Buffering as a Manifestation of Inequality: Impact of Bandwidth Quality on Educational Outcomes

Mukund Sundararajan

University of British Columbia

ECON 490: Seminar in Applied Economics

Dr. Jonathan Graves

December 13, 2021

Abstract

As students This paper examines the impact of differential access to high quality broadband on educational outcomes for secondary school students in British Columbia, Canada. By linking broadband availability information published

Buffering as a Manifestation of Inequality: Impact of Bandwidth Quality on Educational Outcomes

Introduction

In a world stricken by the COVID-19 pandemic, people and institutions have been forced to transition to working, studying, and recreating from home. While this change has been relatively easy to make for some, the gap between those who were able to transition relatively smoothly to online-oriented life and those who experienced severe disruptions often falls along the same lines as the pre-existing digital divide. The digital divide, in its simplest form, is the stratification of society caused by the differential levels of access and use of the internet across different populations. A dimension of this broader issue is the growing disparity between students who have access to high quality internet and those who do not.

With research demonstrating the benefits of online homework platforms (TODO: Cite), teachers are transitioning to online learning tools (TODO: Cite), which only exacerbates this issue. Even prior to the COVID-19 pandemic that forced all students online, the gap in access was a serious issue. [TODO: papers]. While there is some early evidence that the pandemic has alleviated certain areas of the digital divide, the educational gap remains [TODO: Cite].

While much of the literature today agrees upon the existence of such a gap and its impact on student performance and long term outcomes, it tends to be predicated on availability, rather than quality. It is easy to imagine students with low-bandwidth connections taking longer to complete the same tasks as their peers with better quality connections. For example, they cannot access video tutorials or stream their classes without constant stoppages, which also causes disengagement. To address this, this paper will specifically explore the impact of connection quality on student performance. Since broadband expansion is a multi-year project requiring significant public and private investment into infrastructure, identifying areas that would benefit the most from increased attention is highly essential.

The analysis sample is constructed from multiple different datasets published by the Government of Canada, Government of British Columbia (BC), and the Fraser Institute, restricted to secondary school students in the province of British Columbia, Canada.
TODO: Results

TODO: Paper Structure

Existing Literature

Studying the effects of broadband proliferation on key social indicators is not a novel idea. Since its introduction, the internet has fundamentally changed how we live, and its effects can be seen through many metrics. For example, a simple search pulls up studies inspecting the effect of the internet on economic growth (Choi & Hoon Yi, 2009), international trade (Lin, 2015), and interest rates (Luo et al., 2018), among many others. Unsurprisingly then, the effect of internet usage on student performance has been studied extensively.

Studies in this area primarily focus on the relationship between internet usage patterns and academic performance. In their paper, Austin and Totaro (2011) establish that high school students who are exposed to moderate internet use at home outperform their peers who either don't have such access or show intense internet usage at a statistically significant level. Along similar lines, Xu et al. (2019) and Siraj et al. (2015) show that frequent internet usage is positively associated with undergraduate student academic performance. This gap is particularly stark when compared with students who have no access or access only via a cell phone due to living in rural areas. Such students are likely to perform worse not only on exams but also on homework completion and are less likely to attend university Hampton et al. (2020). Thus it is clear that access and moderate use of the internet has some impact on student performance.

The importance of providing high quality connections is backed up by policy objectives from the Government of Canada. The CRTC has a major objective to ensure that all Canadians have access to high quality mobile and broadband internet (Radio-Television, Commission, et al., 2020), supported by the Universal Broadband Fund for a total of $2.75 billion in funding. This is adding to existing provincial initiatives, leading to much improved connectivity in rural areas (Rajabiun & Middleton, 2013).

