)$ for all $t > 0$). VA enters the learning model as an additional signal that enters both principals' expectations, if both principals receive it, or only only the retaining principal's expectation, if it is only accessible to her.[4]

3.2 Bidding

The teacher labor market generally moves in the summer between school years. At that time, teachers may sample two offers, an update from their current school and one outside offer. In many public education systems, strict salary schedules determines teachers' pay. In North Carolina, the state sets a base salary schedule that depends exclusively upon easily observable characteristics, such as education and experience.[5] Districts supplement this base amount with a percentage of the base schedule. In general, this means that principals cannot differentially pay teachers within their school on the basis of perceived performance. [6] While principals cannot adjust salaries to influence whether a teacher stays, principals

[3]The orthogonality assumptions are also not necessary to derive the following predictions. However, relaxing these require a less restrictive, though more complicated set of assumptions, outlining the direction and magnitude of correlations between the errors of the signals.

[4]The EVAAS VA measures included a multi-year average of teachers' VA as well as a history of past year-by-year VA.

[5]As of 2014, North Carolina will move to paying teachers in part based upon teachers' VA.

[6]In Section 6, I discuss policy exceptions to this in North Carolina school districts.

may influence school staffing through non-pecuniary position attributes, such as planning time, teaching assignments, or additional requirements.

In the context of the model, this means teachers take the position that offers the highest total compensation, which is comprised of salary set by district, characteristics of school, and characteristics of position. In the simple model, I assume that each principal presents a bid of total compensation equal to their expectation of the teacher's effectiveness, under a sealed second price auction.[7] Principals formulate these expectations by averaging over the signals they receive (initially just m and P_0^h or P_t^r). In accordance to standard Bayesian updating, they weight each signal by its precision relative to the other signals. I list the bids of hiring (b_{NV}^{h*}) and retaining (b_{NV}^{r*}) principals in equations 1 and 2 respectively.[8]

$$b_{NV}^{h*} = \frac{\sigma_\tau(0)}{Z_{NV}^h} m + \frac{\sigma_\epsilon}{Z_{NV}^h} P_0^h.^{9} \tag{1}$$

$$b_{NV}^{r*} = \frac{\sigma_\tau(t)}{Z_{NV}^r} m + \frac{\sigma_\epsilon}{Z_{NV}^r} P_t^r.^{10} \tag{2}$$

If there is private learning, only retaining principals place more weight to their private signals, (P_t^r), while placing less weight on the prior belief. This is reflected by $\sigma_\tau(t)$ in equation 2, which shrinks with additional private information as opposed to $\sigma_\tau(0)$ from equation 1, which remains constant for hiring principals. Thus, the bids diverge with additional private information.

If a principal's rival is from outside of the district and uninformed of the measure, when a retaining principal receives a teacher's VA, she incorporates it into her private signal (denoted by the subscript RV). The new private signal ($P_{t\nu}^r$) becomes the precision-weighted average of the prior private information and the new VA.[11] In which case, the retaining principal's optimal bid is shown in equation 3, while the hiring principal's bid remains unchanged from equation 1.

$$b_{RV}^{r*} = \frac{\sigma_\tau(t\,V)}{Z_{RV}^r} m + \frac{\sigma_\epsilon}{Z_{RV}^r} P_{t\nu}^r.^{12} \tag{3}$$

Equation 3 is similar to equation 2 except for the replacement of P_t^r by $P_{t\nu}^r$ and of $\sigma_\tau(t)$ by $\sigma_\tau(t\,V)$. If VA is informative, the precision of the cumulative private information must increase, as shown by Lemma 1.

[7] In Appendix 8.2, I relax this assumption allowing each principal to view and update their offers conditioning on the rival's offer. The basic predictions follow under either bidding process.

[8] Subscript NV indicates neither principal received the teacher's VA.

[9] $Z_{NV}^h = \sigma_\tau(0) + \sigma_\epsilon$.

[10] $Z_{NV}^r = \sigma_\tau(t) + \sigma_\epsilon$.

[11] $P_{t\nu}^r = \frac{\sigma_\nu P_t^r + \sigma_\tau(t) V}{\sigma_\nu + \sigma_\tau(t)}$.

[12] $Z_{RV}^r = \sigma_\tau(t\,V) + \sigma_\epsilon$.

Lemma 1: The precision of the private signal increases with the incorporation of VA into the private signal ($\sigma_\tau(t\,V) < \sigma_\tau(t)$).

Proof: Under the orthogonality assumptions, $var(P_{t\nu}) \equiv \sigma_\tau(t\,V) = \frac{\sigma_\nu^2\sigma_\tau(t)+\sigma_\nu\sigma_\tau(t)^2}{(\sigma_\nu+\sigma_\tau(t))^2} = \frac{\sigma_\nu\sigma_\tau(t)}{\sigma_\nu+\sigma_\tau(t)} \cdot \frac{\sigma_\tau(t)(\sigma_\nu+\sigma_\tau(t))}{\sigma_\nu+\sigma_\tau(t)} - \frac{\sigma_\nu\sigma_\tau(t)}{\sigma_\nu+\sigma_\tau(t)} = \frac{\sigma_\tau^2(t)}{\sigma_\nu+\sigma_\tau(t)}$, and $\frac{\sigma_\tau^2(t)}{\sigma_\nu+\sigma_\tau(t)} > 0$, by property of variances.

This decrease in the variance of the private signal decreases the weight retaining principals place on their prior beliefs and the public information, while increasing the relative weight they place on their now fuller private information. Since the hiring principals' expectations do not change, the introduction of VA exacerbates informational asymmetries between prospective employers, and the two principals' bids further diverge.

In contrast, the VA becomes a public signal, if both bidding principals are informed of a teacher's VA (as occurs when both principals are within districts adopting VA). I list the optimal bids of hiring and retaining principals' when both have access to a teacher's VA in equations 4 and 5 respectively.

$$b_{HV}^{h*} = \frac{\sigma_\tau(0)\sigma_\nu}{Z_{HV}^h}m + \frac{\sigma_\tau(0)\sigma_\epsilon}{Z_{HV}^h}V + \frac{\sigma_\epsilon\sigma_\nu}{Z_{HV}^h}P_0^h.^{13} \tag{4}$$

$$b_{HV}^{r*} = \frac{\sigma_\tau(t)\sigma_\nu}{Z_{HV}^r}m + \frac{\sigma_\tau(t)\sigma_\epsilon}{Z_{HV}^r}V + \frac{\sigma_\epsilon\sigma_\nu}{Z_{HV}^r}P_t^r.^{14} \tag{5}$$

Equations 4 and 5 are standard Bayesian expectations with VA (V) serving as a third signal.The introduction of VA (V) shift weight from the private signals (P) to the public information contained in V and the prior (m), so long as neither m nor P perfectly capture true ability (μ). Thus, for bids in which both principals become informed of a teacher's VA, the information between prospective employers becomes more symmetric, and their expectations converge, increasing the chance of mobility.

3.3 Mobility with the introduction of VA

After teachers receive both bids, they move to the school that offers the highest bid.[15] Accordingly, the probability of a move is:

$$P(M) = P\left[b^{h*} - b^{r*} > 0\right]. \tag{6}$$

The availability of VA to some prospective employers, but not others, provides a rare test for the model laid out above. What predictions does this model provide about how teacher

[13] $Z_{HV}^h = \sigma_\tau(0)\sigma_\nu + \sigma_\tau(0)\sigma_\epsilon + \sigma_\epsilon\sigma_\nu$.
[14] $Z_{HV}^r = \sigma_\tau(t)\sigma_\nu + \sigma_\tau(t)\sigma_\epsilon + \sigma_\epsilon\sigma_\nu$.
[15] For simplicity, I model mobility decisions as a spot market. A fixed transition cost or idiosyncratic teacher preferences over schools may be added without additional assumptions.

mobility will change with the adoption of VA? As described in Section 2, both districts' adoptions of VA provide a shock to the information of all principals within the district. Thus, by examining teacher mobility in response to the release of VA, I test whether releasing VA leads toward informational symmetry between employers. However, out-of-district principals cannot directly access the new VA measures. Thus, examining mobility out of adopting districts evidences whether the information spreads to all employers or furthers informational asymmetries between them.

There are two primary ways of thinking about the impact of VA in the model. The first is more in keeping with the prior employer learning literature. Empirically, VA measures serve as difficult-to-observe measures of teacher quality, which researchers may use to proxy for μ about which employers are learning. The information shock primarily comes through the change in variances of employers' signals. In this framework, the model offers predictions of whether better or worse teachers move as response to adopting VA. Equation 7 takes this broad view.[16]

$$\begin{aligned}&\frac{\partial E[b^{h*}_{HV} - b^{r*}_{HV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ \mu]}{\partial \mu} = \\ &\frac{{\sigma_\epsilon}^2(\sigma_\tau(0) - \sigma_\tau(t))}{Z^h_{NV} Z^r_{NV} Z^h_{HV} Z^r_{HV}}(\sigma_\tau(0)\sigma_\tau(t)\sigma_\epsilon + 2\sigma_\nu\sigma_\tau(0)\sigma_\tau(t) + \sigma_\epsilon\sigma_\nu\sigma_\tau(0) + \sigma_\epsilon\sigma_\nu\sigma_\tau(t)) > 0.\end{aligned} \tag{7}$$

Under the assumption that $\sigma_\tau(0) > \sigma_\tau(t)$, which is fundamental to asymmetric employer learning $\frac{\partial E\left[b^{h*}_{HV} - b^{r*}_{HV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ \mu\right]}{\partial \mu} > 0$. Therefore, the model predicts that providing VA to both principals, as occurred within both districts, should raise the probability that good teachers move, all else equal.

Under the second interpretation, EVAAS VA enters the two districts directly as new signals. Accordingly, the model offers predictions on the differential effects of the policy on the probability of moving for teachers receiving different signals, all else equal. After some algebra, equation 8 takes this more narrow view.[17]

$$\frac{\partial E\left[b^{h*}_{HV} - b^{r*}_{HV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ V\ \mu\right]}{\partial V} = \frac{{\sigma_\epsilon}^2\sigma_\nu(\sigma_\tau(0) - \sigma_\tau(t))}{Z^h_{HV} Z^r_{HV}} > 0. \tag{8}$$

Within the districts, where both principals are aware of the signals, the model predicts releasing VA increases the probability that teachers who receive high-VA signals will transfer schools. While the interpretations are subtly different, the comparative statics with respect to VA after the policy takes effect are the same. In both instances, the predicted increase

[16]See Appendix 8.1.1 for proof.
[17]See Appendix 8.1.2 for proof.

mobility of effective (or high-VA) teachers results from informational differences between employers.

Recall from Section 2, that if principals in other districts know of the existence of VA for teachers from Winston-Salem and Guilford, the policy would theoretically alter their information. In this context, the previous predictions would apply to out-of-district moves as well. However, it is plausible that principals in other districts were uninformed about the policy. In which case, VA enters retaining principals' private signals in Guilford and Winston-Salem, making the balance of information more asymmetric between retaining and out-of-district principals.

The same two interpretations of VA apply here. I will first take the broader view of VA with equation 9 demonstrating the predicted change in the relationship between teachers' underlying abilities and the probability of moving to uninformed principals once districts release their teachers' VA. Equation 10 presents the partial derivative of the expected difference in the differences between employers bids with respect to the VA signal itself.

$$\frac{\partial E[b^{h*}_{NV} - b^{r*}_{RV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ \mu]}{\partial \mu} = \frac{\sigma_\epsilon(\sigma_\tau(t\ V) - \sigma_\tau(t))}{Z^r_{NV} Z^r_{RV}} < 0.^{18} \tag{9}$$

$$\frac{\partial E\left[b^{h*}_{HV} - b^{r*}_{HV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ V\ \mu\right]}{\partial V} = \frac{-\sigma_\epsilon \sigma_\tau(t)}{Z^r_{RV}(\sigma_\nu + \sigma_\tau(t))} < 0.^{19} \tag{10}$$

Again the predictions remain consistent. Under lemma 1, $\sigma_\tau(t) > \sigma_\tau(t\ V)$, which implies that $\frac{\partial E\left[b^{h*}_{NV} - b^{r*}_{RV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ \mu\right]}{\partial \mu} < 0$. Therefore, the model predicts that the release of VA to retaining principals increases the likelihood that ineffective teachers move out-of-district, and vice versa. Equation 10 shows the partial derivative with respect to the signal is the negative of positive variances and is accordingly negative. Thus, the policy leads to adverse selection of out-of-district moving teachers on the basis of the signal or underlying effectiveness, all else equal.

It is important to note that good (or high-VA) teachers may choose to reveal their EVAAS report to principals in other districts in an effort to move out-of-district. Accordingly, the furthering of information asymmetries between employers may not universally apply to out-of-district moves. However, as long as some low-VA teachers are able to move out-of-district without being penalized by their EVAAS report (or their unwillingness to reveal it), the model predicts more negative (smaller in magnitude or negative) effects of VA on the probability of moving out-of-district after policy implementation than are produced for moves within-district.[20] Thus, the test between symmetric and asymmetric learning is whether

[19] See Appendix 8.1.4 for proof.

[20] See Appendix 8.2.7 for a more formal treatment of this.

effects of the policy on the selection of out-of district movers are significantly more negative than the effects of adopting VA on the selection of within-district movers.

The different predictions for within and out-of-district mobility are the primary predictions and follow from this simple model. However, in its simplicity this model makes some restrictive assumptions. Most notably that principals bid in accordance to silent second price auctions and that there is no public learning other than through VA. Appendix 8.2 relaxes these assumptions allowing principals to match offers and also incorporates a public signal that evolves with teachers' experience as well as adding some further complexity to better accommodate this setting. In so doing, this richer model also provides additional predictions. For instance, it demonstrates that the mobility predictions apply more strongly to moves to more desirable schools, differences in mobility that may result from differences between districts in VA policy implementation, dynamics regarding observable teacher characteristics, and dynamics with respect to years of tenure at retaining schools. Table 1 presents a summary of the predictions and corresponding key assumptions, tables of evidence, and appendices for proofs and additional description of these predictions.

Table 1: Model Predictions

	Primary Model Predictions	Assumptions: There was prior private learning,VA is informative, and...	Parameterized Predictions	Table	Appendix
1.	Better (higher-VA) teachers will become more likely to move within district (subscript WD) after the adoption of VA.		$\gamma_{14WD} > 0$	3	8.1.1 (8.1.2) 8.2.3 (8.2.4)
2.	Worse (lower-VA) teachers will become more likely to move out of district (subscript OD) after the adoption of VA.	VA may be kept private.	$0 > \gamma_{14OD}$	3	8.1.3 (8.1.3) 8.2.5 (8.2.6)
	Secondary Model Predictions				
3.	The introduction of VA should cause a larger difference in the selection of movers within-district as opposed to out-of-district, for those moves within and out-of Guilford than for those moving within and out-of Winston-Salem.	VA is a more salient signals for teachers moving from Winston-Salem than for teacher moving from Guilford.	$\gamma_{14WD_{GCS}} - \gamma_{14OD_{GCS}} > \gamma_{14WD_{WSF}} - \gamma_{14OD_{WSF}}$	3	8.2.7
4.	The selection effects should be particularly true for moves to higher-performing schools (subscript HP). The positive within-district selection of movers to higher performing schools leads to further within-district sorting (subscript S) of teachers to schools.	Teachers prefer higher-performing schools and principals at lower-performing schools are constrained in attracting talent.	$\gamma_{14WD_{HP}} > 0$ and $\gamma_{14S} > 0$	3, 6	8.2.8
5.	The introduction of VA should lead to a negative change in the selection of movers on the basis of easily observable characteristics for within-district moves, and the change in selection will be less negative (or even positive) for out-of-district moves.	VA may be kept private.	$\gamma_{24WD} < 0$ and $0 < \gamma_{24OD}$	4	8.2.9 8.2.10
6.	The positive change in selection should be particularly true for teachers with more tenure at a given school.		Coefficient on $VA \times Ten \times TreatDist > 0$	5	8.2.11

4 Data and Estimation

In this section, I describe both the data and methods used to generate VA measures of teacher effectiveness, and estimate the effects of the district policies on the teacher mobility. Subsection 4.1 describes the generation of VA. Subsection 4.2 describes the estimation sample. Subsection 4.3 describes the difference-in-differences estimation approach used to identify the effects of the new information on the mobility decisions of teachers and principals.

4.1 Value-Added Measures

While there are other valuable dimensions of teaching, many schools and districts care a great deal about teachers' abilities to raise their students' performance on standardized assessments. This study relies on administrative, longitudinal data, which links students to their teachers and was generously provided by the North Carolina Education Research Data Center (NCERDC) to estimate teachers' abilities to do just that. Though a robust source of data, the NCERDC does not contain the exact VA measures issued to each teacher within the treatment districts, and neither district agreed to release them. Consequently, this study generates the student gains on the North Carolina End of Grade exams attributable to each teacher.

There are two primary ways to go about this. The first is to attempt to model the exact measures that teachers and principals receive. This is primarily useful in explaining the teachers' and principals' observed behavior. The second is to model teacher effectiveness as accurately as possible. This is primarily useful in evaluating the consequences of the policy. In my preferred specification, I model teacher effectiveness rather than attempting to replicate the EVAAS measure.[21] This is because the policy context matters in this setting, while as in the prior employer learning literature, the VA measures need not totally encompass a teacher's effectiveness. Here, VA measures only need to be stronger correlates with teacher effectiveness than are other correlates with productivity, such as educational attainment and level of certification.[22][23]

I present my preferred Dynmaic OLS (DOLS) measure of VA in equation 11. According to Guarino et al. (2012), this DOLS estimator is more robust to nonrandom student assignment,

[21] An element of feasibility also enters this preference. The EVAAS system is proprietary, and the exact data and methods used are not disclosed. Furthermore, SAS uses two different proprietary models, and for large school districts it is unclear which is used.

[22] The extant literature supports this claim. As Rivkin et al. (2005) show, easily observed teacher characteristics are not highly correlated with teacher effectiveness. Recent work shows significant correlation between teachers' VA measures and many future student outcomes, including educational attainment, earnings, and probability of incarceration (Chetty et al., 2011, 2014).

[23] Whereas Farber and Gibbons (1996); Altonji and Pierret (2001); Lange (2007); Schönberg (2007), Pinkston (2009), and Bates (2015) use AFQT score as a strong correlate with productivity about which employers must learn, I use the VA described above in this capacity.

a frequent criticism of the often used Empirical Bayes (EB) estimator, which assumes random assignment of students to teachers.[24] In practice, I use both DOLS and EB in Section 6 and the results do not change much as the measures are highly correlated.[25]

$$A_{ijt} = \mathrm{T}_{\mathbf{t}} + \mathbf{A_{ijt-1}}\boldsymbol{\beta_0} + \mathbf{X_{it}}\boldsymbol{\beta_1} + \boldsymbol{VA_j} + e_{it} \tag{11}$$

Here, A_{ijt} represents student i's mathematics achievement in teacher j's class in year t. Including $\mathbf{A_{it-1}}$ allows for the correlation of previous math and reading test performances with current performance. Additionally, $\mathbf{X_{it}}$ is a vector including demographic attributes of individual students, such as grade, race, gender, special needs, and gifted status. It is $\mathbf{VA_j}$, a vector of teacher indicators, which is of primary interest for this study. Acknowledging that VA measures can be somewhat unstable in any single year, my preferred estimates use data from each year a teacher is teaching 4^{th} through 8^{th} grade during my sample period. This allows me to gain the most precise estimate of teachers' true underlying ability, μ.

