Leveraging Lotteries for School Value-Added: Testing and Estimation

Josh Angrist, MIT and NBER

Peter Hull, MIT

Parag Pathak, MIT and NBER

Chris Walters, UC Berkeley and NBER

July 2015

Value-added Models in Education

- Value-added models (VAMs): Used to estimate causal effects of individual teachers and schools on student test scores
- Typical VAM: OLS regression of test scores on school indicators and controls for previous scores; relies on selection-on-observables assumption
- VAM estimates are central to policy decisions
  - Teacher personnel decisions
  - School report cards
  - School closure/restructuring decisions
- Continuing controversy over bias in VAM estimates (Kane et al., 2008, 2013; Rothstein, 2010, 2014; Chetty et al., 2014, 2015)

Lottery Quasi-experiments

- A parallel literature uses admission lotteries to assess causal effects of groups of schools
- Many admissions systems involve forms of random assignment
  - Independent charter school lotteries
  - Random tie-breaking in centralized assignment mechanisms
- Lottery quasi-experiments used to evaluate first-choice public schools, charter schools, pilot schools, magnet schools, exam schools, small high schools (Cullen et al., 2006; Abdulkadiroglu et al., 2011, 2014; Angrist et al., 2013; Dobbie and Fryer, 2013; Bloom and Unterman, 2014; Deming et al., 2014)
- But there are barriers to quasi-experimental VAM estimation: school-specific lottery estimates can be imprecise, and some schools are undersubscribed

This Paper I: Testing for Bias

- This paper uses admission lotteries from Boston, MA to assess and improve school value-added models
- First, we develop and implement a new test for bias
  - We conduct an overidentification test that checks whether OLS VAMs predict admission effects in every lottery
  - Approach builds on existing "forecast bias" tests (Kane and Staiger, 2008; Chetty et al., 2014; Deutsch, 2012; Deming, 2014).
  - Unlike previous studies, we exploit all overidentifying restrictions implied by the VAM framework
  - We decisively reject the null that OLS predicts lottery impacts. This implies bias in conventional VAMs

This Paper II: Improving Value-added Estimates

- Second, we use lotteries to improve on conventional VAMs
- OLS estimates are biased, but lottery estimates are less precise and some schools are undersubscribed, so lottery-based IV model is underidentified
- We develop a hybrid measure of school value-added that combines the advantages of each approach
  - Estimate a random coefficients model describing joint distribution of value-added, OLS bias and lottery compliance
  - Then form empirical Bayes (EB) posterior predictions of school value-added conditional on both sets of estimates
  - Strategy optimally combines OLS and lottery estimates accounting for bias in OLS, underidentification of IV, and sampling error
- Simulations show that hybrid approach reduces statistical error and improves policy targeting

Setting: Public Schools in Boston

- Boston public schools are diverse, with several competing sectors
  - Traditional district schools
  - Charter schools
  - Pilot schools
- Admission processes differ by sector
  - Charter schools: independent lotteries
  - Traditional and pilot schools: centralized mechanism with random tiebreaking within preference/priority groups
- We combine charter lotteries with first-choice district lotteries for children applying to 6th grade from 2006-2013, along with MA administrative data with test score, enrollment and demographic information

Table 1: Boston 6th grade students and schools

	Full sample		Lottery sample	
	Schools	Students	Schools	Applications
Sector	(1)	(2)	(3)	(4)
Traditional	25	18,167	16	4,849
Charter	17	5,585	5	3,655
Pilot	9	4,112	7	1,303
All	51	27,864	28	9,807

Note: the lottery sample includes 8,718 unique applicants.

Value-added Framework: Causal Model

- Constant-effects causal model:

$$Y_i = \mu_0 + \sum_j \beta_j D_{ij} + X_i'\gamma + \epsilon_i$$

- D_{ij} is an indicator for attending school j
- β_j is school j's value-added: causal effect relative to omitted school
- X_i is a vector of baseline covariates including previous achievement
- ϵ_i is unobserved ability
- This is a causal model so ϵ_i may be correlated with D_{ij}
- We'll revisit the constant-effects assumption later

Value-added Framework: OLS

- Corresponding OLS value-added model:

$$Y_i = \alpha_0 + \sum_j \alpha_j D_{ij} + X_i' \Gamma + v_i$$

- This is a regression model so v_i is uncorrelated with right-hand-side variables by definition
- Selection-on-observables implies $\alpha_j = \beta_j \; \forall \; j$
- Otherwise, OLS may be biased:

$$\alpha_j = \beta_j + b_j$$

Lottery Offer Instruments

- Lotteries at $L < J$ schools, with offers indicated by $Z_i = (Z_{i1}, ..., Z_{iL})'$
- Offers at school ℓ are randomly assigned conditional on a vector of stratifying variables $C_{i\ell}$, which includes a dummy for applying to ℓ
- We maintain the assumption that lottery offers are (conditionally) independent of student ability:

$$E[\epsilon_i | C_i, Z_i] = \lambda_0 + C_i' \lambda_c$$

- Implies that lottery offers are valid instruments for school attendance after conditioning on $C_i = (C_{i1}, ..., C_{iL})'$
- Not enough instruments to identify value-added model with lotteries alone
- But we can still ask: can OLS VAM estimates predict lottery effects?

