Inequality in Household Adaptation to Schooling Shocks: Covid-Induced Online Learning Engagement in Real Time

Andrew Bacher-Hicks, Joshua Goodman, Christine Mulhern

CES**ifo**

Inequality in Household Adaptation to Schooling Shocks: Covid-Induced Online Learning Engagement in Real Time

Abstract

We use high frequency internet search data to study in real time how US households sought out online learning resources as schools closed due to the Covid-19 pandemic. By April 2020, nationwide search intensity for both school- and parent-centered online learning resources had roughly doubled relative to baseline. Areas of the country with higher income, better internet access and fewer rural schools saw substantially larger increases in search intensity. The pandemic will likely widen achievement gaps along these dimensions given schools' and parents' differing engagement with online resources to compensate for lost school-based learning time. Accounting for such differences and promoting more equitable access to online learning could improve the effectiveness of education policy responses to the pandemic. The public availability of internet search data allows our analyses to be updated when schools reopen and to be replicated in other countries.

Andrew Bacher-Hicks
Boston University / MA / USA
abhicks@bu.edu

Joshua Goodman
Boston University / MA / USA
JoshuaGoodman@brandeis.edu

Christine Mulhern
RAND Corporation / Santa Monica / CA / USA
cmulhern@rand.org

1 Introduction

As schools across the United States closed in response to the Covid-19 pandemic, roughly 55 million K-12 students experienced a serious disruption to their school year. Though most schools quickly began offering some type of virtual education (Hamilton et al., 2020; Lake and Dusseault, 2020), there have been growing concerns about the effects of this unprecedented shift (Malkus, 2020; von Hippel, 2020). In particular, there are fears that low-income students will be unequally harmed by the shift to online learning, due to less access to online resources to compensate for lost in-person instruction (Horowitz, 2020). As states and districts consider how to best educate students in the wake of the pandemic, it is critical to better understand the effect of pandemic-induced school closures on students' access to online learning resources, particularly for low-income students.

This paper uses high frequency, nationally representative Google search data to document in real time how parents and students sought out online resources as schools closed in response to the Covid-19 pandemic. Combining the online search measures with measures of demographic characteristics by geography, we use a difference-in-difference strategy to estimate how Covid-induced demand for online resources varied by a range of geographic and socioeconomic factors, including income, internet access and school rurality.

We document three findings that are new to the research literature. First, we show that pre-Covid search intensity for online learning resources falls largely into two categories, which we call "school-centered resources" and "parent-centered resources". School-centered resources are platforms typically used by schools to provide instruction, assign work, or communicate with students (such as Google Classroom or Schoology). Parent-centered resources are more generic search terms likely indicating parents or students are seeking supplemental learning resources (such as home schooling or math worksheets). We show that search intensity for school-centered resources dwarfs that for parent-centered resources and that both follow the school calendar, peaking at the start of each school year and vanishing in the summer.

Second, we show that the onset of Covid dramatically disrupted this usual school calendar cycle of search intensity, as the pandemic triggered a very large increase in demand for online

learning resources. By April 2020, nationwide search intensity for online learning resources had roughly doubled relative to baseline. We find sharp increases in searches for both school- and parent-centered resources, suggesting that increased demand for online support came not only from schools shifting their mode of instruction but also from parents and students seeking additional support to fill learning gaps as schools closed.

Third, we show the pandemic substantially widened socioeconomic gaps in searches for online learning resources. Search intensity rose twice as much in areas with above median socioeconomic status (measured by household income, parental education, and computer and internet access) as in areas with below mean socioeconomic status. Search intensity for school-centered resources, for example, increased by 15 percent for each additional $10,000 in mean household income and by roughly 50 percent for each 10 percentage point increase in the fraction of households with broadband internet and a computer. Areas with more rural schools and Black students saw lower increases in search intensity. Socioeconomic gaps widened both between and within the country's four Census regions (Northeast, Midwest, South, and West).

Our work adds to three strands of the research literature. First, our paper is the first to show that internet search behavior can provide useful, real-time information about education-related actions being taken by households. Prior work has shown the utility of search data in predicting economic and social outcomes such as disease spread (Polgreen et al., 2008), consumer behavior (Choi and Varian, 2012), voting (Stephens-Davidowitz, 2014), and fertility decisions (Kearney and Levine, 2015). Most recently, Goldsmith-Pinkham and Sojourner (2020) use the volume of online search for unemployment benefits to predict post-Covid unemployment claims. Our results suggest that search data contain valuable information about how households react to educational shocks, both in terms of overall use of educational resources and in heterogeneity in such usage by socioeconomic characteristics.

Second, we measure a new aspect of the digital divide, namely the extent to which households seek out online learning resources either prompted by their schools or of their own accord. A large literature documents pre-Covid gaps in access to and proficiency in the use digital technologies by income, education, and family background (Bucy, 2000; Rice and Haythornthwaite, 2006; Jones

et al., 2009; Vigdor et al., 2014). Multiple post-Covid surveys show consistent socioeconomic gaps in self-reported engagement with remote learning at a single point in time (Barnum and Bryan, 2020). We complement this evidence with the first nationally representative revealed preference measure of such engagement, based on households' actual behavior rather than self reports. Ours is also the first high frequency data brought to bear on this issue, allowing study of the evolution over time of engagement with online learning resources.[1]

Third, we provide some of the clearest evidence on one channel through which the Covid-19 pandemic will likely widen socioeconomic educational gaps. Based on prior estimates of school closure effects from natural disasters and summer months, Kuhfeld et al. (2020) predict that Covid-induced closures will generate substantial learning losses, with the largest negative effects concentrated among low-achieving students. Aucejo et al. (2020) surveyed university students and find the pandemic lowered on-time graduation rates and job offers, with larger effects among low-income students. Using data from one online learning platform, Chetty et al. (2020) provide perhaps the only direct measure of Covid-induced learning loss, showing that low-income students experienced substantially larger and more persistent reductions in learning progress relative to high-income students. We show that socioeconomic gaps in engagement with online learning resources are not limited to a single platform or location but are a widespread and fundamental feature of the post-Covid landscape. Accounting for household responses to changing school inputs will be critical for predicting educational effects of the pandemic and policy responses to it, given evidence that parental and school investments are often substitutes, both in the US (Houtenville and Conway, 2008) and developing countries (Das et al., 2013; Pop-Eleches and Urquiola, 2013).

Our findings provide insight into the mechanisms underlying learning losses that are beginning to emerge following pandemic-induced school closures and can help inform future policy responses to schooling disruptions, whether related to the pandemic or not. That search for school-centered resources increases more in high income areas suggests either that those areas' schools are using online platforms more, that those areas' parents are more likely to engage with such

---

[1]The pandemic has also exacerbated socioeconomic gaps in non-online forms of parental compensation for lost schooling time. Jæger and Blaabæk (2020) use Danish register data to show that pre-Covid income-based gaps in library book takeout rates increased post-Covid.

platforms, or both. That search for parent-centered resources increases more in high income areas suggests that, separate from schools' actions, parents are differentially likely to seek out their own ways of compensating for their children's lost learning time.

These results can help policymakers and school leaders formulate more effective responses to the educational disruptions caused by Covid-induced school closures. Students from lower income families and schools may require additional attention and resources given lower engagement with online learning resources during spring 2020. Moreover, because remote learning will likely remain a central piece of the public education system for the foreseeable future (Cleveland, 2020), preventing the widening of achievement gaps may require improving access to home computers and broadband internet for low income and rural students. Schools may also need to improve the deployment of remote learning platforms to more equitably engage students and parents in the use of those platforms.

Whether efforts to close gaps in online learning engagement succeed will only become clear as new data become available in subsequent school years. One advantage of using publicly available search data to measure household behavior is that our analyses can be easily updated in real time when the school year begins in the fall. This will help reveal whether socioeconomic gaps in engagement with online learning have narrowed since the initial shock of schools closing or if different remote learning strategies across regions were particularly successful. In addition, the set of search terms studied can be easily modified to accommodate new online learning resources as they emerge. Finally, our analyses can be replicated in other countries, particularly ones large enough to generate search data at sub-national levels such as provinces and cities. The flexibility of this approach shows promise for understanding the behavioral responses of households to school closures and developing policy responses in real time to address changing student needs.

