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

Michael Bates*

October 19, 2015

Abstract

While a large literature focuses on informational asymmetries between workers and employers, more recent studies focus on asymmetric information between current and prospective employers. Despite the intuitive appeal of the theory, there is little direct, empirical evidence that current employers benefit from an informational advantage. I adapt models of public and private employer learning to the market for teachers. I then use statewide, micro-level, administrative data from North Carolina to formulate value-added measures (VAMs) of teacher productivity. I exploit the adoption of VAMs of teacher performance by two of the largest school districts in the state, a shock to the available information for some, but not all, employers, to provide an initial direct test of asymmetric employer learning. Consistent with a shock to public information, for job moves within the district, I find that the adoption of value-added measures increases the probability that high-VAM teachers move to higher-performing schools. For moves out of the district, I find that the impacts of policy are mitigated and even reversed by teachers with lower value-added measures becoming more likely to move to higher-performing schools. This adverse selection to plausibly less informed principals is consistent with asymmetric employer learning. Further, I find evidence that these moves lead to an increase in the sorting of teachers across schools within district, exacerbating the inequality in access to high quality teaching.

*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 Annual Meeting of the Midwest Economics Association, the 2014 UM-MSU-UWO Labor Day Conference, the Causal Inference in Education Research at the University of Michigan, the Applied Economics Seminar at University of California, Riverside, and the Applied Microeconomics Seminar at Michigan State University 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. This research was supported by the Institute of Education Sciences Grant R305B090011 to Michigan State University. The opinions expressed are those of the author and do not represent the views of the U.S. Department of Education. All errors are my own.

1 Introduction

Gaps in information hinder the efficient 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, a more recent literature focuses on asymmetric information between current and prospective employers. Empirical work uses these 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 monoposonist 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, there is little direct evidence of asymmetric employer learning. This is in part due to the absence of direct measures of productivity, and more importantly, due to a lack of exogenous variation in the informational landscape in which employers operate.

In this paper, I adapt models of public and private employer learning to the market for middle and elementary school teachers. I then use statewide, micro-level, administrative data from North Carolina to formulate value-added measures (VAMs) of teacher productivity.[1] Lastly, I exploit the adoption of teacher VAMs by two of the largest school districts in the state, a shock to the available information for some, but not all, potential employers, to provide an initial direct test of asymmetric employer learning.

The adoption of VAMs in North Carolina provides a rich context for examining employer

[1] VAMs calculate how much a teachers' students learn in comparison to how much those students are expected to learn. There are several methods for estimating VAMs. In econometric terms, I estimate teacher fixed effects in the regression of student test scores on student covariates including past test scores.

learning. Each of the two large districts that adopted VAMs did so in different ways and separately from the rest of the state. This provides three different 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 VAMs; 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.

This study examines how the relationship between teacher quality and the probability of moving schools changes with the adoption of VAMs of teacher effectiveness. If VAMs are informative, they provide teachers with a public signal of their ability. Thus, the model predicts that VAMs increase the likelihood that effective teachers move from one school to another within the district. If the information spreads easily through the market there should be no difference between the impacts of VAMs for moves within-district and out of Guilford and Winston-Salem. However, if retaining principals keep teachers' VAMs 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.

Using differences-in-differences analysis, I find that by releasing VAMs to teachers and principals, both districts increase the probability that high-VAM teachers will move within district to a higher-performing school. I estimate that the release of VAMs increases the probability that a teacher with a one standard deviation higher VAM moves within-district to higher-performing schools by about 10%. I find that the selection of mobile teachers becomes more negative for teachers moving to another school outside of Guilford and Winston-Salem after they adopt VAMs. The policy leads teachers who are a full standard deviation below average to become 15% 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-VAM 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 VAMs is consistent with asymmetric employer learning.

In the primary education context, questions of efficiency and equity are of particular importance. Previous research finds wide variation in the quality of teachers [Rivkin et al., 2005, Chetty et al., 2011, 2014]. Yet, at the point of hire, detecting good teachers is difficult, since easily observable teacher characteristics, such as educational attainment and college selectivity, are not highly correlated with teacher effectiveness [Rivkin et al., 2005, Staiger and Rockoff, 2010]. Informational gaps may lead schools and districts to hire relatively ineffective teachers, while passing on more capable ones. Thus, asymmetric information can have significant ramifications for the students they serve [Chetty et al., 2011, 2014].

After the date of hire, while principals typically do not observe a direct measure of a teachers' effectiveness, they can observe their teachers in action and inspect student outcomes. However, the quality of a teacher may remain difficult for the employing school to uncover, and harder still for other schools to learn. The amount of uncertainty in the market, and with whom the uncertainty lies, can differentially affect not only the initial sorting, but also the resorting of teachers across schools.

Persistent informational gaps between teachers' true effectiveness and employers' perceptions of it may lead schools to undervalue effective teachers and allow ineffective teachers to impede the progress of their pupils. In contrast, complete and public information allows better teachers more choice over where to teach. When teachers are given VAM reports, the VAMs provide them a new credible way to signal their ability.

In the teacher labor market, wages are typically set rigidly and are not tied to performance.[2] Thus, the implications of employer learning are felt primarily through teacher

[2] There are exceptions to this. In Section 7, I discuss two policies (ABC growth and Strategic Staffing) that deviate from this standard wage rigidity. The ABC growth program provides incentives to every teacher in schools that make their growth targets. Strategic staffing policies offer incentives to teach at hard-to-staff schools.

mobility from one school to another. There is a large body of work, which examines teacher preferences [Boyd et al., 2008, Jackson, 2009, Boyd et al., 2013]. They find that teachers in general 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. Consequently, while providing good teachers more choice, better information may also exacerbate the divide in access to high quality education. The degree to which information stays exclusively with current principals may mitigate these effects. This work provides the first examination of whether the release of VAMs leads to further sorting of teachers to schools. Rising inequity may be an important consequence of the policy that has been previously overlooked.

The possibility of growing inequity in access to effective teaching is particularly important given the speed at which states and school districts are adopting VAMs. The entire state of North Carolina adopted teacher-level VAMs in the 2013 school year. As of May, 2014, 38 states have required teacher evaluations to incorporate teachers' impacts on student achievement on standardized exams. Even among the remaining states, many large school districts have already incorporated VAMs into evaluations of their teachers. While these policies have been controversial, the debate has previously ignored the signaling impact of VAMs on the distribution of effective teachers across schools. By examining changes in the sorting of teachers, I evaluate the impact of the information on the distribution of teacher quality across schools. The rising mobility of effective teachers to high-performing schools and the rise in the correlation between teacher VAMs and school-wide student performance in Winston-Salem in particular, evidences rising inequity in access to high quality education as a result of VAM adoption.

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. 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 VAMs 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 value-added measures (VAMs). Consequently for moves within Guilford, the introduction of VAMs 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 principals directly received the VAM reports. The VAMs were not directly given to principals at other schools in the district.

However, the introduction of VAMs 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 VAM 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 VAMs 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 VAMs 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, VAMs 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 VAMs. When Guilford contracted with SAS, it was just two years after the creation of the EVAAS system, and two years before before the passage of No Child Left Behind. VAMs 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.

To summarize the basic intuition of the model in Section 4, if VAMs provide meaningful information to all principals in the district, and teachers in general prefer to teach at better schools, when districts release VAMs, good teachers will more likely move to higher-performing schools. It is also possible that current principals become less able to keep quiet which teachers are really good, while passing off the worse teachers to unwitting employers. Table 1 shows exactly this general pattern for moves within Guilford and Winston-Salem. In both districts, the average VAM of teachers who move within the district increases sharply after releasing VAMs. For moves out of these districts, the average VAM of moving teachers

Table 1: Average VAM of Teachers moving within and out of Winston-Salem and Guilford

		Panel A: Within District Movers			Panel B: Out of District Movers		
		1998-1999	2000-2007	2008-2010	1998-1999	2000-2007	2008-2010
Guilford	Mean VAM	-0.166	0.093	0.246	0.116	-0.174	-0.125
	N	101	463	104	48	206	34
Winston-Salem	Mean VAM	0.009	-0.088	0.031	-0.528	-0.100	-0.243
	N	188	275	63	26	121	21
Rest of State	Mean VAM	-0.069	0.020	0.052	-0.116	-0.118	-0.109
	N	1882	6793	1966	962	4230	833

Note: VAMs are measured in standard deviations. Guilford first adopted VAMs in 2000. Winston-Salem first adopted VAMs in 2008.

drops following the adoption of the policy. These means are not conditional on any easily observable characteristics, and so it is difficult to say whether the changes in information are driving these patterns. However, the increases of 0.259 and 0.119 standard deviations of average VAMs of movers within Guilford and Winston-Salem respectively suggests that the releasing VAMs within the district allows high-VAM teachers to move to other schools. The 0.290 and 0.143 drop in average VAMs of moving out of Guilford and Winston-Salem is indicative of low-VAM teachers moving to plausibly less informed principals outside of the district.

3 Employer Learning, VAMs, and Teacher Mobility

This is the first study directly testing a general model of public and private learning by exploiting information shocks to a large, relevant labor market. However, there is a robust extant literature building models of employer learning and fitting them to stylized empirical facts.

Farber and Gibbons [1996] provides the seminal model and test for employer learning. They assume that employers cannot directly observe the ability of potential workers and must rely on correlates to infer workers' expected value to the firm. They treat a subset of worker characteristics as easily observable to all, another as easily observable to the market (and not to researchers), and yet another subset of potential correlates with productivity as easily observable to the econometricians (but not the market). They (and many after them) uses the percentile from a cognitive ability assessment, the Armed Forces Qualification Test (AFQT), from the National Longitudinal Survey of Youth of 1979 (NLSY79), as this relatively strong correlate with productivity that is veiled to the the market at the time of hire, but is visible to researchers. By assuming a competitive marketplace and that employers all learn at the same rate in the Farber and Gibbons [1996] model, wages track the employers' learning process. Altonji and Pierret [2001] adopt a similar foundation in their examination of statistical discrimination as does Lange [2007] in his study of the speed at which employers learn. Each finds that the correlation between wages and AFQT score increases with experience, while the correlation between wages and easily observable characteristics falls over time.

Recent work in the economics of education presents evidence that principals also learn about teacher quality over time. While Staiger and Rockoff [2010] and Rivkin et al. [2005] point to the difficulty in identifying effective teachers at the point of hire, the strongest evidence of principals learning about teacher quality comes from Rockoff et al. [2012].[3] In a randomized control trial they provided teacher-level VAMs only to teachers' current principals, whom they surveyed before and after they had received the VAMs. Rockoff et al. [2012] find that those who randomly received more precise VAM reports were more responsive to the information, than were principals receiving noisier VAM reports.[4] These results are consistent with the Bayesian updating model used in Farber and Gibbons [1996], Altonji

[3]Jacob and Lefgren [2008] and Chingos and West [2011] also find evidence of principals' learning. Jacob and Lefgren [2008] find that principals are better at identifying the most and least effective teacher. Their observation of slightly higher correlations for principals who have known their teachers for longer suggests a gradual learning process. Chingos and West [2011] find that principals classify their teachers on the basis of effectiveness, and principals of schools under accountability pressure are more likely to move effective teachers into and less effective teachers out of high-stakes teaching assignments.

[4]Rockoff et al. [2012] also finds that providing VAMs to principals cause less effective teachers to leave at a higher rate. While the authors do not directly link these results to either learning hypothesis, these results in the experimental context are consistent with asymmetric employer learning.

and Pierret [2001], and Lange [2007].

Theoretical examinations of asymmetric employer learning came much earlier with Waldman [1984] offering the seminal work. Greenwald [1986] provides the initial prediction of adverse selection of mobile workers as a consequence of informational asymmetries between competing employers. More recently, Schönberg [2007], Pinkston [2009], Kahn [2013], and Bates [2015] each allow for private employer learning in an empirical setting. Also, each use the NLSY79 to test their models against empirical features of the data. Their cumulative evidence regarding asymmetric learning is mixed. Whereas, Schönberg [2007] finds that learning is largely symmetric, Pinkston [2009] finds that learning is largely asymmetric. Their disagreement hinges on whether information passes through job-to-job transitions, with Pinkston [2009] finding that the correlation between wages and ability moves more closely with respect to continuous working spells than with experience. Both Schönberg [2007] and Bates [2015] find that workers are only adversely selected into mobility in job-to-unemployment transitions, whereas asymmetric learning also predicts such selection for job-to-job moves as well. However, Bates [2015] also demonstrates positive selection into mobility on the basis of education, noting that consistent with asymmetric learning, those who attend more competitive colleges are more likely to both switch employers and be laid off. Consistent with asymmetric employer learning, Kahn [2013] finds that movers' wages are more volatile in the immediate aftermath of a transition than are the wages of those who remain in place.

A common criticism of much of the earlier literature asks what AFQT scores are really telling us. There is little evidence that AFQT scores are related to productivity in many jobs held by the largely low-skilled respondents of the NLSY. Similarly, if employers care greatly about AFQT scores, they would simply administer the test themselves. By using a more direct measure of productivity than the assumed correlates, this study avoids such criticism.

Only DeVaro and Waldman [2012] depart from the use of the NLSY. They use administrative personnel files to examine promotion decisions based on private and public information. In support of asymmetric employer learning, they find that conditional on private performance reviews, those with more education are more likely to be promoted than are those with less education. They also find that larger wage increases accompany promotions of less educated workers than accompany promotions of higher-educated workers. This, they argue, is due to the fact that promotions are a stronger public signal for those with lower, easily observable characteristics.

However, the stylized empirical facts given by each as evidence of asymmetric learning are consistent with the theoretical models, but are susceptible to alternative explanations. For instance, it is particularly difficult in each of these contexts to distinguish employer learning from specific human capital accumulation, which may be directed towards those with low general human capital but high ability. Further, post-move wage volatility may be explained by differences in job match quality, education may provide higher level skills leading to faster promotion, and symmetric learning may explain why large wage increases accompany promotions of less-educated workers. The absence of direct asymmetric information shocks has prevented the previous literature from isolating information acquisition from alternate explanations. This work uses 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.