4.2 Estimation Sample

This study restricts attention to the 5,986,132 third through eighth grade student, year observations from 1997 through 2011 to construct the VA measures for 134,219 teachers who teach 4^{th} through 8^{th} grade. I link these data to education, licensing, and work history data of 67,062 lead teachers without teaching assistants for whom the records are complete. These teachers are dispersed across the 2,966 schools in 117 school districts. I further restrict the sample to only those teachers teaching 4^{th} through 8^{th} grade at the time of observation, since they are the only elementary and middle-school teachers to receive VA. This restriction pares down my sample from 416,135 teacher-year observations to 236,018. At the teacher level, the data includes the teachers' race, gender, institution of higher education, degrees earned, experience, and tenure at a given school. Each of these are easily observable to all schools and many are likely used to filter job candidates. I use performance at the school in which the teacher currently works as an additional, easily observable, possible correlate with effectiveness. Table 2 provides summary statistics for my estimation sample.

The districts that adopt VA do not differ substantially from state averages in achievement or percent of student receiving proficiency on the state standardized exams. Given that both districts include urban centers, they do have a higher proportion of black students and teachers than does an average district in the state. While teachers come from colleges of

[24] Given teachers' preferences found in Jackson (2009) and Boyd et al. (2013), it seems unlikely that teacher effects would be uncorrelated with student-level covariates.

[25] Rose et al. (2012) finds 94-95% agreement between the EVAAS measure and DOLS and 95-97% agreement between EVAAS and EB.

Table 2: Sample Summary Statistics

	Guilford		Winston-Salem		Rest of North Carolina	
	Mean	SD	Mean	SD	Mean	SD
Scaled Score	250.38	71.71	249.23	68.86	252.36	70.49
Percent Proficient	0.75	0.14	0.74	0.15	0.76	0.13
Share of black Students	0.42	0.24	0.36	0.24	0.29	0.24
Share of black Teachers	0.25	0.43	0.21	0.41	0.15	0.36
Share of Hispanic Teachers	0.01	0.09	0.00	0.04	0.00	0.06
Share of Teachers with Advanced Degrees	0.30	0.46	0.36	0.48	0.29	0.45
College Selectivity (Barron's)	3.95	1.43	3.92	1.68	3.93	1.44
Experience	11.59	9.76	13.36	9.71	12.19	9.85
Tenure	3.23	3.05	3.59	3.26	3.68	3.35
Job Moves	0.09	0.28	0.08	0.28	0.08	0.27
Within-District Moves	0.06	0.24	0.06	0.24	0.05	0.22
Out-of-District Moves	0.03	0.16	0.02	0.14	0.03	0.16
Left NCPS	0.06	0.23	0.04	0.20	0.06	0.24
VA	0.02	1.01	0.01	0.99	0.00	1.00
N	11,239		8,295		216,484	

Note: VA is measured in standard deviations with the mean centered at 0.
Tenure is generated, and is censored for those already working at a given school in 1995.

comparable selectivity, across districts, in Winston-Salem, a larger share of the teaching-force holds an advanced degree. However, on the basis of VA, teaching quality in both districts is very close to the state average.

4.3 Estimation Strategy

I use a modification of differences-in-differences to compare changes in the relationship between teacher quality and mobility around the adoptions of VA to the changes in the same relationship over the same times in the rest of the state. I estimate the following specification:

$$y_{jdt} = \boldsymbol{T_t} + \mathbf{D_d} + \mathbf{TreatDist_{jd}} \times \mathbf{Post_t}\delta + VA_j\mathbf{DinD_{1dt}} + \mathbf{X_{jdt}}\mathbf{DinD_{2dt}} + \xi_{jdt}, \tag{12}$$

$$\text{where } \mathbf{DinD_{hdt}} = \gamma_{h1} + \mathbf{TreatDist_{jd}}\boldsymbol{\gamma_{h2}} + \mathbf{Post_t}\boldsymbol{\gamma_{h3}} + \mathbf{TreatDist_{jd}} \times \mathbf{Post_t}\boldsymbol{\gamma_{h4}}, h = 1, 2,$$

y_{jdt} is an indicator of a job change for teacher j in district d and in year t. $\mathbf{T_t}$ represents year effects, $\mathbf{D_d}$ represents district fixed effects, and $\mathbf{X_{jdt}}$ is a vector of teacher and school characteristics including teacher experience, tenure,[26] race, highest degree earned and selectivity of bachelor degree granting institution, as well as percent of students who are black and percent of students testing above proficiency at the school level. $\mathbf{DinD_{1dt}}$ captures the differences in the effects of VA on mobility based on whether VA measures were available for teacher j in district d, at time t. Interactions with treatment district indicators separate permanent differences in the impacts of VA measures and other characteristics from confounding the effect of treatment, while interactions with indicators for post years do the same for statewide changes in the effects at the times the policies take effect. Thus, the identifying variation comes from the differences between adopting districts and the rest of the state in the differences in the regression coefficients of VA measures on the probability of moving schools between pre- and post-policy years. Furthermore, easily observable, lower correlates with effectiveness may become less tied to the probability of moving after the introduction of VA. Thus, I relax the restriction that the coefficients on easily observables remain constant throughout the policy adoption by interacting other teacher covariates with the differences-in-differences framework, $\mathbf{DinD_{1dt}}$, as well.

Keeping in mind previously estimated teacher preferences and potential differences in information available, I examine the six types of job changes separately: within district moves, within district moves to higher-performing schools, within district moves to lower-performing schools, out-of-district moves, out-of-district moves to higher-performing schools, and out-of-district moves to lower-performing schools. Given that teachers initiate most moves, moves to worse schools are likely driven by largely by idiosyncratic teacher preferences. Due to the indirect mechanism by which hiring principals in Winston-Salem obtain teachers' VA and the potential additional salience of VA signals to principals outside the district during Winston-Salem's later adoption, I separate treatment by district.

Given how the districts distributed VA, it seems clear that the new information would be public between two principals in Guilford. Perhaps to a lesser extant the same holds for Winston-Salem. Accordingly, the model predicts $\boldsymbol{\gamma_{14WD}} > 0$ (where $\boldsymbol{\gamma_{14WD}}$ is the effect of the interaction of VA with receiving treatment on the probability of moving within-district).

When comparing the expectations of a retaining principal within one of the treatment districts to a hiring principal in another district there is some ambiguity as to whether VA provide a more precise expectation for both principals or only the current one. Thus, the symmetric learning model for out-of-district moves predicts $\boldsymbol{\gamma_{14OD}} = \boldsymbol{\gamma_{14WD}}$ (where $\boldsymbol{\gamma_{14OD}}$ is the effect of the interaction of VA with receiving treatment on the probability of moving out-of-district). If current principals can keep information from employers in other districts,

[26] Because tenure is generated and censored for job matches beginning prior to 1995, an indicator of whether the current match existed in 1995 is included in all regressions.

the signal improves the precision of the current principal's signal about the true quality of the teacher, while the expectation of the out-of-district principal is unaffected. In which case, the asymmetric learning model would apply predicting $\boldsymbol{\gamma}_{\mathbf{14WD}} > \boldsymbol{\gamma}_{\mathbf{14OD}}$ and possibly $\boldsymbol{\gamma}_{\mathbf{14OD}} < 0$ for out-of district moves.

This type of movement may have important implications for the distribution of teacher quality across schools. If better teachers are more able to signal their true quality, and do so in general to move to better schools, the divide in teacher quality between the worst and best schools may widen. Accordingly, I estimate equation 12 substituting percent of students proficient in the school taught at the subsequent year, for the binary variable of whether teachers move. Again, if VA is informative, and teachers do in general prefer to teach at better schools, $\boldsymbol{\gamma}_{\mathbf{14SQ}} > 0$ in this regression as well. ($\boldsymbol{\gamma}_{\mathbf{14SQ}}$ is the effect of the interaction of VA with receiving treatment on the proficiency levels of the school where the teacher works the subsequent year.) Similar to the probability of moving to a better school, we may expect these effects to be somewhat muted for teachers moving later in their careers, in which case hiring principals may already have more complete information.

Furthermore, because there would be more information available on more experienced teachers, if there had previously been some degree of public learning, the model predicts the effects to diminish with teacher experience. Likewise, if there had previously been private learning, the learning model predicts the shock to public information to have larger ramifications for teachers with more tenure at a given school all else equal. In later specifications, I interact VA with experience and the difference-in-differences, **DinD**, interactions.

There are two distinct issues that complicate the estimation of standard errors in this study. First, the policy variation occurs at the district level, meaning the errors may be correlated for teachers moving from or within the same district. Clustering at the district level make the standard errors robust to this cross-sectional dependence. Secondly, the VA measures are estimated and thus inherently suffer from estimation error. Were this a singular issue, it would be appropriate to bootstrap the student data to account for this estimation error.[27]

Accordingly, I adopt a sampling approach that accounts for both the estimation error of VA measures and the clustered nature of the data. First, I sample districts randomly with replacement just as with the standard cluster-bootstrap. I then conduct stratified sampling at the teacher level, such that for every teacher who was originally sampled, I randomly sample student/year observations with replacement. In so doing, this provides generally

[27] It may seem natural to cluster-bootstrap at the district level. However, this samples all students for a every teacher in a sampled district, and does not address the estimation error. The standard errors from the cluster bootstrap are smaller than the non-bootstrap clustered standard errors by about a factor of ten.

more conservative standard errors across parameters. Table A1 in the Appendix 8.6 presents all standard errors for Table 3 for comparison. Throughout the remainder of this paper, I present the more conservative district-clustered-teacher-stratified-bootstrap standard errors (CSB SEs).

5 Results

5.1 Mobility and Sorting

How does mobility change with the adoption of VA and what does that tell us about the way employers learn about their employees? Table 3 presents the estimated impact of revealing EVAAS reports of teacher effectiveness on the relationship between teachers' VA and the probability a teacher moves to another school. Given the evidence that teachers prefer to teach in schools with higher-performing students, Table 3 decomposes effects by whether the receiving school has higher or lower-performing students than the current school.[28] The test between symmetric and asymmetric employer learning focuses on how the effects of VA on the probability of moving within-district differ from the effects of VA on the probability of moving out-of-district after the treatment districts adopt the measures of teacher quality. Panel A restricts attention to within-district moves, and Panel B presents evidence from out-of-district moves.

The first row presents the the relationship between VA measures and the probability of each type of move in the rest of the state, regardless of any districts adopting the policy. In general, there is little relationship between VA and the probability of moving within or out of the district. However, when discerning between moves to more and less proficient schools a familiar pattern emerges. From columns 2 and 3 of Panel A, a teacher with a standard deviation higher VA is about 0.3 percentage points more likely to move to a higher-performing school and 0.2 percentage points less likely to move to a lower-performing school within the district. Panel B exhibits the same pattern regarding moves to schools outside of the current district. A one standard deviation increase in VA before the policy takes effect raises the probability of moving to a higher-performing school by about a tenth of a percentage point and lowers the probability of moving to lower-performing school by about the same magnitude.

Within both Guilford and Winston-Salem, the release of VA intensifies this pattern. From the coefficient on the interactions between policy treatment and VA in both districts, a stan-

[28]Primary effects of VA on different types of moves further supports this distinction. I define a move to a higher performing school as a move in which the school taught at the following year has a higher percentage of students who achieve proficiency than the current school. Proficiency rates are demeaned by year statewide averages, while a move to a lower-performing school is defined in the reverse way.

dard deviation increase in a teacher's VA leads to about a half of a percentage point increase in the probability of moving within district after the district released the value-added information. While the magnitudes of the effects are very close between districts, they are only statistically significant beyond the 95% confidence level for Guilford. Column 2 illustrates that these results are driven by moves to higher-performing schools, as the model predicts. From column 2, the estimated coefficients imply that the adoption of VA raises the probability that a teacher with one standard deviation higher VA will move to a higher-performing school by over 14% (p-value .011) in Guilford and nearly 18% (p-value .009) in Winston-Salem. Column 3 reveals little change in the effects of VA on the probability of moving to a lower-performing school within district. The similarity of the point estimates on the impact of VA post-treatment between Guilford and Winston-Salem provides no evidence that relying upon teachers to voluntarily disclose their VA scores to hiring principals mitigates the effects.

From Section 3, the effect of the policy should be no different whether teachers move to schools within or outside of the district, under the symmetric learning hypothesis. However, asymmetric employer learning predicts the policy to give principals in Guilford and Winston-Salem an informational advantage over principals in other districts. This translates into smaller selection effects for teachers moving to other districts than for within-district moves, and these effects may even be negative. The second column of Panel B presents changes in the effect of teacher quality on the probability of moving to a better, out-of-district school after the adoption of VA. Again, these changes in selection of mobile workers are consistent with the employer learning model.

The change in selection of teachers leaving Guilford provides the strongest evidence of growing informational asymmetries between employers. In Guilford, a teacher who has a standard deviation lower VA, is a full percentage point more likely to move out-of-district. This same, low-VA teacher is about a half a percentage point more likely to move to a better school out-of-district (p-value 0.001). There is also a statistically significant effect on the probability of moving to lower-performing schools out of Guilford. While the model does not predict this type of movement, it is not surprising. Low VA scores may lead current principals to devalue some of their teachers, who may respond by moving to lower-performing schools that are not privy to their value-added scores.

In Winston-Salem, the difference between within- and out-of-district moves is less pronounced, though still consistent with private employer learning. While in Winston-Salem, a teacher with one standard deviation higher VA is more likely to move to a higher-performing school out-of-district after the policy takes effect, the point estimate is only 38% of that from moving within-district and is no longer statistically significant. Were outside principals informed of the signal, we would expect the same positive effects found in the second column of Panel A to be present in in the second column of Panel B.

The fact that effects are more negative in Guilford than Winston-Salem, may be explained by differences in the salience of the signals between teachers moving from Guilford as opposed to those moving from Winston-Salem. Guilford's adoption of the EVAAS measures of teacher effectiveness occurred in 2000. It is unlikely that at that time principals in other districts had much understanding of the measures, or their reliability. In contrast, the rest of the state adopted school-level EVAAS reports simultaneously with Winston-Salem's adoption of teacher level VA. Given this difference in contexts, high VA teachers from Winston-Salem may have been better able to use their VA to obtain positions outside of Winston-Salem, than would a comparable teacher moving earlier from Guilford. In Winston-Salem, the increase in high-VA teachers' ability to signal their effectiveness may mitigate any effects from relatively low VA teachers exploiting the informational asymmetry. The mitigated effects of VA for those moving out of Winston-Salem in addition to the negative selection of teachers moving away from Guilford evidences informational asymmetries between potential employers within as opposed to outside of the district.

Table 3: Probability of Moving Schools Within and Out of District

	Panel A: Within-District Moves			Panel B: Out-Of-District Moves		
VARIABLES	Total	To a higher performing school	To a lower performing school	Total	To a higher performing school	To a lower performing school
VAM	0.0016	0.0032	-0.0016	0.0002	0.0014	-0.0012
	(0.00129)	(0.00091)	(0.00074)	(0.00096)	(0.00072)	(0.00058)
VAM x Treatment GCS	0.0058	0.0051	0.0007	-0.0103	-0.0054	-0.0049
	(0.00265)	(0.00199)	(0.00151)	(0.00261)	(0.00195)	(0.00156)
VAM x Treatment WSF	0.0052	0.0060	-0.0008	0.0009	0.0023	-0.0014
	(0.00286)	(0.00229)	(0.00194)	(0.00241)	(0.00208)	(0.00129)
Treatment GCS	-0.0040	-0.0050	0.0010	-0.0162	-0.0232	0.0070
	(0.00851)	(0.00571)	(0.00679)	(0.00374)	(0.00233)	(0.00268)
Treatment WSF	0.0555	0.0475	0.0080	-0.0020	0.0147	-0.0167
	(0.00499)	(0.00372)	(0.00299)	(0.00274)	(0.00224)	(0.00178)
Observations	236,018	236,018	236,018	236,018	236,018	236,018

CSB standard errors from 500 repetitions appear in brackets. All regressions use a linear functional form, and include teacher level covariates and interactions with treatment indicators, as well as year and district fixed effects. .

5.2 Observables

In addition to predicting mobility dynamics with respect to teacher VA, the model presented in Section 3 also offers predictions regarding easily observable covariates with teacher effectiveness. In instances where the VA measures shock the available public information, the model predicts principals would place less emphasis on easily observable covariates with teacher effectiveness, such as degree attainment and college selectivity. In cases where VA exacerbate informational asymmetries between current and hiring principals, the same covariates expectedly receive additional emphasis on the probability of a move.

To provide ease of interpretation, I generate an index of easily observable teacher quality by taking the fitted values from the OLS regression of teacher VA on teacher covariates. I include as components of this index, an indicator for having an advanced degree, a vector of indicators for Barron's College Competitiveness index, years of experience, years of tenure, an indicator for whether tenure is censored, race, gender, and a vector of year indicators.[29]

In general, those with high observable characteristics are more likely to move within district. That result is driven by moves to higher-performing schools, while those with lower observable characteristics are more likely to move to lower-performing schools. For moves out-of district, the positive relationship between the index and the probability of moving to a better school offsets the negative relationship between the index and the probability of moving to a lower-performing school. These relationships are expected given the sorting of teachers based on observable characteristics as shown in Jackson (2009) among others.

The first two columns of Table 4 do not bear out the predictions for within district moves. While noisy, the point estimates of the effects of the teacher index on the probability of moving schools within-district after the adoption of VA are positive, though only statistically significantly so for moves to better schools within Guilford. While not expected, this result may be explained by the additional churn that accompanies the adoption of VA particularly for moves to better schools within Guilford. More positions may become available as a result of high-VA teachers moving to better schools, and low-VA teachers moving out of district. As a result, those with good observables find it easier to move in addition to those with high VA. Heterogeneous openness among principals to VA may also contribute.[30] In which case, as high-VA teachers move to principals that value VA, those with other favorable attributes move to the principals who value those characteristics.

The change in the relationship between the index and the probability of moving out-of-

[29] The VA measures used in this analysis are the residuals from the projection of my standard VA measures on the components of the index.

[30] Informal conversations with principals in Winston-Salem and Guilford indicate this may be the case, as two current lower elementary principals that I spoke with indicated that teachers' VA played a limited role in their hiring decisions.

district with the adoptions of VA is more supportive of the model. Whereas movers out of Guilford are adversely selected on the basis of the hard-to-observe VA, they are positively selected on the basis of this index of easily observable measures of teacher quality. This is true across moves to higher or lower performing schools, and provides further evidence that the moving teachers with a high index, but low VA were able to keep their VA private, while utilizing their otherwise strong resumés to move to uninformed principals. Given that it is plausible that more teachers moving from Winston-Salem could inform out-of-district principals of their VA, results in either direction may make sense. Accordingly, the results for moves out of Winston-Salem are not very informative. While the results for moves out of Guilford are reassuring, cumulatively, the evidence from changes in the relationship between the index of easily observable teacher characteristics, and the probability of moving schools is too mixed to draw definitive conclusions.[31]

[31] In unreported regressions, with the exception of out-of-Guilford moves the results shown in Table 4 are sensitive to the variable composition of the teacher quality index.