Overidentification Test

- We check predictive accuracy of VAM estimates with a standard overidentification test
- Selection on observables implies $v_i = \epsilon_i$, which in turn implies

$$E[v_i|C_i, Z_i] = \lambda_0 + C_i'\lambda_c$$

- Since v_i is an OLS residual we can test this. Look for $\phi_z = 0$ in the regression

$$v_i = \phi_0 + Z_i'\phi_z + C_i'\phi_c + \omega_i$$

- In practice we see $\hat{v}_i$ rather than v_i so must adjust inference for first-step estimation
- This test has L degrees of freedom. Testable even with a single lottery

Overidentification Test: Other Interpretations

- Note that $v_i = Y_i - \hat{Y}_i$ and consider the regressions

$$Y_i = \rho_0 + Z_i'\rho_z + C_i'\rho_c + \eta_i$$

$$\hat{Y}_i = \psi_0 + Z_i'\psi_z + C_i'\psi_c + u_i$$

- Testing $\phi_z = 0$ is equivalent to testing $\rho_z = \psi_z$
- Captures intuition that effects on test scores should equal effects on VAM-predicted scores

Overidentification Test: Other Interpretations

$$Y_i = \rho_0 + Z_i'\rho_z + C_i'\rho_c + \eta_i$$
$$\hat{Y}_i = \psi_0 + Z_i'\psi_z + C_i'\psi_c + u_i$$

- Finally, note that these equations are the reduced form and first stage of a 2SLS procedure instrumenting $\hat{Y}_i$ with lottery offers in equation for Y_i
- Using all offers as instruments, or using them one at a time, resulting IV coefficients should equal 1
- This links our approach to "forecast bias" tests in Chetty et al. (2014), Kane and Staiger (2008) and Deming (2014), which test whether a single coefficient equals 1 in similar procedures
- But test has L degrees of freedom instead of 1: Number of testable restrictions equals number of quasi-experiments
- Tests that focus only on overidentified IV coefficient leave information on the table

Table 3: Tests for bias in school value-added models

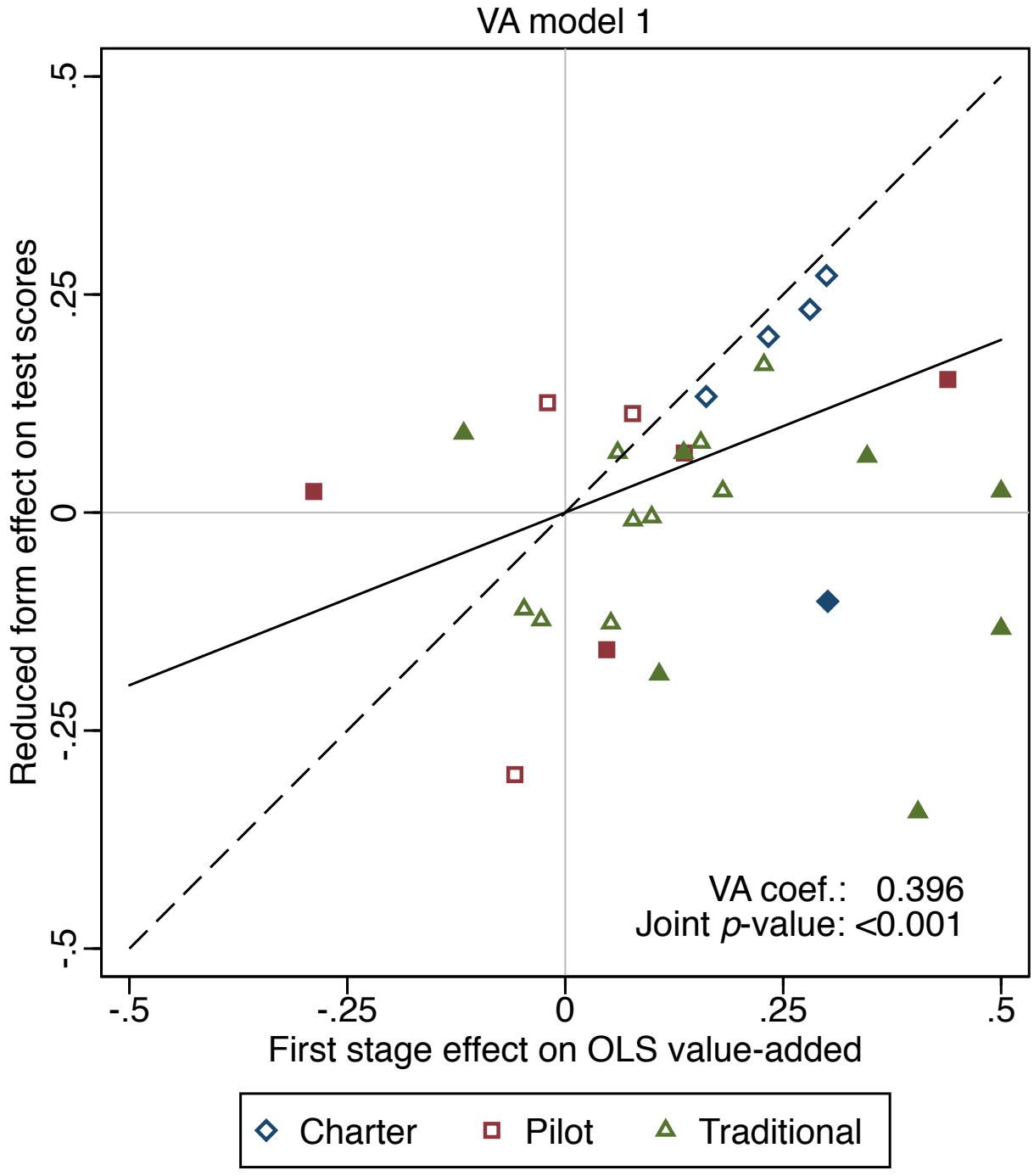
	Model 1 (1)	Model 2 (2)	Model 3 (3)
Value-added coefficient	0.396	0.645	0.864
	(0.056)	(0.065)	(0.075)
First stage *F*-statistic	45.6	36.1	29.6
p-value: VA coef. equals 1	<0.001	<0.001	0.071
p-value: overid. restrictions	<0.001	<0.001	0.003
Joint *p*-value: VA coef. equals 1 and overid. restrictions	<0.001	<0.001	<0.001
Bootstrap-refined joint *p*-value	<0.001	<0.001	<0.001
VAM controls:			
Year effects	X	X	X
Demographics		X	X
Baseline scores			X

