Empirical Methods for Policy Evaluation II

Matteo Bobba

Toulouse School of Economics (TSE)

TSE PhD Program
Fall 2021

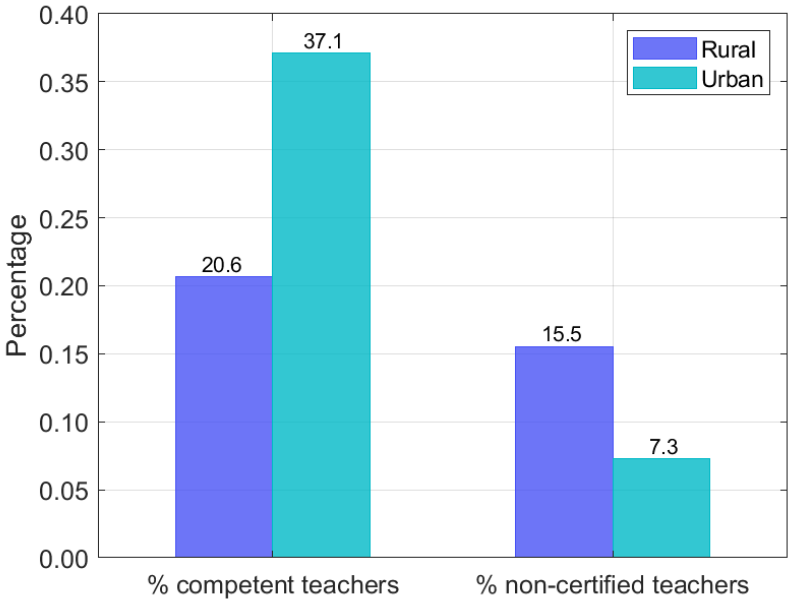
Part 3: Market Design Models & RDDs

- Context
  - **Implemented policy and evaluation**
- Methods
  - Agarwal and Somaini (ARE, 2020)
- Application
  - Bobba et al. (2021)

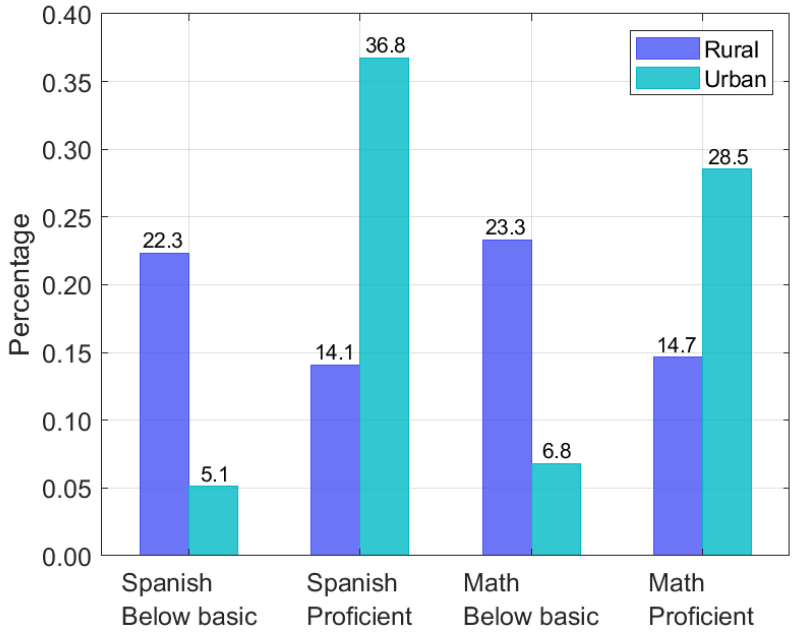
Wage Policies in the Public Sector

- Centralized labor clearinghouses are increasingly used for entry-level positions in professional marketplaces
  - Both sides of the market express preferences over each other while a matching algorithm characterizes their allocation
- Ideal setting to study the provision of services in the public sector
  - Collective wage bargaining often leads workers to sort on non-pecuniary aspects of employment
- How teacher wage policies can alleviate inequality of opportunities for students?
  - Wide learning gaps between urban and rural areas
  - Teachers matter for both education and life-cycle outcomes

Peruvian Public Primary Schools



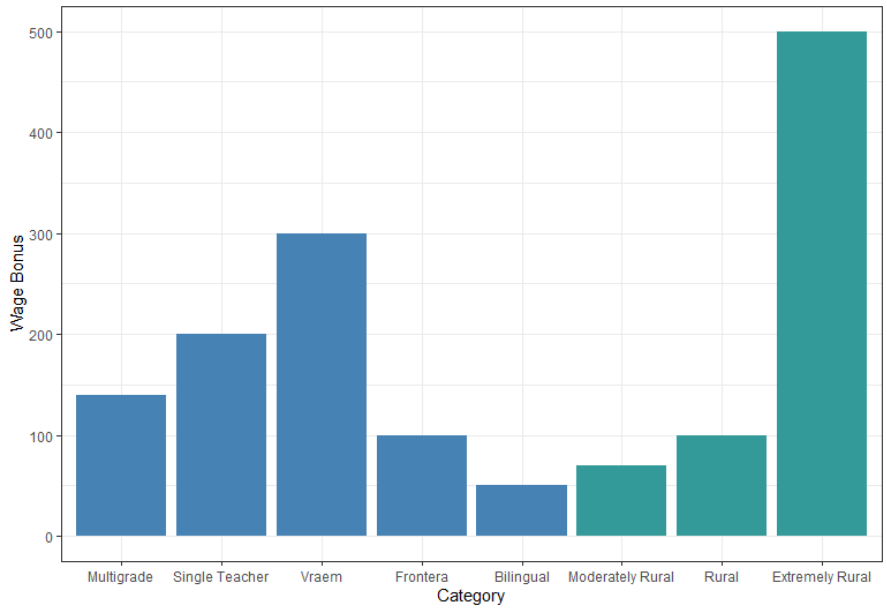
a) Teachers



b) Students

Contracts and Wages of Public Sector Teachers

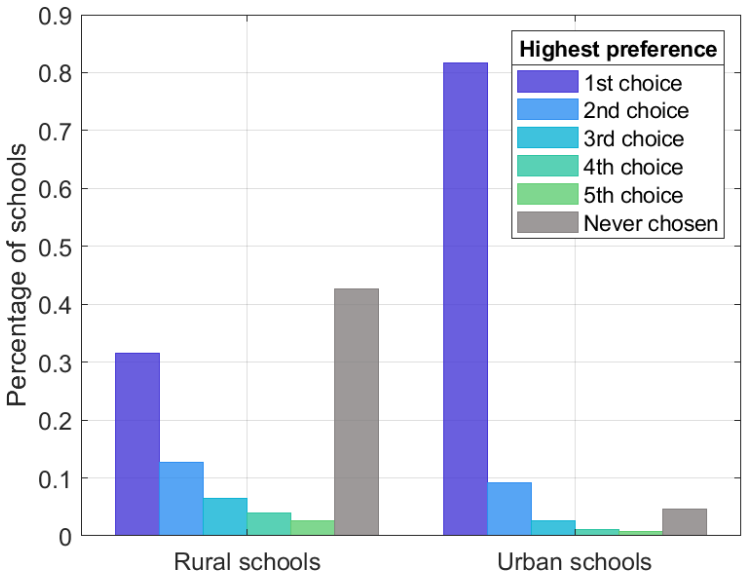
- Teachers are hired under two types of contracts
  - Permanent: civil servants with indefinite contracts
  - Contract: short-term positions, more widespread in rural areas
- Teacher wages are on average lower than college graduate wages and vary by
  - Type of contract
  - Seniority (only if permanent)
  - Location and school characteristics



