Do teacher characteristics matter? New results on the effects of teacher preparation on student achievement

Sharon Kukla-Acevedo

University of Kentucky, Martin School of Public Policy & Administration, 415 Patterson Office Tower, Lexington, KY 40506, USA

1. Introduction

Economists and policy researchers are now demonstrating that teachers “matter.” After many years of research that failed to find systematic relationships between policy variables and student outcomes, recent research illustrates that individual teachers generate differential effects on students’ test scores and other outcomes. Many of these studies are based on empirical results that estimate teacher fixed effects. Rather than identifying measurable and observable characteristics of teachers, the studies use fixed effects to control for teachers and find that the fixed effects are significant in explaining student achievement (Rivkin, Hanushek, & Kain, 2005; Rockoff, 2004). Researchers and policymakers now face the challenge of identifying observable characteristics of teachers that signal quality teaching. The work in this area is extensive and varied, employing a mix of methodology, data, and units of analyses. Despite this variation, the literature is suggestive of some teacher characteristics that are important for student learning.

Recent studies generally report that teacher experience has a positive effect on student test scores (Clotfelter, Ladd, & Vigdor, 2006; Goldhaber & Anthony, 2007; Goldhaber & Brewer, 1997; Jepsen, 2005; Krueger, 1999; Noell, 2005, 2001; Rivkin et al., 2005; Rockoff, 2004; Sanders, Ashton & Wright, 2005). The positive effect also appears to be non-linear in nature as demonstrated by substantial improvements in teaching skill during the first 3–5 years in the classroom with the effects generally tapering off around the fifth year (Rivkin et al., 2005). Despite this fairly consistent result, not all studies find an association between experience and student achievement (Cooper & Cohn, 1997; Ehrenberg & Brewer, 1994; Ferguson & Ladd, 1996). While not specifically acknowledged by the authors, the methodologies employed in these studies provide one possible explanation for the lack of finding. Ehrenberg and Brewer (1994) and Cooper and Cohn (1997) use OLS estimation without fixed effects, making the estimates vulnerable to omitted variables bias. Ferguson and Ladd (1996) use Hierarchical Linear Modeling (HLM), a technique that is becoming increasingly popular in education research because its error structure accounts for the nested nature of the data. However, as noted in Jepsen (2005), it assumes that the variance in achievement is due to class-
room specific factors rather than attributing it to omitted student-level factors such as motivation. Studies that use HLM typically report smaller effect sizes than studies that employ OLS and fixed effects methods.

All of these studies make the implicit assumption that experience operates similarly for all teachers. However, it is likely that the effect of experience varies with teachers’ qualities or abilities. Rather than estimate the effect of this variable independently of other teacher attributes, this paper looks at the joint relationship between teacher experience and teacher qualifications to determine whether experience has a consistently positive effect on student achievement.

There is tenuous evidence that teachers’ content area preparation affects student learning. Using a strong value-added design that includes teacher fixed effects, Goldhaber and Brewer (1997) find that holding either a BA or MA in math has a statistically significant, positive relationship with student math achievement. Monk (1994) presents a nuanced relationship between teacher content preparation and student achievement. He finds that teacher preparation predicts student performance, but the magnitude of the positive effect varies according to subject matter and grade level. Neither of the datasets used in these papers has the capacity to link individual students to teachers, forcing the authors to aggregate to the classroom level. This prevents the authors from exploring the non-random sorting of teachers and students within and across schools, so the results could be biased in unknown ways. Using student–teacher matched data from the San Diego Unified School District, Betts, Zau, and Rice (2003) improves upon the design of these prior studies by including student fixed effects to mitigate omitted variables bias. The study fails to detect a systematic relationship between content area preparation and student achievement, but the generalizability of these findings must be considered since the data represent only one school district in the U.S. No nationally representative dataset contains measures of teacher content preparation and matches students to their teachers over time, so it is important to explore the role of content area preparation in another geographic region of the country.

There are other reasons to examine content area preparation further. Teachers’ skill and knowledge are important factors to consider when measuring the impact of teacher inputs on student achievement, but data limitations typically force researchers to use proxies like number of college courses taken and degree attainment to capture these dimensions. While these proxies should be positively correlated with content knowledge, they may not reflect teachers’ ability to transfer knowledge in the classroom. This paper improves upon past research by including several variables that indicate teacher performance during pre-service training—overall GPA, math GPA, and math education GPA. All else equal, a high achieving college student is likely to be a high achieving teacher.

