Digital Connectivity and Standard of Living

Team 5: Dominique Rever, Faraz Khojasteh Far, Justin Siegel, Jonathan Meystrik, Braden Thorne, Dhruv Modi

1. Introduction

Broadband and internet access are widely recognized as drivers of socioeconomic progress. Yet, access remains uneven, widening the digital divide and exacerbating global disparities in education, opportunity, and quality of life. This project investigates whether greater digital connectivity reduces poverty and improves literacy using a global World Bank panel dataset (2013–2023), supplemented with historical telephone line data for causal inference. We go beyond simple correlations by applying a multifaceted approach: econometric models (OLS, Fixed Effects, SUR, Quantile Regression, 2SLS), machine learning (Random Forests), and Structural Equation Modeling (SEM), supported by imputation and interactive visualizations. We aim to provide policymakers and development stakeholders with robust, actionable insights into broadband's role in advancing human development.

2. Problem Definition

Jargon-Free: Does better broadband access reduce poverty and improve literacy worldwide, particularly in regions with limited infrastructure? We seek to understand whether investments in internet access lead to meaningful improvements in people's lives. This helps inform global efforts to close the digital divide and promote more inclusive development.

Formal Definition: We estimate the causal effect (β) of broadband penetration (B_{it}) on standard-of-living outcomes (Y_{it}), such as poverty and literacy rates for country *i* at time *t*. Our panel model controls for covariates (X_{it}), country-specific fixed effects (α_i), and time fixed effects (δ_t): $\boldsymbol{Y_{it} = \beta B_{it} + \gamma X_{it} + \alpha_i + \delta_t + \epsilon_{it}}$**.** We also explore nonlinearities, heterogeneous impacts (via quantile regression), and mediated pathways using Structural Equation Modeling (SEM). To address potential endogeneity, we implement two-stage least squares (2SLS) with historical fixed-line phone penetration as an instrument. We aim to provide policymakers and development stakeholders with robust, actionable insights.

3. Literature Survey

- *Czernich et al. (2011).* Broadband improves GDP in OECD countries. Reveals connectivity as a driver of prosperity. Focuses on developed economies; we will extend to emerging markets.
- *Atasoy (2013).* Broadband expansion improves employment and wages. Adds a labor market perspective to life quality metrics. Lacks broad well-being indicators that we will integrate.
- *Qiang et al. (2009).* Broadband stimulates economic growth across income levels. Useful for assessing cross-country variations. Uses older data; we will update with recent statistics.
- *Kolko (2012).* Broadband correlates with regional employment and income growth. Micro-level study supports regional analysis. Less applicable to low-income countries; we will adapt globally.
- *Greenstein & McDevitt (2009).* Broadband contributes to new GDP growth in the U.S., quantifying direct and indirect economic impacts. Focuses on U.S. data; we will generalize findings.
- *OECD (2014).* Comprehensive framework for measuring the digital economy. Highlights current measurement gaps and trends. Broad scope lacks country-specific focus; we will tailor analyses.
- *Chinn & Fairlie (2007).* Internet access accelerates GDP growth across countries. Provides empirical evidence for broadband's economic impact. Uses national data, missing local effects.
- *Norris (2001).* Digital divide affects civic engagement and economic equality. Shows non-fiscal impacts of broadband. Focuses on U.S. and qualitative aspects; we will add global data.
- *Minges (2015).* Internet access correlates with economic growth globally. Supports empirical case for broadband-driven development. Provides general data; we will add interactive visualizations.

