Browsers Don't Lie? Gender Differences in the Effects of the Indian COVID-19 Lockdown on Digital Activity and Time Use

Amalia R. Miller
University of Virginia
IZA and NBER

Kamalini Ramdas
London Business School

Alp Sungu
London Business School

July 2021

Abstract

We study gender differences in the impact of the initial COVID-19 lockdown in India on internet activity, using primary data collected during the lockdown via an online survey that ended with consensual sharing of internet browser history records spanning the prior 90 days. The browser data captured online activity before and during the lockdown, without reporting or recall biases. We find that online activity, as measured by time or clicks, increased substantially for both men and women during the lockdown, but that men's online activity increased significantly more. The gender gap in response to the lockdown is present both overall and within particular categories, such as production, leisure, and job search, and it is driven mainly by parenthood. Although men were spending more time online, they also reported significantly larger increases in time devoted to childcare in our survey than women did. Female respondents did not report the same about their spouses, which suggests that gender differences in reporting might be biasing the survey results. This underscores the value of examining objective browser histories to examine time use.

Keywords: COVID-19; gender; online activity; time use; browser data; privacy

We thank the Wheeler Institute for Business and Development at London Business School and the Institute of Labor Economics (IZA) Gender Growth and Labour Markets in Low Income Countries Programme (G2LM|LIC) for financial support; our partners PowrofYou and Dynata for assistance with data collection; and Sanjay Jain, Selin Kesebir, Carmit Segal and participants at the 2020 IZA G2LM|LIC annual research conference for helpful comments. This document is an output from a project funded by the UK Department for International Development (DFID) and IZA for the benefit of developing countries. The views expressed here are not necessarily those of DFID or IZA. Human subjects IRB approval from London Business School (REC-651).

I. Introduction

Although lockdowns imposed to address the COVID-19 pandemic were widely expected to impose more severe burdens on women than on men (United Nations 2020), empirical evidence on the differential effects on time use has been limited because of difficulties in collecting reliable survey data during lockdowns (Egger et al. 2021). We address this challenge by collecting unique online data from over a thousand adults in India that includes two key components: 1) an internet browser history tracing each individual's detailed online activity over the prior 90 days and 2) a survey covering demographics, employment, and time use topics. We fielded our survey during the initial lockdown in India — which started on March 25, 2020 — and collected information on over 30 million website visits. Together, these elements provide a unique perspective on how men and women adapted differently to the pandemic.

Highlighting the greatly increased value of digital access, we find that browser use increased significantly during the lockdown, across a range of activities, including production, leisure, job search, and human capital investment. Men and women both increased their time online. But the increases were significantly larger for men across a range of online activities. The gender gap is larger among parents, suggesting that additional household obligations during the lockdown disproportionally consumed mothers' time.

When we split the sample by working status, we find that among full-time workers, the gender gap in leisure time is more prominent, while the gap in productive activities is larger among those not in full-time jobs. This suggests that working women were sacrificing online leisure to maintain productive time use, while other women were forgoing potential earnings opportunities. Worryingly, we also find an indication that the gendered effects of the lockdown may lead to persistent economic effects: women show a significant decline in online job search activity relative to men, a gap that is particularly large among job seekers.

Our findings of relative declines in women's online activity contribute to the emerging body of work examining gender differences in the effects of the COVID-19 pandemic, and particularly to empirical research on how lockdowns have differentially affected men's and women's time use. The fact that women typically provide the bulk of unpaid domestic labor is a major source of gender inequality in economic outcomes (e.g., Fuchs 1988; Miller 2011). It is also the source underlying concerns raised by policymakers, scholars, and advocates that pandemic school closures and lockdowns — which reduced access to paid domestic service providers and other market substitutes for domestic production — would have more adverse effects on women, and especially on mothers (Alon et al. 2020; Burki 2020). Despite the importance of these concerns, it has been difficult to quantify the gendered effects of the pandemic on time use with available data sources. We contribute in four ways.

Our first contribution is to expand the evidence on gender differences in self-reported time use from our online survey fielded during the pandemic. Researchers attempting to quantify the pandemic's effect on time have typically relied on self-reported survey data (e.g., Giurge, Whillans, and Yemiscigil 2021), and some have asked retrospective questions to obtain a pre-pandemic baseline (e.g., Adams-Prassl et al. 2020 and the ongoing UN Women Rapid Gender Assessments). Our survey included retrospective questions to track changes and further asked about partner's time use to create two subjective measures of gender differences. Longitudinal time use studies that were also fielded before the pandemic can rely entirely on contemporaneous reports from the same individuals (Zhou et al. 2020), but these surveys are rare and they still depend on subjective reporting.[1]

Our second contribution is that we combine our survey with browser histories to obtain objective and detailed external records of activity that avoid recall bias and misreporting issues in subjective reports. This is similar in spirit to the approach taking in studies focused on the effects of the pandemic on worker productivity that combine surveys with objective public measures of individual output to address concerns about subjective self-reports.[2] In our survey, male respondents reported significantly *larger* increases in childcare time during the lockdown than did women. Although this pattern echoes the finding in Zhou et al. (2020), where self-reported housework time increased by more for men (3.5 hours) than for than women (3 hours) at the onset of UK lockdown, it is not reflected in responses to questions about spousal time use in our survey. There we see no relative increase in time that women report their husbands are spending on children. Although this inconsistency could come from differences across households, the fact that men reported greater childcare time while also spending significantly more time online raises concerns about the reliability of the self-reported time use measures. A particular concern is that men and women might perceive or report their activities differently, as suggested by prior work finding that men tend to overreport their household production time (Kan and Pudney 2008).

Our third contribution to understanding the gendered effects of the pandemic is our examination of novel outcomes related to internet activity, made possible by the exceptional depth and detail in browser data relative to surveys. Because we observe every individual web page that is opened, we are able to go beyond overall summary information about usual time allocations to

[1] Detailed time diaries are more reliable but also onerous, and difficult to collect during the lockdowns. The American Time Use Study was suspended between March 19 and May 11, 2020.
[2] For example, Cui et al. (2020) and Myers et al. (2020) find relative declines in academic research productivity for women, with the latter study reporting particularly stark effects on mothers.

examine variation across a range of specific categories, including specific activities such as video watching or job search. We can also measure activity and compare differences across and even within days.

