

More Tutoring or Better Targeting?

Research Proposal

Full title: “More Tutoring or Better Targeting? The Economics of Targeted Support for School Students in Wartime Ukraine.”

English title: *More Tutoring or Better Targeting? Allocating Learning-Recovery Support in Wartime Ukraine.*

Proposed paper: an applied economic study of scarce-resource allocation and the marginal value of additional information. The research question, a reproducible empirical pilot, an economic model, and a plan for the main paper have been prepared. Verification was completed on September 12, 2026. This is a developed research proposal with preliminary results, not a completed program evaluation or an already published paper.

Core Idea

Organizers of learning-recovery programs in Ukraine need to decide whom to offer additional tutoring when funding or places are insufficient to serve everyone. They can use simple information about families and schools, or collect more precise information about students' skills. Additional screening also consumes resources: money, teachers' time, and families' time. This study compares these options within a single budget-allocation problem.

Main question: how well do simple rules identify adolescents who lack basic mathematics proficiency, and what is the maximum justified expenditure on additional information when that spending reduces the amount of tutoring available?

The practical outputs are a framework for comparing programs, a table of maximum justified screening expenditures, and a procedure for handling incomplete information. Completing the main paper does not require interviews, contact with participants, access to school records, or microdata from another researcher's experiment. All essential empirical fields have already been obtained.

Abstract

This project studies the allocation of scarce learning-recovery places in wartime Ukraine. Using public PISA 2025 microdata for the covered Ukrainian regions, it compares transparent assignment rules based on school location and family circumstances at equal expected capacity. The primary estimand is access offered to students below basic mathematics proficiency, rather than the causal effect of tutoring. A budget model translates targeting performance into an upper bound on additional information expenditure and evaluates explicitly hypothetical screening accuracy. Exploratory analysis examines incomplete family information and a mixed assignment rule preserving access outside the priority group. The contribution is a policy-specific comparison of access, information costs and inclusion safeguards under wartime resource constraints. The design distinguishes measured educational need from unobserved treatment gains and avoids projecting the sample to excluded regions or children outside its coverage.

1. The Decision This Research Informs

The intended users are departments within Ukraine's Ministry of Education and Science, local education authorities, and organizations that fund additional tutoring. The study does not assume that any of these organizations has agreed to participate or that its current selection process is inefficient. Its findings can be developed independently and presented as evidence to inform policy options.

The study addresses a specific problem: a budget funds a limited number of offers of additional tutoring for adolescents approximately the age represented by PISA. The objective of the first part is to provide access to as many students below basic mathematics proficiency as possible, subject to equity constraints. This is an explicitly chosen targeting objective. It differs from maximizing aggregate learning gains: the returns to tutoring are not observed in the source data.

The question is relevant to existing learning-recovery programs. In a published Ukrainian experiment, demand exceeded program capacity, and access was randomly assigned at the household level. Brief diagnostic assessments were already used to form learning groups in the second and third waves. Thus, the practical allocation problem exists, but "being the first to propose testing" cannot constitute this paper's contribution. [1]

Users of the findings will be able to compare three decisions: expand the number of places using simple selection rules; keep some places open to everyone; or reduce the number of places to obtain additional information. For the third option, the paper will specify the maximum justified cost for a given level of information accuracy. If even perfect additional information cannot justify the proposed price, that supports retaining the funds for tutoring under the chosen objective and cost model.

The study does not measure what proportion of the observed skill shortfall was caused specifically by the war. The pandemic, prewar differences, and selection into the survey also matter. The wartime context defines the recovery problem and budget constraints; it does not automatically identify the cause of every low score.

2. The Proposed Contribution

The proposed contribution is a joint assessment of **targeting performance, information costs, and the consequences of incomplete family information** for a specific Ukrainian policy problem. It should provide a quantitative answer about the choice of allocation rule and the use of a limited budget.