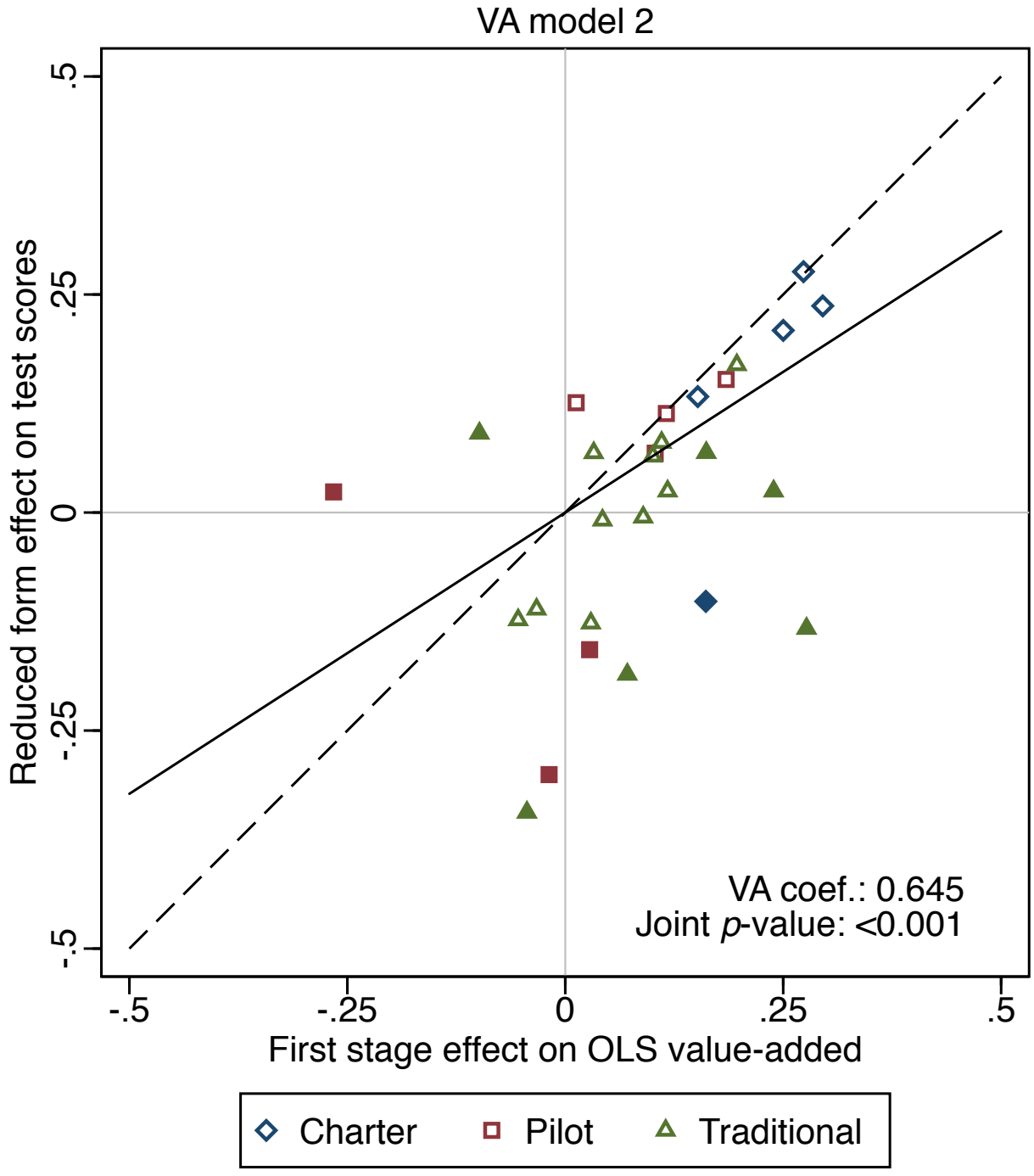
Table 3: Tests for bias in school value-added models

