

From Value Added to Welfare Added: A Social Planner Approach to Education Policy and Statistics

Tanner S Eastmond - UCSD

Nathan Mather - Intensity

Michael Ricks - University of Nebraska, Lincoln

Julian Betts - UCSD and NBER

January 3, 2024

Averages “mean” well, but don’t capture everything

- What is the effect of a teacher, hospital, school, or policy?
 - ▶ Whether in teacher value added, school and hospital report cards, or policy evaluations researchers use mean-oriented statistics

Averages “mean” well, but don’t capture everything

- What is the effect of a teacher, hospital, school, or policy?
 - ▶ Whether in teacher value added, school and hospital report cards, or policy evaluations researchers use mean-oriented statistics
- But mean masks important heterogeneity of multiple types

Averages “mean” well, but don’t capture everything

- What is the effect of a teacher, hospital, school, or policy?
 - ▶ Whether in teacher value added, school and hospital report cards, or policy evaluations researchers use mean-oriented statistics
- But mean masks important heterogeneity of multiple types
 - ▶ **Impacts:** If policies have **heterogeneous effects** on different groups

Averages “mean” well, but don’t capture everything

- What is the effect of a teacher, hospital, school, or policy?
 - ▶ Whether in teacher value added, school and hospital report cards, or policy evaluations researchers use mean-oriented statistics
- But mean masks important heterogeneity of multiple types
 - ▶ **Impacts:** If policies have **heterogeneous effects** on different groups
 - ▶ **Outcomes:** If policies affect **multiple outcomes** in different ways

Averages “mean” well, but don’t capture everything

- What is the effect of a teacher, hospital, school, or policy?
 - ▶ Whether in teacher value added, school and hospital report cards, or policy evaluations researchers use mean-oriented statistics
- But mean masks important heterogeneity of multiple types
 - ▶ **Impacts:** If policies have **heterogeneous effects** on different groups
 - ▶ **Outcomes:** If policies affect **multiple outcomes** in different ways
 - ▶ **Preferences:** If policymakers have **heterogeneous value** over groups (e.g., levels and gaps)

Averages “mean” well, but don’t capture everything

- What is the effect of a teacher, hospital, school, or policy?
 - ▶ Whether in teacher value added, school and hospital report cards, or policy evaluations researchers use mean-oriented statistics
- But mean masks important heterogeneity of multiple types
 - ▶ **Impacts:** If policies have **heterogeneous effects** on different groups
 - ▶ **Outcomes:** If policies affect **multiple outcomes** in different ways
 - ▶ **Preferences:** If policymakers have **heterogeneous value** over groups (e.g., levels and gaps)
- Theoretically, when does heterogeneity matter for maximizing a social objective?

Averages “mean” well, but don’t capture everything

- What is the effect of a teacher, hospital, school, or policy?
 - ▶ Whether in teacher value added, school and hospital report cards, or policy evaluations researchers use mean-oriented statistics
- But mean masks important heterogeneity of multiple types
 - ▶ **Impacts:** If policies have **heterogeneous effects** on different groups
 - ▶ **Outcomes:** If policies affect **multiple outcomes** in different ways
 - ▶ **Preferences:** If policymakers have **heterogeneous value** over groups (e.g., levels and gaps)
- Theoretically, when does heterogeneity matter for maximizing a social objective?
- Empirically, how large are the welfare gains from accounting for heterogeneity
 - ▶ Application: Getting “welfare added” from teacher value added measures (today’s focus for time)

Emerging evidence suggests heterogeneity matters

- Researchers promote “value added”-like estimates for policy in many contexts

Emerging evidence suggests heterogeneity matters

- Researchers promote “value added”-like estimates for policy in many contexts
 - ▶ Heterogeneous and multidimensional value added are the emerging frontier

(Jackson, 2018, Petek and Pope, forthcoming, Delgado, 2022, Ahn et al., 2021, Bates et al., 2022, Dahlstrand, 2022)

Emerging evidence suggests heterogeneity matters

- Researchers promote “value added”-like estimates for policy in many contexts

- ▶ Heterogeneous and multidimensional value added are the emerging frontier

(Jackson, 2018, Petek and Pope, forthcoming, Delgado, 2022, Ahn et al., 2021, Bates et al., 2022, Dahlstrand, 2022)

- ▶ But both heterogeneity and multidimensionality undermine ranking units and allocations

(Condie et al., 2014)

Emerging evidence suggests heterogeneity matters

- Researchers promote “value added”-like estimates for policy in many contexts

- ▶ Heterogeneous and multidimensional value added are the emerging frontier

(Jackson, 2018, Petek and Pope, forthcoming, Delgado, 2022, Ahn et al., 2021, Bates et al., 2022, Dahlstrand, 2022)

- ▶ But both heterogeneity and multidimensionality undermine ranking units and allocations

(Condie et al., 2014)

- In the real world, numerous policies have distributional objectives

Emerging evidence suggests heterogeneity matters

- Researchers promote “value added”-like estimates for policy in many contexts

- ▶ Heterogeneous and multidimensional value added are the emerging frontier

(Jackson, 2018, Petek and Pope, forthcoming, Delgado, 2022, Ahn et al., 2021, Bates et al., 2022, Dahlstrand, 2022)

- ▶ But both heterogeneity and multidimensionality undermine ranking units and allocations

(Condie et al., 2014)

- In the real world, numerous policies have distributional objectives

- ▶ Heterogeneous weights from public finance can extend to cost-benefit analysis

(Adler, 2016, Fleurbaey and Abi-Rafeh, 2016)

Emerging evidence suggests heterogeneity matters

■ Researchers promote “value added”-like estimates for policy in many contexts

- ▶ Heterogeneous and multidimensional value added are the emerging frontier

(Jackson, 2018, Petek and Pope, forthcoming, Delgado, 2022, Ahn et al., 2021, Bates et al., 2022, Dahlstrand, 2022)

- ▶ But both heterogeneity and multidimensionality undermine ranking units and allocations

(Condie et al., 2014)

■ In the real world, numerous policies have distributional objectives

- ▶ Heterogeneous weights from public finance can extend to cost-benefit analysis

(Adler, 2016, Fleurbaey and Abi-Rafteh, 2016)

- ▶ Can welfare theory also make value added useful despite heterogeneity?

The key is moving from value added to “welfare-added”

- We use a welfare-theoretic framework to approach a policymaker's objective

The key is moving from value added to “welfare-added”

- We use a welfare-theoretic framework to approach a policymaker's objective
 - ▶ Framework fully ranks allocations, given weights on outcomes and groups

The key is moving from value added to “welfare-added”

- We use a welfare-theoretic framework to approach a policymaker's objective
 - ▶ Framework fully ranks allocations, given weights on outcomes and groups
 - ▶ We characterize three gains from including heterogeneity in estimation

The key is moving from value added to “welfare-added”

- We use a welfare-theoretic framework to approach a policymaker's objective
 - ▶ Framework fully ranks allocations, given weights on outcomes and groups
 - ▶ We characterize three gains from including heterogeneity in estimation
1. Using **comparative advantage** makes better allocations than only absolute

The key is moving from value added to “welfare-added”

- We use a welfare-theoretic framework to approach a policymaker's objective
 - ▶ Framework fully ranks allocations, given weights on outcomes and groups
 - ▶ We characterize three gains from including heterogeneity in estimation
- 1. Using **comparative advantage** makes better allocations than only absolute
 - ▶ Related to literatures on targeting for health and public programs

The key is moving from value added to “welfare-added”

- We use a welfare-theoretic framework to approach a policymaker's objective
 - ▶ Framework fully ranks allocations, given weights on outcomes and groups
 - ▶ We characterize three gains from including heterogeneity in estimation
- 1. Using **comparative advantage** makes better allocations than only absolute
 - ▶ Related to literatures on targeting for health and public programs
- 2. Directly addressing distributional or **equity concerns** controls relevant gaps

The key is moving from value added to “welfare-added”

- We use a welfare-theoretic framework to approach a policymaker's objective
 - ▶ Framework fully ranks allocations, given weights on outcomes and groups
 - ▶ We characterize three gains from including heterogeneity in estimation
- 1. Using **comparative advantage** makes better allocations than only absolute
 - ▶ Related to literatures on targeting for health and public programs
- 2. Directly addressing distributional or **equity concerns** controls relevant gaps
 - ▶ As in research in optimal tax, discrimination, and environmental justice

The key is moving from value added to “welfare-added”

- We use a welfare-theoretic framework to approach a policymaker's objective
 - ▶ Framework fully ranks allocations, given weights on outcomes and groups
 - ▶ We characterize three gains from including heterogeneity in estimation
- 1. Using **comparative advantage** makes better allocations than only absolute
 - ▶ Related to literatures on targeting for health and public programs
- 2. Directly addressing distributional or **equity concerns** controls relevant gaps
 - ▶ As in research in optimal tax, discrimination, and environmental justice
- 3. Heterogeneous estimators may be more **externally valid**

The key is moving from value added to “welfare-added”

- We use a welfare-theoretic framework to approach a policymaker's objective
 - ▶ Framework fully ranks allocations, given weights on outcomes and groups
 - ▶ We characterize three gains from including heterogeneity in estimation
- 1. Using **comparative advantage** makes better allocations than only absolute
 - ▶ Related to literatures on targeting for health and public programs
- 2. Directly addressing distributional or **equity concerns** controls relevant gaps
 - ▶ As in research in optimal tax, discrimination, and environmental justice
- 3. Heterogeneous estimators may be more **externally valid**
 - ▶ Like papers on drug trials and selection on unobservables

This theory of heterogeneity is crucial in teacher value added

- We estimate **heterogeneous, multidimensional** value added in San Diego Unified
 - ▶ Over 1,700 teachers in 3-5th grade classrooms between 2002-2013

This theory of heterogeneity is crucial in teacher value added

- We estimate **heterogeneous**, **multidimensional** value added in San Diego Unified
 - ▶ Over 1,700 teachers in 3-5th grade classrooms between 2002-2013

This theory of heterogeneity is crucial in teacher value added

- We estimate **heterogeneous**, **multidimensional** value added in San Diego Unified
 - ▶ Over 1,700 teachers in 3-5th grade classrooms between 2002-2013

Group A

--	--

Group B

--	--

This theory of heterogeneity is crucial in teacher value added

- We estimate **heterogeneous, multidimensional** value added in San Diego Unified
 - ▶ Over 1,700 teachers in 3-5th grade classrooms between 2002-2013

Group A

--	--

Group B

--	--

This theory of heterogeneity is crucial in teacher value added

- We estimate **heterogeneous, multidimensional** value added in San Diego Unified
 - ▶ Over 1,700 teachers in 3-5th grade classrooms between 2002-2013

	Outcome 1	Outcome 2
Group A		
Group B		

This theory of heterogeneity is crucial in teacher value added

- We estimate **heterogeneous, multidimensional** value added in San Diego Unified
 - ▶ Over 1,700 teachers in 3-5th grade classrooms between 2002-2013

	Outcome 1	Outcome 2
Group A		
Group B		

This theory of heterogeneity is crucial in teacher value added

- We estimate **heterogeneous, multidimensional** value added in San Diego Unified
 - ▶ Over 1,700 teachers in 3-5th grade classrooms between 2002-2013
 - ▶ Our main application focuses on **achievement** impacts on **multiple subjects**

	Outcome 1	Outcome 2
Group A		
Group B		

This theory of heterogeneity is crucial in teacher value added

- We estimate **heterogeneous, multidimensional** value added in San Diego Unified
 - ▶ Over 1,700 teachers in 3-5th grade classrooms between 2002-2013
 - ▶ Our main application focuses on **achievement** impacts on **multiple subjects**

	Outcome 1	Outcome 2
Group A	High Achieving Math	
Group B		

This theory of heterogeneity is crucial in teacher value added

- We estimate **heterogeneous, multidimensional** value added in San Diego Unified
 - ▶ Over 1,700 teachers in 3-5th grade classrooms between 2002-2013
 - ▶ Our main application focuses on **achievement** impacts on **multiple subjects**

	Outcome 1	Outcome 2
Group A	High Achieving Math	High Achieving ELA
Group B		

This theory of heterogeneity is crucial in teacher value added

- We estimate **heterogeneous, multidimensional** value added in San Diego Unified
 - ▶ Over 1,700 teachers in 3-5th grade classrooms between 2002-2013
 - ▶ Our main application focuses on **achievement** impacts on **multiple subjects**

	Outcome 1	Outcome 2
Group A	High Achieving Math	High Achieving ELA
Group B	Low Achieving Math	

This theory of heterogeneity is crucial in teacher value added

- We estimate **heterogeneous, multidimensional** value added in San Diego Unified
 - ▶ Over 1,700 teachers in 3-5th grade classrooms between 2002-2013
 - ▶ Our main application focuses on **achievement** impacts on **multiple subjects**

	Outcome 1	Outcome 2
Group A	High Achieving Math	High Achieving ELA
Group B	Low Achieving Math	Low Achieving ELA

This theory of heterogeneity is crucial in teacher value added

- We estimate **heterogeneous, multidimensional** value added in San Diego Unified
 - ▶ Over 1,700 teachers in 3-5th grade classrooms between 2002-2013
 - ▶ Our main application focuses on **achievement** impacts on **multiple subjects**
- Accounting for heterogeneity results in large welfare gains