While there is a large literature examining the mobility patters of higher- or lower-VAM teachers, none have previously considered the signaling effects of VAMs on teacher mobility and the distribution of teacher quality within the market. Earlier work shows that students in poor, low-achieving schools face teachers who are in general less experienced, less educated, and less effective than their counterparts in more affluent and higher achieving schools [Lankford et al., 2002, Clotfelter et al., 2005, Sass et al., 2012].[5] Though their

[5]Sass et al. [2012] also notes that there is huge variation in teacher quality within high poverty schools.

is significant churn within the teacher labor market, Hanushek et al. [2005], Krieg [2006], Goldhaber et al. [2007] and Boyd et al. [2008] each note that teachers whom they estimate to have higher VAMs tend to stay in the profession longer, and are no more likely to transfer between schools than their less-effective counterparts.[6] There is more disagreement about distributional effects of this turnover. Boyd et al. [2008] finds that, conditional on moving, high-VAM teachers are more likely to move to high-performing schools than are low-VAM teachers, whereas Hanushek et al. [2005] and Goldhaber et al. [2007] find no evidence of this resorting of teachers. While descriptions of where effective teachers have traditionally moved from and to have important implications for education inequity, they have little power to predict how the adoption of VAMs will alter the allocation of teachers across schools.

If VAMs provide new and credible information to principals, this new signal may expand the number of schools willing to hire high-VAM teachers. Jackson [2009] and Boyd et al. [2013] each find that on average white teachers prefer not to teach in schools with a large proportion of black students. Boyd et al. [2013] also find that teachers prefer schools that are closer, are suburban, and have a smaller proportion of students in poverty. Taking these estimated preferences as given, the new signals of teacher quality may lead high-VAM teachers to move to schools that have lower proportions of minorities, are more affluent, and are higher achieving. Guilford and Winston-Salem's release of VAMs, allows this work to provide the first exploration of this previously ignored consequence of this controversial policy.

4 Model

This section describes a model and provides predictions for which workers move, and where they go–and how each may change in response to an information shock. Table 2 presents a summary of the predictions and corresponding key assumptions, tables of evidence, and

[6] Boyd et al. [2008] finds that ineffective teachers are more likely to leave the profession only in their first year of teaching.

appendices for proofs of these predictions. The remaining section focuses on describing the intuition, model structure, and the primary predictions. Please see Appendix 9.1 for proofs and discussion of secondary predictions. This model builds 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 tweaks 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.

4.1 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, 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.[7] 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 $\epsilon \sim N(0, \sigma_\epsilon)$.[8]
2. The public signal is given by $R_x = \mu + \xi_x$, where $\xi \sim N(0, \sigma_\xi(x))$, and $\frac{\partial \sigma_\xi(x)}{\partial x} < 0$.
3. 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.

[7]Principals face rigid budget constraints, which translate to a fixed number of positions.
[8]The normality assumptions are not necessary, but are useful in deriving the comparative statics.

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

4. 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 $P_{t\nu}^r = \frac{\sigma_\nu P_t^r + \sigma_\tau(t) V}{\sigma_\nu + \sigma_\tau(t)}$. The variance of $P_{t\nu}^r$ is denoted as $\sigma_\tau(t\, V)$. The hiring principal's signal remains unchanged.

5. The noise of each signal is orthogonal to the noise of the other signals.[9]

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).

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 $\left(\frac{\partial \sigma_\xi(x)}{\partial x} < 0\right)$.

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

[9]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.

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 $t > 0$).

VAMs enter the learning model as an additional piece of information that may enter either the public or private signal. VAMs influence the public signal if they are accessible to both principals as occured within Guilford adn Winston-Salem school districts. If VAMs enter both principals' public signal, $R_{x\nu} = \frac{\sigma_\nu R_x + \sigma_\xi(x) V}{\sigma_\nu + \sigma_\xi(x)}$ replaces R_x. VAMs impact the private signal, if they are accessible to only current principals, as is more likely to occur when hiring principals are from different districts. If VAMs enter retaining principals' private signal, $P^r_{t\nu} = \frac{\sigma_\nu P^r_t + \sigma_\tau(t) V}{\sigma_\nu + \sigma_\tau(t)}$ replaces P^r_t. The introduction of VAMs alter these expectations by changing both the content of the signal and the signal's precision, and thus the weight that principals ascribe to it.

4.2 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.[10] 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.[11] Further, cumbersome dismissal processes result in teachers initiating much of the mobility. While principals cannot adjust salaries to influence whether a teacher stays, principals 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 (C_{isd}), which is comprised of salary (w_d) set by district d, characteristics

[10]As of 2014, North Carolina will move to paying teachers in part based upon teachers' VAMs.
[11]In Section 7, I discuss policy exceptions to this in North Carolina school districts.

of school s (S_{sd}), and characteristics of position i (J_{isd}). Thus, $C_{isd} = w_d + S_{sd} + J_{isd}$. It may be unrealistic to presume that large differences in pay or school attributes may be overcome by position-specific attributes. Thus, I impose a maximum quality of job attributes $\overline{J}$ that will be useful in predicting to where teachers move.

For simplicity, I assume that each principal presents a sealed bid for the teacher and pays the minimum of the two bids.[12] In such sealed-bid, second-price auctions, principals' optimal strategy is to offer the their expectation of the teacher's effectiveness (assuming that principals seek to maximize teacher effectiveness within their schools).[13] Principals formulate these expectations by averaging over their prior belief of quality (m), the public signal (R_x), and their private signals (P_0^h or P_t^r). They weight each signal by its precision relative to the other signals, similar to a standard Bayesian updating model. I list the bids of hiring (b_{isdNV}^{h*}) and retaining (b_{isdNV}^{r*}) principals in equations 1 and 2 respectively.[14]

$$b_{isdNV}^{h*} = \frac{\sigma_\tau(0)\sigma_\xi(x)}{Z_{NV}^h} m + \frac{\sigma_\tau(0)\sigma_\epsilon}{Z_{NV}^h} R_x + \frac{\sigma_\epsilon\sigma_\xi(x)}{Z_{NV}^h} P_0^h.^{15} \tag{1}$$

$$b_{isdNV}^{r*} = \frac{\sigma_\tau(t)\sigma_\xi(x)}{Z_{NV}^r} m + \frac{\sigma_\tau(t)\sigma_\epsilon}{Z_{NV}^r} R_x + \frac{\sigma_\epsilon\sigma_\xi(x)}{Z_{NV}^r} P_t^r.^{16} \tag{2}$$

If there is public learning, increases in experience leads to a more precise public signal. As $\sigma_\xi(x)$ declines, both principals place relatively more weight on the public signal, R_x, and less weight on the prior belief (m) and the private signals (P_0^h or P_t^r). Thus, the bids of the hiring and retaining principals converge. 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

[12] Previous versions modeled open continuous bidding, which permits the adoption of optimal bidding strategies from Milgrom and Weber [1982]. This allows each school to update the optimal bid conditioning on the rival's bidding behavior. However, both bidding processes result in the same predictions.

[13] Prior work shows principals care about teacher effectiveness, particularly in schools under accountability pressure. Other work shows that high-VAM teachers also lead to a wide array of better future outcomes for their students, giving further reason to suggest principals may maximize these short-run measures of effectiveness.

[14] Subscript NV indicates neither principal received the teacher's VAM.

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

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

and public signal. 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.

The introduction of VAMs alters the information available to principals, but the optimal bids that incorporate VAMs have similar form to those shown in equation 1 and 2. Whether the VAMs are public or private are particularly important for determining retaining principals' expectations of a given teacher in the adopting districts.

If a principal's rival is from outside of the district and uninformed of the measure, when a retaining principal receives a teacher's VAM, 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 VAM. 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.

$$b^{r*}_{isdRV} = \frac{\sigma_\tau(t\,V)\sigma_\xi(x)}{Z^r_{RV}}m + \frac{\sigma_\tau(t\,V)\sigma_\epsilon}{Z^r_{RV}}R_x + \frac{\sigma_\epsilon\sigma_\xi(x)}{Z^r_{RV}}P^r_{t\nu}.^{17} \tag{3}$$

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

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

Proof: Under the orthogonality assumptions, $var(P_{t\nu}) \equiv \sigma_\tau(t\,V) = \frac{\sigma^2_\nu\sigma_\tau(t)+\sigma_\nu\sigma_\tau(t)^2}{(\sigma_\nu+\sigma_\tau(t))^2} = \frac{\sigma_\nu\sigma_\tau(t)}{\sigma_\nu+\sigma_\tau(t)}$. $\frac{\sigma_\tau(t)(\sigma_\nu+\sigma_\tau(t))}{\sigma_\nu+\sigma_\tau(t)} - \frac{\sigma_\nu\sigma_\tau(t)}{\sigma_\nu+\sigma_\tau(t)} = \frac{\sigma^2_\tau(t)}{\sigma_\nu+\sigma_\tau(t)}$, and $\frac{\sigma^2_\tau(t)}{\sigma_\nu+\sigma_\tau(t)} > 0$, 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 VAMs exacerbates informational asymmetries between

[17] $Z^r_{RV} = \sigma_\tau(t\,V)\sigma_\xi(x) + \sigma_\tau(t\,V)\sigma_\epsilon + \sigma_\epsilon\sigma_\xi(x)$.

prospective employers, and the two principals' bids further diverge.

In contrast, the VAM enters the public signal, if both bidding principals are informed of a teacher's VAM (as occurred when both principals were within districts that adopted VAMs). I list the optimal bids of hiring and retaining principals' when both have access to a teacher's VAM in equations 4 and 5 respectively.

$$b^{h*}_{isdHV} = \frac{\sigma_\tau(0)\sigma_\xi(x\,V)}{Z^r_{HV}}m + \frac{\sigma_\tau(0)\sigma_\epsilon}{Z^r_{HV}}R_{x\nu} + \frac{\sigma_\epsilon\sigma_\xi(x\,V)}{Z^r_{HV}}P^h_0.^{18} \tag{4}$$

$$b^{r*}_{isdHV} = \frac{\sigma_\tau(t)\sigma_\xi(x\,V)}{Z^r_{HV}}m + \frac{\sigma_\tau(t)\sigma_\epsilon}{Z^r_{HV}}R_{x\nu} + \frac{\sigma_\epsilon\sigma_\xi(x\,V)}{Z^r_{HV}}P^r_t.^{19} \tag{5}$$

Equations 4 and 5 are similar to equations 1 and 2 with the exception that R_x is replaced by $R_{x\nu}$, as VAMs enter the public signal. While in expectation the magnitude of the public signal is the same with or without VAMs, 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 VAMs into the public signal ($\sigma_\xi(x\,V) < \sigma_\xi(x)$).

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

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

[18] $Z^h_{HV} = \sigma_\tau(0)\sigma_\xi(x\,V) + \sigma_\tau(0)\sigma_\epsilon + \sigma_\epsilon\sigma_\xi(x\,V)$.
[19] $Z^r_{HV} = \sigma_\tau(t)\sigma_\xi(x\,V) + \sigma_\tau(t)\sigma_\epsilon + \sigma_\epsilon\sigma_\xi(x\,V)$.

4.3 Mobility with the introduction of VAMs

After teachers receive both bids, they move to the school that offers the highest bid.[20] Accordingly, the probability of a move is:

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

The availability of VAMs 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 mobility will change with the adoption of VAMs? As described in Section 2, both districts' adoptions of VAMs provide a shock to the information of all principals within the district. Thus, by examining teacher mobility in response to the release of VAMs, I test whether releasing VAMs leads toward informational symmetry between employers. However, out-of-district principals cannot directly access these new VAMs. 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 VAMs in the model. The first is more in keeping with the prior employer learning literature. Empirically, VAMs 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 these VAMs. Equation 7 takes this broad view.[21]

$$\begin{aligned}\frac{\partial E[b_{HV}^{h*} - b_{HV}^{r*} - (b_{NV}^{h*} - b_{NV}^{r*})|m\,\mu]}{\partial \mu} &= \frac{{\sigma_\epsilon}^2(\sigma_\tau(0) - \sigma_\tau(t))(\sigma_\xi(x) - \sigma_\xi(x\,V))}{Z_{NV}^h Z_{NV}^r Z_{HV}^h Z_{HV}^r} \\ &[\sigma_\xi(x)\sigma_\xi(x\,V)(2\sigma_\tau(t)\sigma_\tau(0) + \sigma_\epsilon\sigma_\tau(0) + \sigma_\tau(t)\sigma_\epsilon) + (\sigma_\xi(x\,V) + \sigma_\xi(x))\sigma_\tau(t)\sigma_\epsilon\sigma_\tau(0))] > 0.\end{aligned} \tag{7}$$

[20] 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.

[21] See Appendix 9.1.2 for proof.

Under the assumption that $\sigma_\tau(0) > \sigma_\tau(t)$, which is fundamental to asymmetric employer learning and by $\sigma_\xi(x) > \sigma_\xi(xV)$, which was shown in lemma 2, $\frac{\partial E\left[b^{h*}_{HV} - b^{r*}_{HV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ \mu\right]}{\partial \mu} > 0$. Therefore, the model predicts that providing VAMs to both principals, as occurred within both districts, should raise the probability that good teachers move, all else equal.

Under the second interpretation, EVAAS VAMs enter 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.[22]

$$\frac{\partial E\left[b^{h*}_{HV} - b^{r*}_{HV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ V\ \mu\right]}{\partial V} = \frac{1}{Z^h_{HV} Z^r_{HV}} \frac{\sigma_\xi(x)}{\sigma_\nu + \sigma_\xi(x)} > 0 \tag{8}$$

Therefore, while the interpretations are subtly different, the comparative statics with respect to VAMs after the policy takes effect are the same. Within the districts, where both principals are aware of the signals, the model predicts releasing VAMs increases the probability that teachers who receive high-VAM signals will transfer schools.