Fourth, by studying data from India, we shed light on a developing country where data collection challenges during pandemic lockdowns have been particularly severe and where research about the internet overall has been constrained by restrictive data access and limited availability.[3] India is also an important country for the analysis of gender differences in time use, because of the prevalence of gender discrimination and range of barriers to women's labor force participation (e.g., Heath and Tan 2020; Hyland et al. 2020).[4] Another benefit of studying India is that we can examine the impact of the initial nationwide lockdown, which was exceptional in its scale and scope. The lockdown was arguably the world's largest (in terms of population affected) and strictest.[5] It precluded all non-essential workers from stepping out of their houses and closed all educational institutions (MHA 2020).

By providing new insight into the differential effects of the pandemic on men and women, this paper also adds to the broader literature in economics on gender differences in time allocation (e.g., Becker 1985; Blau and Kahn 2017). We do this in part by exploiting the exogenous shock induced by the strict national COVID-19 lockdown in India. We also contribute by examining a novel set of outcomes related to internet use. The internet is a technology that profoundly affects people's lives, but it remains an area relatively unexamined by researchers interested in gender, particularly in developing countries. Fletcher, Pande and Moore (2017) recommend time use

[3] https://www.epw.in/engage/article/where-data-study-internet-india
[4] India ranks 140th out of 156 countries on the World Economic Forum's 2021 Global Gender Gap Index, available at http://www3.weforum.org/docs/WEF_GGGR_2021.pdf
[5] For comparisons of lockdown stringency, see, e.g., https://ourworldindata.org/grapher/covid-stringency-index?tab=table&stackMode=absolute&time=2020-03-25&region=World.

surveys as a tool to understand how Indian women spend their time and how much they are engaging in the labor market. We offer a detailed and objective view on this issue.

Finally, this study contributes to the literature using online and digital data to study behavior generally, and in response to the COVID-19 pandemic in particular.[6] Like other studies, we leverage digital footprints left by individuals in their ordinary activities. However, we differ from much of the literature that examines aggregate or de-contextualized data in that we explicitly ask individuals to grant us one-time access to obtain a current snapshot of their recent browser history. Responses to our survey provide context for the digital histories of individual respondents. Although this approach has obvious limitations – the scale is smaller, and the sample may be less representative because individuals self-select into it – there are also important advantages.

The practical advantage is that we are able to link the browser information with survey information, which is precisely what enables us to study gender, family status and employment. More fundamentally, our data collection approach represents another way to balance between the competing interests of detailed digital data collection and protecting the privacy of individual users. Rather than embedding or exploiting trackers on individual computers, we worked in partnership with *Powrofyou*, a technology platform that emphasizes consensual and minimally invasive digital data sharing. To the extent that privacy concerns affect the willingness of individuals to join the study (e.g., Athey, Catalini and Tucker 2017; Lin 2021), recruiting costs will be higher in such an approach, and the sample less representative. We can offset these challenges by collecting anonymized histories that enable us to study changes in the activities of individuals over time.

[6] For example, Bacher-Hicks et al. (2020) study online learning in the US using aggregate search trends and DeFilippis et al. (2020) examine de-identified, aggregated email and meeting meta-data. Smartphone geolocation data has been put to great use in academic studies (e.g., Chiou and Tucker 2020; Chen, Chevalier, and Long 2020) and as part of surveillance and prediction models for the pandemic.

The fact that our data collection is entirely backward-looking has the ethical advantage of increasing the control that respondents have over the extent of data sharing. In countries with weaker institutions and less oversight on how data is used, data access may come hand in hand with ethical concerns. It also addresses the methodological concern that forward-looking data collection with informed consent could affect online activity, for example, if people alter their online behavior when they know they are being tracked for a study. As concerns about digital privacy increase in prominence among regulators and the public (Schwartz 2019; Goldfarb and Tucker 2019; Acquisiti, Taylor and Wagman 2016), and data privacy laws are increasingly being adopted in developing countries,[7] the importance of considering alternative models for ethical and privacy-protecting digital data collection will increase as well.

II. Data

A. Primary Data Collection

We collaborated with *Powrofyou*, an internet-browser analytics platform, and *Dynata*, a global first-party data platform to field the survey between mid-May and early June 2020. Individuals drawn from *Dynata*'s marketing pools in India were invited to participate in the survey and directed to an online survey that ended with a consensual browser data upload using the *Powrofyou* software. Participants with valid data were compensated for their effort. *Powrofyou* has an internet browser extension that collects retrospective data stored in each user's browser account history. These records cover up to 90 days of recent activity on a particular browser account. Activity is collected for individuals across their electronic devices (e.g., computer, smartphone, tablet), as long as they are logged into the same user account and using the same browser type (e.g., Google

[7] https://unctad.org/page/data-protection-and-privacy-legislation-worldwide

Chrome). The data include observations from every website visit, including the URL (uniform resource locator, i.e., web address) and timestamp of the visit.[8] Although the browsing history data covers usage spanning multiple electronic devices, the data collection process itself needed to be conducted on a personal computer, because of technical requirements of the *Powrofyou* program. We do not collect information from private browsing or Incognito mode, and personal identifiers are removed.

Each URL has an associated *title*, which conveys additional meaningful information, such as a Google search phrase, the headline of a newspaper article, or a YouTube video title. Using the URL, title, and timestamp for each website visit, *Powrofyou* provides a detailed categorization scheme and calculates the session's usage duration in seconds. The categories are obtained from the Google Cloud Platform, using a Natural Language Processing algorithm that makes a probabilistic category assignment to each website domain.[9] In addition to these categories, we also code websites as being primarily related to *leisure* (entertainment) or *production* (non-recreational) uses, based on these categories.[10]

We obtained data that met our quality control standards from 1,118 individuals aged 22 to 54 located in 28 states across India. We prevented individuals from participating with data from a new browser account or a secondary browser type that is not used regularly by requiring that the browser history include at least 30 days of data. We dropped one user who preferred not to state

[8] The software only captures retrospective data. Once the data transfer is over, it automatically deletes itself and redirects participants to the survey platform.
[9] The universe of categories can be found at https://cloud.google.com/natural-language/docs/categories.
[10] Our *Leisure* category includes Adults, Arts & Entertainment, Games, Online Communities (e.g., social media), and Shopping. *Production* category includes Business & Industrial, Computers & Electronics (e.g., digital repositories), Finance, Internet & Telecom (e.g., e-mail and search engines), Jobs & Education, Law & Government, News, Science, and References (e.g., Calculator). Other Google Cloud categories combined covers 0.8% of our data. Some websites – such as spam webpages – are labelled as "other". Median "other" category usage on a day covers 7% of total time use.

their gender. We took two steps to ensure that respondents are human and not computer bots. First, we included an attention test question in the survey, aimed at verifying a human is reading the questions. However, because the question was multiple choice, some bots are expected to get through by chance. We therefore also dropped all users with more than an average of 3,000 URL visits per day. Inspecting the activity of the 19 users who failed the latter requirement confirmed that they are unlikely to be human. For example, on certain days, they repeatedly visited the same handful of business websites, refreshing every 5 seconds throughout the day.