	Model 1 (1)	Model 2 (2)	Model 3 (3)
Value-added coefficient	0.396	0.645	0.864
	(0.056)	(0.065)	(0.075)
First stage *F*-statistic	45.6	36.1	29.6
p-value: VA coef. equals 1	<0.001	<0.001	0.071
p-value: overid. restrictions	<0.001	<0.001	0.003
Joint *p*-value: VA coef. equals 1 and overid. restrictions	<0.001	<0.001	<0.001
Bootstrap-refined joint *p*-value	<0.001	<0.001	<0.001
VAM controls:			
Year effects	X	X	X
Demographics		X	X
Baseline scores			X

Table 3: Tests for bias in school value-added models

	Model 1 (1)	Model 2 (2)	Model 3 (3)
Value-added coefficient	0.396 (0.056)	0.645 (0.065)	0.864 (0.075)
First stage *F*-statistic	45.6	36.1	29.6
p-value: VA coef. equals 1	<0.001	<0.001	0.071
p-value: overid. restrictions	<0.001	<0.001	0.003
Joint *p*-value: VA coef. equals 1 and overid. restrictions	<0.001	<0.001	<0.001
Bootstrap-refined joint *p*-value	<0.001	<0.001	<0.001
VAM controls:			
Year effects	X	X	X
Demographics		X	X
Baseline scores			X


VA model 1
Reduced form effect on test scores
.5
.25
0
-.25
-.5
-.5
-.25
0
.25
.5
First stage effect on OLS value-added
VA coef.: 0.396
Joint p-value: <0.001
Charter
Pilot
Traditional


VA model 2
Reduced form effect on test scores
.5
.25
0
-.25
-.5
-.5
-.25
0
.25
.5
First stage effect on OLS value-added
VA coef.: 0.645
Joint p-value: <0.001
Charter
Pilot
Traditional

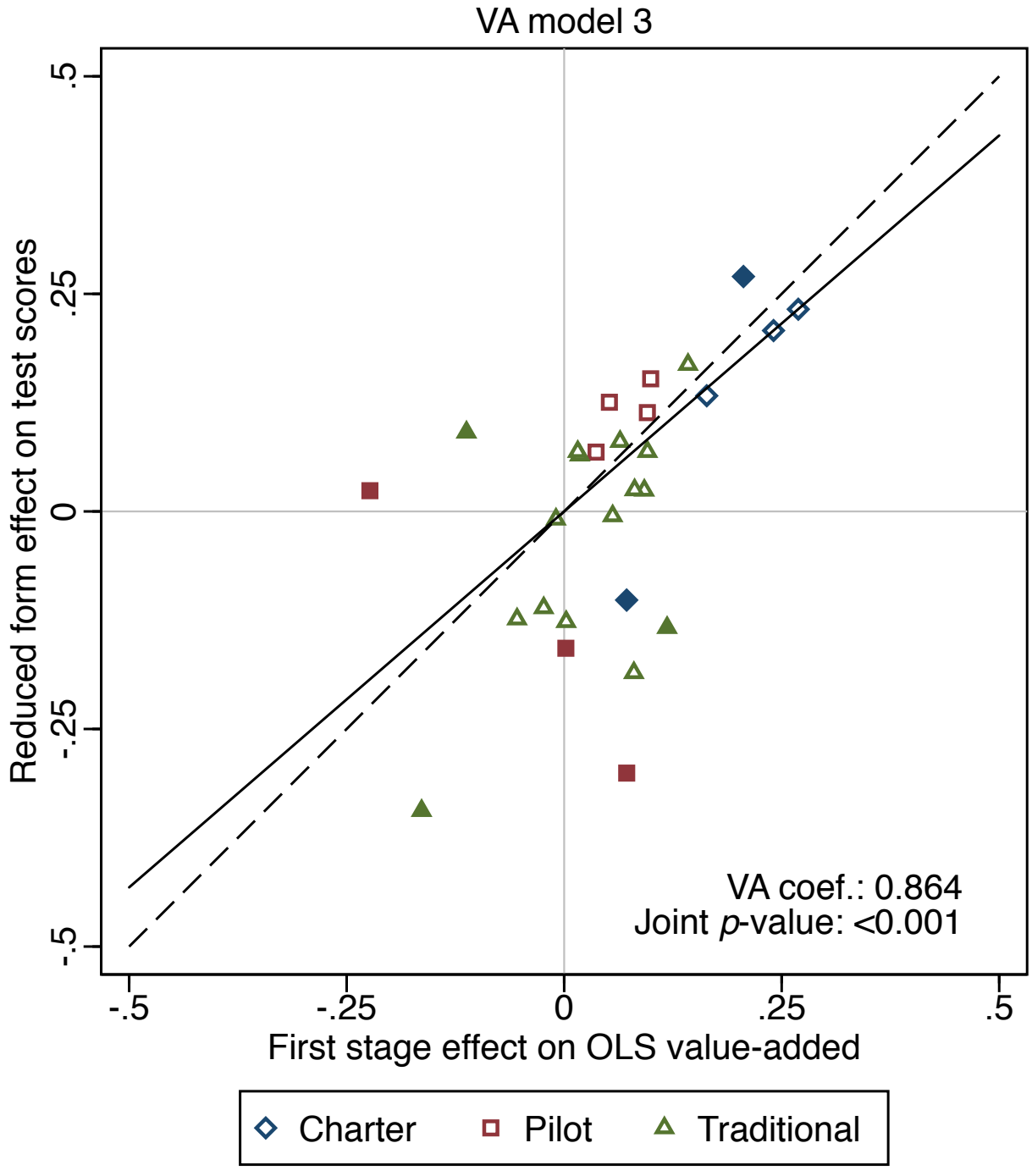

VA model 3
Reduced form effect on test scores
.5
.25
0
-.25
-.5
-.5
-.25
0
.25
.5
First stage effect on OLS value-added
VA coef.: 0.864
Joint p-value: <0.001
Charter
Pilot
Traditional