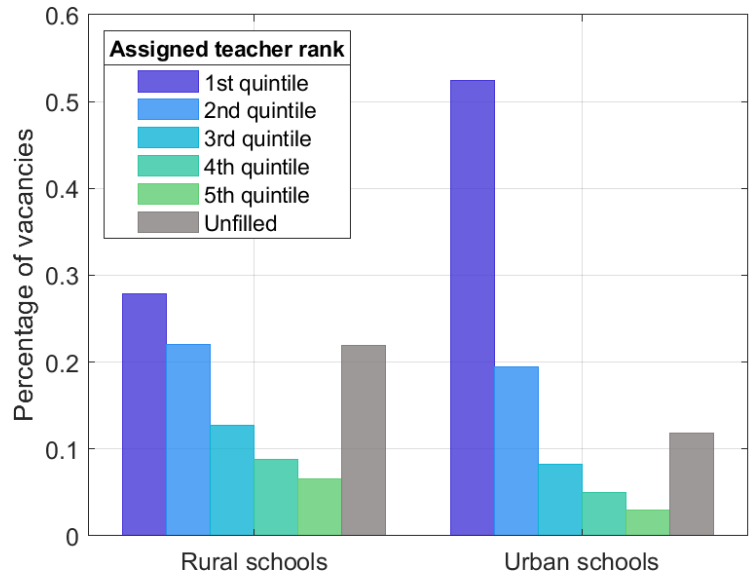
National Recruitment Drives

- Recently established centralized assignment mechanism
  1. All certified applicants take a national competency test
  2. Those eligible for permanent position (score>60%) rank up to 5 jobs in a school-district
  3. Schools interview up to 20 apps (by score) and select
  4. Short-term teaching vacancies are posted together with unfilled vacancies
  5. Teachers ranked by score in each district and assigned by serial-dictatorship
  6. Unfilled vacancies resolved in a decentralized way to un-certified personnel

Teacher Sorting and Assignment Outcomes



a) Permanent Teachers



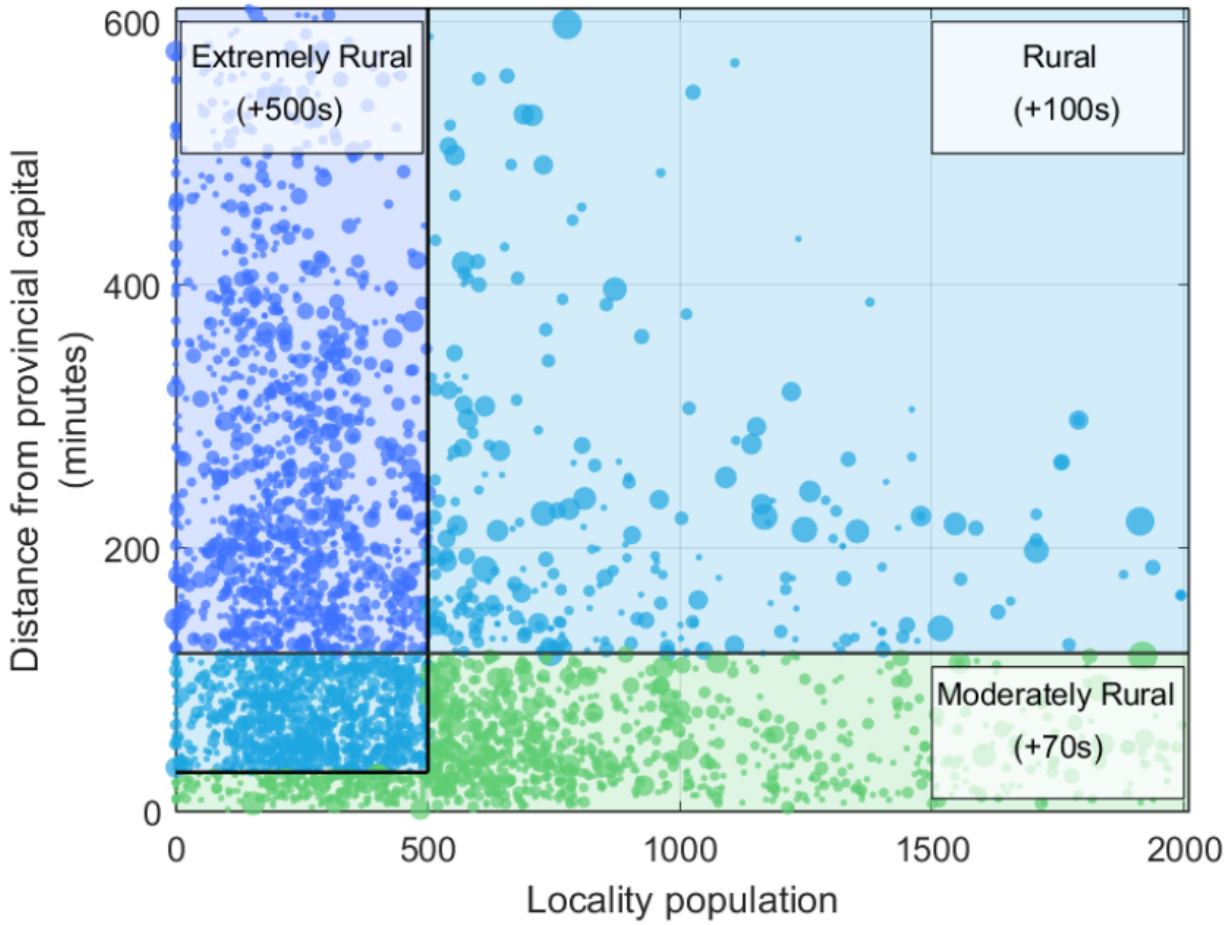
b) Contract Teachers

⇒ More than two thirds of permanent vacancies in rural schools remain unfilled

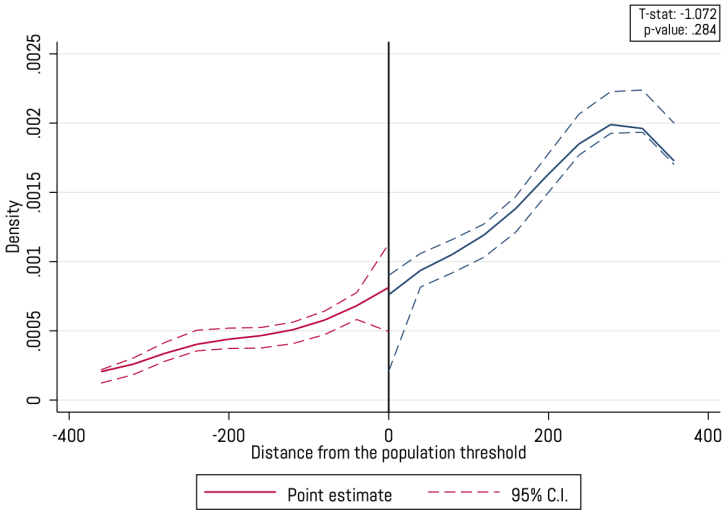
⇒ Rural vacancies are filled by teachers with 0.5 SD lower scores on average