2. Data and measures

This paper uses unique data from a school district in Kentucky that matched individual teachers to 5th grade math students. They were compiled with the cooperation of the district and the Kentucky Education Professional Standards Board (EPSB). EPSB compiles annual data on all teachers in the Commonwealth and also maintains detailed records of the teachers’ pre-service training. The agency provided 5th grade data for the 2000–2001, 2001–2002 and 2002–2003 academic years. After accounting for missing information, the dataset contains 3812 students, 46 schools, and 120 teachers.

The outcome measure is the standardized 5th grade math test score from the state testing program, Kentucky Core Content Test (KCCT). The KCCT is a criterion-referenced test that assesses individual student performance against a specified set of state educational goals and consists of both multiple-choice and open-response questions. The test scores are converted to statewide grade-by-year Z-scores with a mean of 0 and standard deviation of 1. Kentucky does not test the same subject in subsequent years, therefore, the students’ KCCT reading score is used as a measure of prior achievement in the analyses. The average math score is 0.08 with a standard deviation of 1.07, suggesting that this sample of students performs slightly higher than other 5th grade math students in the state.

Demographic information on the students, including gender and race, is included in the models. Indicator variables specify whether the student is female, African American, Latino/a, Asian American, or other. Male and European American students provide the reference category. Students also report on subsidized lunch status, allowing the creation of three variables that indicate whether a student receives federally subsidized lunch, partially subsidized lunch, or does not qualify for subsidized lunch. Table 1 provides means and standard deviations for the student and teacher characteristics. The table indicates a racially diverse district with 62.4% European American students and 32.6% African American students. Asian American and Latino/a students combined constitute roughly 3%, but these are both growing segments of the population in this district. Female students make up 50.8% of the population, 48.9% receive some form of subsidized lunch.

The dataset contains detailed information on the teachers’ college coursework and GPAs. The numbers of math content and math education hours taken during pre-service training are included as distinct variables in the models. GPA is separated into overall, math content, and math education hours taken during pre-service training—overall GPA, math GPA, and math education GPA. All else equal, a high achieving college student is likely to be a high achieving teacher.

---

1 The use of a reading test score as a measure of prior achievement for a math outcome is fairly unconventional. However, Eberists and Hollebeck (2001) look at the properties of different subject area tests and determine that this is a feasible option in value-added models, such as those employed in this paper. Additional sensitivity analyses are conducted to determine whether the effects of content area preparation and pre-service performance are robust to alternative specifications.
relationship between experience and teacher attributes, the coursework and GPA variables are multiplied by years of experience. These five interaction terms show the effects of teacher attributes on student achievement over time. Teacher demographic variables are incorporated in the same fashion as the student demographic variables, with male and European American teachers serving as the omitted categories. On average, teachers take slightly more math content hours than math education hours and they earn higher GPAs in math education courses than in math content courses. An overwhelming majority of teachers in the sample are European American, female, and have about 16 years of experience.

The dataset also includes information on school characteristics. The percent of students in the school that are European American, African American, Asian American, Latino/a, or some other race, and the percent of students that receive federally subsidized lunch are incorporated into the models to control for the effect of school composition on student achievement.

3. Empirical results

Non-random sorting occurs in this district, so OLS models will provide biased estimates of the teacher characteristics. To control for time-invariant, unmeasured variables, such as motivation or parental support, a series of fixed effects models are employed. Using this technique, the variation within students is used to estimate the effects of teacher qualifications. The appropriate interpretations of these coefficients relate changes in a teacher’s level of a particular qualification to the change in a single student’s academic performance. The fixed effects model can be expressed as:

\[ A_t - A_{t-1} = \Delta A_t = \beta_1 \text{Stu}_{it} + \beta_2 \text{Tch}_{ijmt} + \beta_3 \text{Sch}_{mt} + \gamma_i + \delta_j + \lambda_m + u_t \]

where \( A_t \) and \( A_{t-1} \) are measures of student achievement, \( \text{Stu}_{it} \) is a vector of student characteristics, \( \text{Tch}_{ijmt} \) is a vector of teacher characteristics, such as years of experience, college GPA and coursework, \( \text{Sch}_{mt} \) is a vector of school characteristics, including the racial and socioeconomic composition of the school, and \( u_t \) is an error term. The subscripts denote students (i), teachers (j), schools (m) and time (t), while \( \gamma_i \) is a student fixed effect, \( \delta_j \) is a teacher fixed effect, and \( \lambda_m \) is a school fixed effect. To eliminate the three levels of fixed effects, the student effect is demeaned, while the teacher and school effects are modeled by including indicator variable regressors. Finally, standard errors are clustered by teacher to account for the nested nature of the data.