2 Data and Empirical Strategy

2.1 Data

Our outcome measures of search intensity come from Google Trends, which makes publicly available weekly measures of internet search behavior both nationally and by Designated Market Area (DMA). Each of the 210 United States DMAs comprise a mutually exclusive small group of counties and are the finest geographic level at which search data is consistently available.[2] The publicly available measure of search behavior for a given term or topic is "search intensity", which calculates the fraction of a given area's Google searches devoted to that term or topic. Raw search volume and raw search intensity are not available. Instead, Google Trends normalizes measures of search intensity to allow for comparison of relative intensity over time and across DMAs. Given the challenge of interpreting such magnitudes, we often use the logarithm of search intensity so that estimates can be interpreted as percent changes.

We focus on measuring search intensity for terms related to online learning. We first assembled a list of 45 potential such keywords, downloaded five years of weekly search data from June 2015 through May 2020, and then ranked the keywords by their nationwide popularity. We then identified the 10 most popular keywords related to branded online learning resources (such as "Google Classroom" or "Khan Academy") and the ten most popular keywords for general online learning resources (such as "online learning", "home school", and "math worksheet"). Table 1 shows these lists and their popularity.[3] The most popular of these search terms by far is Google Classroom. Khan Academy, one of the next most popular terms, was roughly 13 percent as popular as Google Classroom.[4] General (non-branded) learning resources saw substantially lower search intensity than that.

We combine the ten items in each of these two lists to create two primary measures of search intensity for learning resources. We do so for three reasons. First, Google Trends does not publicly

---

[2] DMAs are defined by Nielsen. For more information, see: `https://www.nielsen.com/us/en/intl-campaigns/dma-maps/`.

[3] For the full list of 45 keywords, see see Table A.1.

[4] We normalize search intensity across keywords relative to the pre-Covid (March-May 2019) search intensity of Google Classroom.

report search intensity if the total number of searches in a given area and week falls below a certain threshold, so combining search terms substantially reduces the amount of missing data. Second, we are more interested in the general level of demand for online learning resources rather than the popularity of any one resource. Third, Google Trends restricts the number of keywords whose total search intensity can be downloaded at one time, which requires us to focus on the most popular keywords.[5] Therefore, our measure of search intensity aggregates the total popularity for the ten most popular keywords in DMA d in week t as follows

$$SearchIntensity_{dt} = \sum_{i=1}^{10} \frac{Google\ searches\ for\ [keyword\ i]_{dt}}{Total\ Google\ searches_{dt}}. \tag{1}$$

Given the ten keywords in each list, we refer to branded learning resources as "school-centered resources" and general learning resources as "parent-centered resources". The former list largely consists of educational platforms, such as Google Classroom and Schoology, that schools use to connect with students and are typically not used without the school's involvement. The latter list consists of general learning resources such as "math worksheets" and "home school", which we interpret as parents (or guardians or students themselves) searching for online learning resources on their own, without particular guidance from a school.[6]

To characterize the pre-Covid demographics of each DMA, we use county-level data from the 2016 American Community Survey (ACS) and then generate DMA-level measures as the population-weighted average across the counties within each DMA.[7] We can therefore characterize DMAs by a variety of measures, including: mean household income; median household income; fraction of adults with a B.A.; fraction of households with broadband internet; and fraction of households with a computer. These five measures are so highly correlated that we refer to their population-weighted first principal component, which nearly equally weights each measure

---

[5]Nearly all of the search intensity for online learning is concentrated in the top 10 terms in each list as shown in Table A.1.

[6]Khan Academy is perhaps the most ambiguous search term, in that many schools use it to deliver curriculum to students, but parents and students can also access it without the school's involvement. We place it in the list of school-centered resources but often show separate results for it and for Google Classroom, the two most popular search terms in the post-Covid period.

[7]In seven cases (0.2 percent of counties), the county is within more than one DMA. In these cases, we include the county in both DMAs.

and explains 86 percent of their joint variation, as DMA-level socioeconomic status (SES). This allows us to do simple comparisons of search intensity across high and low SES areas of the country. We supplement ACS data with county-level data from the Stanford Education Data Archive (SEDA), which provides additional information such as the fraction of schools in rural areas and the racial composition of the school-age population. As of 2016, 27 percent of adults had a B.A., 23 percent lived in households without broadband internet access, and 25 percent of schools were in rural areas.[8]

2.2 Empirical Strategy

We first estimate changes in nationwide search intensity for learning resources as a result of Covid-induced school closures, using both a week-by-week event study specification and a two-period before-after specification. The event study specification is

$$Log(SearchIntensity_{twy}) = \sum_{i=-39}^{-1} \beta_i \, iBefore_t + \sum_{i=1}^{12} \beta_i \, iAfter_t + \alpha PriorYears_t + \mu_w + \lambda_y + \varepsilon_{twy}, \quad (2)$$

where we regress on a vector of week indicators the logarithm of nationwide search intensity in week t, week of year number w (1-52), and school year y (2015-2020). Here, $iBefore$ and $iAfter$ are indicators for week t falling i weeks before or after March 1, 2020, when states began discussing the possibility of school closures. Inclusion of week of year and school year fixed effects, which we call calendar effects, means that coefficients β_i can be interpreted as differences in search intensity compared to the same weeks in prior years. The vector of weekly indicators cover all 52 weeks (other than that starting March 1, 2020) in the most recent year of our data, leaving the prior four years of data to identify week of year effects. Exclusion of the March 1, 2020 indicator, and inclusion of a $PriorYears$ indicator for observations in the first four years of the data, means the coefficients β_i can be interpreted as week i's deviation from calendar-predicted search intensity relative to March 1, 2020.

The nationwide before-after specification, which replaces the vector of weekly pre- and post-

---

[8]See Table A.2 for these and other summary statistics.

Covid indicators with a single post-Covid indicator, is

$$Log(SearchIntensity_{twy}) = \beta PostCovid_t + \mu_w + \lambda_y + \varepsilon_{twy}. \qquad (3)$$

Inclusion of calendar effects means β can be interpreted as the overall post-Covid increase in search intensity compared to those same weeks in prior years. We exclude the ramp-up period of March 2020 from the before-after analysis because school closure discussions began in early March and nearly all schools were initially closed by states between March 16 and March 23.[9] The post-Covid coefficient thus estimates the average increase in search intensity in April and May of 2020, when schools and parents were fully aware of the scope of educational disruption. The difference-in-difference estimator β from Equation 3 is the average of the April-May weekly event study coefficients β_4 through β_{12} from Equation 2.

To study how search intensity changed differentially by SES, we modify our nationwide specifications by interacting the pre- and post-Covid indicators with indicators for whether a DMA is of above or below median SES. Our modified event study specification is

$$
\begin{aligned}
Log(SearchIntensity_{dtwy}) = & \sum_{i=-39}^{-1} (\beta_i \, iBefore_t + \gamma_i \, iBefore_t * HighSES_d) \\
& + \sum_{i=1}^{12} (\beta_i \, iAfter_t + \gamma_i \, iAfter_t * HighSES_d) \\
& + \alpha_1 PriorYears_t + \alpha_2 PriorYears_t * HighSES_d + \mu_w + \lambda_y + \varepsilon_{twy}, \quad (4)
\end{aligned}
$$

so that the coefficients γ_i estimate the difference in weekly search intensity between high and low SES DMAs.

We modify the nationwide before-after specification to become a difference-in-difference of the

---

[9]See Table A.3 for a list of school closure dates by state.

8

form

$$Log(SearchIntensity_{dtwy}) = \beta_1 PostCovid_t * HighSES_d + \beta_2 PostCovid_t * LowSES_d$$

$$+ \sigma HighSES_d + \mu_w + \lambda_y + \varepsilon_{twy}. \quad (5)$$

Here, β_1 and β_2 measure the overall post-Covid change in search intensity for high and low SES DMAs respectively. We then compute the difference between those two coefficients, which corresponds to the average over time of the event study coefficients γ_4 through γ_{12}. We also run difference-in-difference specifications where the post-Covid indicator is interacted with continuous measures of DMA characteristics, such as income, broadband penetration rates, and school rurality.

All DMA-level regressions use standard errors clustered by DMA and are weighted by DMA population to be nationally representative at the individual level. Nation-level regressions use heteroskedasticity robust standard errors. All regressions exclude weeks with the major school holidays of Thanksgiving, Christmas and New Year's. This has no effect on any estimates of interest but improves the scale of event-study graphs given the increased noise associated with such weeks.

