

POLICY BRIEF

Race and Gender Differences in Teacher Evaluation Ratings and Teacher Employment Outcomes

Steven Drake, *Michigan State University*

Joshua M. Cowen, *Michigan State University*

Amy Auletto, *Michigan State University*

May 2019

EPIC

Education Policy
Innovation Collaborative
RESEARCH WITH CONSEQUENCE

ACKNOWLEDGMENTS

This policy brief is based on a published paper "Grading Teachers: Race and Gender Differences in Low Evaluation Ratings and Teacher Employment Outcomes," in the American Educational Research Journal - <https://journals.sagepub.com/doi/full/10.3102/0002831219835776>

DISCLAIMER

This research result used data collected and maintained by the Michigan Department of Education (MDE) and/or Michigan's Center for Educational Performance and Information (CEPI). Results, information and opinions solely represent the analysis, information and opinions of the author(s) and are not endorsed by, or reflect the views or positions of, grantors, MDE and CEPI or any employee thereof.



Race and Gender Differences in Teacher Evaluation Ratings and Teacher Employment Outcomes

By Steven Drake, Joshua M. Cowen, Amy Auletto

Overview

Over the past decade, Michigan has made changes to a number of laws related to teachers and teaching in the state's public school system. These changes include new requirements for teacher evaluation, and the use of teacher performance ratings for decisions such as teacher tenure, dismissal and retention. Since 2011, state law has required local education agencies (districts) to rate teachers as highly effective, effective, minimally effective or ineffective. These ratings must be based on classroom observations and, for most teachers, a measure of achievement growth for students assigned to each teacher. Since 2016, state law has also required districts to adopt valid and reliable classroom observation protocols, and new training for teacher evaluators.¹

In this brief, we examine differences in teacher evaluation ratings in schools across the state from the 2011-12 through the 2015-16 academic years. We pay special attention to differences in ratings between male and female teachers, White teachers and teachers of color, and teachers in traditional public schools and public school academies (PSAs, or charter schools). These comparisons are important because evaluation ratings can significantly affect teachers' job security, and because the overall supply of teachers in Michigan—and especially of Black teachers—has declined considerably over the decade.²

KEY FINDINGS INCLUDE:

- Statewide, nearly 19% of Black teachers in Michigan received a low rating from 2011-12 to 2015-16, compared to just 7% of White teachers.
- Teachers of color—especially Black teachers—are 50% more likely to receive low evaluation ratings than White teachers within the same school.
- Teachers of color in schools with high numbers of White teachers are more likely to receive low ratings.
- Male teachers are more likely than female teachers to receive low ratings.
- First-year teachers more likely to receive lower ratings than more experienced teachers.
- Teachers in charter schools are more likely to receive low ratings than those in traditional schools.
- Teachers rated below effective are more likely to leave their school after receiving their rating; low-rated teachers of color are not more likely to leave than low rated White teachers.
- Colleagues matter:
 - o Black teachers are less likely to receive low ratings in schools with higher percentages of Black teachers.
 - o Male teachers are less likely to receive low ratings in schools with male administrators.

BACKGROUND

In recent years, many states have implemented performance-based teacher evaluation systems. Although the 2015 Every Student Succeeds Act relaxed the extent to which the federal government dictates state teacher evaluation policy, the majority of states continue to require teachers to be formally evaluated.³ Like other states, Michigan has implemented a series of teacher evaluation reforms dating back to 2010. The most important of these was PA 102 in July 2011, which created a ratings system from ineffective to highly effective beginning in the 2011-12 academic year, and PA 173, which further

refined requirements for classroom observation and evaluator training.⁴ Michigan's teacher evaluation laws were meant to engage and reshape its teacher workforce through implementation of "high-stakes" teacher assessments. In theory, teachers can respond to their ratings and work to improve practice. Teachers who do not improve could be removed from practice.