Recall from Section 2, that if principals in other districts know of the existence of VAMs for teachers from Winston-Salem and Guilford, the policy would theoretically alter their informationl. 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, VAMs enter 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 VAMs' role apply here. I will first take the broader view of VAMs as a measure of μ. Equation 9 demonstrates the predicted change in the relationship between teachers' underlying abilities and the probability of moving to uninformed principals

[22] See Appendix 9.1.3 for proof.

once districts release their teachers' VAMs.[23]

$$\frac{\partial E[b^{h*}_{NV} - b^{r*}_{RV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ \mu]}{\partial \mu} = \frac{\sigma_\epsilon \sigma_\xi(x)^2(\sigma_\tau(t\ V) - \sigma_\tau(t))}{Z^r_{NV} Z^r_{RV}} < 0. \tag{9}$$

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 VAMs to retaining principals increases the likelihood that ineffective teachers move out-of-district, and vice versa.

Taking the more narrow view of VAMs as only pertaining to the signal itself, again the predictions remain consistent. Equation 10 presents the partial derivative of the expected difference in the differences between employers bids with respect to the VAM signal itself.[24]

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

Here, the policy leads to adverse selection of out-of-district moves on the basis of the signal, all else equal.

It is important to note that good (or high-VAM) 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-VAM 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 VAM on the probability of moving out-of-district after policy implementation than are produced for moves within-district.[25] Thus, the test between symmetric and asymmetric learning is whether effects of the policy on the selection of out-of district movers are significantly more negative than the effects of adopting VAMs on the selection of within-district movers.

[23] See Appendix 9.1.4 for proof.

[24] See Appendix 9.1.5 for proof.

[25] See Appendix 9.1.6 for a more formal treatment of this.

Table 2: Model Predictions

	Primary Model Predictions	Assumptions: There was prior private learning,VAMs are informative, and...	Parameterized Predictions	Table	Appendix
1.	Better (higher-VAM) teachers will become more likely to move within district (subscript WD) after the adoption of VAMs.		$\gamma_{14WD} > 0$	4	9.1.2 (9.1.3) :
2.	Worse (lower-VAM) teachers will become more likely to move out of district (subscript OD) after the adoption of VAMs.	VAMs may be kept private.	$0 > \gamma_{14OD}$	4	9.1.4 (9.1.5)
	Secondary Model Predictions				
3.	The introduction of VAMs 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.	VAMS are a more salient signals for teachers moving from Winston-Salem than for teacher moving from Guilford.	$\gamma_{14WD_{GCS}} - \gamma_{14OD_{GCS}} > \gamma_{14WD_{WSF}} - \gamma_{14OD_{WSF}}$	4	9.1.6
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 $\gamma_{14S} > 0$	4, 5	9.1.7
5.	The introduction of VAMs 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.	VAMs may be kept private.	$\gamma_{24WD} < 0$ and $0 < \gamma_{24OD}$	6	9.1.8 9.1.9
6.	The positive change in selection should be particularly true for teachers with more tenure at a given school.		Coefficient on $VAM \times Ten \times TreatDist > 0$	7	9.1.10

5 Data and Estimation

In this section, I describe both the data and methods used to generate VAMs of teacher effectiveness, and estimate the effects of the district policies on the teacher mobility. Subsection 5.1 describes the generation of VAMs. Subsection 5.2 describes the estimation sample. Subsection 5.3 describes the difference-in-differences estimation approach used to identify the effects of the new information on the mobility decisions of teachers and principals.

5.1 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 VAMs issued to each teacher within the treatment districts, and neither district agreed to release them. Consequently, this study will measure 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. To illustrate this distinction, suppose that the EVAAS score were totally uninformative. Observing mobility based on them would clearly illustrate the impact of the additional signal, but would offer no insight into the effect on educational equity. In contrast, using a measure of true effectiveness provides direct policy implications. Further, following earlier studies of employer learning in presuming that the researcher may access information, which is more closely tied to productivity, but is originally unavailable to market participants, this broader approach to VAMs as a proxy for ability is also useful in testing the learning hypotheses.[26] Accordingly, I prefer this second, broader approach. In my preferred specification, I model teacher effectiveness rather than attempting to replicate the EVAAS measure.[27] In actuality, the resulting measures from either approach are likely be highly correlated, and in Section 7, I check the robustness of my results against other specifications.[28] In this context, the VAMs need not totally encompass a teacher's effectiveness. Here, VAMs only need to be stronger correlates with teacher effectiveness than are other correlates with productivity, such as educational attainment and level of certification.[29] While VAMs likely do not measure all traits that principals may seek in their teachers, they do directly measure one component of teaching quality that is important to principals and policy makers.

My preferred measure of VAMs is what Guarino et al. [2012] call the Dynamic OLS (DOLS) estimator presented in equation 11. According to Guarino et al. [2012], this DOLS estimator is more robust to nonrandom student assignment, a frequent criticism of the often used Empirical Bayes estimator, which assumes random assignment of students to teachers.[30]

$$A_{ijt} = \tau_{\mathbf{t}} + \mathbf{A_{ijt-1}}\boldsymbol{\beta_0} + \mathbf{X_{it}}\boldsymbol{\beta_1} + \boldsymbol{VAM_j} + e_{it} \tag{11}$$

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

[26] 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 VAM described above in this capacity.

[27] 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.

[28] Rose et al. [2012] finds a 0.91 correlation between one EVAAS measure and Dynamic OLS.

[29] 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' VAMs and many future student outcomes, including educational attainment, earnings, and probability of incarceration [Chetty et al., 2011, 2014].

[30] 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.

individual students, such as grade, race, gender, special needs, and gifted status. It is $\mathbf{VAM_j}$, a vector of teacher indicators, which is of primary interest for this study. Acknowledging that VAMs 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.

5.2 Estimation Sample

This study restricts attention to the 5,986,132 elementary and middle school student, year observations from 1997 through 2011 to construct the VAMs 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 VAMs. 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 3 provides summary statistics for my estimation sample.

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

Table 3: 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
VAM	0.02	1.01	0.01	0.99	0.00	1.00
N	11,239		8,295		216,484	

Note: VAM 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.

5.3 Estimation Strategy

I use differences-in-differences to compare changes in mobility around the adoptions of VAMs to the changes in mobility over the same times in the rest of the state. I estimate the following specification:

$$y^*_{jdt} = \boldsymbol{T_t} + \mathbf{d_d} + \mathbf{TreatDist_{jd}}\delta + VAM_j\mathbf{G_{1dt}} + \mathbf{X_{jdt}}\mathbf{G_{2dt}} + \xi_{jdt}, \tag{12}$$

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

where y^*_{jdt} is the latent probability of a job change for teacher j in district d and in year t. I only observe the binary outcome of when a move occurs. $\mathbf{T_t}$ 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,[31] 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. $\mathbf{G_{1dt}}$ and $\mathbf{G_{2dt}}$ capture the differences in the effects of VAMs on mobility based on whether VAMs were available for teacher j in district d, at time t. Interactions with treatment district indicators separate permanent differences in the impacts of VAMs 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 VAMs 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 VAMs, which necessitates interacting other teacher covariates with the differences-in-differences framework 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' VAMs and the potential additional salience of VAM signals to principals outside the district during Winston-Salem's later adoption, I separate treatment by district.

Given how the districts distributed VAMs, 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

[31] 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.

of the interaction of VAM with receiving treatment on the probability of moving within-district). Furthermore, because there would be more information available on more experienced teachers, if there 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 VAM with experience and the difference-in-differences, $\mathbf{G}$, interactions.

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 VAMs 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 VAM with receiving treatment on the probability of moving out-of-district). If current principals can keep information from employers in other districts, 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 $\boldsymbol{\gamma_{14WD}} > \boldsymbol{\gamma_{14OD}}$ and possibly $\boldsymbol{\gamma_{14OD}} < 0$ 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 VAMs are informative, and teachers do in general prefer to teach at better schools, $\boldsymbol{\gamma_{14SQ}} > 0$ in this regression as well. ($\boldsymbol{\gamma_{14SQ}}$ is the effect of the interaction of VAM 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.

There are two distinct issues that complicate the estimation of standard errors in this study. First, the policy variation occurs at the district level. As a result, the errors may be correlated for teachers moving from or within the same district. The appropriate response to this single issue is to cluster the standard errors at the district level. The second, issue results from the fact that the teacher VAMs are estimated. By simply clustering the standard errors, the VAMs are treated as though they are known, and thus, they do not account for the inherent variability due to estimation error. Were this a singular issue, it would be appropriate to bootstrap the student data to account for this estimation error. It may seem natural to then cluster-bootstrap at the district level. However, this samples all students for a every teacher in a sampled district, and as a result, does not actually address the estimation error. In fact, the standard errors from the cluster bootstrap are smaller than the non-bootstrap clustered standard errors by about a factor of ten.

Accordingly, I adopt a sampling approach that accounts for both the estimation error of VAM 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 more conservative standard errors across parameters. The standard errors on the effects of the policy on the relationship between VAMs and the probability of moving schools are comparable to the standard bootstrapped standard errors, and the standard errors on all other estimated coefficients are comparable to the non-bootstrapped district-clustered standard errors. Table 17 in the Appendix 9.6 presents all standard errors for Table 4 for comparison. Throughout the remainder of this paper, I present the more conservative district-clustered-teacher-stratified-bootstrap standard errors (CSB SEs).

6 Results

6.1 Mobility and Sorting

How does mobility change with the adoption of VAMs and what does that tell us about the way employers learn about their employees? Table 4 presents the estimated impact of revealing EVAAS reports of teacher effectiveness on the relationship between teachers' VAMs and the probability a teacher moves to another school. Given the evidence that teachers prefer to teach in schools with higher-performing students, Table 4 decomposes effects by whether the receiving school has higher or lower-performing students.[32] The test between symmetric and asymmetric employer learning focuses on how the effects of VAMs on the probability of moving within-district differ from the effects of VAMs 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 VAMs 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 VAMs 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 VAM 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 VAM 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

[32]Primary effects of VAMs 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.

about the same magnitude.

Within both Guilford and Winston-Salem, the release of VAMs intensifies this pattern. From the coefficient on the interactions between policy treatment and VAMs in both districts, a standard deviation increase in a teacher's VAM 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 VAMs raises the probability that a teacher with one standard deviation higher VAM 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 VAMs on the probability of moving to a lower-performing school within district. The similarity of the point estimates on the impact of VAMs post-treatment between Guilford and Winston-Salem provides no evidence that relying upon teachers to voluntarily disclose their VAMs to hiring principals mitigates the effects.

From Section 4, 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 VAMs. 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 VAM, is a full percentage point more likely to move out-of-district. This same 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 VAMs 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 VAM 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 VAMs. Given this difference in contexts, high VAM teachers from Winston-Salem may have been better able to use their VAMs to obtain positions outside of Winston-Salem, than would a comparable teacher moving earlier from Guilford. In Winston-Salem, the increase in high-VAM teachers' ability to signal their effectiveness may mitigate any effects from relatively low VAM teachers exploiting the informational asymmetry. The mitigated effects of VAM for those moving out of Winston-Salem in addition to the negative selection of teachers moving away from Guilford evidences informational asymmetries

Table 4: Probability of Moving Schools Within and Out of District

	Panal A: Within-District Moves			Panal 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. *** p<0.01, ** p<0.05, * p<0.1.

between potential employers within as opposed to outside of the district.

6.2 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 5 presents these results in panel A, as well as, the effects of VAM adoption on the sorting of teachers to schools with respect to students' race (panel B) and students' performance (panel C).[33]

The coefficient on VAM 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. Moving

[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.

down the column shows that the release of VAMs magnifies that sorting in both adopting districts. VAM adoption in Winston-Salem leads to a 1.3 percentage point increase in the probability that a teacher with a standard deviation higher VAM 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 VAM describes the general relationship between teachers' VAMs 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 VAMs. 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 VAMs 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 VAM 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 VAM 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]

Next, I turn to the change in sorting with VAM 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 VAMs of teacher effectiveness 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 VAMs leads a teacher with one standard deviation higher VAM 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 VAMs. 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. The large effects in Winston-Salem taken together with the mobility patterns from Table 4 and panel A of Table 5 evidence rising inequality in the distribution of effective teachers as an unintended consequence of VAM adoption.

6.3 Observables

In addition to predicting mobility dynamics with respect to teacher VAMs, the model presented in Section 4 also offers predictions regarding easily observable covariates with teacher effectiveness. In instances where the VAMs shock the available public information, the model

[34] The result that students in better schools also get better teachers is consistent with findings in Boyd et al. [2005] and Boyd et al. [2008].

Table 5: 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
VAM	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]
VAM 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]
VAM 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 interactionswith treatment indicators. *** p<0.01, ** p<0.05, * p<0.1.

predicts principals would place less emphasis on easily observable covariates with teacher effectiveness, such as degree attainment and college selectivity. In cases where VAMs 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 VAMs 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.[35]

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

[35]The VAMs used in this analysis are the residuals from the projection of my standard VAMs on the components of the index.

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 6 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 VAMs 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 VAMs particularly for moves to better schools within Guilford. More positions may become available as a result of high-VAM teachers moving to better schools, and low-VAM teachers moving out of district. As a result, those with good observables find it easier to move in addition to those with high VAMs. Heterogeneous openness among principals to VAMs may also contribute.[36] In which case, as high-VAM teachers move to principals that value VAMs 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-district with the adoptions of VAMs is more supportive of the model. Whereas movers out of Guilford are adversely selected on the basis of the hard-to-observe VAM, 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 VAM were able to keep their VAM 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 VAMs, 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

[36] 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' VAMs played a limited role in their hiring decisions.

Table 6: Effects of teacher quality index on the probability of moving

	Within-District Moves			Out-of-District Moves		
Variables	Total	To a higher performing schools	To a lower performing schools	Total	To a higher performing schools	To a lower performing schools
VAM	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]
VAM 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]
VAM 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 VAMs used in this analysis are the residuals from the projection of my standard VAMs on the components of the index. *** p<0.01, ** p<0.05, * p<0.1

is too mixed to draw definitive conclusions.[37]

6.4 Differential Effects With Respect to Experience and Tenure

The final piece of primary analysis examines the effects of the policy on the correlation between teacher VAMs and the probability of moving with respect to years of tenure. Were private learning already prevalent in the market, the model predicts the effects of the policy to be larger for those who have taught at the same school for longer, all else being equal. This is due to large existing asymmetries in information between current and hiring principals,

[37]In unreported regressions, with the exception of out-of-Guilford moves the results shown in Table 6 are very sensitive to the variable composition of the teacher quality index.

which releasing VAMs largely undo. Thus within-district, the model predicts larger positive selection of movers who have more years of tenure. 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-VAM 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 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, the release of VAMs would not serve as much of a shock for teachers about whom there already exists a great deal of information. Thus, we may expect smaller results for less experienced teachers. While Table 7 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 learning, though the mixed evidence on public learning makes me hesitant to draw definitive conclusions on the prior learning environment.