Policy Changes to Compensation in Rural Locations

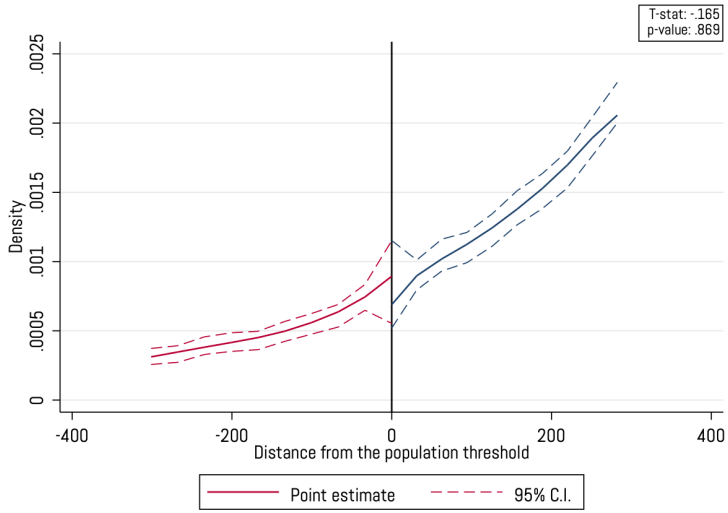
⇒ New wage bonus for teachers at ∼25,000 rural schools



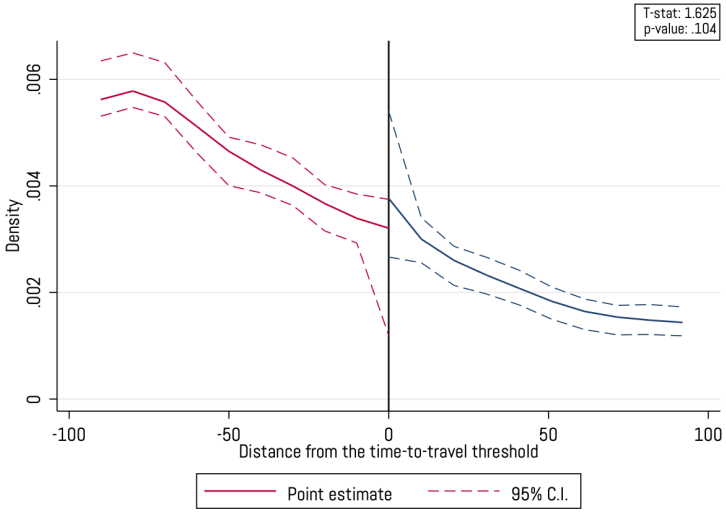
Density Tests Around Extremely Rural Cutoff



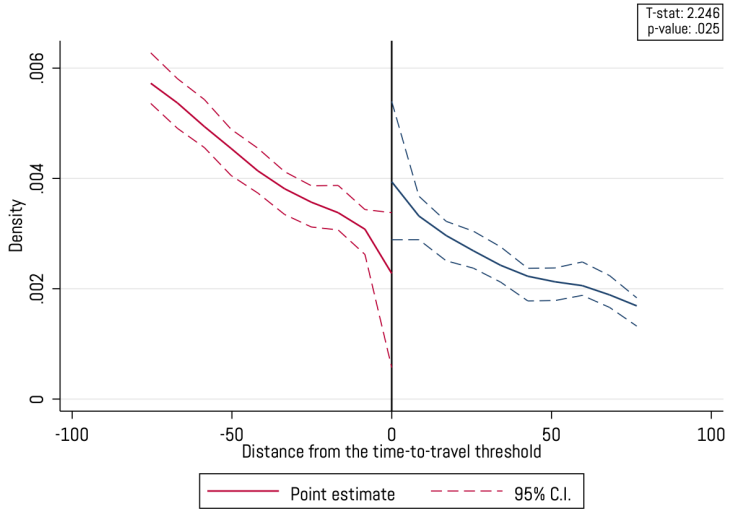
a) Population (2015)



b) Population (2017)



c) Time-to-travel (2015)



d) Time-to-travel (2017)

Sharp RD Along Population Cutoff

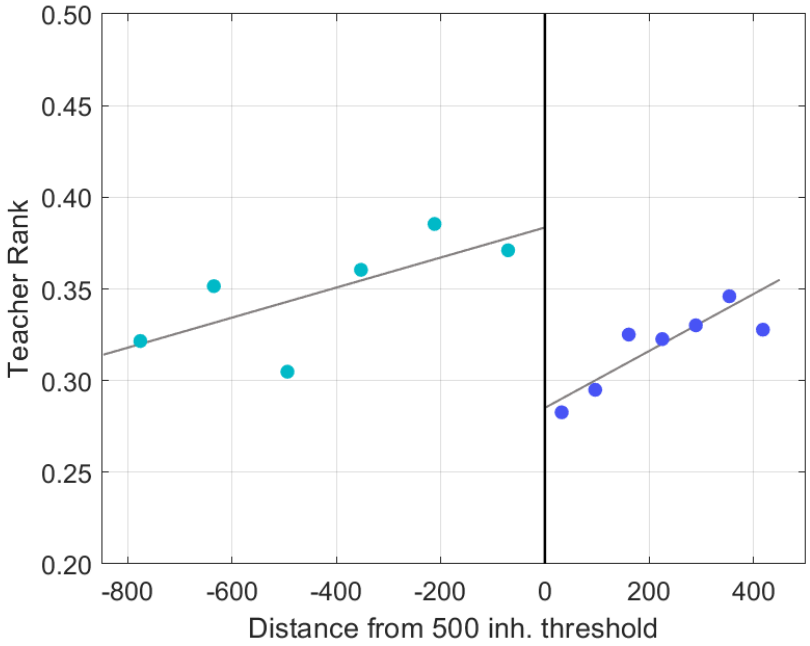
- We solely rely on population-based assignment rule for rural schools with time-to-travel>30min
  $\Rightarrow$ Weighted average increase in wages of 15%
- Given continuity of potential outcomes around the population cutoff

$$y_{jt} = \gamma_0 + \gamma_1 \mathbf{1}(pop_{jt} < pop_c) + g(pop_c - pop_{jt}) + \delta_t + u_{jt}$$