Of primary interest is the estimation of \( \beta_2 \), which, if correctly modeled, can be interpreted as the impact of teacher qualifications on subsequent achievement. Modeling student achievement is challenging because of the threat of omitted variables bias, which is likely to arise if unobserved family or student characteristics are correlated with teacher ability and student achievement. The model guards against this by controlling for unmeasured time-invariant factors.

The fixed effects model is run on the full sample of students, as well as several subgroups. African American and European American students are analyzed separately because there is evidence that these two groups respond differently to school and teacher inputs. Ehrenberg and Brewer (1994) find that teachers’ college selectivity affects the achievement scores of European American students only, while advanced degrees impact the achievement scores of African American students. The peer effects literature provides motivation for looking at the effects of teacher ability on groups of students categorized by achievement levels. This literature indicates that the composition of students in a classroom has implications for learning, especially for certain groups of students (Hoxby, 2000). Low-performing students are affected by their peers’ characteristics more than high-performing students, so classrooms composed of a diverse set of abilities bene-

---

2 Average years of teaching experience is slightly higher than is found in other datasets. Using statewide data from North Carolina, Goldhaber and Anthony (2007) report average years of teaching experience to be roughly 13 years; Jepsen (2005) shows average experience to be roughly 14 years in the Prospects dataset, and in the Longitudinal Survey of American Youth. However, Monk (1994) reports average teacher experience to be just over 16 years, similar to that found in this sample.

3 In the interest of space, results are not included here. Empirical results are available from the author upon request.
fit students at the bottom of the distribution, while those at the top remain largely unaffected (Zimmer & Toma, 2000). It is reasonable to extend this argument to determine whether teacher characteristics have differential impacts on below-average and above-average performing students, as measured by their prior test scores. Since family income is generally a predictor of student performance, the model is also run separately on students that receive federally subsidized lunch and that are not eligible.

Results from the fixed effects models are listed in Table 2. The direct effects of overall GPA, math GPA, math hours, math education GPA, and math education hours are not listed in the table because fixed effects regression does not provide estimates for time-invariant variables. The second part of the analysis incorporates between effects models to provide estimates for time-invariant variables. The second listed in the table because fixed effects regression does not arrive at conclusions about the direct and marginal effects of the time-invariant factors. The positive coefficients on experience tend to have negative coefficients, suggesting a reductive effect as teachers gain experience. Math GPA × experience is generally positive across the groups, indicating that teachers with more math hours outperform other teachers with each additional year of experience.

The interaction terms and their component variables should not be interpreted individually due to the extrapolation to unlikely scenarios. For example, the coefficient on experience demonstrates the effect of experience when a teacher’s GPA, math GPA, math education GPA, math hours, and math education hours are equal to zero. In this sample, the lowest of all GPAs is 1.0 and the minimum number of math and math education hours is 3. Joint tests of hypotheses must be conducted to determine if the suite of variables containing the interaction term is jointly equal to zero instead of the more common case that concludes whether an individual coefficient is equal to zero. Table 3 lists the p-values from the F-tests of joint significance and reveals that teacher characteristics are predictive of achievement for every student group. Further, all six teacher qualifications are significant for the pooled, African American, subsidized lunch, regular lunch and above-average performing samples.

An important step after identifying the teacher qualities that affect achievement is to quantify the magnitude of the effect. This is explored by calculating the change in student math achievement associated with a standard deviation increase in the teacher characteristic. The marginal effects for teacher experience are listed in Table 4. The magnitude, sign, and statistical significance are similar across the seven groups. The effect size of approximately one-tenth of a standard deviation is similar to the size of teacher effects reported in the literature, although most studies report a positive coefficient for teacher experience. Further analysis shows that, in this sample, the positive effect of experience

---

4 Since GPA, math hours, math GPA, math education hours, and math education GPA drop out of the fixed effects estimation, marginal effects cannot be provided for these variables.
peaks at 14 years and then begins to negatively influence student learning. The largest effects of experience are found for students who are above-average performing (−0.158) and African American (−0.150). This result lends no support to the peer effects hypothesis that teacher effects operate most strongly for students who are African American, receive subsidized lunch, and are below-average performing.