3   Results

Nationwide search intensity for learning resources followed regular annual patterns until March 2020, at which point search intensity dramatically increased relative to similar months in prior years. As seen in Figure 1, search intensity for school-centered resources typically peaks near the start of each school year and declines steadily until summer, when it largely vanishes. Search intensity for parent-centered resources is steadier throughout the school year but also declines substantially in summer. Covid-related school closures altered these patterns, with nationwide search intensity for both types of learning resources roughly doubling by late March.[10] Search intensity then starts to decay, likely due to households successfully locating their desired online

---

[10] A logarithmic increase of 0.7 represents a roughly 100 percent increase ($e^{0.7} - 1 \approx 1$).

resources and to school years ending in May and June. Accounting for usual calendar effects confirms all of these changes observed in the raw data, with no pre-trends prior to March 2020 and then a large and statistically significant rise in search intensity that then fades by late May.[11]

Growth in post-Covid search intensity varied substantially by geography and socioeconomic status. Figure 2 maps DMAs by quartiles of SES (panel A) and post-Covid changes in search intensity (panel B), defined as April-May search intensity differences between 2020 and prior years. Areas with high income, parental education and technological access are concentrated in the Northeast and West coast, as well as Utah and Colorado. Post-Covid search intensity also increases most noticeably, though not exclusively, in the Northeast and West coast, suggesting that high SES areas see larger spikes in search intensity for learning resources.

Event study analyses show that weekly search intensity for learning resources increased significantly more in areas with higher income and better technological access. Figure 3 shows Equation 4's regression-based weekly difference in search intensity between DMAs above and below median SES, controlling for calendar effects. By mid-March, high SES areas see jumps in search intensity for both school- and parent-centered learning resources roughly 30 percent higher than low SES areas. These weekly differences are statistically significant and do not decay with time. Such differences are also visible in the unadjusted data.[12]

Difference-in-difference analyses show that the average post-Covid increase in search intensity for learning resources was roughly twice as high in high SES areas as in low ones. Table 2 shows, in panel A's before-after analysis based on Equation 3, that nationwide search intensity increased 67 percent (51 log points) for school-based resources, driven largely by a 114 percent (76 log point) increase for Google Classroom, and 40 percent (34 log points) for parent-centered resources. Panel B, based on Equation 5, shows that search intensity for all of these resources increased substantially in low SES areas but much more so in high SES areas. Low SES areas searched 36 percent (31 log points) more intensely for school-centered resources but high SES areas searched an additional 48 percent (39 log points) more intensely, a difference that is highly statistically significant. Search intensity for parent-centered resources increased by 25 percent (22 log points) in low SES

---

[11]See Figure A.1, based on Equation 2.

[12]See Figure A.2.

10

areas but by an additional 27 percent (24 log points) in high SES areas. These patterns hold for broad resource categories and for individual products like Google Classroom and Khan Academy. Interestingly, the high SES indicator in panel B's regressions suggests few clear differences by SES in search intensity prior to the pandemic.

The pandemic widened gaps in search intensity for learning resources not only by broad measures of SES but also by specific measures such as income, technological access, rurality and race. Panel C of Table 2 shows difference-in-difference analyses where continuous DMA-level demographic measures have been interacted with a post-Covid indicator. Search intensity for school-centered resources increases by an additional 15 percent (14 log points) with every additional $10,000 in mean household income and by a roughly additional 50 percent (36 and 44 log points) with every 10 percentage point increase in the fraction of households with broadband internet and a computer. A 10 percentage point increase in the fraction of schools located in rural areas is associated with a 17 percent (16 log point) lower change in search intensity, while a 10 percentage point increase in the fraction of students who are Black is associated with a nine percent lower change in search intensity. Parent-based resources show similar patterns, though with somewhat smaller magnitudes. Measures of SES (such as income or broadband access), school rurality and race are independently associated with search behavior even when controlling for each other.[13]

Finally, we note three additional analyses that further illuminate or strengthen the prior results. First, Table A.5 presents effects by SES quartile. This analysis shows that the observed post-Covid widening of search intensity gaps is fairly linear in SES and not just driven by the most and least disadvantaged areas. Second, the widening of gaps by SES is not only the result of SES gaps between regions of the country but gaps within regions. Table A.6 reports effects for each of the four U.S. Census regions (Northeast, Midwest, South, West) and shows that SES-based search intensity gaps appear even when making within-region comparisons. Finally, because search intensity is coded as zero for terms with low search volume in a given DMA-week, Table A.7 excludes DMAs with low search volume. The results are unchanged, which is unsurprising given that such areas tend to have low population.

---

[13]See Table A.4.

4 Discussion

We document a sharp increase in searches for learning resources as schools closed in response to the Covid-19 pandemic. By April 2020, nationwide search intensity for online learning resources had roughly doubled relative to baseline. The shock of school closures increased demand both for the specific online platforms schools shifted instruction to (such as Google Classroom) and for the supplemental resources that households sought out to fill gaps in their learning (such as math worksheets). The likelihood of future school closures or partial re-openings implies these supplemental online resources are likely to become important drivers of student learning.

Though demand for online resources increased in both high and low SES areas, the increase was substantially larger in high SES areas. Areas of the country with higher income, greater internet access, and fewer rural schools had substantially larger increases than less advantaged areas. Along with results from several contemporaneous studies, these results suggest that academic gaps across students will be wider than normal in future school years.

Our results suggest the potential value of policy responses that directly address these documented inequalities in engagement with online learning resources. Students in low SES areas and rural communities are likely to need additional support to overcome the educational challenges created by Covid-19. Because online learning will likely remain a key component of school systems in the near future, school leaders and policymakers may want to prioritize access to home computers and broadband internet. Improving access to and engagement with online learning platforms will likely be an important step to equalizing learning opportunities and preventing a widening of achievement gaps.

Publicly available, high frequency internet search data helps illuminate the evolution of educational choices made by households, as well as socioeconomic inequalities in those choices. Our analyses can be updated in real time to study future changes in engagement with online learning, can be modified to study different search terms, and can be replicated in other countries. Household adaptation to schooling shocks is an understudied phenomenon that can be readily observed in internet search data. Understanding and accounting for such behavioral responses by parents and students will be critical to predicting the long-term effects of the pandemic.

References

Aucejo, E. M., J. F. French, M. P. U. Araya, and B. Zafar (2020). The impact of COVID-19 on student experiences and expectations: Evidence from a survey. *NBER Working Paper 27392*.

Barnum, M. and C. Bryan (2020, June 26). America's great remote-learning experiment: What surveys of teachers and parents tell us about how it went. *Chalkbeat*.

Bucy, E. P. (2000). Social access to the internet. *Harvard International Journal of Press/Politics 5*(1), 50–61.

Chetty, R., J. N. Friedman, N. Hendren, and M. Stepner (2020). Real-time economics: A new platform to track the impacts of COVID-19 on people, businesses, and communities using private sector data. *NBER Working Paper 27431*.

Choi, H. and H. Varian (2012). Predicting the present with Google Trends. *Economic Record 88*, 2–9.

Cleveland, C. (2020, June 29). Toward reopening: What will school look like this fall? *EducationNext*.

Das, J., S. Dercon, J. Habyarimana, P. Krishnan, K. Muralidharan, and V. Sundararaman (2013). School inputs, household substitution, and test scores. *American Economic Journal: Applied Economics 5*(2), 29–57.

Goldsmith-Pinkham, P. and A. Sojourner (2020). Predicting initial unemployment insurance claims using Google Trends. *Working Paper*.

Hamilton, L. S., D. Grant, J. H. Kaufman, M. Diliberti, H. L. Schwartz, G. P. Hunter, C. Messan Setodji, and C. J. Young (2020). COVID-19 and the state of K–12 schools: Results and technical documentation from the spring 2020 American educator panels COVID-19 surveys. *RAND Research Report*.

Horowitz, J. (2020, April 15). Lower-income parents most concerned about their children falling behind amid COVID-19 school closures. *Pew Research Center*.

Houtenville, A. J. and K. S. Conway (2008). Parental effort, school resources, and student achievement. *Journal of Human Resources 43*(2), 437–453.

Jæger, M. M. and E. H. Blaabæk (2020). Inequality in learning opportunities during Covid-19: Evidence from library takeout. *Research in Social Stratification and Mobility*, 100524.

