Who Benefits from Remote Schooling?
Self-selection and Match Effects

Jesse Bruhn\textsuperscript{1}  Christopher Campos\textsuperscript{2}  Eric Chyn\textsuperscript{3}

\textsuperscript{1}Brown University  \textsuperscript{2}University of Chicago and NBER  \textsuperscript{3}UT-Austin and NBER

NBER Spring Education Meeting May 2023
Motivation

- School choice expansions are hypothesized to lead to improvements in student-school match quality
  - Families can more efficiently sort into schools that suit their students’ needs (Hoxby 2003)
Motivation

- School choice expansions are hypothesized to lead to improvements in student-school match quality.
- Match effects are theoretically important but evidence is thin.
  - Do parents know their match quality? Perhaps not, due to lack of information. (Abdulkadiroglu et al. 2020; Clark, Martorell, and Wiswall 2022)
  - Sorting on match quality? Limited evidence. (Bruhn 2020; Campos and Kearns 2023)
  - Horizontal differentiation is important in some settings. (Bau 2021; Gilraine, Petronijevic, and Singleton 2021)
Motivation

• School choice expansions are hypothesized to lead to improvements in student-school match quality

• Match effects are theoretically important but evidence is thin

• The pandemic provides a unique context to study sorting and match effects
  → Families compelled to switch to remote learning options
  → Families learn about their match quality with respect to remote learning
Motivation

- School choice expansions are hypothesized to lead to improvements in student-school match quality
- Match effects are theoretically important but evidence is thin
- The pandemic provides a unique context to study sorting and match effects
  - Families compelled to switch to remote learning options
  - Families learn about their match quality with respect to remote learning
- Post-pandemic demand for remote learning is common across numerous cities
  - Pandemic changed enrollment patterns (Dee and Murphy 2021; Musaddiq et al. 2022; Dee 2023)
  - A host of explanations for why families may prefer remote learning (Bacher-Hicks et al. 2022)
  - Common Core data reveals sizable increase in students enrolling in exclusively virtual schools
Motivation: Trends in Remote Learning (Common Core)

Definition of Remote School: All instruction offered by the school is virtual.
Motivation: In the Press

Sticking around: Most big districts will offer virtual learning this fall, a sign of pandemic’s effect

By Kalyn Belsha and Matt Barnum | Jun 8, 2022, 7:00am EDT
Sticking around: Most big districts will offer virtual learning this fall, a sign of pandemic’s effect

By Kalyn Belsha and Matt Barnum | Jun 8, 2022, 7:00am EDT

NYC rolling out 2 virtual learning programs with aim to turn them into fully remote schools by 2023

By Cayla Bamberger | June 23, 2022 | 6:03pm | Updated
Sticking around: Most big districts will offer virtual learning this fall, a sign of pandemic's effect

By Kalyn Belsha and Matt Barnum | Jun 8, 2022, 7:00am EDT

NYC rolling out 2 virtual learning programs with aim to turn them into fully remote schools by 2023

By Cayla Bamberger | June 23, 2022 | 6:03pm | Updated

LAUSD Virtual Academy: Pathway to the Future

Welcome to Los Angeles Unified's Virtual Academy! While the Los Angeles Unified School District (LAUSD) believes in-person learning is generally best for most students, some families may want or need other options. LAUSD is excited to announce new Virtual Academy Schools for the 2022-2023 school year. The online academies will explore and expand independent study in a broader way using technology and promoting creativity.
This paper

- Leverage the pandemic as a shock to families’ awareness of their remote learning match quality
- Focus on Los Angeles where families were allowed to self-select into remote options in 2021-2022
- Collect novel survey data containing rich information about families’ preferences for remote learning
  - Build on existing working using conjoint experiments (Mas and Pallais 2017; Wiswall and Zafar 2018)
  - Innovation: Link experimental preference estimates to potential outcome model to identify treatment effects
- Use experimental preference estimates to assess Roy-style selection into remote learning in the post-pandemic landscape

Open questions we focus on:

1. What are families’ preferences for remote learning? Heterogeneity?
2. What are the selection patterns into remote learning and how does that affect outcomes?
Preview of Results