While fixed effects models are generally acknowledged as a strong estimation technique for education production research, the method is not without its shortcomings. Specifically, the estimates are confounded by time-varying factors that are constant across students and they do not provide coefficient estimates for time-invariant factors of interest. Between effects models remedy these challenges. This strategy provides estimates for GPA, math hours, math GPA, math education hours, and math education GPA, allowing the calculation of marginal effects of these variables. Additionally, by taking simple differences of the component variables, the coefficients on overall GPA again, the effect of experience is negative, although it is only significant for two cases for which they are statistically significant. Of the component variables, the coefficients on overall GPA are positive, substantively large, and statistically significant for the less advantaged student types (African American, subsidized lunch, and below-average performing).

Marginal effects are calculated for those characteristics that are jointly significant, and presented in Table 6. Once again, the effect of experience is negative, although it is insignificant in both samples. Overall GPA is most often predictive of student achievement, and the marginal effect is different from zero in the pooled sample, as well as for stu-

### Table 3

<table>
<thead>
<tr>
<th></th>
<th>Pooled</th>
<th>African American</th>
<th>European American</th>
<th>Subsidized lunch</th>
<th>Regular lunch</th>
<th>Below-average performing</th>
<th>Above-average performing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>0.027</td>
<td>0.002</td>
<td>0.001</td>
<td>0.032</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Overall GPA</td>
<td>&lt;0.001</td>
<td>0.003</td>
<td>0.046</td>
<td>0.009</td>
<td>0.007</td>
<td>0.167</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Math hours</td>
<td>0.002</td>
<td>0.005</td>
<td>0.132</td>
<td>0.018</td>
<td>0.011</td>
<td>0.100</td>
<td>0.002</td>
</tr>
<tr>
<td>Math GPA</td>
<td>0.001</td>
<td>0.006</td>
<td>0.095</td>
<td>0.035</td>
<td>0.002</td>
<td>0.119</td>
<td>0.002</td>
</tr>
<tr>
<td>Math education hours</td>
<td>0.003</td>
<td>0.005</td>
<td>0.166</td>
<td>0.022</td>
<td>0.016</td>
<td>0.096</td>
<td>0.003</td>
</tr>
<tr>
<td>Math education GPA</td>
<td>0.001</td>
<td>0.001</td>
<td>0.166</td>
<td>0.004</td>
<td>0.028</td>
<td>0.011</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Values of F-tests of joint significance for fixed effects models, by student group:

<table>
<thead>
<tr>
<th></th>
<th>Pooled</th>
<th>African American</th>
<th>European American</th>
<th>Subsidized lunch</th>
<th>Regular lunch</th>
<th>Below-average performing</th>
<th>Above-average performing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>0.027</td>
<td>0.002</td>
<td>0.001</td>
<td>0.032</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Overall GPA</td>
<td>&lt;0.001</td>
<td>0.003</td>
<td>0.046</td>
<td>0.009</td>
<td>0.007</td>
<td>0.167</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Math hours</td>
<td>0.002</td>
<td>0.005</td>
<td>0.132</td>
<td>0.018</td>
<td>0.011</td>
<td>0.100</td>
<td>0.002</td>
</tr>
<tr>
<td>Math GPA</td>
<td>0.001</td>
<td>0.006</td>
<td>0.095</td>
<td>0.035</td>
<td>0.002</td>
<td>0.119</td>
<td>0.002</td>
</tr>
<tr>
<td>Math education hours</td>
<td>0.003</td>
<td>0.005</td>
<td>0.166</td>
<td>0.022</td>
<td>0.016</td>
<td>0.096</td>
<td>0.003</td>
</tr>
<tr>
<td>Math education GPA</td>
<td>0.001</td>
<td>0.001</td>
<td>0.166</td>
<td>0.004</td>
<td>0.028</td>
<td>0.011</td>
<td>0.003</td>
</tr>
</tbody>
</table>