Closely related study	Existing findings	Distinction of the proposed paper
Dinarte-Diaz et al., 2025, Ukraine [1]	Tutoring experiments, diagnostic assessment for group formation, and heterogeneous effects	Comparison of rules for offering scarce places and the budget ceiling for additional screening
Zakotiuk, Oksamytna, Bondar, 2025 [2]	Changes in social disparities in PISA 2018/2022	A decision about allocating places, rather than another description of the social gradient

Closely related study	Existing findings	Distinction of the proposed paper
Scott-Clayton, Crosta, Belfield, 2012 [3]	Errors in placement into remedial education using different information sources	Ukraine's school population during wartime, without assuming that effects transfer from other settings
Johnston et al., 2017 [4]	The economic value of information for targeted school-based preventive programs	An expenditure bound based on observed need when the performance of a specific screening tool is unknown

General theories of targeting and the value of information already exist. The PISA 2025 vintage alone does not establish novelty. The closely related sources reviewed did not reveal this exact combination of a Ukrainian budget-allocation problem, rule comparisons, incomplete information, and a maximum price for additional screening. This is a limited, targeted review, not proof that no competing work exists.

A strong paper should establish a result that changes the policy choice: for example, no meaningful advantage of a more complex social index over a simple geographic rule; substantial sensitivity to missing information; or a small justified budget for additional screening. Any of these findings could be useful, even if it does not confirm the author's initial expectation.

3. Data and Scope of Applicability

The source is the OECD's official PISA 2025 Student Public Use File. The archive has been downloaded and the source SAV file read. Filtering by Ukraine's country code yields 7,105 unique students, 294 schools, and 33 strata across 17 covered regions. The source SAV file contains 2,168,666,883 bytes; its SHA-256 hash is recorded in the pilot log. [5]

Field	Purpose	Verification completed
PV1MATH ... PV10MATH	Distribution of mathematics proficiency	All ten values are present for 7,105 students
W_FSTUWT and 80 W_FSTURWT	Weighting and uncertainty estimation	81 columns with no missing values
STRATUM	Urban or rural school stratum	All 33 labels read; 1,619 rural records
ESCS	Research index of family resources	495 missing values; weighted share 7.11%
HISCED	Highest level of parental education	801 missing values; weighted share 11.46%

The rural stratum refers to the school sample, not a verified current place of residence for the child. ESCS is not a ready-to-use income record in an administrative register. Parental education could potentially be collected through a short questionnaire, but its availability to program organizers at no cost has not been established. Rules based on these fields are therefore research alternatives, not a ready-made register of beneficiaries.

The study population consists of students of the target age represented by the survey in the covered regions. Results cannot automatically be extended to excluded territories, children outside school,

all refugees, or applicants who independently enroll in an online program. Survey weights do not recover the unknown outcomes of these groups. Other PISA waves and air-raid alert data are not essential for the main paper.

4. Hypotheses and Exact Comparison Rules

The primary capacity is $q = 20\%$ of the represented population. This is a specified research scenario, not an existing Ukrainian quota. All rules have the same expected number of place offers. If demand within the priority group exceeds capacity, everyone within that group has an equal probability of selection. If the group is smaller than capacity, the remaining places are allocated among the other students.

The four baseline rules are: random selection from all students; priority for the rural stratum; priority for the bottom weighted ESCS quartile among nonmissing observations; and priority for students whose highest level of parental education is below short-cycle tertiary education, identified by HISCED labels 1-5. Observations with missing information remain in the sample as a separate group. In the baseline ESCS/HISCED rules, missing values do not confer priority; this assumption is examined separately.

H1: rules based on family circumstances or school characteristics cover more baseline need than random assignment at the same capacity. **H2:** the more complex ESCS index delivers a substantively larger improvement than a simple geographic rule. **H3:** the treatment of incomplete information materially changes targeting performance. **H4:** preserving some access outside the priority group has a measurable cost under the primary need criterion.

The four baseline rules were specified before their results were examined in this pilot. The missing-information checks, hybrid rule, and 10%/30% capacity scenarios were added after the initial calculation and are exploratory. The study is not preregistered. The main specification for the paper should be fixed before further expansion of the analysis; results that have already been examined cannot be presented as new confirmatory tests.