This theory of heterogeneity is crucial in teacher value added

- We estimate **heterogeneous, multidimensional** value added in San Diego Unified
 - ▶ Over 1,700 teachers in 3-5th grade classrooms between 2002-2013
 - ▶ Our main application focuses on **achievement** impacts on **multiple subjects**
- Accounting for heterogeneity results in large welfare gains

This theory of heterogeneity is crucial in teacher value added

- We estimate **heterogeneous, multidimensional** value added in San Diego Unified
 - ▶ Over 1,700 teachers in 3-5th grade classrooms between 2002-2013
 - ▶ Our main application focuses on **achievement** impacts on **multiple subjects**
- Accounting for heterogeneity results in large welfare gains
 1. **Substantial heterogeneity**: average comparative advantage is 48-53% of a standard deviation in absolute advantage

This theory of heterogeneity is crucial in teacher value added

- We estimate **heterogeneous, multidimensional** value added in San Diego Unified
 - ▶ Over 1,700 teachers in 3-5th grade classrooms between 2002-2013
 - ▶ Our main application focuses on **achievement** impacts on **multiple subjects**
- Accounting for heterogeneity results in large welfare gains
 1. **Substantial heterogeneity**: average comparative advantage is 48-53% of a standard deviation in absolute advantage
 2. **Efficiency gains**: teacher reallocations using comparative advantage raise average scores 34-97% more than absolute advantage

This theory of heterogeneity is crucial in teacher value added

- We estimate **heterogeneous, multidimensional** value added in San Diego Unified
 - ▶ Over 1,700 teachers in 3-5th grade classrooms between 2002-2013
 - ▶ Our main application focuses on **achievement** impacts on **multiple subjects**

- Accounting for heterogeneity results in large welfare gains
 1. **Substantial heterogeneity**: average comparative advantage is 48-53% of a standard deviation in absolute advantage
 2. **Efficiency gains**: teacher reallocations using comparative advantage raise average scores 34-97% more than absolute advantage
 3. **Equity gains**: Reallocations using comparative advantage could close gaps and raise average scores

This theory of heterogeneity is crucial in teacher value added

- We estimate **heterogeneous, multidimensional** value added in San Diego Unified
 - ▶ Over 1,700 teachers in 3-5th grade classrooms between 2002-2013
 - ▶ Our main application focuses on **achievement** impacts on **multiple subjects**

- Accounting for heterogeneity results in large welfare gains
 1. **Substantial heterogeneity:** average comparative advantage is 48-53% of a standard deviation in absolute advantage
 2. **Efficiency gains:** teacher reallocations using comparative advantage raise average scores 34-97% more than absolute advantage
 3. **Equity gains:** Reallocations using comparative advantage could close gaps and raise average scores
 4. **Welfare impacts:** Combined optimization could increase present valued earnings by \$83 M (34% above ignoring multidimensionality)

Today's Talk

1. When Will Heterogeneity Matter for Welfare?

- ▶ Graphical Intuition of Theory

2. What Are the Implications for Value Added?

- ▶ Heterogeneous Impacts
- ▶ Reallocations and Decompositions

3. What About Welfare and Policy?



When Will Heterogeneity Matter for Welfare?

Comparative advantage and distributional concerns affect welfare



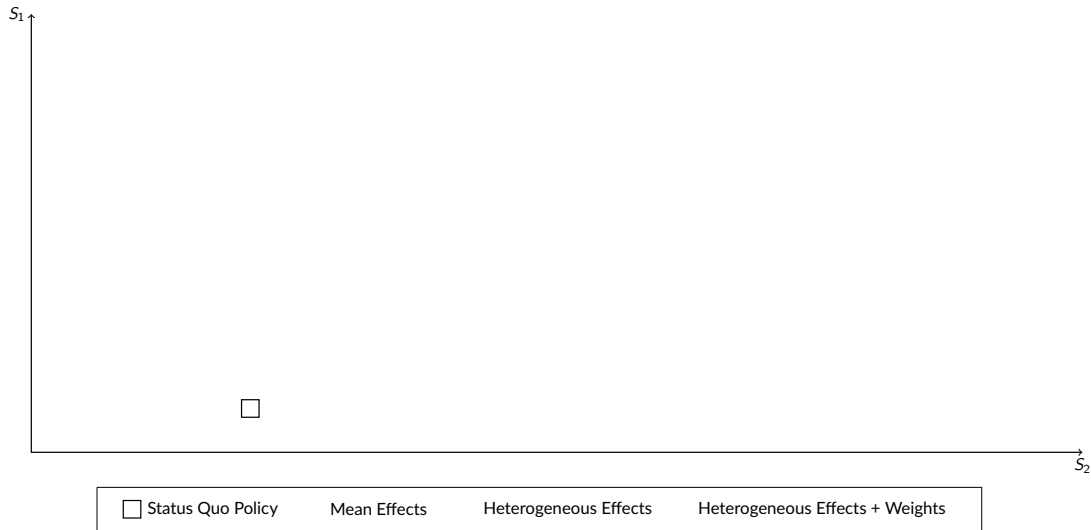
Status Quo Policy

Mean Effects

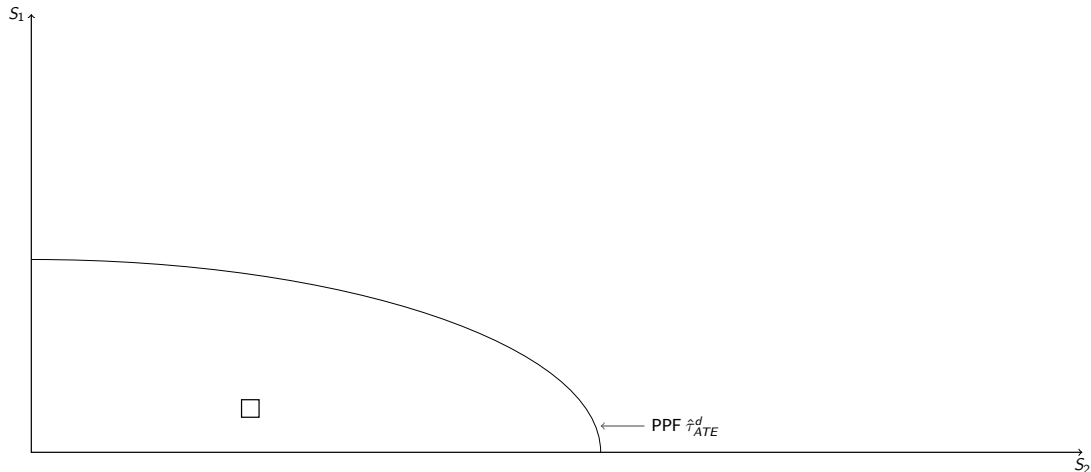
Heterogeneous Effects

Heterogeneous Effects + Weights

Comparative advantage and distributional concerns affect welfare



Comparative advantage and distributional concerns affect welfare



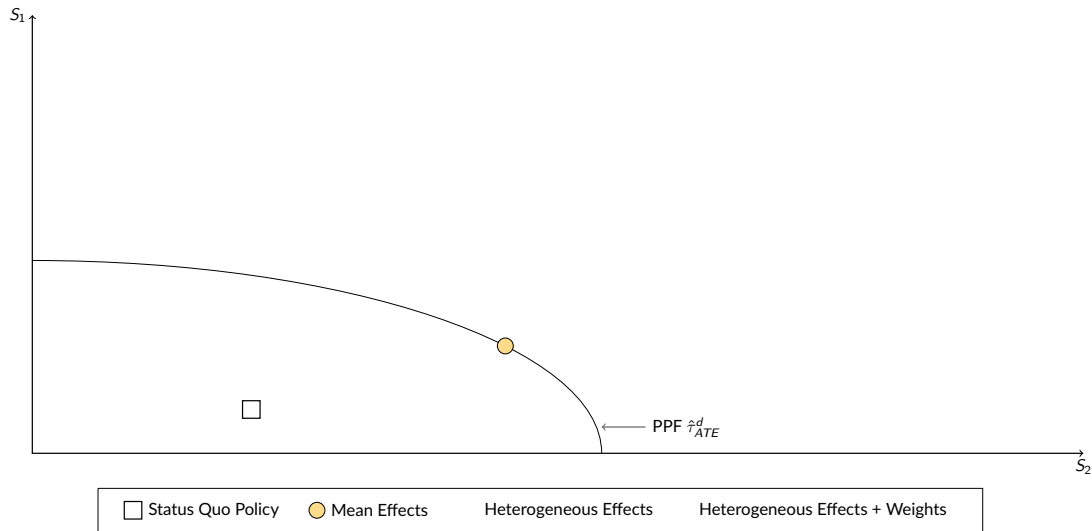
Status Quo Policy

Mean Effects

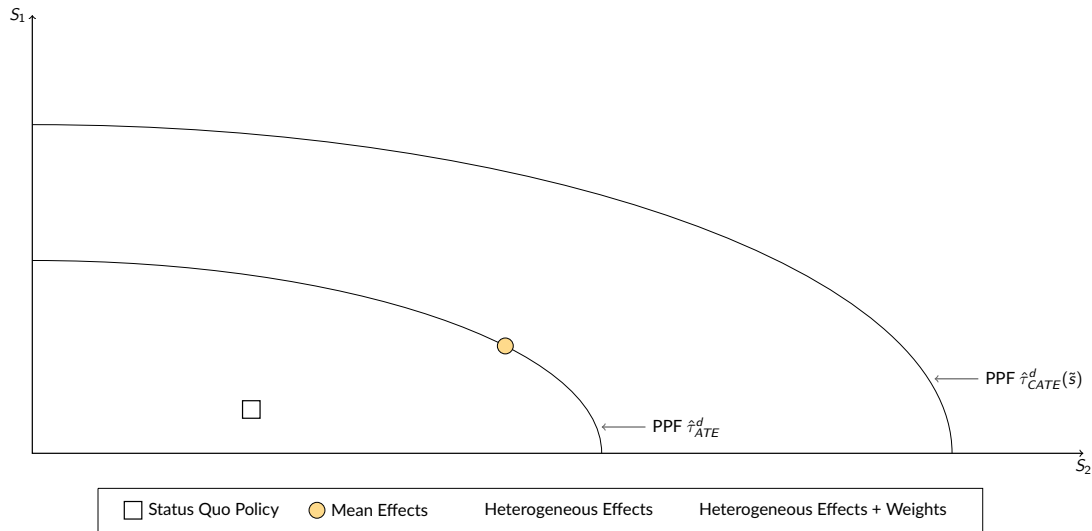
Heterogeneous Effects

Heterogeneous Effects + Weights

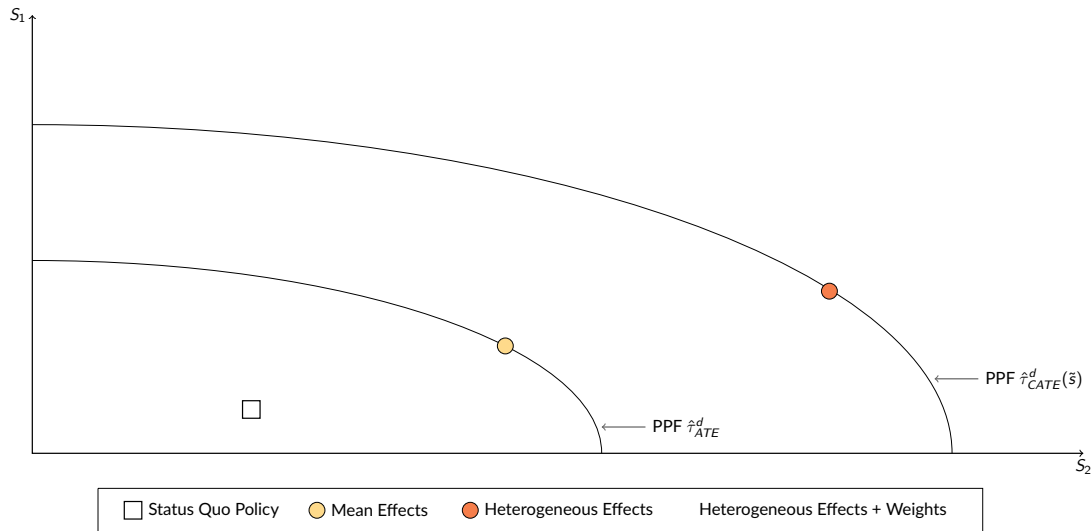
Comparative advantage and distributional concerns affect welfare