N: 2477 754 1522 1075 1237 662 1815

### Table 4

<table>
<thead>
<tr>
<th></th>
<th>Pooled</th>
<th>African American</th>
<th>European American</th>
<th>Subsidized lunch</th>
<th>Regular lunch</th>
<th>Below-average performing</th>
<th>Above-average performing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience</td>
<td>−0.116***</td>
<td>−0.150**</td>
<td>−0.117***</td>
<td>−0.110***</td>
<td>−0.119***</td>
<td>−0.100***</td>
<td>−0.158***</td>
</tr>
<tr>
<td>N</td>
<td>2477</td>
<td>754</td>
<td>1522</td>
<td>1075</td>
<td>1237</td>
<td>662</td>
<td>1815</td>
</tr>
</tbody>
</table>

Estimates from fixed effects models, by student group. Note: ***p < 0.01; **p < 0.05.
students who are African American, receive subsidized lunch, and are below-average performing. The marginal effects of math hours and math education hours are two to three times larger than those of the other teacher characteristics, but they are found in the pooled sample only. This implies that there is low power to detect the effects of teacher characteristics on achievement when the data are split in this manner.

As mentioned above, unmeasured factors confound the estimates in different ways in both estimation strategies. The simple difference of the between and within effects estimates of experience and the interaction terms provides estimates that are presumably free of both types of bias. These differences are listed in Table 7. The coefficients for GPA × experience, math education hours × experience, and math education GPA × experience tend to be negative across all student groups, suggesting that the effects of the qualifications diminish as teachers gain experience. Teachers with lower overall GPAs, higher math education GPAs, and more math education hours initially start teaching at a disadvantage, but their students’ test scores catch up to those of their peers over time. The most interesting results are those regarding math content courses, in which the interaction terms are generally positive. Those teachers who took more math content courses, and scored well in them, produce higher student math scores initially in their careers. More importantly, this effect increases as they gain years of experience, and the student test scores of their lower achieving peers do not adequately catch up.

The structure of the Kentucky state testing plan for this time period precludes the possibility of conducting this analysis with two consecutive years of math scores. This is potentially problematic because students may be naturally inclined in one academic area, while struggling in the other. To explore the use of the KCCT Reading score as an appropriate prior achievement score, the models are run on a sample of students whose performance on the reading and math tests are similar. Students were selected into the sample if the difference in their reading and math test scores falls within one standard deviation of the mean difference. Despite the smaller sample size, the results of the sensitivity analysis are similar to the results found in the full sample.5

### 4. Discussion

Out of all the teacher qualifications, only overall GPA consistently, positively impacts students’ math achievement across student group and model specification, making it an important teacher characteristic to include in models that predict math achievement. This is consistent with the theory that teacher motivation, as demonstrated by college performance, impacts student test scores. The marginal effect, calculated as a standard deviation increase in GPA, ranges from about 0.034 standard deviations in the pooled sample to 0.084 standard deviations for students who are African American. While this effect is not overwhelming for any given year, the cumulative effect of a student being

---

5 Results are available from the author upon request.
assigned to teachers with higher college GPAs over multiple years could be quite substantial. In terms of their students’ academic performance, teachers with higher GPAs start out their careers at an advantage over their low-performing peers, but the effect is not constant over time. Specifically, college performance and real-world teaching experience interact, reducing the gap in their teaching effectiveness.

The preferred specifications indicate that math education GPA consistently predicts math achievement across group. There are signals that the marginal effect is larger than that found for GPA (0.385 standard deviations), but this should be interpreted with caution as the joint test of hypothesis was only significant in the pooled sample and it cannot be determined whether the between effects models accounted for all sources of bias. The effect of math education GPA is initially negative, but this diminishes over time, and students start making positive gains during the teachers’ fifth year in the classroom. The number of math content hours also impacts math scores, save for students who are European American. Again, the marginal effect is large (between 0.108 and 0.281 standard deviations), but should be viewed cautiously. The interaction between experience and the number of math content courses taken is positive and implies that teachers who took more math content hours are initially more effective in comparison to other teachers and the effectiveness gap grows as they gain experience.

One of the distinguishing features of this study is the testing of teacher characteristics on different student groups. By splitting the students into specific racial, income, and academic performance groups, it is possible to see that experience and teacher characteristics reliably predict the performance of African American students. Studies that examine teacher qualifications on pooled samples alone may miss important relationships between teacher characteristics and student achievement that exist within important student subgroups. The findings from two separate literatures provided the motivation to split the full sample into groups based on income, race, and academic performance. Studies on peer effects generally conclude that groups of student achieve differently based on their peers’ characteristics, making it plausible that they will also achieve differently according to their teachers’ characteristics. Based on this research, the expectation was that lower-performing students, those from minority racial groups, and those from lower income families would be most affected by their teachers’ characteristics. The second line of research indicates that teacher and school characteristics affect students differentially along racial lines.