Survey Evidence:
1. The typical family in Los Angeles dislikes remote learning
   → Need to be compensated with a 40 percentage point higher school proficiency rate to be indifferent between remote and in-person learning
2. Roughly 25% of families anticipate demanding remote options in the future
3. Roughly 20% of families perceived their kids perform better in remote relative to in-person

Treatment Effects and Heterogeneity:
4. Use experimental preference estimates to construct propensity scores summarizing selection into remote
   → Preferences predict actual choices
   → Balance lagged achievement conditional on implied propensity score
5. Remote ATT: \(-0.11\sigma\) (SE: 0.028) for reading and \(-0.13\sigma\) (SE: 0.027) for math
6. Positive remote effects for families with large estimated tastes for remote learning; evidence of Roy-style selection
Related Literature

Match Effects


• Contribution: Evidence of parents sorting children into schools that is consistent with theory

Remote Learning

• Bueno 2020, Kofoed et al. 2021, Goldhaber et al. 2022, Jack et al. 2022, Singh et al. 2022

• Contribution: Study heterogeneity in remote learning effects in a post-pandemic landscape and implications for educational inequality and efficiency

Using choice models to estimate treatment effects

• Heckman 1979, Heckman et al. 2006, Abdulkadiroğlu et al. 2020, Mas and Pallais 2017, Wiswall and Zafar 2018

• Contribution: Import insights from conjoint experiments to program evaluation framework
Related Literature

**Match Effects**


- **Contribution:** Evidence of parents sorting children into schools that is consistent with theory

**Remote Learning**

- Bueno 2020, Kofoed et al. 2021, Goldhaber et al. 2022, Jack et al. 2022, Singh et al. 2022

- **Contribution:** Study heterogeneity in remote learning effects in a post-pandemic landscape and implications for educational inequality and efficiency
Related Literature

**Match Effects**
- **Contribution:** Evidence of parents sorting children into schools that is consistent with theory

**Remote Learning**
- Bueno 2020, Kofoed et al. 2021, Goldhaber et al. 2022, Jack et al. 2022, Singh et al. 2022
- **Contribution:** Study heterogeneity in remote learning effects in a post-pandemic landscape and implications for educational inequality and efficiency

**Using choice models to estimate treatment effects**
- Heckman 1979, Heckman et al. 2006, Abdulkadiroğlu et al. 2020, Mas and Pallais 2017, Wiswall and Zafar 2018
- **Contribution:** Import insights from conjoint experiments to program evaluation framework
1. Setting: Remote learning during the pandemic and post-pandemic

2. Survey evidence
   → Eliciting Preferences and Estimation
   → Evidence

3. Conceptual Framework
   → Potential outcome model
   → Map conjoint experiment to treatment effects
   → Treatment Effects and Heterogeneity

4. Concluding Thoughts
Remote Schooling in Los Angeles

- LAUSD announced two week closure on March 13, 2020
- The 2019-2020 academic year finishes entirely remote
  - 150,000 computer shortage
  - 20% of families without computer or laptop and 16% without access to internet
- 83% of LAUSD teachers voted to start the 2020-2021 year virtually
- LAUSD was criticized for staying virtual longer than neighboring districts
- LAUSD stayed remote until April 2021
  - 30% of elementary school students return when schools reopen
  - 12% of middle school students return
  - 7% of high school students return

-USC Annenberg Report
“Post-Pandemic” Remote Schooling in Los Angeles

• California mandates schools to offer a remote option for the 2021-2022 academic year

• LAUSD creates a virtual City of Angels school to accommodate families choosing remote

• 4.7% of students enroll in the remote option

• LAUSD opens six virtual academies for 2022-2023 academic year

• Other large districts also expanding remote offerings
Data

LAUSD Student Data 2019-2022

- Demographics
- Addresses
- Achievement

Novel Survey Data

- $\sim 3,400$ survey responses
  - Highly selected sample
  - We model preference heterogeneity parametrically in extrapolation
  - Extrapolation produces estimates that are forecast unbiased
- Link survey data with administrative data
The Survey