7 Robustness

In the following section, I examine the robustness of the effects of VAM adoption. Section 7.1 considers changes in effects when using only prior years of student data when constructing VAMs. Section 7.2 considers whether other district policies that paid teachers to work in hard-to-staff schools impact the estimated effects. Appendix 9.2 considers teacher mobility in accordance with the state ABC growth bonus-pay system. Within-district, year-by-year analysis of the changing effects of VAMs on mobility and sorting are presented in Appendix 9.3. In Appendix 9.4 and Appendix 9.5, I consider alternate functional forms for the mobility analysis. In Appendix 9.4, I take seriously the normality assumptions, and perform normal Maximum Likelihood Estimation. In Appendix 9.5, I use competing risks regression to

Table 7: Differential Effects With Respect to Experience and Tenure

VARIABLES	Within District Total	Within District Higher Performing	Out of District Total	Out of District Higher Performing
VAM	-0.0001	0.0028*	-0.0001	0.0023
	[0.0023]	[0.00161]	[0.00244]	[0.00173]
Experience x VAM	-0.0000	0.0000	-0.0000	-0.0000
	[0.00011]	[0.00008]	[0.00011]	[0.00008]
Tenure x VAM	0.0020**	0.0006	0.0006	0.0005
	[0.0008]	[0.00059]	[0.00073]	[0.00058]
VAM x Treatment GCS	0.0033	0.0050	-0.0181***	-0.0095*
	[0.00568]	[0.00465]	[0.00693]	[0.00514]
Experience x VAM x Treatment GCS	0.0016***	0.0010***	0.0002	0.0003
	[0.00026]	[0.0002]	[0.00032]	[0.00026]
Tenure x VAM x Treatment GCS	0.0056***	0.0004	0.0008	0.0014
	[0.00179]	[0.00146]	[0.00217]	[0.00178]
VAM x Treatment WSF	-0.0003	-0.0010	-0.0073	-0.0051
	[0.00551]	[0.00431]	[0.00503]	[0.00452]
Experience x VAM x Treatment WSF	0.0003	0.0005	0.0002	0.0002
	[0.00043]	[0.00036]	[0.00029]	[0.00025]
Tenure x VAM 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. *** p<0.01, ** p<0.05, * p<0.1

examine the possibility of correlated errors between types of moves.[38]

7.1 Sensitivity to VAM Construction

The possibility that teachers may have different VAMs after moving to other schools, may present issues for using VAMs constructed from student data from a teacher's entire career. This could result from moves leading to higher match quality between teachers and schools, as Jackson [2013] finds. It may also result from transitory adjustment costs, giving a theoretically ambiguous direction of potential bias.[39]

Consequently, in Table 8, I allow teachers VAM scores to vary each year, using only data from the current and previous years to construct a teacher's VAM 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 VAMs 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 VAM 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 4. Panel C in Table 8 corresponds with Table 5. While the effect on teacher sorting doubles in Winston-Salem, the results become more negative and statistically insignificant in Guilford.

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 VAM prior to the their release is identified

[38]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.

[39]More closely approximating the information that teachers and principals receive is another rationale for restricting the data used in generating teacher VAMs. In which case using Empirical Bayes estimation provides what is believed to be a closer approximation to the algorithm used in creating the EVAAS measures. Table 19 in Appendix 9.6 provides results using Empirical Bayes estimation on the restricted sample of student test scores in calculating teacher VAMs. The results are very similar.

Table 8: Probability of moving schools within-district using restricted data VAM

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
VAM	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]
VAM 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]
VAM 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.
*** p<0.01, ** p<0.05, * p<0.1

off of just two years of data. As a result, the estimates of teachers VAMs 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 getting around this issue is to use a fixed number of years prior to the current period when constructing VAMs. Unfortunately, the adoption of VAMs by Guilford comes just three years into the student data sample. Since the construction of VAMs requires at least one prior year of student data, this gives just two years at which I could fix my VAM estimate. Not only would this force a noisier estimate of each teacher's VAM 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 VAMs, I drop Guilford from the analysis and vary the number of prior years of data I use to construct the VAMs from 2 to 8. Table 9 demonstrates that though the relationship between years used and the effect of the interaction of the policy in Winston-Salem and VAM is not monotonic as the sample used varies, the estimates using

Table 9: Effect of VAMs constructed using various number of years on the probability of moving to a "better" school

VARIABLES	2yr VAM	3yr VAM	4yr VAM	5yr VAM	6yr VAM	7yr VAM	8yr VAM
VAM	0.0020***	0.0023***	0.0024***	0.0023***	0.0025***	0.0027***	0.0040***
	[0.00054]	[0.0005]	[0.00051]	[0.00073]	[0.00076]	[0.00072]	[0.00083]
VAM x Treatment Winston-Salem	0.0103***	0.0087***	0.0076***	0.0064**	0.0099***	0.0118***	0.0150***
	[0.00241]	[0.00233]	[0.00245]	[0.00287]	[0.00293]	[0.003]	[0.00323]
Treatment Winston-Salem	0.0555***	0.0540***	0.0550***	0.0480***	0.0427***	0.0457***	0.0407***
	[0.00382]	[0.00373]	[0.00362]	[0.00385]	[0.00396]	[0.00427]	[0.00434]
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. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

more years of data are clearly the largest. This further suggests correlated measurement error presents a problem for this approach.

7.2 Strategic Staffing

A possible complication arises due to alternate teacher compenstion 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 occured in treatment districts during the sample period and could potentially alter teacher preferences over schools.[40] 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

[40]"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.

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 VAM. 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-VAM teachers may opt to work at low performing school, which is in fact the intent of the policy.

Panels A and B of table 10 reports similar information as is provided in Table 4, with the difference that the binary dependent variable in Table 10 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 4, 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 VAMs and the probability of moving within Winston-Salem. Column 2 shows that releasing VAMs raises the probability that a teacher with one standard deviation higher VAM 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 VAMs 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 10 presents the impacts of the policy on teacher sorting within-district and within-district among non-strategic staffing schools. Column 1 of panel C is identical to column 2 of panel C in Table 5. I include it here for ease of comparison. Column

Table 10: 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
VAM	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]
VAM 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]
VAM 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.
*** p<0.01, ** p<0.05, * p<0.1

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-VAM 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.[41] Table 10 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 much larger impacts of releasing VAMs.

8 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

[41] Table 20 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.

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 VAMs of teacher effectiveness, and often contentiously, this signaling role of the measures has avoided discussion. The policy implication of this finding is not to universally avoid using VAMs. However, it would be useful to provide policy makers an estimate of the cost of retaining high-VAM 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 7.2, several districts in North Carolina are implementing a range of staffing policies designed to induce teachers to work in low-performing schools. Some incorporate VAMs 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 VAMs, policy makers should be aware of the potential consequences of these policies on educational equity, as well as the costs of offsetting these effects.

References

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

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

Joseph G. Altonji and Charles R. Pierret. Employer learning and statistical discrimination. *The Quarterly Journal of Economics*, 116(1):313–350, February 2001. ISSN 0033-5533, 1531-4650. doi: 10.1162/003355301556329. URL http://qje.oxfordjournals.org/content/116/1/313.

Michael David Bates. Job seperation under asymmetric employer learning. Working paper, University of California at Riverside, July 2015. URL http://www.nber.org/papers/w5605.

Donald Boyd, Hamilton Lankford, Susanna Loeb, and James Wyckoff. Explaining the short careers of high-achieving teachers in schools with low-performing students. *The American Economic Review*, 95(2):166–171, May 2005. ISSN 0002-8282. doi: 10.2307/4132810. URL http://www.jstor.org/stable/4132810. ArticleType: research-article / Issue Title: Papers and Proceedings of the One Hundred Seventeenth Annual Meeting of the American Economic Association, Philadelphia, PA, January 7-9, 2005 / Full publication date: May, 2005 / Copyright Â© 2005 American Economic Association.

Donald Boyd, Pam Grossman, Hamilton Lankford, Susanna Loeb, and James Wyckoff. Who leaves? teacher attrition and student achievement. Working Paper 14022, National Bureau of Economic Research, May 2008. URL http://www.nber.org/papers/w14022.

Donald Boyd, Hamilton Lankford, Loeb, and James Wyckoff. 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, January 2013. ISSN 0734-306X. doi: 10.1086/666725. URL http://www.jstor.org/stable/10.1086/666725. ArticleType: research-article / Full publication date: January 2013 / Copyright Â© 2013 The University of Chicago.

Raj Chetty, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. How does your kindergarten classroom affect your earnings? evidence from project star. *The Quarterly Journal of Economics*, 126(4): 1593–1660, November 2011. ISSN 0033-5533, 1531-4650. doi: 10.1093/qje/qjr041. URL http://qje.oxfordjournals.org/content/126/4/1593.

Raj Chetty, John N. Friedman, and Jonah E. Rockoff. Measuring the impacts of teachers ii: Teacher value-added and student outcomes in adulthood. *American Economic Review*, 104(9):2633–2679, September 2014. ISSN 0002-8282.

Matthew M. Chingos and Martin R. West. Promotion and reassignment in public school districts: How do schools respond to differences in teacher effectiveness? *Economics of Education Review*, 30(3):419–433, June 2011. ISSN 0272-7757. doi: 10.1016/j.econedurev.2010.12.011. URL http://www.sciencedirect.com/science/article/pii/S0272775710001767.

Charles T. Clotfelter, Helen F. Ladd, and Jacob Vigdor. Who teaches whom? race and the distribution of novice teachers. *Economics of Education Review*, 24(4):377–392, August 2005. ISSN 0272-7757. doi: 10.1016/j.econedurev.2004.06.008. URL http://www.sciencedirect.com/science/article/pii/S0272775704001153.

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

Jed DeVaro and Michael Waldman. The signaling role of promotions: Further theory and empirical evidence. *Journal of Labor Economics*, 30(1):91–147, January 2012. ISSN 0734-306X. doi: 10.1086/662072. URL http://www.jstor.org/stable/10.1086/662072. ArticleType: research-article / Full publication date: January 2012 / Copyright Â© 2012 The University of Chicago.

Henry S. Farber and Robert Gibbons. Learning and wage dynamics. *The Quarterly Journal of Economics*, 111(4):1007–1047, November 1996. ISSN 0033-5533, 1531-4650. doi: 10.2307/2946706. URL http://qje.oxfordjournals.org/content/111/4/1007.

Jason P. Fine and Robert J. Gray. A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association*, 94(446): 496–509, June 1999. ISSN 0162-1459. doi: 10.1080/01621459.1999.10474144. URL http://amstat.tandfonline.com/doi/abs/10.1080/01621459.1999.10474144.

Robert Gibbons and Lawrence Katz. Layoffs and lemons. Working Paper 2968, National Bureau of Economic Research, December 1991. URL http://www.nber.org/papers/w2968.

Steven Glazerman, Ali Protik, Bing-ru Teh, Julie Bruch, and Neil Seftor. Moving high-performing teachers: Implementation of transfer incentives in seven districts. Technical report, Mathematica Policy Research, 2012.

Dan Goldhaber, Betheny Gross, and Daniel Player. *Are Public Schools Really Losing Their Best? Assessing the Career Transitions of Teachers and Their Implications for the Quality of the Teacher Workforce. Working Paper 12.* National Center for Analysis of Longitudinal Data in Education Research, October 2007.

Robert J. Gray. A class of k-sample tests for comparing the cumulative incidence of a competing risk. *The Annals of Statistics*, 16(3):1141–1154, September 1988. ISSN 0090-5364. URL http://www.jstor.org/stable/2241622.

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

Sanford J. Grossman. The informational role of warranties and private disclosure about product quality. *Journal of Law and Economics*, 24(3):461–483, December 1981. URL http://www.jstor.org/stable/725273.

Cassandra Guarino, Mark D. Reckase, and Jeffrey M. Wooldridge. Can value-added measures of teacher performance be trusted? Technical Report 6602, Discussion Paper series, Forschungsinstitut zur Zukunft der Arbeit, 2012. URL http://www.econstor.eu/handle/10419/62407.

Davis Guggenheim. *Waiting for Superman*. Paramount Home Entertainment, Hollywood, CA, 2011.

Eric A. Hanushek, John F. Kain, Daniel M. O'Brien, and Steven G. Rivkin. The market for teacher quality. Working Paper 11154, National Bureau of Economic Research, February 2005. URL http://www.nber.org/papers/w11154.

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

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

Brian A. Jacob and Lars Lefgren. Can principals identify effective teachers? evidence on subjective performance evaluation in education. *Journal of Labor Economics*, 26(1):101–136, January 2008. ISSN 0734-306X, 1537-5307.

Boyan Jovanovic. Job matching and the theory of turnover. *Journal of Political Economy*, 87(5):972–990, October 1979. ISSN 0022-3808. doi: 10.2307/1833078. URL http://www.jstor.org/stable/1833078. ArticleType: research-article / Issue Title: Part 1 / Full publication date: Oct., 1979 / Copyright Â© 1979 The University of Chicago Press.

Lisa B Kahn. Asymmetric information between employers. *American Economic Journal: Applied Economics*, 5(4):165–205, October 2013. ISSN 1945-7782, 1945-7790.

John M. Krieg. Teacher quality and attrition. *Economics of Education Review*, 25(1): 13–27, February 2006. ISSN 0272-7757. doi: 10.1016/j.econedurev.2004.09.004. URL http://www.sciencedirect.com/science/article/pii/S0272775704001281.

Fabian Lange. The speed of employer learning. *Journal of Labor Economics*, 25(1):1–35, January 2007. ISSN 0734-306X. doi: 10.1086/508730. URL http://www.jstor.org/stable/10.1086/508730. ArticleType: research-article / Full publication date: January 2007 / Copyright Â© 2007 The University of Chicago.