- *Deloitte (2021)*. Broadband investment drives economic competitiveness. Recent (post-2020) perspective. U.S.-centric; we will expand to global contexts.
- *Katz & Koutroumpis (2012).* Broadband fuels economic growth in Senegal. Provides a case study of broadband's impact. Findings may not generalize beyond Senegal.
- *Isley & Low (2022).* Broadband adoption supports rural employment. Demonstrates connectivity's role in reducing urban-rural disparity. Limited to rural U.S.; we will generalize findings.
- *Shah & Jimenez-Duran (2023).* Broadband improves student test scores. Connects internet access to educational outcomes. Focuses on standardized tests; we will explore broader impacts.
- *Koutroumpis & Sarri (2023).* Broadband has greater economic benefits in developing regions. Highlights the potential for poverty alleviation. Findings may not apply to well-connected regions.
- *Tchamyou (2018).* Digital divide exacerbates educational inequality. Demonstrates broadband's role in addressing socioeconomic gaps. Focuses on Africa; we will expand to other regions.
- *Arvin & Pradhan (2014).* Broadband penetration correlates with GDP growth. Shows broadband as a catalyst for innovation. Use macro-level indicators; we will add local-level data.
- *Falch & Henten (2018).* Compares broadband policies across countries. Provides insight into regulatory strategies. Lacks social impact analysis; we will connect policies to outcomes.
- *Nishijima et al. (2017).* The digital divide in Brazil is driven by income and location. Examine barriers to broadband adoption in developing regions. Does not connect to education or poverty outcomes; we will bridge this gap.

While prior research often relies on older data, limited geographic scope, or narrow GDP-focused outcomes, our project addresses these gaps through recent global data, advanced causal methods to reduce endogeneity, and a broader emphasis on fundamental quality-of-life indicators such as poverty and literacy.

4. Proposed Method

- **Intuition**: Why Our Approach is Well-Suited

Broader Scope: While most studies focus on GDP or developed nations, we examine poverty and adult literacy across a diverse global panel (2013–2023), capturing both economic and social dimensions of development, especially in underrepresented regions.
Integrated Framework: We combine econometric models (Fixed Effects, 2SLS), machine learning (Random Forests), and Structural Equation Modeling (SEM) to assess average, nonlinear, and indirect effects within a unified framework, offering a more realistic picture than descriptive methods alone.
Robust Data Strategy: We apply a two-stage imputation method—time-series interpolation and group-wise median filling by development status—to address missing data. To mitigate endogeneity, we use historical fixed-line telephone penetration (2005) as an instrument.
Actionable Insights: Quantile regression and Random Forests uncover distributional and nonlinear effects, while SEM clarifies indirect pathways (e.g., via literacy). Interactive visualizations enhance accessibility for both academic and policy audiences.

- **Detailed Description:**

Data Construction and Preprocessing: We built a panel dataset using World Bank indicators for 217 countries (2013–2023). After removing aggregates and non-country entries, we applied two-stage imputations: (1) time-series interpolation within countries, and (2) median imputation by development status (using a $20,000 GDP per capita threshold). GDP per capita was log-transformed for interpretability, and the dataset was indexed by country and year to support cross-sectional and longitudinal analysis.

  - **Baseline Analytics:**

Descriptive Statistics and Visualization: We began with exploratory analysis to assess distributions, correlations, and nonlinearities. Fixed broadband averaged 14.98 subscriptions per 100 people, but its highly skewed distribution (Figure 1) highlights global disparities. Literacy was high overall (mean = 87.1%), while poverty showed broader dispersion (mean = 24.4%). (Figure 2) Showed broadband positively correlated with literacy ($r = 0.56$) and negatively with poverty ($r = –0.53$), along with strong ties to internet usage, urbanization, and life expectancy, raising multicollinearity concerns. Scatter plots offered deeper insights. Broadband and internet usage (Figure 5) showed saturation at high access levels. Broadband's link to literacy (Figure 6) was weaker and clustered near the upper bound. The relationship with poverty (Figure 4) was nonlinear and more varied, particularly at higher access levels. (Figure 3) shows uneven global progress in broadband and development outcomes, while the 2023 map (Figure 3) reveals persistent digital inequality. These patterns informed our choice of panel regressions, quantile models, and machine learning approaches.

OLS Regression: We implemented OLS models as a baseline to estimate average effects, assuming linearity and exogeneity. Separate regressions were run for poverty and literacy, using broadband as the key independent variable, and controlling for internet usage, log-transformed GDP per capita, and urban population share. The poverty model showed a statistically insignificant effect of broadband (Coef = 0.0399, $p = 0.110$), while the literacy model revealed a significant negative association (Coef = –0.1822, $p < 0.001$). These results contradict theoretical expectations and highlight the limitations of pooled cross-sectional models in capturing dynamic or structural relationships in development outcomes.