Data

Datasets

Broadband information for BC is a subset of the National Broadband Data published by Innovation, Science, and Economic Development Canada under Canada's Open Government Initiative (Winters & Lacharité, 2021). This dataset aggregates current bandwidth availability information over a hexagonal grid of Canada, where each hexagon is approximately 25 squared kilometres. The bandwidth availability consists of boolean markers for each speed category: (notated in download/upload megabits per second) <5/1, 5/1, 10/2, 25/5, and 50/10; technology type: wired, wireless, and combined; and LTE availability. A hexagon is considered to have a certain speed available to them if over 75% of the population in that hexagon have access to that speed. Since I'm not examining the availability by technology type, I'm discarding the maximum speeds per medium and focusing solely on the maximum combined speed that is available in each hexagon. Each speed category was turned into a dummy variable, with value 1 if the hexagon had that speed available as its maximum. LTE availability is also ignored as LTE is unlikely to be used for educational content consumption due to low data caps. This dataset is cross-referenced with Pseudo-Household Demographic Distribution Dataset, published by the same organisation, to retrieve population counts for each hexagon (Winters & Pilon, 2020).

Information regarding secondary school performance is scraped from an annual report cataloguing various variables of the top 250 ranked schools in BC, published by the

Fraser Institute (Emes & Cowley, 2020). From this datset, we can retrieve the average exam marks from 2015 to 2019, % of students who have special needs, % of students who are ESL speakers, as well as whether the school is public or independent. It is important to note that the average exam mark is for mandatory exams in language arts (Communications 12, English 12, English 12 First Peoples, and Français langue première 12) for 12th grade students, as BC has eliminated other province-wide exams. This is merged with a dataset from Education Analytics and published by the BC Provincial Government that contains all BC schools with indicators for French programs ("BC Schools - K-12 with Francophone Indicators," 2020), but importantly contains the geographic co-ordinates for each school.

Since income is likely to be a significant determinant in educational outcomes, median household income from the 2016 Census conducted by Statistics Canada is used ("2016 Census," 2017). The total median household income across all family types for every census subdivision, which in BC corresponds roughly a municipality, is used to determine the income level at each school.

Analysis Sample

As there were some differences between the BC government's dataset on schools and the Fraser Institute report, any conflicting or missing records were dropped, and duplicate records that were created due to multiple schools with the same name were manually removed after confirming the addresses. Following this procedure, the final analysis sample contains academic performance as well as school level information for 235 schools in BC. Since the broadband data is accurate up to March 2021 and no historical data is available, only the latest school year from the school report, 2018-2019, is used. While this may cause some slight inconsistency, due to the Income is assigned by determining which census subdivision the school lies in, and then assigned the median household income in that census subdivsion. Since the subdivisons are not generated to a standard size or population, they can vary dramatically as they reflect local boundaries. Instead of directly using the median income, I use the log in order to scale down the values. Table 1 contains summary statistics for these school level variables.

Table 1

Summary Statistics for School Data in 2019 with Log Median Income in 2016

	Count	Mean	Std Dev	Min	25%	50%	75%	Max
AVERAGE_EXAM_MARK	235.000	69.148	4.838	54.200	66.300	68.600	71.650	85.400
PERCENT_ESL	235.000	3.380	4.283	0.000	0.150	1.900	5.100	21.400
PERCENT_SPECIAL_NEEDS	235.000	11.887	4.780	0.000	8.900	11.500	14.650	29.100
type_Public	235.000	0.843	0.365	0.000	1.000	1.000	1.000	1.000
LOG_MEDIAN_INCOME	235.000	11.186	0.176	10.672	11.086	11.191	11.290	11.612

In order to assign the hexagons, distance from the centre of every hexagon to every school was computed, and each populated hexagon was assigned its closest school. If the distance between the hexagon's centre and the school was over 50 kilometres, the hexagon was dropped. This was done to avoid over-assignment due to missing data: since the school report only contains data for the top 250 schools, it's possible that students in these far away regions are either being home-schooled due to the distance, or go to a local school that does not show up in the top 250 list. Figure [TODO] illustrates the distribution of speeds geographically, while Table [TODO] presents the summary statistics for the same.