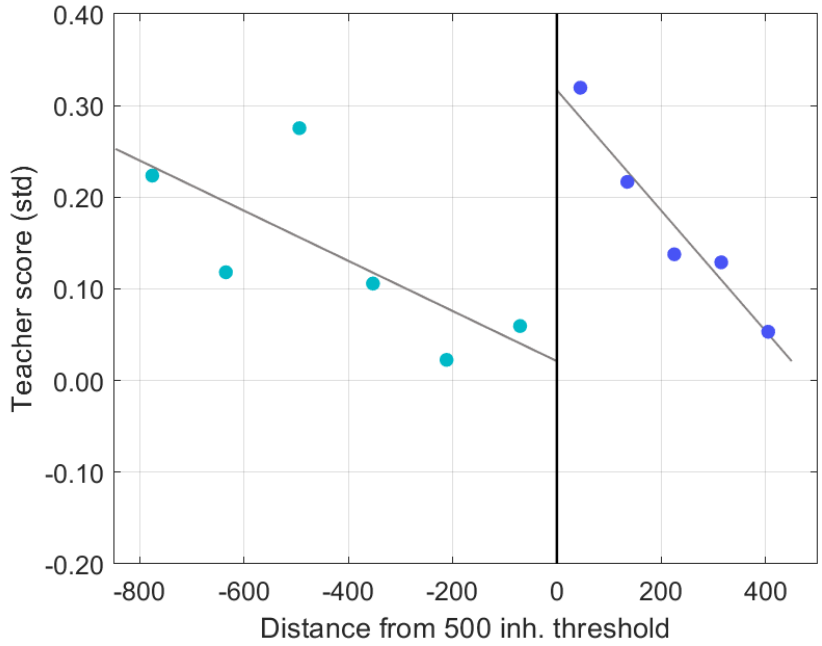
  - $g(\cdot)$: flexible polynomial on population
  - δ_t: indicator for year of assignment
  - u_{jt}: error term, clustered at the school-year level

$\Rightarrow$ Estimate γ_1 non-parametrically within MSE-optimal bandwidths

Wage Increment and Teacher Labor Supply/Quality



a) Preferences



b) Teacher score

NOTES. Each marker in the Figure indicates the average of the outcome variable within each bin, defined following the IMSE-optimal evenly spaced method by Calonico, Cattaneo and Titiunik (2015). Solid lines represent the predictions from linear regressions estimated separately for observations to the left and to the right of the cutoff.

Wage Increment and Student Achievement (Spanish Test Scores)

	(1) No vacancy	(2) Any vacancy	(3) Permanent vacancy	(4) Short-term vacancy
Above cutoff	0.014	0.298**	-0.057	0.317**
	(0.157)	(0.127)	(0.190)	(0.137)
Mean dep. var. (LHS)	-0.471	-0.470	-0.382	-0.491
BW	124.095	108.453	175.257	114.883
Schools	372	691	292	622
Observations	3948	9700	3409	8966

NOTES. This table reports the effect of crossing the population threshold on the student-level average in the 4th Spanish test score taken in 2018 for different sub-samples of primary schools. Each cell reports the bias-corrected regression-discontinuity estimates obtained using the robust estimator proposed by Calonico, Cattaneo, and Titiunik (2014). Regressions are defined within a mean-square error optimal bandwidth (BW), which are reported at the bottom part of the table. The table also reports the mean of the dependent variable computed within the interval $(-BW, 0]$ (left-hand-side of the cutoff). SE are clustered at the school $\times$ year level. *** p$<$ 0.01, ** p$<$0.05, and *p$<$0.10.

Takeaways from RD Analysis

1. The increase in compensation shifted labor supply of teacher across schools
   $\Rightarrow$ Vacancies become more desirable: filled by teachers with higher scores, but no effects at extensive margin

2. The increase in supply toward disadvantaged locations raised pool of measurable teacher quality
   $\Rightarrow$ But only when assignment ignored schools discretionary step in hiring (Duflo et al., 2015)

3. Observed increase in productivity is highly correlated with the inflow of new teachers across schools
   $\Rightarrow$ Little scope of efficiency wages (de Ree et al., 2018)

Part 3: Market Design Models & RDDs

- Context
  - Implemented policy and evaluation
- Methods
  - **Agarwal and Somaini (ARE, 2020)**
- Application
  - Bobba et al. (2021)

Revealed Preference Analysis of School Choice Models

- Unlike traditional consumer settings, education markets typically do not use prices as clearing mechanism
  - Empirical approaches to estimate student preferences crucially depend on the systems for rationing or admitting students to schools/colleges
- Revealed preference approach: use data from decisions and outcomes realized within centralized assignment to estimate the distribution of student preferences over schools
- Leverage estimated preferences to generate counterfactual matching equilibria
  1. Changes in the design/features of the assignment mechanism in place
  2. Changes in school characteristics

School Choice Mechanisms

- R_{ik} is student i's submitted ranked-order list (RoL), where k is the position of the ranked school
- t_{ij} is student i's priority (exam score) in school j
- q_j is a vector of schools' capacities
- A mechanism ϕ maps the tuple (R, t) into school assignments μ

$$\mu_{ij} = \begin{cases} 1 \text{ if student } i \text{ is assigned to school } j \\ 0 \text{ otherwise} \end{cases}$$

$$\sum_i \mu_{ij} \leq q_j, \ \forall j$$

Deferred Acceptance (Gale and Shapley 1962)

- Step 0: if priorities are coarse, a tie-breaker ν_i is generated
- Step 1: each student applies to their highest ranked school R_{i1}. Schools tentatively hold the students with the highest priority up to q_j while the remaining students are rejected
- Step $k > 1$: previously rejected students apply to their highest ranked school that has not yet rejected them. At each school j, the previously held students are considered along with the new applicant in round k. Schools tentatively hold the highest priority students while the remaining students are rejected
- The algorithm terminates whenever no students are rejected, or if all school slots q_j are filled, or if all rejected students have applied to each of their ranked schools

DA Properties

- DA generates an allocation determined by a vector of cutoff priorities p_j and test scores s_i such that each student is assigned to the highest ranked school in her choice set:

$$\Omega(s_i, p) = \{j : s_i > p_j\}$$

- DA is strategy-proof: submitting RoLs that coincide with ordinal preferences is a weakly dominant strategy
- DA is stable: there is no pair (i, j) such that: (i) i strictly prefers j over μ_{ik} and (ii) if school does not have empty slots, then $s_i > p_j$

Part 3: Market Design Models & RDDs

- Context
  - Implemented policy and evaluation
- Methods
  - Agarwal and Somaini (ARE, 2020)
- Application
  - **Bobba et al. (2021)**