Jones, S., C. Johnson-Yale, S. Millermaier, and F. S. Pérez (2009). US college students' internet use: Race, gender and digital divides. *Journal of Computer-Mediated Communication 14*(2), 244–264.

Kearney, M. S. and P. B. Levine (2015). Media influences on social outcomes: The impact of MTV's 16 and Pregnant on teen childbearing. *American Economic Review 105*(12), 3597–3632.

Kuhfeld, M., J. Soland, B. Tarasawa, A. Johnson, E. Ruzek, and J. Liu (2020). Projecting the potential impacts of COVID-19 school closures on academic achievement. *Annenberg Institute Working Paper 20-226*.

Lake, R. and B. Dusseault (2020, April 27). Remote classes are in session for more school districts, but attendance plans are still absent. *Center on Reinventing Public Education*.

Malkus, N. (2020, June 16). School districts' remote-learning plans may widen student achievement gap. *EducationNext*.

Polgreen, P. M., Y. Chen, D. M. Pennock, F. D. Nelson, and R. A. Weinstein (2008). Using internet searches for influenza surveillance. *Clinical Infectious Diseases 47*(11), 1443–1448.

Pop-Eleches, C. and M. Urquiola (2013). Going to a better school: Effects and behavioral responses. *American Economic Review 103*(4), 1289–1324.

Rice, R. E. and C. Haythornthwaite (2006). Perspectives on internet use: Access, involvement and interaction. *The Handbook of New Media*, 92–113.

Stephens-Davidowitz, S. (2014). The cost of racial animus on a black candidate: Evidence using Google search data. *Journal of Public Economics 118*, 26–40.

Vigdor, J. L., H. F. Ladd, and E. Martinez (2014). Scaling the digital divide: Home computer technology and student achievement. *Economic Inquiry 52*(3), 1103–1119.

von Hippel, P. (2020, April 9). How will the coronavirus crisis affect children's learning? Unequally. *EducationNext*.

Figure 1: Weekly Nationwide Search Intensity for Learning Resources

(A) School-Centered Resources

(B) Parent-Centered Resources
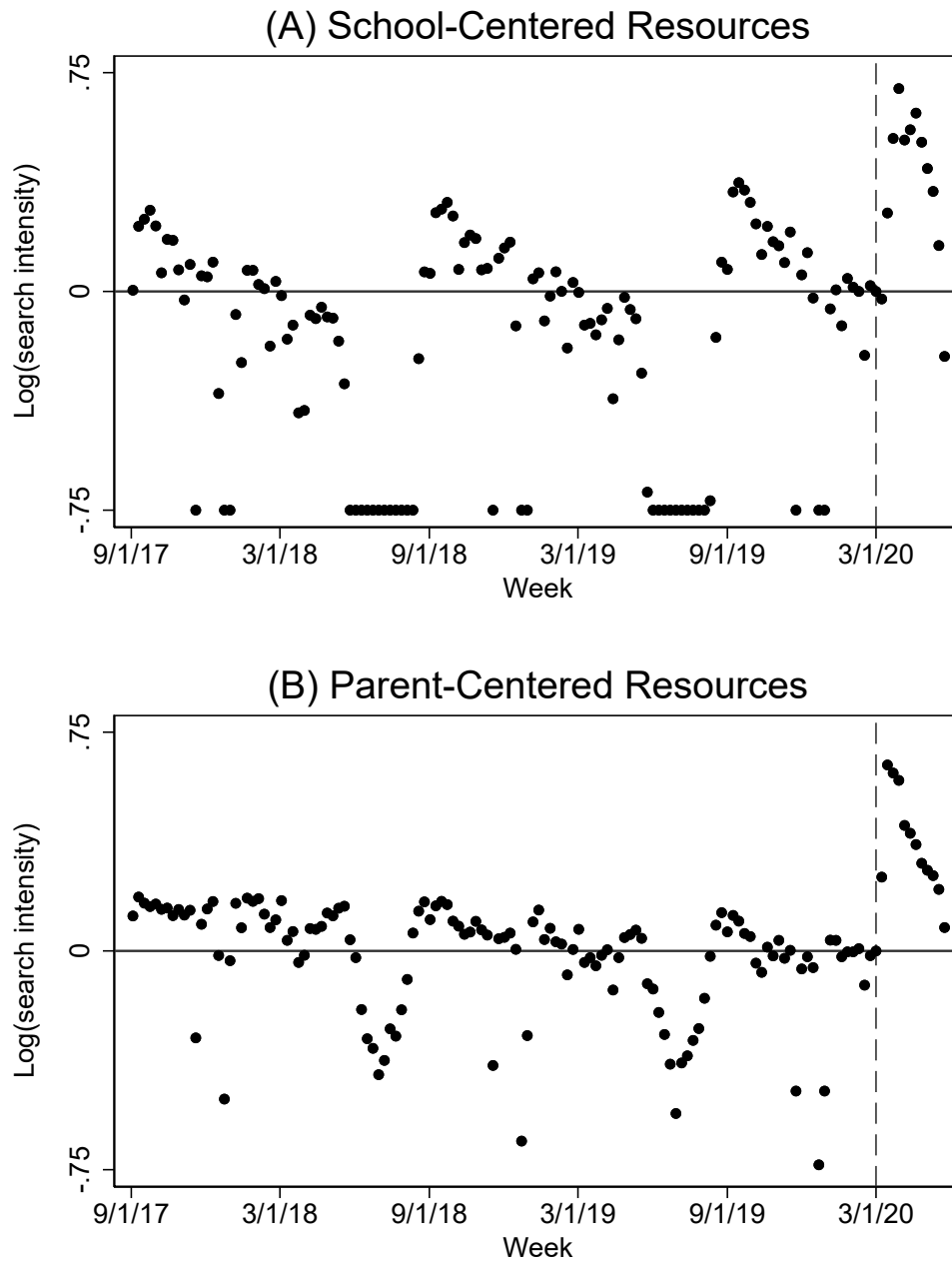
Notes: The figure above shows weekly measures of the logarithm of nationwide search intensity relative to intensity on March 1, 2020. Panel A shows search intensity for school-centered resources and panel B for parent-centered resources. Search intensity is bottom-coded at -0.75 for scaling purposes.

Figure 2: Geography of Socioeconomic Status and Change in Search Intensity

(A) Socioeconomic Status

Top quartile
3rd quartile
2nd quartile
Bottom quartile

(B) Post-Covid Change in Search Intensity
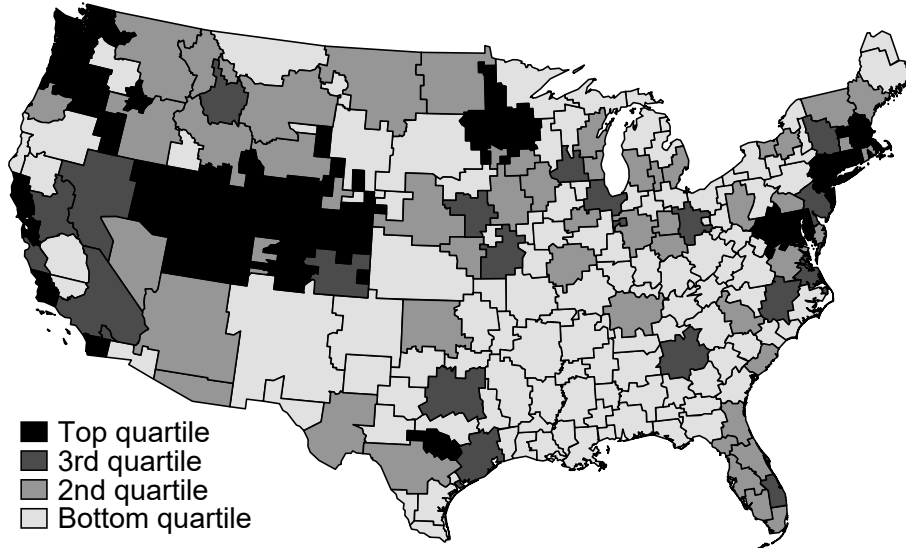
Top quartile
3rd quartile
2nd quartile
Bottom quartile
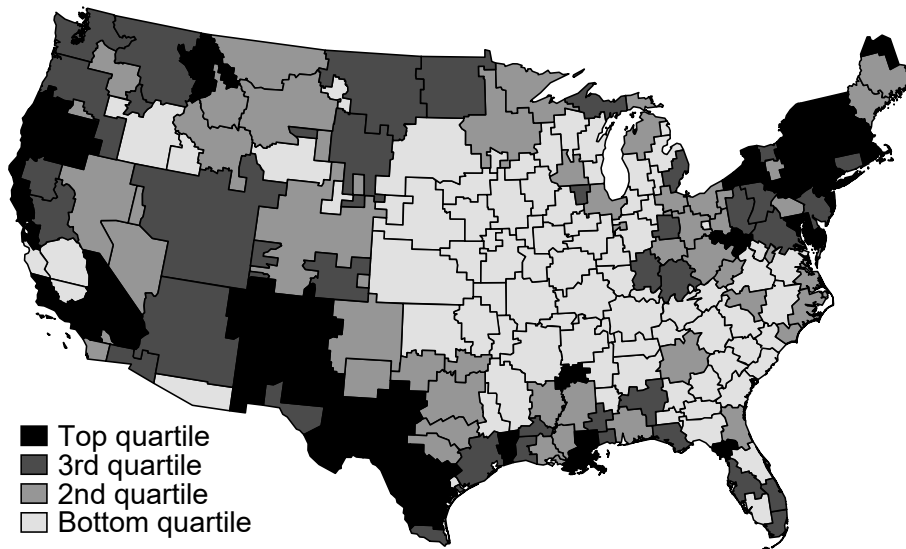
Notes: The figure above maps Designated Market Areas (DMA) across the nation. Panel A sorts DMAs into quartiles of socioeconomic status, which is defined as the first principal component of: mean household income; median household income fraction of adults with a four-year college degree; fraction of households with a computer; and fraction of households with broadband access. Panel B sorts DMAs into quartiles of post-Covid changes in search intensity for school-centered resources, defined as the difference in the logarithm of search intensity between April-May 2020 and April-May of 2015-19.