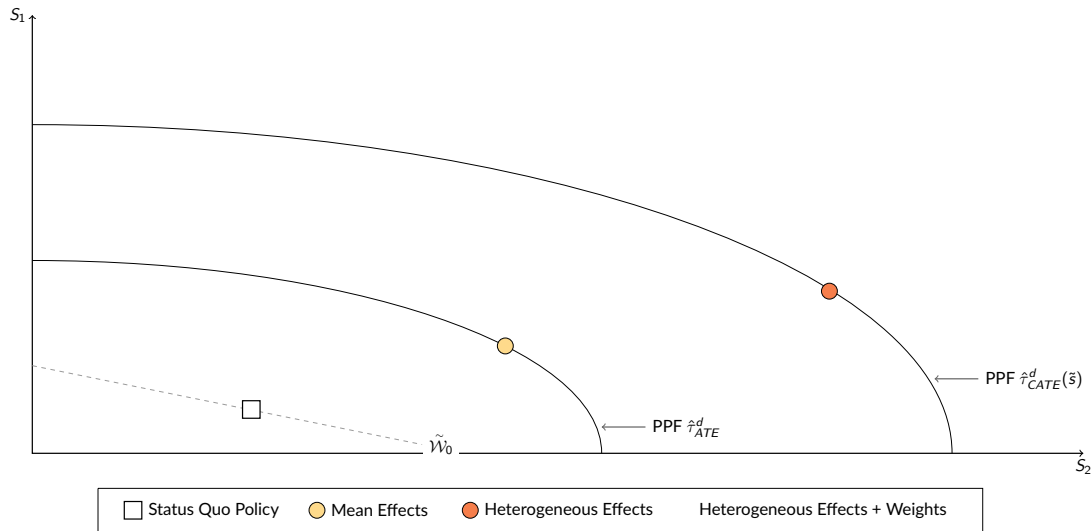
Comparative advantage and distributional concerns affect welfare



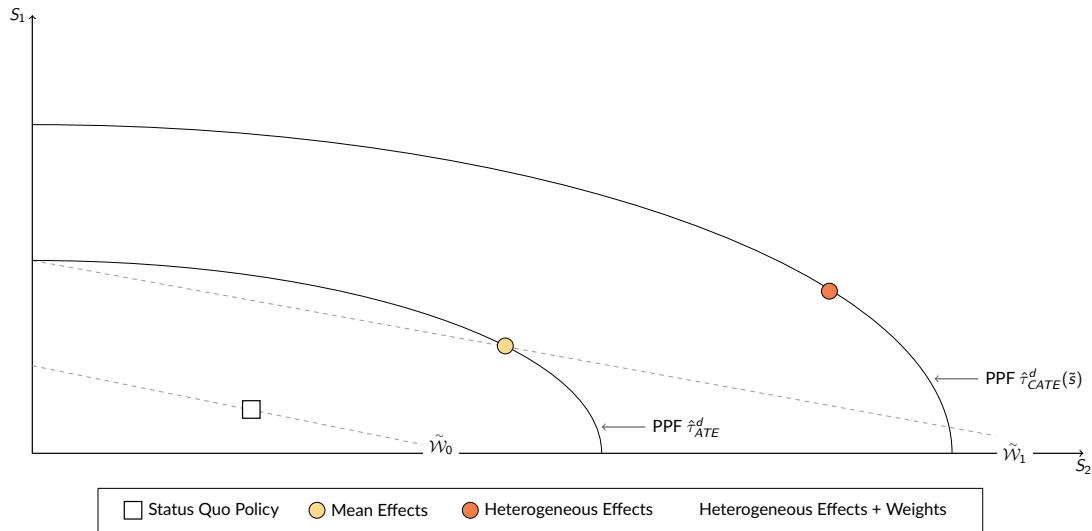
Comparative advantage and distributional concerns affect welfare



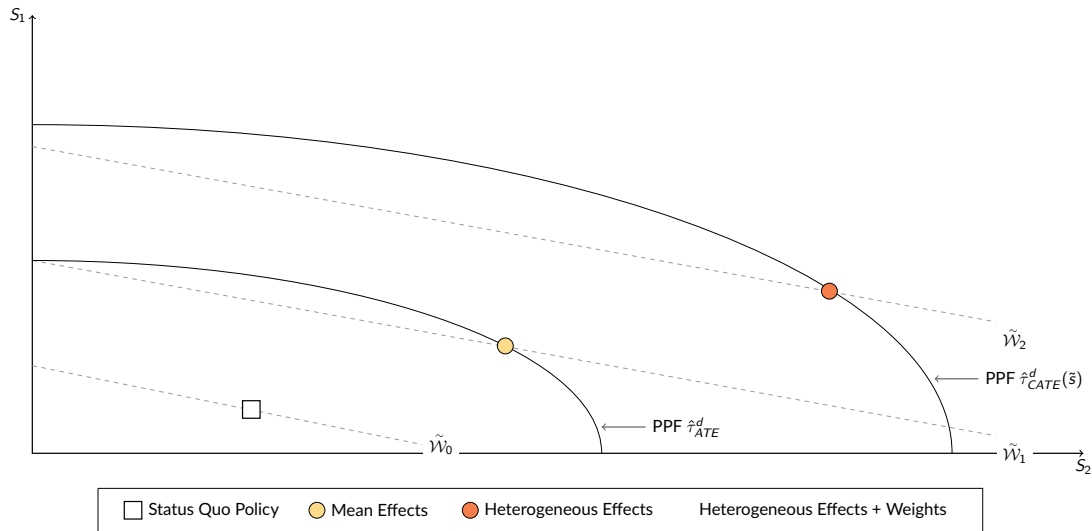
Comparative advantage and distributional concerns affect welfare



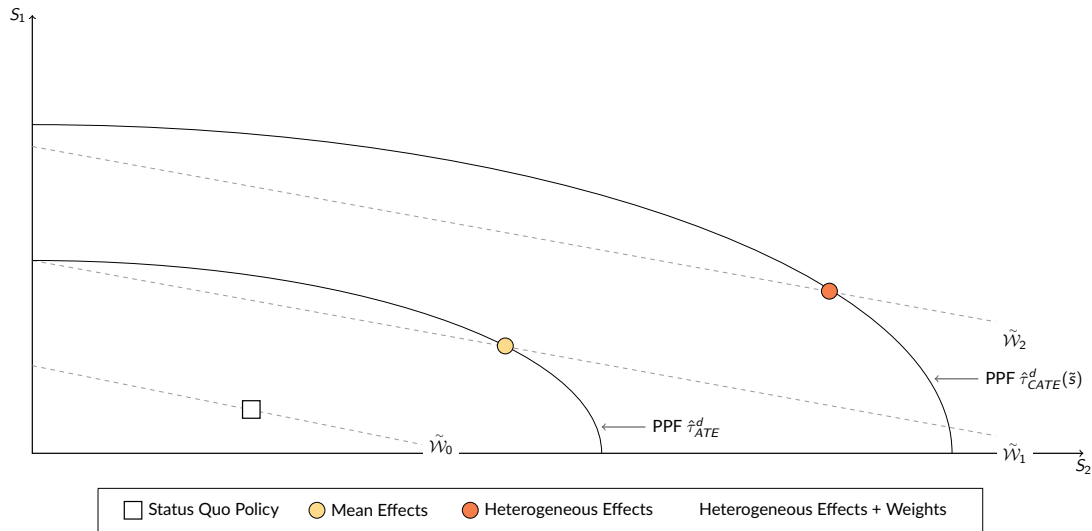
Comparative advantage and distributional concerns affect welfare



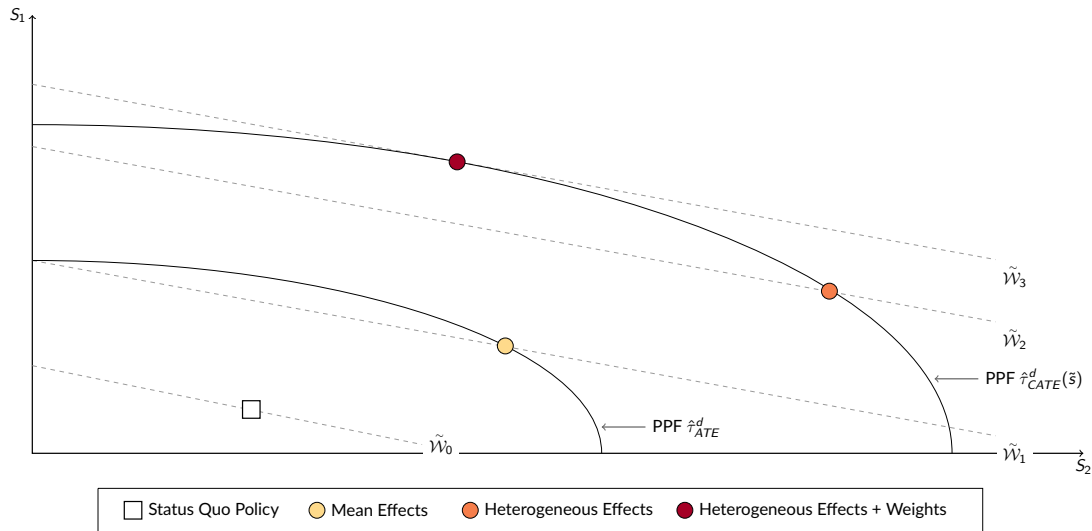
Comparative advantage and distributional concerns affect welfare



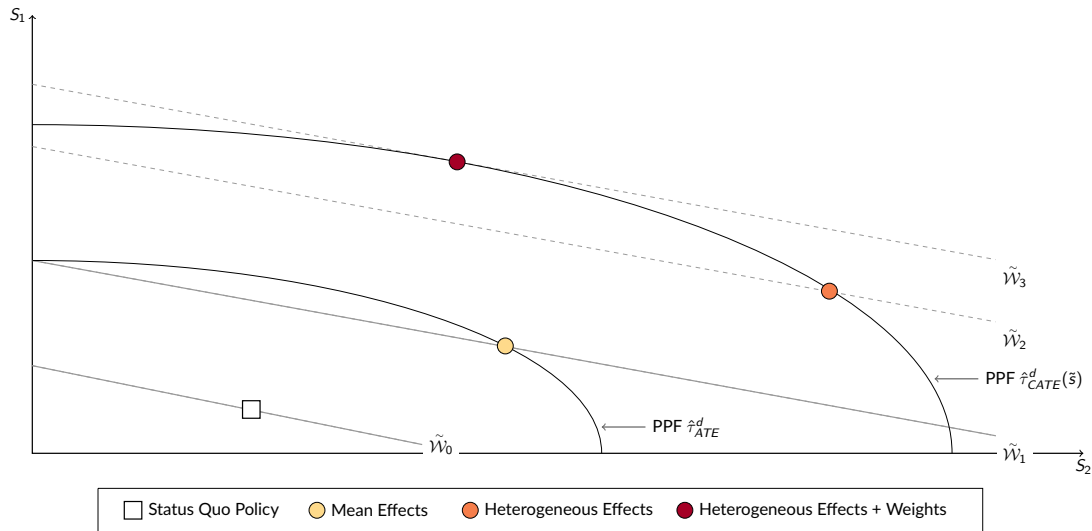
Comparative advantage and distributional concerns affect welfare



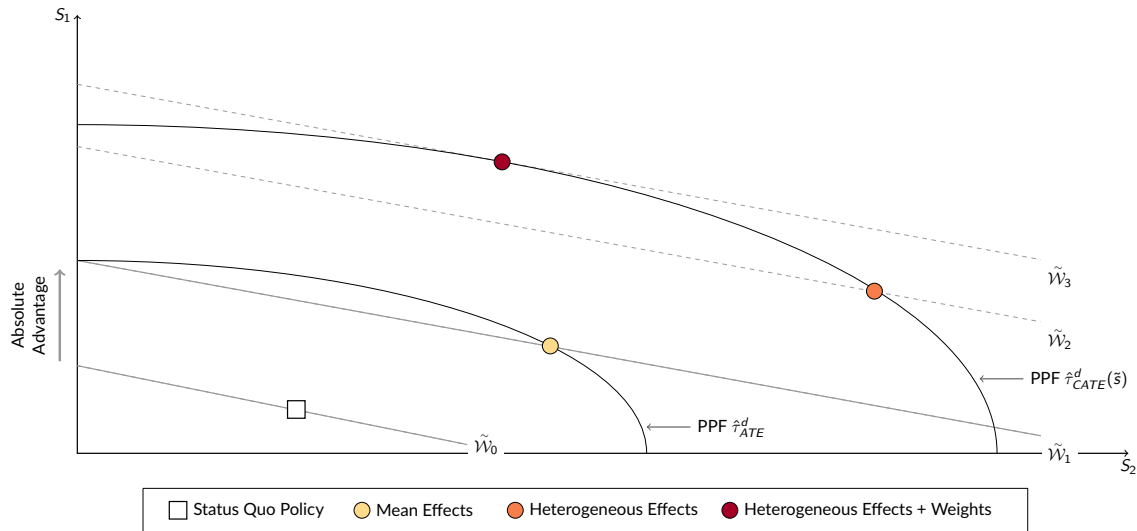
Comparative advantage and distributional concerns affect welfare