An Empirical Model of Teachers School Choice

$$v_{ij} = \alpha_i w_j + \boldsymbol{\beta}_i' \boldsymbol{z}_j + \boldsymbol{\delta}' \boldsymbol{d}_{ij} + \boldsymbol{\lambda}' \boldsymbol{m}_{ij} + \epsilon_{ij}$$
$$v_{i0} = \boldsymbol{\eta}_0 + \boldsymbol{\eta}_1' \boldsymbol{q}_i + \epsilon_{i0},$$

- Heterogeneous preferences for loc./school characteristics: $\boldsymbol{\beta}_i = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1' \boldsymbol{x}_i$
- Heterogeneous preferences over compensation: $\alpha_i = \alpha_0 + \boldsymbol{\alpha}_1' \boldsymbol{x}_i + \sigma \nu_i$

Use stability of the matching equilibrium to infer revealed preferences in short-term contract jobs

$\Rightarrow$ Discrete choice problem with individualized choice sets $\Omega(s_i)$

$$L(\boldsymbol{\theta}) = \frac{1}{n} \sum_{i=1}^{n} \log \left\{ \int_0^{\infty} \left(\frac{\exp \tilde{v}_{i\mu(i)}}{\sum_{k \in \Omega(s_i) \cup \{0\}} \exp \tilde{v}_{ik}} \right) dF(\nu_i) \right\}$$

Identification

Preference parameters $\boldsymbol{\theta}$ are identified if

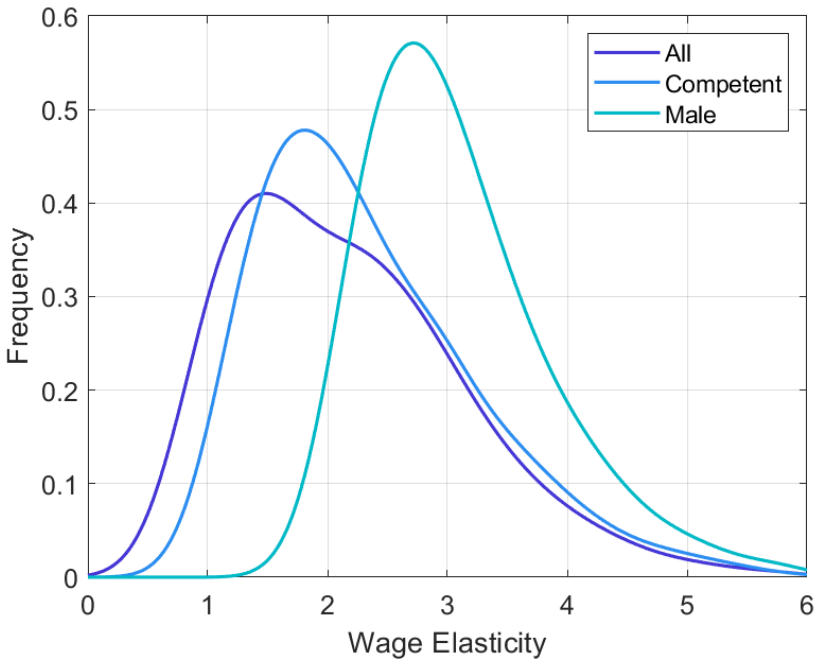
1. Observables are independent from both taste shifter ϵ_{ij} and random coeff. ν_i

   $\Rightarrow$ Wages are set exogenously via deterministic rules and we are controlling flexibly for all relevant wage determinants through $\boldsymbol{z}_j$

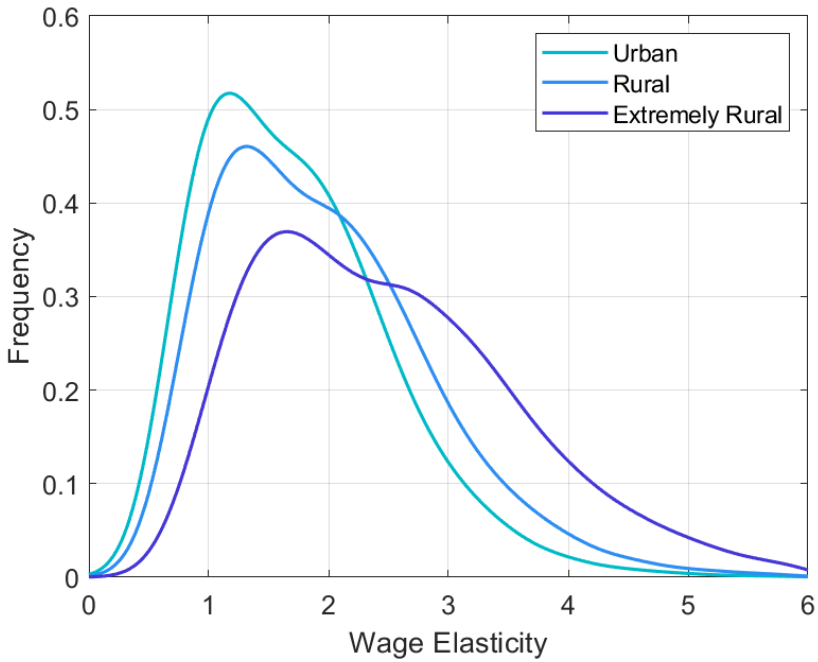
2. Choice sets $\Omega(s_i)$ are independent from taste shifter ϵ_{ij} conditional on observables

   $\Rightarrow$ Vertical preferences of the schools over applicants (by the competency score) rule out the possibility that the decision by teacher i to accept or reject a given job posting may trigger a chain of acceptance or rejections by other teachers that may feed back into teacher i's set of feasible schools