Table 4: Effects of teacher quality index on the probability of moving

	Within-District Moves			Out-of-District Moves		
Variables	Total	To higher performing schools	To lower performing schools	Total	To higher performing schools	To lower performing schools
VA	0.0018	0.0039	-0.0021	-0.0002	0.0014	-0.0016
	(0.00111)	(0.00078)	(0.00073)	(0.00091)	(0.00068)	(0.00053)
Teacher Quality Index (TQ Index)	0.005	0.0071	-0.0021	-0.0005	0.0031	-0.0035
	(0.00233)	(0.00173)	(0.00105)	(0.00186)	(0.00115)	(0.00096)
VA x Treatment GCS	0.0083	0.0069	0.0014	-0.0109	-0.0053	-0.0056
	(0.00237)	(0.00177)	(0.0014)	(0.00249)	(0.00189)	(0.00145)
VA x Treatment WSF	0.0063	0.0062	0.0000	0.0001	0.0018	-0.0017
	(0.00248)	(0.00199)	(0.00193)	(0.00212)	(0.00189)	(0.00115)
TQ Index x Treatment GCS	0.0040	0.0043	-0.0003	0.0076	0.0061	0.0015
	(0.00246)	(0.00153)	(0.00145)	(0.00116)	(0.00088)	(0.00088)
TQ Index x Treatment WSF	0.0029	0.0027	0.0002	-0.0011	-0.0026	0.0015
	(0.00254)	(0.00192)	(0.00131)	(0.00097)	(0.00078)	(0.00063)
Treatment GCS	0.0142	0.0253	-0.0111	-0.0120	-0.0132	0.0011
	(0.00595)	(0.00449)	(0.00405)	(0.00258)	(0.00167)	(0.00189)
Treatment WSF	-0.0015	0.0091	-0.0106	0.0118	0.0177	-0.0059
	(0.00383)	(0.00242)	(0.00253)	(0.00251)	(0.00136)	(0.00139)
Observations	236,018	236,018	236,018	236,018	236,018	236,018

CSB standard errors from 500 repetitions appear in brackets. All regressions use a linear functional form, and include teacher level covariates and interactions with treatment indicators. The VA measures used in this analysis are the residuals from the projection of my standard VA measures on the components of the index.

5.3 Differential Effects With Respect to Experience and Tenure

Examination of differential effects with respect to years of experience and tenure in a given school may provide insight into the type of learning that previously prevailed. Were private learning already prevalent in the market, the model predicts larger positive selection of movers who have more years of tenure.[32] This is because the information gaps between retaining and hiring principals grows with time a teacher teaches in the same school. The introduction of VA would be a larger shock to the information gap on these teachers.

The results in columns 1 and 2 are consistent with this prediction of prior private learning. For each additional year of tenure a standard-deviation-higher-VA teacher has, he is about 0.6 a percentage point more likely to move within Guilford and 0.3 a percentage point more likely to move within Winston-Salem. From column 2, the economic and statistical

[32] See prediction 6 with proof provided in Appendix 8.2.11.

significance falls when focusing on moves to better schools, providing reason to pause before concluding that the learning was previously asymmetric.

While ambiguity in the model prevents me from making a formal prediction regarding experience, if there was previous public learning, intuitively the release of VA would serve as less of a shock for teachers about whom there already existed a great deal of information. Thus, we may expect smaller results for less experienced teachers. While Table 5 exhibits this relationship for teachers moving out of the district (though not statistically significantly so), the same is not true for teachers moving within district. Cumulatively, these results largely suggest prior private over public learning.

Table 5: Differential Effects With Respect to Experience and Tenure

	Within District		Out of District	
VARIABLES	Total	Higher Performing	Total	Higher Performing
VA	-0.0001	0.0028	-0.0001	0.0023
	(0.0023)	(0.00161)	(0.00244)	(0.00173)
Experience x VA	-0.0000	0.0000	-0.0000	-0.0000
	(0.00011)	(0.00008)	(0.00011)	(0.00008)
Tenure x VA	0.0020	0.0006	0.0006	0.0005
	(0.0008)	(0.00059)	(0.00073)	(0.00058)
VA x Treatment GCS	0.0033	0.0050	-0.0181	-0.0095
	(0.00568)	(0.00465)	(0.00693)	(0.00514)
Experience x VA x Treatment GCS	0.0016	0.0010	0.0002	0.0003
	(0.00026)	(0.0002)	(0.00032)	(0.00026)
Tenure x VA x Treatment GCS	0.0056	0.0004	0.0008	0.0014
	(0.00179)	(0.00146)	(0.00217)	(0.00178)
VA x Treatment WSF	-0.0003	-0.0010	-0.0073	-0.0051
	(0.00551)	(0.00431)	(0.00503)	(0.00452)
Experience x VA x Treatment WSF	0.0003	0.0005	0.0002	0.0002
	(0.00043)	(0.00036)	(0.00029)	(0.00025)
Tenure x VA x Treatment WSF	0.0028	0.0009	0.0004	0.0004
	(0.00078)	(0.00055)	(0.00053)	(0.00046)
Observations	236,018	236,018	236,018	236,018

CSB standard errors from 500 repetitions appear in brackets. All regressions use a linear functional form, and include teacher level covariates and interactions with treatment indicators.

5.4 Educational Equity

The increases in the mobility of effective teachers to higher performing schools is concerning for educational equality. Depending on the district, the mobility results are similar or even stronger when looking at the change in the relationship between teacher effectiveness and mobility with respect to the student body's racial composition. Table 6 presents these results in panel A, as well as, the effects of VA adoption on the sorting of teachers to schools with respect to students' race (panel B) and students' performance (panel C).[33]

The coefficient on VA in column 1 of panel A demonstrates that in general more effective teachers are more likely to move to schools with smaller shares of black students than their current school. Moving down the column shows that the release of VA magnifies that sorting in both adopting districts. VA adoption in Winston-Salem leads to a 1.3 percentage point increase in the probability that a teacher with a standard deviation higher VA moves within-district to a school with a lower share of black students. This is more than double the effect size found for moving to higher-performing schools. It is worth mentioning that this is accompanied by a 0.8 percentage point drop in the probability that a similarly effective teacher moves to a school with a higher proportion of black students. For moves within Guilford, the effects are smaller, but still statistically significantly positive. For out-of-district moves, there continues to be no statistically significant effect for Winston-Salem, and in Guilford there continues to be adverse selection to schools with higher and lower shares of black students.

Turning to panels B and C, the coefficient on VA describes the general relationship between teachers' VA and the share of black or proficient students at the school they teach at the subsequent year. Since all regressions control for the current share of black students and proficient students at the current school, it can be thought of as the relationship between teacher effectiveness and year-by-year change in school proficiency level or racial composition in the absence of observable VA. The first columns of panels B and C examine sorting for all teachers in the sample who remain teaching in North Carolina the following year. The second columns of panels B and C restrict the sample to those who remain within their current district. These second columns may be more informative for predicting the effects of the policy in the rest of the state after the adoption of EVAAS VA measures becomes statewide. The effects may be more pronounced for the state as a whole, because the costs of moving out of state are in general higher than those of moving out of a school district.

From the first row in panel B, a standard deviation increase in a teacher's VA is associated with about a tenth of a percentage point decrease in the the percent of black students. Across both columns of panel C, the same standard deviation higher VA is associated with a quarter of a percentage point increase in the percent of students who are proficient in the school in which he teaches the subsequent year.[34]

[33] The data on free and reduced price lunch status (FRL) do not permit me to examine the effect of the policy on mobility with respect to FRL for Guilford. However, unreported regressions show that in Winston-Salem the mobility patterns with respect to FRL are very similar to those regarding students' race.

[34] The result that students in better schools also get better teachers is consistent with findings in Boyd

Next, I turn to the change in sorting with VA adoption in rows 3 and 4. Including teachers who move within and out of district, it seems from the first columns of panels B and C that releasing VA has opposite effects in the two districts on the distribution of teacher quality across schools. However, this can be explained by the adverse selection of teachers moving from Guilford after the policy takes effect.

Turning to the sample of teachers who remain in the same district, the second column of both panels provides evidence of further sorting in Winston-Salem. From the second column of panel B, the release of VA leads a teacher with one standard deviation higher VA to be at a school with 0.3 percentage points lower share of black students. From the second column of panel C, the same teacher will be at a school that has 0.2 percentage points higher proficiency rates after the district releases VA. Taken literally, this translates to 70 and 300 percent increases in the sorting of teacher quality towards high achieving students and away from black students respectively. However, each estimate is noisy, and is only marginally statistically significant (respective p-values of 0.096 and 0.099), and should be treated accordingly. In Guilford, the positive coefficient estimate suggests that the policy leads better teachers to move to schools with higher proportion of black students, but has essentially no effect on sorting with regard to student performance. However, neither effect is statistically significant.[35] The large effects in Winston-Salem taken together with the mobility patterns from Table 3 and panel A of Table 6 evidence rising inequality in the distribution of effective teachers as an unintended consequence of VA adoption.

et al. (2005) and Boyd et al. (2008).

[35]Contextually, it is important to note that both district offered teachers financial incentives to teach in lower-performing schools. Analysis in Section 6.3 provides insight about the effects of VA adoption on the re-sorting of teachers between schools in which no compensating differentials were in place. Further, I find no evidence of more low-VA teachers leaving teaching in response to district adopting VA. In unreported regressions, the probability of leaving North Carolina Public Schools from WSF were statistically unrelated to the teachers' VA and from Guilford, better teachers were more likely to leave.

Table 6: Educational Equity

Panel:	A: Moves based on share of students who are black				B: Growth in percent black		C: Growth in percent proficient	
	Within-District		Out-of-District					
VARIABLES	To lower percent black	To higher percent black	To lower percent black	To higher percent black	Total	Stay Within district	Total	Stay Within district
VA	0.0021	-0.0005	0.0009	-0.0007	-0.0018	-0.0011	0.0028	0.0024
	(.00088)	(.00086)	(.00078)	(.00059)	(.00046)	(.00038)	(0.00033)	(0.00033)
VA x Treatment GCS	0.0037	0.0021	-0.0067	-0.0035	0.005	0.0026	-0.0005	-0.0000
	(.0019)	(.00167)	(.00217)	(.00143)	(.00198)	(.002)	(0.00074)	(0.0007)
VA x Treatment WSF	0.0133	-0.0082	-0.0007	0.0017	-0.0034	-0.0033	0.0007	0.0017
	(.00228)	(.00188)	(.00192)	(.00129)	(.00235)	(.002)	(0.00114)	(0.00102)
Treatment GCS	0.0040	-0.0088	-0.0043	-0.0119	0.0354	0.0290	-0.0195	-0.0157
	(.00513)	(.00738)	(.00251)	(.00278)	(.00319)	(.00302)	(0.00211)	(0.00216)
Treatment WSF	0.0277	0.0280	-0.0041	0.0020	-0.0198	-0.0245	0.0290	0.0231
	(.00355)	(.00292)	(.00233)	(.00164)	(.00318)	(.00328)	(0.00172)	(0.00168)
Observations	236,018	236,018	236,018	236,018	209,424	202,943	209,424	202,943

CSB standard errors from 500 repetitions appear in brackets. All regressions use a linear functional form, and include teacher level covariates, and their interactions with treatment indicators. .

6 Robustness

In the following section, I examine the robustness of the effects of VA adoption. Section 6.2 considers changes in effects when using only prior years of student data when constructing VA measures. Section 6.3 considers whether other district policies that paid teachers to work in hard-to-staff schools impact the estimated effects. Appendix 8.3 considers teacher mobility in accordance with the state ABC growth bonus-pay system. Within-district, year-by-year analysis of the changing effects of VA on mobility and sorting are presented in Appendix 6.1. In Appendix 8.4 and Appendix 8.5, I consider alternate functional forms for the mobility analysis. In Appendix 8.4, I take seriously the normality assumptions, and perform normal Maximum Likelihood Estimation. In Appendix 8.5, I use competing risks regression to examine the possibility of correlated errors between types of moves.[36]

6.1 Differential trends

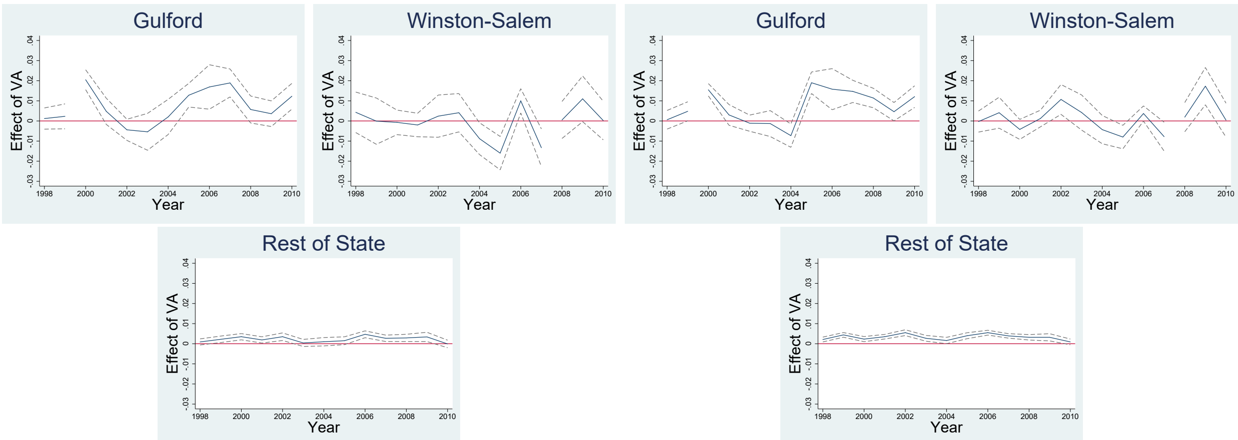
To investigate the potential differences in time trends between Guilford and Winston-Salem loading onto the policy change, within each treatment district separately and within the rest of the state I estimate the impact of VA on the probability of moving at each year. Figure 1 shows the evolution of VA coefficient estimates by year on within-district mobility

[36]Because job mobility is often localized, I also restricted analysis to districts which share a border with Guilford and Winston-Salem. The results from this restriction were noisy and uninformative, and are unreported here.

Figure 1: The effects of VA on the probability of moving schools within-district by year.

Panel A: Within-district total moves

Panel B: Within-district moves to higher-performing schools

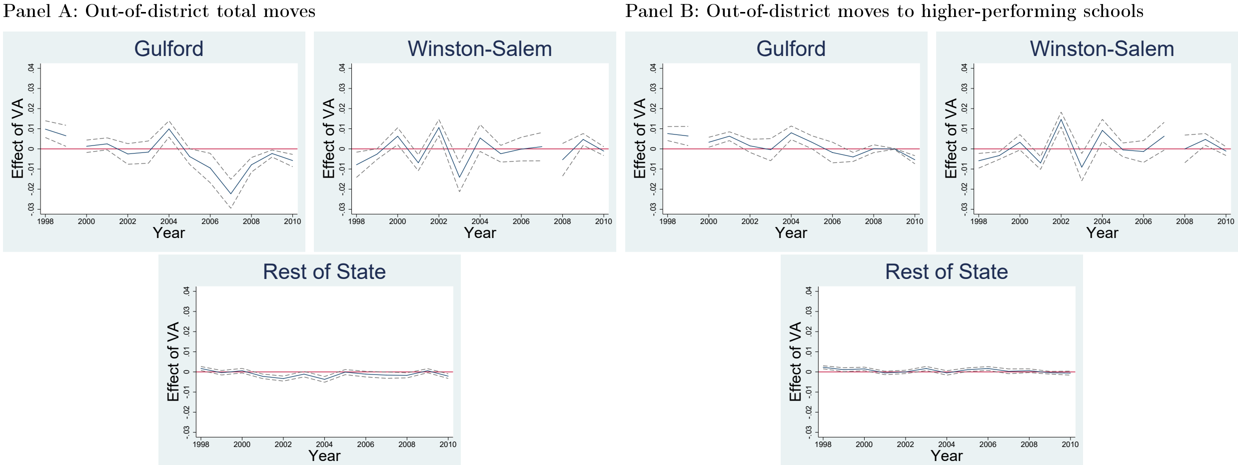


Note: Solid blue line reflects the point estimates within district or rest-of state on the interactions between year indicators and VA. The dotted lines indicate the 90% confidence interval from within-district bootstrap with 500 replications.

Figure 2: The effect of VA on the probability of moving schools out-of-district by year.

Panel A: Out-of-district total moves

Panel B: Out-of-district moves to higher-performing schools



Note: Solid blue line reflects the point estimates within district or rest-of state on the interactions between year indicators and VA. The dotted lines indicate the 90% confidence interval from within-district bootstrap with 500 replications.

with 90% confidence intervals and breaks at policy adoption. In panel A the dependent variable is an indicator for moving within-district, and in panel B the outcome is an indicator for moving to higher-performing schools within-district. Figures 2 and 3, illustrate the same coefficient evolution on out-of-district mobility and growth in school performance. Tables A2, A3, and A4, provide the accompanying tables to the figures in text.

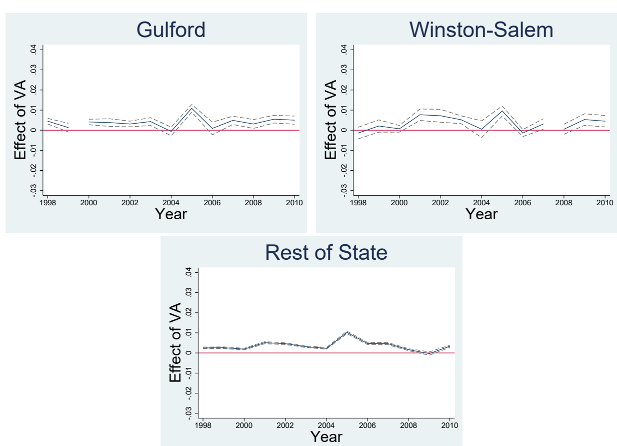
In both districts there is a spike in the correlation of VA with the probability of moving within-district soon after the policy takes effect. While the estimates are too noisy to say anything conclusive, the pre-policy trends do not seem diverge in a way that would bias up the results.

6.2 Sensitivity to VA Construction

As discussed in Sections 4.1 and 3, when constructing VA estimates for each teacher it may make sense to set the objective of approximating the signals teachers and principals receive as opposed to the true effectiveness of the teacher. Panel A of Table 7 reflects similar regressions as does Table 3 except that I use Empirical Bayes estimates of teachers' VA rather than DOLS. Using Empirical Bayes, the results remain remarkably similar both in magnitude and precision.

The possibility that teachers may have differences in VA after moving to other schools, may present issues for using VA measures constructed from student data from a teacher's entire career. This could result from moves leading to higher match quality between teachers

Figure 3: The effect of VA on teacher sorting within-district by year.



Note: Solid blue line reflects the point estimates within district or rest-of state on the interactions between year indicators and VA. The dotted lines indicate the 90% confidence interval from within-district bootstrap with 500 replications.

and schools, as Jackson (2013) finds. It may also result from transitory adjustment costs, giving a theoretically ambiguous direction of potential bias

Consequently, in Panel B of Table 7, I allow teachers VA scores to vary each year, using only data from the current and previous years to construct a teacher's VA in any given year. The main effects hold, though they are in general somewhat exaggerated in Winston-Salem and smaller in Guilford. Still, the adoption of VA raises the probability that good teachers move to better schools. Whereas in Winston-Salem, the effect grows to a full percentage point, in Guilford, a teacher with an one standard deviation higher VA becomes 0.3 percentage points more likely to move to better school post-policy. From the middle column of Panel B, the negative selection of teachers moving out of Guilford falls to just 30% of the estimate given in Table 3. Panel C in Table 7 corresponds with Table 6. While the effect on teacher sorting doubles in Winston-Salem, the results become more negative and statistically insignificant in Guilford.