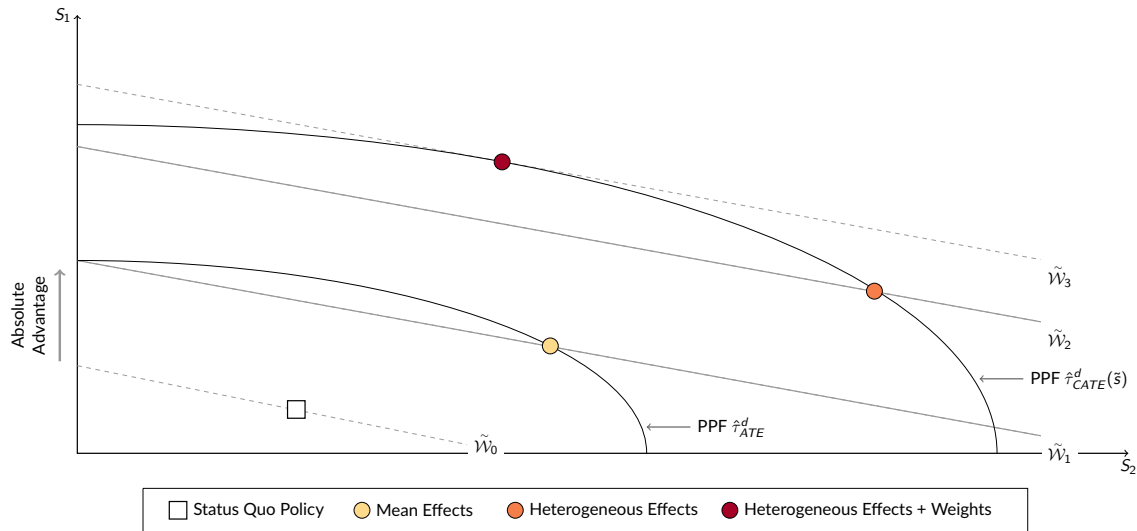
Comparative advantage and distributional concerns affect welfare



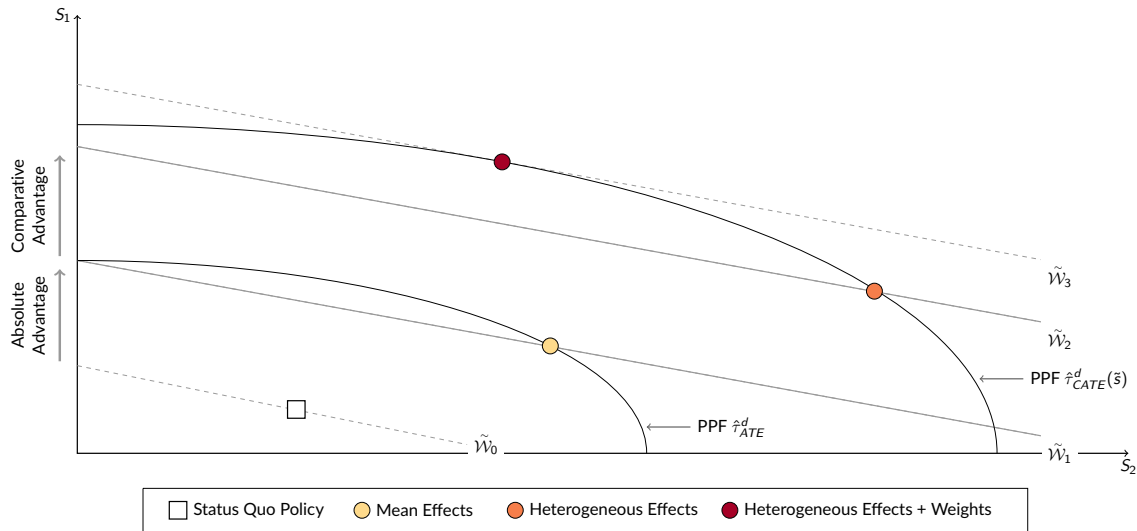
Comparative advantage and distributional concerns affect welfare



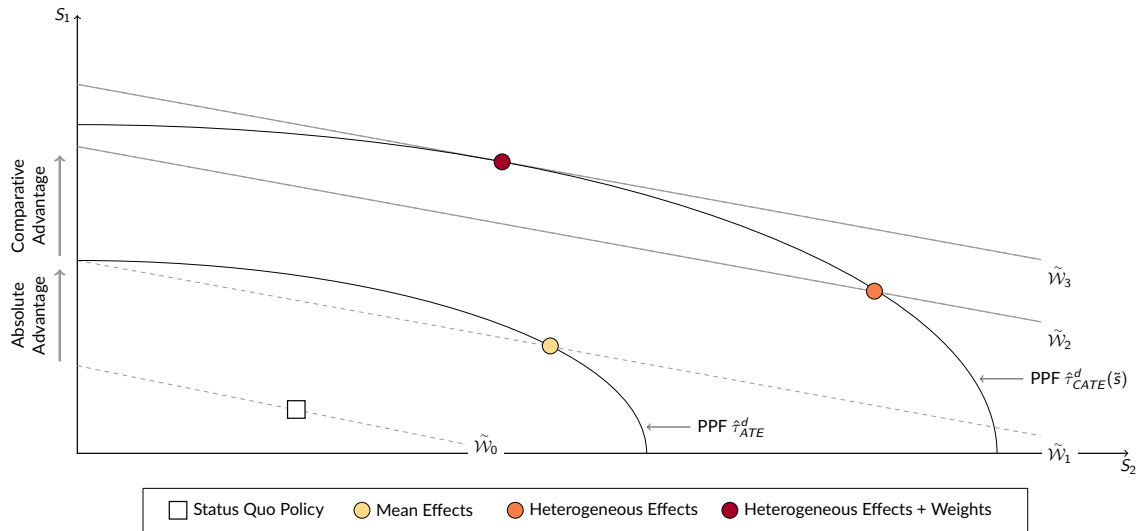
Comparative advantage and distributional concerns affect welfare



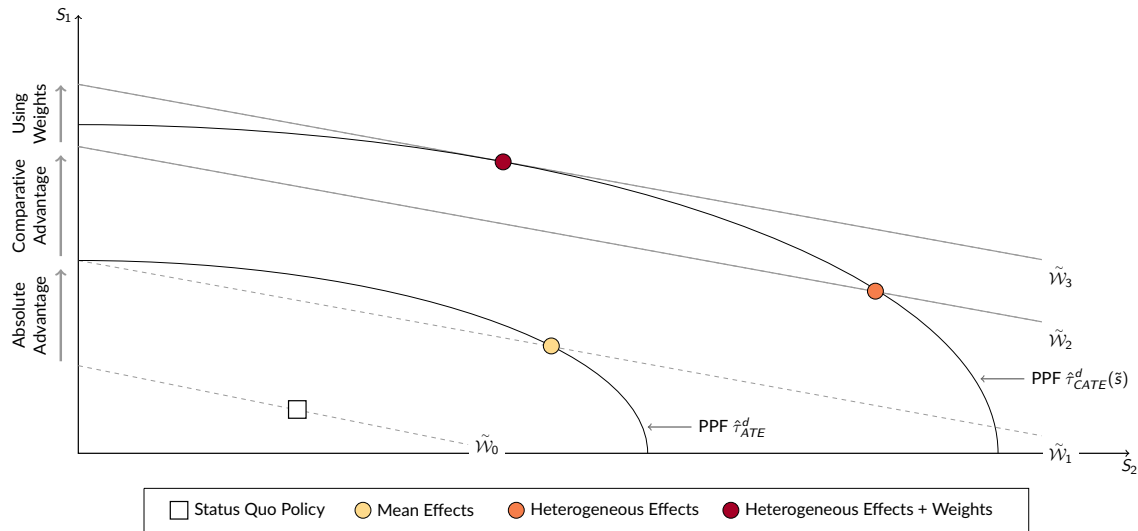
Comparative advantage and distributional concerns affect welfare



Comparative advantage and distributional concerns affect welfare



Comparative advantage and distributional concerns affect welfare



Application: Welfare and Heterogeneous Value Added

Estimating Heterogeneous Teacher Value Added

Estimating value added in San Deigo Unified

- Empirical setting is the San Diego Unified School District
 - ▶ Second largest district in CA, broadly representative of the state
 - ▶ 1,791 teachers in grades 3-5 classrooms from 2002-2013

Estimating value added in San Deigo Unified

- Empirical setting is the San Diego Unified School District
 - ▶ Second largest district in CA, broadly representative of the state
 - ▶ 1,791 teachers in grades 3-5 classrooms from 2002-2013
- We explore **heterogeneity in value added** along the achievement distribution

Estimating value added in San Deigo Unified

- Empirical setting is the San Diego Unified School District
 - ▶ Second largest district in CA, broadly representative of the state
 - ▶ 1,791 teachers in grades 3-5 classrooms from 2002-2013
- We explore **heterogeneity in value added** along the achievement distribution
 - ▶ Achievement seems like the most natural dimension of **heterogeneity in teacher effects** (think: teaching differentiation) and **social preferences** (think: policy examples)

Estimating value added in San Deigo Unified

- Empirical setting is the San Diego Unified School District
 - ▶ Second largest district in CA, broadly representative of the state
 - ▶ 1,791 teachers in grades 3-5 classrooms from 2002-2013
- We explore **heterogeneity in value added** along the achievement distribution
 - ▶ Achievement seems like the most natural dimension of **heterogeneity in teacher effects** (think: teaching differentiation) and **social preferences** (think: policy examples)
 - ▶ Split estimates by median score in the district separately for **math and ELA**

Estimating value added in San Deigo Unified

- Empirical setting is the San Diego Unified School District
 - ▶ Second largest district in CA, broadly representative of the state
 - ▶ 1,791 teachers in grades 3-5 classrooms from 2002-2013
- We explore **heterogeneity in value added** along the achievement distribution
 - ▶ Achievement seems like the most natural dimension of **heterogeneity in teacher effects** (think: teaching differentiation) and **social preferences** (think: policy examples)
 - ▶ Split estimates by median score in the district separately for **math and ELA**
- Very standard VA estimation adapted for heterogeneity following Chetty et al. (2014a), Delgado (2022), and Bates et al. (2022)

VA Estimation Details

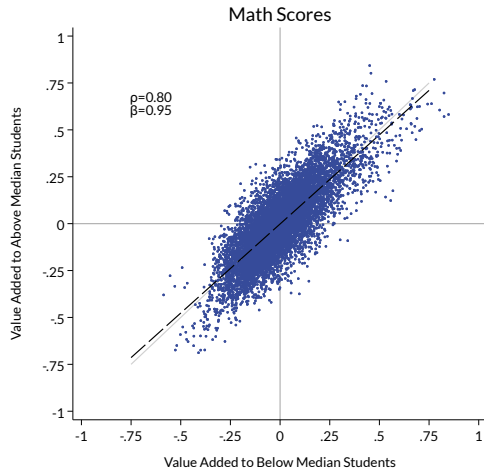
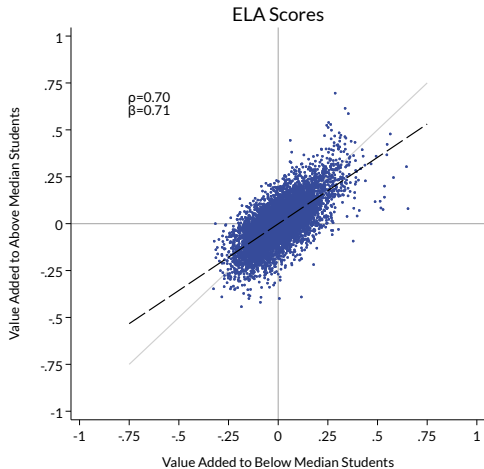
Estimating value added in San Deigo Unified

- Empirical setting is the San Diego Unified School District
 - ▶ Second largest district in CA, broadly representative of the state
 - ▶ 1,791 teachers in grades 3-5 classrooms from 2002-2013
- We explore **heterogeneity in value added** along the achievement distribution
 - ▶ Achievement seems like the most natural dimension of **heterogeneity in teacher effects** (think: teaching differentiation) and **social preferences** (think: policy examples)
 - ▶ Split estimates by median score in the district separately for **math and ELA**
- Very standard VA estimation adapted for heterogeneity following Chetty et al. (2014a), Delgado (2022), and Bates et al. (2022)

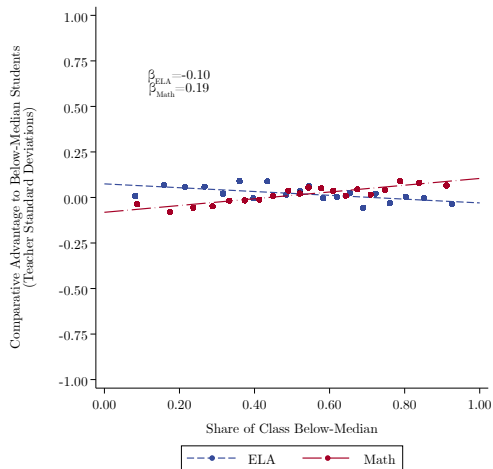
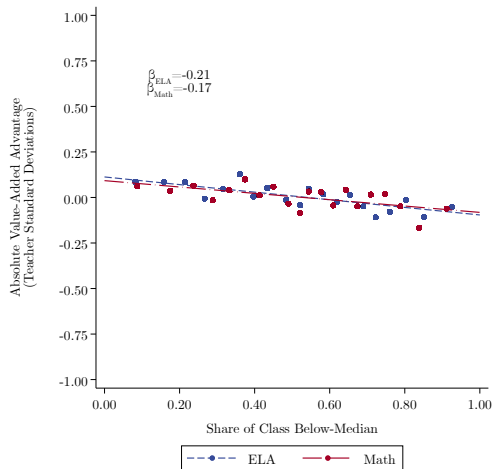
VA Estimation Details

- ▶ **Key Assumption:** classroom-by-achievement-type shocks and idiosyncratic shocks are conditionally independent (with restrictions about stationarity)

Teacher effects are correlated but dispersed



Teacher value added only varies slightly with class composition



Reallocating Teachers to Classrooms

What allocations maximize student achievement?