5. What Is Estimated

Basic mathematics proficiency is defined by the threshold $z = 420.07$ on the PISA scale. The indicator $Y < z$ and the depth of the shortfall $\max(z - Y, 0)$ are calculated separately for each plausible value. Plausible values reflect uncertainty in the proficiency distribution; they cannot be averaged for each child and treated as a precise diagnostic assessment. The threshold corresponds to the official mathematics proficiency scale. [6]

The primary measure is the **share of students below basic proficiency who are offered a place**. Additional measures are the share of these students among all offer recipients and the share of the total baseline proficiency shortfall attributable to recipients. The latter measures coverage of need, not the shortfall eliminated.

Calculations use ten plausible values and each of the 81 sets of weights. The ESCS quartile, priority-group size, and selection probabilities are recalculated for every replicate. BRR-Fay with a parameter of 0.5 is used: the within-imputation variance equals 0.05 times the sum of squared deviations across the 80 replicates. Between-imputation variance is then added with a multiplier of

1.1. Differences between rules are calculated on a paired basis before the variances are combined. [7]

Normal-approximation intervals are reported as estimate ± 1.96 SE. They describe sampling and proficiency uncertainty, but not randomness in a future lottery for places, tutoring effects, the accuracy of a new test, or incomplete national coverage. The numerous exploratory comparisons should not be interpreted as a set of independent confirmations.

6. An Empirical Pilot: Results That Support the Research Design

The estimated share of students below basic mathematics proficiency is 45.96%, with a preliminary 95% interval of 43.47-48.45%. The comparisons below allocate offers of places to 20% of the represented population.

Rule	Coverage of students below basic proficiency	Share below basic mathematics proficiency among recipients	Share of baseline proficiency shortfall attributable to recipients
Random assignment	20.00%	45.96%	20.00%
Rural-stratum priority	25.31%	58.17%	28.83%
Low-ESCS priority	25.43%	58.44%	28.48%
Parental-education priority	23.39%	53.74%	24.47%

Under the rural rule, coverage increases by 5.31 percentage points relative to random selection, with a paired interval of 3.55-7.08 percentage points. However, the difference between the rural and ESCS rules is only -0.12 percentage points, with an interval from -2.22 to 1.99. A statistically discernible difference has not been established; equivalence has not been demonstrated either.

For illustration, in a hypothetical population of 1,000 students with 200 place offers, the random rule would reach approximately 92 adolescents below basic proficiency, while the rural rule would reach approximately 116. This converts weighted shares to a hypothetical group size; it is not a new observed sample or evidence of 24 additional children receiving tutoring. Actual participation and learning gains were not measured.

These findings motivate an economically useful hypothesis: collecting more complex information should be justified by additional targeting gains, not merely by a more detailed questionnaire. The pilot does not yet support recommending ESCS or rural status as the sole condition for actual access.

7. Incomplete Information and Preserving Access

When students with missing ESCS are added to the priority group, coverage of students below basic proficiency rises from 25.43% to 27.29%. The paired difference is 1.86 percentage points, with an interval of 1.27-2.45. The share of the baseline shortfall attributable to recipients rises from 28.48% to 32.02%. Thus, the treatment of missing values materially affects this rule's performance.

This is not evidence that missing information indicates poverty, nor a proposal to reward incomplete questionnaires. The reasons for missing survey responses and people's reactions to real program

rules may differ. The practical implication for program design is to provide a transparent process for clarifying information and separately reviewing such cases, rather than automatically excluding them from consideration. The cost and effect of such a process have not yet been measured.

An exploratory hybrid allocates 25% of expected capacity through selection open to all students and 75% through rural priority. At $q = 20\%$, this gives a student in an urban stratum a 5% probability of receiving a place offer, and a student in a rural stratum a 68.54% probability. The rule reaches 23.99% of students below basic proficiency, compared with 25.31% under pure rural priority. The cost of this broader access is 1.33 percentage points of coverage, with a paired interval of 0.89-1.77 percentage points.