- **Econometric Models:**

Fixed Effects Panel Model: To isolate causal relationships, we used a two-way fixed effects model via PanelOLS, controlling for time-invariant country traits (e.g., governance, geography) and global shocks. Estimation used robust standard errors. For poverty, broadband had a significant negative effect (Coef = –0.1270, $p < 0.001$), suggesting increased access within countries reduces poverty. Its effect on literacy was insignificant (Coef = –0.0251, $p = 0.201$), indicating no measurable change. These results highlight broadband's role in poverty reduction, even when education outcomes remain stable.

Seemingly Unrelated Regression (SUR): We estimated a SUR model to jointly examine poverty and literacy while accounting for correlated error terms across equations, especially relevant when shared unobservables (e.g., infrastructure) may influence both outcomes. Using a system OLS approach, broadband showed a marginally significant positive association with poverty (Coef = 0.0399, $p \approx 0.054$) and a consistent negative effect on literacy. These estimates diverge from the FE model, suggesting potential bias from unmodeled country-specific characteristics. SUR serves as a useful diagnostic but lacks the causal interpretability of panel methods.

Two-Stage Least Squares (2SLS): To address potential endogeneity in broadband adoption, we used historicalphonesper100_2005 as an instrument in a 2SLS model. This variable proxies for historical infrastructure conditions, assumed to influence current broadband but not poverty directly. The instrument showed strong first-stage relevance (F-stat > 10). In the second stage, broadband had a positive but only marginally significant effect on poverty (Coef = 0.0361, $p = 0.088$). This contradicts the FE results and underscores IV models' sensitivity to instrument validity, particularly in contexts with overlapping infrastructure and development drivers.

Quantile Regression: To capture heterogeneous effects, we applied quantile regression at the 20th, 50th, and 80th percentiles of poverty distribution. This method avoids OLS's constant-variance assumption and allows richer distributional insights. Results showed a significant positive coefficient at lower and median poverty levels (e.g., Coef = 0.0406 at the median, $p = 0.017$), but an insignificant effect on higher poverty quantiles. These unexpected findings suggest broadband access may coincide with other poverty-reducing institutions in moderately poor settings, while offering limited benefit in severely impoverished contexts.

- **Machine Learning Model:**

Random Forest Regression: To capture nonlinearities and interactions, we trained a Random Forest Regressor using an 80/20 train-test split. The model achieved a strong out-of-sample R^2 of 0.896, indicating high predictive accuracy. However, broadband ranked only fifth in feature importance (0.06), well behind life expectancy (0.47), literacy (0.19), urbanization (0.11), and GDP per capita (0.09). These results suggest that while broadband contributes to predictive performance, its impact on poverty is modest compared to more structural development indicators.

- **Structural Modeling:**

Structural Equation Modeling (SEM): We used SEMOPY to estimate a Structural Equation Model examining whether broadband reduces poverty indirectly through education. SEM enables simultaneous estimation of direct and mediated effects. Results showed a significant direct negative effect of broadband on poverty (Coef = –0.102, $p < 0.05$) and a strong negative effect of literacy on poverty (Coef = –0.456, $p < 0.001$). However, broadband's effect on literacy was only marginally significant (Coef = 0.055, $p \approx 0.055$), suggesting that its poverty-reducing impact is only partially mediated through education.