Previous research has noted the challenges of implementing such policies even under favorable, narrowly constructed conditions, and has largely studied their effect on measures of student achievement.⁵ One concern is how current and prospective teachers necessary to staff public schools perceive the extent to which evaluation systems are fair and objective. Earlier research has demonstrated the potential for high-stakes evaluation to increase the number of teachers exiting public schools, and decrease the number of new teachers entering the profession—especially when teacher pay is not raised to offset new job insecurity.⁶ In addition, if teachers from different demographic groups are affected differently by evaluation policies, the composition of the teacher workforce may change in ways unanticipated by policy makers through affecting teachers' and prospective teachers' understandings of fairness and their respective risks.

In this policy brief we examine the evaluation ratings of Michigan public school teachers from 2011 to 2015, and describe which teacher and school characteristics are especially related to low evaluation scores. We look for differences between traditional public schools and public school academies (PSAs or charter schools), as charter teachers tend to have higher rates of exit from their schools than those in traditional settings.⁷ Also, because other researchers working in individual school districts elsewhere in the United States have found evidence that teachers of color are especially likely to receive low performance ratings from their supervisors, we consider that possibility in Michigan as well.⁸

Such a pattern would be problematic if replicated and found not to be specific to particular districts' evaluation procedures or administrators. This brief is also set against a broader backdrop of a roughly 27% decline in Michigan's African American teaching force from 2011-2015.

We examine evaluation ratings of teachers and describe which teacher and school characteristics are especially related to low evaluation scores.

HOW THIS ANALYSIS WAS CONDUCTED

We focus on the earlier years of Michigan’s teacher evaluation system, using data on teachers and schools compiled by the Michigan Department of Education (MDE) and the Center for Educational Performance and Information (CEPI) between 2011-12 and 2015-16. These data contain all public K-12 employees, including each teacher’s summative end-of-year rating (highly effective, effective, minimally effective, or ineffective, described above).⁹ We **analyzed 364,658 teacher-year ratings** representing the evaluations of 97,446 licensed classroom teachers working in an instructional capacity* for the 2011-12 and 2015-16 school years.

We analyze these data in three steps. First, we calculate the frequency (percentage) of each rating category from 2011-2015—highly effective, effective, minimally effective, and ineffective—for teachers of different demographic backgrounds and teaching assignment. We also calculate each rating frequency for schools with different characteristics. Second, we calculate the extent to which teachers receiving each rating category left their teaching assignments immediately after receiving a low rating. Third, we calculate rating differences within each school, which allows us to rule out the possibility that different ratings are explained by the schools in which teachers work.

RESULTS

In Table 1 we show the results of our first analysis: the frequency of each teacher evaluation rating across the 2011-12 to 2015-16 school years. Several patterns stand out in this table. The first is the relative lack of low ratings—minimally effective or ineffective—for all teachers in Michigan. Overall, an average of only 2.6% of Michigan teachers during this time period received a low rating in a given year. The second pattern, however, is how different that frequency is for teachers of color. On average, 4.5% of Latino/Latina teachers received a low rating in a given year—more than twice the frequency for White teachers. Meanwhile, 7.5% of Black teachers received minimally effective ratings in a given year—more than three times the frequency of White teachers. Other noteworthy patterns include: probationary teachers, teachers in public school academies, teachers in Title 1 schools, and teachers in state-designated turnaround schools are more likely to receive low ratings.

**Note: We used the State of Michigan’s system of job assignment descriptions given to all school employees to determine teaching status. We determined whether a school employee was acting as a teacher in an instructional capacity by inspecting these codes and designating them as a teacher or non-teacher. We exclude teacher’s aides and other paraprofessionals, coaches, and reading and other specialists with an objective of only investigating those whose job is to lead classroom instruction in core subjects, the arts, physical education, and vocational education.*

TABLE 1. Frequencies of teacher evaluation ratings across selected categories 2011-12 to 2015 -16.