- Combining theory and estimates, we can see how much heterogeneity matters

What allocations maximize student achievement?

- Combining theory and estimates, we can see how much heterogeneity matters
- We consider allocating teachers to classes (within or across schools) $\mathcal{J} : C \rightarrow J$

What allocations maximize student achievement?

- Combining theory and estimates, we can see how much heterogeneity matters
- We consider allocating teachers to classes (within or across schools) $\mathcal{J} : C \rightarrow J$

$$\max_{\mathcal{J} \in \mathcal{J}} \tilde{\mathcal{W}}(\mathcal{J}; \omega) = \max_{\mathcal{J} \in \mathcal{J}} \frac{1}{N_{i,t}} \sum_{(i,t)} \omega_L L_{i,t} \hat{\tau}_L^{\mathcal{J}(i,t)} + (1 - \omega_L) (1 - L_{i,t}) \hat{\tau}_H^{\mathcal{J}(i,t)}$$

What allocations maximize student achievement?

- Combining theory and estimates, we can see how much heterogeneity matters
- We consider allocating teachers to classes (within or across schools) $\mathcal{J} : C \rightarrow J$

$$\max_{\mathcal{J} \in \mathcal{J}} \tilde{\mathcal{W}}(\mathcal{J}; \omega) = \max_{\mathcal{J} \in \mathcal{J}} \frac{1}{N_{i,t}} \sum_{(i,t)} \omega_L L_{i,t} \hat{\tau}_L^{\mathcal{J}(i,t)} + (1 - \omega_L) (1 - L_{i,t}) \hat{\tau}_H^{\mathcal{J}(i,t)}$$

- ▶ Because teachers differ you cannot just assign by **absolute advantage**

What allocations maximize student achievement?

- Combining theory and estimates, we can see how much heterogeneity matters
- We consider allocating teachers to classes (within or across schools) $\mathcal{J} : C \rightarrow J$

$$\max_{\mathcal{J} \in \mathcal{J}} \tilde{\mathcal{W}}(\mathcal{J}; \omega) = \max_{\mathcal{J} \in \mathcal{J}} \frac{1}{N_{i,t}} \sum_{(i,t)} \omega_L L_{i,t} \hat{\tau}_L^{\mathcal{J}(i,t)} + (1 - \omega_L) (1 - L_{i,t}) \hat{\tau}_H^{\mathcal{J}(i,t)}$$

- ▶ Because teachers differ you cannot just assign by **absolute advantage**
- ▶ Because class sizes differ you cannot just assign by **comparative advantage**

What allocations maximize student achievement?

- Combining theory and estimates, we can see how much heterogeneity matters
- We consider allocating teachers to classes (within or across schools) $\mathcal{J} : C \rightarrow J$

$$\max_{\mathcal{J} \in \mathcal{J}} \tilde{\mathcal{W}}(\mathcal{J}; \omega) = \max_{\mathcal{J} \in \mathcal{J}} \frac{1}{N_{i,t}} \sum_{(i,t)} \omega_L L_{i,t} \hat{\tau}_L^{\mathcal{J}(i,t)} + (1 - \omega_L) (1 - L_{i,t}) \hat{\tau}_H^{\mathcal{J}(i,t)}$$

- ▶ Because teachers differ you cannot just assign by **absolute advantage**
- ▶ Because class sizes differ you cannot just assign by **comparative advantage**
- ▶ The result is an optimization problem with over $2.5 \cdot 10^{1846}$ permutations

What allocations maximize student achievement?

- Combining theory and estimates, we can see how much heterogeneity matters
- We consider allocating teachers to classes (within or across schools) $\mathcal{J} : C \rightarrow J$

$$\max_{\mathcal{J} \in \mathcal{J}} \tilde{\mathcal{W}}(\mathcal{J}; \omega) = \max_{\mathcal{J} \in \mathcal{J}} \frac{1}{N_{i,t}} \sum_{(i,t)} \omega_L L_{i,t} \hat{\tau}_L^{\mathcal{J}(i,t)} + (1 - \omega_L) (1 - L_{i,t}) \hat{\tau}_H^{\mathcal{J}(i,t)}$$

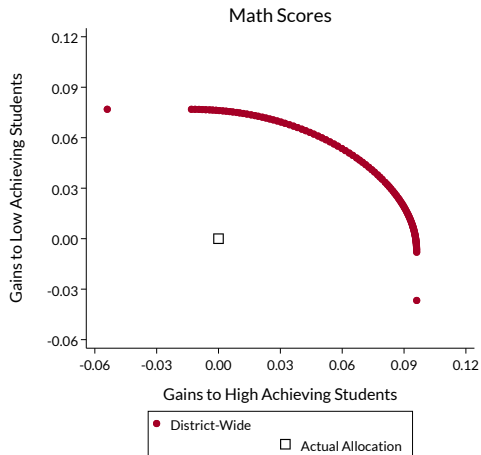
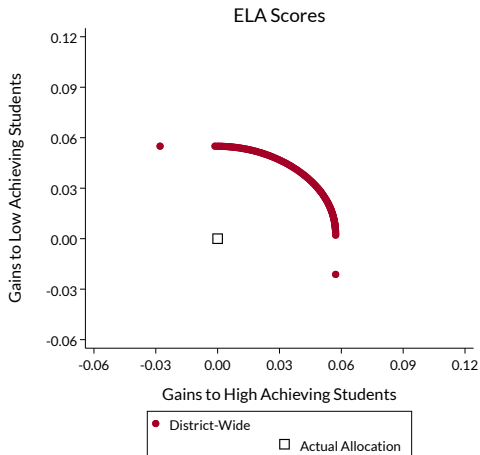
- ▶ Because teachers differ you cannot just assign by **absolute advantage**
 - ▶ Because class sizes differ you cannot just assign by **comparative advantage**
 - ▶ The result is an optimization problem with over $2.5 \cdot 10^{1846}$ permutations
- So we recharacterize the problem as a mixed integer linear programming problem

One year of reallocating teachers to classes creates large gains

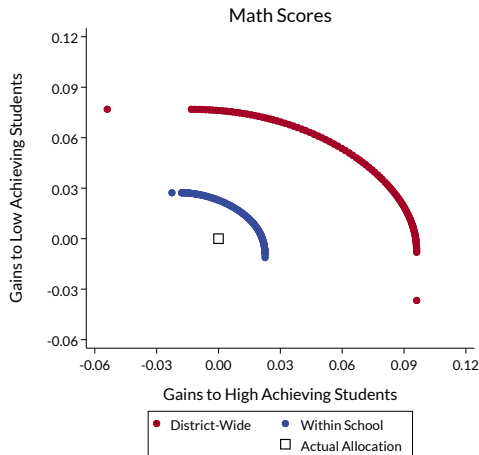
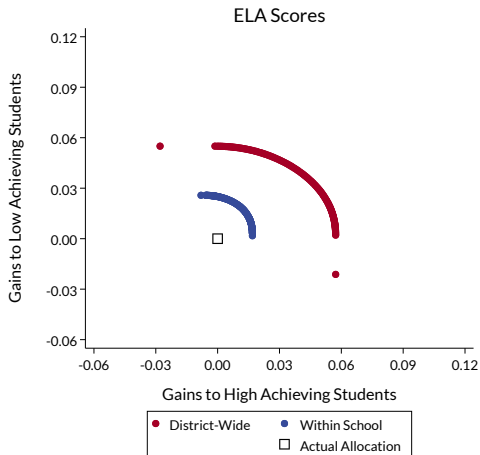
Gains from Heterogeneity

Equity Gains

One year of reallocating teachers to classes creates large gains



One year of reallocating teachers to classes creates large gains



Welfare and Policy

What does it take to get to “welfare added”?

- Can we move from evaluating the “social objective” to true “welfare”?

What does it take to get to “welfare added”?

- Can we move from evaluating the “social objective” to true “welfare”?
- **Families:** Assume families only value long-run labor-market effects from teachers

What does it take to get to “welfare added”?

- Can we move from evaluating the “social objective” to true “welfare”?
- **Families:** Assume families only value long-run labor-market effects from teachers
 - ▶ Use estimates of the **subject-specific** effects of higher value-added teachers on earnings

(Chetty et al., 2014b)

What does it take to get to “welfare added”?

- Can we move from evaluating the “social objective” to true “welfare”?
- **Families:** Assume families only value long-run labor-market effects from teachers
 - ▶ Use estimates of the **subject-specific** effects of higher value-added teachers on earnings
(Chetty et al., 2014b)
- **Teachers:** We calculate the MVPF for dollars spent on added compensation

What does it take to get to “welfare added”?

- Can we move from evaluating the “social objective” to true “welfare”?
- **Families:** Assume families only value long-run labor-market effects from teachers
 - ▶ Use estimates of the **subject-specific** effects of higher value-added teachers on earnings
(Chetty et al., 2014b)
- **Teachers:** We calculate the MVPF for dollars spent on added compensation
 - ▶ Compare to existing calibrations/estimates of compensating variation
(Rothstein, 2015, Bates et al., 2022)

What does it take to get to “welfare added”?

- Can we move from evaluating the “social objective” to true “welfare”?
- **Families:** Assume families only value long-run labor-market effects from teachers
 - ▶ Use estimates of the **subject-specific** effects of higher value-added teachers on earnings
(Chetty et al., 2014b)
- **Teachers:** We calculate the MVPF for dollars spent on added compensation
 - ▶ Compare to existing calibrations/estimates of compensating variation
(Rothstein, 2015, Bates et al., 2022)

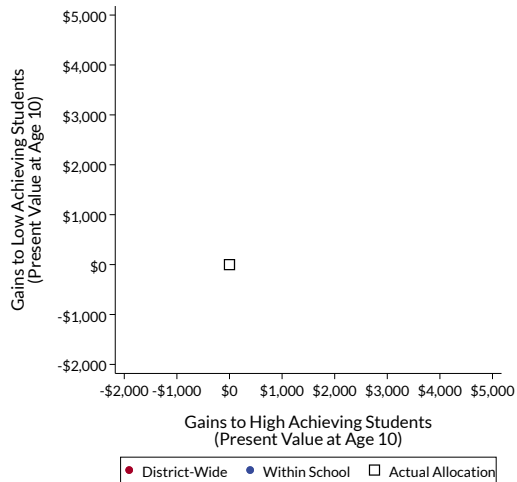
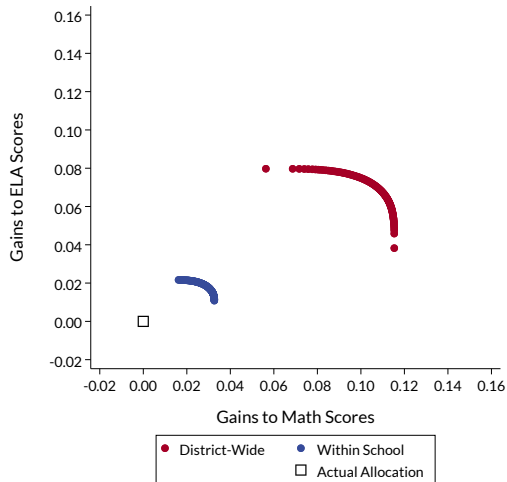
What does it take to get to “welfare added”?

- Can we move from evaluating the “social objective” to true “welfare”?
- **Families:** Assume families only value long-run labor-market effects from teachers
 - ▶ Use estimates of the **subject-specific** effects of higher value-added teachers on earnings
(Chetty et al., 2014b)
- **Teachers:** We calculate the MVPF for dollars spent on added compensation
 - ▶ Compare to existing calibrations/estimates of compensating variation
(Rothstein, 2015, Bates et al., 2022)

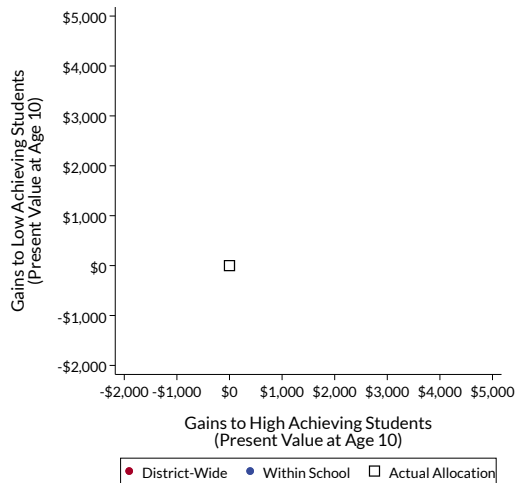
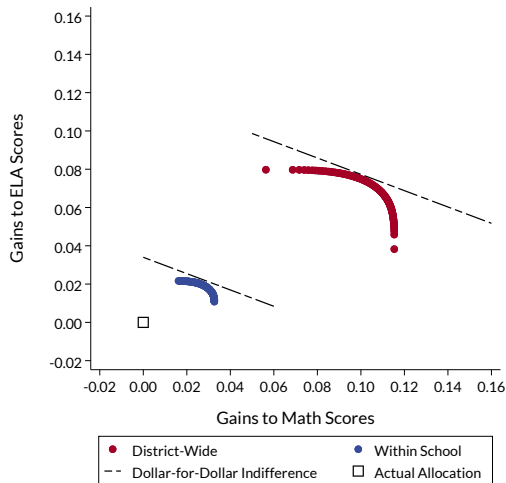
What does it take to get to “welfare added”?