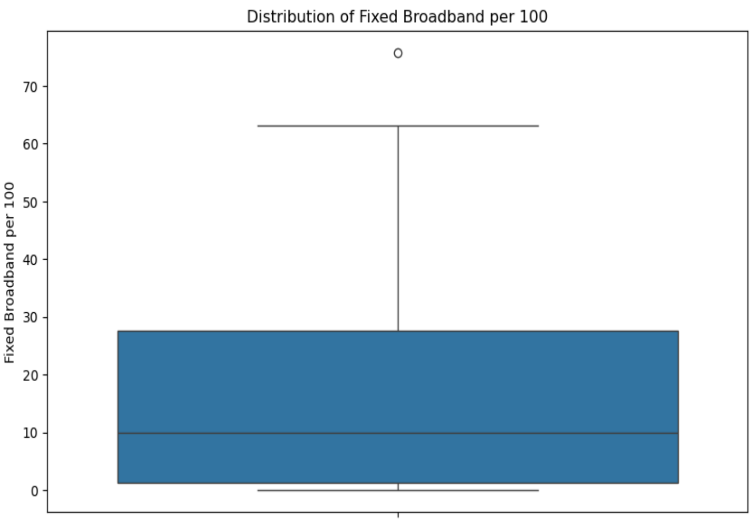

Figure 1: Skewed Distribution of Fixed Broadband per 100

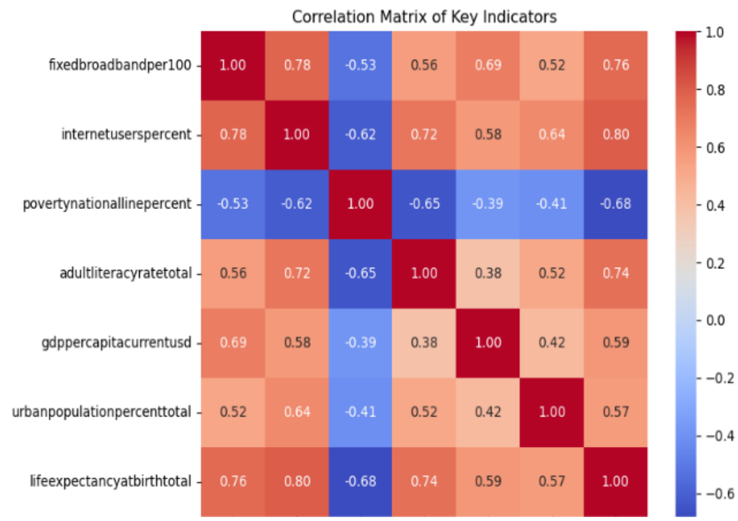

Figure 2: Correlation Matrix of Key Indicators

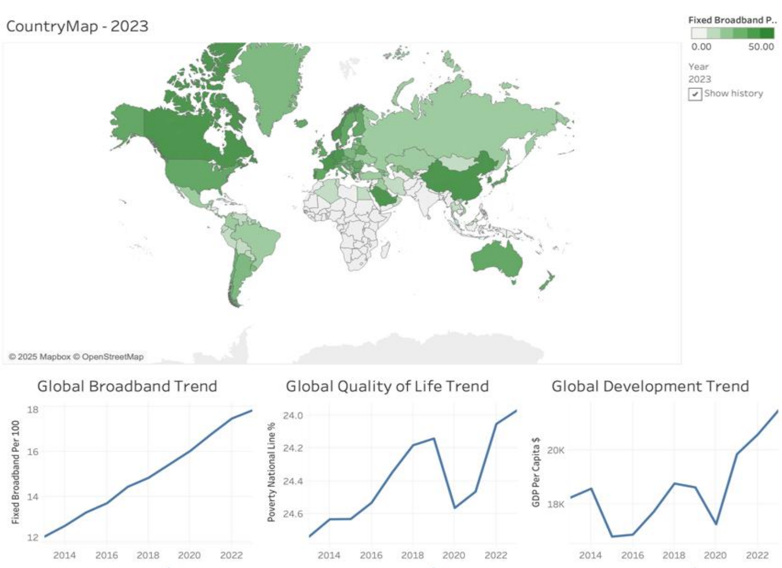

Figure 3: 2023 World Choropleth map of Broadband connection per country. Avg. Global Broadband, Quality of Life and Global Development trends.