Teacher Evaluation Ratings					
	Highly Effective	Effective	Minimally Effective	Ineffective	Any Low Score
<i>Teacher Demographic Type</i>					
Male	31.3%	65.4%	2.6%	0.7%	3.2%
Female	38.7%	59.0%	1.9%	0.4%	2.3%
Black	41.2%	51.3%	5.7%	1.8%	7.5%
Latino/Latina	31.7%	63.7%	3.7%	0.8%	4.6%
White	36.7%	61.1%	1.8%	0.4%	2.2%
Other Race	34.1%	62.6%	2.7%	0.6%	3.3%
<i>Teacher Role</i>					
Title I Teacher	36.4%	57.6%	4.8%	1.2%	6.0%
Probationary Status	25.6%	69.8%	3.8%	0.8%	4.6%
Professional Status	38.7%	59.1%	1.8%	0.5%	2.2%
High Stakes Assignment	36.8%	60.6%	2.1%	0.5%	2.6%
Low Stakes Assignment	37.3%	60.3%	2.0%	0.5%	2.5%
Secondary	35.2%	62.1%	2.1%	0.6%	2.7%
Elementary	37.7%	59.8%	2.0%	0.5%	2.5%
<i>School Type</i>					
Public School Academy (PSA or charter)	22.0%	67.9%	8.4%	1.6%	10.1%
Traditional Public School	38.1%	60.0%	1.5%	0.4%	1.9%
Public School Academy Probationary	18.0%	70.2%	9.9%	1.9%	11.7%
Public School Academy Professional	24.8%	66.3%	7.4%	1.4%	8.9%
Traditional Public School Probationary	27.8%	69.7%	2.0%	0.4%	2.4%
Traditional Public School Professional	39.5%	58.7%	1.4%	0.4%	1.9%
> 50th Percentile Minority Student Body	32.6%	63.5%	3.2%	0.8%	3.9%
> 75th Percentile Minority Student Body	32.2%	61.6%	5.0%	1.3%	6.3%
> 50th Percentile Poverty Student Body	32.3%	63.6%	3.4%	0.8%	4.2%
> 75th Percentile Poverty Student Body	32.3%	62.3%	4.3%	1.1%	5.4%
Priority School	32.7%	59.7%	5.8%	1.8%	7.6%
Focus School	27.0%	71.0%	1.7%	0.3%	2.0%
Reward School	34.2%	63.4%	1.9%	0.5%	2.4%
Turnaround School	40.0%	49.4%	7.8%	2.7%	10.6%
City Locale	36.7%	58.5%	3.8%	1.0%	4.8%
Suburb Locale	36.6%	61.7%	1.5%	0.3%	1.8%
Town Locale	37.3%	61.0%	1.3%	0.4%	1.7%
Rural Locale	37.5%	60.5%	1.6%	0.4%	2.0%
AVERAGE	36.8%	60.6%	2.1%	0.5%	2.6%

Notes: Probationary teachers defined as those with provisional licenses; teachers with non-provisional and non-temporary licenses defined as "professional" status.

High-stakes assignment is defined as assignment to a state-tested subject/grade; low-stakes assignments are teachers in non-tested subjects or grades.

Source: Michigan administrative data. Any low score is the sum of minimally effective and ineffective. Priority, Focus and Reward schools are designated by the State of Michigan.

The low frequency of low evaluations in each year overall masks important data over the longer time period we examine here. Table 2 provides calculations of the cumulative frequency of low ratings for certain teachers over the 2011-2015 time period. This is important because key components of the teacher evaluation laws described above link job security to the number of low ratings teachers receive over time (although thus far we have seen few teachers receive enough consecutive ratings to warrant formal dismissal). Overall, we see in Table 2 that the chance a Black teacher in Michigan received at least one low rating during this timeframe is nearly 19%, compared to 6.5% for White teachers. Latino/Latina, male, and public school academy school teachers are also far more likely to receive at least one low rating. And the chances of receiving multiple low ratings are also far higher for Black, Latino/Latina, male and public school academy teachers as well.