- Can we move from evaluating the “social objective” to true “welfare”?
- **Families:** Assume families only value long-run labor-market effects from teachers
 - ▶ Use estimates of the **subject-specific** effects of higher value-added teachers on earnings
(Chetty et al., 2014b)
- **Teachers:** We calculate the MVPF for dollars spent on added compensation
 - ▶ Compare to existing calibrations/estimates of compensating variation
(Rothstein, 2015, Bates et al., 2022)

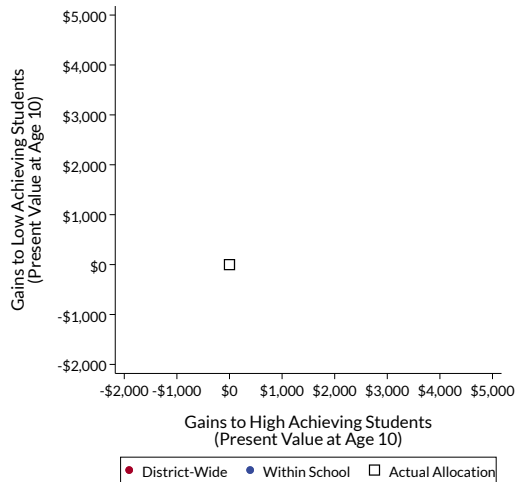
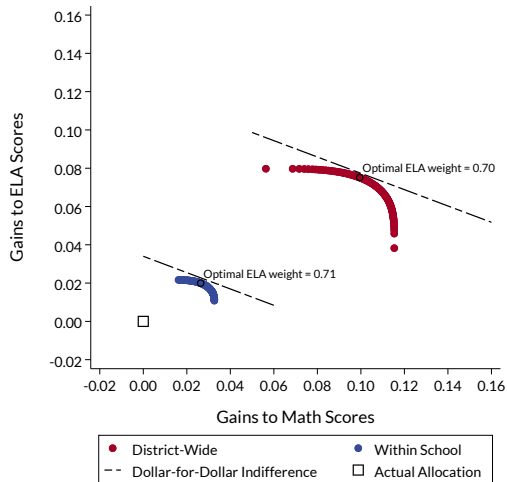
Calculating optimal earnings gains from each year of reallocations



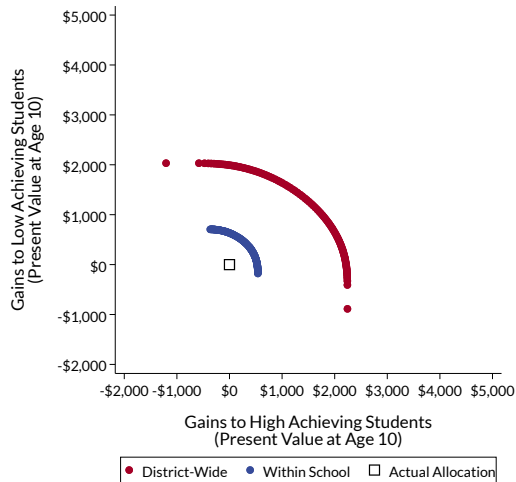
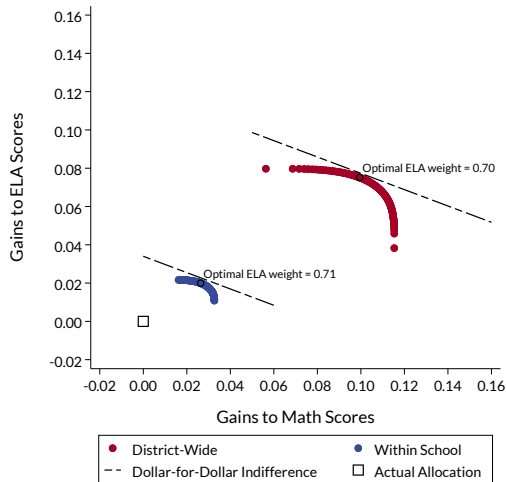
Calculating optimal earnings gains from each year of reallocations



Calculating optimal earnings gains from each year of reallocations



Calculating optimal earnings gains from each year of reallocations



Are gains like this really feasible?

- At a district level this would create \$7.4-27.9M dollars of earnings **every year**

Are gains like this really feasible?

- At a district level this would create \$7.4-27.9M dollars of earnings **every year**
 - ▶ Because ELA matters, gains are larger (\$1.5-7M) than only allocating by math

Are gains like this really feasible?

- At a district level this would create \$7.4-27.9M dollars of earnings **every year**
 - ▶ Because ELA matters, gains are larger (\$1.5-7M) than only allocating by math
- But this policy isn't costless (teachers might hate it), so is it really feasible?

Are gains like this really feasible?

- At a district level this would create \$7.4-27.9M dollars of earnings **every year**
 - ▶ Because ELA matters, gains are larger (\$1.5-7M) than only allocating by math
- But this policy isn't costless (teachers might hate it), so is it really feasible?
 - ▶ **Thought experiment:** Give teachers a bonus for the possibility of reallocation

Are gains like this really feasible?

- At a district level this would create \$7.4-27.9M dollars of earnings **every year**
 - ▶ Because ELA matters, gains are larger (\$1.5-7M) than only allocating by math
- But this policy isn't costless (teachers might hate it), so is it really feasible?
 - ▶ **Thought experiment:** Give teachers a bonus for the possibility of reallocation
 - ▶ How much would it be worth paying them?

Are gains like this really feasible?

- At a district level this would create \$7.4-27.9M dollars of earnings **every year**
 - ▶ Because ELA matters, gains are larger (\$1.5-7M) than only allocating by math
- But this policy isn't costless (teachers might hate it), so is it really feasible?
 - ▶ **Thought experiment:** Give teachers a bonus for the possibility of reallocation
 - ▶ How much would it be worth paying them?

$$MVPF = \frac{\text{Benefits_to_Students}}{\text{Net_cost_to_budget}} = \frac{\sum \Delta \$^j_i \cdot (1 - t)}{N_J \cdot \text{Bonus} - \sum \Delta \$^j_i \cdot t}$$

Are gains like this really feasible?

- At a district level this would create \$7.4-27.9M dollars of earnings **every year**
 - ▶ Because ELA matters, gains are larger (\$1.5-7M) than only allocating by math
- But this policy isn't costless (teachers might hate it), so is it really feasible?
 - ▶ **Thought experiment:** Give teachers a bonus for the possibility of reallocation
 - ▶ How much would it be worth paying them?

$$MVPF = \frac{\text{Benefits_to_Students}}{\text{Net_cost_to_budget}} = \frac{\sum \Delta \$^j_i \cdot (1 - t)}{N_J \cdot \text{Bonus} - \sum \Delta \$^j_i \cdot t}$$

- In this framework, many bonus policies have an infinite MVPF (net cost to budget is zero because discounted tax gains more than pay for costs)

MVPF Results

So what did we learn?

- Good reason to believe heterogeneity (and estimating heterogeneity) matters

So what did we learn?

- Good reason to believe heterogeneity (and estimating heterogeneity) matters
- **Teacher effectiveness varies** with student achievement with long-term implications

So what did we learn?

- Good reason to believe heterogeneity (and estimating heterogeneity) matters
- **Teacher effectiveness varies** with student achievement with long-term implications
- Lots of room for **gains through reallocations**—far beyond standard VA

So what did we learn?

- Good reason to believe heterogeneity (and estimating heterogeneity) matters
- **Teacher effectiveness varies** with student achievement with long-term implications
- Lots of room for **gains through reallocations**—far beyond standard VA
 - ▶ Bigger gains (overall and relative to standard VA), with egalitarian social weights

So what did we learn?

- Good reason to believe heterogeneity (and estimating heterogeneity) matters
- **Teacher effectiveness varies** with student achievement with long-term implications
- Lots of room for **gains through reallocations**—far beyond standard VA
 - ▶ Bigger gains (overall and relative to standard VA), with egalitarian social weights
- Many teacher reallocation programs could pay for themselves in the long run

Thank You!

teastmond@ucsd.edu

njmather@umich.edu

mricks4@unl.edu

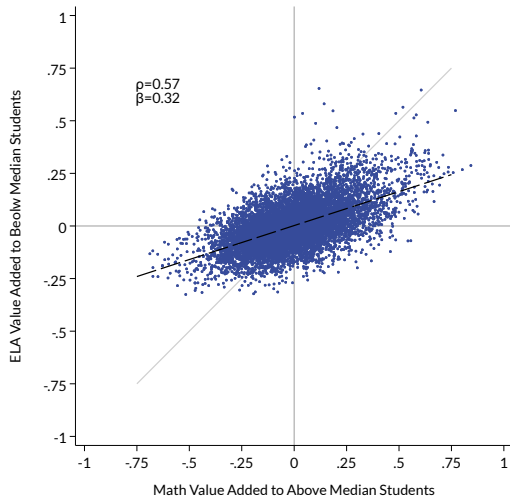
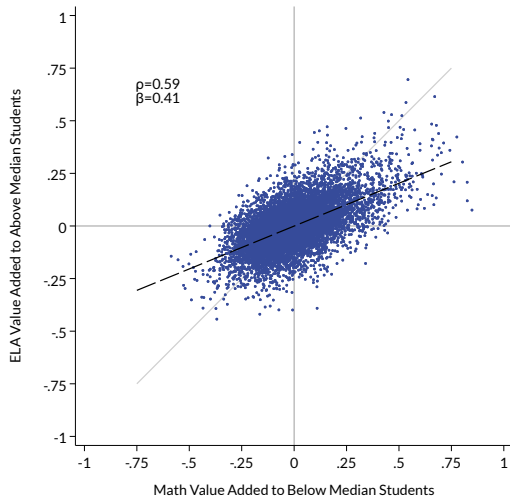
jbetts@ucsd.edu

Appendix

References

- Adler, M. D. (2016): "Benefit-cost analysis and distributional weights: an overview," *Review of Environmental Economics and Policy*.
- Ahn, T., E. Aucejo, J. James, et al. (2021): "The Importance of Matching Effects for Labor Productivity: Evidence from Teacher-Student Interactions," Tech. rep.
- Bates, M. D., M. Dinerstein, A. C. Johnston, and I. Sorkin (2022): "Teacher Labor Market Equilibrium and Student Achievement," Tech. rep., National Bureau of Economic Research.
- Chetty, R., J. N. Friedman, and J. E. Rockoff (2014a): "Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates," *American Economic Review*, 104, 2593–2632.
- (2014b): "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood," *American Economic Review*, 104, 2633–79.
- Condie, S., L. Lefgren, and D. Sims (2014): "Teacher heterogeneity, value-added and education policy," *Economics of Education Review*, 40, 76–92.
- Dahlstrand, A. (2022): "Defying Distance? The Provision of Services in the Digital Age," Tech. rep.
- Delgado, W. (2022): "Heterogeneous Teacher Effects, Comparative Advantage, and Match Quality: Evidence from Chicago Public Schools," .
- Fleurbaey, M. and R. Abi-Rafteh (2016): "The Use of Distributional Weights in Benefit-Cost Analysis: Insights from Welfare Economics¹," *Review of Environmental Economics and Policy*.
- Jackson, C. K. (2018): "What do test scores miss? The importance of teacher effects on non-test score outcomes," *Journal of Political Economy*, 126, 2072–2107.
- Petek, N. and N. G. Pope (forthcoming): "The Multidimensional Impact of Teachers on Students," *Journal of Political Economy*.
- Rothstein, J. (2015): "Teacher quality policy when supply matters," *American Economic Review*, 105, 100–130.

Cross subject correlations are much weaker



Estimation extensions and robustness checks

- Heterogeneous effects leave lots of room for **comparative advantage**
 - ▶ Cross subject correlations are weaker than within-subject correlations
 - ▶ Similar to cross-race correlation and less correlated than cross-SES

Cross-Subject Scatter

(Delgado, 2022, Bates et al., 2022)

Estimation extensions and robustness checks

- Heterogeneous effects leave lots of room for **comparative advantage**

- ▶ Cross subject correlations are weaker than within-subject correlations

Cross-Subject Scatter

- ▶ Similar to cross-race correlation and less correlated than cross-SES

(Delgado, 2022, Bates et al., 2022)

- **Key Concern:** Are the differences noise or actual comparative advantage?

Estimation extensions and robustness checks

- Heterogeneous effects leave lots of room for **comparative advantage**

- ▶ Cross subject correlations are weaker than within-subject correlations

Cross-Subject Scatter

- ▶ Similar to cross-race correlation and less correlated than cross-SES

(Delgado, 2022, Bates et al., 2022)

- **Key Concern:** Are the differences noise or actual comparative advantage?

- Heterogeneity seems to reliably estimate economically meaningful differences

Persistence

Long-Run Effects

Estimation extensions and robustness checks

- Heterogeneous effects leave lots of room for **comparative advantage**

- ▶ Cross subject correlations are weaker than within-subject correlations

Cross-Subject Scatter

- ▶ Similar to cross-race correlation and less correlated than cross-SES

(Delgado, 2022, Bates et al., 2022)

- **Key Concern:** Are the differences noise or actual comparative advantage?