Table 3: Tests for bias in school value-added models

	Model 1 (1)	Model 2 (2)	Model 3 (3)	Model 3 without charter lotteries (4)
Value-added coefficient	0.396	0.645	0.864	0.549
	(0.056)	(0.065)	(0.075)	(0.164)
First stage *F*-statistic	45.6	36.1	29.6	11.2
p-value: VA coef. equals 1	<0.001	<0.001	0.071	0.006
p-value: overid. restrictions	<0.001	<0.001	0.003	0.043
Joint *p*-value: VA coef. equals 1 and overid. restrictions	<0.001	<0.001	<0.001	0.002
Bootstrap-refined joint *p*-value	<0.001	<0.001	<0.001	0.002
VAM controls:				
Year effects	X	X	X	X
Demographics		X	X	X
Baseline scores			X	X

Bias or Heterogeneity?

- Test results show that VAM predictions differ from lottery effects, implying bias in our constant-effects model
- OLS and lottery estimates may also differ because of effect heterogeneity
- Heterogeneous effects are also a problem for VAMs: OLS identifies variance-of-treatment weighted average effect that may not have predictive value for specific individuals
- Nonetheless, useful to know if rejection is due to bias or heterogeneity. Two explorations:
  - Allow school effects to vary across observable subgroups
  - Limit VAM sample to randomized lottery applicants

Table 4: Robustness of bias tests to effect heterogeneity

	Constant effects	Allowing heterogeneity by:				
		Year	Subsidized lunch	Special education	Baseline score tercile	Interacted groups
	(1)	(2)	(3)	(4)	(5)	(6)
	A. VAM estimates from full sample					
Value-added coefficient	0.864	0.916	0.849	0.863	0.866	0.930
	(0.075)	(0.072)	(0.075)	(0.074)	(0.075)	(0.061)
Bootstrap-refined joint *p*-value	<0.001	<0.001	<0.001	<0.001	<0.001	0.002
	B. VAM estimates from lottery sample					
Value-added coefficient	0.868	0.962	0.851	0.872	0.873	0.934
	(0.070)	(0.068)	(0.069)	(0.070)	(0.070)	(0.052)
Bootstrap-refined joint *p*-value	<0.001	0.002	<0.001	<0.001	<0.001	0.004

Table 4: Robustness of bias tests to effect heterogeneity

		Allowing heterogeneity by:				
	Constant effects	Year	Subsidized lunch	Special education	Baseline score tercile	Interacted groups
	(1)	(2)	(3)	(4)	(5)	(6)
		A. VAM estimates from full sample				
Value-added coefficient	0.864	0.916	0.849	0.863	0.866	0.930
	(0.075)	(0.072)	(0.075)	(0.074)	(0.075)	(0.061)
Bootstrap-refined joint *p*-value	<0.001	<0.001	<0.001	<0.001	<0.001	0.002
		B. VAM estimates from lottery sample				
Value-added coefficient	0.868	0.962	0.851	0.872	0.873	0.934
	(0.070)	(0.068)	(0.069)	(0.070)	(0.070)	(0.052)
Bootstrap-refined joint *p*-value	<0.001	0.002	<0.001	<0.001	<0.001	0.004

Improving Value-added

- Lottery-based tests show that VAM estimates are biased
- But they are also highly predictive – how much does bias matter?
- Next, develop and estimate a model allowing us to:
  1. Quantify magnitudes of variation in value-added and bias
  2. Construct an improved value-added measure combining OLS and lottery estimates

Random Coefficients Building Blocks

- Our random coefficients model fits the joint distribution of OLS, lottery reduced form, and first-stage estimates
- OLS parameter α_j identifies $\beta_j + b_j$
- Reduced form for lottery ℓ identifies a linear combination of value-added parameters:

$$\rho_z^\ell = \sum_j \pi_j^\ell \beta_j$$

- Parameters $\{\pi_j^\ell\}_{\ell=1}^L$ are first-stage coefficients from a regression of D_{ij} on Z_i and C_i

Random Coefficients (Empirical Bayes)

- First level of the model treats observed estimates as noisy measures of school-specific parameters:

$$\hat{\alpha}_j = \beta_j + b_j + e_j^{\alpha}$$

$$\hat{\rho}_z^{\ell} = \sum_j \pi_j^{\ell} \beta_j + e_{\ell}^{\rho}$$

$$\hat{\pi}_j^{\ell} = \pi_j^{\ell} + e_{\ell j}^{\pi}$$

- The e_j are mean-zero estimation errors with distributions approximated by standard asymptotics
- Second level treats school-specific parameters as draws from a distribution in the population of schools: $(\beta_j, b_j, \pi_j) \sim F(\beta, b, \pi; \theta)$
- $F(\cdot; \theta)$ can be viewed as prior distribution in a Bayesian hierarchical model, with hyperparameter θ
- Empirical Bayes (EB) approach (Morris, 1983): Estimate the prior

