Rural-Urban Disparities in Student Achievement

Ashton Jaubert Ryan Orr Sal Patlan

Agenda

- Introduction & Motivation

- Literature Context

- Research Questions & Hypotheses

- Data & Methods

- Key Results

- Discussion & Implications

- Conclusion & Next Steps

Introduction

- Educational equity varies by geographic location. Geographic location influences access to resources such as teacher quality and support services. Learning environments and academic expectations are strongly affected by a student's setting.

- Rural areas face unique challenges, including:

- Fewer resources

- Teacher shortages

- Isolation

These challenges often stem from lower population density, leaving rural areas behind in learning and creating a persistent educational gap.

Research Objective

Using key variables such as unemployment, economic disparities, and rural locales to first find out if

-Rural locales are showing a decrease in test scores below the national average.

-How, if the previous is true, that these rural locales are being affected by economic disparities.

-Conduct further research into how and why it is common for these locales to be underdeveloped in terms of education

Literature Review

- Past research focuses heavily on urban/suburban disparities

- Recent shift: examining **rural-specific challenges** (e.g., isolation, limited teacher quality)

- **Drescher et al. (2022)** – rural subgroups show variation

- **Johnson et al. (2021)** – rural students often overlooked in growth and achievement data

Methodology

- The study uses data from the 2009–2019 Educational Opportunity Project, covering nearly all U.S. schools and districts. After merging geographic and demographic datasets, the final sample includes **600,028 observations**. This enables detailed analysis of how student achievement varies by location. **Math test scores for grades 3–8** were used as the measure of student achievement.

Key variables include math test scores (representing student achievement),

Locale Categories such as urban, suburb, town, rural,

Demographic Variables: poverty rate, unemployment, and comparison percentages of different races and ethnicities.

Dummy Variables

We create lower 25% and upper 25% for each locales, This allows us to have a better representation to see how the increase in proportions have an impact on test scores.

A single district may have several different schools that each have different classifications, for example rural may have fringe remote and distant schools. The difference could be so dramatic that you have suburban schools within city, or rural schools within city.

Dummy Variables:
urban_lower25 urban_upper25   suburb_lower25 suburb_middle25
suburb_upper25 town_lower25   town_middle25   town_upper25   r
ural_lower25   rural_middle25   rural_upper25

Regression

- **Overall Trend:** A slight negative correlation (-0.0734) exists between test scores (relative to the national average) and the proportion of students in rural schools. Simple regression confirms that a higher proportion of rural students significantly predicts lower test scores relative to the national average.
- **Specific Locale Impacts (Regression):** Rural Fringe locales show a positive association with test scores (compared to the national average).
- Rural Distant and Rural Remote locales show negative associations with test scores.
- Urban and Town locales are also associated with lower test scores relative to the national average.

cs_mn_all	Coefficient	Std. err.	t	P>|t|	[95% conf. interval]	
locale_city_large	.0584657	.0052451	11.15	0.000	.0481855	.068746
locale_city_midsize	0	(omitted)				
locale_city_small	.0966721	.0044976	21.49	0.000	.087857	.1054873
locale_suburb_large	.0843312	.0039958	21.10	0.000	.0764995	.0921629
locale_suburb_midsize	.0536275	.0047023	11.40	0.000	.044411	.0628439
locale_suburb_small	.0220498	.0049362	4.47	0.000	.0123749	.0317246
locale_town_fringe	.0286467	.0044109	6.49	0.000	.0200014	.0372919
locale_town_distant	.0602441	.0040487	14.88	0.000	.0523088	.0681794
locale_town_remote	.0551944	.00414	13.33	0.000	.0470802	.0633085
locale_rural_fringe	.0336092	.004056	8.29	0.000	.0256595	.0415589
locale_rural_distant	.0148598	.0039229	3.79	0.000	.0071711	.0225485
locale_rural_remote	.052188	.0039929	13.07	0.000	.044362	.0600139
perecd	-.9989874	.002637	-378.84	0.000	-1.004156	-.9938189
povertyall	-.4431305	.0106216	-41.72	0.000	-.4639486	-.4223125
unempall	-1.410037	.0176604	-79.84	0.000	-1.444651	-1.375423
_cons	.6282186	.0040702	154.35	0.000	.6202412	.6361961