Figure 3: Event Study of Search Intensity Gap by Socioeconomic Status
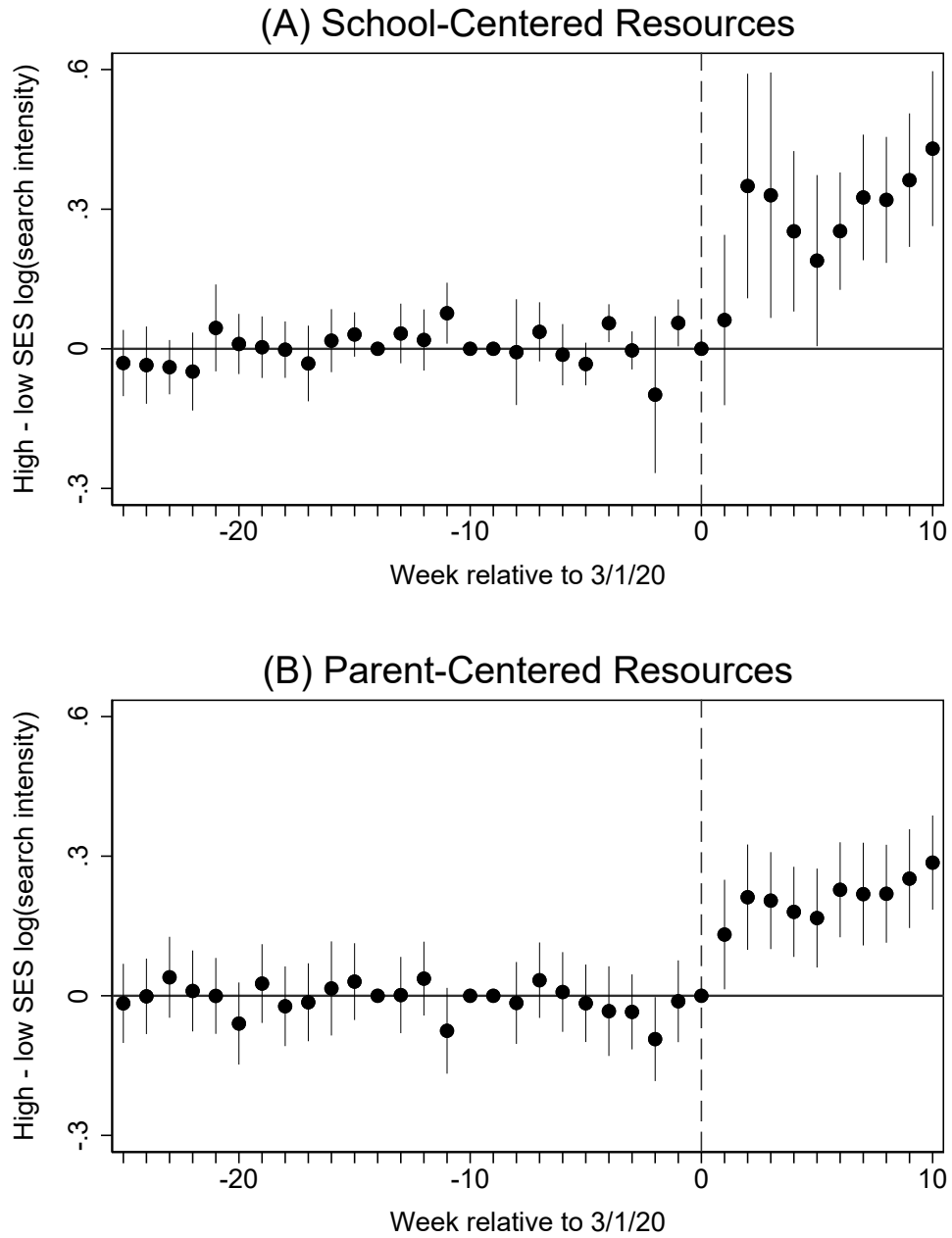
Notes: The figure above shows event study coefficients estimating the difference in the logarithm of weekly search intensity between Designated Market Areas (DMA) of above and below median socioeconomic status. The regressions are weighted by DMA population and include fixed effects for week of year (1-52) and school year (2015-2020). Also shown are 95 percent confidence intervals calculated with standard errors clustered by DMA. Panel A shows search intensity for school-centered resources and panel B for parent-centered resources. The sample contains search data from June 2015 through May 2020, excluding weeks with major school holidays.

Table 1: Search Intensity of Top 10 Individual Keywords

| School-Centered Resources				Parent-Centered Resources		
Keyword	Pre-Covid	Post-Covid		Keyword	Pre-Covid	Post-Covid
Google Classroom	1.00	1.95		Online school	0.04	0.06
Khan Academy	0.13	0.20		Online classes	0.03	0.05
Kahoot	0.33	0.19		Home school	0.03	0.03
Seesaw	0.02	0.15		Online class	0.00	0.02
Schoology	0.07	0.12		Math game	0.03	0.02
Class Dojo	0.01	0.06		Distance learning	0.00	0.02
Flipgrid	0.00	0.05		Math worksheets	0.00	0.02
D2L	0.05	0.05		Online math	0.00	0.01
Nearpod	0.02	0.02		Math problem	0.00	0.01
Edmodo	0.02	0.02		Online reading	0.00	0.00

Notes: Mean nationwide search intensity is shown for March-May 2019 (pre-Covid) and March-May 2020 (post-Covid). Search intensity of each term is measured relative to the pre-Covid search intensity for "Google Classroom". Searches are not case sensitive.