Table 7: Probability of moving schools using Empirical Bayes VA

VARIABLES	Within District: Total	Within District: To a higher performing school	Within District: To a lower performing school	Out of District: Total	Out of District: To a higher performing school	Out of District: To a lower performing school
Panel A: Full sample of student test scores						
VAM	0.0006	0.0028	-0.0022	-0.0006	0.0014	-0.0020
	(0.00141)	(0.00097)	(0.00079)	(0.00094)	(0.00064)	(0.00059)
VAM x Treatment GCS	0.0048	0.0059	-0.0011	-0.0130	-0.0078	-0.0051
	(0.00256)	(0.002)	(0.00135)	(0.00229)	(0.00179)	(0.00148)
VAM x Treatment WSF	0.0066	0.0085	-0.0020	0.0009	0.0023	-0.0013
	(0.00288)	(0.00225)	(0.00178)	(0.00235)	(0.00212)	(0.00121)
Treatment GCS	-0.0048	-0.0055	0.0007	-0.0174	-0.0245	0.0072
	(0.00743)	(0.00478)	(0.00652)	(0.00326)	(0.00233)	(0.00177)
Treatment WSF	0.0553	0.0471	0.0082	-0.0022	0.0144	-0.0167
	(0.00453)	(0.0032)	(0.00282)	(0.00233)	(0.00209)	(0.0014)
Panel B: Restricted to prior student test scores						
VAM	-0.0015	0.0000	-0.0015	-0.0021	-0.0011	-0.0010
	(0.00169)	(0.00141)	(0.00093)	(0.00098)	(0.00073)	(0.00062)
VAM x Treatment GCS	0.0035	0.0037	-0.0001	-0.0063	-0.0041	-0.0023
	(0.00331)	(0.00252)	(0.00221)	(0.00232)	(0.00195)	(0.00129)
VAM x Treatment WSF	0.0090	0.0129	-0.0039	0.0020	0.0019	0.0001
	(0.003)	(0.00236)	(0.00186)	(0.0023)	(0.00202)	(0.00113)
Treatment GCS	-0.0032	-0.004	0.0008	-0.0162	-0.0239	0.0077
	(0.01311)	(0.00855)	(0.01071)	(0.00515)	(0.00281)	(0.00431)
Treatment WSF	0.0555	0.0477	0.0078	-0.0021	0.0147	-0.0167
	(0.00496)	(0.00346)	(0.00294)	(0.00234)	(0.00208)	(0.00142)
Observations	236,018	236,018	236,018	236,018	236,018	236,018

CSB standard errors from 500 repetitions appear in brackets. All regressions use a linear functional form, and include teacher level covariates and interactions with treatment indicators, as well as year and district fixed effects.

While it is possible subsequent match quality increases for teachers from Guilford and decreases for teachers in Winston-Salem, I believe measurement error may provide a more plausible explanation. In Guilford, the effect of VA prior to the their release is identified off of just two years of data. As a result, the estimates of teachers' VA are noisier for this period as well as in the immediate aftermath of the policy. Measurement error in the primary variable of interest may attenuate the estimates in Guilford where there is little data prior to the adoption of the policy, while the effects in Winston-Salem become relatively stronger.

One way of shedding light on this issue is to use a fixed number of years prior to the current period when constructing VA measures. Unfortunately, the adoption of VA by Guilford comes just three years into the student data sample. Since the construction of VA measures requires at least one prior year of student data, this gives just two years at which I could fix my VA estimate. Not only would this force a noisier estimate of each teacher's VA for the entire sample, it also provides merely one year of data prior to the adoption of the policy in Guilford. To demonstrate the changes of the estimates with varying the number of years of data used in constructing VA measures, I drop Guilford from the analysis and vary the number of prior years of data I use to construct the VA measures from 2 to 8. Table 8 demonstrates that though the relationship between years used and the effect of the interaction of the policy in Winston-Salem and VA is not monotonic as the sample used varies, the estimates using more years of data are clearly the largest. This further suggests correlated measurement error presents a problem for restricting VA construction to prior years of data.

Table 8: Sensitivity of VA estimates to using various number of years of student data in VA construction

VARIABLES	2yr VA	3yr VA	4yr VA	5yr VA	6yr VA	7yr VA	8yr VA
VA	0.0010	0.0015	0.0017	0.0019	0.0021	0.0023	0.0035
	(0.00032)	(0.00047)	(0.00047)	(0.00058)	(0.00063)	(0.00066)	(0.00072)
VA x Treatment WSF	0.0119	0.0114	0.0108	0.0116	0.0142	0.0163	0.0181
	(0.00614)	(0.00613)	(0.00609)	(0.00621)	(0.0063)	(0.00655)	(0.00685)
Treatment WSF	0.0550	0.0534	0.0542	0.0473	0.0416	0.0439	0.0401
	(0.01873)	(0.01855)	(0.01856)	(0.0181872)	(0.01911)	(0.02002)	(0.0227)
Observations	207,673	189,531	170,598	151,067	131,567	111,786	94,884

CSB standard errors from 500 repetitions appear in brackets. All regressions use a linear functional form, and include teacher level covariates and interactions with treatment indicators. Observations from GCS are omitted from the above analysis.

6.3 Strategic Staffing

A possible complication arises due to alternate teacher compensation plans. District strategic staffing policies, which aim to attract more capable teachers to teach in and stay at hard-to-staff schools may be problematic because they occurred in treatment districts during the sample period and could potentially alter teacher preferences over schools.[37] Charlotte-Mecklenburg Schools (CMS) and Winston-Salem were by far the earliest adopters of these initiatives with CMS beginning its Equity Plus program in 1999 and Winston-Salem following suit in 2000. By 2012 each major district in North Carolina adopted some program to attract teachers to hard-to-staff schools. In CMS, teachers received a signing bonus to enter a targeted school and teachers with a masters degree could receive up to $2,500 per year to remain in the school. A smaller incentive was offered to teachers enrolled in masters programs, though the district also offered tuition reimbursement. Winston-Salem awarded 20% of the district salary supplement ($500-$1,500) to each teacher in targeted schools. Furthermore the entire state offered $1,800 bonuses to math, science, and special education teachers who taught in high poverty or low achieving schools during the three year period 2002-2004. In 2007, Guilford adopted its own strategic staffing program, in which bonuses ranged from $5,000-$25,500 depending on subject taught, grade level, and VA. Cumberland County Schools gave stipends to 30 "master teachers" across their 10 most difficult school. In 2008, CMS began tailoring their plan more towards targeting better teachers and Winston-Salem, followed suit in 2012. These programs may reverse which schools are most desirable to teachers. With large enough incentives, high-VA teachers may opt to work at low performing school, which is in fact the intent of the policy.

Panels A and B of table 9 reports similar information as is provided in Table 3, with the difference that the binary dependent variable in Table 9 is equal to one if a move occurs and the receiving school is not classified as strategic staffing. As might be expected, the results are quite similar to those in Table 3, as teachers working in strategic staffing schools comprise just 4% of the sample. However, the policy has a much larger effect on the correlation between VA and the probability of moving within Winston-Salem. Column 2 shows that releasing VA raises the probability that a teacher with one standard deviation higher VA will move within Winston-Salem by a full percentage point, which is nearly double the effect found when examining all schools together. Also, the effect of the policy on the correlation between VA and the probability of moving out of Winston-Salem drops by 40%, when restricting analysis to moves to non-strategic staffing schools. Both changes serve to widen the gap in the estimates between moves within and out of Winston-Salem, providing further evidence of private learning.

Panel C of table 9 presents the impacts of the policy on teacher sorting within-district and out-of-district among non-strategic staffing schools. Column 1 of panel C is identical

[37]"Strategic Staffing" is the official term for later policies with the same objectives. Earlier policies had a variety of different names; Equity Plus (1 and 2), Focus School, and Mission Possible.

to column 2 of panel C in Table 6. I include it here for ease of comparison. Column 2 restricts the sample further to only include non-strategic staffing schools. Moving from column 1 to 2, in both districts, the estimated effect of the policy on the degree to which high-VA teachers sort into high performing schools becomes more positive. For Guilford, the coefficient becomes positive, though not statistically significantly so. In Winston-Salem, the point estimate of the sorting effects moves from a 60% increase in the level of within-district sorting in the rest of the state between non-strategic staffing schools to an over 75% increase.[38] Table 9 provides no evidence that strategic staffing policies are driving the earlier results. If anything, it seems that these pay policies may have muted what would otherwise have been larger impacts of releasing VA.

Table 9: Mobility between non-strategic-staffing schools with respect to school proficiency

	Panel A: Within-District Moves to non-strategic staffing schools			Panel B: Out-Of-District Moves to non-strategic staffing schools			Panel C: School Quality Growth staying within-district	
VARIABLES	Total	To a higher performing school	To a lower performing school	Total	To a higher performing school	To a lower performing school	Total	Excluding strategic staffing
VA	0.0014	0.0031	-0.0018	0.0002	0.0013	-0.0011	0.0024	0.0026
	(0.00127)	(0.00086)	(0.00076)	(0.00098)	(0.00072)	(0.00059)	(0.00033)	(0.00034)
VA x Treatment GCS	0.0043	0.0041	0.0002	-0.0111	-0.0054	-0.0057	-0.0000	0.0009
	(0.00244)	(0.00197)	(0.00148)	(0.00248)	(0.00194)	(0.0014)	(0.0007)	(0.00072)
VA x Treatment WSF	0.0100	0.0103	-0.0004	-0.0007	0.0014	-0.0021	0.0017	0.0020
	(0.00233)	(0.00176)	(0.00148)	(0.00208)	(0.00196)	(0.00113)	(0.00102)	(0.00114)
Treatment GCS	-0.0118	-0.0084	-0.0034	-0.0158	-0.0238	0.0079	-0.0157	0.0029
	(0.00848)	(0.00552)	(0.00728)	(0.00362)	(0.00221)	(0.00272)	(0.00216)	(0.00222)
Treatment WSF	0.0241	0.0390	-0.0149	-0.0027	0.0114	-0.0141	0.0231	0.0196
	(0.0049)	(0.00345)	(0.00287)	(0.00255)	(0.00233)	(0.00142)	(0.00168)	(0.0018)
Observations	236,018	236,018	236,018	236,018	236,018	236,018	202,943	197,364

CSB standard errors from 500 repetitions appear in brackets.
All regressions include teacher level covariates and interactions with treatment indicators.

[38] Table A6 provides a similar inspection instead focusing on the racial composition of the schools from which and towards which teachers move. The results are similar, except that sorting with respect to race becomes more significant in both districts when focusing only on non-strategic staffing schools and the magnitude of the mobility effects are somewhat muted.

7 Conclusion

If employers are unable to learn accurate information about their teaching force over time, their subsequent personnel decisions regarding teachers would be no better at identifying effective teachers than at the point of hire. If learning is entirely asymmetric, that is other schools are no better able to tell the effectiveness of an experienced applicant than of a novice applicant, effective teachers become trapped in schools in which they do not wish to teach, while principals shuffle their less capable teachers to other schools in what the documentary Waiting for Superman terms "The Lemon Dance" (Guggenheim, 2011). The release of value-added measures of teacher effectiveness does seem to provide actionable information to those who are aware of them. The evidence above suggests that the new information provides effective teachers with more mobility, while "The Lemon Dance" becomes focused on the uninformed.

Additionally, the evidence from subsequent teacher sorting suggests that the increase in mobility leads to increased inequity in the distribution of teacher quality across schools. Despite the fact that 38 states have adopted teacher VA, and often contentiously, this signaling role of the measures has avoided discussion. The policy implication of this finding is not to universally avoid using VA. However, it would be useful to provide policy makers an estimate of the cost of retaining high-VA teachers in hard-to-staff schools. The analysis excluding strategic staffing schools implies that the sorting may have been larger without the incentives to induce teachers to work in lower-performing schools. As mentioned in Section 6.3, several districts in North Carolina are implementing a range of staffing policies designed to induce teachers to work in low-performing schools. Some incorporate VA into the incentive schemes.

Clotfelter et al. (2011) and Glazerman et al. (2012) have examined the question of attracting teachers to understaffed schools. Further work is needed to estimate the costs and effectiveness of these policies in retaining effective teachers in low-performing schools, which may cost substantially less. As states and districts continue to adopt teacher VA, policy makers should be aware of the potential consequences of these policies on educational equity, as well as the costs of offsetting these effects.

Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Disclosure

This work was conducted under expedited IRB approval through both Michigan State University and the University of California at Riverside. The author disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through R305B090011 to Michigan State University. The opinions expressed are those of the authors and do not represent the views of the Institute or the U.S. Department of Education.

References

Ahn, T. and J. Vigdor (2012). How salient are performance incentives in education? evidence from north carolina. Technical report, Working paper, Univ. Kentucky and Duke Univ.

Ai, C. and E. C. Norton (2003). Interaction terms in logit and probit models. *Economics letters 80*(1), 123–129.

Altonji, J. G. and C. R. Pierret (2001, February). Employer learning and statistical discrimination. *The Quarterly Journal of Economics 116*(1), 313–350.

Bates, M. D. (2015, July). Job seperation under asymmetric employer learning. Working paper, University of California at Riverside.

Boyd, D., P. Grossman, H. Lankford, S. Loeb, and J. Wyckoff (2008, May). Who leaves? teacher attrition and student achievement. Working Paper 14022, National Bureau of Economic Research.

Boyd, D., H. Lankford, Loeb, and J. Wyckoff (2013, January). Analyzing the determinants of the matching of public school teachers to jobs: Disentangling the preferences of teachers and employers. *Journal of Labor Economics 31*(1), 83–117.

Boyd, D., H. Lankford, S. Loeb, and J. Wyckoff (2005, May). Explaining the short careers of high-achieving teachers in schools with low-performing students. *The American Economic Review 95*(2), 166–171.

Chetty, R., J. N. Friedman, N. Hilger, E. Saez, D. W. Schanzenbach, and D. Yagan (2011, November). How does your kindergarten classroom affect your earnings? evidence from project star. *The Quarterly Journal of Economics 126*(4), 1593–1660.

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014, September). Measuring the impacts of teachers ii: Teacher value-added and student outcomes in adulthood. *American Economic Review 104*(9), 2633–2679.

Chingos, M. M. and M. R. West (2011, June). Promotion and reassignment in public school districts: How do schools respond to differences in teacher effectiveness? *Economics of Education Review 30*(3), 419–433.

Clotfelter, C. T., H. F. Ladd, and J. L. Vigdor (2011). Teacher mobility, school segregation, and pay-based policies to level the playing field. *Education 6*(3), 399–438.

DeVaro, J. and M. Waldman (2012, January). The signaling role of promotions: Further theory and empirical evidence. *Journal of Labor Economics 30*(1), 91–147.

Farber, H. S. and R. Gibbons (1996, November). Learning and wage dynamics. *The Quarterly Journal of Economics 111*(4), 1007–1047.

Fine, J. P. and R. J. Gray (1999, June). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association 94*(446), 496–509.

Gibbons, R. and L. Katz (1991, December). Layoffs and lemons. Working Paper 2968, National Bureau of Economic Research.

Glazerman, S., A. Protik, B.-r. Teh, J. Bruch, and N. Seftor (2012). Moving high-performing teachers: Implementation of transfer incentives in seven districts. Technical report, Mathematica Policy Research.

Gray, R. J. (1988, September). A class of k-sample tests for comparing the cumulative incidence of a competing risk. *The Annals of Statistics 16*(3), 1141–1154.

Greenwald, B. C. (1986). Adverse selection in the labour market. *The Review of Economic Studies*, 325–347.

Grossman, S. J. (1981, December). The informational role of warranties and private disclosure about product quality. *Journal of Law and Economics 24*(3), 461–483.

Guarino, C., M. D. Reckase, and J. M. Wooldridge (2012). Can value-added measures of teacher performance be trusted? Technical Report 6602, Discussion Paper series, Forschungsinstitut zur Zukunft der Arbeit.

Guggenheim, D. (2011). *Waiting for Superman*. Hollywood, CA: Paramount Home Entertainment.

Jackson, C. K. (2009). Student demographics, teacher sorting, and teacher quality: Evidence from the end of school desegregation. *The Journal of Labor Economics 27*(2), 213–256.

Jackson, C. K. (2013). Match quality, worker productivity, and worker mobility: Direct evidence from teachers. *Review of Economics and Statistics*.

Jacob, B. A. and L. Lefgren (2008, January). Can principals identify effective teachers? evidence on subjective performance evaluation in education. *Journal of Labor Economics 26*(1), 101–136.

Jovanovic, B. (1979, October). Job matching and the theory of turnover. *Journal of Political Economy 87*(5), 972–990.

Kahn, L. B. (2013, October). Asymmetric information between employers. *American Economic Journal: Applied Economics 5*(4), 165–205.

Lange, F. (2007, January). The speed of employer learning. *Journal of Labor Economics 25*(1), 1–35.

Milgrom, P. and S. Oster (1987). Job discrimination, market forces, and the invisibility hypothesis. *The Quarterly Journal of Economics*, 453–476.

Milgrom, P. R. (1981, October). Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics 12*(2), 380–391.

Milgrom, P. R. and R. J. Weber (1982, September). A theory of auctions and competitive bidding. *Econometrica 50*(5), 1089–1122.

Mundlak, Y. (1978). On the pooling of time series and cross-sectional data. *Econometrica 46*, 69–86.

NCES (2016). Fast facts: Back to school statistics. *National Center for Education Statistics.* Accessed: 2016-11-30, http://nces.ed.gov/fastfacts/display.asp?id=372.

Pinkston, J. C. (2009). A model of asymmetric employer learning with testable implications. *The Review of Economic Studies 76*(1), 367–394.

Rivkin, S. G., E. A. Hanushek, and J. F. Kain (2005). Teachers, schools, and academic achievement. *Econometrica 73*(2), 417–458.

Rockoff, J. E., D. O. Staiger, T. J. Kane, and E. S. Taylor (2012, December). Information and employee evaluation: Evidence from a randomized intervention in public schools. *The American Economic Review 102*(7), 3184–3213.

Rose, R. A., G. T. Henry, and D. L. Lauen (2012, August). Comparing value-added models for estimating individual teacher effects on a statewide basis simulations and empirical analyses. Technical report, Consortium for Educational Research and Evaluation, North Carolina.

Sanders, W. L., A. M. Saxton, and S. P. Horn (1997). The tennessee value-added assessment system: A quantitative outcomes-based approach to educational assessment. *Grading teachers, grading schools: Is student achievement a valid evaluational measure*, 137–162.

Schönberg, U. (2007, October). Testing for asymmetric employer learning. *Journal of Labor Economics 25*(4), 651–691.

Spence, M. (1973, August). Job market signaling. *The Quarterly Journal of Economics 87*(3), 355–374.

Staiger, D. O. and J. E. Rockoff (2010, July). Searching for effective teachers with imperfect information. *The Journal of Economic Perspectives 24*(3), 97–117.

Vigdor, J. L. (2008). Teacher salary bonuses in north carolina. In *Conference paper, National Center on Performance Incentives.-0.026.*

Waldman, M. (1984). Job assignments, signalling, and efficiency. *The RAND Journal of Economics 15*(2), 255–267.