Regression model 1

- Test
Scores=b0 +b1 ·urban_lower25+b2 ·suburb_lower25+b3 ·town_lower25+b4 ·rural_lower25+ε

```
. regress cs_mn_all urban_lower25 suburb_lower25 town_lower25 rural_lower25
note: urban_lower25 omitted because of collinearity.
```

Source	SS	df	MS
Model	4262.93805	3	1420.97935
Residual	99986.3276	600,024	.166637214
Total	104249.266	600,027	.173740958

Number of obs	= 600,028
F(3, 600024)	= 8527.38
Prob > F	= 0.0000
R-squared	= 0.0409
Adj R-squared	= 0.0409
Root MSE	= .40821

cs_mn_all	Coefficient	Std. err.	t	P>|t|	[95% conf. interval]	
urban_lower25	0	(omitted)				
suburb_lower25	-.1632192	.0013719	-118.97	0.000	-.1659081	-.1605304
town_lower25	.0417735	.0012747	32.77	0.000	.0392752	.0442719
rural_lower25	.0007327	.0012929	0.57	0.571	-.0018013	.0032667
_cons	.1115401	.0018338	60.82	0.000	.1079459	.1151343

.

Regression model 2

- Test Scores=b0 +b1 ·urban_upper25+b2 ·suburb_upper25+b3 ·town_upper25 +b4 ·rural_upper25+ε

```
. reg cs_mn_all urban_upper25 suburb_upper25 town_upper25 rural_upper25
note: urban_upper25 omitted because of collinearity.
```

Source	SS	df	MS
Model	528.736108	3	176.245369
Residual	103720.53	600,024	.172860635
Total	104249.266	600,027	.173740958

Number of obs	= 600,028
F(3, 600024)	= 1019.58
Prob > F	= 0.0000
R-squared	= 0.0051
Adj R-squared	= 0.0051
Root MSE	= .41577

cs_mn_all	Coefficient	Std. err.	t	P>|t|	[95% conf. interval]	
urban_upper25	0	(omitted)				
suburb_upper25	-.0006693	.0022429	-0.30	0.765	-.0050653	.0037266
town_upper25	-.0326516	.0027559	-11.85	0.000	-.0380531	-.0272502
rural_upper25	-.0563935	.0010828	-52.08	0.000	-.0585158	-.0542712
_cons	.0586888	.0007753	75.70	0.000	.0571692	.0602084

.

Regression model 3

Test Scores = $b_0 + b_1 \cdot perecd + b_2 \cdot povertyall + b_3 \cdot unempall + \varepsilon$

```
. reg cs_mn_all perecd povertyall unempall
```

Source	SS	df	MS		
Model	45619.7404	3	15206.5801		
Residual	55271.814	593,087	.093193434		
Total	100891.554	593,090	.170111711		

Number of obs	=	593,091
F(3, 593087)	>	99999.00
Prob > F	=	0.0000
R-squared	=	0.4522
Adj R-squared	=	0.4522
Root MSE	=	.30528

cs_mn_all	Coefficient	Std. err.	t	P>|t|	[95% conf. interval]	
perecd	-1.010941	.0026132	-386.85	0.000	-1.016063	-1.005819
povertyall	-.4570064	.0101003	-45.25	0.000	-.4768026	-.4372102
unempall	-1.307002	.0171409	-76.25	0.000	-1.340598	-1.273407
_cons	.6759685	.0011709	577.29	0.000	.6736735	.6782635

IQR Analysis

- **Achievement Patterns (IQR Analysis):** The impact of locale varies across the achievement distribution: **Rural Remote & Town Remote Areas:** Districts with the highest proportion (upper 25%) tend to have average test scores *below* the national average, suggesting challenges for higher achievers. For instance, in Rural Remote areas, the mean score for the upper 25% group was -0.07 (below national avg.), while the lower 25% was +0.047.
- **Large Suburban Areas:** Show the opposite pattern. Districts with the highest proportion (upper 25%) have average test scores *well above* the national average (e.g., +0.18 in Suburb Large), indicating these locales may better support high achievement relative to national standards.
- **Large City Areas:** Follow a pattern like rural/town remote, with the upper 25% scoring below the national average (-0.08).