- Heterogeneity seems to reliably estimate economically meaningful differences

Persistence

Long-Run Effects

- ▶ Heterogeneous effects tend to be just as **forecast unbiased**

Estimation extensions and robustness checks

- Heterogeneous effects leave lots of room for **comparative advantage**

- ▶ Cross subject correlations are weaker than within-subject correlations

Cross-Subject Scatter

- ▶ Similar to cross-race correlation and less correlated than cross-SES

(Delgado, 2022, Bates et al., 2022)

- **Key Concern:** Are the differences noise or actual comparative advantage?

- Heterogeneity seems to reliably estimate economically meaningful differences

Persistence

Long-Run Effects

- ▶ Heterogeneous effects tend to be just as **forecast unbiased**
- ▶ We find that **comparative advantage** is extremely persistent over time

Estimation extensions and robustness checks

- Heterogeneous effects leave lots of room for **comparative advantage**

- ▶ Cross subject correlations are weaker than within-subject correlations

Cross-Subject Scatter

- ▶ Similar to cross-race correlation and less correlated than cross-SES

(Delgado, 2022, Bates et al., 2022)

- **Key Concern:** Are the differences noise or actual comparative advantage?

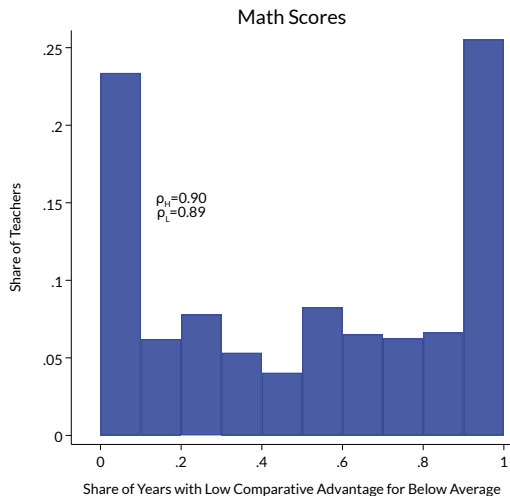
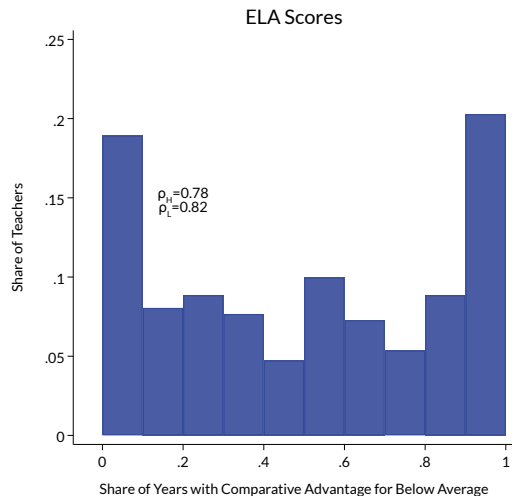
- Heterogeneity seems to reliably estimate economically meaningful differences

Persistence

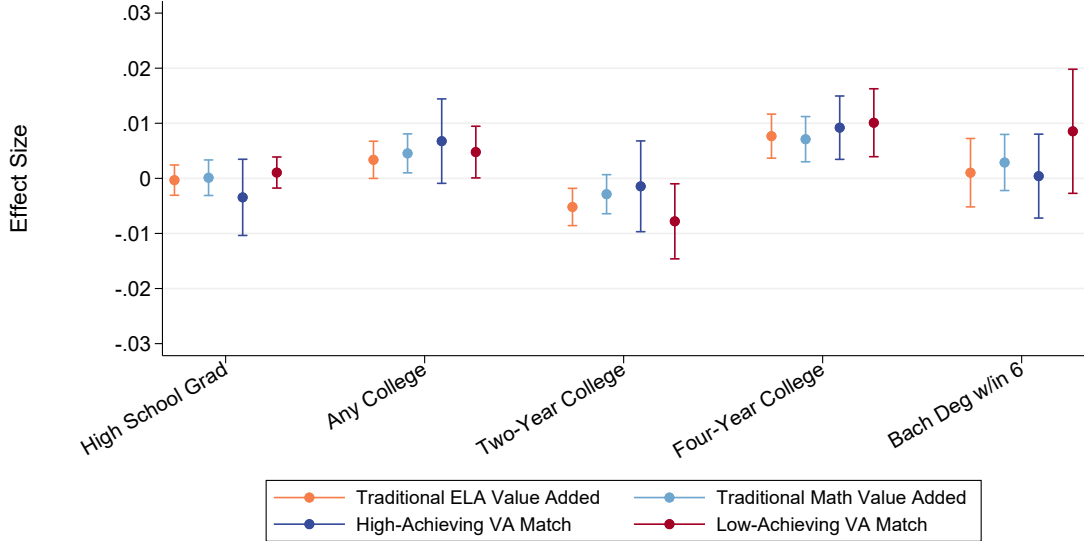
Long-Run Effects

- ▶ Heterogeneous effects tend to be just as **forecast unbiased**
- ▶ We find that **comparative advantage** is extremely persistent over time
- ▶ Heterogeneous value added captures all the information from **long-run outcomes**

Comparative Advantage is persistent over time



Heterogeneous value added predicts long-term outcomes



Quantifying the gains from comparative advantage and equity

- We compare the gains of the optimal allocation with the gains from using standard VA

VA versus Heterogeneity

Distributional Gains

Egalitarian Socres

Back

Quantifying the gains from comparative advantage and equity

- We compare the gains of the optimal allocation with the gains from using standard VA

VA versus Heterogeneity

Distributional Gains

Egalitarian Socres

Back

Quantifying the gains from comparative advantage and equity

- We compare the gains of the optimal allocation with the gains from using standard VA

VA versus Heterogeneity

Distributional Gains

Egalitarian Scores

$$\max_{\mathcal{J} \in \mathcal{J}} \tilde{\mathcal{W}}(\mathcal{J}; \omega) = \max_{\mathcal{J} \in \mathcal{J}} \frac{1}{N_{i,t}} \sum_{(i,t)} \omega_L L_{i,t} \hat{\tau}_H^{j(c)} + (1 - \omega_L) (1 - L_{i,t}) \hat{\tau}_L^{j(c)}$$

- ▶ Heterogeneous VA raises average scores by 66-97% more than standard VA
- ▶ While always valuable, heterogeneity is particularly useful when both groups matter

Quantifying the gains from comparative advantage and equity

- We compare the gains of the optimal allocation with the gains from using standard VA

VA versus Heterogeneity

Distributional Gains

Egalitarian Scores

$$\max_{\mathcal{J} \in \mathcal{J}} \tilde{\mathcal{W}}(\mathcal{J}; \omega) = \max_{\mathcal{J} \in \mathcal{J}} \frac{1}{N_{i,t}} \sum_{(i,t)} \omega_L L_{i,t} \hat{\tau}_{VA}^{j(c)} + (1 - \omega_L) (1 - L_{i,t}) \hat{\tau}_{VA}^{j(c)}$$

- ▶ Heterogeneous VA raises average scores by 66-97% more than standard VA
- ▶ While always valuable, heterogeneity is particularly useful when both groups matter
- ▶ With egalitarian social objectives, heterogeneity is key to maintaining average scores

Quantifying the gains from comparative advantage and equity

- We compare the gains of the optimal allocation with the gains from using standard VA

VA versus Heterogeneity

Distributional Gains

Egalitarian Scores

$$\max_{\mathcal{J} \in \mathcal{J}} \tilde{\mathcal{W}}(\mathcal{J}; \omega) = \max_{\mathcal{J} \in \mathcal{J}} \frac{1}{N_{i,t}} \sum_{(i,t)} \omega_L L_{i,t} \hat{\tau}_{VA}^{j(c)} + (1 - \omega_L) (1 - L_{i,t}) \hat{\tau}_{VA}^{j(c)}$$

- ▶ Heterogeneous VA raises average scores by 66-97% more than standard VA
- ▶ While always valuable, heterogeneity is particularly useful when both groups matter
- ▶ With egalitarian social objectives, heterogeneity is key to maintaining average scores

- Our teacher reallocations can shrink achievement and racial gaps

Achievement Gaps

Racial Gaps

Winners & Losers

Quantifying the gains from comparative advantage and equity

- We compare the gains of the optimal allocation with the gains from using standard VA

VA versus Heterogeneity

Distributional Gains

Egalitarian Scores

$$\max_{\mathcal{J} \in \mathcal{J}} \tilde{\mathcal{W}}(\mathcal{J}; \omega) = \max_{\mathcal{J} \in \mathcal{J}} \frac{1}{N_{i,t}} \sum_{(i,t)} \omega_L L_{i,t} \hat{\tau}_{VA}^{j(c)} + (1 - \omega_L) (1 - L_{i,t}) \hat{\tau}_{VA}^{j(c)}$$

- ▶ Heterogeneous VA raises average scores by 66-97% more than standard VA
- ▶ While always valuable, heterogeneity is particularly useful when both groups matter
- ▶ With egalitarian social objectives, heterogeneity is key to maintaining average scores

- Our teacher reallocations can shrink achievement and racial gaps

Achievement Gaps

Racial Gaps

Winners & Losers

- ▶ Requires egalitarian weights (54-72%) on low achievers in district reallocations

Quantifying the gains from comparative advantage and equity

- We compare the gains of the optimal allocation with the gains from using standard VA

VA versus Heterogeneity

Distributional Gains

Egalitarian Scores

$$\max_{\mathcal{J} \in \mathcal{J}} \tilde{\mathcal{W}}(\mathcal{J}; \omega) = \max_{\mathcal{J} \in \mathcal{J}} \frac{1}{N_{i,t}} \sum_{(i,t)} \omega_L L_{i,t} \hat{\tau}_{VA}^{j(c)} + (1 - \omega_L) (1 - L_{i,t}) \hat{\tau}_{VA}^{j(c)}$$

- ▶ Heterogeneous VA raises average scores by 66-97% more than standard VA
- ▶ While always valuable, heterogeneity is particularly useful when both groups matter
- ▶ With egalitarian social objectives, heterogeneity is key to maintaining average scores

- Our teacher reallocations can shrink achievement and racial gaps

Achievement Gaps

Racial Gaps

Winners & Losers

- ▶ Requires egalitarian weights (54-72%) on low achievers in district reallocations
- ▶ Within-school reallocations don't meaningfully change racial gaps

Quantifying the gains from comparative advantage and equity

- We compare the gains of the optimal allocation with the gains from using standard VA

VA versus Heterogeneity

Distributional Gains

Egalitarian Scores

$$\max_{\mathcal{J} \in \mathcal{J}} \tilde{W}(\mathcal{J}; \omega) = \max_{\mathcal{J} \in \mathcal{J}} \frac{1}{N_{i,t}} \sum_{(i,t)} \omega_L L_{i,t} \hat{\tau}_{VA}^{j(c)} + (1 - \omega_L) (1 - L_{i,t}) \hat{\tau}_{VA}^{j(c)}$$

- ▶ Heterogeneous VA raises average scores by 66-97% more than standard VA
- ▶ While always valuable, heterogeneity is particularly useful when both groups matter
- ▶ With egalitarian social objectives, heterogeneity is key to maintaining average scores

- Our teacher reallocations can shrink achievement and racial gaps

Achievement Gaps

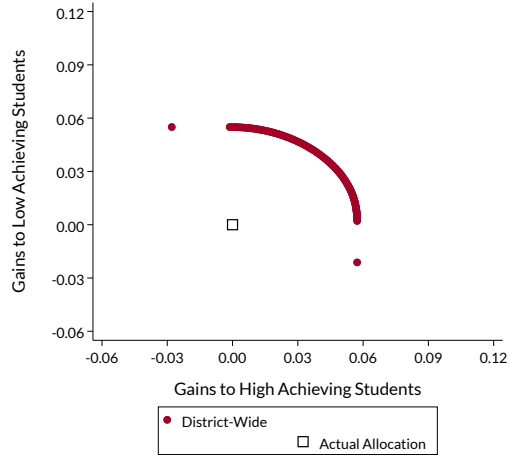
Racial Gaps

Winners & Losers

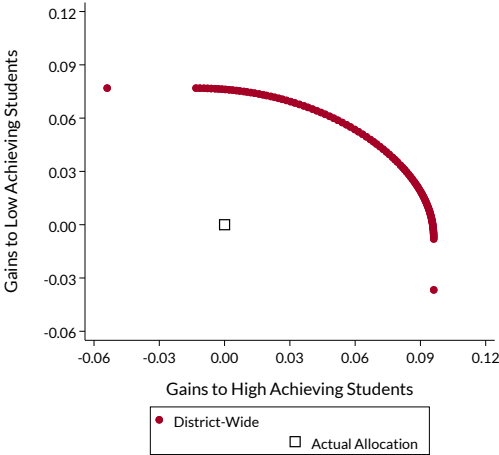
- ▶ Requires egalitarian weights (54-72%) on low achievers in district reallocations
- ▶ Within-school reallocations don't meaningfully change racial gaps
- ▶ Note these are average not "Pareto" gains: there are winners and losers