Random Coefficients Parameterization

- Parameterization of first-stage coefficients:

$$\pi_\ell^\ell = \frac{\exp(\delta_\ell)}{1+\exp(\delta_\ell)}$$

- δ_ℓ can be viewed as mean utility in a binary logit model predicting compliance with the offer in lottery ℓ

$$\pi_j^\ell = -\pi_\ell^\ell \cdot \frac{\exp(\xi_j + \nu_j^\ell)}{1+\sum_{k\neq\ell}\exp(\xi_k+\nu_k^\ell)} \text{ for } j \neq \ell$$

- $\xi_j + \nu_j^\ell$ is mean utility for school j in a multinomial logit model predicting fallback school choices among compliers in lottery ℓ
- Lottery-invariant term ξ_j allows some schools to be consistently likely to serve as fallback options
- Parameterization ensures that theoretical restrictions on first stage are satisfied

Random Coefficients Model

- Each school is characterized by parameters $(\beta_j, b_j, \delta_j, \xi_j)$
- Distributional assumptions:

$$(\beta_j, b_j, \delta_j, \xi_j) \sim N\left((\beta_0, b_0, \delta_0, \xi_0), \Sigma\right),$$

$$\nu_{\ell j} \sim N\left(0, \sigma_\nu^2\right)$$

- $\theta = (\beta_0, b_0, \delta_0, \xi_0, \Sigma, \sigma_\nu^2)$ is vector of hyperparameters governing the distribution of value-added, bias, and lottery compliance
- In some specifications we also allow the mean vector $(\beta_0, b_0, \delta_0, \xi_0)$ to vary by school sector (traditional, charter or pilot)

Simulated Minimum Distance Estimation

- We estimate hyperparameters via simulated minimum distance, using means, variances and covariances of functions of OLS, reduced form and first stage estimates
- Vector of model-predicted moments:

$$\tilde{m}(\theta) = \frac{1}{S} \sum_{s} \tilde{m}^{s}(\theta)$$

- Simulated minimum distance estimator:

$$\hat{\theta}_{SMD} = \arg\min_{\theta} J \left(\hat{m} - \tilde{m}(\theta)\right)' W \left(\hat{m} - \tilde{m}(\theta)\right)$$

- Two-step estimation with estimate of optimal weighting matrix based on independent simulations

Table 5: Joint distribution of causal value-added and OLS bias

Parameter	Description	Model 1 (1)	Model 2 (2)	Model 3 (3)
σ_β	Std. dev. of causal value-added	0.237 (0.064)	0.191 (0.072)	0.211 (0.063)
σ_b	Std. dev. of bias in OLS value-added	0.438 (0.076)	0.338 (0.102)	0.158 (0.067)
ρ	Correlation of value-added and bias	0.008 (0.321)	0.069 (0.479)	-0.227 (0.455)
σ_α	Std. dev. of OLS value-added	0.499 (0.051)	0.399 (0.052)	0.232 (0.025)
	J-statistic (d.f.):	5.76 (4)	7.35 (4)	5.32 (4)
	Overid. *p*-value:	0.218	0.119	0.256

Table 5: Joint distribution of causal value-added and OLS bias

Parameter	Description	Model 1 (1)	Model 2 (2)	Model 3 (3)
σ_β	Std. dev. of causal value-added	0.237 (0.064)	0.191 (0.072)	0.211 (0.063)
σ_b	Std. dev. of bias in OLS value-added	0.438 (0.076)	0.338 (0.102)	0.158 (0.067)
ρ	Correlation of value-added and bias	0.008 (0.321)	0.069 (0.479)	-0.227 (0.455)
σ_α	Std. dev. of OLS value-added	0.499 (0.051)	0.399 (0.052)	0.232 (0.025)
	J-statistic (d.f.):	5.76 (4)	7.35 (4)	5.32 (4)
	Overid. *p*-value:	0.218	0.119	0.256

Table 6: Random coefficient estimates with sector effects

Mean shifters	β_j (1)	b_j (2)
Charter	0.376	0.012
	(0.122)	(0.118)
Pilot	0.062	-0.055
	(0.141)	(0.136)
Within-sector variances		
Std. dev.	0.163	0.129
	(0.065)	(0.056)
Correlation	-0.454	
	(0.423)	

Table 6: Random coefficient estimates with sector effects

Mean shifters	β_j (1)	b_j (2)
Charter	0.376	0.012
	(0.122)	(0.118)
Pilot	0.062	-0.055
	(0.141)	(0.136)
Within-sector variances		
Std. dev.	0.163	0.129
	(0.065)	(0.056)
Correlation	-0.454	
	(0.423)	

Table 6: Random coefficient estimates with sector effects

Mean shifters	β_j (1)	b_j (2)
Charter	0.376	0.012
	(0.122)	(0.118)
Pilot	0.062	-0.055
	(0.141)	(0.136)
Within-sector variances		
Std. dev.	0.163	0.129
	(0.065)	(0.056)
Correlation	-0.454	
	(0.423)	

Empirical Bayes Posteriors

- Once we have the prior, can compute posteriors for individual schools
- EB posterior predictions are widely used to measure effects of individual teachers and schools (Kane et al., 2008; Chetty et al., 2014)
- Conventional shrinkage formula:

$$\alpha_j^* = \lambda_j \hat{\alpha}_j + (1 - \lambda_j)\mu_\alpha,$$

$$\lambda_j = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + Var(e_j^\alpha)}$$