Methodology and Results

Initial Model

To identify the impact on exam marks (AVERAGE_EXAM_MARK), I use the following regression specification:

$$\begin{aligned}\text{AVERAGE\_EXAM\_MARK}_i = \beta_0 &+ \beta_1(\text{COMBINED\_50\_10}_i) + \beta_2(\text{COMBINED\_25\_5}_i) + \\ &\beta_3(\text{COMBINED\_10\_2}_i) + \beta_4(\text{COMBINED\_5\_1}_i) + \\ &\beta_5(\text{type\_PUBLIC}_i) + \beta_6(\text{PERCENT\_ESL}_i) + \\ &\beta_7(\text{PERCENT\_SPECIAL\_NEEDS}_i) + \\ &\beta_8(\log(\text{MEDIAN\_INCOME}_i)) + \epsilon_i\end{aligned}$$

Each explanatory variable is associated to a school i, while β_0 is the intercept and ϵ_i is the error term. The bandwidth variables, which in the original dataset were categorical variables, are transformed into proportions due to the aggregation process that linked each hexagon to a school; each school has a percentage of assigned areas that have combined speeds of 50/10, 25/5, 10/2, and 5/1. The COMBINED_LT_5_1 was omitted to avoid multicollinearity. Conducting tests for heteroskedasticity and

References

2016 census (98-400-X2016099). (2017, September). Statistics Canada.

Austin, W., & Totaro, M. W. (2011). High school students' academic performance and internet usage. *Journal of economics and economic education research*, *12*(1), 41.

Bc schools - k-12 with francophone indicators. (2020). Education Analytics. https://catalogue.data.gov.bc.ca/dataset/bc-schools-k-12-with-francophone-indicators

Choi, C., & Hoon Yi, M. (2009). The effect of the internet on economic growth: Evidence from cross-country panel data. *Economics Letters*, *105*(1), 39–41. https://doi.org/https://doi.org/10.1016/j.econlet.2009.03.028

Emes, J., & Cowley, P. (2020). *Report card on british columbia's secondary schools 2020* (Technical Report). Fraser Institute.

Hampton, K., Fernandez, L., Robertson, C., & Bauer, J. M. (2020). Broadband and student performance gaps.

Lin, F. (2015). Estimating the effect of the internet on international trade. *The Journal of International Trade & Economic Development*, *24*(3), https://doi.org/10.1080/09638199.2014.881906, 409–428. https://doi.org/10.1080/09638199.2014.881906

Luo, C., Li, M., Peng, P., & Fan, S. (2018). How does internet finance influence the interest rate? evidence from chinese financial markets. *Dutch Journal of Finance and Management*, *2*(1).

Radio-Television, C., Commission, T. Et al. (2020). Broadband fund: Closing the digital divide in canada. *Consumer Information, April*, *27*.

Rajabiun, R., & Middleton, C. (2013). Rural broadband development in canada's provinces: An overview of policy approaches. *Journal of Rural and Community Development*, *8*(2).

Siraj, H. H., Salam, A., Hasan, N., Jin, T. H., Roslan, R. B., & Othman, M. N. B. (2015). Internet usage and academic performance: A study in a malaysian public university. *International Medical Journal*, *22*(2), 83–86.

Winters, D., & Lacharité, Y. (2021). *National broadband data* (NBD.7.2). Government of Canada. https://open.canada.ca/data/en/dataset/00a331db-121b-445d-b119-35dbbe3eedd9

Winters, D., & Pilon, P. (2020). *Pseudo-household demographic distribution* (PHH.3.5). Government of Canada. https://open.canada.ca/data/en/dataset/b3a1d603-19ca-466c-ae95-b5185e56addf

Xu, X., Wang, J., Peng, H., & Wu, R. (2019). Prediction of academic performance associated with internet usage behaviors using machine learning algorithms. *Computers in Human Behavior*, *98*, 166–173. https://doi.org/https://doi.org/10.1016/j.chb.2019.04.015