Hamilton Lankford, Susanna Loeb, and James Wyckoff. Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational Evaluation and Policy Analysis*, 24(1): 37–62, March 2002. ISSN 0162-3737, 1935-1062. doi: 10.3102/01623737024001037. URL http://epa.sagepub.com/content/24/1/37.

Paul Milgrom and Sharon Oster. Job discrimination, market forces, and the invisibility hypothesis. *The Quarterly Journal of Economics*, pages 453–476, 1987.

Paul R. Milgrom. Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics*, 12(2):380–391, October 1981. doi: 10.2307/3003562. URL `http://www.jstor.org/stable/3003562`.

Paul R. Milgrom and Robert J. Weber. A theory of auctions and competitive bidding. *Econometrica*, 50(5):1089–1122, September 1982. ISSN 0012-9682. doi: 10.2307/1911865. URL `http://www.jstor.org/stable/1911865`. ArticleType: research-article / Full publication date: Sep., 1982 / Copyright Â© 1982 The Econometric Society.

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

Joshua C. Pinkston. A model of asymmetric employer learning with testable implications. *The Review of Economic Studies*, 76(1):367–394, 2009. doi: 10.1111/j.1467-937X.2008.00507.x.

Steven G. Rivkin, Eric A. Hanushek, and John F. Kain. Teachers, schools, and academic achievement. *Econometrica*, 73(2):417–458, 2005. ISSN 1468-0262.

Jonah E. Rockoff, Douglas O. Staiger, Thomas J. Kane, and Eric S. Taylor. Information and employee evaluation: Evidence from a randomized intervention in public schools. *The American Economic Review*, 102(7):3184–3213, December 2012. doi: 10.1257/aer.102.7.3184.

Roderick A. Rose, Gary T. Henry, and Douglas L. Lauen. 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, August 2012.

William L Sanders, Arnold M Saxton, and Sandra P Horn. 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*, pages 137–162, 1997.

Tim R. Sass, Jane Hannaway, Zeyu Xu, David N. Figlio, and Li Feng. Value added of teachers in high-poverty schools and lower poverty schools. *Journal of Urban Economics*, 72(2-3):104–122, September 2012. ISSN 0094-1190.

Uta Schönberg. Testing for asymmetric employer learning. *Journal of Labor Economics*, 25(4):651–691, October 2007. ISSN 0734-306X. URL `http://www.jstor.org/stable/10.1086/522905`. ArticleType: research-article / Full publication date: October 2007 / Copyright Â© 2007 The University of Chicago.

Michael Spence. Job market signaling. *The Quarterly Journal of Economics*, 87 (3):355–374, August 1973. ISSN 0033-5533. doi: 10.2307/1882010. URL http://www.jstor.org/stable/1882010. ArticleType: research-article / Full publication date: Aug., 1973 / Copyright Â© 1973 Oxford University Press.

Douglas O. Staiger and Jonah E. Rockoff. Searching for effective teachers with imperfect information. *The Journal of Economic Perspectives*, 24(3):97–117, July 2010. doi: 10.1257/jep.24.3.97.

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

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

9 Appendix

9.1 Comparative Statics

The probability of transferring schools if given by the following equation (equation 6 in text):

$$P(M) = P[b^{h*} - b^{r*} > 0]$$

9.1.1 Base Probability of Moving

For simplicity, these first derivations adopt the notation of bidding in the absence of VAMs. Substituting the hiring and retaining principals bids provides the following:

$$P(M) = P\left[\frac{\sigma_\tau(0)\sigma_\xi(x)}{Z^h_{NV}}m + \frac{\sigma_\tau(0)\sigma_\epsilon}{Z^h_{NV}}R_x + \frac{\sigma_\epsilon\sigma_\xi(x)}{Z^h_{NV}}P^h_0 - \left(\frac{\sigma_\tau(t)\sigma_\xi(x)}{Z^r_{NV}}m + \frac{\sigma_\tau(t)\sigma_\epsilon}{Z^r_{NV}}R_x + \frac{\sigma_\epsilon\sigma_\xi(x)}{Z^r_{NV}}P^r_t\right) > 0\right], \tag{13}$$

where $Z^h_{NV} = \sigma_\tau(0)\sigma_\xi(x) + \sigma_\tau(0)\sigma_\epsilon + \sigma_\epsilon\sigma_\xi(x)$ and $Z^r_{NV} = \sigma_\tau(t)\sigma_\xi(x) + \sigma_\tau(t)\sigma_\epsilon + \sigma_\epsilon\sigma_\xi(x)$

After some algebra, equation 13 becomes the following:

$$\begin{aligned}=&P\{\frac{\sigma_\xi(x)}{Z^h_{NV}Z^r_{NV}}[(m-\mu)\sigma_\xi(x)(\sigma_\tau(0)-\sigma_\tau(t)) + (\sigma_\epsilon\sigma_\tau(t) + \sigma_\xi(x)\sigma_\tau(t) + \sigma_\epsilon\sigma_\xi(x))\tau^h_0 \\ &- (\sigma_\epsilon\sigma_\tau(0) + \sigma_\xi(x)\sigma_\tau(0) + \sigma_\epsilon\sigma_\xi(x))\tau^r_t + \sigma_\epsilon(\sigma_\tau(0)-\sigma_\tau(t))\xi] > 0\}\end{aligned}$$

Letting $\psi \equiv (\sigma_\epsilon\sigma_\tau(t) + \sigma_\xi(x)\sigma_\tau(t) + \sigma_\epsilon\sigma_\xi(x))\tau^h_0 - (\sigma_\epsilon\sigma_\tau(0) + \sigma_\xi(x)\sigma_\tau(0) + \sigma_\epsilon\sigma_\xi(x))\tau^r_t + \sigma_\epsilon(\sigma_v - \sigma_\tau(t))\xi$, be the composite error term, simplifies the above, to equation 9.1.1 from within text, presented below:

$$P(M) = P\left\{\psi > \sigma_{\xi}(x)[\sigma_{\tau}(0) - \sigma_{\tau}(t)](\mu - m)\right\}.$$

Under the assumptions that τ^r, τ^h and ξ are each orthogonal to one another,

$$\begin{aligned}
\sigma_{\psi} \equiv var(\psi) =& var[(\sigma_{\epsilon}\sigma_{\tau}(t) + \sigma_{\xi}(x)\sigma_{\tau}(t) + \sigma_{\epsilon}\sigma_{\xi}(x))\tau_0^h \\
&- (\sigma_{\epsilon}\sigma_{\tau}(0) + \sigma_{\xi}(x)\sigma_{v} + \sigma_{\epsilon}\sigma_{\xi}(x))\tau_t^r + \sigma_{\epsilon}(\sigma_{\tau}(0) - \sigma_{\tau}(t))\xi] \\
&= \sigma_{\tau}(t)(\sigma_{\epsilon}\sigma_{\tau}(0) + \sigma_{\xi}(x)\sigma_{\tau}(0) + \sigma_{\epsilon}\sigma_{\xi}(x))^2 \\
&+ \sigma_{\tau}(0)(\sigma_{\epsilon}\sigma_{\tau}(t) + \sigma_{\xi}(x)\sigma_{\tau}(t) + \sigma_{\epsilon}\sigma_{\xi}(x))^2 + \sigma_{\xi}(x)\sigma_{\epsilon}^2(\sigma_{\tau}(0) - \sigma_{\tau}(t))^2
\end{aligned} \tag{14}$$

Assuming normality of the error terms, the probability of a school-to-school transition may be written as:

$$\begin{aligned}
P(M) &= \Phi\left\{\frac{-1}{\sqrt{\sigma_{\psi}}}\left[\sigma_{\xi}(x)[\sigma_{\tau}(0) - \sigma_{\tau}(t)](\mu - m)\right]\right\} \\
&= \Phi\left\{-\beta_{xt}(\mu - m)\right\}.
\end{aligned} \tag{15}$$

9.1.2 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 $\frac{\partial E[b^{h*}_{HV} - b^{r*}_{HV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ \mu]}{\partial \mu}$. Here, the subscript HV denotes that hiring principals may access a teacher's VAM, while the subscript NV denotes that there are no VAMs informing the bidding. I present the conditional expectation in equation 16 below.

$$
\begin{aligned}
E[b^{h*}_{HV} - b^{r*}_{HV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ \mu] =& \frac{\sigma_\tau(0)\sigma_\xi(x\ V)}{Z^h_{HV}}m + \frac{\sigma_\tau(0)\sigma_\epsilon}{Z^h_{HV}}\mu + \frac{\sigma_\epsilon\sigma_\xi(x\ V)}{Z^h_{HV}}\mu \\
&- \left(\frac{\sigma_\tau(t)\sigma_\xi(x\ V)}{Z^r_{HV}}m + \frac{\sigma_\tau(t)\sigma_\epsilon}{Z^r_{HV}}\mu + \frac{\sigma_\epsilon\sigma_\xi(x\ V)}{Z^r_{HV}}\mu\right) \\
&- \left(\frac{\sigma_\tau(0)\sigma_\xi(x)}{Z^h_{NV}}m + \frac{\sigma_\tau(0)\sigma_\epsilon}{Z^h_{NV}}\mu + \frac{\sigma_\epsilon\sigma_\xi(x)}{Z^h_{NV}}\mu\right) \\
&+ \left(\frac{\sigma_\tau(t)\sigma_\xi(x)}{Z^r_{NV}}m + \frac{\sigma_\tau(t)\sigma_\epsilon}{Z^r_{NV}}\mu + \frac{\sigma_\epsilon\sigma_\xi(x)}{Z^r_{NV}}\mu\right).
\end{aligned}
\tag{16}
$$

Taking the derivative of equation 16 with respect to μ gives the following:

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

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

$$\begin{aligned}\frac{\partial E[b_{HV}^{h*} - b_{HV}^{r*} - (b_{NV}^{h*} - b_{NV}^{r*})|m\,\mu]}{\partial \mu} &= \frac{{\sigma_\epsilon}^2(\sigma_\tau(0) - \sigma_\tau(t))(\sigma_\xi(x) - \sigma_\xi(x\,V))}{Z_{NV}^h Z_{NV}^r Z_{HV}^h Z_{HV}^r} \\ &[\sigma_\xi(x)\sigma_\xi(x\,V)(2\sigma_\tau(t)\sigma_\tau(0) + \sigma_\epsilon\sigma_\tau(0) + \sigma_\tau(t)\sigma_\epsilon) \\ &+ (\sigma_\xi(x\,V) + \sigma_\xi(x))\sigma_\tau(t)\sigma_\epsilon\sigma_\tau(0))].\end{aligned} \tag{17}$$

The above appears as equation 7 in text. $\frac{1}{Z_{HV}^h Z_{HV}^r Z_{NV}^h Z_{NV}^r}$ 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 VAMs are at all informative, lemma 2 shows that $\sigma_\xi(x V) - \sigma_\xi(x) < 0$. All other terms are positive variances, which implies that $\frac{\partial E[b_{HV}^{h*} - b_{HV}^{r*} - (b_{NV}^{h*} - b_{NV}^{r*})|m\,\mu]}{\partial \mu} > 0$, which in turn implies that the probability of moving within-district increases with increases in μ.

9.1.3 Comparative statics for within-district moves with respect to VAMs (V)

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

$$\begin{aligned}
b^{h*}_{HV}-b^{r*}_{HV} =& \frac{\sigma_\tau(0)\sigma_\xi(x\,V)}{Z^h_{HV}}m+\frac{\sigma_\tau(0)\sigma_\epsilon}{Z^h_{HV}}R_{x\nu}+\frac{\sigma_\epsilon\sigma_\xi(x\,V)}{Z^h_{HV}}P^h_0\\
&-\left(\frac{\sigma_\tau(t)\sigma_\xi(x\,V)}{Z^r_{HV}}m+\frac{\sigma_\tau(t)\sigma_\epsilon}{Z^r_{HV}}R_{x\nu}+\frac{\sigma_\epsilon\sigma_\xi(x\,V)}{Z^r_{HV}}P^r_t\right)\\
=& \frac{1}{Z^h_{HV}Z^r_{HV}}[\sigma_\xi(x\,V)\sigma_\epsilon(\sigma_\tau(0)-\sigma_\tau(t))(\sigma_\xi(x\,V)(m-\mu)+\sigma_\epsilon\frac{\sigma_\nu\xi+\sigma_\xi(x)\nu}{\sigma_\nu+\sigma_\xi(x)})\\
&+\tau^hZ^r_{HV}\sigma_\xi(x\,V)\sigma_\epsilon-\tau^r_tZ^h_{HV}\sigma_\xi(x\,V)\sigma_\epsilon]
\end{aligned}$$

Substituting in the VAM (V) and prior public signal (R_x) separately provides equation 18

$$\begin{aligned}
=&\frac{1}{Z^h_{HV}Z^r_{HV}}[\sigma_\xi(x\,V)\sigma_\epsilon(\sigma_\tau(0)-\sigma_\tau(t))(\sigma_\xi(x\,V)(m-(1+\sigma_\epsilon)\mu)+\sigma_\epsilon\frac{\sigma_\nu R_x+\sigma_\xi(x)V}{\sigma_\nu+\sigma_\xi(x)})\\
&+\tau^hZ^r_{HV}\sigma_\xi(x\,V)\sigma_\epsilon-\tau^r_tZ^h_{HV}\sigma_\xi(x\,V)\sigma_\epsilon]
\end{aligned}\tag{18}$$

Turning back to the probability of moving in absence of VAMs,

$$\begin{aligned}
b^{h*}_{NV}-b^{r*}_{NV} =& \frac{\sigma_\tau(0)\sigma_\xi(x)}{Z^h_{NV}}m+\frac{\sigma_\tau(0)\sigma_\epsilon}{Z^h_{NV}}R_x+\frac{\sigma_\epsilon\sigma_\xi(x)}{Z^h_{NV}}P^H_0\\
&-\left(\frac{\sigma_\tau(t)\sigma_\xi(x)}{Z^r_{NV}}m+\frac{\sigma_\tau(t)\sigma_\epsilon}{Z^r_{NV}}R_x+\frac{\sigma_\epsilon\sigma_\xi(x)}{Z^r_{NV}}P^R_t\right)\\
=& \frac{1}{Z^h_{NV}Z^r_{NV}}[\sigma_\xi(x)\sigma_\epsilon(\sigma_\tau(0)-\sigma_\tau(t))(\sigma_\xi(x)(m-\mu)+\sigma_\epsilon\xi)\\
&+\tau^hZ^r_{NV}\sigma_\xi(x)\sigma_\epsilon-\tau^r_tZ^h_{NV}\sigma_\xi(x)\sigma_\epsilon]
\end{aligned}\tag{19}$$

Combining equation 18 with equation 19 and taking the expectation conditional on prior beliefs and VAMs provides equation 20:

$$\begin{aligned}E[b^{h*}_{HV} - b^{r*}_{HV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ V\ \mu] =& \frac{1}{Z^h_{HV} Z^r_{HV}}[\sigma_\xi(x\ V)\sigma_\epsilon(\sigma_\tau(0) - \sigma_\tau(t))(\sigma_\xi(x\ V) \\ & (m - (1+\sigma_\epsilon)\mu) + \sigma_\epsilon \frac{\sigma_\nu \mu + \sigma_\xi(x)V}{\sigma_\nu + \sigma_\xi(x)})] \\ & - \frac{1}{Z^h_{NV} Z^r_{NV}} \sigma_\xi(x)\sigma_\epsilon(\sigma_\tau(0) - \sigma_\tau(t))(\sigma_\xi(x)(m - \mu)\end{aligned} \tag{20}$$

Taking the derivative with respect to VAMs (V) provides equation 8 from the text.[42]

$$\frac{\partial E\left[b^{h*}_{HV} - b^{r*}_{HV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ V\ \mu\right]}{\partial V} = \frac{1}{Z^h_{HV} Z^r_{HV}} \frac{\sigma_\xi(x)}{\sigma_\nu + \sigma_\xi(x)} > 0$$

As $\frac{1}{Z^h_{HV} Z^r_{HV}} \frac{\sigma_\xi(x)}{\sigma_\nu + \sigma_\xi(x)}$ is function of variances, it must be positive. Meaning that releasing VAMs raises the probability that high-VAM teachers move schools.