This hybrid is not claimed to be optimal. It illustrates how the study can quantify the trade-off between targeting and preserving access outside the priority category. Baseline calculations have also been completed at capacities of 10% and 30%; all tables and uncertainty estimates are available in the accompanying aggregate files.

8. An Economic Model of Information Expenditure

Let N be the size of the target population, B the budget after common baseline administrative expenses, and c the uniform cost of reserving one place. Then $q_0 = B/(N \times c)$. If additional information is collected from all N individuals at a cost of s per person, the share that can still be offered places is $q(s) = q_0 - s/c$. Charging screening costs only to future recipients would be incorrect: those recipients have not yet been identified.

The main calculation is the maximum price of additional information that does not reduce the number of offers to students below basic mathematics proficiency relative to the baseline. Under perfect knowledge of proficiency status, when the share below basic proficiency exceeds q_0 , the formula is $s_{\max}/c = q_0 \times (1 - p_0)$, where p_0 is the share below basic proficiency among recipients under the baseline rule.

Relative to the rural rule at $q_0 = 20\%$, the ceiling is **8.37% of the cost of one place per person screened**, with a preliminary interval of 7.39-9.34%. This is an upper bound under perfect knowledge of proficiency status and the chosen objective. It is not a recommended test price, demonstrated savings, or an evaluation of an existing diagnostic instrument. At $q_0 = 20\%$, this ceiling corresponds to approximately 41.8% of the initial budget for places when summed across everyone screened, not 8.37% of the total budget. Lower information accuracy reduces the justified expenditure.

For a test scenario with sensitivity a and specificity b , the share of positive results is $g = aL + (1 - b)(1 - L)$, where L is the share below basic proficiency. The positive predictive value is $pD = aL/g$. If $pD > p_0$ and there are enough positive results to fill the required number of places, the ceiling is $s_{\max}/c = q_0 \times (1 - p_0/pD)$. If positive results are insufficient, a separate rule is needed to fill the remaining places; the simple formula no longer applies. The appendix calculates a grid of 25 hypothetical combinations of a and b for each baseline rule.

If tutoring has the same effect for every student, this model does not justify paying for screening to increase aggregate learning gains: screening reduces the number of places. Larger effects for students below basic proficiency could change the result, but transferring such effects from other

experiments cannot be treated as established. Prioritizing these students has independent value only as an explicitly chosen access objective or under additional assumptions about returns.

9. Recommendations Supported by Different Findings

Research finding	Decision it supports	What the finding does not establish
A complex characteristic provides no convincing improvement over a simple one	Do not justify purchasing a questionnaire solely because it describes families in greater detail; compare its price with the range of potential gains	A nonsignificant difference cannot establish exact equality between rules
Incomplete information changes targeting performance	Provide for verification and clarification of incomplete cases; report their coverage separately	Missing information should not automatically confer priority
Additional screening costs exceed even the perfect-information ceiling	Preserve funds for places under the chosen objective and cost structure	Other useful functions of diagnostic assessment are outside this narrow model
The price is below the ceiling, but accuracy is unknown	Obtain evidence of accuracy before deciding to implement	A low price alone does not establish that a test is useful
A mixed rule preserves access outside the priority category	Choose an explicit trade-off between access and targeting	The scenario's 25% share is not an established optimum

Diagnostic assessment can also help form learning groups and adapt instruction. These pedagogical benefits are not included in the calculated ceiling for its selection function. The finding therefore cannot be used as a general argument against diagnostic assessment. It concerns additional expenditure justified specifically by better allocation of scarce places.

The paper may conclude with a conditional recommendation and bounds within which a decision can be made confidently. Actual implementation requires checking applicability to the program's own applicant pool, as well as the cost and accuracy of the instrument used. These are subsequent policy steps, not hidden prerequisites for obtaining this paper's main findings.

10. Plan for Completing the Paper

Research design and reproducibility are already in place: all essential data have been read; the primary pilot, 10%/30% capacity scenarios, missing-information checks, and one hybrid rule have been calculated; the main estimates and standard errors have been independently reproduced from the source SAV file using separate code. The maximum absolute discrepancy was 1.4×10^{-16} .