**Goals:**
- Collect descriptive facts about families’ experiences during Covid remote era
- Collect descriptive facts about anticipated demand for remote learning
- Experimentally estimate preferences for remote learning

**Issues in practice:**
- Survey responses may not reflect real choices
  - Our experimental estimates are consistent with estimates using actual choices
- Low survey response rates introduce selection bias *(Dutz et al. 2021)*
  - Strong covariate support allows us to account for preference heterogeneity parametrically
Eliciting Preferences

- Well-developed literature using conjoint survey instruments to experimentally identify preferences (Mas and Pallais 2017; Wiswall and Zafar 2018; Aucejo, French, and Zafar 2022)

- Respondents presented with sequence of hypothetical choices randomly varying product attributes

- Experimental manipulation of attributes credibly identifies preferences

- We elicit preferences for three important school attributes
  - Travel time
  - Academic Quality
  - Learning modality (in-person and remote)
Eliciting Preferences

6. You will now see a sequence of scenarios, each with three school options that the school district could offer you in Fall 2022. For each set of three, indicate the one you prefer the most (Best) and the one you prefer the least (Worst).

Recall that a fully remote option is entirely virtual (100% remote) and traditional in-person instruction is 0% remote.

Travel time corresponds to the commute time in minutes from your home to the school. For traditional in-person instruction, students make the trip to school every day.

Assume pandemic-related safety issues are as they were in 2019 before COVID.

Besides the characteristics shown, assume that these schools are otherwise identical in terms of their academic instruction and quality.

There are no right or wrong answers to these questions. We only want to know which of the options you would most prefer.
## Eliciting Preferences

<table>
<thead>
<tr>
<th>Type of Instruction</th>
<th>Fully Remote</th>
<th>In Person</th>
<th>Fully Remote</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent of students meeting state</td>
<td>20</td>
<td>60</td>
<td>30</td>
</tr>
<tr>
<td>academic standards</td>
<td>Zero</td>
<td>15</td>
<td>Zero</td>
</tr>
<tr>
<td>Travel time to school (minutes)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Best</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Worst</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Eliciting Preferences

<table>
<thead>
<tr>
<th>Type of Instruction</th>
<th>In Person</th>
<th>In Person</th>
<th>Fully Remote</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent of students meeting state academic standards</td>
<td>10</td>
<td>50</td>
<td>70</td>
</tr>
<tr>
<td>Travel time to school (minutes)</td>
<td>60</td>
<td>15</td>
<td>Zero</td>
</tr>
<tr>
<td>Best</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Worst</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Preference Estimation

Indirect utility of school $j$ for respondent $i$ at hypothetical choice $k$ is

\[ U_{ikj} = \omega Q_j + \omega_R \text{Remote}_j + \omega_d d_{ij} + \varepsilon_{ikj} \]

- $Q_j$: Academic quality of hypothetical school $j$
- $\text{Remote}_j$: Remote learning indicator
- $d_{ij}$: Travel time to hypothetical school $j$
- $\varepsilon_{ikj}$: iid logit shock
Preference Estimation

Indirect utility of school $j$ for respondent $i$ at hypothetical choice $k$ is

$$U_{ikj} = \omega Q_j + \omega R Remote_j + \omega_{d}d_{ij} + \varepsilon_{ikj}$$

- We collect a rank-ordered list $R_{ik} = (R_{1k}, R_{2k}, R_{3k})$ in each hypothetical scenario, with top-ranked option satisfying

$$R_{1ik} = \arg\max_{j \in J_k} U_{ikj}$$

and the following options satisfy

$$R_{irk} = \arg\max_{j \in J_k \{R_{imk} : m < r\}} U_{ikj} \quad r > 1, \quad k = 1, \ldots, 10$$
Preference Estimation

Indirect utility of school $j$ for respondent $i$ at hypothetical choice $k$ is

$$U_{ikj} = \omega Q_j + \omega_R \text{Remote}_j + \omega_d d_{ij} + \varepsilon_{ikj}$$