Preference Estimates: Wage Elasticities



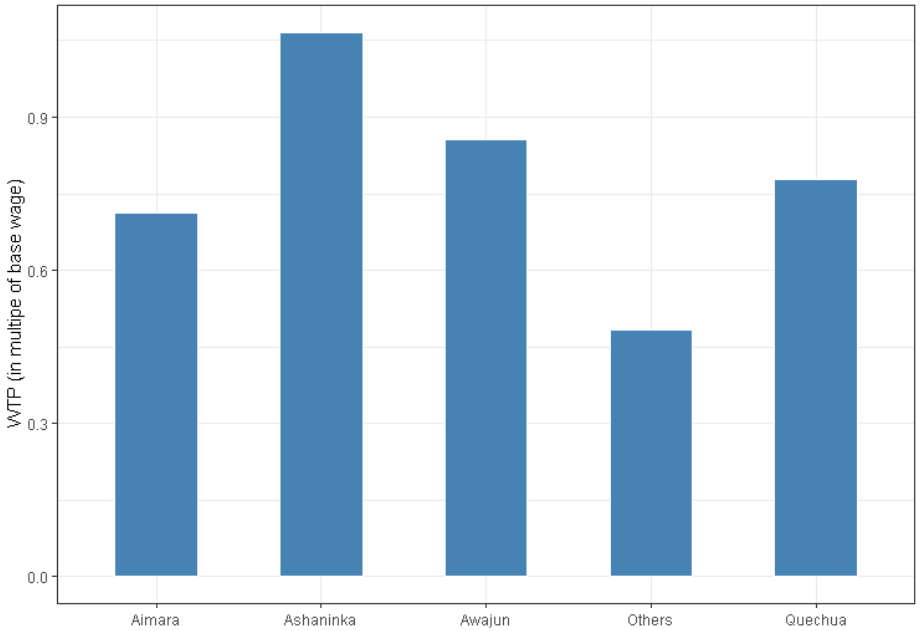
a) Teacher Characteristics



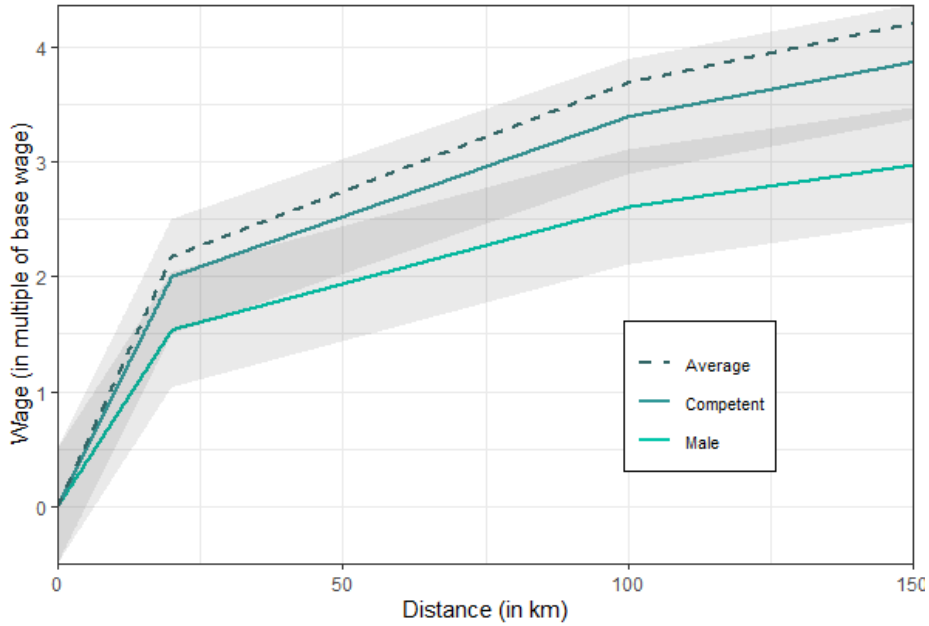
b) School Characteristics

NOTES. These elasticities give the % change in the conditional probability that teacher i chooses school j, which we denote P_{ij}, resulting from a 1% increase in the wage proposed in school j: $\frac{\partial P_{ij}}{\partial w_j} \frac{w_j}{P_{ij}} = \alpha_i w_j (1 - P_{ij})$

Preference Estimates: Match effects



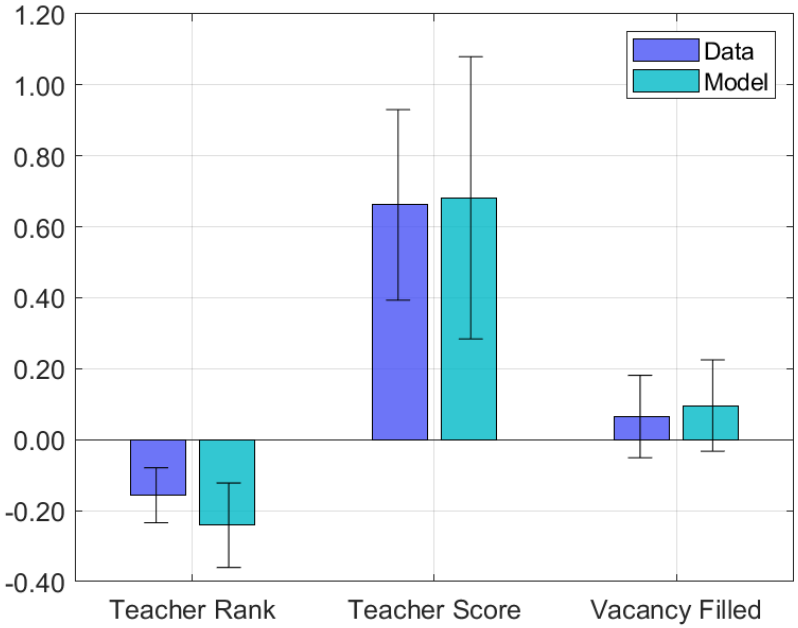
a) Ethnolinguistic Match



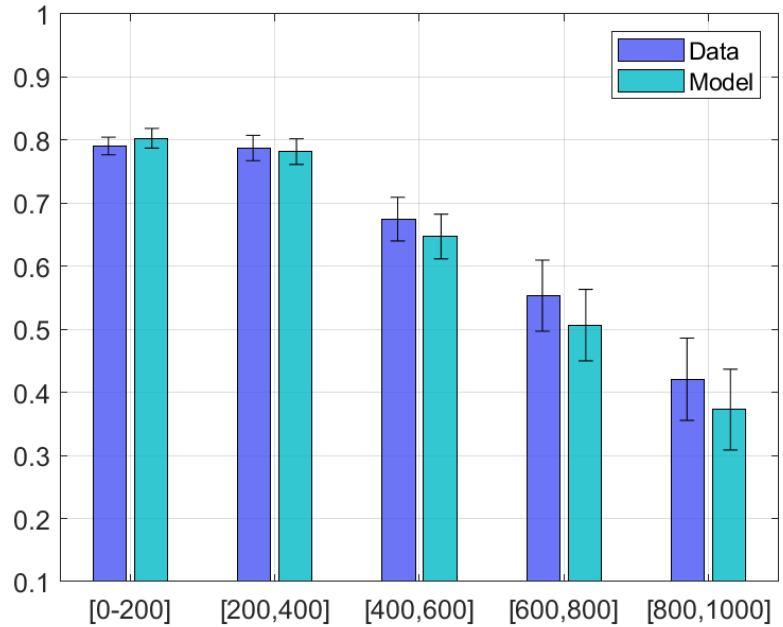
b) Disutility of Distance

NOTES. Panel A depicts the estimated ethnolinguistic match effects in terms of willingness to pay. Panel B shows how much extra money teachers would need to get paid on average in order to make them indifferent between a school located where they live and a school located x km away keeping other observables constant. The dashed line displays this relationship for the average teacher. The other lines display the average for specific groups of teachers (male and competent). The shaded areas around each line displays how the unobserved wage preference heterogeneity affects this relationship by showing confidence bands of the size of one standard deviation of the wage coefficient.