TABLE 2. Cumulative percentages of teachers receiving low ratings by sector and demographic 2011-12 to 2015-16.

Teachers Receiving Low Ratings by Sector and Demographic						
Number of Low Ratings Received	White	Black	Latino/ Latina	Female	Male	All
<i>All Schools</i>						
1	5.0%	14.7%	10.8%	5.2%	7.2%	5.7%
≤ 2	6.1%	18.0%	12.2%	6.4%	8.9%	7.0%
≤ 5	6.5%	18.9%	12.8%	6.7%	9.4%	7.4%
<i>Traditional Public Schools</i>						
1	3.9%	12.8%	8.9%	4.0%	6.0%	4.5%
≤ 2	4.8%	15.6%	10.3%	4.9%	7.4%	5.5%
≤ 5	5.2%	16.4%	10.8%	5.2%	7.9%	5.9%
<i>Charter Schools</i>						
1	15.1%	21.4%	20.8%	15.1%	19.6%	16.1%
≤ 2	18.1%	26.3%	22.5%	18.3%	23.6%	19.4%
≤ 5	19.2 %	27.4%	23.7%	19.3%	24.7%	20.4%

Note: Probationary teachers defined as those with provisional licenses; teachers with non-provisional and non-temporary licenses defined as “professional” status.

Source: Calculations by authors using Michigan administrative data.

Teachers of any school context are more likely to leave immediately after receiving a low rating. Table 3 indicates that for traditional public schools, more than 40% of teachers receiving an ineffective rating, and more than 20% of teachers rated as minimally effective, exited their schools after the year in which they received those evaluations. For public school academy teachers, those numbers are comparable for minimally effective teachers and about 33% - or one-third - for teachers receiving the lowest score (ineffective) available. We note that in follow-up analyses (unreported in this brief for reasons of space),¹⁰ the increased chances that a Black teacher exits is due primarily to their increased risk of a low evaluation rating. In fact, Black teachers are marginally more likely to persist in their schools after a low rating than their non-Black colleagues with a similarly low rating.

TABLE 3. Probability of immediate exit* by effectiveness rating for post-reform years commencing 2011-12 to 2014-15.

Probability of Exit by Effectiveness Rating				
Effectiveness Rating	Traditional Public Schools		Public School Academies	
<i>All Schools</i>	Probationary Status	Professional Status	Probationary Status	Professional Status
Highly Effective	4.5%	4.7%	10.5%	8.5%
Effective	6.4%	6.8%	12.5%	12.7%
Minimally Effective	22.0%	23.8%	17.6%	22.1%
Ineffective	41.4%	40.6%	32.4%	32.9%
AVERAGE AMONG ALL TEACHERS	6.4%	6.4%	13.1%	12.7%

*Notes: Immediate exit does not guarantee permanent exit from teaching. Instead it measures exit for at least one year. Correlation between immediate exit and permanent exit is high (> .9).

Source: Michigan administrative data.

Finally, in Table 4, we report the frequencies of low evaluation ratings using within-school calculations, meaning that unlike Tables 1-3, which compare teachers across Michigan to each other regardless of where they teach, the analysis in Table 4 focuses on differences within each Michigan school. Table 4 shows that Black teachers and male teachers remain far more likely to receive a low evaluation rating even compared to White or female teachers within their schools. This is true, as the table depicts, even for Black teachers who had previously received a lower score, which provides some indication that these differences are not reflections of teacher effectiveness only. Table 4 also shows that these differences are smaller for Black teachers in schools with more Black colleagues. The same is true for male teachers with male supervisors.

TABLE 4. Low Effectiveness Rate Difference by Race and Gender after Controlling for Differences in Low Rating Issuance Across Schools.

