Appendix

This Appendix provides more details about the data and statistical models used to analyze the empirical data from the Stanford Education Data Archive (SEDA; Reardon et al., 2018) data in Section 2.

Data

As noted in the main text, the data analyzed here are from SEDA version 2.1 (Fahle et al., 2018). We make the following sample restrictions to the SEDA version 2.1 database for the purposes of our analyses. First, some states reported proficiency data in only two categories during some years, requiring that HOMOP models were fit to these datasets. We drop these observations because all districts within that particular state, grade, subject, and year were constrained to have equal standard deviations. Second, in cases where data were reported in three or more proficiency categories (the majority of data), PHOP models were fit by constraining the logged standard deviation for districts with fewer than 50 students to be equal to the average logged standard deviation of all districts with more than 50 students in the same state, grade, year, and subject. We therefore drop all district observations with estimates based on fewer than 50 students. These restrictions ensure that the remaining standard deviation estimates were estimated without constraints. After these restrictions, we drop all states with estimates for fewer than 50 districts. Final sample sizes are presented in the main text of the paper.

Statistical Models

SEDA contains estimates of σ'_{grt} (standardized within states, grades, years, and subjects) with an associated standard error for each district g in grade r and year t in each state and subject. The SEDA data also contain estimated standard errors of the $\hat{\sigma}'_{grt}$'s. We use the delta method to estimate the standard error of $\hat{\gamma}'_{grt} = \ln\left(\hat{\sigma}'_{grt}\right)$ as:

$$SE\left(\hat{\gamma}'_{grt}\right) = \sqrt{\frac{1}{{\hat{\sigma}'_{grt}}^2} SE\left(\hat{\sigma}'_{grt}\right)^2} = \frac{1}{\hat{\sigma}'_{grt}} SE\left(\hat{\sigma}'_{grt}\right). \quad \text{(A1)}$$

We use the estimated sampling variances of the $\hat{\gamma}'_{grt}$ values in a variance-known model (Raudenbush & Bryk, 2002) that accounts for the sampling error in the estimates. For each state-subject dataset, the general form of the model begins with an equation for the estimated $\hat{\gamma}'_{grt}$ values

$$\begin{gathered} \hat{\gamma}'_{grt} = \gamma'_{grt} + \epsilon_{grt} \\ \epsilon_{grt} \sim N\left(0, \hat{V}_{grt}\right) \end{gathered} \tag{A2}$$

where $\hat{V}_{grt}$ is the square of the estimated standard error of $\hat{\gamma}'_{grt}$. We then fit two models for the γ'_{grt} values in each state-subject dataset:

Model 1: $\gamma'_{grt} \ = \beta_{0g} + \Gamma_{rt} + e_{grt}$, where $e_{grt} \sim N(0, \omega_1^2)$ and $\beta_{0g} \sim N(0, \nu_{00})$;

Model 2: $\gamma'_{grt} \ = \beta_{0g} + \beta_{1g} r + \beta_{2g} t + \Gamma_{rt} + e_{grt}$, where $e_{grt} \sim N(0, \omega_2^2)$ and (A3)

$$\begin{bmatrix} \beta_{0g} \\ \beta_{1g} \\ \beta_{2g} \end{bmatrix} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{T} = \begin{bmatrix} \tau_{00} & & \\ \tau_{10} & \tau_{11} & \\ \tau_{20} & \tau_{21} & \tau_{22} \end{bmatrix} \right).$$

The Γ_{rt} are grade by year fixed effects. Grade and year variables were centered at the mean value within each state-subject dataset. All models were fit using the software HLM 7 (Raudenbush, Bryk, & Congdon, 2013).

Model 1 includes a random intercept for each district and represents the fully pooled HETOP model structure in each state-subject dataset. In this model, ν_{00} is the variance between districts in γ_{grt} values, while ω_1^2 is the variance within districts across grades and years. Model 2 includes random grade and year linear trends for each district and represents the linear trend pooled HETOP model structure in each state-subject dataset; ω_2^2 is the unexplained within-district variance in γ_{grt} values and the elements of $\mathbf{T}$ indicate the variance of district-specific intercepts and linear trends. We can use estimates from Model 1 to estimate the ratio $\rho = \hat{\nu}_{00}/(\hat{\nu}_{00} + \hat{\omega}_1^2)$ to quantify the proportion of variation in γ_{grt} that is between districts. We then calculate the quantity $\Delta_{12} = 1 - \hat{\omega}_2^2/\hat{\omega}_1^2$, the percent of unexplained, within-district variance in γ_{grt} values that can be explained by adding district-specific linear grade and year trends. Table A1 summarizes the average values across the 80 state-subject models. Each row represents 40 models estimated using ELA test score data and 40 models estimated using Math test score data. Interpretation of the results is presented in the main text.

References

Fahle, E. M., Shear, B. R., Kalogrides, D., Reardon, S. F., DiSalvo, R., & Ho, A. D. (2018). *Stanford Education Data Archive: Technical documentation (Version 2.1)*. Retrieved from https://stacks.stanford.edu/file/druid:db586ns4974/SEDA_documentation_v21.pdf

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publications, Inc.

Raudenbush, S. W., Bryk, A. S., & Congdon, R. (2013). *HLM 7.01 for Windows*. Skokie, IL: Scientific Software International, Inc.

Reardon, S. F., Ho, A. D., Shear, B. R., Fahle, E. M., Kalogrides, D., & DiSalvo, R. (2018). Stanford Education Data Archive (Version 2.1). Retrieved from https://cepa.stanford.edu/seda/overview

Table A1. Summary of HLM Model Estimates by Subject.

	ELA			Math		
Statistic	Min	Mean	Max	Min	Mean	Max
ω_1^2	0.0004	0.0018	0.0045	0.0008	0.0028	0.0064
ν_{00}	0.0015	0.0033	0.0058	0.0015	0.0049	0.0103
ρ	0.4072	0.6519	0.8814	0.4280	0.6374	0.8846
ω_2^2	0.0000	0.0012	0.0036	0.0005	0.0020	0.0051
τ_{00}	0.0015	0.0033	0.0059	0.0015	0.0049	0.0104
τ_{11}	0.0000	0.0001	0.0004	0.0001	0.0002	0.0004
τ_{22}	0.0000	0.0001	0.0002	0.0000	0.0001	0.0002
Δ_{12}	0.1544	0.3463	0.9997	0.2071	0.2935	0.4260

Note: This table summarizes estimates from Models 1 and 2 estimated in each state-subject dataset. ω_1^2 = residual variance in γ_{grt} values in Model 1; ν_{00} = between-district variance in γ_{grt} values in Model 1; $\rho = \nu_{00}/(\nu_{00} + \omega_1^2)$ is the percent of total variance between rather than within districts; ω_2^2 = residual variance in γ_{grt} values in Model 2; τ_{00} = between-district variance in γ_{grt} values in Model 2; τ_{11} and τ_{22} are between-district variances in grade and year trends, respectively; $\Delta_{12} = (\omega_1^2 - \omega_2^2)/\omega_1^2$ is the percent of unexplained variance in Model 1 that is explained by including linear grade and year trends in Model 2.