8 Appendices for online publication

8.1 Simplified Model Proofs

8.1.1 Comparative statics for within-district moves with respect to teacher effectiveness (μ)

Assuming the probability of moving schools is monotonically increasing in the difference between b^{h*} and b^{r*}, the sign of $\frac{\partial P[b^{h*}_{HV}-b^{r*}_{HV}>0|m\,\mu]-P[b^{h*}_{NV}-b^{r*}_{NV}>0|m\,\mu]}{\partial\mu}$ is implied by the sign of $\frac{\partial E[b^{h*}_{HV}-b^{r*}_{HV}-(b^{h*}_{NV}-b^{r*}_{NV})|m\,\mu]}{\partial\mu}$. Further, the latter requires no distributional assumptions. Here, the subscript HV denotes that hiring principals may access a teacher's VA, while the subscript NV denotes that there is no VA informing the bidding. I present the conditional expectation in equation 22 below.[39]

$$E[b^{h*}_{HV}-b^{r*}_{HV}-(b^{h*}_{NV}-b^{r*}_{NV})|m\,\mu]=\frac{\sigma_\tau(0)\sigma_\nu}{Z^h_{HV}}m+\frac{\sigma_\tau(0)\sigma_\epsilon}{Z^h_{HV}}\mu+\frac{\sigma_\epsilon\sigma_\nu}{Z^h_{HV}}\mu- \left(\frac{\sigma_\tau(t)\sigma_\nu}{Z^r_{HV}}m+\frac{\sigma_\tau(t)\sigma_\epsilon}{Z^r_{HV}}\mu+\frac{\sigma_\epsilon\sigma_\nu}{Z^r_{HV}}\mu\right)-\left(\frac{\sigma_\tau(0)}{Z^h_{NV}}m+\frac{\sigma_\epsilon}{Z^h_{NV}}\mu\right)+\left(\frac{\sigma_\tau(t)}{Z^r_{NV}}m+\frac{\sigma_\epsilon}{Z^r_{NV}}\mu\right). \tag{13}$$

Taking the derivative of equation 22 with respect to μ gives the following:

$$\frac{\partial E[b^{h*}_{HV}-b^{r*}_{HV}-(b^{h*}_{NV}-b^{r*}_{NV})|m\,\mu]}{\partial\mu}=\frac{\sigma_\tau(0)\sigma_\epsilon}{Z^h_{HV}}+\frac{\sigma_\epsilon\sigma_\nu}{Z^h_{HV}}-\left(\frac{\sigma_\tau(t)\sigma_\epsilon}{Z^r_{HV}}+\frac{\sigma_\epsilon\sigma_\nu}{Z^r_{HV}}\right)-\frac{\sigma_\epsilon}{Z^h_{NV}}+\frac{\sigma_\epsilon}{Z^r_{NV}}.$$

$$=\frac{\sigma_\epsilon(\sigma_\tau(0)-\sigma_\tau(t))}{Z^h_{NV}Z^r_{NV}}-\frac{\sigma^2_\nu\sigma_\epsilon(\sigma_\tau(0)-\sigma_\tau(t))}{Z^h_{HV}Z^r_{HV}}$$

$$=\frac{\sigma_\epsilon(\sigma_\tau(0)-\sigma_\tau(t))[Z^h_{HV}Z^r_{HV}-Z^h_{NV}Z^r_{NV}\sigma^2_\nu]}{Z^h_{NV}Z^r_{NV}Z^h_{HV}Z^r_{HV}}$$

$$\frac{\partial E[b^{h*}_{HV}-b^{r*}_{HV}-(b^{h*}_{NV}-b^{r*}_{NV})|m\,\mu]}{\partial\mu}= \frac{{\sigma_\epsilon}^2(\sigma_\tau(0)-\sigma_\tau(t))}{Z^h_{NV}Z^r_{NV}Z^h_{HV}Z^r_{HV}}(\sigma_\tau(0)\sigma_\tau(t)\sigma_\epsilon+2\sigma_\nu\sigma_\tau(0)\sigma_\tau(t)+\sigma_\epsilon\sigma_\nu\sigma_\tau(0)+\sigma_\epsilon\sigma_\nu\sigma_\tau(t)) \tag{14}$$

The above appears as equation 7 in text. The key assumption driving this prediction is $\sigma_\tau(0)-\sigma_\tau(t)>0$ which implies asymmetric employer learning. All other terms are positive variances, which implies that $\frac{\partial E[b^{h*}_{HV}-b^{r*}_{HV}-(b^{h*}_{NV}-b^{r*}_{NV})|m\,\mu]}{\partial\mu}>0$, which in turn implies that the probability of moving within-district increases with the policy and increases in μ.

[39] Recall that $Z^r_{HV}=\sigma_\tau(t)\sigma_\nu+\sigma_\tau(t)\sigma_\epsilon+\sigma_\epsilon\sigma_\nu$, $Z^h_{HV}=\sigma_\tau(0)\sigma_\nu+\sigma_\tau(0)\sigma_\epsilon+\sigma_\epsilon\sigma_\nu$, $Z^r_{NV}=\sigma_\tau(t)+\sigma_\epsilon$, and $Z^h_{NV}=\sigma_\tau(0)+\sigma_\epsilon$.

8.1.2 Comparative statics for within-district moves with respect to VA (V)

In determining the comparative statics with regard to the VA signal, I seek to sign $\frac{\partial E[b^{h*}_{HV} - b^{r*}_{HV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ V\ \mu]}{\partial V}$.

$$\begin{aligned}
\frac{\partial E[b^{h*}_{HV} - b^{r*}_{HV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ V\ \mu]}{\partial V} =& \frac{\sigma_\tau(0)\sigma_\epsilon}{Z^h_{HV}} - \frac{\sigma_\tau(t)\sigma_\epsilon}{Z^r_{HV}} \\
=& \frac{\sigma_\tau(0)\sigma_\epsilon(\sigma_\tau(t)\sigma_\nu + \sigma_\tau(t)\sigma_\epsilon + \sigma_\epsilon\sigma_\nu)}{Z^h_{HV} Z^r_{HV}} \\
-& \frac{\sigma_\tau(t)\sigma_\epsilon(\sigma_\tau(0)\sigma_\nu + \sigma_\tau(0)\sigma_\epsilon + \sigma_\epsilon\sigma_\nu)}{Z^h_{HV} Z^r_{HV}} \\
=& \frac{{\sigma_\epsilon}^2\sigma_\nu(\sigma_\tau(0) - \sigma_\tau(t))}{Z^h_{HV} Z^r_{HV}}.
\end{aligned}$$

Again, the assumption that $sigto > sigt$ forces the expression to positive. Thus, releasing VA raises the probability that high-VA teachers move schools.

8.1.3 Comparative statics for out-of-district moves with respect to teacher effectiveness (μ)

In determining the prediction for out-of-district moves, I seek to sign $\frac{\partial E[b^{h*}_{NV} - b^{r*}_{RV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ \mu]}{\partial \mu}$. Here, the subscript RV denotes that only retaining principals may access a teacher's VA, while the subscript NV denotes that there is no VA informing the bidding. The first thing to note is that hiring principals bids cancel each other. Thus, I focus on retaining principals' bids with and without VA. Letting $Z^r_{RV} = \sigma_\xi(x)\sigma_\tau(t\ V) + \sigma_\epsilon\sigma_\tau(t\ V) + \sigma_\epsilon\sigma_\xi(x)$, equation 25 gives the conditional expectation of this difference.

$$E[b^{h*}_{NV} - b^{r*}_{RV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ \mu] = \frac{\sigma_\tau(t)}{Z^r_{NV}}m + \frac{\sigma_\epsilon}{Z^r_{NV}}\mu - \left(\frac{\sigma_\tau(t\ V)}{Z^r_{RV}}m + \frac{\sigma_\epsilon}{Z^r_{RV}}\mu\right) \tag{15}$$

Taking the derivative with respect to μ gives:

$$\begin{aligned}
\frac{\partial E[b^{h*}_{NV} - b^{r*}_{RV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ \mu]}{\partial \mu} =& \frac{\sigma_\epsilon}{Z^r_{NV}} - \frac{\sigma_\epsilon}{Z^r_{RV}} \\
=& \frac{\sigma_\epsilon(\sigma_\epsilon + \sigma_\tau(t\ V)) - (\sigma_\epsilon + \sigma_\tau(t))]}{Z^r_{NV} Z^r_{RV}}
\end{aligned}$$

$$\frac{\partial E[b^{h*}_{NV} - b^{r*}_{RV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ \mu]}{\partial \mu} = \frac{\sigma_\epsilon(\sigma_\tau(t\ V) - \sigma_\tau(t))}{Z^r_{NV} Z^r_{RV}}.$$

The above appears as equation 9 in text. Lemma 1 demonstrates that $\sigma_\tau(t) - \sigma_\tau(t\ V) > 0$. All other terms are positive variances, implying that $\frac{\partial E[b^{h*}_{RV} - b^{r*}_{RV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ \mu]}{\partial \mu} < 0$, which in turn implies that the probability of transitions to uninformed principals increases with declines in teacher effectiveness (μ).

8.1.4 Comparative statics for out-of-district moves with respect to VA (V)

In determining the comparative statics with regard to the VA signal, I seek to sign $\frac{\partial E[b^{h*}_{RV} - b^{r*}_{RV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\,V\,\mu]}{\partial V}$. Recall that:

$$b^{r*}_{RV} = \frac{\sigma_\tau(t\,V)}{Z^r_{RV}} m + \frac{\sigma_\epsilon}{Z^r_{RV}} \left(\frac{\sigma_\nu P^r_t + \sigma_\tau(t) V}{\sigma_\nu + \sigma_\tau(t)} \right).$$

$$\frac{\partial E\left[b^{h*}_{NV} - b^{r*}_{RV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\,V\,\mu\right]}{\partial V} = \frac{-\sigma_\epsilon \sigma_\tau(t)}{Z^r_{RV}(\sigma_\nu + \sigma_\tau(t))}.$$

As equation 10 is the negative of a function of variances, it is less than zero. Thus after VA is released, as a teacher's VA decreases, the probability of moving to uniformed principals increases.

8.2 Full Model

8.2.1 Structure

The full model has much of the same structure as the simplified version with three primary exceptions. First, I intorduce a public signal R_x to nest the symmetric model of employer learning within this broader model. Second, recognizing that hiring may be a dynamic process, I allow each principal to update expectations of teacher quality, and thus bids, depending on the actions of the rival principal. Third, it may be unrealistic to presume that principals may overcome large differences in pay or school attributes with position-specific attributes. Thus, I impose a school-level, proportional constraint on principals bids, which increases in school quality. I outline the structure of the full model below.

1. The public signal is given by $R_x = \mu + \xi_x$, where $\xi \sim N(0, \sigma_\xi(x))$, and $\frac{\partial \sigma_\xi(x)}{\partial x} < 0$.
2. Private signal:
   (a) For hiring principals (denoted by the superscript h), the private signal is given by$P^h = \mu + \tau^h$ where $\tau^h \sim N(0, \sigma_\tau(0))$. $\sigma_\tau(0)$ is fixed over time.
   (b) For a retaining principal (denoted by the superscript r), the private signal is given by $P^r_t = \mu + \tau^r_t$ where $\tau^r_t \sim N(0, \sigma_\tau(t))$ and $\frac{\partial \sigma_\tau(t)}{\partial t} < 0$.
3. The VAM serve as an additional piece of information that may alter both the mean and precision of the public or private signal depending on whether it is available to both bidding principals. It has the form $V = \mu + \nu$, where $\nu \sim N(0, \sigma_\nu)$.
   (a) When both principals are informed by VAMs, the public signal becomes $R_{x\nu} = \frac{\sigma_\nu R_x + \sigma_\xi(x) V}{\sigma_\nu + \sigma_\xi(x)}$. The variance of $R_{x\nu}$ is denoted as $\sigma_\xi(x\,V)$.

(b) When only the retaining principal is informed by VAMs, her private signal becomes $P^r_{t\nu} = \frac{\sigma_\nu P^r_t + \sigma_\tau(t)V}{\sigma_\nu + \sigma_\tau(t)}$. The variance of $P^r_{t\nu}$ is denoted as $\sigma_\tau(t\,V)$. The hiring principal's signal remains unchanged.

4. $c \sim N(0, \sigma_c)$ represents an idiosyncratic cost.

5. The noise of each signal is orthogonal to the noise of the other signals.[40]

Teachers know their effectiveness (μ), but cannot credibly reveal it. As a teacher begins his career, all principals begin with the prior belief that he is as good as the average teacher with his same characteristics (m). The teacher encounters two principals to whom he may privately (but noisily) signal his ability akin to an interview, (denoted by P^h_0 where 0 indicates no additional private information).

Over time, teachers may draw on their experience to bolster their public signals denoted by R_x (for examples consider resumés and networks of references). If there is public learning, the variance of the public signal ($\sigma_\xi(x)$) will shrink with teacher experience (x), as more information comes into the market $\left(\frac{\partial \sigma_\xi(x)}{\partial x} < 0\right)$.

Retaining principals may obtain private information unavailable to rival employers (P^r_t) the longer a teacher teaches within the school (t). If such private learning occurs, the precision of the current principal's signal ($\sigma_\tau(t)$) increases the longer a teacher works in the school, while hiring principals' private signals from interviewing the teacher have a constantly high variance ($\sigma_\tau(0)$). Thus, the accumulation of private information leads to $\sigma_\tau(t) < \sigma_\tau(0)$ for all $t > 0$. In order to nest symmetric learning within the more flexible model, I maintain that that even in this special case, employers receive a private signal each period, but the variance of the signal is constant over years of tenure ($\sigma_\tau(t) = \sigma_\tau(0)$ for all $t > 0$).

VA enters the learning model as an additional piece of information influencing either the public or private signal. VAs influence the public signal if they are accessible to both principals, transforming the public signal into $R_{x\nu} = \frac{\sigma_\nu R_x + \sigma_\xi(x)V}{\sigma_\nu + \sigma_\xi(x)}$. VAs impact the private signal, if they are accessible to only current principals, making the private signal become $P^r_{t\nu} = \frac{\sigma_\nu P^r_t + \sigma_\tau(t)V}{\sigma_\nu + \sigma_\tau(t)}$.

8.2.2 Bids

While the components of compensation remain largely unchanged from the model in text, the public signals and more dynamic bidding structure alters principals' optimal bids. Allowing principals expectations of a candidate to change as she sees outside demand for the candidate may be more realistic than the straightforward sealed bid second price auction. To show the other extreme, I model the bidding as an open and continuous English auction. This permits the adoption of optimal bidding strategies from Milgrom and Weber (1982), in which each principal updates her expectations and thus optimal bid conditioning on the rival's bidding behavior. As the presence of a rival bidder verifies that one other employer received a signal at least as positive as the private signal of the principal, she places additional weight upon her private signal.

[40] The orthogonality assumptions are also not necessary to derive the following predictions. However, relaxing these require a less restrictive, though more complicated set of assumptions, outlining the direction and magnitude of correlations between the errors of the signals.

Below, I enumerate the structures of the optimal bids.

- Hiring principals' bid in absence of VA:

$$b^{h*}_{isdNV} = \frac{\sigma_\tau(0)\sigma_\xi(x)}{Z^h_{NV}} m + \frac{\sigma_\tau(0)\sigma_\epsilon}{Z^h_{NV}} R_x + \frac{2\sigma_\epsilon\sigma_\xi(x)}{Z^h_{NV}} P^h_0.\text{[41]} \tag{16}$$

- Retaining principals' bid in absence of VA:

$$b^{r*}_{isdNV} = \frac{\sigma_\tau(t)\sigma_\xi(x)}{Z^r_{NV}} m + \frac{\sigma_\tau(t)\sigma_\epsilon}{Z^r_{NV}} R_x + \frac{2\sigma_\epsilon\sigma_\xi(x)}{Z^r_{NV}} P^r_t.\text{[42]} \tag{17}$$

- Retaining principals' bid with private VA:

$$b^{r*}_{isdRV} = \frac{\sigma_\tau(t\,V)\sigma_\xi(x)}{Z^r_{RV}} m + \frac{\sigma_\tau(t\,V)\sigma_\epsilon}{Z^r_{RV}} R_x + \frac{2\sigma_\epsilon\sigma_\xi(x)}{Z^r_{RV}} P^r_{t\nu}.\text{[43]} \tag{18}$$

- Hiring principals' bid with public VA:

$$b^{h*}_{isdHV} = \frac{\sigma_\tau(0)\sigma_\xi(x\,V)}{Z^r_{HV}} m + \frac{\sigma_\tau(0)\sigma_\epsilon}{Z^r_{HV}} R_{x\nu} + \frac{2\sigma_\epsilon\sigma_\xi(x\,V)}{Z^r_{HV}} P^h_0.\text{[44]} \tag{19}$$

- Retaining principals' bid with public VA:

$$b^{r*}_{isdHV} = \frac{\sigma_\tau(t)\sigma_\xi(x\,V)}{Z^r_{HV}} m + \frac{\sigma_\tau(t)\sigma_\epsilon}{Z^r_{HV}} R_{x\nu} + \frac{2\sigma_\epsilon\sigma_\xi(x\,V)}{Z^r_{HV}} P^r_t.\text{[45]} \tag{20}$$

Again, whether VA information is public or private determines whether the expectations and bids of hiring and retaining principals converge or diverge with the introduction of VA. If only a retaining principal receives a teacher's VA, she incorporates it into her private signal (denoted by the subscript RV). The new private signal ($P^r_{t\nu}$) becomes the precision-weighted average of the prior private information and the new VA. In which case, the retaining principal's optimal bid is shown in equation 18, while the hiring principal's bid remains unchanged from equation 16. If VAs are informative, the precision of the cumulative private information must increase, as was shown by Lemma 1. Since the hiring principals' expectations do not change, the introduction of VAs exacerbates informational asymmetries between prospective employers, and the two principals' bids further diverge.

In contrast, if both bidding principals are informed of a teacher's VA, the new information enters the public signal. Equations 19 and 20 reflect the new public signal, $R_{x\nu}$. While in expectation the magnitude of the public signal is the same with or without VAs, Lemma 2 shows that the variance of the public signal must change as a result.

Lemma 2: The precision of the public signal increases with the incorporation of VAs into the public signal ($\sigma_\xi(x\,V) < \sigma_\xi(x)$).

[41] $Z^h_{NV} = \sigma_\tau(0)\sigma_\xi(x) + \sigma_\tau(0)\sigma_\epsilon + 2\sigma_\epsilon\sigma_\xi(x)$.
[42] $Z^r_{NV} = \sigma_\tau(t)\sigma_\xi(x) + \sigma_\tau(t)\sigma_\epsilon + 2\sigma_\epsilon\sigma_\xi(x)$.
[43] $Z^r_{RV} = \sigma_\tau(t\,V)\sigma_\xi(x) + \sigma_\tau(t\,V)\sigma_\epsilon + 2\sigma_\epsilon\sigma_\xi(x)$.
[44] $Z^h_{HV} = \sigma_\tau(0)\sigma_\xi(x\,V) + \sigma_\tau(0)\sigma_\epsilon + 2\sigma_\epsilon\sigma_\xi(x\,V)$.
[45] $Z^r_{HV} = \sigma_\tau(t)\sigma_\xi(x\,V) + \sigma_\tau(t)\sigma_\epsilon + 2\sigma_\epsilon\sigma_\xi(x\,V)$.

Proof: Under the orthogonality assumptions, $var(R_{x\nu}) \equiv \sigma_\xi(x\,V) = \frac{\sigma_\nu^2\sigma_\xi(x)+\sigma_\nu\sigma_\xi(x)^2}{(\sigma_\nu+\sigma_\xi(x))^2} = \frac{\sigma_\nu\sigma_\xi(x)}{\sigma_\nu+\sigma_\xi(x)} \cdot \frac{\sigma_\xi(x)(\sigma_\nu+\sigma_\xi(x))}{\sigma_\nu+\sigma_\xi(x)} - \frac{\sigma_\nu\sigma_\xi(x)}{\sigma_\nu+\sigma_\xi(x)} = \frac{\sigma_\xi^2(x)}{\sigma_\nu+\sigma_\xi(x)} \cdot \frac{\sigma_\xi^2(x)}{\sigma_\nu+\sigma_\xi(x)} > 0$, by property of variances.

Using the finding from Lemma 2, that the variance of the public signal drops with the introduction of VAs, once hiring and retaining principals may access a teacher's VA, they shift weight from their prior beliefs and their private information and place it onto the public information that now includes a teacher's VA. For bids in which both principals become informed of a teacher's VA, the information between prospective employers becomes more symmetric, and their expectations converge.