9.1.4 Comparative statics for out-of-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 $\frac{\partial P[b^{h*}_{NV} - b^{r*}_{RV} > 0|m\ \mu] - P[b^{h*}_{NV} - b^{r*}_{NV} > 0|m\ \mu]}{\partial \mu}$ is implied by the sign of $\frac{\partial E[b^{h*}_{NV} - b^{r*}_{RV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ \mu]}{\partial \mu}$. Here, the subscript RV denotes that only retaining principals may access a teacher's VAM, while the subscript NV denotes that there are no VAMs 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 VAMs. Letting $Z^r_{RV} = \sigma_\xi(x)\sigma_\tau(t\ V) + \sigma_\epsilon\sigma_\tau(t\ V) + \sigma_\epsilon\sigma_\xi(x)$, equation 21 gives the conditional expectation of this difference.

$$\begin{aligned}E[b^{h*}_{NV} - b^{r*}_{RV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\ \mu] =& \frac{\sigma_\tau(t)\sigma_\xi(x)}{Z^r_{NV}} m + \frac{\sigma_\tau(t)\sigma_\epsilon}{Z^r_{NV}}\mu + \frac{\sigma_\epsilon\sigma_\xi(x)}{Z^r_{NV}}\mu \\ & - \left(\frac{\sigma_\tau(t\ V)\sigma_\xi(x)}{Z^r_{RV}} m + \frac{\sigma_\tau(t\ V)\sigma_\epsilon}{Z^r_{RV}}\mu + \frac{\sigma_\epsilon\sigma_\xi(x)}{Z^r_{RV}}\mu\right)\end{aligned} \tag{21}$$

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

Taking the derivative with respect to μ gives:

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

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

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

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

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).

9.1.5 Comparative statics for out-of-district moves with respect to VAMs (V)

In determining the comparative statics with regard to the VAM signal, I seek to sign $\frac{\partial E[b^{h*}_{RV} - b^{r*}_{RV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\,V\,\mu]}{\partial V}$. Turning back to the probability of moving in absence of VAMs, equation 19 provides:

$$b^{h*}_{NV} - b^{r*}_{NV} = \frac{1}{Z^h_{NV}Z^r_{NV}}[\sigma_\xi(x)\sigma_\epsilon(\sigma_\tau(0) - \sigma_\tau(t))(\sigma_\xi(x)(m-\mu) + \sigma_\epsilon\xi) + \tau^h Z^r_{NV}\sigma_\xi(x)\sigma_\epsilon$$
$$- \tau^r_t Z^h_{NV}\sigma_\xi(x)\sigma_\epsilon]$$

In the case where only retaining principals may access a teacher's VAM, as is plausible for out-of-district moves, the difference between hiring and retaining principals bids is given by equation 22:

$$
\begin{aligned}
b^{h*}_{RV} - b^{r*}_{RV} =& \frac{\sigma_\tau(0)\sigma_\xi(x)}{Z^r_{RV}} m + \frac{\sigma_\tau(0)\sigma_\epsilon}{Z^r_{RV}} R_x + \frac{\sigma_\epsilon\sigma_\xi(x)}{Z^r_{RV}} P^h_0 \\
& - \left(\frac{\sigma_\tau(t\,V)\sigma_\xi(x)}{Z^r_{RV}} m + \frac{\sigma_\tau(t\,V)\sigma_\epsilon}{Z^r_{RV}} R_x + \frac{\sigma_\epsilon\sigma_\xi(x\,V)}{Z^r_{RV}} P^r_{t\nu} \right) \\
& = \frac{1}{Z^h_{RV} Z^r_{RV}} [\sigma_\xi(x)\sigma_\epsilon(\sigma_\tau(0) - \sigma_{\tau\,\nu}(t\,V))(\sigma_\xi(x)(m-\mu) + \sigma_\epsilon \xi) \\
& + \tau^h Z^r_{HV} \sigma_\xi(x)\sigma_\epsilon - \sigma_\xi(x)\sigma_\epsilon Z^h_{RV} \frac{\sigma_\nu \tau^r_t + \sigma_\tau(t)\nu}{\sigma_\nu + \sigma_\tau(t)}] \\
& = \frac{1}{Z^h_{RV} Z^r_{RV}} [\sigma_\xi(x)\sigma_\epsilon(\sigma_\tau(0) - \sigma_{\tau\,\nu}(t\,V))(\sigma_\xi(x)(m-\mu) + \sigma_\epsilon \xi) \\
& + \tau^h Z^r_{RV} \sigma_\xi(x)\sigma_\epsilon - \sigma_\xi(x)\sigma_\epsilon Z^h_{RV} \frac{\sigma_\nu \tau^r_t + \sigma_\tau(t)(V-\mu)}{\sigma_\nu + \sigma_\tau(t)}]
\end{aligned}
\tag{22}
$$

The derivative of equation 22 with respect to the VAM signal (V) is referred to in text as equation 10, and is presented below:

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

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

9.1.6 Informed out-of-district principals

It is important to note that good (or high-VAM) teachers may choose to reveal their VAMs 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-VAM 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 VAM. Equation 23 gives the conditional expectation of this difference.

$$
\begin{aligned}
&E[\delta_d(b^{h*}_{HV} - b^{r*}_{HV}) + (1-\delta_d)(b^{h*}_{NV} - b^{r*}_{RV}) - (b^{h*}_{NV} - b^{r*}_{NV})|m\ \mu] = \\
&\delta_d \frac{\sigma_\tau(0)\sigma_\xi(x\ V)}{Z^h_{HV}} m + \frac{\sigma_\tau(0)\sigma_\epsilon}{Z^h_{HV}}\mu + \frac{\sigma_\epsilon\sigma_\xi(x\ V)}{Z^h_{HV}}\mu - \delta_d\left(\frac{\sigma_\tau(t)\sigma_\xi(x\ V)}{Z^r_{HV}}m + \frac{\sigma_\tau(t)\sigma_\epsilon}{Z^r_{HV}}\mu + \frac{\sigma_\epsilon\sigma_\xi(x\ V)}{Z^r_{HV}}\mu\right) \\
&-(1-\delta_d)\left(\frac{\sigma_\tau(t\ V)\sigma_\xi(x)}{Z^r_{RV}}m + \frac{\sigma_\tau(t\ V)\sigma_\epsilon}{Z^r_{RV}}\mu + \frac{\sigma_\epsilon\sigma_\xi(x)}{Z^r_{RV}}\mu\right) \\
&-\delta_d\left(\frac{\sigma_\tau(0)\sigma_\xi(x)}{Z^h_{NV}}m + \frac{\sigma_\tau(0)\sigma_\epsilon}{Z^h_{NV}}\mu + \frac{\sigma_\epsilon\sigma_\xi(x)}{Z^h_{NV}}\mu\right) + \left(\frac{\sigma_\tau(t)\sigma_\xi(x)}{Z^r_{NV}}m + \frac{\sigma_\tau(t)\sigma_\epsilon}{Z^r_{NV}}\mu + \frac{\sigma_\epsilon\sigma_\xi(x)}{Z^r_{NV}}\mu\right).
\end{aligned}
\tag{23}
$$

Taking the derivative of equation 23 with respect to μ gives the weighted average of equations 7 and 9 where each is weighted by δ_d and $(1-\delta_d)$ respectively. This derivative is shown in equation 24.

$$
\begin{aligned}
&\frac{\partial E[\delta_d(b^{h*}_{HV} - b^{r*}_{HV}) + (1-\delta_d)(b^{h*}_{NV} - b^{r*}_{RV}) - (b^{h*}_{NV} - b^{r*}_{NV})|m\ \mu]}{\partial\mu} = \\
&\delta_d \frac{{\sigma_\epsilon}^2(\sigma_\tau(0)-\sigma_\tau(t))(\sigma_\xi(x)-\sigma_\xi(x\ V))}{Z^h_{NV}Z^r_{NV}Z^h_{HV}Z^r_{HV}}[\sigma_\xi(x)\sigma_\xi(x\ V)(2\sigma_\tau(t)\sigma_\tau(0) + \sigma_\epsilon\sigma_\tau(0) + \sigma_\tau(t)\sigma_\epsilon) \\
&+(\sigma_\xi(x\ V) + \sigma_\xi(x))\sigma_\tau(t)\sigma_\epsilon\sigma_\tau(0))] + (1-\delta_d)\frac{\sigma_\epsilon{\sigma_\xi(x)}^2(\sigma_\tau(t\ V)-\sigma_\tau(t))}{Z^r_{NV}Z^r_{RV}}.
\end{aligned}
\tag{24}
$$

Equation 25 shows that taking the derivative of equation 24 with respect to δ_d demonstrates that as the share of informed principals increases the probability that good teachers increases as well.

$$
\begin{aligned}
&\frac{\partial^2 E[\delta_d(b^{h*}_{HV} - b^{r*}_{HV}) + (1-\delta_d)(b^{h*}_{NV} - b^{r*}_{RV}) - (b^{h*}_{NV} - b^{r*}_{NV})|m\ \mu]}{\partial\mu\partial\delta_d} = \\
&\frac{{\sigma_\epsilon}^2(\sigma_\tau(0)-\sigma_\tau(t))(\sigma_\xi(x)-\sigma_\xi(x\ V))}{Z^h_{NV}Z^r_{NV}Z^h_{HV}Z^r_{HV}}[\sigma_\xi(x)\sigma_\xi(x\ V)(2\sigma_\tau(t)\sigma_\tau(0) + \sigma_\epsilon\sigma_\tau(0) + \sigma_\tau(t)\sigma_\epsilon) \\
&+(\sigma_\xi(x\ V) + \sigma_\xi(x))\sigma_\tau(t)\sigma_\epsilon\sigma_\tau(0))] - \frac{\sigma_\epsilon{\sigma_\xi(x)}^2(\sigma_\tau(t\ V)-\sigma_\tau(t))}{Z^r_{NV}Z^r_{RV}} > 0.
\end{aligned}
\tag{25}
$$

As noted previously, VAMs were 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 VAM of her school in 2008, when Winston-Salem began using EVAAS VAMs of teacher effectiveness. These different settings lead the share of out-of-district principals who are informed of VAMs 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 VAMs and the probability of moving from Winston-Salem to be more positive after Winston-Salem adopts VAMs than is the relationship between VAMs and the probability of moving from Guilford after Guilford adopts VAMs. 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' VAMs whereas in Guilford they did. However, it is likely that principals still inferred something when a teacher chose not to reveal his VAM. If the share of informed principals was lower within Winston-Salem than within Guilford ($\delta_{WSF} < \delta_{GCS}$), A safer prediction may be, $\gamma_{14WD_{GCS}} - \gamma_{14OD_{GCS}} > \gamma_{14WD_{WSF}} - \gamma_{14OD_{WSF}}$.

9.1.7 Comparative statics with respect to VAMs (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 $E\left[\rho^h b_{HV}^{h*} - \rho^r b_{HV}^{r*} - (\rho^h b_{NV}^{h*} - \rho^r b_{NV}^{r*})|m\ V\ \mu\right]$ with respect to VAMs (V) and S^s. I present these cross partials below.

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

$$\frac{\partial^2 E\left[\rho^h b_{HV}^{h*} - \rho^r b_{HV}^{r*} - (\rho^h b_{NV}^{h*} - \rho^r b_{NV}^{r*})|m\, V\, \mu\right]}{\partial V \partial S^h} = \frac{\partial \rho^h}{\partial S^h} \frac{\sigma_\tau(0)\sigma_\epsilon\sigma_\xi(x)}{Z_{HV}^h(\sigma_v + \sigma_\xi(x))} \tag{26}$$

As everything else is a function of variances, $\frac{\partial \rho^h}{\partial S^h} > 0$ implies that equation 26 is positive.

$$\frac{\partial^2 E\left[\rho^h b_{HV}^{h*} - \rho^r b_{HV}^{r*} - (\rho^h b_{NV}^{h*} - \rho^r b_{NV}^{r*})|m\, V\, \mu\right]}{\partial V \partial S^r} = -\frac{\partial \rho^r}{\partial S^r} \frac{\sigma_\tau(0)\sigma_\epsilon\sigma_\xi(x)}{Z_{HV}^h(\sigma_v + \sigma_\xi(x))} \tag{27}$$

Conversely, $\frac{\partial \rho^r}{\partial S^r} > 0$ implies that equation 27 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.