Model Validation (Local Vs. Global Policy Effects)



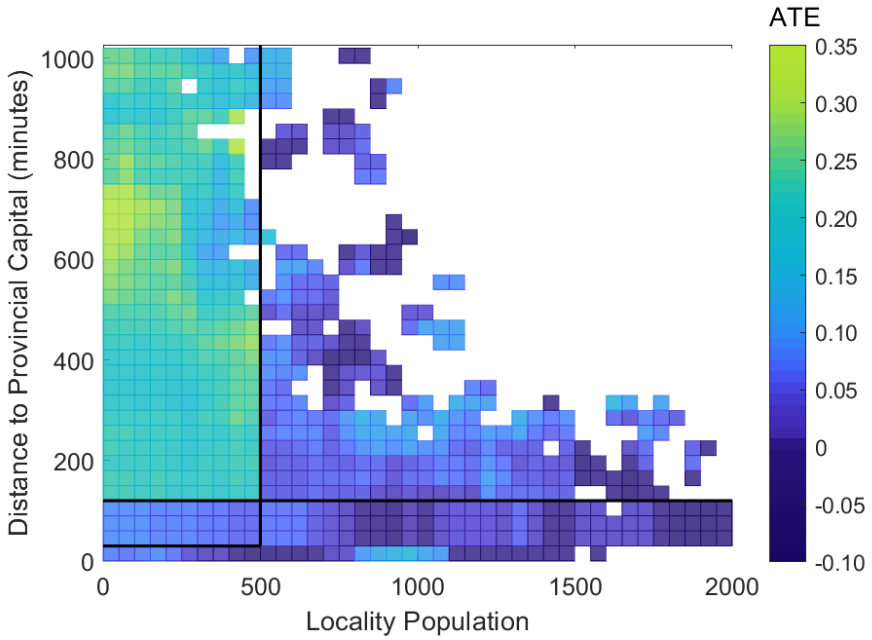
a) Simulated Threshold Crossing Effects



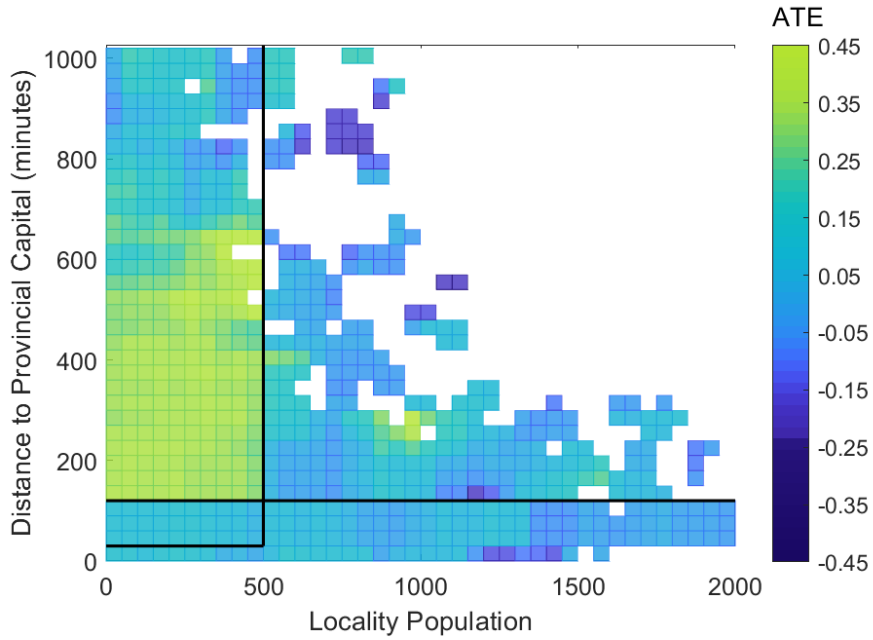
b) Share of Filled Vacancies by Remoteness

NOTES. To assess model fit in Panel A, we predict indirect utilities for each teacher and simulate the match using the serial dictatorship algorithm. We then estimate on simulated data the change in average outcomes at the population cutoff and compare it with the estimated RD effects observed in the data. Panel B compares the share of vacancies filled observed in the actual data and the simulated data depending on how far the schools posting the vacancies are located from the provincial capital.

Model-Based Policy Evaluation



a) Share of Filled Vacancies



b) Teacher Scores

NOTES. This Figure uses simulated assignment data which is generated by running the serial dictatorship algorithm using predicting utilities computed from the estimated preference parameters as well as a randomly drawn set of taste shocks ϵ_{ij}. Panel A plots smooth-weighted cell average differences in the share of vacancies filled between the assignment simulated under the current policy and the counterfactual scenario where we would remove all wage bonuses. Panel B plots smooth-weighted cell-average percentage differences in teacher score at the vacancy level between the assignment simulated under the current policy and the counterfactual scenario where we would remove all wage bonuses.

Takeaways from Model Estimates

1. Wage bonuses have been partly effective in shifting teacher labor supply beyond the population cutoff

   $\Rightarrow$ Reduce the gap in teacher quality between rural/urban schools

   $\Rightarrow$ Fill very undesirable vacancies that would not have been filled otherwise

2. Current wage schedule features limited variation for schools in Extremely Rural areas

   $\Rightarrow$ Scope for improvement with more flexible targeting that leverages heterogenous preferences

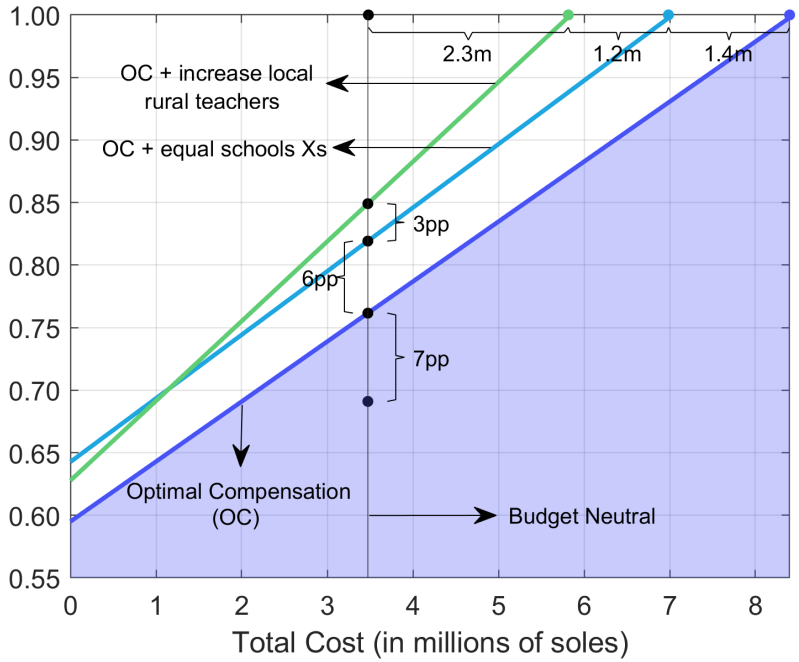
Optimal Wage Policy

- We adapt the matching-with-contracts framework (Hatfield-Milgrom, 2005)
  - Contract: teacher-wage pair
  - Allow schools to poach teachers by posting higher wages
  - Stability: no teacher-school pair would find rematch profitable at given wage
- Iterative algorithm that delivers the lowest equilibrium wage schedule that satisfies a given objective
  1. All vacancies are filled
  2. Fill all schools with above-average-quality teacher in urban areas