After teachers receive both bids, they move to the school that maximize their utility. Accordingly, the probability of a move is:

$$P(M) = P\left[b_{isd}^{h*} - b_{isd}^{r*} > 0\right]. \tag{21}$$

I investigate how we expect this probability to change in the different contexts the setting provides. Primarily, by examining teacher mobility in response to the release of VAs, I test whether releasing VAs provides greater informational symmetry between employers. However, out-of-district principals cannot directly access these new VAs. Thus, examining mobility out of adopting districts evidences whether the information spreads to all employers or furthers informational asymmetries between them. I then consider heterogeneity in moves to higher- and lower-performing schools and with respect to teacher tenure within current school. I derive relevant predictions below.

8.2.3 Comparative statics for within-district moves with respect to teacher effectiveness (μ)

Assuming the probability of moving schools is monotonically increasing in the difference between b^{h*} and b^{r*}, the sign of $\frac{\partial P[b_{HV}^{h*}-b_{HV}^{r*}>0|m\,\mu]-P[b_{NV}^{h*}-b_{NV}^{r*}>0|m\,\mu]}{\partial\mu}$ is implied by the sign of $\frac{\partial E[b_{HV}^{h*}-b_{HV}^{r*}-(b_{NV}^{h*}-b_{NV}^{r*})|m\,\mu]}{\partial\mu}$. Here, the subscript HV denotes that hiring principals may access a teacher's VA, while the subscript NV denotes that there is no VA informing the bidding. I present the conditional expectation in equation 22 below.

$$\begin{aligned}
E[b_{HV}^{h*} - b_{HV}^{r*} - (b_{NV}^{h*} - b_{NV}^{r*})|m\,\mu] =& \frac{\sigma_\tau(0)\sigma_\xi(x\,V)}{Z_{HV}^h}m + \frac{\sigma_\tau(0)\sigma_\epsilon}{Z_{HV}^h}\mu + \frac{2\sigma_\epsilon\sigma_\xi(x\,V)}{Z_{HV}^h}\mu \\
&- \left(\frac{\sigma_\tau(t)\sigma_\xi(x\,V)}{Z_{HV}^r}m + \frac{\sigma_\tau(t)\sigma_\epsilon}{Z_{HV}^r}\mu + \frac{2\sigma_\epsilon\sigma_\xi(x\,V)}{Z_{HV}^r}\mu\right) \\
&- \left(\frac{\sigma_\tau(0)\sigma_\xi(x)}{Z_{NV}^h}m + \frac{\sigma_\tau(0)\sigma_\epsilon}{Z_{NV}^h}\mu + \frac{2\sigma_\epsilon\sigma_\xi(x)}{Z_{NV}^h}\mu\right) \\
&+ \left(\frac{\sigma_\tau(t)\sigma_\xi(x)}{Z_{NV}^r}m + \frac{\sigma_\tau(t)\sigma_\epsilon}{Z_{NV}^r}\mu + \frac{2\sigma_\epsilon\sigma_\xi(x)}{Z_{NV}^r}\mu\right).
\end{aligned} \tag{22}$$

Taking the derivative of equation 22 with respect to μ gives the following:

$$\begin{aligned}\frac{\partial E[b^{h*}_{HV}-b^{r*}_{HV}-(b^{h*}_{NV}-b^{r*}_{NV})|m\,\mu]}{\partial\mu} &= \frac{\sigma_\tau(0)\sigma_\epsilon}{Z^h_{HV}}+\frac{2\sigma_\epsilon\sigma_\xi(x\,V)}{Z^h_{HV}}-\left(\frac{\sigma_\tau(t)\sigma_\epsilon}{Z^r_{HV}}+\frac{2\sigma_\epsilon\sigma_\xi(x\,V)}{Z^r_{HV}}\right)\\ &-\left(\frac{\sigma_\tau(0)\sigma_\epsilon}{Z^h_{NV}}+\frac{2\sigma_\epsilon\sigma_\xi(x)}{Z^h_{NV}}\right)+\left(\frac{\sigma_\tau(t)\sigma_\epsilon}{Z^r_{NV}}+\frac{2\sigma_\epsilon\sigma_\xi(x)}{Z^r_{NV}}\right).\end{aligned}$$

$$=\frac{2\sigma_\xi(x)^2\sigma_\epsilon(\sigma_\tau(0)-\sigma_\tau(t))}{Z^h_{NV}Z^r_{NV}}-\frac{\sigma_\xi(x\,V)^2\sigma_\epsilon(\sigma_\tau(0)-\sigma_\tau(t))}{Z^h_{HV}Z^r_{HV}}$$

$$=\frac{2\sigma_\epsilon(\sigma_\tau(0)-\sigma_\tau(t))[Z^h_{HV}Z^r_{HV}\sigma_\xi(x)^2-Z^h_{NV}Z^r_{NV}\sigma_\xi(x\,V)^2]}{Z^h_{NV}Z^r_{NV}Z^h_{HV}Z^r_{HV}}$$

$$\begin{aligned}\frac{\partial E[b^{h*}_{HV}-b^{r*}_{HV}-(b^{h*}_{NV}-b^{r*}_{NV})|m\,\mu]}{\partial\mu} &= \frac{2\sigma_\epsilon(\sigma_\tau(0)-\sigma_\tau(t))}{Z^h_{NV}Z^r_{NV}Z^h_{HV}Z^r_{HV}}[\sigma_\xi(x)^2(\sigma_\tau(t)\sigma_\xi(x\,V)^2\sigma_\tau(0)\\ &+4\sigma_\epsilon^2\sigma_\xi(x\,V)^2+2\sigma_\epsilon\sigma_\xi(x\,V)^2\sigma_\tau(0)+\sigma_\tau(t)\sigma_\epsilon^2\sigma_\tau(0)\\ &+2\sigma_\epsilon^2\sigma_\xi(x\,V)\sigma_\tau(0)+2\sigma_\tau(t)\sigma_\epsilon\sigma_\tau(0)\sigma_\xi(x\,V)\\ &+2\sigma_\tau(t)\sigma_\xi(x\,V)^2\sigma_\epsilon)+2\sigma_\tau(t)\sigma_\epsilon^2\sigma_\xi(x\,V))\\ &-\sigma_\xi(x\,V)^2(4\sigma_\epsilon^2\sigma_\xi(x)^2+2\sigma_\tau(t)\sigma_\xi(x)\sigma_\tau(0)\sigma_\epsilon\\ &+\sigma_\tau(t)\sigma_\epsilon^2\sigma_\tau(0)+2\sigma_\epsilon\sigma_\xi(x)^2\sigma_\tau(0)+\sigma_\tau(t)\sigma_\tau(0)\sigma_\xi(x)^2\\ &+2\sigma_\xi(x)\sigma_\tau(0)\sigma_\epsilon^2+2\sigma_\tau(t)\sigma_\xi(x)^2\sigma_\epsilon+2\sigma_\tau(t)\sigma_\epsilon^2\sigma_\xi(x))]\end{aligned}$$

$$\begin{aligned}&=\frac{2\sigma_\epsilon(\sigma_\tau(0)-\sigma_\tau(t))}{Z^h_{NV}Z^r_{NV}Z^h_{HV}Z^r_{HV}}[\sigma_\xi(x)^2(2\sigma_\tau(t)\sigma_\epsilon\sigma_\tau(0)\sigma_\xi(x\,V)\\ &+\sigma_\tau(t)\sigma_\epsilon^2\sigma_\tau(0)+2\sigma_\epsilon^2\sigma_\xi(x\,V)\sigma_\tau(0)+2\sigma_\tau(t)\sigma_\epsilon^2\sigma_\xi(x\,V))\\ &-\sigma_\xi(x\,V)^2(2\sigma_\tau(t)\sigma_\epsilon\sigma_\tau(0)\sigma_\xi(x)+\sigma_\tau(t)\sigma_\epsilon^2\sigma_\tau(0)\\ &+2\sigma_\epsilon^2\sigma_\xi(x)\sigma_\tau(0)+2\sigma_\tau(t)\sigma_\epsilon^2\sigma_\xi(x))]\end{aligned}$$

$$\begin{aligned}\frac{\partial E[b^{h*}_{HV}-b^{r*}_{HV}-(b^{h*}_{NV}-b^{r*}_{NV})|m\,\mu]}{\partial\mu} &= \frac{2\sigma_\epsilon{}^2(\sigma_\tau(0)-\sigma_\tau(t))(\sigma_\xi(x)-\sigma_\xi(x\,V))}{Z^h_{NV}Z^r_{NV}Z^h_{HV}Z^r_{HV}}\\ &\quad[2\sigma_\xi(x)\sigma_\xi(x\,V)(\sigma_\tau(t)\sigma_\tau(0)+\sigma_\epsilon\sigma_\tau(0)+\sigma_\tau(t)\sigma_\epsilon)\\ &\quad+(\sigma_\xi(x\,V)+\sigma_\xi(x))\sigma_\tau(t)\sigma_\epsilon\sigma_\tau(0))].\end{aligned}\tag{23}$$

$\frac{1}{Z^h_{HV}Z^r_{HV}Z^h_{NV}Z^r_{NV}}$ is positive, as it is purely a function of variances. As a fundamental component of asymmetric employer learning, it is assumed that $\sigma_\tau(0)-\sigma_\tau(t)>0$. If VA is at all informative, lemma 2 shows that $\sigma_\xi(x\,V)-\sigma_\xi(x)<0$. All other terms are positive variances, which implies that $\frac{\partial E[b^{h*}_{HV}-b^{r*}_{HV}-(b^{h*}_{NV}-b^{r*}_{NV})|m\,\mu]}{\partial\mu}>0$, which in turn implies that the probability of moving within-district increases with increases in μ.

8.2.4 Comparative statics for within-district moves with respect to VA (V)

In determining the comparative statics with regard to the VA signal, I seek to sign $\frac{\partial E[b^{h*}_{HV} - b^{r*}_{HV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ V\ \mu]}{\partial V}$. Explicitly showing V allows equation 22 to be written as follows.

$$
\begin{aligned}
E[b^{h*}_{HV} - b^{r*}_{HV} - &(b^{h*}_{NV} - b^{r*}_{NV})|m\ \mu] = \\
&\frac{\sigma_\tau(0)\sigma_\xi(x\ V)}{Z^h_{HV}}m + \frac{\sigma_\tau(0)\sigma_\epsilon}{Z^h_{HV}}\frac{\sigma_\nu\mu + \sigma_\xi(x)V}{\sigma_\nu + \sigma_\xi(x)} + \frac{2\sigma_\epsilon\sigma_\xi(x\ V)}{Z^h_{HV}}\mu \\
&- \left(\frac{\sigma_\tau(t)\sigma_\xi(x\ V)}{Z^r_{HV}}m + \frac{\sigma_\tau(t)\sigma_\epsilon}{Z^r_{HV}}\frac{\sigma_\nu\mu + \sigma_\xi(x)V}{\sigma_\nu + \sigma_\xi(x)} + \frac{2\sigma_\epsilon\sigma_\xi(x\ V)}{Z^r_{HV}}\mu\right) \\
&- \left(\frac{\sigma_\tau(0)\sigma_\xi(x)}{Z^h_{NV}}m + \frac{\sigma_\tau(0)\sigma_\epsilon}{Z^h_{NV}}\mu + \frac{2\sigma_\epsilon\sigma_\xi(x)}{Z^h_{NV}}\mu\right) \\
&+ \left(\frac{\sigma_\tau(t)\sigma_\xi(x)}{Z^r_{NV}}m + \frac{\sigma_\tau(t)\sigma_\epsilon}{Z^r_{NV}}\mu + \frac{2\sigma_\epsilon\sigma_\xi(x)}{Z^r_{NV}}\mu\right).
\end{aligned}
\tag{24}
$$

Taking the derivative with respect to VA (V) provides the following.[46]

$$
\frac{\partial E\left[b^{h*}_{HV} - b^{r*}_{HV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ V\ \mu\right]}{\partial V} = \frac{2{\sigma_\epsilon}^2\sigma_\xi(x)(\sigma_\tau(0) - \sigma_\tau(t))}{Z^h_{HV}Z^r_{HV}} > 0
$$

As a fundamental component of asymmetric employer learning, it is assumed that $\sigma_\tau(0) - \sigma_\tau(t) > 0$. Meaning that releasing VA raises the probability that high-VA teachers move schools.

8.2.5 Comparative statics for out-of-district moves with respect to teacher effectiveness (μ)

Here, the subscript RV denotes that only retaining principals may access a teacher's VA, while the subscript NV denotes that there is no VA informing the bidding. The first thing to note is that hiring principals bids cancel each other. Thus, I focus on retaining principals' bids with and without VA. Letting $Z^r_{RV} = \sigma_\xi(x)\sigma_\tau(t\ V) + \sigma_\epsilon\sigma_\tau(t\ V) + \sigma_\epsilon\sigma_\xi(x)$, equation 25 gives the conditional expectation of this difference.

$$
\begin{aligned}
E[b^{h*}_{NV} - b^{r*}_{RV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ \mu] = &\frac{\sigma_\tau(t)\sigma_\xi(x)}{Z^r_{NV}}m + \frac{\sigma_\tau(t)\sigma_\epsilon}{Z^r_{NV}}\mu + \frac{2\sigma_\epsilon\sigma_\xi(x)}{Z^r_{NV}}\mu \\
&- \left(\frac{\sigma_\tau(t\ V)\sigma_\xi(x)}{Z^r_{RV}}m + \frac{\sigma_\tau(t\ V)\sigma_\epsilon}{Z^r_{RV}}\mu + \frac{2\sigma_\epsilon\sigma_\xi(x)}{Z^r_{RV}}\mu\right)
\end{aligned}
\tag{25}
$$

[46] $\frac{\partial\sigma_\xi(x\ V)}{\partial V} = 0$, since the variance of the signal does not depend on the magnitude of the signal.

Taking the derivative with respect to μ gives:

$$\frac{\partial E[b^{h*}_{NV} - b^{r*}_{RV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ \mu]}{\partial \mu} = \frac{\sigma_\tau(t)\sigma_\epsilon}{Z^r_{NV}} + \frac{2\sigma_\epsilon\sigma_\xi(x)}{Z^r_{NV}} - \left(\frac{\sigma_\tau(t\ V)\sigma_\epsilon}{Z^r_{RV}} + \frac{2\sigma_\epsilon\sigma_\xi(x)}{Z^r_{RV}}\right)$$

$$= \frac{\sigma_\epsilon[(\sigma_\tau(t) + 2\sigma_\xi(x))Z^r_{RV} - (\sigma_\tau(t\ V) + 2\sigma_\xi(x))Z^r_{NV}]}{Z^r_{NV} Z^r_{RV}}$$

$$\begin{aligned} &= \frac{\sigma_\epsilon}{Z^r_{NV} Z^r_{RV}}[(\sigma_\tau(t) + 2\sigma_\xi(x))(\sigma_\xi(x)\sigma_\tau(t\ V) \\ &+ \sigma_\epsilon\sigma_\tau(t\ V) + 2\sigma_\epsilon\sigma_\xi(x)) - (\sigma_\tau(t\ V) + 2\sigma_\xi(x)) \\ &(\sigma_\xi(x)\sigma_\tau(t) + \sigma_\epsilon\sigma_\tau(t) + 2\sigma_\epsilon\sigma_\xi(x))] \end{aligned}$$

$$\frac{\partial E[b^{h*}_{NV} - b^{r*}_{RV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ \mu]}{\partial \mu} = \frac{2\sigma_\epsilon\sigma_\xi(x)^2(\sigma_\tau(t\ V) - \sigma_\tau(t))}{Z^r_{NV} Z^r_{RV}}.$$

The above appears as equation 9 in text. Lemma 1 demonstrates that $\sigma_\tau(t) - \sigma_\tau(t\ V) > 0$. All other terms are positive variances, implying that $\frac{\partial E[b^{h*}_{RV} - b^{r*}_{RV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ \mu]}{\partial \mu} < 0$, which in turn implies that the probability of transitions to uninformed principals increases with declines in teacher effectiveness (μ).

8.2.6 Comparative statics for out-of-district moves with respect to VA (V)

In determining the comparative statics with regard to the VA signal, I seek to sign $\frac{\partial E[b^{h*}_{RV} - b^{r*}_{RV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ V\ \mu]}{\partial V}$.

$$\begin{aligned} E[b^{h*}_{HV} - b^{r*}_{HV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ \mu] = \\ &\frac{\sigma_\tau(0)\sigma_\xi(x)}{Z^h_{NV}}m + \frac{\sigma_\tau(0)\sigma_\epsilon}{Z^h_{NV}}\mu + \frac{2\sigma_\epsilon\sigma_\xi(x)}{Z^h_{NV}}\mu \\ &- \left(\frac{\sigma_\tau(t\ V)\sigma_\xi(x)}{Z^r_{RV}}m + \frac{\sigma_\tau(t\ V)\sigma_\epsilon}{Z^r_{RV}}\mu + \frac{2\sigma_\epsilon\sigma_\xi(x)}{Z^r_{RV}}\frac{\sigma_\nu P^r_t + \sigma_\tau(t)V}{\sigma_\nu + \sigma_\tau(t)}\right) \\ &- \left(\frac{\sigma_\tau(0)\sigma_\xi(x)}{Z^h_{NV}}m + \frac{\sigma_\tau(0)\sigma_\epsilon}{Z^h_{NV}}\mu + \frac{2\sigma_\epsilon\sigma_\xi(x)}{Z^h_{NV}}\mu\right) \\ &+ \left(\frac{\sigma_\tau(t)\sigma_\xi(x)}{Z^r_{NV}}m + \frac{\sigma_\tau(t)\sigma_\epsilon}{Z^r_{NV}}\mu + \frac{2\sigma_\epsilon\sigma_\xi(x)}{Z^r_{NV}}\mu\right). \end{aligned} \tag{26}$$

The derivative of equation 26 with respect to the VA signal (V) is presented below:

$$\frac{\partial E\left[b^{h*}_{NV} - b^{r*}_{RV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ V\ \mu\right]}{\partial V} = \frac{-2\sigma_\xi(x)\sigma_\epsilon\sigma_\tau(t)}{Z^r_{RV}(\sigma_\nu + \sigma_\tau(t))} < 0$$

As equation 8.2.6 is the negative of a function of variances, it is less than zero. Thus after VA is released, as a teacher's VA decreases, the probability of moving to uniformed principals increases.