Reallocation exercises

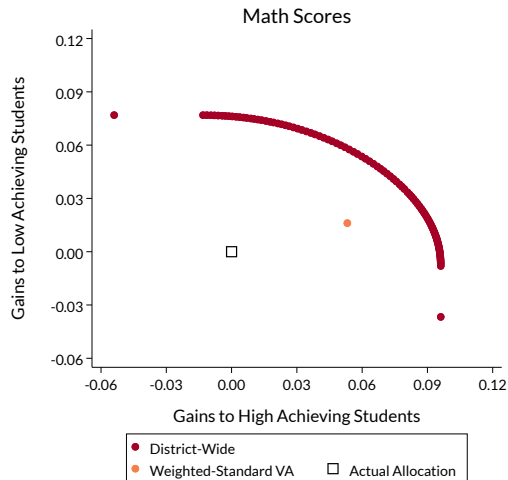
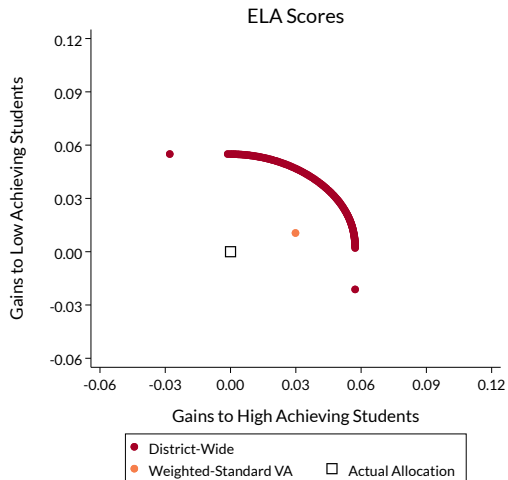
ELA Scores



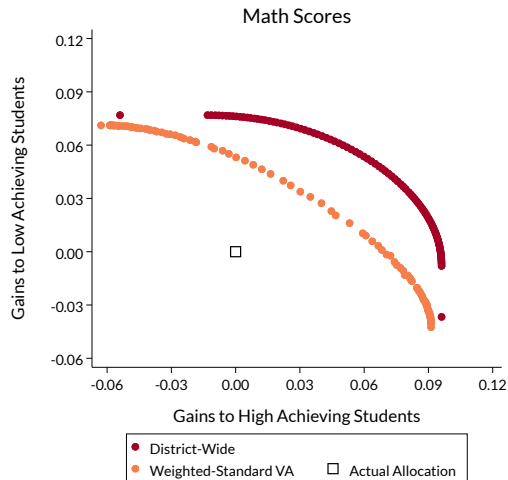
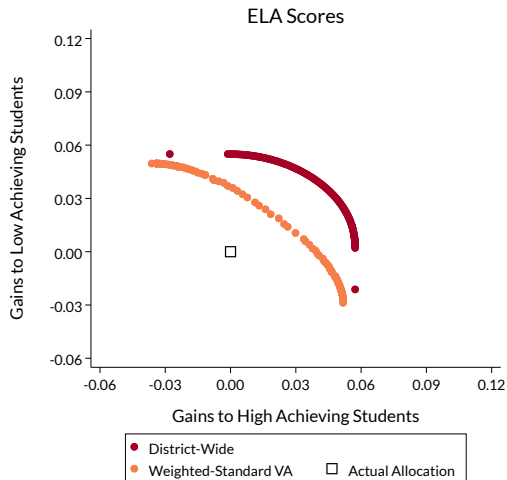
Math Scores



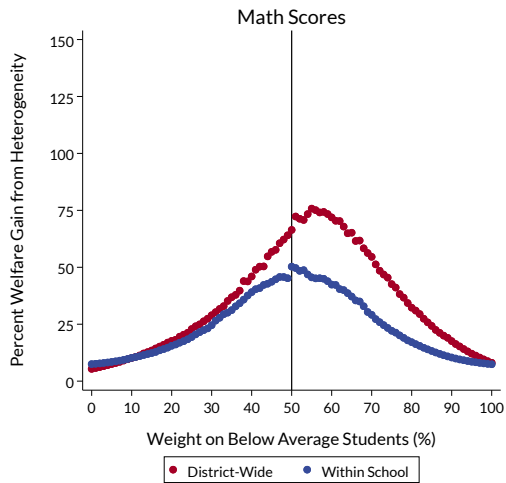
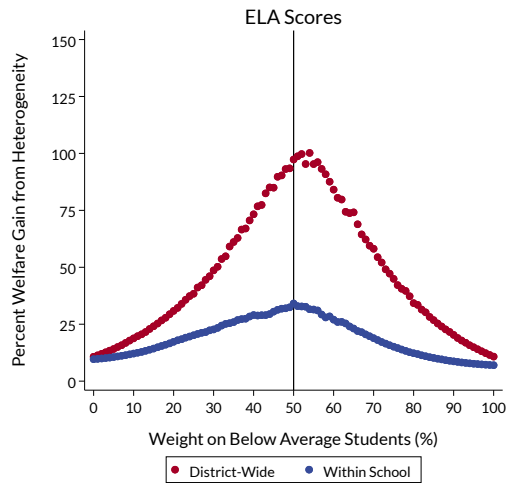
Reallocation exercises



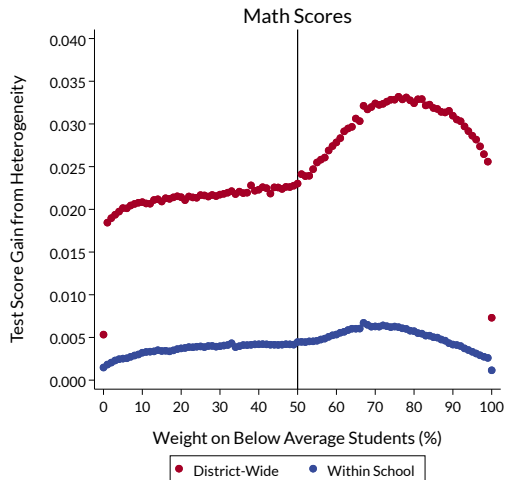
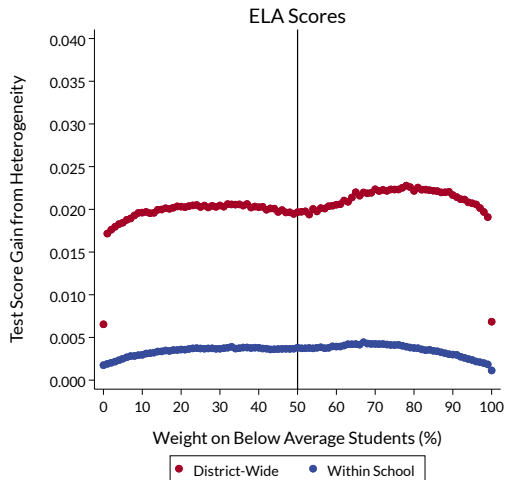
Reallocation exercises



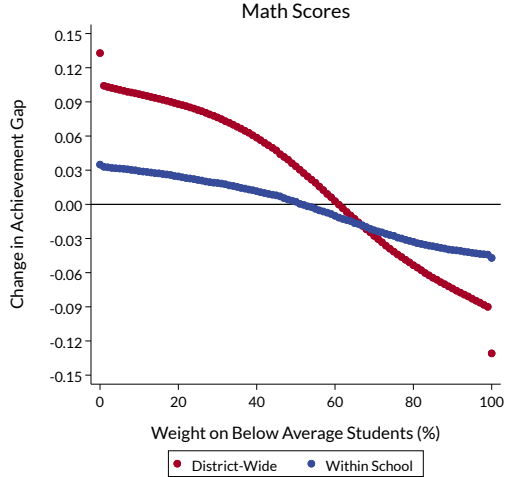
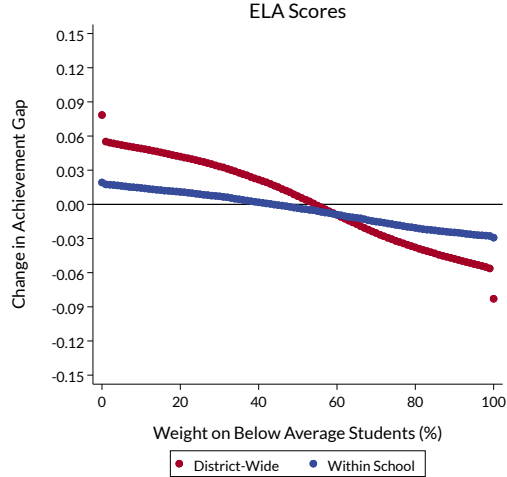
Gains are the largest when the social objective is egalitarian



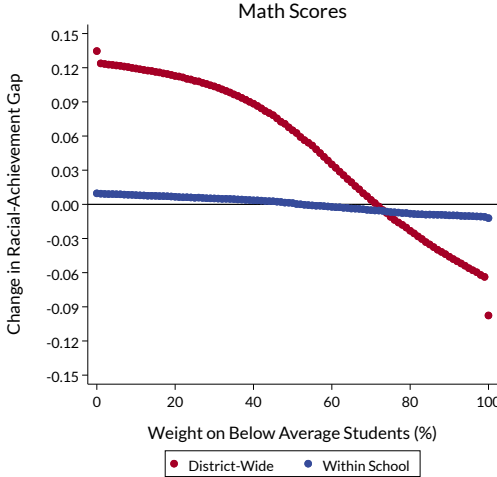
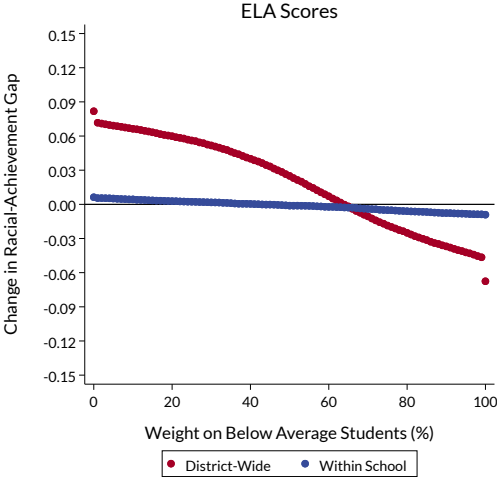
With distributional preferences, heterogeneity raises scores



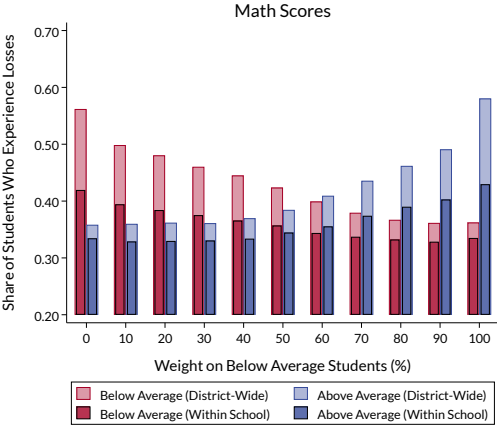
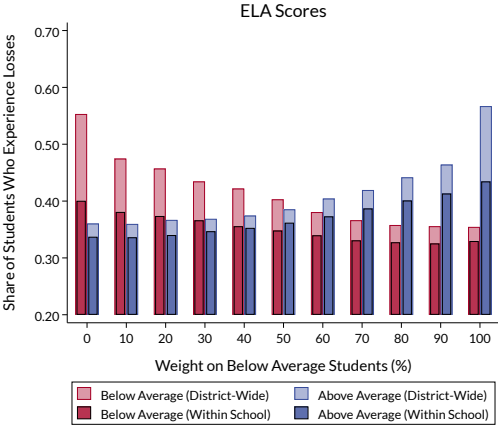
Reallocations can strongly shape achievement gaps



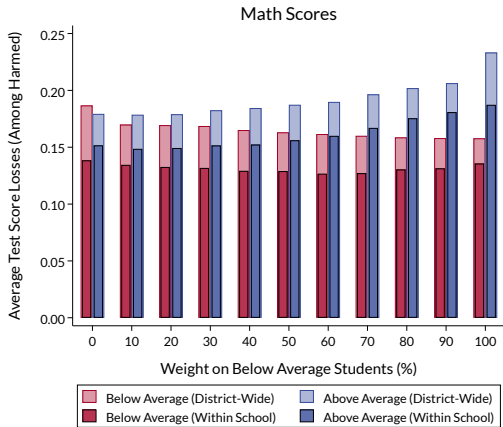
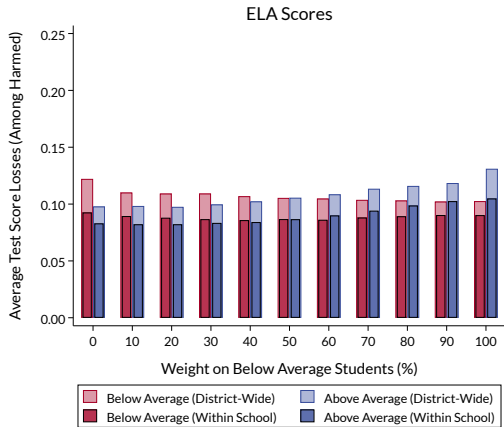
Only moving teachers across schools can change racial gaps



The share of each group harmed depends on the welfare weight



While shares change the size of losses is fairly constant



Reallocations have high MVPF even for large bonus programs