Low Effectiveness Rate Difference by Race and Gender			
	All Public Schools	Traditional Public Schools	Public School Academies (Charter)
White teachers	2.5%	1.8%	10.0%
Black teachers	3.5%	2.8%	11.2%
-- with few other Black faculty	4.6%	3.8%	12.4%
-- with moderate Black faculty	3.7%	3.0%	11.4%
-- with large Black faculty	2.1%	1.5%	9.0%
Latino/Latina teachers	3.8%	3.1%	12.0%
Other racial/ethnic groups	2.8%	2.1%	10.8%
Female teachers	2.3%	1.6%	9.3%
Male teachers	3.5%	2.6%	13.4%
-- with primarily male administration	3.1%	2.3%	13.5%
-- with primarily mixed gender administration	3.3%	2.6%	12.7%
-- with primarily female administration	4.2%	3.3%	13.8%

Notes: Results represent the synthesis of several independent regression models, each predicting the occurrence of an ineffective or minimally effective rating.

"Few other Black faculty" means less than 10% of faculty in school are Black. "Moderate Black faculty" means between 10 and 50% of faculty are Black. "Large Black faculty" means greater than 50% of faculty are Black.

"Primarily male" administrations are 80% or more male. "Mixed gender" administrations are between 20% and 80% male. "Primarily female" administration is less than 20% male. Primarily male and primarily female administrations are very often all female or all male.

Results represent the synthesis of several independent regression models.

SUMMARY

In this policy brief we document substantial variation in the receipt of low teacher evaluation ratings in Michigan. Although minimally effective and ineffective ratings are rare in Michigan (less than three percent annually to date) those low ratings that do occur fall disproportionately on teachers with particular demographic characteristics and those working in particular schools. Teachers of color in particular are more likely to be labeled “minimally effective” or “ineffective” than White teachers in their same school. Between

2011-12 and 2015-16, nearly 19% of Black teachers in Michigan received a low rating compared to just 6.5% of White teachers. Significantly, Black teachers who worked in traditional public schools staffed by a higher proportion of White colleagues have an increased risk of receiving a low rating as compared to Black teachers who work with more same-race colleagues. Black teachers whose schools have Black administrators—our best calculation of who is supervising these teachers—are less likely to receive a low rating. Male teachers are also more likely to receive low ratings and, similar to the pattern we see for Black teachers, that risk decreases when male teachers have male supervisors.

In addition to the variation in low rating assignment based on teacher characteristics, the substantial variation across school type is notable. Public school academies (PSAs or charter schools) assigned five times as many low ratings as did their traditional counterparts, though it does not appear that low-rated charter teachers were correspondingly more likely to leave those schools.

Overall, in fact, the differences that we calculate in the relative frequency of low ratings between teachers of color, male teachers, and charter teachers do not translate into higher rates of exit compared to White female teachers in traditional public schools also receiving low ratings. In other words, the evaluation rating itself is a more meaningful predictor of teacher exit than race or gender or school-type.

IMPLICATIONS

The two commonly agreed upon purposes of teacher evaluation are to measure and develop teachers.¹¹ Although teacher evaluation systems intend to measure how well teachers are performing and, when necessary, remove them from their position or encourage other employment, evaluation is also a tool that can encourage professional growth and offer feedback to teachers on how they can improve their instruction. However, if evaluation ratings are not being consistently assigned, as seen in this study, the broader system's reputation for fairness may be harmed - ultimately affecting teacher retention and recruitment. In particular, is the potential raised in our results - a story we can ultimately neither confirm nor reject—that a flexible, locally determined evaluation system offers at least one context for (intentional or unintentional) discrimination on the basis of race and gender. Such a pattern has been found in states elsewhere.¹² As Michigan and other states continue to adopt and adapt high-stakes teacher evaluation systems, the results of this brief suggest that the design, implementation, and monitoring of these systems include routine checks to ensure equitable outcomes.