8.2.7 Informed out-of-district principals

It is important to note that good (or high-VA) teachers may choose to reveal their VA to out-of-district principals. Accordingly, the furthering of information asymmetries between employers may not universally apply to out-of-district moves. It may be truer to the setting to examine the expected difference in differences of bids between pre- and post-VA years, allowing for a mix between informed and uninformed out-of-district principals. In this context let δ_d be the home-district-specific probability that the outside principal is informed of the teacher's VA. Equation 27 gives the conditional expectation of this difference.

$$
\begin{aligned}
E[\delta_d(b^{h*}_{HV} - b^{r*}_{HV}) + &(1-\delta_d)(b^{h*}_{NV} - b^{r*}_{RV}) - (b^{h*}_{NV} - b^{r*}_{NV})|m\,\mu] = \\
&\delta_d\left(\frac{\sigma_\tau(0)\sigma_\xi(x\,V)}{Z^h_{HV}}m + \frac{\sigma_\tau(0)\sigma_\epsilon}{Z^h_{HV}}\mu + \frac{2\sigma_\epsilon\sigma_\xi(x\,V)}{Z^h_{HV}}\mu\right) \\
&-\delta_d\left(\frac{\sigma_\tau(t)\sigma_\xi(x\,V)}{Z^r_{HV}}m + \frac{\sigma_\tau(t)\sigma_\epsilon}{Z^r_{HV}}\mu + \frac{2\sigma_\epsilon\sigma_\xi(x\,V)}{Z^r_{HV}}\mu\right) \\
&-(1-\delta_d)\left(\frac{\sigma_\tau(t\,V)\sigma_\xi(x)}{Z^r_{RV}}m + \frac{\sigma_\tau(t\,V)\sigma_\epsilon}{Z^r_{RV}}\mu + \frac{2\sigma_\epsilon\sigma_\xi(x)}{Z^r_{RV}}\mu\right) \\
&-\left(\frac{\sigma_\tau(0)\sigma_\xi(x)}{Z^h_{NV}}m + \frac{\sigma_\tau(0)\sigma_\epsilon}{Z^h_{NV}}\mu + \frac{2\sigma_\epsilon\sigma_\xi(x)}{Z^h_{NV}}\mu\right) \\
&+\left(\frac{\sigma_\tau(t)\sigma_\xi(x)}{Z^r_{NV}}m + \frac{\sigma_\tau(t)\sigma_\epsilon}{Z^r_{NV}}\mu + \frac{2\sigma_\epsilon\sigma_\xi(x)}{Z^r_{NV}}\mu\right).
\end{aligned}
\tag{27}
$$

Taking the derivative of equation 27 with respect to μ gives the weighted average of symmetric and asymmetric introductions of VA.

$$
\begin{aligned}
&\frac{\partial E[\delta_d(b^{h*}_{HV} - b^{r*}_{HV}) + (1-\delta_d)(b^{h*}_{NV} - b^{r*}_{RV}) - (b^{h*}_{NV} - b^{r*}_{NV})|m\,\mu]}{\partial\mu} = \\
&\delta_d\frac{2{\sigma_\epsilon}^2(\sigma_\tau(0)-\sigma_\tau(t))(\sigma_\xi(x)-\sigma_\xi(x\,V))}{Z^h_{NV}Z^r_{NV}Z^h_{HV}Z^r_{HV}}[\sigma_\xi(x)\sigma_\xi(x\,V)(2\sigma_\tau(t)\sigma_\tau(0) + \sigma_\epsilon\sigma_\tau(0) + \sigma_\tau(t)\sigma_\epsilon) \\
&+2(\sigma_\xi(x\,V)+\sigma_\xi(x))\sigma_\tau(t)\sigma_\epsilon\sigma_\tau(0))] + (1-\delta_d)\frac{\sigma_\epsilon\sigma_\xi(x)^2(\sigma_\tau(t\,V)-\sigma_\tau(t))}{Z^r_{NV}Z^r_{RV}}.
\end{aligned}
\tag{28}
$$

Equation 29 shows that taking the derivative of equation 28 with respect to δ_d demonstrates that as the share of informed principals increases the probability that good teachers increases as well.

$$\frac{\partial^2 E[\delta_d(b^{h*}_{HV} - b^{r*}_{HV}) + (1-\delta_d)(b^{h*}_{NV} - b^{r*}_{RV}) - (b^{h*}_{NV} - b^{r*}_{NV})|m\ \mu]}{\partial\mu\partial\delta_d} =$$

$$\frac{2{\sigma_\epsilon}^2(\sigma_\tau(0) - \sigma_\tau(t))(\sigma_\xi(x) - \sigma_\xi(x\ V))}{Z^h_{NV}Z^r_{NV}Z^h_{HV}Z^r_{HV}}[\sigma_\xi(x)\sigma_\xi(x\ V)(2\sigma_\tau(t)\sigma_\tau(0) + \sigma_\epsilon\sigma_\tau(0) + \sigma_\tau(t)\sigma_\epsilon) \tag{29}$$

$$+ 2(\sigma_\xi(x\ V) + \sigma_\xi(x))\sigma_\tau(t)\sigma_\epsilon\sigma_\tau(0))] - \frac{\sigma_\epsilon\sigma_\xi(x)^2(\sigma_\tau(t\ V) - \sigma_\tau(t))}{Z^r_{NV}Z^r_{RV}} > 0.$$

As noted previously, VA was little known when Guilford adopted their usage in 2000. If principals place no value on the measure, it is the same being uninformed of its content. Conversely, every out-of-district principal received an EVAAS VA of her school in 2008, when Winston-Salem began using EVAAS VA measures of teacher effectiveness. These different settings lead the share of out-of-district principals who are informed of VA to be higher for those leaving from Winston-Salem than for those moving from Guilford ($\delta_{WSF} > \delta_{GCS}$). Consequently, I expect the relationship between VA and the probability of moving from Winston-Salem to be more positive after Winston-Salem adopts VA than is the relationship between VA and the probability of moving from Guilford after Guilford adopts VA. Empirically, I expect $\gamma_{14OD_{GCS}} < \gamma_{14OD_{WSF}}$. The same logic can be applied to the fact that within Winston-Salem hiring principals did not directly receive teachers' VA whereas in Guilford they did. However, it is likely that principals still inferred something when a teacher chose not to reveal his VA. If the share of informed principals was lower within Winston-Salem than within Guilford ($\delta_{WSF} < \delta_{GCS}$), A safer prediction may be, $\gamma_{14WD_{GCS}} - \gamma_{14OD_{GCS}} > \gamma_{14WD_{WSF}} - \gamma_{14OD_{WSF}}$.

8.2.8 Comparative statics with respect to VA (V) and school quality (S)

It may not be realistic to suppose that all schools can bid for teachers in accordance with how the principal expects teachers to perform. Large differences in pay or school quality may be to great for a principal to overcome with position-specific, non-pecuniary benefits (J_{isd}). In this subsection I introduce a school-level, proportional constraint on principals bids ($\rho^s < 1$ where superscript $s = r, h$ indicates retaining and hiring principals) reflecting the costs to principals of providing these position-specific attributes. The key feature of ρ^s is that it is increasing in school quality (S^s) $\left[\frac{\partial\rho^s}{\partial S^s} > 0\right]$. In order to gain predictions regarding the probability of moving within-district in this framework, I take the cross partial of $E\left[\rho^h b^{h*}_{HV} - \rho^r b^{r*}_{HV} - (\rho^h b^{h*}_{NV} - \rho^r b^{r*}_{NV})|m\ V\ \mu\right]$ with respect to VA (V) and S^s. I present these cross partials below.

$$\frac{\partial E\left[\rho^h b^{h*}_{HV} - \rho^r b^{r*}_{HV} - (\rho^h b^{h*}_{NV} - \rho^r b^{r*}_{NV})|m\ V\ \mu\right]}{\partial V} = \rho^h \frac{\sigma_\tau(0)\sigma_\epsilon\sigma_\xi(x)}{Z^h_{HV}(\sigma_v + \sigma_\xi(x))} - \rho^r \frac{\sigma_\tau(0)\sigma_\epsilon\sigma_\xi(x)}{Z^r_{HV}(\sigma_v + \sigma_\xi(x))}$$

$$\frac{\partial^2 E\left[\rho^h b^{h*}_{HV} - \rho^r b^{r*}_{HV} - (\rho^h b^{h*}_{NV} - \rho^r b^{r*}_{NV})|m\, V\, \mu\right]}{\partial V \partial S^h} = \frac{\partial \rho^h}{\partial S^h}\frac{2\sigma_\tau(0)\sigma_\epsilon\sigma_\xi(x)}{Z^h_{HV}(\sigma_v + \sigma_\xi(x))} \tag{30}$$

As everything else is a function of variances, $\frac{\partial \rho^h}{\partial S^h} > 0$ implies that equation 30 is positive.

$$\frac{\partial^2 E\left[\rho^h b^{h*}_{HV} - \rho^r b^{r*}_{HV} - (\rho^h b^{h*}_{NV} - \rho^r b^{r*}_{NV})|m\, V\, \mu\right]}{\partial V \partial S^r} = -\frac{\partial \rho^r}{\partial S^r}\frac{2\sigma_\tau(0)\sigma_\epsilon\sigma_\xi(x)}{Z^h_{HV}(\sigma_v + \sigma_\xi(x))} \tag{31}$$

Conversely, $\frac{\partial \rho^r}{\partial S^r} > 0$ implies that equation 31 is negative. Thus, the probability of a move within district increases as the hiring school quality rises relative to the quality of the retaining school.

8.2.9 Comparative statics for within-district moves with respect to easily observable teacher characteristics (m)

The introduction of new information may also change the weighting principals formerly applied to easily observable teacher characteristics. Throughout the model m stands as summary measure of easily observable correlates with teacher effectiveness. I derive the predicted change in the relationship between a teacher's easily observable traits and the probability of moving within district with the introduction of VA, taking the derivative of $E\left[b^{h*}_{HV} - b^{r*}_{HV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\, \mu\right]$ shown in equation 22 with respect to (m).

$$\begin{aligned}\frac{\partial E\left[b^{h*}_{HV} - b^{r*}_{HV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\, \mu\right]}{\partial m} = {} & \frac{2\sigma_\epsilon(\sigma_\tau(0) - \sigma_\tau(t))(\sigma_\xi(x\, V) - \sigma_\xi(x))}{Z^r_{HV}Z^h_{HV}Z^r_{NV}Z^h_{NV}} \\ & [2\sigma_\tau(t)\sigma_\tau(0)\sigma_\xi(x)\sigma_\xi(x\, V) + 2\sigma_\xi(x\, V)\sigma_\epsilon\sigma_\xi(x)(\sigma_\tau(0) \\ & + \sigma_\tau(t)) + (\sigma_\xi(x\, V) + \sigma_\xi(x))\sigma_\tau(t)\sigma_\epsilon\sigma_\tau(0)].\end{aligned} \tag{32}$$

Under the assumptions of prior private learning ($\sigma_\tau(0) - \sigma_\tau(t) > 0$), and informative VA ($\sigma_\xi(x\, V) - \sigma_\xi(x) < 0$), equation 32 implies that $\frac{\partial E[b^{h*}_{HV} - b^{r*}_{HV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\, \mu]}{\partial m} < 0$. Thus, the model predicts the probability of moving after the introductions of VA decreases as a teacher's VA increases, or empirically, $\gamma_{24WD} < 0$.

8.2.10 Comparative statics for out-of-district moves with respect to easily observable teacher characteristics (m)

I derive the predicted change in the relationship between a teacher's easily observable traits and the probability of moving out-of district with the introduction of VA, taking the derivative of $E\left[b^{h*}_{HV} - b^{r*}_{HV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\,\mu\right]$ with respect to (m).

$$\frac{\partial E\left[b^{h*}_{RV} - b^{r*}_{RV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\,\mu\right]}{\partial m} = \frac{2\sigma_\xi(x)^2\sigma_\epsilon(\sigma_\tau(t) - \sigma_\tau(t\,V))}{Z^r_{RV}Z^r_{NV}} \tag{33}$$

Under the assumption that VA is informative to current principals ($\sigma_\tau(t) - \sigma_\tau(t\,V) > 0$), $\frac{\partial E[b^{h*}_{RV} - b^{r*}_{RV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\,\mu]}{\partial\mu} > 0$. This implies that the probability of out-of-district transitions increases with declines in teacher effectiveness .

8.2.11 Comparative statics for within-district moves with respect to ability (μ) and tenure (t)

In order to investigate the learning environment that prevailed in the absence of VA, I extend the model to provide differential predictions for workers who have been employed by the same school for a longer period of time or who are simply more experienced. In order to examine whether there was prior private learning, I take the cross partial of $E[b^{h*}_{HV} - b^{r*}_{HV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\,\mu]$ with respect to μ and t. Below is the derivative of equation 22 with respect to μ.

$$\begin{aligned}\frac{\partial E[b^{h*}_{HV} - b^{r*}_{HV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\,\mu]}{\partial\mu} &= \frac{\sigma_\tau(0)\sigma_\epsilon}{Z^h_{HV}} + \frac{2\sigma_\xi(x\,V)\sigma_\epsilon}{Z^h_{HV}} - \left(\frac{\sigma_\tau(t)\sigma_\epsilon}{Z^r_{HV}} + \frac{2\sigma_\epsilon\sigma_\xi(x\,V)}{Z^r_{HV}}\right) \\ &- \left[\frac{\sigma_\tau(0)\sigma_\epsilon}{Z^h_{NV}} + \frac{2\sigma_\xi(x)\sigma_\epsilon}{Z^h_{NV}} - \left(\frac{\sigma_\tau(t)\sigma_\epsilon}{Z^r_{NV}} + \frac{2\sigma_\epsilon\sigma_\xi(x)}{Z^r_{NV}}\right)\right]\end{aligned} \tag{34}$$

Taking the derivative of equation 34 with respect to t gives the following:

$$\begin{aligned}\frac{\partial^2 E[b^{h*}_{HV} - b^{r*}_{HV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\,\mu]}{\partial\mu\partial t} &= \frac{\partial\sigma_\tau(t)}{\partial t}\frac{\sigma_\epsilon Z^r_{NV} - 2(\sigma_\tau(t)\sigma_\epsilon + \sigma_\epsilon\sigma_\xi(x))(\sigma_\epsilon + \sigma_\xi(x))}{{Z^r_{NV}}^2} \\ &- \frac{\partial\sigma_\tau(t)}{\partial t}\frac{\sigma_\epsilon Z^r_{HV} - 2(\sigma_\tau(t)\sigma_\epsilon + \sigma_\epsilon\sigma_\xi(x\,V))(\sigma_\epsilon + \sigma_\xi(x\,V))}{{Z^r_{HV}}^2} \\ &= -2\frac{\partial\sigma_\tau(t)}{\partial t}\frac{\sigma_\xi(x)^2\sigma_\epsilon{Z^r_{HV}}^2 - \sigma_\xi(x\,V)^2\sigma_\epsilon{Z^r_{NV}}^2}{{Z^r_{HV}}^2{Z^r_{NV}}^2} \\ &= \frac{\partial\sigma_\tau(t)}{\partial t}\frac{2\sigma_\epsilon\sigma_\tau(t)(\sigma_\xi(x\,V) - \sigma_\xi(x))}{{Z^r_{HV}}^2{Z^r_{NV}}^2} \\ &2\sigma_\xi(x)\sigma_\xi(x\,V)(2\sigma_\epsilon + \sigma_\tau(t)) + \sigma_\tau(t)\sigma_\epsilon(\sigma_\xi(x) + \sigma_\xi(x\,V))\end{aligned} \tag{35}$$

The assumptions of prior private learning $\left(\frac{\partial \sigma_\tau(t)}{\partial t} < 0\right)$ and informative VA $(\sigma_\xi(x\ V) < \sigma_\xi(x))$, imply that equation 35 is positive. Thus, the positive change in selection with the introduction of VA should be more positive for those with more tenure. Empirically, the model predicts the coefficient on the interaction between adopting VA, the VA measures, and tenure to be positive $(VA \times Ten \times TreatDist > 0)$.

8.2.12 Comparative statics for within-district moves with respect to VA (V) and tenure (t)

In order to investigate the learning environment that prevailed in the absence of VA, I extend the model to provide differential predictions for workers who have been employed by the same school for a longer period of time or who are simply more experienced. In order to examine whether there was prior private learning, I take the cross partial of $E[b^{h*}_{HV} - b^{r*}_{HV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ V\ \mu]$ with respect to VA (V) and years of tenure (t).

$$\frac{\partial^2 E\left[b^{h*}_{HV} - b^{r*}_{HV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ V\ \mu\right]}{\partial V \partial t} = -\frac{\partial \sigma_\tau(t)}{\partial t}\frac{\sigma_\xi(x\ V) + \sigma_\epsilon}{Z^h_{HV}(Z^r_{HV})^2}\frac{2\sigma_\xi(x)}{\sigma_\nu + \sigma_\xi(x)} > 0 \qquad (36)$$

The assumption of prior private learning provides $(\frac{\partial \sigma_\tau(t)}{\partial t} < 0)$ leads equation 36 to be positive. This means that the model predicts larger positive effects of the introduction of VA on the probability that high-VA teachers move, when those teachers have more tenure, all else equal. Empirically, this means the model predicts that the coefficient on the triple interaction of $VA \times TreatDist \times tenure$ to be positive.

8.3 Robustness: Mobility based on ABC Growth Policies

In the 1996/1997 school year the state of North Carolina began rewarding teachers who worked in schools in which the students made substantial growth. The state awarded bonuses of either \$750 or \$1,500 based on whether the school achieved growth in student test scores beyond predetermined tiered thresholds. These bonuses were given to all teachers in qualifying schools. For additional detail about the policy please see Vigdor (2008) and Ahn and Vigdor (2012).

As a result, teaching in high growth schools may be additionally attractive to teachers since the bonuses depended upon school performance. Table 10 is comparable to Table 3 except that the dependent variable here is whether the teacher moves to higher (lower) growth school as opposed to a higher (lower) performing school within and out of district. The total within and out-of districts mobility estimates in columns 1 and 4 of Table 3 are unaffected, and so they are omitted.

When examining this alternate school attribute on which teachers may sort, the primary findings remain intact. The within district mobility is driven by moves to more favorable schools for both districts. Though the results are attenuated here as a teacher with a full standard deviation higher VA is 0.3 percentage point more likely to move within district to a higher ABC growth school for teachers whose VA are released, the estimates remain statistically significantly positive for both districts. Though these estimates are not statistically different from the estimated effect on the probability of moving to higher performing schools, perhaps they suggest that school performance may be a stronger motivator for teacher mobility than student growth.

The estimated effects for moves outside the district are remarkably close between Table 3 and Table 10. The adverse selection of movers out of Guilford County Schools holds for moves to both better and worse schools, while moves from Winston-Salem to better schools remain unrelated to teachers' VA after the policy takes effect.

Table 10: Probability of moving to higher or lower growth schools

	Panel A: Within-District Moves		Panel B: Out-Of-District Moves	
VARIABLES	To a higher ABC growth school	To a lower ABC growth school	To a higher ABC growth school	To a lower ABC growth school
VA	0.0024	-0.0006	0.0008	-0.0005
	(0.00073)	(0.00077)	(0.00056)	(0.0006)
VA x Treatment GCS	0.0031	0.0013	-0.0048	-0.0052
	(0.00152)	(0.00153)	(0.00139)	(0.002)
VA x Treatment WSF	0.003	0.0017	0	0.0014
	(0.0015)	(0.00155)	(0.00131)	(0.001)
Treatment GCS	0.0074	-0.0023	0.0057	-0.0129
	(0.00385)	(0.00612)	(0.00187)	(0.00219)
Treatment WSF	0.0156	0.0074	-0.001	-0.0093
	(0.00206)	(0.00297)	(0.00126)	(0.00209)
Observations	236,018	236,018	236,018	236,018

CSB standard errors from 500 repetitions appear in brackets.

All regressions include teacher level covariates and interactions with treatment indicators.

8.4 Normal Maximum Likelihood Estimation

The results in Table 3 are from a linear probability model, which are more straight forward both computationally and in interpretation. Taking the normality and orthogonality assumptions from Section 3 seriously would suggest normal Maximum Likelihood Estimation (probit estimation). As noted in Ai and Norton (2003), the functional form of probit estimation incorporates an interaction term, even when one is not specifically modeled. As a result, if the researcher is interested in estimating the average partial effect (APE) of an interaction additionally programming is necessary. Table 11 in Appendix 8.6 provides the APEs in accordance with Ai and Norton (2003). Comparison between Table 3 and Table 11 provides very similar results.

Table 11: Probability of moving schools using normal maximum likelihood estimation.