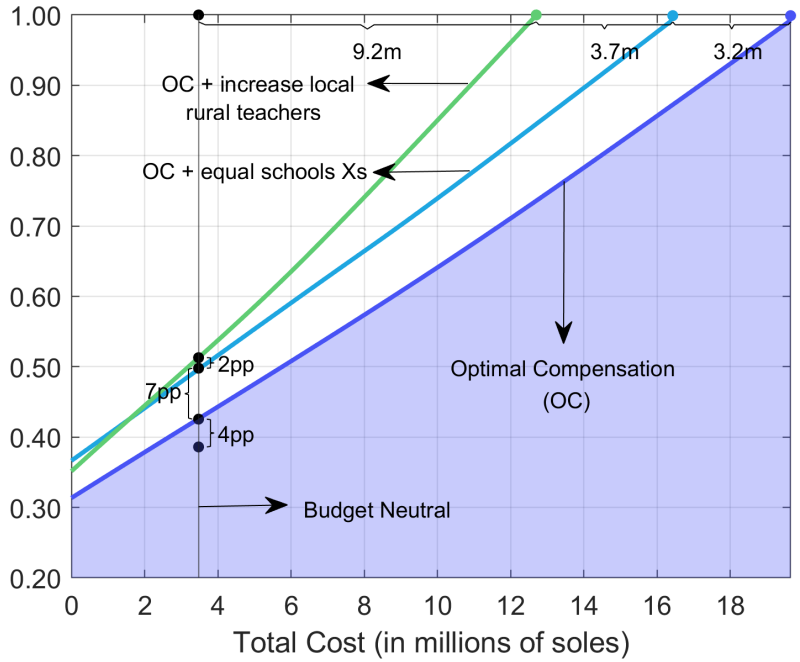
School-proposing Generalized Deferred Acceptance (GDA)

- Step 1: Each school proposes to the highest scoring teacher the lowest wage possible. This teacher is tentatively assigned to her preferred school
- Step k: Schools with remaining vacancies/lower quality teachers propose the same wage to the teacher ranked just below. If they cannot do so, they propose a slightly higher wage to the highest scoring teacher
- Iterate until objective is reached
- Each step of the algorithm with change in wages is optimal – i.e. maximizes the underlying objective given the total wage bill at that round

Cost-Efficiency Frontiers



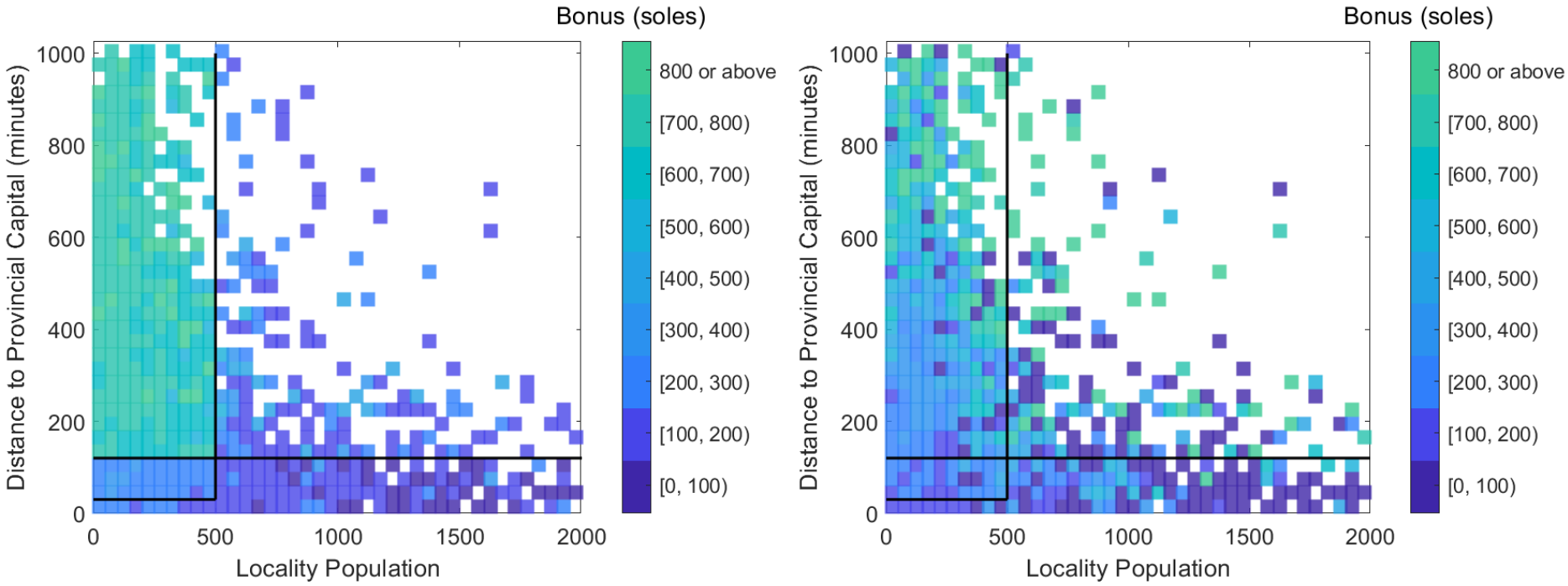
a) Share of Filled Vacancies



b) Share of Schools with Competent Teachers

NOTES. The Figure plots the implied cost in terms of the total wage bill, as derived by the equilibrium matching environment described above, of increasing the share of schools with at least one competent teacher (Panel A) or increasing the share of vacancies filled (Panel B). The dark line shows the benchmark equilibrium, whereas the orange and dotted blue lines consider complementary interventions that shift outward the cost efficiency frontiers.

Budget-Neutral Wage Changes



NOTES. This Figure shows the distribution of wages under the current policy (Panel A) and the distribution of the wages resulting from our algorithm that maximize the share of filled vacancies under the same budget as the current policy (Panel B).

Takeways from Counterfactual Analysis

- Inferring individual preferences is crucial to design cost-efficient policies:
  - Current wage bonus policy is largely misallocated in order to appropriately compensate teachers
  - Complementary place-based policies can also be effective
- Centralized assignment mechanism for public-sector teachers can be used in order to increase equity without fundamentally altering the efficiency of the market allocation

Market Design Meets Research Design

- In this specific application, the RD variation is only used for in-sample validation of the model
  - Equilibrium nature of the model restricts the scope of research design in this case
- A recent literature uses quasi-experimental variation embedded in school assignment mechanisms to study the causal effects of attending particular school types
  - Abdulkadiroglu et al, ECMA 2021
  - Abdulkadiroglu et al, AER 2020