Table 2: Changes in Search Intensity by Socioeconomic Status and Other Measures

	School-centered resources (1)	Parent-centered resources (2)	Google Classroom (3)	Khan Academy (4)
(A) Nationwide				
Post-Covid	0.51***	0.34***	0.76***	0.41***
	(0.05)	(0.02)	(0.05)	(0.03)
(B) By median SES				
Post-Covid * High SES	0.70***	0.46***	0.92***	0.50***
	(0.08)	(0.03)	(0.08)	(0.05)
Post-Covid * Low SES	0.31***	0.22***	0.59***	0.32***
	(0.05)	(0.02)	(0.07)	(0.03)
High SES	0.02	0.01	0.10	0.05
	(0.06)	(0.05)	(0.10)	(0.05)
High-Low SES Change	0.39***	0.24***	0.33***	0.19***
	(0.11)	(0.03)	(0.11)	(0.05)
(C) By income, online access, race				
Post-Covid * Household income	0.14***	0.08***	0.12***	0.07***
	(0.04)	(0.01)	(0.03)	(0.02)
Post-Covid * Broadband penetration rate	0.36***	0.30***	0.31***	0.27***
	(0.08)	(0.03)	(0.09)	(0.04)
Post-Covid * Computer penetration rate	0.44***	0.41***	0.41***	0.33***
	(0.10)	(0.04)	(0.10)	(0.07)
Post-Covid * Fraction of schools in rural areas	-0.16***	-0.08***	-0.18***	-0.09***
	(0.03)	(0.01)	(0.03)	(0.02)
Post-Covid * Fraction of students who are Black	-0.09***	-0.06***	-0.03	-0.06***
	(0.03)	(0.01)	(0.04)	(0.02)
N	50,400	50,400	50,400	50,400

Notes: Heteroskedasticity robust standard errors clustered by DMA are in parentheses (* p<.10 ** p<.05 *** p<.01). In panel A, each column regresses the logarithm of the listed topic's search intensity on a post-March 1 indicator. Panel B interacts the post indicator with above and below median SES indicators. Below each pair of interaction coefficients in panel B is the estimated difference between the high and low SES interaction coefficients. Each coefficient in Panel C comes from a separate regression that interacts the post indicator with DMA-level household mean income or the other listed resource and demographic measures. Income is measured per $10,000, while the remaining variables are scaled to interpret coefficients as predicted effects of a 10 percentage point increase in the interacted characteristic. The regressions are weighted by DMA population and include fixed effects for week of year (1-52) and school year (2015-2020). The sample contains search data from June 2015 through May 2020, excluding March 2020 and weeks with major school holidays.

Figure A.1: Event Study of Nationwide Search Intensity for Learning Resources
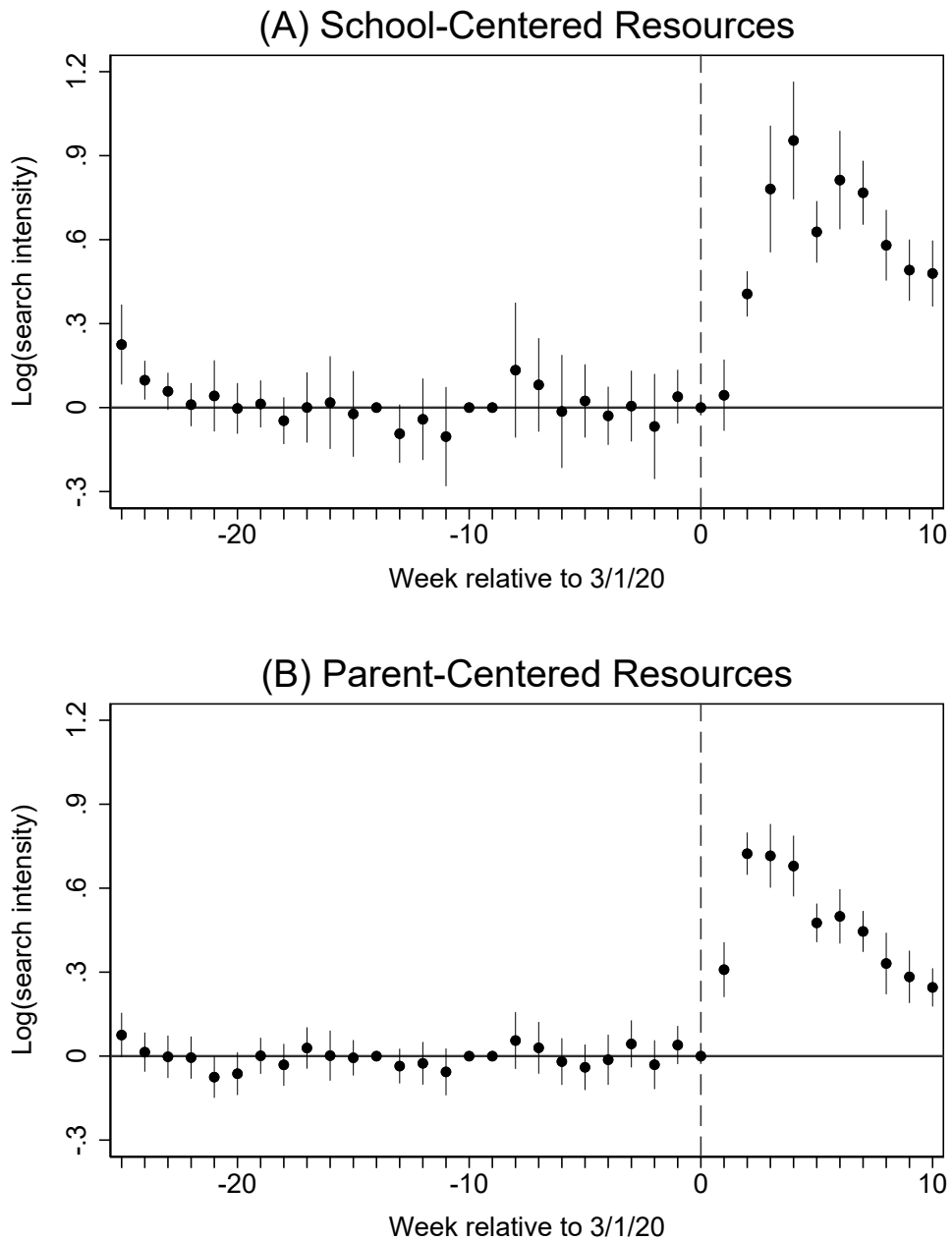
(A) School-Centered Resources

(B) Parent-Centered Resources

Notes: The figure above shows event study coefficients estimating the difference in the logarithm of nationwide weekly search intensity relative to the same weeks in prior years. The regressions are weighted by DMA population and include fixed effects for week of year (1-52), school year (2015-2020), and weeks containing major school holidays. Also shown are 95 percent confidence intervals calculated with heteroskedasticity robust standard errors. Panel A shows search intensity for school-centered resources and panel B for parent-centered resources. The sample contains search data from June 2015 through May 2020, excluding weeks with major school holidays.

Figure A.2: Weekly Search Intensity for Learning Resources by Socioeconomic Status

(A) School-Centered Resources

(B) Parent-Centered Resources
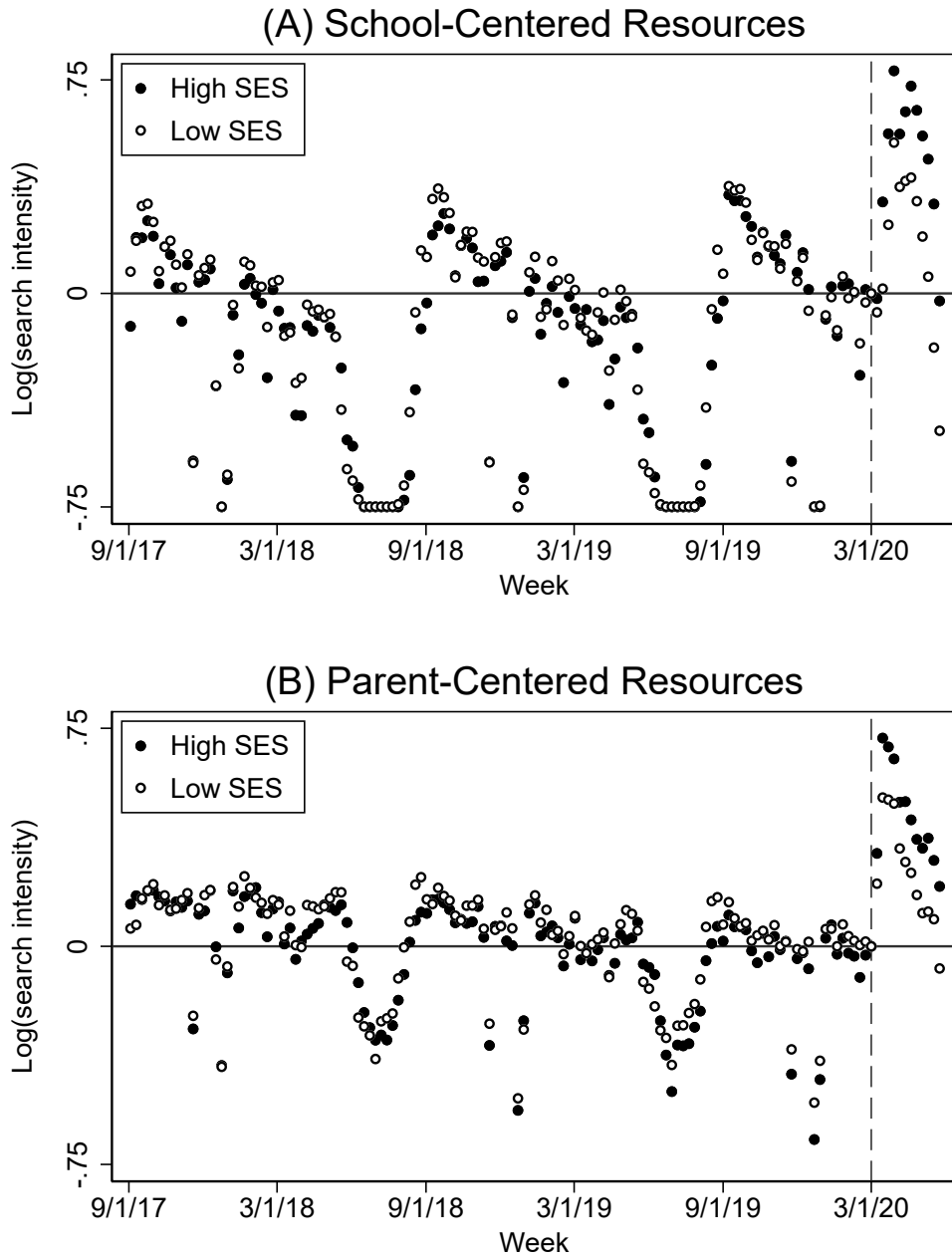
Notes: The figure above separates Designated Market Areas into those of above and below median socioeconomic status, then shows weekly measures of the logarithm of search intensity relative to intensity on March 1, 2020. Panel A shows search intensity for school-centered resources and panel B for parent-centered resources. Search intensity is bottom-coded at -0.75 for scaling purposes.