In total, we collected over 31.5 million webpage visits to 134,123 unique websites in our data. We aggregated these data to the daily level for each participant, using different categories of activity. We also limited our analytical sample to the period between February 22 and May 10, 2020, to avoid dates with few observations, coming from the slightly staggered enrollment timing. Our final dataset includes 81,929 days of individual browser usage data with 52,690 days coming from 702 men and 29,239 days from 396 women.

B. Sample Composition

Table 1 reports summary statistics on the composition of our sample. Although we targeted equal gender balance, 64% of our respondents were male. This may come from the requirement that respondents use a computer to complete our survey, combined with the overall gender gap in digital access. The requirement for computer access likely also contributes to the other ways in which the respondents are not representative of the general population of India. Particularly striking is the high educational attainment in our sample, with over 90% of both men and women being college graduates. Full-time employment (including self-employment) is somewhat low for men, at 77%, but higher than for women (64%). The men in our sample are also more likely than the women to be married (64%) and have children (61%), but the majority of the female respondents are married (60%) and have children (58%).

The lack of representativeness of this sample for the overall population of India is less of a problem for the internal validity of the within-person changes that we measure (using objective data that does not depend on personal recollection).[11] However, to the extent that the impact of the lockdown varied across individuals, sample selection will mean that the averages that we compute, and the differences by gender in those averages, will be affected by the composition of the sample. This means it is important to interpret the estimates as applying to individuals of the type that would and could complete this survey. This is a relatively advantaged group within India that is literate in English and that has access to an internet-connected computer. Our estimates would clearly not apply to populations with low literacy or minimal digital access. Nonetheless, they provide insight into an interesting and rapidly growing subpopulation.[12]

C. Browser Data Raw Trends by Gender

The average user in our sample spent about 4 hours a day using the browser, of which over an hour was spent on YouTube watching (on average 5 different) videos, and which includes an average of 5 Google searches (Table 2).

Both men and women increased their online time use during the lockdown relative to the period immediately prior. However, the increase for women was smaller than for men. This is

[11] We should also note that the difficulty of recruiting a representative sample, not unique to this study, has been greatly exacerbated by the COVID-19 pandemic and lockdowns. Response rates have been shown to have declined overall, and in systematic ways related to observable characteristics, even in well-established surveys in the US, such as the Current Population Survey used to compute official unemployment rates (Heffetz and Reeves, 2020). The challenges have been immense in lower-income countries (e.g., Egger et al. 2021).

[12] Computer penetration in India is estimated to be about 3 percent and growing about 15 percent a year. Source at https://www.idc.com/getdoc.jsp?containerId=prAP45648319

apparent in Figure 1, which plots daily average browser time use by men (dashed blue line) and women (solid red line) in our sample. The blue shaded region reflects the pre-lockdown period. Time online was similar between men and women at the start of the sample, including the pattern of daily fluctuations. Men started to show slight increases in usage right before the Indian lockdown, following the March 11, 2020, official World Health Organization (WHO) declaration of the COVID-19 outbreak as a global pandemic. During the lockdown, men and women both gradually increased their browser time, but the increase for men is noticeably greater. This is the basic pattern that we examine in our regression analyses.

When we examine internet use by purpose, we find that users spend an average of about 2 hours a day on leisure and an hour and a half on production websites (Table 2). We see sizable increases in both activities for men and women (p<0.01), but the increases in leisure are larger in both absolute terms and relative to pre-lockdown mean. These changes are also shown in daily plots by category in Appendix Figure A1.[13] The general increase in leisure time online could reflect a common pattern of increasing leisure time during periods of economic hardship (as found in Aguiar, Hurst, and Karabarbounis 2013 during the Great Recession). It could also reflect a shift to online leisure to replace the variety of offline social and leisure activities prevented by the lockdown. Similarly, some of the increased productive time online may be from a shift to remote work during the lockdown, which affected 44.5 percent of our sample (Table 1). Some productive activity took place offline or without a browser. Nevertheless, it is reassuring that production is

[13] Figure A1 shows relative declines in women's time online across a range of outcomes. The category of online shopping is worth mentioning because it confirms the usual pattern of a relative decline in women's time but is unusual in that women's usage greatly exceeded men's in the pre-lockdown period. The sharp drop in shopping activity from the outset of the lockdown is consistent with the severity of the Indian lockdown that prevented deliveries of goods to homes.

the unique browsing category in our data showing a strong cyclical pattern of weekly usage with regular drops on Sundays (Figure A1).

We note two exceptions to the general pattern in Table 2. The first is online job search, which we revisit in Section III.C. It resembles other categories in that there is a relative increase in men's time use but differs in women showing both relative and absolute reductions during the lockdown. The second is online learning time. This category shows large increases for both men and women that are not statistically distinguishable from one another.[14]

III. Estimated Gender Differences in the Effects of the Lockdown on Browser Activity

This section presents results from a series of simple regression models that capture the differential effect of the lockdown by gender with panel data and two-way fixed effects. We split the observations by time into the pre-lockdown baseline until March 24, 2020, and the lockdown period starting on March 25, 2020. This date corresponds to the first national COVID-19 lockdown in India. It was imposed suddenly and strictly curtailed activities outside the home.[15] Our unit of analysis is a person-day and our estimation equation takes the form:

$$Y_{it} = \beta\ Lockdown_t \times\ Female_i +\ \gamma_i + \delta_t +\ \epsilon_{it} \tag{1}$$

Y_{it} is the outcome of interest for individual i on date t. $Lockdown_t$ is a binary variable indicating that date t occurs during the lockdown, and $Female_i$ is a binary variable equal to 1 if individual i is female. γ_i is a vector of individual fixed effects and δ_t is a vector of date fixed effects. Standard errors are clustered at the individual level. Our coefficient of interest is β, which captures the

[14] The coding in Table 2 is based on domain names. If we instead examine time spent on YouTube videos in the "educational" category, we also find increases for both genders with no significant difference between them (Table A8).