The tables indicating joint significance provide the evidence necessary to explore these hypotheses. The between effects models provide tenuous support for the peer effects hypothesis, but on the whole, both types of models are more consistent with a racial effect. The peer effects support is found in Table 6 where overall GPA is shown to impact the achievement of students who are African American, receive federally subsidized lunch, or have below-average prior test scores. However, it is the only teacher characteristic exhibiting this pattern.

On the other hand, both types of models support the hypothesis that African American students are affected differently by teacher characteristics than European American students. There are no differences when comparing students who receive federally subsidized lunch and those who do not, and in the case of prior achievement, the fixed effects models show that teacher qualities affect above-average performing students, but not their below-average performing peers. Table 3 shows that every teacher characteristic affects students who are African American, while only three variables affect European American students. Similarly, Table 6 indicates that three teacher attributes are jointly statistically significant for students who are African American, while the rest of the groups are relatively unaffected by these factors. All of this evidence indicates that teacher qualifications differentially affect student math achievement along racial lines, but not according to family income levels or previous achievement.

5. Policy discussion

The findings of this study inform the current policy debate regarding traditional and alternative paths to teacher certification. Advocates for the traditional pathway argue that education school coursework provides important pedagogical and classroom management skills that are integral to teaching success, while supporters of alternative programs assert that content knowledge is the most important attribute of a quality teacher. These data do not allow the direct testing of the pathway effects, but they do provide support that both content and pedagogical knowledge are important to effective teaching. Content knowledge is a key component of both traditional and alternative pathways to teaching, and the findings for math content hours justify this focus. The
Table 7
Unbiased estimates of teacher quality variables

<table>
<thead>
<tr>
<th></th>
<th>Below-average performing</th>
<th>Regular lunch</th>
<th>Subsidized lunch</th>
<th>Above-average performing</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Experience</strong></td>
<td>0.543*** (0.011)</td>
<td>0.974*** (0.016)</td>
<td>0.003 (0.003)</td>
<td>0.894*** (0.008)</td>
</tr>
<tr>
<td><strong>GPA</strong></td>
<td>-0.073*** (0.004)</td>
<td>0.000 (0.004)</td>
<td>0.003 (0.003)</td>
<td>-0.112*** (0.004)</td>
</tr>
<tr>
<td><strong>Math hours</strong></td>
<td>0.000 (0.000)</td>
<td>0.022*** (0.000)</td>
<td>0.002 (0.002)</td>
<td>-0.025*** (0.000)</td>
</tr>
<tr>
<td><strong>Math GPA</strong></td>
<td>0.092*** (0.001)</td>
<td>0.124*** (0.002)</td>
<td>0.002 (0.002)</td>
<td>-0.006*** (0.000)</td>
</tr>
</tbody>
</table>

**Note:** *** < 0.01.

Marginal effect of this characteristic is slightly smaller than that of math education hours, but it is somewhat more robust. More importantly, it has a critical relationship with experience. The positive effects of math content hours grow each successive year the teacher is in the classroom. All else equal, a teacher who took 11 h of math content will have higher student math scores than a teacher who took 10 h of math content and will have incrementally higher student math scores over the years.

The number of math education hours has the largest marginal effect of any teacher characteristic, although this effect is negative until teachers gain between 10 and 14 years of experience. The importance of this finding should not be overstated since it is not robust across student group, but it does illustrate the value of math education coursework. This provides support for the traditional pathway to teaching, in which students graduate from an accredited program, which includes coursework in pedagogy, content, student teaching, and passing the appropriate state licensure exam. There is nearly universal agreement that teacher education needs to be improved, but there is no consensus on how or why this can be achieved (Cochran-Smith & Fries, 2005). While this study cannot provide answers to this question, it does suggest that maintaining standards for qualified prospective teachers is a suitable approach.

**Acknowledgments**

I am grateful to the Kentucky Education Professional Standards Board for providing the data used in this project. I thank Tim Sass for valuable comments on an earlier draft of this manuscript.

**References**