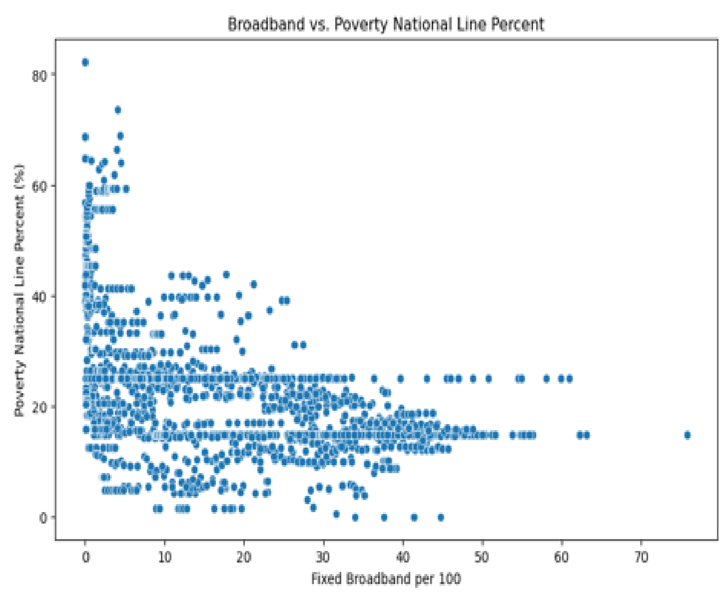

Figure 4: Scatterplot of Broadband vs. National Poverty Line

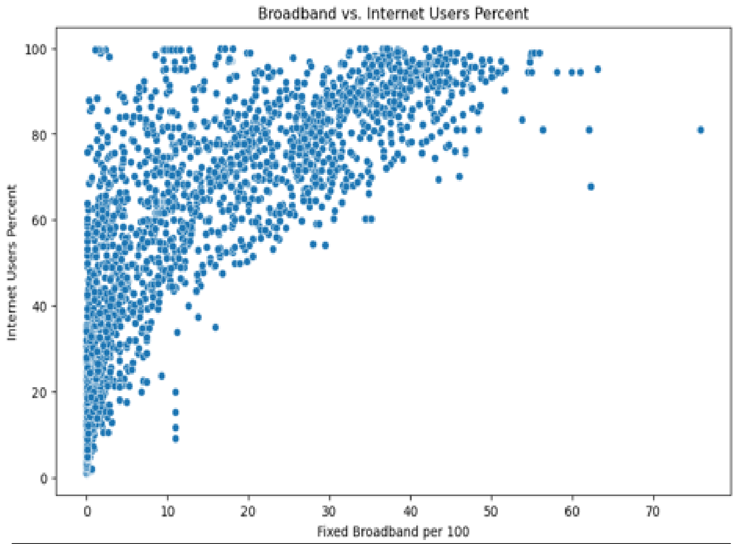

Figure 5: Scatterplot of Broadband vs. Internet Users

Figure 6: Scatterplot of Broadband vs. Adult Literacy Rate

5. Evaluation

- **Experimental Testbed and Evaluation Questions:**

Our testbed is a panel dataset of 217 countries (2013–2023) from World Bank indicators. After cleaning and development-status-aware imputation, it includes 2,387 country-year observations covering broadband, poverty, literacy, GDP, urbanization, internet use, and related metrics. The panel structure supports both cross-sectional and longitudinal modeling. To evaluate our methods and broadband's development role, we designed the following experiments:

  - **Model Comparison for Predictive Accuracy**: Do broadband indicators improve predictions of poverty and literacy relative to traditional development metrics?
  - **Panel Data Analysis for Within-Country Effects**: Does increased broadband access within a country over time reduce poverty or improve literacy?
  - **Distributional and Nonlinear Insights**: Are broadband's effects consistent across the poverty distribution, or do they exhibit thresholds or diminishing returns?
  - **Dashboard Usability:** Are our visualizations intuitive and actionable for non-technical audiences?

- **Detailed Description of the Experiments and Observations:**
  - **Model Comparison for Predictive Accuracy:** We tested whether broadband access improves the prediction of poverty and literacy beyond traditional development indicators. OLS and Random Forest models were trained using GDP, urbanization, and internet usage, with and without broadband. In OLS models, broadband added minimal value: poverty R^2 rose slightly (0.461 → 0.462), and literacy R^2 improved marginally (0.568 → 0.578). In contrast, Random Forests showed stronger gains: poverty R^2 increased from 0.736 to 0.820, and literacy from 0.809 to 0.862. These results suggest broadband's predictive power is more effectively captured in nonlinear models with interaction effects.
  - **Panel Data Analysis for Within-Country Effects:** Fixed effects panel models assessed whether broadband expansion over time reduces poverty or improves literacy, controlling for country- and year-specific confounders. Broadband had a significant negative effect on poverty (Coef = –0.127, $p < 0.001$), indicating that increased access is associated with lower poverty within countries. However, its effect on literacy was statistically insignificant (Coef = –0.025, $p = 0.201$), suggesting limited direct influence on national literacy rates. These findings support broadband's potential role in poverty reduction, while highlighting that education impacts may require broader systemic interventions.
  - **Distributional and Nonlinear Insights:** To examine variation in broadband's effects across poverty levels, we applied quantile regression and Random Forests. Quantile regression showed a positive, significant association at the 20th and 50th percentiles, but no effect at the 80th. This suggests broadband may be more impactful in moderately poor countries and less effective where