- α_j^* is posterior mean when α_j and e_j^α are normal; otherwise it's the best linear unbiased predictor (BLUP)
- EB posterior mean plugs in estimates of hyperparameters μ_α and σ_α^2

Empirical Bayes Posteriors with Lotteries

- Lotteries provide information that can improve posterior predictions
- First consider scenario with all schools oversubscribed, so $J = L$
- In this case we can estimate a just-identified IV value-added model that uses only lottery variation
- Resulting IV estimates, $\hat{\beta}_j$, are not biased by student sorting
- But IV estimates are likely to be considerably less precise than OLS
- Creates a tradeoff between bias and variance – may want to put some weight on OLS on MSE grounds (Judge and Mittelhammer, 2007)

Hybrid Posterior Means with Just-ID IV

- Hybrid posterior mean in just-identified case:

$$\beta_j^* = w_{1j}\hat{\beta}_j + w_{2j}(\hat{\alpha}_j - b_0) + (1 - w_{1j} - w_{2j})\beta_0$$

- β_j^* is a weighted average of IV ($\hat{\beta}_j$), biased OLS net of mean bias ($\hat{\alpha}_j - b_0$), and grand mean (β_0)
- Shrinkage factors w_{1j} and w_{2j} depend on sampling variance of (e_j^β, e_j^α) along with prior variance Σ; chosen to minimize MSE
- As before, EB version plugs in estimates of β_0, b_0 and Σ

Hybrid Posteriors with Underidentified IV

- When $J > L$ the order condition fails and the IV model is not identified
- Same intuition applies, however: incorporate all available information into posterior distribution
- OLS is always available. Posterior mean conditional on OLS:

$$E\left[\beta_j|\hat{\alpha}_j\right] = \left(\frac{\sigma_\beta^2 + \sigma_{\beta b}}{\sigma_\beta^2 + \sigma_b^2 + 2\sigma_{\beta b} + Var(e_j^\alpha)}\right)(\hat{\alpha}_j - b_0)$$

$$+ \left(\frac{\sigma_b^2 + \sigma_{\beta b} + Var(e_j^\alpha)}{\sigma_\beta^2 + \sigma_b^2 + 2\sigma_{\beta b} + Var(e_j^\alpha)}\right)\beta_0$$

- This is already an improvement over conventional methods because shrinkage factor accounts for bias

Hybrid Posteriors with Underidentified IV

- If $\pi_j^\ell \neq 0$, the reduced form $\hat{\rho}_z^\ell$ puts weight on β_j and can be used to further improve the posterior prediction
- The $\hat{\pi}_j^\ell$ give reduced form weights and contribute additional information due to correlation between δ_j, ξ_j and β_j

Hybrid Posterior Modes

- No closed form solution for hybrid posterior mean $E[\beta_j|\hat{\alpha}, \hat{\rho}, \hat{\pi}]$, but can approximate via MCMC
- We instead work with posterior modes, aka maximum a posteriori estimates
- Modes are much quicker to compute. Solve regularized MLE problem:

$$\Theta^*_{mode} = \arg\max_{\Theta} \log f\left(\hat{\Omega}|\Theta\right) + \log f\left(\Theta|\hat{\theta}_{SMD}\right)$$

- Here $\Theta = (\beta_1, ..., \nu_J^\ell)'$ are the random coefficients and $\hat{\Omega} = (\hat{\alpha}_1 .... \hat{\pi}_J^L)'$ are the observed estimates

Assessing Improvements

- How much do lotteries improve value-added predictions?
- We simulate data from the model implied by our Boston estimates
- Then apply conventional and hybrid strategies to simulated data
- Metrics for evaluating improvements:
  1. Mean squared error
  2. Mistakes in school classification schemes
  3. Predicted test score improvements generated by VA-based policies
- In principle could use optimal functions of posterior for decisionmaking. We use conventional means and hybrid modes for simplicity