Table A.1: Search Intensity of Individual Keywords

Branded Learning Resources			General Learning Resources		
Keyword	Pre-Covid	Post-Covid	Keyword	Pre-Covid	Post-Covid
Google Classroom	1.00	1.95	Online school	0.04	0.06
Khan Academy	0.13	0.20	Online classes	0.03	0.05
Kahoot	0.33	0.19	Home school	0.03	0.03
Seesaw	0.02	0.15	Online class	0.00	0.02
Schoology	0.07	0.12	Math game	0.03	0.02
Class Dojo	0.01	0.06	Distance learning	0.00	0.02
Flipgrid	0.00	0.05	Math worksheets	0.00	0.02
D2L	0.05	0.05	Online math	0.00	0.01
Nearpod	0.02	0.02	Math problem	0.00	0.01
Edmodo	0.02	0.02	Online reading	0.00	0.00
Flocabulary	0.02	0.02	Educational game	0.00	0.00
Starfall	0.03	0.02	Education game	0.00	0.00
GoNoodle	0.00	0.02	Online lessons	0.00	0.00
ClassDojo	0.00	0.02	Free preschool worksheets	0.00	0.00
Socrative	0.02	0.00	Educational apps	0.00	0.00
			Education apps	0.00	0.00
			Educational games	0.00	0.00
			Vocabulary game	0.00	0.00
			School worksheets	0.00	0.00
			Reading game	0.00	0.00
			Online tutoring	0.00	0.00
			Virtual education	0.00	0.00
			Online lesson	0.00	0.00
			Virtual school	0.00	0.00
			Educational videos	0.00	0.00
			Educational app	0.00	0.00
			Free school worksheets	0.00	0.00
			Education app	0.00	0.00
			Online science	0.00	0.00
			Online social studies	0.00	0.00
			Education games	0.00	0.00

Notes: The pre-Covid sample contains search data from March 2019 through May 2019 in the United States. The post-Covid sample contains search data from March 2020 through May 2020 in the United States. Magnitudes are interpreted as search popularity relative to the popularity of "Google Classroom" in the pre-Covid time period. Search terms are not case sensitive, so "Google Classroom" is equivalent to "google classroom." The focus of this paper is K-12 online learning resources, so we do not include keywords related to professional servives (e.g., Webex), textbooks (e.g., Pearson), or postsecondary (e.g., Canvas), or adult learning (e.g., Masterclass).

Table A.2: Summary Statistics of DMAs

	Mean	SD
(A) 2016 American Community Survey		
Mean household income (1,000s)	73.36	12.92
Median household income (1,000s)	54.89	10.20
Fraction of adults with a B.A.	26.74	6.46
Fraction of households with broadband internet	77.22	5.46
Fraction of households with a computer	86.74	3.84
(B) 2016 Stanford Education Data Archive		
Fraction of schools in rural areas	0.25	0.17
Fraction of students who are Black	0.14	0.15
N	210	

Notes: Panel A presents summary statistics of data from the 2016 American Community Survey (ACS). We use county-level ACS data and generate DMA-level measures as the population-weighted average across the counties within each DMA. Panel B presents summary statistics of 2016 data from the Stanford Education Data Archive (SEDA). We again use county-level SEDA data and generate DMA-level measures as the population-weighted average.

Table A.3: School Closure Dates by State

State	Legal status	State closure start date	Date closed for the year	Public school enrollment
Alabama	Ordered	March 19	April 6	744,930
Alaska	Ordered	March 16	April 9	132,737
Arizona	Ordered	March 16	March 30	1,123,137
Arkansas	Ordered	March 17	April 6	493,447
California	Recommended	March 19	April 1	6,309,138
Colorado	Ordered	March 23	April 20	905,019
Connecticut	Ordered	March 17	May 5	535,118
Delaware	Ordered	March 16	April 24	136,264
District of Columbia	Ordered	March 16	April 17	85,850
Florida	Recommended	March 16	April 18	2,816,791
Georgia	Ordered	March 18	April 1	1,764,346
Hawaii	Ordered	March 23	April 17	181,550
Idaho	Recommended	March 24	April 6	297,200
Illinois	Ordered	March 17	April 17	2,026,718
Indiana	Ordered	March 20	April 2	1,049,547
Iowa	Ordered	March 16	April 17	509,831
Kansas	Ordered	March 18	March 17	494,347
Kentucky	Recommended	March 16	April 20	684,017
Louisiana	Ordered	March 16	April 15	716,293
Maine	Recommended	March 16	March 31	180,512
Maryland	Ordered	March 16	May 6	886,221
Massachusetts	Ordered	March 17	April 21	964,514
Michigan	Ordered	March 16	April 2	1,528,666
Minnesota	Ordered	March 18	April 23	875,021
Mississippi	Ordered	March 20	April 14	483,150
Missouri	Ordered	March 23	April 9	915,040
Montana	Closure expired	March 16	n/a	146,375
Nebraska	Ordered	March 23	April 3	319,194
Nevada	Ordered	March 16	April 21	473,744
New Hampshire	Ordered	March 16	April 16	180,888
New Jersey	Ordered	March 18	May 4	1,410,421
New Mexico	Ordered	March 16	March 26	336,263
New York	Ordered	March 18	May 1	2,729,776
North Carolina	Ordered	March 16	April 24	1,550,062
North Dakota	Ordered	March 16	May 1	109,706
Ohio	Ordered	March 17	April 20	1,710,143
Oklahoma	Ordered	March 17	March 25	693,903
Oregon	Ordered	March 16	April 8	606,277
Pennsylvania	Ordered	March 16	April 9	1,727,497
Puerto Rico	Ordered	March 16	April 24	365,181
Rhode Island	Ordered	March 23	April 23	142,150
South Carolina	Ordered	March 16	April 22	771,250
South Dakota	Recommended	March 16	April 6	136,302
Tennessee	Recommended	March 20	April 15	1,001,562
Texas	Ordered	March 23	April 17	5,360,849
Utah	Ordered	March 16	April 14	659,801
Vermont	Ordered	March 18	March 26	88,428
Virginia	Ordered	March 16	March 23	1,287,026
Washington	Ordered	March 17	April 6	1,101,711
West Virginia	Ordered	March 16	April 21	273,855
Wisconsin	Ordered	March 18	April 16	864,432
Wyoming	Closure expired	March 16	n/a	94,170

Notes: Data come from Education Week's "Coronavirus and School Closures" website, last updated on May 15, 2020. All closure dates refer to 2020.