- Logit assumption implies that the likelihood function for a given individual $i$ and hypothetical scenario $k$ is

$$\mathcal{L}(R_{ik}|Q_j, \text{Remote}_j, d_{ij}, X_i) = \frac{\exp(V_{iR_{i1k}})}{\sum_{m \in \{R_{i1k}, R_{i2k}, R_{i3k}\}} \exp(V_{im})} \frac{\exp(V_{iR_{i2k}})}{\sum_{m \in \{R_{i2k}, R_{i3k}\}} \exp(V_{im})}$$
<table>
<thead>
<tr>
<th></th>
<th>All Students</th>
<th>Survey Respondents</th>
<th>Conjoint Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lagged ELA</td>
<td>0</td>
<td>.18</td>
<td>.46</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(1.05)</td>
<td>(1.05)</td>
</tr>
<tr>
<td>Lagged Math</td>
<td>.01</td>
<td>.17</td>
<td>.44</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(1.01)</td>
<td>(1.02)</td>
</tr>
<tr>
<td>Female</td>
<td>.48</td>
<td>.48</td>
<td>.49</td>
</tr>
<tr>
<td></td>
<td>(.5)</td>
<td>(.5)</td>
<td>(.5)</td>
</tr>
<tr>
<td>Special Education</td>
<td>.12</td>
<td>.1</td>
<td>.09</td>
</tr>
<tr>
<td></td>
<td>(.32)</td>
<td>(.3)</td>
<td>(.29)</td>
</tr>
<tr>
<td>URM</td>
<td>.8</td>
<td>.76</td>
<td>.65</td>
</tr>
<tr>
<td></td>
<td>(.4)</td>
<td>(.43)</td>
<td>(.48)</td>
</tr>
<tr>
<td>N</td>
<td>225,696</td>
<td>3,447</td>
<td>1,148</td>
</tr>
</tbody>
</table>
## Survey Evidence

<table>
<thead>
<tr>
<th></th>
<th>All Students</th>
<th>Survey Respondents</th>
<th>Conjoint Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lagged ELA</td>
<td>0</td>
<td>.18</td>
<td>.46</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(1.05)</td>
<td>(1.05)</td>
</tr>
<tr>
<td>Lagged Math</td>
<td>.01</td>
<td>.17</td>
<td>.44</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(1.01)</td>
<td>(1.02)</td>
</tr>
<tr>
<td>Female</td>
<td>.48</td>
<td>.48</td>
<td>.49</td>
</tr>
<tr>
<td></td>
<td>(.5)</td>
<td>(.5)</td>
<td>(.5)</td>
</tr>
<tr>
<td>Special Education</td>
<td>.12</td>
<td>.1</td>
<td>.09</td>
</tr>
<tr>
<td></td>
<td>(.32)</td>
<td>(.3)</td>
<td>(.29)</td>
</tr>
<tr>
<td>URM</td>
<td>.8</td>
<td>.76</td>
<td>.65</td>
</tr>
<tr>
<td></td>
<td>(.4)</td>
<td>(.43)</td>
<td>(.48)</td>
</tr>
<tr>
<td>N</td>
<td>225,696</td>
<td>3,447</td>
<td>1,148</td>
</tr>
</tbody>
</table>
Survey Evidence: Families value academic quality

![Bar chart showing minutes spent on academic activities by different school levels and enrollment plans.](chart.png)
Survey Evidence: Families dislike remote learning

![Bar chart showing required increase in achievement for different school levels and current remote status.](image-url)
Taking Stock of Results and Next Steps

Survey Evidence

• 23% of respondents are likely to choose remote learning options in the future
• 20% of respondents believe their students excelled in remote relative to in-person
• Families tend to value academic quality and dislike remote learning

Next Steps

• Pivot to causal remote learning effects
• Use propensity scores implied by choice model to select correct estimates
• Do families with a higher proclivity for remote learning have larger treatment effects?
Empirical Framework

\[ Y_i = \alpha + X_i'\gamma + \beta D_i + u_i, \]

- \( Y_i \): 2022 math or ELA achievement
- \( D_i \): Remote learning indicator
- \( X_i \): Baseline covariates, including lagged achievement