Rural and Non-rural Disparities

- Student achievement varies significantly by geographic locale. Overall, a higher proportion of rural students is linked to slightly lower test scores compared to the national average (correlation = -0.0734). Regression analysis shows that Rural Fringe areas perform better than average, while Rural Distant, Rural Remote, Urban, and Town locales are associated with lower scores.

Achievement patterns differ across the distribution:

- **Rural Remote & Town Remote:** Upper 25% of districts score below the national average, indicating challenges for high achievers.

- **Large Suburban Areas:** Upper 25% score well above the national average (+0.18), suggesting strong support for high achievement.

- **Large Cities:** Upper 25% also score below the national average (-0.08).

- **Core Finding:** Geographic location strongly influences test scores, with rural and remote areas facing unique challenges, particularly for higher-achieving students.

Socioeconomic Factors & Student Achievement

Key Variables

- 👤 Economically Disadvantaged
- ⓢ Poverty Rate
- 💼 Unemployment Rate

Impact Across Distribution

Correlation with Test Scores

perecd	−0.66	−0.35
povertyall	−0.55	−0.26
unempall	−0.39	−0.08
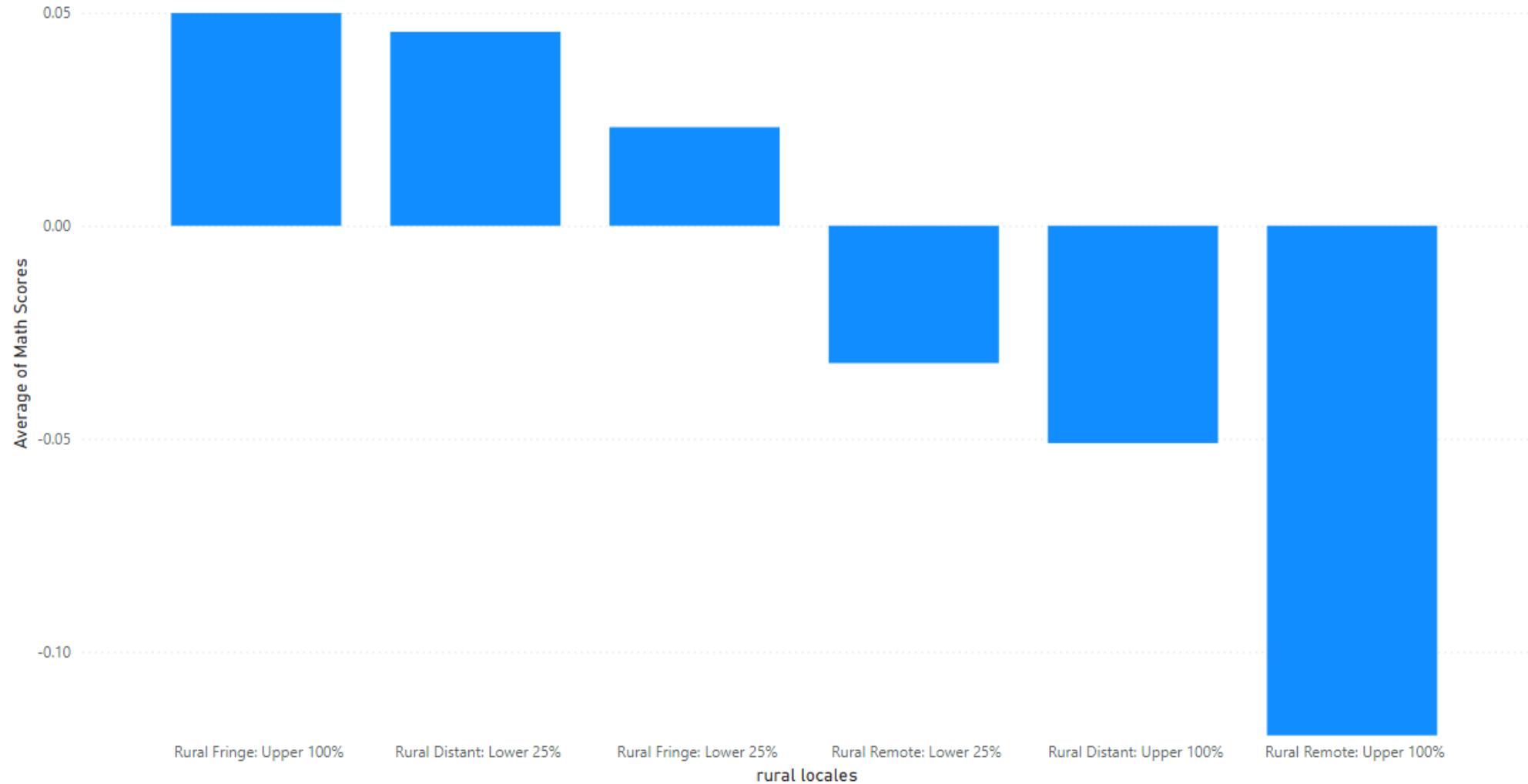
Lowest 25% vs. Highest 25%