poverty is extreme. Random Forests ranked broadband below life expectancy, literacy, urbanization, and GDP, but above internet usage, highlighting its intermediate importance. These results indicate broadband's development effects are not uniform and may depend on broader structural conditions and baseline levels of poverty.

- **Usability Testing of Visual Dashboards:** We built an interactive dashboard with a global choropleth map shaded by 2023 broadband subscriptions per 100 people. Hovering over a country reveals 2013–2023 trends in broadband, poverty, and GDP per capita. Users found the dashboard intuitive and informative, identifying low-access regions and linking them to development challenges. Feedback included suggestions for improved country research, annotated outliers (e.g., high broadband but high poverty), and added literacy data. Overall, the dashboard effectively conveyed key patterns and supported our findings, enhancing accessibility for non-technical audiences.

6. Conclusions and Discussion

Key Findings

- **Within-Country Poverty Reduction**: Fixed effects models show that broadband expansion is significantly associated with lower poverty over time (Coef = –0.1270, $p < 0.001$), supporting its role in poverty alleviation.
- **Conflicting Causal Estimates**: 2SLS using historical phone lines yields a positive and significant coefficient (Coef = +0.1113, $p = 0.0001$), contradicting FE results and suggesting possible instrument validity issues.
- **Predictive Value**: In Random Forests, broadband contributed modestly to poverty prediction but ranked below life expectancy, literacy, and GDP.
- **Mixed Model Evidence**: OLS, SUR, and quantile regression produced inconsistent results; SEM suggested an indirect path from broadband to poverty via literacy, but broadband's effect on literacy was weak.

Overall Conclusion

Our analysis suggests broadband access can support poverty reduction within countries, though its impact is often indirect and varies by context. Fixed effects models show strong potential, while results across predictive and causal methods were mixed. Literacy outcomes appeared less responsive, underscoring the need for complementary investments in education, affordability, and governance. Broadband is a powerful but insufficient tool on its own, requiring integrated, cross-sectoral development strategies.

- **Limitations -** Our analysis faces several constraints:
- **Endogeneity concerns** remain despite attempts to address them via instrumental variables.
- **Instrumental variable validity** is questionable due to possible correlation with unobserved factors.
- **Data aggregation and access metrics** may mask within-country inequalities in usage or quality.
- **Causal inference limitations** restrict the strength of policy recommendations.

Implications and Future Work

- **Policy implications**: The FE results offer cautious support for investing in broadband infrastructure to reduce poverty. However, the lack of consistent evidence across models, particularly the 2SLS findings, suggests that broadband access alone may not be sufficient. Greater attention should be placed on complementary policies promoting affordability, digital skills, and effective usage.
- **Research directions**: Future work should explore better instruments for broadband access, incorporate dynamic panels or sub-national data, and distinguish between mobile and fixed-line broadband. Further refinement of the SEM framework could clarify indirect pathways through education, health, and employment.

Effort Distribution Statement

All team members have contributed a similar amount of effort across all phases of the project, including literature review, data cleaning, modeling, analysis, visualization, and report writing.