	Panel A: Within-District Moves			Panel B: Out-of-District Moves		
VARIABLES	Total	To a higher performing school	To a lower performing school	Total	To a higher performing school	To a lower performing school
VA	0.0022	0.0030	-0.0011	-0.0011	0.0005	-0.0018
	(0.00114)	(0.00079)	(0.00068)	(0.00083)	(0.0006)	(0.0005)
VA x Treatment GCS	0.0046	0.0040	0.0021	-0.0117	-0.0065	-0.0053
	(0.0025)	(0.00172)	(0.00185)	(0.00274)	(0.00203)	(0.0017)
VA x Treatment WSF	0.0029	0.0038	-0.0010	0.0002	0.0026	-0.0020
	(0.00268)	(0.00193)	(0.00221)	(0.00313)	(0.00238)	(0.00324)
Treatment GCS	0.0110	0.0112	0.0001	-0.0009	-0.0036	0.0027
	(0.00268)	(0.0019)	(0.00177)	(0.0019)	(0.00161)	(0.00101)
Treatment WSF	-0.0149	-0.0103	-0.0080	0.0022	-0.0011	-0.0226
	(0.00441)	(0.00369)	(0.0031)	(0.00493)	(0.00342)	(0.00679)
Observations	236,018	236,018	236,018	236,018	236,018	236,018

CSB standard errors from 500 repetitions appear in brackets.
All regressions include teacher level covariates and interactions with treatment indicators.

8.5 Competing Risks Analysis

By performing separate regressions for each type of school transfer, the above analysis treats each type of move as independent of the others. However, it is possible that the propensity of a teacher to move within-district to a higher-performing school is related to the propensity of moving to a higher-performing school in another district. The same could be said with any combination of outcomes. To test the sensitivity of my earlier results to these possibilities, I adopt a competing risks approach, as proposed by Fine and Gray (1999).

Competing risks survival analysis models the subdistribution hazard ($\lambda_E(t)$) of a particular type of event, such as a move within a school district ($E = WD$), as a function of an unspecified baseline hazard ($\lambda_{E0}(t)$), as well as a vector of time-varying covariates ($\mathbf{Z}(t)$). [47]

$$\lambda_{WD}(t|\mathbf{Z}) = \lambda_{WD0}(t)exp\{\mathbf{Z}(t)\boldsymbol{\beta}_0\}, \tag{37}$$

In the context of this study, time at risk (t) is defined as the difference between the current year and the year at which the teacher first appears matched with the current school.[48] $\mathbf{Z}(t)$ is a vector including all covariates used in Table 3, with the exception of tenure, which is perfectly correlated with t. I additionally include district averages of all within-district-varying covariates to control for unobserved, district-wide effects, as in Mundlak (1978)[49].

Table 12 reports the coefficient estimates for each type of transfer between schools. Accordingly, $\beta \times 100$ may be interpreted as the percent change in the marginal probability of a particular type of mobility due to a one unit change in the covariate. Columns 1 and 4, examine transfers within and out of the district respectively, with the other broad type of transfer serving as a competing risk. Columns 2, 3, 5, and 6, examine transfers to higher and lower-performing schools, within and out of the district, with the other types of transfers serving as competing risks.

In this framework, results remain largely consistent. From columns 1 and 2, the probability of moving within Guilford for a teacher with a one standard deviation higher VA score increases by 9% with the release of teacher VA, and for moves within-district to better schools, the probability increases by 13%. Both effects are significantly different from zero and are within a percentage point estimates shown in Table 3. For moves within Winston Salem, the results are somewhat more sensitive. Using competing risks analysis drops the point estimate of the effect of the policy by teacher VA on within district moves by half and the estimate loses significance. The point estimate on moves to a higher-performing school is more stable staying between 10-15%, though the significance level drops with this specification to a p-value of 0.106. From columns 4 and 5, a teacher with a one standard deviation lower VA becomes 33.6% (29.5%) more likely to move out of Guilford (to a higher-performing school) after the policy takes effect. In Winston-Salem, there remain no statistically significant effects of the policy on which teachers move. In general, the public and private learning

[47] Gray (1988) defines the subdistribution hazard as, $\lambda_{WD}(t) = lim_{\Delta t \to 0} \frac{P(t<T\leq t+\Delta t, E=WD | t\leq T \bigcup t<T, E\neq WD)}{\Delta t}$, where T is the timing of the event occurrence of which there are different types.

[48] I use teacher to school matches as the basis of this survival analysis. Though this forces me to assume independence of matches, it allows me to retain the original sample making it easier to compare the results.

[49] Unreported regression results show little difference depending on whether or not district averages are included

results are further verified in Guilford with this competing risks analysis, and while the point estimates in Winston Salem are noisier, I believe they are sufficiently stable to avoid concern.

Table 12: Changes in the marginal probability of each type of transfer between schools

	Panel A: Within-District Moves			Panel B: Out-Of-District Moves		
VARIABLES	Total	To a higher performing school	To a lower performing school	Total	To a higher performing school	To a lower performing school
VA	0.03	0.09	-0.07	0.01	0.08	-0.10
	(0.021)	(0.024)	(0.030)	(0.028)	(0.035)	(0.042)
VA x Treatment GCS	0.09	0.13	0.10	-0.41	-0.35	-0.40
	(0.045)	(0.051)	(0.076)	(0.104)	(0.111)	(0.164)
VA x Treatment WSF	0.04	0.11	-0.08	0.02	0.15	-0.21
	(0.050)	(0.068)	(0.095)	(0.116)	(0.141)	(0.238)
Treatment GCS	0.01	0.22	-0.23	0.24	-0.12	0.49
	(0.116)	(0.107)	(0.113)	(0.122)	(0.130)	(0.160)
Treatment WSF	0.56	0.27	0.87	-0.87	0.18	-7.22
	(0.118)	(0.145)	(0.144)	(0.167)	(0.219)	(0.587)
Observations	236,018	236,018	236,018	236,018	236,018	236,018

CSB standard errors from 500 repetitions appear in brackets.

All regressions include teacher level covariates and interactions with treatment indicators.

8.6 Appendices Tables

Table A1: Probability of moving schools using alternate standard errors

	Within-District Moves			**Out-of-District Moves**		
	Total	**To higher performing schools**	**To lower performing schools**	**Total**	**To higher performing schools**	**To lower performing schools**
VA	**0.0016**	**0.0032**	**-0.0016**	**0.0002**	**0.0014**	**-0.0012**
	[0.00139]	[0.00091]	[0.00083]	[0.00084]	[0.00057]	[0.00050)
	{0.00056}	{0.0004}	{0.00036}	{0.00039}	{0.00031}	{0.00022}
	(0.00129)	(0.00091)	(0.00074)	(0.00096)	(0.00072)	(0.00058)
VA x Treatment GCS	**0.0058**	**0.0051**	**0.0007**	**-0.0103**	**-0.0054**	**-0.0049**
	[0.00168]	[0.00115]	[0.00091]	[0.00090]	[0.00061]	[0.00057)
	{0.00262}	{0.00204}	{0.00153}	{0.00192}	{0.00164}	{0.00106}
	(0.00265)	(0.00199)	(0.00151)	(0.00261)	(0.00195)	(0.00156)
VA x Treatment WSF	**0.0052**	**0.006**	**-0.0008**	**0.0009**	**0.0023**	**-0.0014**
	[0.00147]	[0.00094]	[0.00125]	[0.00084]	[0.00068]	[0.00051)
	{0.00323}	{0.00255}	{0.00204}	{0.00186}	{0.00167}	{0.00096}
	(0.00286)	(0.00229)	(0.00194)	(0.00241)	(0.00208)	(0.00129)
Treatment GCS	**-0.004**	**-0.005**	**0.001**	**-0.0162**	**-0.0232**	**0.007**
	[0.00829]	[0.00608]	[0.00537]	[0.00402]	[0.00319]	[0.00214)
	{0.00583}	{0.00436}	{0.00444}	{0.00261}	{0.00114}	{0.0024}
	(0.00851)	(0.00571)	(0.00679)	(0.00374)	(0.00233)	(0.00268)
Treatment WSF	**0.0555**	**0.0475**	**0.008**	**-0.002**	**0.0147**	**-0.0167**
	[0.00579]	[0.00417]	[0.00311]	[0.00258]	[0.00199]	[0.00184)
	{0.00314}	{0.00253}	{0.00215}	{0.0029}	{0.0022}	{0.00171}
	(0.00499)	(0.00372)	(0.00299)	(0.00274)	(0.00224)	(0.00178)
Observations	236,018	236,018	236,018	236,018	236,018	236,018

Clustered standard errors in brackets. Bootstrapped standard errors in braces. District-cluster-bootstrapped-teacher-stratified standard errors in parentheses.

Table A2: The effects of VA on the probability of moving schools within-district by year.

	Total			To a more proficient school		
VARIABLES	Rest of NC	Guilford	Winston-Salem	Rest of NC	Guilford	Winston-Salem
year 1998 x VA	0.0009	0.0012	0.0043	0.0021	0.0006	-0.0003
	(0.00077)	(0.00269)	(0.00513)	(0.00061)	(0.00236)	(0.00267)
year 1999 x VA	0.0022	0.0023	-0.0001	0.0044	0.0048	0.0041
	(0.00083)	(0.00316)	(0.00587)	(0.00059)	(0.00242)	(0.00393)
year 2000 x VA	0.0035	0.0205	-0.0007	0.0023	0.0155	-0.0042
	(0.00079)	(0.00252)	(0.00311)	(0.00065)	(0.00156)	(0.00253)
year 2001 x VA	0.0019	0.0048	-0.0020	0.0035	0.0030	0.0012
	(0.00079)	(0.00332)	(0.00298)	(0.00058)	(0.00262)	(0.00211)
year 2002 x VA	0.0035	-0.0044	0.0024	0.0055	-0.0011	0.0107
	(0.00096)	(0.00268)	(0.00535)	(0.00073)	(0.00205)	(0.00378)
year 2003 x VA	0.0004	-0.0054	0.0041	0.0027	-0.0013	0.0042
	(0.00089)	(0.00467)	(0.00486)	(0.00073)	(0.00329)	(0.00445)
year 2004 x VA	0.0010	0.0020	-0.0088	0.0016	-0.0073	-0.0043
	(0.00106)	(0.00446)	(0.00403)	(0.0008)	(0.00296)	(0.00358)
year 2005 x VA	0.0015	0.0128	-0.0160	0.0040	0.0190	-0.0080
	(0.00099)	(0.00300)	(0.00423)	(0.00075)	(0.00273)	(0.00297)
year 2006 x VA	0.0047	0.0169	0.0100	0.0055	0.0158	0.0037
	(0.00087)	(0.00563)	(0.00308)	(0.00061)	(0.00521)	(0.00193)
year 2007 x VA	0.0027	0.0189	-0.0133	0.0039	0.0147	-0.0078
	(0.00081)	(0.00355)	(0.00478)	(0.00056)	(0.00282)	(0.00366)
year 2008 x VA	0.0029	0.0057	0.0005	0.0032	0.0114	0.0019
	(0.00092)	(0.00342)	(0.00469)	(0.00069)	(0.00247)	(0.00370)
year 2009 x VA	0.0034	0.0036	0.0110	0.0032	0.0046	0.0173
	(0.00118)	(0.00325)	(0.00579)	(0.00091)	(0.00233)	(0.00473)
year 2010 x VA	-0.0001	0.0123	0.0002	0.0009	0.0121	0.0004
	(0.00095)	(0.00326)	(0.00489)	(0.00073)	(0.00274)	(0.00431)
Observations	216,484	11,239	8,295	216,484	11,239	8,295

Standard errors are bootstrapped at the student-year level and appear in brackets.
All regressions include teacher level covariates and interactions with year indicators.

Table A3: The effect of VA on the probability of moving schools out-of-district by year.

	Total			To a more proficient school		
VARIABLES	Rest of NC	Guilford	Winston-Salem	Rest of NC	Guilford	Winston-Salem
year 1998 x VA	0.0017	0.0098	-0.0079	0.0023	0.0076	-0.0059
	(0.0005)	(0.00212)	(0.0032)	(0.00039)	(0.00178)	(0.00187)
year 1999 x VA	-0.0004	0.0065	-0.0026	0.0011	0.0064	-0.0033
	(0.00057)	(0.00267)	(0.00136)	(0.00049)	(0.00243)	(0.00096)
year 2000 x VA	0.0006	0.0013	0.0063	0.0015	0.0033	0.0033
	(0.00057)	(0.00157)	(0.00215)	(0.00045)	(0.00126)	(0.00195)
year 2001 x VA	-0.0022	0.0025	-0.0069	-0.0005	0.0063	-0.0070
	(0.00057)	(0.00152)	(0.00202)	(0.00044)	(0.00112)	(0.00163)
year 2002 x VA	-0.0033	-0.0025	0.0106	0.0000	0.0015	0.0146
	(0.00063)	(0.00261)	(0.00203)	(0.00042)	(0.00167)	(0.00187)
year 2003 x VA	-0.0011	-0.0016	-0.0141	0.0017	-0.0004	-0.0091
	(0.00071)	(0.00282)	(0.00367)	(0.00052)	(0.0028)	(0.00346)
year 2004 x VA	-0.0037	0.0099	0.0054	-0.0005	0.0080	0.0092
	(0.00073)	(0.00206)	(0.0034)	(0.00056)	(0.00172)	(0.00281)
year 2005 x VA	-0.0001	-0.0038	-0.0024	0.0011	0.0033	-0.0005
	(0.00064)	(0.00197)	(0.00212)	(0.00047)	(0.00164)	(0.00176)
year 2006 x VA	-0.0011	-0.0095	-0.0001	0.0017	-0.0018	-0.0013
	(0.00071)	(0.00372)	(0.003)	(0.00048)	(0.00262)	(0.00276)
year 2007 x VA	-0.0016	-0.0223	0.0011	0.0003	-0.0040	0.0063
	(0.00081)	(0.00367)	(0.00358)	(0.00061)	(0.00114)	(0.00352)
year 2008 x VA	-0.0017	-0.0079	-0.0054	0.0006	0.0001	-0.0000
	(0.00064)	(0.00185)	(0.00414)	(0.00047)	(0.00099)	(0.0035)
year 2009 x VA	0.0006	-0.0023	0.0047	-0.0004	0.0000	0.0047
	(0.00051)	(0.00089)	(0.00149)	(0.00035)	(0.00012)	(0.00148)
year 2010 x VA	-0.0021	-0.0058	-0.0011	-0.0006	-0.0054	-0.0011
	(0.00058)	(0.00156)	(0.00113)	(0.00051)	(0.00103)	(0.00112)
Observations	216,484	11,239	8,295	216,484	11,239	8,295

Standard errors are bootstrapped at the student-year level and appear in brackets.
All regressions include teacher level covariates and interactions with year indicators.

Table A4: The effect of VA on teacher sorting within-district by year.

VARIABLES	Rest of NC	Guilford	Winston-Salem
year 1998 x VA	0.0025	0.0045	-0.0014
	(0.00021)	(0.00071)	(0.00146)
year 1999 x VA	0.0026	0.0013	0.0021
	(0.00021)	(0.00109)	(0.00156)
year 2000 x VA	0.0019	0.0041	0.0007
	(0.0002)	(0.00069)	(0.00084)
year 2001 x VA	0.0051	0.0038	0.0077
	(0.00026)	(0.00097)	(0.00146)
year 2002 x VA	0.0046	0.0031	0.0072
	(0.0002)	(0.00072)	(0.00164)
year 2003 x VA	0.0031	0.0043	0.0052
	(0.00019)	(0.00099)	(0.001)
year 2004 x VA	0.0023	-0.0006	0.0005
	(0.00021)	(0.00109)	(0.00212)
year 2005 x VA	0.0102	0.0109	0.0096
	(0.00032)	(0.00097)	(0.00126)
year 2006 x VA	0.0047	0.0009	-0.0014
	(0.00027)	(0.00161)	(0.00089)
year 2007 x VA	0.0046	0.0049	0.0031
	(0.00026)	(0.00105)	(0.00133)
year 2008 x VA	0.0016	0.0031	0.0005
	(0.00025)	(0.00112)	(0.00127)
year 2009 x VA	-0.0003	0.0055	0.0053
	(0.00042)	(0.00097)	(0.00146)
year 2010 x VA	0.0033	0.0050	0.0045
	(0.00027)	(0.00104)	(0.00145)
Observations	185,977	9,616	7,35

Standard errors are bootstrapped at the student-year level and appear in brackets. All regressions include teacher level covariates and interactions with treatment indicators.

Table A5: Probability of moving schools within-district using restricted data VA

Panel	A: Within-District Moves			B: Out-Of-District Moves			C: School Quality Growth	
VARIABLES	Total	To a higher performing school	To a lower performing school	Total	To a higher performing school	To a lower performing school	Total	Within District
VA	0.0003	0.0011	-0.0008	-0.0013	-0.0006	-0.0007	0.0005	0.0004
	(0.00109)	(0.00097)	(0.00063)	(0.00079)	(0.00056)	(0.00043)	(0.00032)	(0.00033)
VA x Treatment GCS	0.0034	0.0030	0.0004	-0.0027	-0.0016	-0.0011	-0.0015	-0.0010
	(0.00249)	(0.002)	(0.00152)	(0.00201)	(0.00167)	(0.00102)	(0.00083)	(0.00076)
VA x Treatment WSF	0.0061	0.0099	-0.0038	0.0019	0.0025	-0.0005	0.0025	0.0037
	(0.00312)	(0.00241)	(0.00216)	(0.00247)	(0.00224)	(0.00122)	(0.00131)	(0.00109)
Treatment GCS	-0.0034	-0.0042	0.0008	-0.0137	-0.0220	0.0082	-0.0196	-0.0156
	(0.00848)	(0.00545)	(0.00717)	(0.00365)	(0.00243)	(0.00275)	(0.0022)	(0.00225)
Treatment WSF	0.0555	0.0486	0.0068	-0.0017	0.0151	-0.0168	0.0299	0.0241
	(0.00533)	(0.00386)	(0.0033)	(0.00283)	(0.00217)	(0.0019)	(0.00165)	(0.00165)
Observations	236,018	236,018	236,018	236,018	236,018	236,018	209,424	202,943

CSB standard errors from 500 repetitions appear in brackets.
All regressions include teacher level covariates and interactions with treatment indicators.

Table A6: Mobility between non-strategic-staffing schools with respect to students' race

	Panel A: Within-District Moves			**Panel B: Out-Of-District Moves**		
VARIABLES	Total	To a higher performing school	To a lower performing school	Total	To a higher performing school	To a lower performing school
VA	-0.0015	0.0000	-0.0015	-0.0021	-0.0011	-0.0010
	(0.00042)	(0.00032)	(0.00027)	(0.0003)	(0.00023)	(0.00019)
VA x Treatment GCS	0.0035	0.0037	-0.0001	-0.0063	-0.0041	-0.0023
	(0.00206)	(0.00162)	(0.00121)	(0.00148)	(0.00111)	(0.00097)
VA x Treatment WSF	0.0090	0.0129	-0.0039	0.002	0.0019	0.0001
	(0.00276)	(0.00216)	(0.00166)	(0.00166)	(0.00143)	(0.00084)
Treatment GCS	-0.0032	-0.0040	0.0008	-0.0162	-0.0239	0.0077
	(0.00408)	(0.00109)	(0.00409)	(0.00121)	(0.00098)	(0.00064)
Treatment WSF	0.0555	0.0476	0.0078	-0.0021	0.0147	-0.0167
	(0.00232)	(0.00173)	(0.00162)	(0.00194)	(0.00193)	(0.00028)
Observations	236,018	236,018	236,018	236,018	236,018	236,018

CSB standard errors from 500 repetitions appear in brackets.
All regressions include teacher level covariates and interactions with treatment indicators.