+0.24	0.35 → 0.26
+0.35	0.26 → 0.08
+0.09	0.09 → 0.08

Conclusion: Higher poverty, unemployment, and economic disadvantage → lower achievement
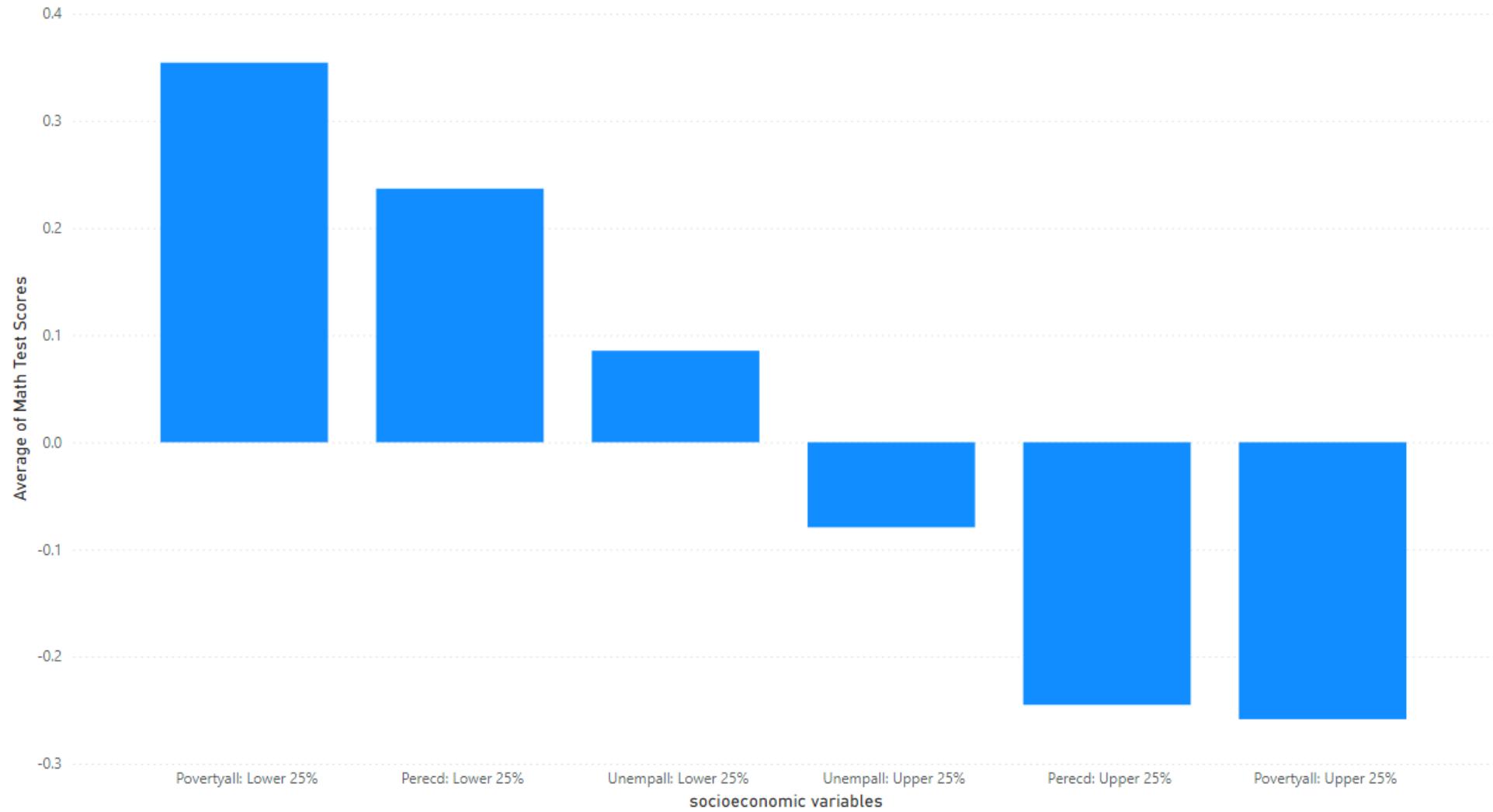
Average of Math Scores
BY RURAL LOCALES