[15] The first lockdown announcement was made on March 24, 2020, and the lockdown started after the midnight on that day. The official guidelines are at https://www.mohfw.gov.in/pdf/Annexure_MHA.pdf

differential average impact of the lockdown on women relative to men. We use a natural logarithmic transformation on our outcomes, after adding 1 second to all daily observations to retain zero values. Because there is only a single cross-section, our survey does not provide time-varying information that we can include as controls in our models. Rather, we use the survey responses to identify gender and other groupings for our examinations of heterogeneous effects.

A. Internet Browser Activity

The regression estimates in Table 3 show significant relative declines in women's online time use following the imposition of the lockdown. This confirms the patterns in the raw data, after accounting for the individual and date fixed effects and employing the log transformation. The decline is sizable and significant across the key outcome measures. Women's total browser time dropped by 25.5 percent compared to men's, which translates into nearly half an hour less time per day.[16] Women's online time use decreased relative to men's by 32.1 percent for leisure and by 28.4 percent for production websites. These relative declines found in time-based measures are also present for the count-based measures in Table 4. Women's daily count of unique URLs visited dropped significantly relative to men's, with a 24.6 percent relative drop in URLs, amounting to about 40 fewer URLs visited per day. When we examine website domains, we observe a significant usage drop for women, relative to men, for video streaming (YouTube), social media (Facebook) and search engine (Google) websites, as measured in time (Table 3) or counts (Table 4).

Although we aggregate activity up to a 24-hour period for most of our analysis, we also explored variation in the overall impacts by time of day. In particular, we divided each day into twelve 2-hour intervals and ran separate regressions on total browser time use for each interval.

[16] The outcome is logged, so the coefficient of -0.294 implies a change of -25.5% = 100*($e^{-0.294}$-1).

The results are in Figure 2 with estimates starting from 6 AM on the left. The effects are largest midday and in the late evening. These times coincide with lunch and dinner times, both typically hot meals in an Indian household. South Asian women bear greater cooking costs than men do (Duflo, Greenstone and Hanna 2008; Miller and Mobarak 2013; Dhar, Jain and Jayachandran 2018). Because of these social mores, women in our sample are more likely than men to be involved in meal preparing, service and washup around these times.

We also conducted several robustness checks to confirm our main results.

First, we checked that the procedure of adding 1 second to avoid dropping zeros was not affecting the estimates. Appendix Table A1 shows estimates for the extensive margin of time use by using an indicator for any browser use of the specified type as the outcome variable. We find no extensive margin effect for overall browser activity in a day. The other outcomes show significant negative estimates that point in the same direction as the main estimates in Table 3. Appendix Tables A2 and A3 focus on the intensive margin by applying the log transformation without first adding 1. The effects again point to relative declines in time use (Table A2) and visits (Table A3) across all categories, though the estimates are not statistically significant for time on Facebook or conducting Google searches.

In our second robustness check, we considered the possibility that the relative increase in men's time use is coming from them more frequently sharing their electronic devices with others in the household during the lockdown. Because of this concern, we included in our survey questions about device sharing and are therefore able to estimate separate effects for the sub-group of people who do not share their smartphone, computer, or tablet with anybody else. Consistently across all regressions, we find larger effect sizes for the sub-group who do not share their electronic devices. On this sample, women's total time online drops 40.7 percent relative to men's, whereas the full sample estimate is 25.5 percent. This difference suggests that women in the overall sample are sharing their electronic devices more intensively than men, and a greater share of their browser activity is being consumed by others. In that case, our full-sample estimates are conservative measures of the relative decline in women's time online.

Because YouTube accounts for almost 20 percent of total browser time in our sample, our final robustness check further parses the video content of 308,497 unique YouTube URLs using Google Cloud's YouTube Data API.[17] Rather than assigning all time on YouTube to the leisure category, we use YouTube video categories to identify content that is more leisure- or production-related. Confirming our initial categorization, two thirds of YouTube time is devoted to leisure. The results of our main analysis are unchanged if we revise the category-level usage by moving YouTube content tagged as productive into that category as well (Appendix Table A7). We are also able to confirm that the pattern of results from all browsing data is also present within YouTube videos: women's time devoted to both leisure and production videos drop considerably relative to men's during the lockdown (Appendix Tables A8 and A9).

B. Heterogeneous Effects by Family and Employment Status

The relative decline in women's online activity is consistent with the hypothesis that women experienced a greater increase in caretaking responsibilities and household obligations after the lockdown that prevented them from spending as much time online. A natural implication of this line of reasoning is that the gender gap in the impact of the lockdown would be larger for parents, who experienced greater shocks to household production. We investigate this prediction

[17] For each URL, the YouTube API returns an array of information about the video, such as the video category, description, and channel name. Video categories are not shared with viewers. Details at https://developers.google.com/youtube/v3. We collected the YouTube API data on 2021 and information about the videos that were removed from YouTube were not available.

by splitting the sample based on parental status (summary statistics for the sub-samples are in Appendix Table A6).

Table 5 presents the separate estimates for individuals with at least one child and for those with no children. We observe a significant drop in total, leisure and production time use for mothers, relative to fathers. Among childless adults, we find no significant gender differences in any of these measures. Although the coefficients have the same sign, the estimates are smaller and noisier on the childless sample. The difference between the two samples is greatest (and statistically significant) for leisure time. Mothers experience a relative drop in online leisure of 46.6 percent ($1 - e^{-0.627}$) compared to fathers, while childless women have a relative drop of less than 5 percent. The disproportionate effect of the lockdown on mothers is primarily manifesting in our data as a reduction in leisure time.

We next split our sample by employment status, to test the prediction that the effects are stronger for women who lack full-time employment and have less economic power and autonomy. Consistent with the prediction, the estimates for total time use and production time are smaller and less significant in the full-time employed sub-sample (columns 1 and 3 of Table 6) than in the sample of individuals not employed full-time (columns 4 and 6). Nevertheless, we do find significant gender differences among full-time workers in leisure time. Relative to full-time employed men, women employed full time had a substantial and significant 43.7 percent decrease in leisure time online. This suggests that the increased household burden that women faced because of the lockdown was not limited to those working part time or less; women with full-time jobs were also affected. In the sample of part-time and non-employed individuals, we see no significant gender gap in the impact of the lockdown on leisure time online. Instead, that sample shows a significant 49.1 percent drop in women's production time online. This pattern is consistent with full-time employed women having less flexibility than other women to reduce their production time online and choosing instead to sacrifice leisure time. It is also consistent with women with weaker ties to employers being less capable than similarly situated men of expanding their productive time online during the lockdown.