Figure 6: Root mean squared error for value-added posterior predictions

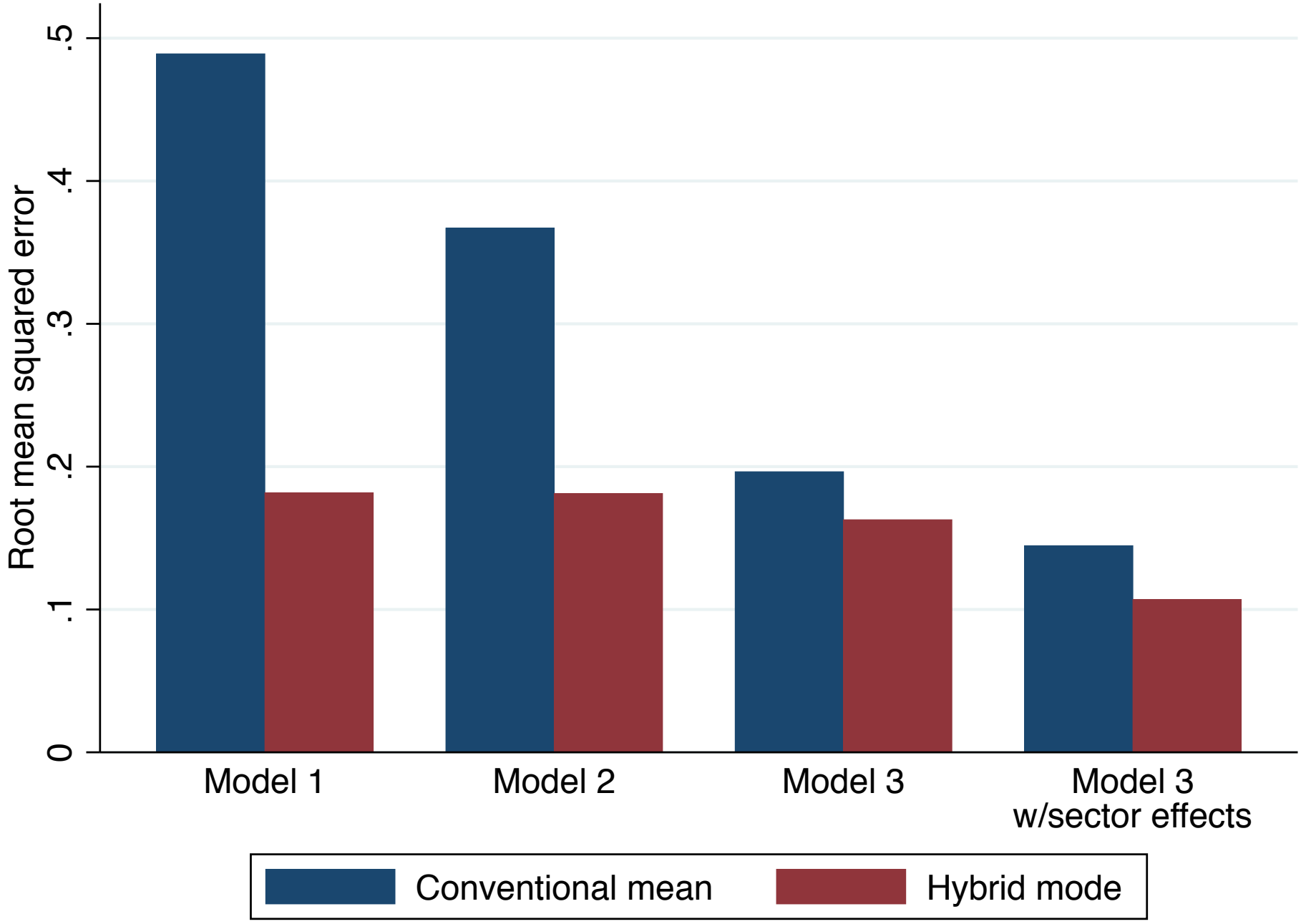


Table 8: Misclassification rates for policies based on value-added predictions

Value-added model	Posterior method	Fraction of failing schools not classified as failing (1)	Fraction of non-failing schools classified as failing (2)
Model 1	Conventional	0.692	0.173
	Hybrid	0.537	0.134
Model 3 w/sector effects	Conventional	0.402	0.101
	Hybrid	0.283	0.071
Model 3 (district schools only)	Conventional	0.458	0.115
	Hybrid	0.337	0.084

Note: Failing schools are those in the bottom value-added quintile.

Table 8: Misclassification rates for policies based on value-added predictions

Value-added model	Posterior method	Fraction of failing schools not classified as failing (1)	Fraction of non-failing schools classified as failing (2)
Model 1	Conventional	0.692	0.173
	Hybrid	0.537	0.134
Model 3 w/sector effects	Conventional	0.402	0.101
	Hybrid	0.283	0.071
Model 3 (district schools only)	Conventional	0.458	0.115
	Hybrid	0.337	0.084

Note: Failing schools are those in the bottom value-added quintile.

Table 8: Misclassification rates for policies based on value-added predictions

Value-added model	Posterior method	Fraction of failing schools not classified as failing (1)	Fraction of non-failing schools classified as failing (2)
Model 1	Conventional	0.692	0.173
	Hybrid	0.537	0.134
Model 3 w/sector effects	Conventional	0.402	0.101
	Hybrid	0.283	0.071
Model 3 (district schools only)	Conventional	0.458	0.115
	Hybrid	0.337	0.084

Note: Failing schools are those in the bottom value-added quintile.

Table 9: Test score impacts of policies based on value-added predictions

Closed school	Replacement school	Conventional (1)	Hybrid (2)	True VA (3)
Lowest-ranked district school	Average district school	0.261	0.313	0.379
	Average charter school	0.621	0.672	0.739
	Average top-quintile district school	0.382	0.492	0.600
	Highest-ranked district school	0.477	0.606	0.736

Table 9: Test score impacts of policies based on value-added predictions

Closed school	Replacement school	Conventional (1)	Hybrid (2)	True VA (3)
Lowest-ranked district school	Average district school	0.261	0.313	0.379
	Average charter school	0.621	0.672	0.739
	Average top-quintile district school	0.382	0.492	0.600
	Highest-ranked district school	0.477	0.606	0.736

Conclusion

- We use school lotteries to assess and improve school value-added estimates
- New test exploiting all available restrictions reveals clear evidence of bias in conventional value-added models
- We develop an estimation procedure that combines OLS and lottery estimates to improve on conventional VAMs
- Simulations show that this hybrid strategy reduces MSE and improves policy targeting
- Proliferation of centralized school assignment mechanisms provides the opportunity for similar validation and hybrid estimation strategies in many districts
- Methods may also be useful in other settings where observational estimates can be combined with quasi-experiments