Next analytical steps: fix one primary criterion and a family of rules; add robustness checks for alternative proficiency thresholds, the sequential exclusion of individual regions, and access constraints. Alternative thresholds should not retrospectively replace the primary definition. Small strata should be pooled using a substantively justified rule. Compare findings with the characteristics of groups excluded from coverage, without assigning the calculated values to those groups.

Economic analysis: extend the cost comparison to unequal costs per place, different participation rates, and explicitly specified effect heterogeneity. The result will be a set of scenarios in which the decision is robust. If small, plausible changes in assumptions change the decision, that should be reported as uncertainty about the decision.

Paper structure: introduction and the organizer's decision; institutional context; data; model and targeting criterion; main results; information expenditure; robustness checks; applicability and recommendations. The technical appendix will include a data dictionary, variance calculations, scenario parameters, and a limitations log. These sections do not require new interviews.

The empirical project remains a study of access to support in the represented population. Causal estimates of the war's effects, returns in future earnings, GDP effects, and the effectiveness of a particular diagnostic test are outside the promised scope. Adding these questions would require a different research design.

11. Materials for the Professor and a Future Website

The professor will receive a coherent research concept with a completed pilot and code. It does not rest on the vague assumption that "the student will find the data later." The directory contains the original file URLs, hash, exact rules, intervals, and a separate log of independent replication. Pilot aggregates are available in CSV format, with hypothetical scenarios provided separately.

A future website could include an explanation of the decision; a comparison of how 200 places are allocated in a hypothetical group of 1,000 students; a capacity selector; the implications of rules for incomplete questionnaires; a model of diagnostic costs and accuracy; and a table of decisions and sources. Every chart should identify the population, unit of measurement, and status of each number: a data-based estimate or a specified scenario.

External presentation will use aggregate results and the project's own code. Student identifiers and attempts to identify PISA schools are unnecessary. The research website should not appear to be a system for assigning actual aid recipients.

Assessment of the choice: this option is recommended as the most developed project on academic achievement with a specific economic decision at its core. The alternative on coordinating grants and minimum tuition prices in higher education has the strength of a direct financial decision, but its broad panel of programs and the precise applicability of the rules have not yet been verified. It is retained separately and is not combined with the main study.

Sources and Appendices

[1] Dinarte-Diaz, L., Gresham, J., Lemos, R., Patrinos, H., Rodriguez-Ramirez, R. (2025). *Investing in Human Capital During Wartime: Experimental Evidence from Ukraine*. World Bank Policy Research Working Paper 11196. [Full text](#). The design, selection, diagnostic assessment, and limitations of comparisons across experiments were checked. That study's results are not attributed to this pilot.

[2] Zakotiuk, Oksamytna, Bondar (2025). Study of learning losses and inequality using PISA 2018/2022. [Article](#).

[3] Scott-Clayton, J., Crosta, P., Belfield, C. (2012). *Improving the Targeting of Treatment: Evidence from College Remediation*. NBER WP 18457. [Authors' working version](#).

[4] Johnston et al. (2017). *Estimating the Economic Value of Information for Screening in Disseminating and Targeting Effective School-based Preventive Interventions: An Illustrative Example*. [Publication](#). Bibliographic details and the abstract were checked; automated access to the full text during this verification was restricted by a browser-verification check.

[5] OECD. [PISA 2025 Database](#); [official Student PUF archive](#); [description of the Ukrainian sample](#).

[6] OECD. [PISA 2022 Technical Report, Annex 17.A](#); [PISA FAQ on scales](#).

[7] OECD. [Preparing and analyzing the PISA database; technical notes for the 2025 results](#). The published BRR-Fay and plausible-value pooling approach is used; no claim is made to have read the full forthcoming 2025 technical report.

Local reproduction: [pilot methodology and execution](#), [main estimates CSV](#), [sensitivity checks CSV](#), [paired contrasts](#), [diagnostic scenarios CSV](#), [independent replication](#), [economic-model review](#), [alternative proposal](#).