Citations

- ¹ State of Michigan. (2011). 96th Regular Session of 2011 Public Acts 100-103, Retrieved from: [http://www.legislature.mi.gov/\(S\(shpnxyzelc5gk4bdylhw3wuz\)\)/documents/2011-2012/publicact/pdf/2011-PA-0100.pdf](http://www.legislature.mi.gov/(S(shpnxyzelc5gk4bdylhw3wuz))/documents/2011-2012/publicact/pdf/2011-PA-0100.pdf); [http://www.legislature.mi.gov/\(S\(z1qtzld2isehc0wpzfqeppzt\)\)/documents/2011-2012/publicact/pdf/2011-PA-0101.pdf](http://www.legislature.mi.gov/(S(z1qtzld2isehc0wpzfqeppzt))/documents/2011-2012/publicact/pdf/2011-PA-0101.pdf); <https://www.legislature.mi.gov/documents/2011-2012/publicact/htm/2011-PA-0102.htm>; [http://www.legislature.mi.gov/\(S\(f4zpyrmryn3defdbjquqjzwy\)\)/documents/2011-2012/publicact/pdf/2011-pa-0103.pdf](http://www.legislature.mi.gov/(S(f4zpyrmryn3defdbjquqjzwy))/documents/2011-2012/publicact/pdf/2011-pa-0103.pdf), Accessed 10/30/2016.
- State of Michigan. (2015). 98th Regular Session of 2011 Public Acts 173, Retrieved from <http://www.legislature.mi.gov/documents/2015-2016/publicact/pdf/2015-PA-0173.pdf>, Accessed 10/30/2016
- ² Drake, S. & Cowen, J. (2019). Teacher Diversity in a Changing Public School Market: Evidence on African American Employment in Michigan since 2005; Office of Postsecondary Education. (2017). Title II Reports from U.S. Department of Education <https://title2.ed.gov/Public/Home.aspx>
- ³ Doherty, K.M. & Jacobs, S. (2015). State of the states: Evaluating teaching, leading and learning. Washington, D.C.: National Council on Teacher Quality. Retrieved from <http://www.nctq.org/dmsView/StateofStates2015>
- ⁴ State of Michigan. (2011). 96th Regular Session of 2011 Public Acts 100-103, Retrieved from: [http://www.legislature.mi.gov/\(S\(shpnxyzelc5gk4bdylhw3wuz\)\)/documents/2011-2012/publicact/pdf/2011-PA-0100.pdf](http://www.legislature.mi.gov/(S(shpnxyzelc5gk4bdylhw3wuz))/documents/2011-2012/publicact/pdf/2011-PA-0100.pdf); [http://www.legislature.mi.gov/\(S\(z1qtzld2isehc0wpzfqeppzt\)\)/documents/2011-2012/publicact/pdf/2011-PA-0101.pdf](http://www.legislature.mi.gov/(S(z1qtzld2isehc0wpzfqeppzt))/documents/2011-2012/publicact/pdf/2011-PA-0101.pdf); <https://www.legislature.mi.gov/documents/2011-2012/publicact/htm/2011-PA-0102.htm>; [http://www.legislature.mi.gov/\(S\(f4zpyrmryn3defdbjquqjzwy\)\)/documents/2011-2012/publicact/pdf/2011-pa-0103.pdf](http://www.legislature.mi.gov/(S(f4zpyrmryn3defdbjquqjzwy))/documents/2011-2012/publicact/pdf/2011-pa-0103.pdf), Accessed 10/30/2016.
- State of Michigan. (2015). 98th Regular Session of 2011 Public Acts 173, Retrieved from <http://www.legislature.mi.gov/documents/2015-2016/publicact/pdf/2015-PA-0173.pdf>, Accessed 10/30/2016.
- ⁵ e.g., Adnot, M., Dee, T., Katz, V., & Wyckoff, J. (2017). Teacher turnover, teacher quality, and student achievement in DCPS. *Educational Evaluation and Policy Analysis*, 39(1), 54-76; Dee, T. & Wyckoff, J. (2015). Incentives, selection, and teacher performance: Evidence from IMPACT. *Journal of Policy Analysis and Management*, 34(2), 267-297. doi: 10.1002/pam.21818; Fryer, R. (2013). Teacher incentives and student achievement: Evidence from New York City public schools. *Journal of Labor Economics*, 31, 373-427; Glazerman, S., & Seifullah, A. (2012). An evaluation of the Chicago Teacher Advancement Program (Chicago TAP) after four years. Princeton, NJ: Mathematica Policy Research. Researchers have emphasized the difficulty of implementation of objective and formatively constructive high-stakes evaluation (Cohen, J. & Goldhaber, D. (2015). Observations on evaluating teacher performance: Assessing the strengths and weaknesses of classroom observations and value-added measures. In J.A. Grissom & P. Youngs (Eds.), *Improving teacher evaluation systems: Making the most of multiple measures* (pp. 8-21). New York, NY: Teachers College Press; Cohen, J. & Goldhaber, D. (2016). Building a more complete understanding of teacher evaluation using classroom observation. *Educational Researcher*, 45(6), 378-387. doi: 10.3102/0013189X16659442
- ⁶ Rothstein, J. (2014). Teacher quality policy when supply matters. *The American Economic Review*, 105(1), 100-130. doi: 10.1257/aer.20121242; Kraft, M., Brunner, E., Doherty, S., & Schwegman, D. (2018). Teacher accountability reforms and the supply of new teachers. Unpublished manuscript.
- ⁷ Between 2011 and 2015, 6.9% of teachers on average left traditional districts in Michigan, while 13.2% left public school academies. Cowen, J.M., & Winters, M.A. (2013). Do charters retain teachers differently? Evidence from elementary schools in Florida. *Education Finance and Policy*, 8(1), 14-42. doi: <https://www-jstor-org.proxy1.cl.msu.edu/stable/educfinapoli.8.1.14>; Stuit, D.A., & Smith, T.M. (2012) Explaining the gap in charter and traditional public school teacher turnover rates. *Economics of Education Review* 31(2):268-279