References

1. Arvin, M. B., & Pradhan, R. P. (2014). Broadband penetration and economic growth nexus: Evidence from cross-country panel data. *Applied Economics*, 46(35), 4360–4369. https://doi.org/10.1080/00036846.2014.957444
2. Atasoy, H. (2013). The effects of broadband Internet expansion on labor market outcomes. *ILR Review*, 66(2), 315–345.
3. Chinn, M. D., & Fairlie, R. W. (2007). Does the Internet Increase Growth? Evidence from Cross-Country Panel Data. *Journal of Development Economics*, 83(2), 271–289. https://doi.org/10.1016/j.jdeveco.2006.01.003 (*Adjust number if combining lists changes order*)
4. Czernich, N., Falck, O., Kretschmer, T., & Woessmann, L. (2011). Broadband infrastructure and economic growth. *The Economic Journal*, 121(552), 505-532. https://doi.org/10.1111/j.1468-0297.2011.02420.x
5. Deloitte. (2021). *Broadband for All: Charting a Path to Economic Growth*. Deloitte Insights. https://www2.deloitte.com/content/dam/Deloitte/us/Documents/process-and-operations/us-charting-a-path-to-economic-growth.pdf
6. Falch, M., & Henten, A. (2018). Dimensions of broadband policies and developments. *Telecommunications Policy*, 42(9), 715–725. https://doi.org/10.1016/j.telpol.2017.11.004
7. Greenstein, S., & McDevitt, R. C. (2009). *The Broadband Bonus: Accounting for Broadband Internet's Impact on U.S. GDP*. Working Paper, Kellogg School of Management, Northwestern University. https://www.itu.int/net/wsis/stocktaking/docs/activities/1287047007/greenstein_broadband_bonus.pdf
8. Isley, C. and Low, S. (2022). Broadband adoption and availability: impacts on rural employment during covid-19. *Telecommunications Policy*, 46(7), 102310. https://doi.org/10.1016/j.telpol.2022.102310
9. Katz, R. L., & Koutroumpis, P. (2012). The Economic Impact of Telecommunications in Senegal. *Communications & Strategies*, 86(2), 21-42. https://ssrn.com/abstract=2270502
10. Kolko, J. (2012). Broadband and local growth. *Journal of Urban Economics*, 71(1), 100–113. https://doi.org/10.1016/j.jue.2011.07.004
11. Koutroumpis, P. and Sarri, D. (2023). The economic impact of broadband access for small firms. *World Economy*, 47(4), 1642-1681. https://doi.org/10.1111/twec.13485
12. Minges, M. (2015). *Exploring the Relationship between Broadband and Economic Growth*. World Development Report 2016 Digital Dividends Background Paper. Washington, DC: World Bank. https://documents1.worldbank.org/curated/zh/178701467988875888/pdf/102955-WP-Box394845B-PUBLIC-WDR16-BP-Exploring-the-Relationship-between-Broadband-and-Economic-Growth-Minges.pdf
13. National Bureau of Economic Research (NBER). (2024). *Broadband Internet Access, Economic Growth, and Wellbeing*. NBER Working Paper No.

32517. https://www.nber.org/papers/w32517 (*Note: Check if this NBER paper was actually used or just from the proposal's list*)

14. Nishijima, M., Ivanauskas, T. M., & Sarti, F. M. (2017). Evolution and determinants of digital divide in Brazil (2005–2013). *Telecommunications Policy*, 41(1), 12–24. https://doi.org/10.1016/j.telpol.2016.10.004
15. Norris, P. (2001). *Digital Divide: Civic Engagement, Information Poverty, and the Internet Worldwide*. Cambridge University Press.
16. OECD. (2014). *Measuring the Digital Economy: A New Perspective*. OECD Publishing. http://dx.doi.org/10.1787/9789264221796-en
17. Qiang, C. Z.-W., Rossotto, C. M., & Kimura, K. (2009). Economic impacts of broadband. In *Information and Communications for Development 2009: Extending Reach and Increasing Impact* (pp. 35–50). World Bank. https://doi.org/10.1596/978-0-8213-7605-8_ch02
18. Shah, A. and Jimenez-Duran, R. (2023). Broadband access and standardized test scores: a causal analysis. *Journal of Student Research*, 12(4). https://doi.org/10.47611/jsrhs.v12i4.5449
19. Tchamyou, V. (2018). Education, lifelong learning, inequality, and financial access: evidence from African countries. *Contemporary Social Science*, 15(1), 7-25. https://doi.org/10.1080/21582041.2018.1433314
20. Whitacre, B. E., Gallardo, R., & Strover, S. (2014). Broadband's contribution to economic growth in rural areas: Moving towards a causal relationship. *Telecommunications Policy*, 38(11), 1011–1023. https://doi.org/10.1016/j.telpol.2014.05.005