C. Effects on Online Job Search

We next consider differences in online job search, an activity that can have lasting effects on labor market outcomes and economic wellbeing. Over three-quarters of job applications worldwide are submitted online (Statista, 2020). India has a growing online job market, which remained active during the lockdown, during which time in-person networking and job applications were strictly disallowed. Our URL-level browser data – covering periods before and during the lockdown – allow us to examine changes in online job-seeking behavior.

We did this by creating a comprehensive list of 60 job search websites frequented in India and classifying website visits in our sample as relating to job search if their URL domain is included in our list of job search websites. We defined two daily outcome variables for online job search. The first is a binary variable indicating whether or not a person visits any job-search websites on that day. The second is a measure of time spent on job search sites.

We find striking gender differences in the impact of the lockdown on job search in the overall sample. While men's time devoted to online job search increased by about 40 percent during the lockdown, women's job search time decreased by a similar amount (Table 2). Regression estimates for the full sample in Table 7 show significant relative decreases in women's online job search relative to men's during the lockdown: a 2.4 percentage point drop on the extensive margin (column 1) and a 13.7 percent decrease in duration (column 2).

One concern with investigating job-seeking behavior on the entire sample is that the lockdown may have caused more unemployment for men than women.[18] In that case, the relative increase in men's time devoted to job search may come from men's greater need for work rather than women's increased household obligations. We address this by identifying individuals who are more likely to be job seekers throughout the 90-day lookback window because they did not have a full-time job and had no change in employment status over the 90 days preceding their survey date. This sub-sample comprises only about a quarter of our full sample. Nevertheless, we detect statistically significant and larger decreases in both measures of job search activity for women relative to men on this sub-sample (columns 3 and 4).

These results for job seeking are particularly concerning as they could exacerbate gender gaps in Indian labor markets beyond the pandemic. Women's full-time labor force participation in India is low and has seen a decline in recent years despite the country's economic growth, lower fertility rates and higher education levels for women. Women in India often lack access to information about available jobs, and search for jobs less efficiently than men, leading to a mismatch between the jobs sought and jobs available (Fletcher, Pande and Moore 2017). Women in our sample stated that it was more important to invest in their partner's careers while men stated the reverse. The job seeking patterns we document during the lockdown may both reflect and reinforce these preferences and social norms.

IV. Gender Differences in Effects of the Lockdown on Household Production

Our focus in this paper is on internet browser activity, where we have the most robust data. Because housework and childcare activities are difficult to capture with browser data alone, we also asked

[18] In our sample 4.3 percent of men and 3.5 percent of women lost a job during the 90-day period prior to the data collection.

survey questions about time use for those categories. Although time-use data can suffer from biases related to subjective reporting and imperfect recall that are not present in the browser data, they may provide direct evidence of the hypothesized mechanism underlying women's relative decreases in productive and leisure time online, a relative increase in women's time spent on household duties. However, that is not what we find.

Table 8 presents the results of our analysis of survey-based measures of time use, comparing men and women, before and during the lockdown. We asked about two time periods in the survey – present and pre-lockdown – and separately covered own and partner's (if married) usual daily time spent on childcare (if they have children) and housework. We converted our interval measure of time use (using 2-hour buckets) into a continuous measure by taking the mid-point of each bin and assigning 10 hours to participants who selected 8 or more hours. For each outcome, we report estimates for own time use in the first column and for partner's time use in the second, using a common sample of married individuals.

Men reported spending an average of 2.6 hours on childcare (constant term in column 1) and 2.6 hours on housework (column 3) per day in the pre-lockdown period. Women reported spending 0.6 more hours than men on childcare (*Female* coefficient in column 1) and 1.2 more hours on housework (column 3) during the pre-lockdown period. In the pre-lockdown period, the gender difference is also consistent between self-reports for own time use and the corresponding self-reports on partner's time use: women reported that their partners devoted less time to both childcare (1.2 hours less, column 2) and housework (1.9 hours less, column 4) than men reported about their partners. The positive and highly significant lockdown coefficients show sharp increase in self-reported measures of both childcare and housework time, which is not surprising given the conditions of the lockdown.

What is surprising is that men report significantly larger increases in their own time devoted to childcare during the lockdown than women do. Men reported spending over 1.5 hours more per day on childcare, double the increase reported by women (0.755 = 1.52 – 0.765; column 1). With this additional 46 minutes a day, men report devoting as much time as women do to childcare during the lockdown.

One concern about this result is that the relative increase in men's self-reported time spent with children is not matched in the reports from partners. Women and men reported nearly identical increases in their partners' time spent on childcare (Table 8, column 2) and the gender differences remained highly significant during the lockdown. While it is true that the men and women in the sample are not necessarily married to one another, the inconsistency between the two measures casts doubt on the reliability of the self-reported relative increase in men's time with children. The relative increase in male household production is also limited to childcare. There are no significant gender differences in housework time during the lockdown for either own or partner's time. These features suggest that the gender difference in the impact of the lockdown on self-reported time devoted to children could derive in part from men and women differing in how they define time spent caring for children and what types of activities that includes or excludes (as discussed, e.g., in Kan and Pudney 2008).

One possibility that would reconcile men's increased time online and time with children would be if men devoted more of their time to child-related browsing, such as browsing child-targeted content with their child. We found no empirical support for this in our data. Because we are not able to identify internet activity that is shared with children, we focused on identifying webpages and videos that are aimed at children using textual analysis of webpage titles and YouTube video descriptions and channels. Across our various approaches to measuring time spent browsing childcare-related content, we found small (ranging from < 10 to 50 seconds) and statistically insignificant gender differences in the effect of the lockdown on child-related internet use (Appendix Table A10).[19]

Another possibility that could reconcile the findings is that men spent more time consuming online content while caring for their children during the lockdown. It could have been content that was not focused primarily on children but that they consumed together with their children. It is also possible that men were pursuing their own leisure and productive activities online while caring for their children. This could produce the gender differences we observe if men are more likely to describe time spent on a device engaged in activities unrelated to children, when in the vicinity of children, as "childcare" time, while women tend to reserve that term for time actively spent caring for children (feeding, cleaning, teaching, etc.) or supervising their activities. Although the browser data are not able to resolve this conflict, they do suggest that the simple self-reported data on childcare time use paint an incomplete picture at best.

V. Conclusion

All over the world, with the curtailment of face-to-face activities during the pandemic, the internet has been an important avenue for production and leisure activities. Through a combination of survey data and consensually obtained anonymized browser histories starting prior to the initial Indian lockdown in March 2020, we gained a unique window into how people's lives have changed during the pandemic, and in particular how men's and women's time use has been differentially affected by the lockdown.