9.1.8 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 VAMs, taking the derivative of $E\left[b_{HV}^{h*} - b_{HV}^{r*} - (b_{NV}^{h*} - b_{NV}^{r*})|m\, \mu\right]$ shown in equation 16 with respect to (m).

$$\begin{aligned}\frac{\partial E\left[b_{HV}^{h*} - b_{HV}^{r*} - (b_{NV}^{h*} - b_{NV}^{r*})|m\, \mu\right]}{\partial m} = &\frac{\sigma_\epsilon(\sigma_\tau(0) - \sigma_\tau(t))(\sigma_\xi(x\, V) - \sigma_\xi(x))}{Z_{HV}^r Z_{HV}^h Z_{NV}^r Z_{NV}^h} \\ &[\sigma_\tau(t)\sigma_\tau(0)\sigma_\xi(x)\sigma_\xi(x\, V) + \sigma_\xi(x\, V)\sigma_\epsilon\sigma_\xi(x)(\sigma_\tau(0) \\ &+ \sigma_\tau(t)) + (\sigma_\xi(x\, V) + \sigma_\xi(x))\sigma_\tau(t)\sigma_\epsilon\sigma_\tau(0)].\end{aligned} \tag{28}$$

Under the assumptions of prior private learning ($\sigma_\tau(0) - \sigma_\tau(t) > 0$), and informative VAMs ($\sigma_\xi(x\, V) - \sigma_\xi(x) < 0$), equation 28 implies that $\frac{\partial E[b_{HV}^{h*} - b_{HV}^{r*} - (b_{NV}^{h*} - b_{NV}^{r*})|m\, \mu]}{\partial m} < 0$. Thus, the model predicts the probability of moving after the introductions of VAMs decreases as a teacher's VAM increases, or empirically, $\gamma_{24WD} < 0$.

9.1.9 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 VAMs, taking the derivative of $E\left[b^{h*}_{HV} - b^{r*}_{HV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\,\mu\right]$ shown in equation **??** with respect to (m).

$$\frac{\partial E\left[b^{h*}_{RV} - b^{r*}_{RV} - (b^{h*}_{NV} - b^{r*}_{NV})|m\,\mu\right]}{\partial m} = \frac{\sigma_\xi(x)^2\sigma_\epsilon}{Z^h_{RV}Z^r_{RV}Z^h_{NV}Z^r_{NV}}(\sigma_\tau(t) - \sigma_\tau(t\,V))$$
$$(\sigma_\tau(0)^2\sigma_\epsilon^2 + \sigma_\tau(0)^2\sigma_\xi(x)^2 + \sigma_\tau(0)^2\sigma_\epsilon\sigma_\xi(x) + \sigma_\xi(x)^2\sigma_\epsilon^2. \tag{29}$$

Under the assumption that VAMs are 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 .

9.1.10 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 VAMs, 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 16 with respect to μ.

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

Taking the derivative of equation 30 with respect to t gives the following:

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

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

$$= \frac{\partial \sigma_\tau(t)}{\partial t} \frac{\sigma_\epsilon \sigma_\tau(t)(\sigma_\xi(x\ V) - \sigma_\xi(x))}{{Z^r_{HV}}^2 {Z^r_{NV}}^2}$$
$$2\sigma_\xi(x)\sigma_\xi(x\ V)(\sigma_\epsilon + \sigma_\tau(t)) + \sigma_\tau(t)\sigma_\epsilon(\sigma_\xi(x) + \sigma_\xi(x\ V)) \tag{31}$$

The assumptions of prior private learning $\left(\frac{\partial \sigma_\tau(t)}{\partial t} < 0\right)$ and informative VAMs ($\sigma_\xi(x\ V) < \sigma_\xi(x)$), imply that equation 31 is positive. Thus, the positive change in selection with the introduction of VAMs should be more positive for those with more tenure. Empirically, the model predicts the coefficient on the interaction between adopting VAMs, the VAMs themselves, and tenure to be positive ($VAM \times Ten \times TreatDist > 0$).

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

In order to investigate the learning environment that prevailed in the absence of VAMs, 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 VAMs (V) and years of tenure (t).

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

The assumption of prior private learning provides ($\frac{\partial \sigma_\tau(t)}{\partial t} < 0$) leads equation 32 to be positive. This means that the model predicts larger positive effects of the introduction of VAMs on the probability that high-VAM 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 $VAM \times TreatDist \times tenure$ to be positive.

9.2 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 11 is comparable to Table 4 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 4 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 VAM is 0.3 percentage point more likely to move within district to a higher ABC growth school for teachers whose VAMs 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 4 and Table 11. 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' VAMs after the policy takes effect.

9.3 Robustness: Year interactions with VAM

The primary threat to validity for difference-in-difference analysis is differential trends. The tables below provide year interactions with the VAM within both treatment districts as well as the rest of the state. 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 my results. It is also noteworthy that is both districts there is a spike in the correlation of VAM with the probability of moving within-district soon after the policy takes effect.

Table 11: Probability of moving to higher or lower growth schools

	Panal A: Within-District Moves		Panal 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**
VAM	0.0024***	-0.0006	0.0008	-0.0005
	[0.00073]	[0.00077]	[0.00056]	[0.0006]
VAM x Treatment GCS	0.0031**	0.0013	-0.0048***	-0.0052***
	[0.00152]	[0.00153]	[0.00139]	[0.002]
VAM 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.
*** p<0.01, ** p<0.05, * p<0.1

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

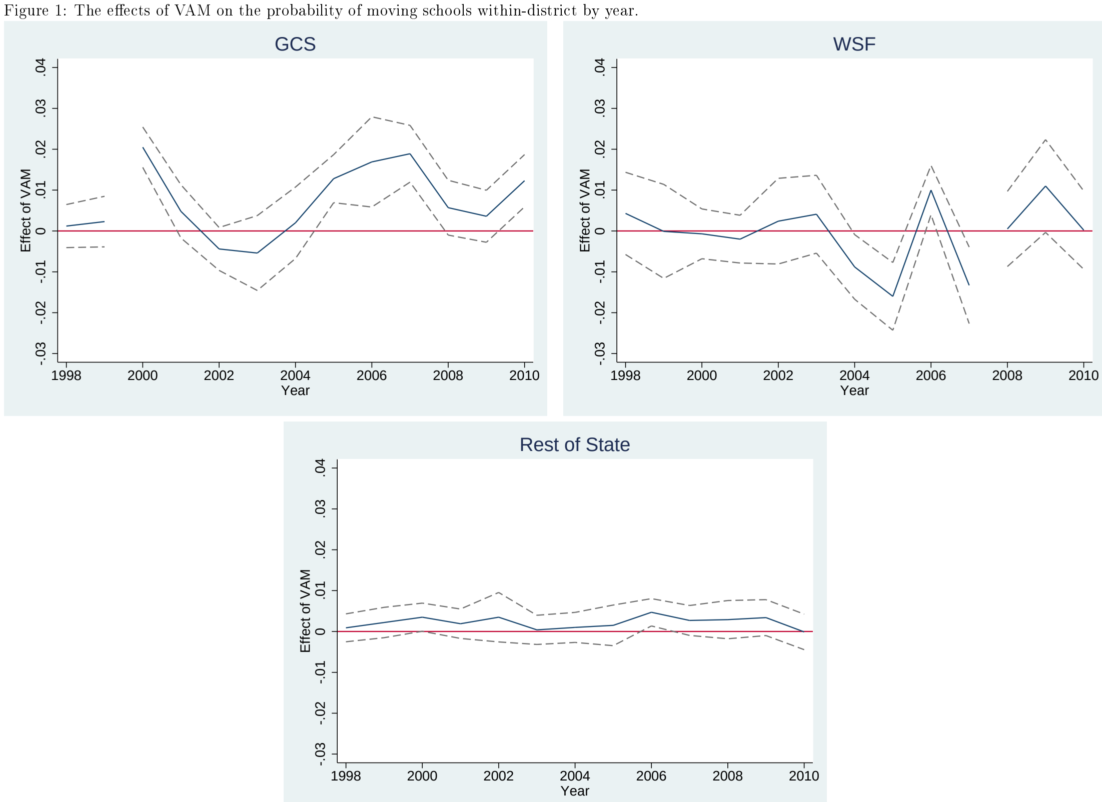

Table 12: The effects of VAM on the probability of moving schools within-district by year.

	Total			To a more proficient school		
VARIABLES	Rest of NC	Guilford	Winston-Salem	Rest of NC	Guilford	Winston-Salem
year 1998 x VAM	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 VAM	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 VAM	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 VAM	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 VAM	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 VAM	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 VAM	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 VAM	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 VAM	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 VAM	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 VAM	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 VAM	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 VAM	-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.
*** p<0.01, ** p<0.05, * p<0.1

Table 13: The effect of VAM 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 VAM	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 VAM	-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 VAM	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 VAM	-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 VAM	-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 VAM	-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 VAM	-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 VAM	-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 VAM	-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 VAM	-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 VAM	-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 VAM	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 VAM	-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.
*** p<0.01, ** p<0.05, * p<0.1

Table 14: The effect of VAM on teacher sorting within-district by year.

VARIABLES	Rest of NC	Guilford	Winston-Salem
year 1998 x VAM	0.0025***	0.0045**	-0.0014
	[0.00021]	[0.00071]	[0.00146]
year 1999 x VAM	0.0026***	0.0013	0.0021
	[0.00021]	[0.00109]	[0.00156]
year 2000 x VAM	0.0019***	0.0041***	0.0007
	[0.0002]	[0.00069]	[0.00084]
year 2001 x VAM	0.0051***	0.0038***	0.0077***
	[0.00026]	[0.00097]	[0.00146]
year 2002 x VAM	0.0046***	0.0031***	0.0072***
	[0.0002]	[0.00072]	[0.00164]
year 2003 x VAM	0.0031***	0.0043***	0.0052***
	[0.00019]	[0.00099]	[0.001]
year 2004 x VAM	0.0023***	-0.0006	0.0005
	[0.00021]	[0.00109]	[0.00212]
year 2005 x VAM	0.0102***	0.0109***	0.0096***
	[0.00032]	[0.00097]	[0.00126]
year 2006 x VAM	0.0047***	0.0009	-0.0014
	[0.00027]	[0.00161]	[0.00089]
year 2007 x VAM	0.0046***	0.0049***	0.0031**
	[0.00026]	[0.00105]	[0.00133]
year 2008 x VAM	0.0016***	0.0031***	0.0005
	[0.00025]	[0.00112]	[0.00127]
year 2009 x VAM	-0.0003	0.0055***	0.0053***
	[0.00042]	[0.00097]	[0.00146]
year 2010 x VAM	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.
*** p<0.01, ** p<0.05, * p<0.1

Figure 2: The effects of VAM on the probability of moving to a a better school within-district by year.

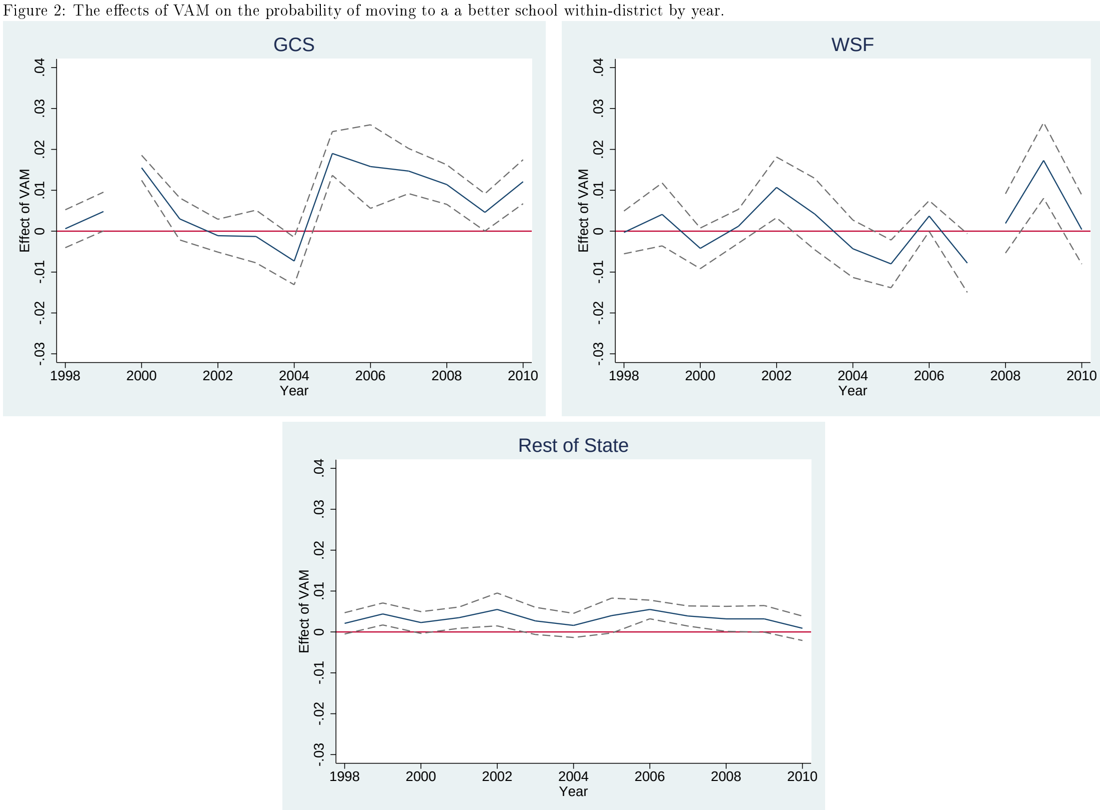

Figure 3: The effect of VAM on the probability of moving schools out-of-district by year