**Issues:**
- Selection into remote is not random
- Preferences likely correlated with potential outcomes
We construct propensity scores using the conjoint experiment

The indirect utility of family $i$ enrolling in the remote option relative to their neighborhood school is

$$U_i = \omega_R + \omega_Q Q_{j(i)} - \omega_d d_{j(i)} + \varepsilon_i$$
We construct propensity scores using the conjoint experiment

The indirect utility of family $i$ enrolling in the remote option relative to their neighborhood school is

$$U_i = \omega_{Rc}(x_i) + \omega_{Qc}(x_i)Q_{j(i)} - \omega_{dc}(x_i)d_{j(i)} + \varepsilon_i$$
We construct propensity scores using the conjoint experiment

The indirect utility of family $i$ enrolling in the remote option relative to their neighborhood school is

$$U_i = \omega_{Rc}(X_i) + \omega_{Qc}(X_i)Q_{j(i)} - \omega_{dc}(X_i)d_{j(i)} + \varepsilon_i$$

- Allow for heterogeneous tastes for quality, remote, and travel time
- Heterogeneity in terms of baseline covariates $c(X_i)$ which depend on grade, gender, baseline achievement, and URM status
- $Q_{j(i)}$: remote achievement relative to neighborhood school $j(i)$
- $d_{j(i)}$: travel time to neighborhood school $j(i)$
We construct propensity scores using the conjoint experiment

The indirect utility of family $i$ enrolling in the remote option relative to their neighborhood school is

$$U_i = \omega_{Rc}(X_i) + \omega_{Qc}(X_i) Q_j(i) - \omega_{dc}(X_i) d_j(i) + \varepsilon_i$$

- Conjoint experiment identifies $(\omega_{Qc}, \omega_{Rc}, \omega_{dc})$
- We observe families’ neighborhood school $Q_j(i)$ and $d_j(i)$
- With estimates of $(\omega_{Qc}, \omega_{Rc}, \omega_{dc})$ and $Q_j(i)$ and $d_j(i)$, we can construct individual-level propensity scores of choosing remote relative to the neighborhood school for the entire sample:

$$P(v_i) = \frac{\exp(\omega_{Rc}(X_i) + \omega_{Qc}(X_i) Q_j(i) - \omega_{dc}(X_i) d_j(i))}{1 + \exp(\omega_{Rc}(X_i) + \omega_{Qc}(X_i) Q_j(i) - \omega_{dc}(X_i) d_j(i))}$$
Propensity scores are forecast unbiased and predict actual enrollment

- Concern: Survey respondents are not representative
  - We estimate preferences separately by covariate cell
  - There is common support among respondents and non-respondents

- Concern: Extrapolation to the entire sample
  - Propensity scores are forecast unbiased

- Concern: Hypothetical choices do not reflect real choices
  - Propensity scores predict real enrollment patterns
Identifying Assumptions

\[ E[Y_i|X_i, D_i, P(v_i)] = \alpha_{c(X_i)} + \beta D_i + \theta P(v_i) + \psi P(v_i) \times D_i. \]  

(1)

- \( \theta \) governs selection on levels
- \( \psi \) governs selection on gains; match effects

Assumptions and implications

- Propensity score summarizes selection into remote
  → Also try a related control function approach (Abdulkadiroglu et al. 2020)
- Linear treatment effect heterogeneity
  → Assumption relaxed as a robustness check
- Testable Implication:

\[ E[X_i|D_i = 1, P(v_i)] - E[X_i|D_i = 0, P(v_i)] = 0 \]
Propensity score mostly balances lagged achievement

<table>
<thead>
<tr>
<th></th>
<th>ELA</th>
<th>Math</th>
<th>ELA</th>
<th>Math</th>
<th>ELA</th>
<th>Math</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uncond.</td>
<td>-.2</td>
<td>-.15</td>
<td>-.1</td>
<td>-.05</td>
<td>0</td>
<td>.05</td>
</tr>
<tr>
<td>Prop. Score</td>
<td>.05</td>
<td>.05</td>
<td>.05</td>
<td>.05</td>
<td>.05</td>
<td>.05</td>
</tr>
</tbody>
</table>

Difference in Lagged Achievement

![Graph showing the difference in lagged achievement](image-url)
Observational propensity score does not balanced lagged achievement
Treatment Effects

Remote Effect

ELA  Math  ELA  Math  ELA  Math

Canonical VAM  Propensity Score  Control Function
Heterogeneity

![Diagram showing the relationship between Treatment Effect and Propensity Score (demeaned)].