Average of Math Test Scores

BY SOCIOECONOMIC VARIABLES

Average of Math Test Scores
BY SUBURB LARGE
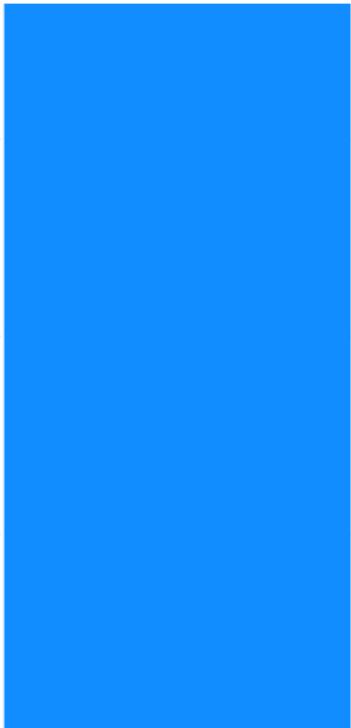
Average of Math Scores
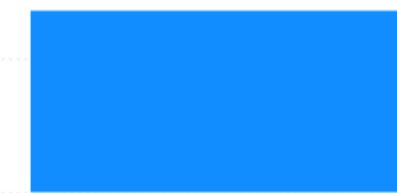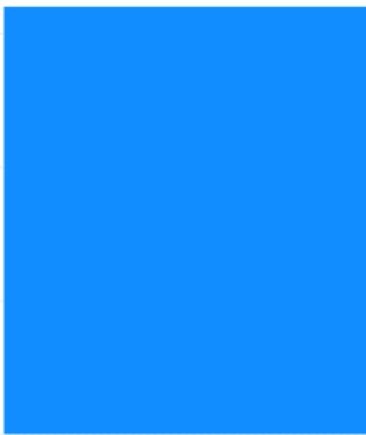
BY TOWN LOCALE

Limitations

- Only math scores (grades 3-8) from 2009–2019 were included

- Lacked variables like school funding, alumni involvement, and policy reform data

- Locale types may not reflect true differences in access or quality

- Findings are correlational, not causal

- Socioeconomic proxies do not capture all structural inequities

Findings/Discussion

- **Observation:** Rural locales are negatively associated with test scores compared to the national average.

- **Nuanced Finding (IQR):** Rural remote, town remote, and large city areas show pronounced negative impact for higher-achieving students (upper 25% score below national average).

- **Contrast:** Suburban areas tend to support higher achievement relative to the national average.

- **Possible Explanations:** Disparities may result from differences in resources, teacher expertise, funding, and access to advanced courses.

- **Implications:** Tailored interventions and policies are needed to improve equity, especially in rural remote areas.

- **Future Research:** Further study on resources, support structures, and long-term impacts is recommended.

References

- Drescher, J., Podolsky, A., Reardon, S. F., & Torrance, G. (2022). The geography of rural educational opportunity. *RSF: The Russell Sage Foundation Journal of the Social Sciences, 8*(3), 123-149.
- Johnson, A., Kuhfeld, M., & Soland, J. (2021). The Forgotten 20%: Achievement and Growth in Rural Schools Across the Nation. *AERA Open, 7*(1), 1-17.
- Reardon, S. F., Ho, A. D., Shear, B. R., Fahle, E. M., Kalogrides, D., & Saliba, J. (2024). Stanford Education Data Archive (Version 5.0).