Citations (continued)

⁸ Grissom, J.A. & Loeb, S. (2016). Assessing principals' assessments: Subjective evaluations of teacher effectiveness in low- and high-stakes environments. *Education Finance and Policy*, 1-53. doi: 0.1162/EDFP_a_00210

Bailey, J., Bocala, C., Shakman, K., & Zweig, J. (2016). Teacher demographics and evaluation: A descriptive study in a large urban district (REL 2017-189). Retrieved from Washington DC: <http://ies.ed.gov/ncee/edlabs>

⁹ All supplemental school and district descriptors are sourced from web portals for the Michigan Department of Education (2017a) and the National Center for Education Statistics. (2017). Elementary/Secondary Information System. Retrieved from <https://nces.ed.gov/ccd/elsi/>

¹⁰ These additional calculations are available on request from EPICedpolicy@msu.edu

¹¹ Papay, J.P. (2012). Refocusing the debate: Assessing the purposes and tools of teacher evaluation. *Harvard Educational Review*, 82(1), 123-141. Retrieved from <http://hepg.org/her-home/home>; Sartain, L. & Steinberg, M.P. (2014). Teachers' labor market responses to performance evaluation reform: Experimental evidence from Chicago Public Schools. *The Journal of Human Resources*, 51(3), 615-655. doi: 10.3368/jhr.51.3.0514-639OR1

¹² Grissom, J.A. & Loeb, S. (2016). Assessing principals' assessments: Subjective evaluations of teacher effectiveness in low- and high-stakes environments. *Education Finance and Policy*, 1-53. doi: 0.1162/EDFP_a_00210



Education Policy Innovation Collaborative

MICHIGAN STATE UNIVERSITY
201 Erickson Hall | 620 Farm Lane
East Lansing, MI 48824

(517) 884-0377
EPICedpolicy@msu.edu
www.EPICedpolicy.msu.edu