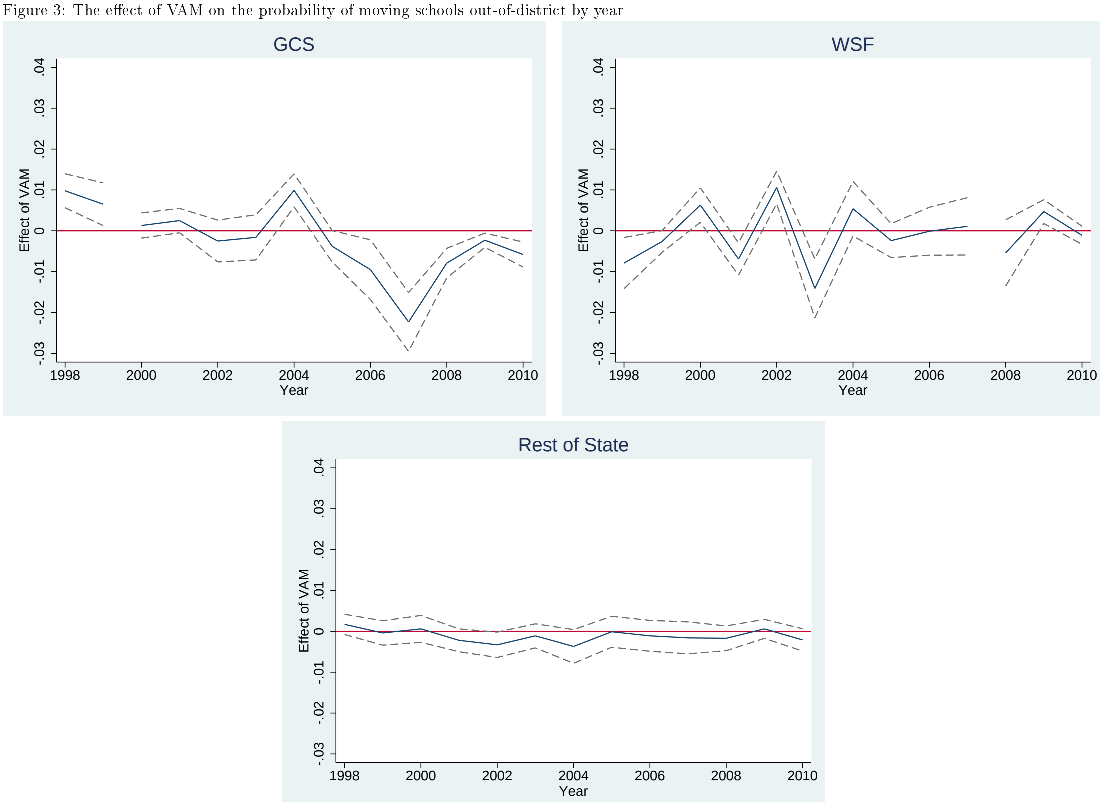

Figure 4: The effects of VAM on the probability of moving to a better school out-of-district by year

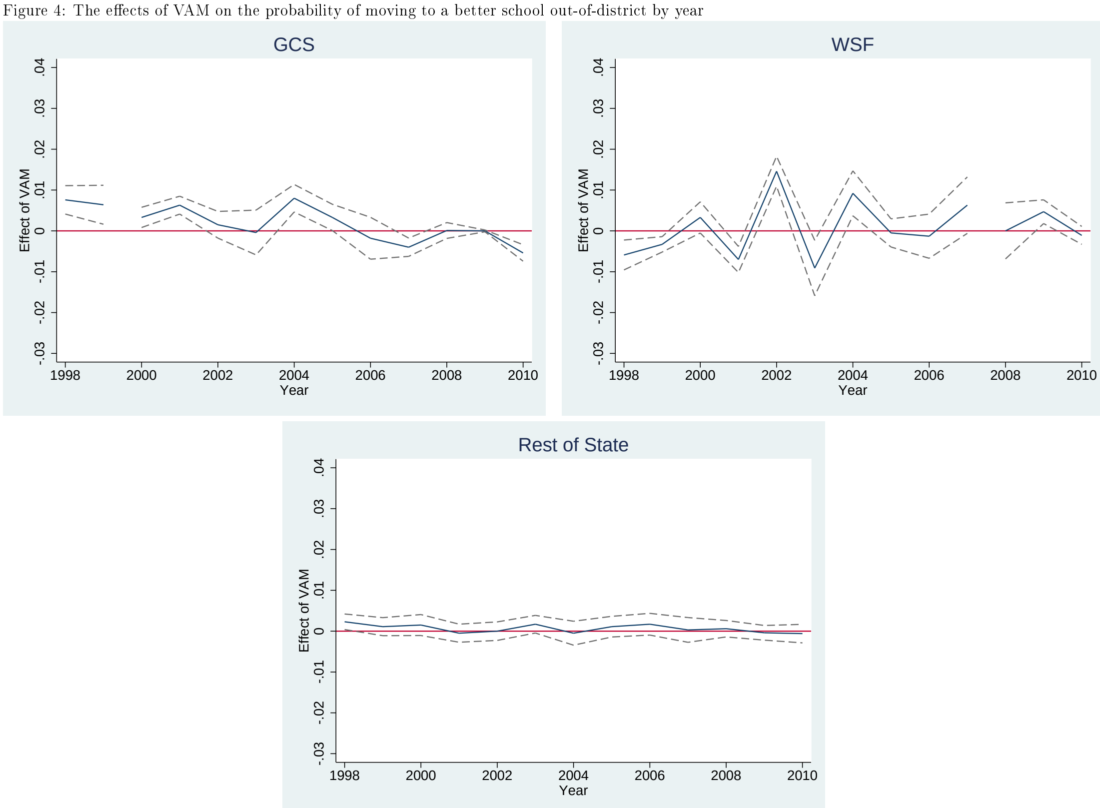

Figure 5: The effect of VAM on teacher sorting within-district by year.

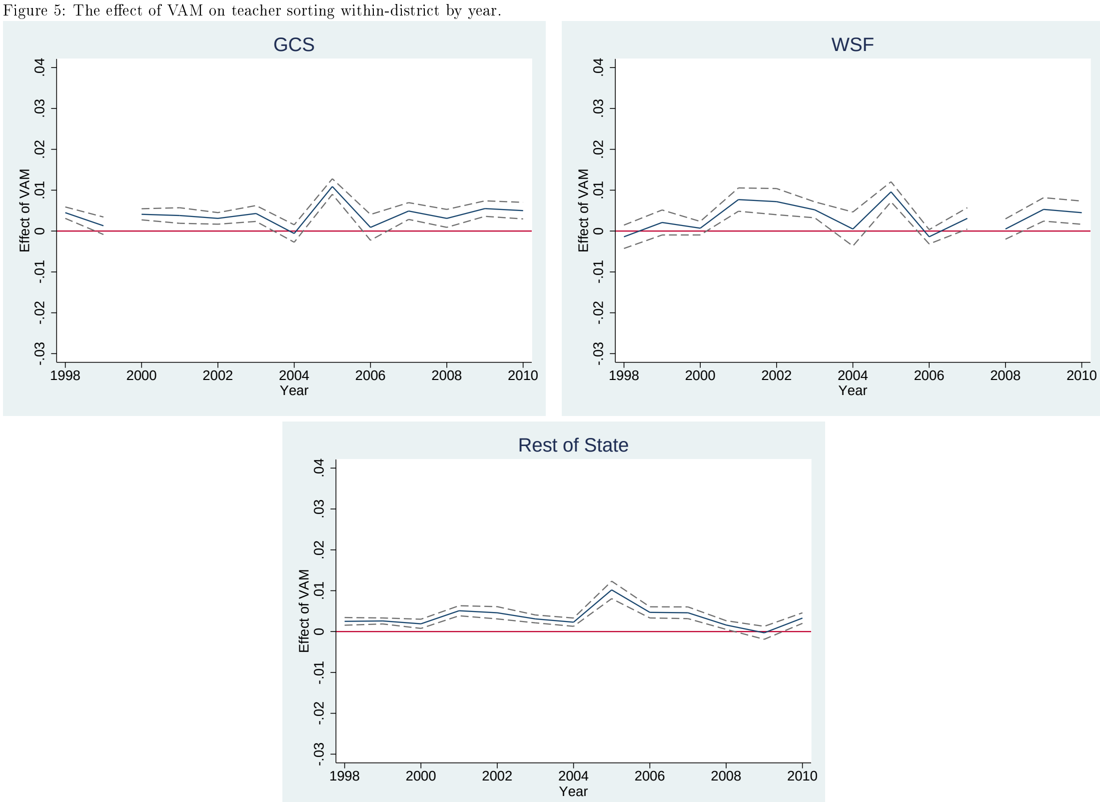

Table 15: Probability of moving schools using normal maximum likelihood estimation.

	Panal A: Within-District Moves			Panal 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	0.0022**	0.0030***	-0.0011	-0.0011	0.0005	-0.0018***
	[0.00114]	[0.00079]	[0.00068]	[0.00083]	[0.0006]	[0.0005]
VAM 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]
VAM 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.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

9.4 Normal Maximum Likelihood Estimation

The results in Table 4 are from a linear probability model, which are more straight forward both computationally and in interpretation. Taking the normality and orthogonality assumptions from Section 4 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 15 in Appendix 9.6 provides the APEs in accordance with Ai and Norton [2003]. Comparison between Table 4 and Table 15 provides very similar results.

9.5 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 ($\mathbf{Z}(t)$). [43]

$$\lambda_{WD}(t|\mathbf{Z}) = \lambda_{WD0}(t)exp\{\mathbf{Z}(t)\boldsymbol{\beta}_0\}, \tag{33}$$

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.[44] $\mathbf{Z}(t)$ is a vector including all covariates used in Table 4, 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][45].

Table 16 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 remarkably consistent. From columns 1 and 2, the probability of moving within-district for a teacher with a one standard deviation higher VAM score increases by 9% with the release of teacher VAMs, and for moves within-district

[43] 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.

[44] 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.

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

to better school, the probability increases by 13%. Both effects are significantly different from zero and are within a percentage point estimates shown in Table 4. From columns 4 and 5, a teacher with a one standard deviation lower VAM 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, the results from Table 4 are muted for total within-district mobility. Column 1 shows a smaller point estimate than appears in Table 4, and the the estimate loses statistical significance. The impact of the policy in Winston-Salem on moves to higher performing schools within district are more stable. The introduction of VAMs raises the probability that a teacher with a one standard deviation higher VAM moves to a higher-performing school by about 11%, though the significance level drops with this specification. For out-of-district moves to higher-performing schools, the point estimate corresponds with a 15% increase in the probability a high-VAM teacher moves out of Winston-Salem to a higher-performing school, though this estimate is very noisy and should be interpreted accordingly. In general, while the public and private learning results are further verified in Guilford with this competing risks analysis, the same cannot be said for Winston-Salem.

9.6 Supplemental Tables

Table 16: Changes in the marginal probability of each type of transfer between schools

	Panal A: Within-District Moves			Panal 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	**0.03**	**0.09*** **	**-0.07** **	**0.01**	**0.08** **	**-0.10** **
	[0.021]	[0.024]	[0.030]	[0.028]	[0.035]	[0.042]
VAM 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]
VAM 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.
*** p<0.01, ** p<0.05, * p<0.1

Table 17: 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
VAM	**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]
VAM 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]
VAM 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 parentheses. Bootstrapped standard errors in braces. District-cluster-bootstrapped-teacher-stratified standard errors in brackets.

Table 18: Probability of moving schools using Empirical Bayes VAM

	Panal A: Within-District Moves			Panal 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	**0.0006**	**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]
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.
*** p<0.01, ** p<0.05, * p<0.1

Table 19: Probability of moving schools using restricted-data, Empirical Bayes VAM

	Panal A: Within-District Moves			**Panal 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	0.0006	0.0028**	-0.0022**	-0.0006	0.0014*	-0.0020***
	[0.00169]	[0.00141]	[0.00093]	[0.00098]	[0.00073]	[0.00062]
VAM x Treatment GCS	0.0048	0.0059**	-0.0011	-0.0130***	-0.0078***	-0.0051***
	[0.00331]	[0.00252]	[0.00221]	[0.00232]	[0.00195]	[0.00129]
VAM x Treatment WSF	0.0066**	0.0085***	-0.0020	0.0009	0.0023	-0.0013
	[0.003]	[0.00236]	[0.00186]	[0.0023]	[0.00202]	[0.00113]
Treatment GCS	-0.0048	-0.0055	0.0007	-0.0174***	-0.0245***	0.0072*
	[0.01311]	[0.00855]	[0.01071]	[0.00515]	[0.00281]	[0.00431]
Treatment WSF	0.0553***	0.0471***	0.0082***	-0.0022	0.0144***	-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.
*** p<0.01, ** p<0.05, * p<0.1

Table 20: Mobility between non-strategic-staffing schools with respect to students' race

Panel	A: Within-District Moves		B: Out-of-District Moves		C: Growth in Percent Black	
Variables	To lower percent Black	To higher percent Black	To a lower percent Black	To higher percent Black	Total excluding strategic staffing schools	Within-district excluding strategic staffing schools
VAM	0.0021**	-0.0007	0.0009	-0.0006	-0.0018***	-0.0012***
	[0.00089]	[0.00086]	[0.00077]	[0.00059]	[0.00045]	[0.0004]
VAM x Treatment GCS	0.0025	0.0014	-0.0057***	-0.0047***	-0.0014	-0.0034***
	[0.00191]	[0.00162]	[0.00208]	[0.00133]	[0.00153]	[0.0013]
VAM x Treatment WSF	0.0086***	-0.0009	-0.0010	0.0002	-0.0079***	-0.0069***
	[0.00182]	[0.00148]	[0.00181]	[0.00125]	[0.00223]	[0.00161]
Treatment GCS	0.0072	-0.0195***	-0.0052**	-0.0108***	0.0261***	0.0216***
	[0.00515]	[0.00741]	[0.00254]	[0.00274]	[0.00317]	[0.00302]
Treatment WSF	0.0282***	0.0034	0.0007	0.0029**	-0.0086***	-0.0081**
	[0.00345]	[0.00258]	[0.00209]	[0.00147]	[0.00328]	[0.00335]
Observations	236,018	236,018	236,018	236,018	201,468	195,208

CSB standard errors from 500 repetitions appear in brackets.
All regressions include teacher level covariates and interactions with treatment indicators.
*** p<0.01, ** p<0.05, * p<0.1