Table A.4: Income, School Rurality and Race

	School-centered resources (1)	Parent-centered resources (2)	Google Classroom (3)	Khan Academy (4)
Post-Covid * Household income	0.10**	0.07***	0.06	0.04**
	(0.04)	(0.01)	(0.04)	(0.02)
Post-Covid * Fraction of schools in rural areas	-0.08***	-0.02**	-0.13***	-0.05***
	(0.03)	(0.01)	(0.04)	(0.02)
Post-Covid * Fraction of students who are Black	-0.07**	-0.04***	-0.00	-0.05***
	(0.03)	(0.01)	(0.03)	(0.01)
N	50,400	50,400	50,400	50,400

Notes: Heteroskedasticity robust standard errors clustered by DMA are in parentheses (* p<.10 ** p<.05 *** p<.01). Each column is a separate regression of the logarithm of the listed topic's search intensity on a post-March 1 indicator interacted with the DMA-level characteristics shown (and includes their direct effects). Income is measured per $10,000, while the remaining variables are scaled to interpret coefficients as predicted effects of a 10 percentage point increase in the interacted characteristic. The regressions are weighted by DMA population and include fixed effects for week of year (1-52) and school year (2015-2020). The sample contains search data from June 2015 through May 2020, excluding March 2020 and weeks with major school holidays.

Table A.5: Changes in Search Intensity, by Socioeconomic Status Quartile

	School-centered resources (1)	Parent-centered resources (2)	Google Classroom (3)	Khan Academy (4)
Post-Covid * Top SES quartile	0.87***	0.54***	1.07***	0.57***
	(0.13)	(0.03)	(0.13)	(0.06)
Post-Covid * 3rd SES quartile	0.55***	0.38***	0.78***	0.44***
	(0.09)	(0.04)	(0.07)	(0.05)
Post-Covid * 2nd SES quartile	0.39***	0.31***	0.72***	0.43***
	(0.07)	(0.02)	(0.11)	(0.04)
Post-Covid * Bottom SES quartile	0.23***	0.14***	0.46***	0.21***
	(0.06)	(0.02)	(0.06)	(0.04)
Top SES quartile	-0.17**	-0.11**	-0.15	0.01
	(0.08)	(0.04)	(0.11)	(0.06)
3rd SES quartile	0.06	-0.09*	-0.03	0.01
	(0.06)	(0.05)	(0.12)	(0.09)
2nd SES quartile	-0.14*	-0.22***	-0.37**	-0.07
	(0.08)	(0.07)	(0.14)	(0.05)
N	50,400	50,400	50,400	50,400

Notes: Heteroskedasticity robust standard errors clustered by DMA are in parentheses (* $p<.10$ ** $p<.05$ *** $p<.01$). Each column regresses the logarithm of DMA-level search intensity for the listed topic on a post-March 1 indicator interacted with indicators for each DMA's SES quartile, as well as directly including those SES quartile indicators. The regressions are weighted by DMA population and include fixed effects for week of year (1-52) and school year (2015-2020). The sample contains search data from June 2015 through May 2020, excluding March 2020 and weeks with major school holidays.

Table A.6: Changes in Search Intensity by Socioeconomic Status and Census Region

	Northeast (1)	Midwest (2)	South (3)	West (4)
(A) School-centered resources				
Post-Covid * High SES	1.12***	0.47***	0.42***	0.67***
	(0.08)	(0.07)	(0.08)	(0.08)
Post-Covid * Low SES	0.72***	0.24***	0.20**	0.63***
	(0.06)	(0.07)	(0.08)	(0.07)
High-Low SES Change	0.40***	0.23**	0.22*	0.04
	(0.12)	(0.11)	(0.13)	(0.13)
(B) Parent-centered resources				
Post-Covid * High SES	0.52***	0.45***	0.32***	0.54***
	(0.05)	(0.04)	(0.05)	(0.05)
Post-Covid * Low SES	0.41***	0.19***	0.11***	0.42***
	(0.05)	(0.04)	(0.03)	(0.04)
High-Low SES Change	0.11	0.26***	0.21***	0.13**
	(0.07)	(0.05)	(0.06)	(0.06)
(C) Google Classroom				
Post-Covid * High SES	1.33***	0.72***	0.82***	0.85***
	(0.08)	(0.10)	(0.13)	(0.09)
Post-Covid * Low SES	1.01***	0.42***	0.48***	0.75***
	(0.06)	(0.09)	(0.10)	(0.07)
High-Low SES Change	0.32**	0.30**	0.34**	0.10
	(0.12)	(0.15)	(0.17)	(0.13)
(D) Khan Academy				
Post-Covid * High SES	0.73***	0.43***	0.29***	0.56***
	(0.01)	(0.08)	(0.04)	(0.10)
Post-Covid * Low SES	0.62***	0.31***	0.13***	0.56***
	(0.04)	(0.09)	(0.05)	(0.05)
High-Low SES Change	0.11***	0.12	0.16**	0.00
	(0.04)	(0.08)	(0.06)	(0.11)
N	5,520	13,680	20,880	10,320

Notes: Heteroskedasticity robust standard errors clustered by DMA are in parentheses (* $p<.10$ ** $p<.05$ *** $p<.01$). Each column in each panel regresses the logarithm of DMA-level search intensity for the listed topic on a post-March 1 indicator interacted with indicators for above and below median DMA-level SES in that Census region, as well as directly including an above median SES indicator. Below each pair of interaction coefficients is the estimated difference between the high and low SES interaction coefficients. The regressions are weighted by DMA population and include fixed effects for week of year (1-52) and school year (2015-2020). The sample contains search data from June 2015 through May 2020, excluding March 2020 and weeks with major school holidays.

Table A.7: Changes in Search Intensity, Excluding Low Search Areas

	School-centered resources (1)	Parent-centered resources (2)	Google Classroom (3)	Khan Academy (4)
(A) Nationwide				
Post-Covid	0.52***	0.35***	0.76***	0.41***
	(0.05)	(0.02)	(0.05)	(0.03)
(B) By median SES				
Post-Covid * High SES	0.74***	0.48***	0.94***	0.51***
	(0.08)	(0.03)	(0.08)	(0.05)
Post-Covid * Low SES	0.30***	0.21***	0.59***	0.30***
	(0.05)	(0.02)	(0.07)	(0.03)
High SES	-0.04	-0.12***	0.07	0.01
	(0.06)	(0.03)	(0.10)	(0.06)
High-Low SES Change	0.44***	0.27***	0.35***	0.21***
	(0.10)	(0.03)	(0.11)	(0.05)
(C) By income, online access, race				
Post-Covid * Household income	0.15***	0.09***	0.13***	0.08***
	(0.03)	(0.01)	(0.03)	(0.02)
Post-Covid * Broadband penetration rate	0.43***	0.34***	0.36***	0.30***
	(0.08)	(0.03)	(0.09)	(0.05)
Post-Covid * Computer penetration rate	0.54***	0.46***	0.47***	0.38***
	(0.10)	(0.04)	(0.10)	(0.07)
Post-Covid * Fraction of schools in rural areas	-0.18***	-0.10***	-0.19***	-0.10***
	(0.03)	(0.01)	(0.03)	(0.02)
Post-Covid * Fraction of students who are Black	-0.09***	-0.04***	-0.02	-0.06***
	(0.03)	(0.02)	(0.04)	(0.02)
N	45,845	40,829	45,682	44,240

Notes: Heteroskedasticity robust standard errors clustered by DMA are in parentheses (* $p<.10$ ** $p<.05$ *** $p<.01$). In panel A, each column regresses the logarithm of the listed topic's search intensity on a post-March 1 indicator. Panel B interacts the post indicator with above and below median SES indicators. Below each pair of interaction coefficients in panel B is the estimated difference between the high and low SES interaction coefficients. Each coefficient in Panel C comes from a separate regression that interacts the post indicator with DMA-level household mean income or the other listed resource and demographic measures. Income is measured per \$10,000, while the remaining variables are scaled to interpret coefficients as predicted effects of a 10 percentage point increase in the interacted characteristic. The regressions are weighted by DMA population and include fixed effects for week of year (1-52) and school year (2015-2020). The sample contains search data from June 2015 through May 2020, excluding March 2020 and weeks with major school holidays, as well as DMAs with low search intensity in any given week.