- Linear
- Quadratic

The graph illustrates the treatment effect against the propensity score demeaned, with linear and quadratic trends indicated.
Concluding Thoughts

Background

- There is an apparent sustained increase in demand for remote learning in the K-12 sector
- Families self-selecting into remote in the “post-pandemic” landscape arguably have improved signals of their remote match quality

Findings

- The typical LAUSD family dislikes remote learning; families have heterogeneous preferences
- We find evidence of sorting on match quality, despite negative average remote learning effects

Future Research

- For what reasons do some families benefit? Bullying? Own-pace learning?
- Similar patterns in other districts?
- Formalize the link between conjoint experiments and program evaluation frameworks
Thank you!
Trends in Virtual Learning (Common Core)

Number of Schools with Supplementary Virtual

Number of Virtual Schools

Year

Exclusively Virtual

Supplementary Virtual
Trends in Virtual Learning (Common Core)
Remote Learning Market Shares by County

![Map showing market shares by county with various color codes indicating different ranges of market shares.]

- Dark red: 0.12 – 1.00
- Orange: 0.07 – 0.12
- Light orange: 0.04 – 0.07
- Yellow: 0.02 – 0.04
- Light yellow: 0.00 – 0.02
- White: No data
Distance Learning Gap

Internet and Device Adoption, from USC Annenberg Analysis

Distance Learning Gap
% K-12 Households with PC and Fixed Broadband
- > 90% - 100%
- > 80% - 90%
- > 70% - 80%
- > 60% - 70%
- > 50% - 60%
- 43% - 50%

Go Back
Survey Evidence: The Remote Experience in the Past and the Future

- Share of Respondents Agreeing
  - I Enjoyed the Remote Experience
  - I Want the District to Expand Remote Offerings
  - I am Likely to Opt for Remote Options in the Future
  - My Child Excelled in the Remote Setting Relative to In-Person
Survey Evidence: Reasons for remaining in remote
Survey Evidence: Heterogeneous valuation of academic quality

The diagram illustrates the valuation of academic quality across different groups and performance levels. The x-axis represents various categories: All, Low Achievement, Average Achievement, High Achievement, Black, Hispanic, White, Parent College, SPED. The y-axis shows the valuation in minutes, with a range from 0 to 40 minutes. The bars indicate the average valuation, with error bars showing the variability. The diagram highlights disparities in valuation across different groups and performance levels.
Survey Evidence: Families dislike remote learning
Common Support

Predicted Test Scores

Density

Conjoint

Non-conjoint
Propensity scores are forecast unbiased
Propensity scores predict actual choices

![Graph showing the relationship between propensity scores and the share of remote work.]
Identifying Assumptions for Control Function Approach

\[ E[Y_i|X_i, D_i, P(v_i)] = \alpha + X'_i\gamma + \beta D_i + \theta \lambda(v_i, X_i) + \psi \lambda(v_i, X_i) \times D_i. \] (2)

- \( \theta \) governs selection on levels
- \( \psi \) governs selection on gains; match effects

**Assumptions and implications**

- Omitted variable bias entirely due to *unobserved* preference heterogeneity
- Testable Implication:

\[ E[X_i|D_i = 1, \lambda(v_i, X_i)] - E[X_i|D_i = 0, \lambda(v_i, X_i)] = 0 \]


References II


Bruhn, Jesse (2019). “The consequences of sorting for understanding school quality”.


Einav, Liran, Amy Finkelstein, and Neale Mahoney (2022). “Producing Health: Measuring Value Added of Nursing Homes”.


Larroucau, Tomás and Ignacio Rios (2020). *Dynamic college admissions and the determinants of students’ college retention*.


