School performance measures sometimes change due to random luck rather than true performance. This means these measures can be misleading—especially when they involve small numbers of students.

The highest peaks and lowest valleys are often misleading indicators of school performance, driven partly by random luck (good or bad).

Stabilization uses historical and contextual data to improve accuracy.

Stabilization increases accuracy by reducing the influence of random luck.

The rollercoaster isn’t as scary!
Improving accuracy of measures for students in historically underserved subgroups is essential to promote equity and help students grow.

Data should be used as a flashlight, not a hammer.

Miguel Cardona, U.S. Secretary of Education

States measure subgroup performance in schools to shine a light on inequities.

But subgroups in schools are often small, giving random luck an outsized influence.

Stabilized performance measures serve as a better flashlight, helping states identify schools and subgroups that most need support.

Our stabilization method is called Bayesian partial pooling. Curious about the nuts and bolts? Check out these resources:

Introduction to the foundational stabilization method, partial pooling:
https://blog.conductrics.com/prediction-pooling-and-shrinkage/

Introduction to general Bayesian methods for beginners:

An application of Bayesian methods for stabilizing results in education science:

This infographic was prepared under contract 91990022C0012 by Regional Educational Laboratory Mid-Atlantic, administered by Mathematica. The content does not necessarily reflect the views or policies of IES or the U.S. Department of Education, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

Please visit the Regional Educational Laboratory Mid-Atlantic website at https://ies.ed.gov/ncee/rel/region/midatlantic