[19] Two of the approaches (manual and *Word2Vec* word embedding-based dictionaries) use textual analysis of webpage titles and YouTube descriptions, and the third focuses on YouTube channels that exclusively produce content for children. See Online Appendix B for a detailed description of the procedures.

Even among our relatively privileged sample of highly-educated individuals in India who owned computers and had internet access, we find significant declines in women's internet use, relative to men's, during the lockdown. These differences are present overall and across a range of leisure and productive activities. The gender difference is larger for parents, particularly in the domain of leisure time, which suggests changes in online activity may come from differential caretaking burdens on women that have been exacerbated by the pandemic. In contrast to the browser data, our time-use survey results show men reporting a relative increase in time spent caring for children, but this is not echoed in reports from women about their spouses.

In addition to providing evidence on how the immediate effects of the initial COVID-19 lockdown in India differed by gender, our results also suggest effects that may persist beyond the pandemic. The large reduction full time working women's online leisure time, relative to their male counterparts, may not be sustainable and have long term consequences, such as burnout, that drive women to part time work or to exit the labor force. The relative decrease in women's online job seeking activity is similarly concerning, as it could produce lower employment rates and worse job matches for women after lockdowns.

References

Acquisti, A., C. Taylor, and L. Wagman. 2016. "The Economics of Privacy." *Journal of Economic Literature* 54(2), 442–492.

Adams-Prassl, Abi, Teodora Boneva, Marta Golin, and Christopher Rauh. 2020. "Inequality in the Impact of the Coronavirus Shock: Evidence from Real Time Surveys." *Journal of Public Economics* 189: 104245.

Aguiar, Mark, Erik Hurst, and Loukas Karabarbounis. 2013. "Time Use During the Great Recession." *American Economic Review* 103 (5): 1664–96.

Alon, Titan M., Matthias Doepke, Jane Olmstead-Rumsey, and Michele Tertilt. 2020. "The Impact of COVID-19 on Gender Equality." No. w26947. National Bureau of Economic Research.

Athey, Susan, Christian Catalini, and Catherine Tucker. 2017. "The Digital Privacy Paradox: Small Money, Small Costs, Small Talk." No. w23488. National Bureau of Economic Research.

Bacher-Hicks, Andrew, Joshua Goodman, and Christine Mulhern. 2021. "Inequality in Household Adaptation to Schooling Shocks: Covid-Induced Online Learning Engagement in Real Time." *Journal of Public Economics* 193: 104345.

Becker, Gary S. 1985. "Human Capital, Effort, and the Sexual Division of Labor." *Journal of Labor Economics* 3, no. 1, Part 2: S33-S58.

Blau, Francine D., and Lawrence M. Kahn. 2017. "The Gender Wage Gap: Extent, Trends, and Explanations." *Journal of Economic Literature*, 55 (3): 789-865.

Burki, Talha. 2020. "The Indirect Impact of COVID-19 on Women." *The Lancet. Infectious Diseases* 20 (8): 904–5.

Chen, M. Keith, Judith A. Chevalier, and Elisa F. Long. 2021. "Nursing Home Staff Networks and COVID-19." *Proceedings of the National Academy of Sciences* 118, no. 1

Chiou, Lesley, and Catherine Tucker. 2020. "Social Distancing, Internet Access and Inequality." No. w26982. National Bureau of Economic Research.

Cui, Ruomeng, Hao Ding, and Feng Zhu. 2020. "Gender Inequality in Research Productivity During the COVID-19 Pandemic." *arXiv preprint arXiv:2006.10194*.

DeFilippis, Evan, Stephen Michael Impink, Madison Singell, Jeffrey T. Polzer, and Raffaella Sadun. 2020. "Collaborating During Coronavirus: The Impact of COVID-19 on the Nature of Work." No. w27612. National Bureau of Economic Research.

Dhar, Diva, Tarun Jain, and Seema Jayachandran. 2018 "Reshaping Adolescents' Gender Attitudes: Evidence from a School-based Experiment in India." No. w25331. National Bureau of Economic Research.

Duflo, Esther, Michael Greenstone, and Rema Hanna. 2008. "Cooking Stoves, Indoor Air Pollution and Respiratory Health in Rural Orissa." *Economic and Political Weekly*: 71-76.

Egger, Dennis, Edward Miguel, Shana S. Warren, Ashish Shenoy, Elliott Collins, Dean Karlan, Doug Parkerson, et al. 2021. "Falling Living Standards during the COVID-19 Crisis: Quantitative Evidence from Nine Developing Countries." *Science Advances* 7 (6): eabe0997.

Fletcher, Erin K., Rohini Pande, and Charity Troyer Moore. 2017. "Women and Work in India: Descriptive Evidence and a Review of Potential Policies." CID Working Paper Series 2017.339, Harvard University, Cambridge, MA.

Fuchs, Victor R. 1989. "Women's Quest for Economic Equality." *Journal of Economic Perspectives*, 3 (1): 25-41.

Giurge, Laura M., Ashley V. Whillans, and Ayse Yemiscigil. 2021. "A Multicountry Perspective on Gender Differences in Time Use During COVID-19." *Proceedings of the National Academy of Sciences* 118, no. 12.

Goldfarb, Avi, and Catherine Tucker. 2019. "Digital Economics." *Journal of Economic Literature* 57, no. 13-43.

Heath, Rachel, and Xu Tan. 2020. "Intrahousehold Bargaining, Female Autonomy, and Labor Supply: Theory and Evidence from India." *Journal of the European Economic Association* 18, no. 4: 1928-1968.

Heffetz, Ori, and Daniel Reeves. 2021. "Measuring Unemployment in Crisis: Effects of COVID-19 on Potential Biases in the CPS." No. w28310. National Bureau of Economic Research.

Hyland, Marie, Simeon Djankov, and Pinelopi Koujianou Goldberg. 2020. "Gendered Laws and Women in the Workforce." *American Economic Review: Insights* 2 (4): 475–9.

Kan, Man Yee, and Stephen Pudney. 2008. "Stylized and Diary Data." *Sociological Methodology* 38: 101–32.

Lin, Tesary. 2020. "Valuing Intrinsic and Instrumental Preferences for Privacy." Available at SSRN 3406412.

MHA. 2020. "Guidelines for Phase-1 Lockdown (Order Number 40-3/2020-D) Dated-24/03/2020."

Miller, Amalia R. 2011. "The Effects of Motherhood Timing on Career Path." *Journal of Population Economics*, 24 (3): 1071–1100.

Miller, Grant, and A. Mushfiq Mobarak. 2013. "Gender Differences in Preferences, Intra-household Externalities, and Low Demand for Improved Cookstoves." No. w18964. National Bureau of Economic Research.

Myers, Kyle R., Wei Yang Tham, Yian Yin, Nina Cohodes, Jerry G. Thursby, Marie C. Thursby, Peter Schiffer, Joseph T. Walsh, Karim R. Lakhani, and Dashun Wang. 2020. "Unequal Effects of the COVID-19 Pandemic on Scientists." *Nature Human Behaviour* 4, no. 9: 880-883.

Schwartz, Paul. M. 2019. "Global Data Privacy: The EU Way." *New York University Law Review* 94, 771.

United Nations. 2020. "Policy Brief: The Impact of COVID-19 on Women." *UN Women.*

Zhou, Muzhi, Ekaterina Hertog, Kamila Kolpashnikova, and Man-Yee Kan. 2020. "Gender Inequalities: Changes in Income, Time Use and Well-being Before and During the UK COVID-19 Lockdown." SocArXiv. June 2. doi:10.31235/osf.io/u8ytc.

FIGURES AND TABLES

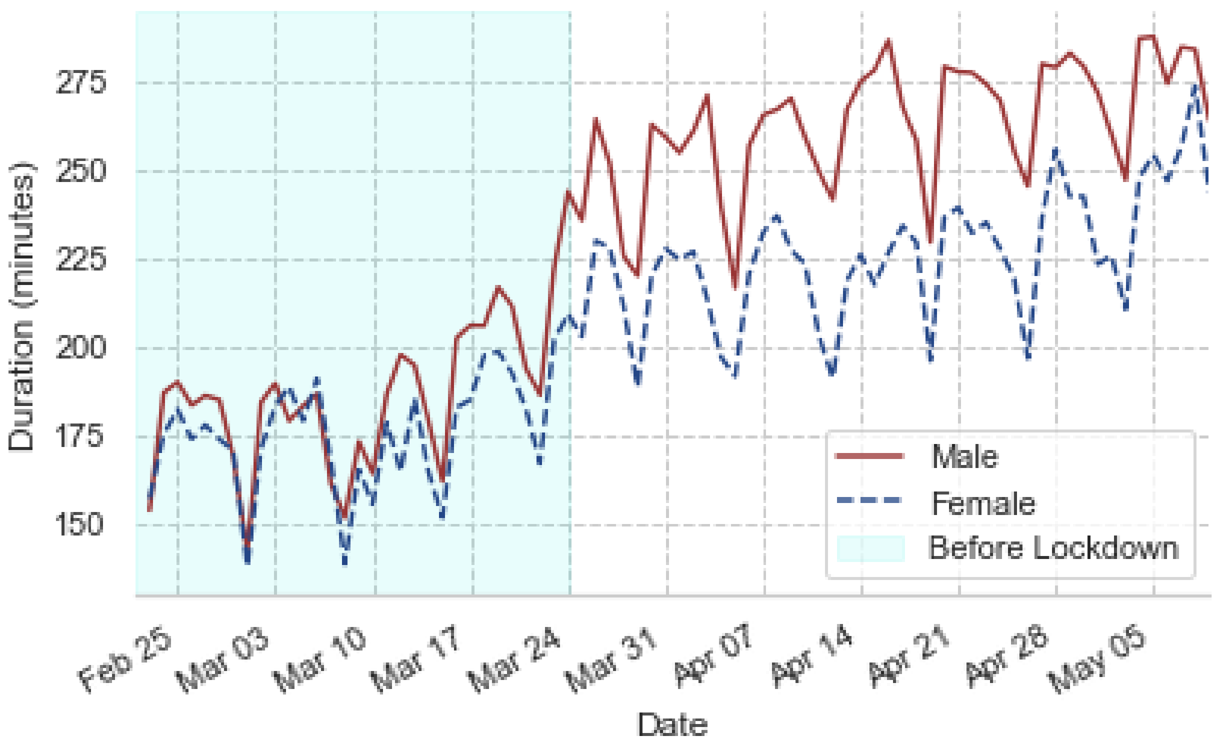

FIGURE 1. DAILY INTERNET BROWSER TIME USE FOR MEN AND WOMEN

Notes: This graph depicts average daily internet browser time use separately for men and women in our sample. The COVID-19 lockdown in India started on March 25, 2020, and continued through the end of the sample period. The official WHO declaration of COVID-19 as a global pandemic was on March 11, 2020.

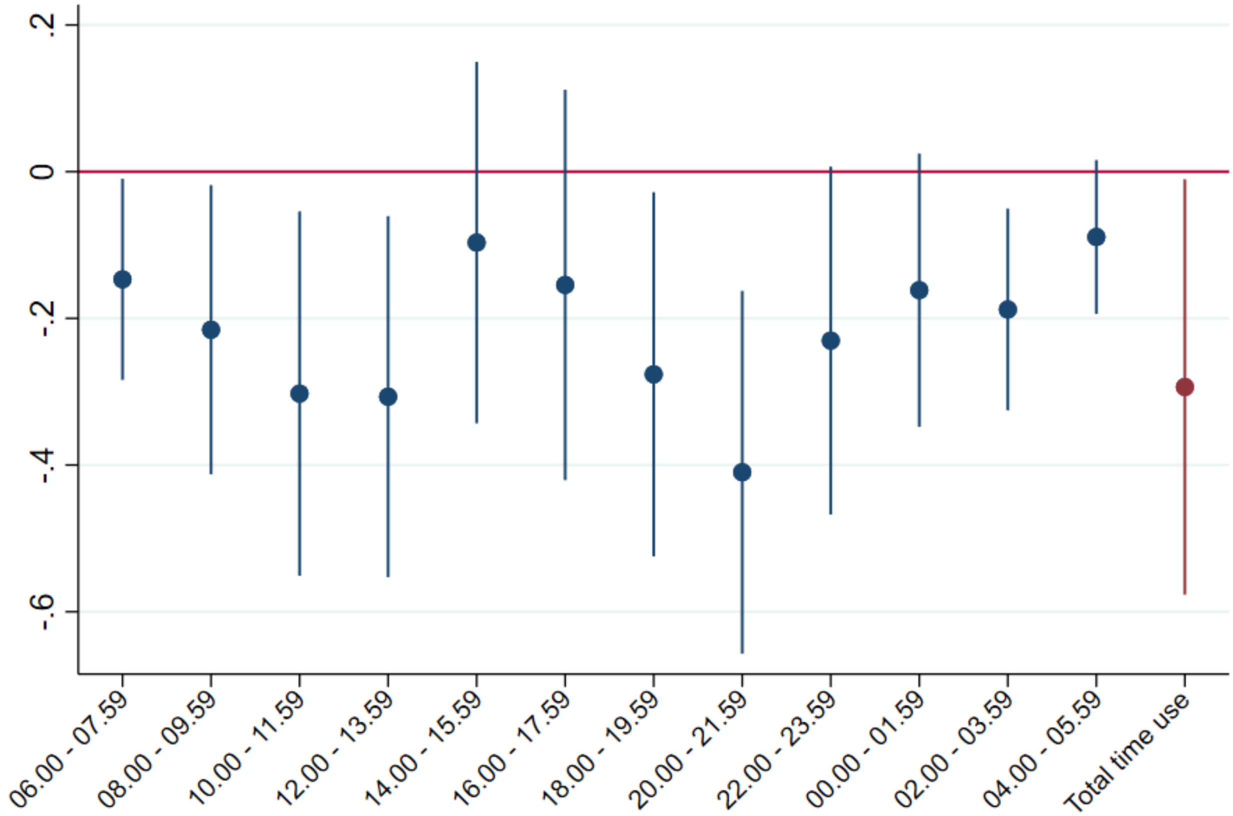

FIGURE 2. EFFECTS OF THE LOCKDOWN ON GENDER GAPS IN BROWSER TIME USE BY HOUR

Notes: This graph depicts separate results for the effects of the lockdown on the gender gap in total daily time use by time of day. The dependent variable is the natural log transformation of total daily browser time plus 1 second. We divided each day into twelve 2-hour intervals and ran separate regressions for each interval using our model with individual and date fixed effects. The dots depict regression estimates for each of the interaction terms between female and Lockdown indicators; bars show 95-percent confidence intervals, with standard errors clustered at the individual level.

TABLE 1. SAMPLE COMPOSITION

Variables	Women Mean	Women Std. Dev	Men Mean	Men Std. Dev	Female—Male Mean	Female—Male Std. Error
Age	30.755	7.366	33.118	7.86	-2.363	0.483***
Any Children	0.583	0.494	0.613	0.488	-0.029	0.031
Any Children Under 8	0.424	0.495	0.486	0.500	-0.061	0.031**
Married	0.606	0.489	0.642	0.48	-0.036	0.031
College Graduate	0.919	0.273	0.916	0.278	0.003	0.017
Employed Full Time	0.636	0.482	0.772	0.42	-0.136	0.029***
White-Collar Occupation	0.215	0.411	0.269	0.444	-0.055	0.027**
Self-Employed	0.124	0.33	0.181	0.385	-0.057	0.022***
Work from Home	0.437	0.497	0.45	0.498	-0.132	0.031
Number of Individuals	396		702		1,098	

Source: Survey responses from 1,098 individuals in India, between May 10 and June 4, 2020.

TABLE 2. DAILY BROWSER USE BY GENDER AND TIME PERIOD

	Full Sample	Female Sample		Male Sample	
		Before Lockdown	During Lockdown	Before Lockdown	During Lockdown
	Mean (Std. Dev.)	Mean (Std. Dev.)	Mean (Std. Dev.)	Mean (Std. Dev.)	Mean (Std. Dev.)
Total Time	224.1	175.8	225.0	186.4	262.0
	(230.3)	(206.0)	(223.8)	(216.8)	(242.3)
Total Unique URLs	255.1	192.0	259.2	198.8	307.0
	(631.4)	(428.9)	(769.7)	(417.1)	(698.9)
Leisure Time	111.6	88.66	114.8	91.19	129.5
	(185.1)	(167.0)	(185.2)	(172.0)	(195.9)
Production Time	82.78	64.55	82.47	69.62	96.79
	(102.8)	(84.37)	(103.8)	(94.38)	(110.3)
YouTube Time	73.63	52.25	68.85	62.33	90.06
	(155.1)	(133.3)	(142.6)	(148.2)	(170.1)
Unique YouTube Videos	6.366	4.101	5.300	5.175	8.417
	(16.40)	(11.33)	(12.13)	(14.83)	(20.14)
Unique Google Searches	5.094	4.104	5.164	4.232	5.895
	(10.18)	(8.893)	(9.471)	(9.492)	(11.24)
Facebook Time	4.955	4.272	4.834	3.905	5.875
	(20.59)	(19.64)	(21.90)	(18.75)	(21.12)
Job Search Time	1.565	1.754	1.221	1.320	1.844
	(10.13)	(15.33)	(7.247)	(8.188)	(10.43)
Online Learning Time	2.762	2.093	3.470	1.692	3.223
	(15.47)	(14.25)	(17.63)	(9.848)	(17.17)
Observations	81,929	10,709	18,530	19,762	32,928

Notes: Outcomes are at the person-day level and reported here in levels (minutes or counts).

TABLE 3. EFFECTS OF THE LOCKDOWN ON GENDER GAPS IN DAILY BROWSER TIME

	Total (1)	Leisure (2)	Production (3)	YouTube (4)	Facebook (5)	Google Search (6)
Lockdown × Female	-0.294**	-0.385**	-0.334**	-0.386***	-0.299***	-0.218**
	(0.144)	(0.151)	(0.138)	(0.139)	(0.0721)	(0.106)
Observations	81,929	81,929	81,929	81,929	81,929	81,929
Number of Individuals	1,098	1,098	1,098	1,098	1,098	1,098
Individual fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Date fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table presents the main estimates for daily internet browser time use. Standard errors are clustered at the individual level. Dependent variables are the natural log transformation of the outcome of interest plus 1 second. Significance at *** p<0.01, ** p<0.05, * p<0.1.

TABLE 4. EFFECTS OF THE LOCKDOWN ON GENDER GAPS IN DAILY WEBSITE VISITS

	Total Unique URLs (1)	Unique Leisure URLs (2)	Unique Production URLs (3)	Unique YouTube Videos (4)	Unique Google Searches (5)
Lockdown × Female	-0.283***	-0.280***	-0.267***	-0.171***	-0.0865**
	(0.0885)	(0.0684)	(0.0854)	(0.0421)	(0.0390)
Observations	81,929	81,929	81,929	81,929	81,929
Number of Individuals	1,098	1,098	1,098	1,098	1,098
Individual fixed effects	Yes	Yes	Yes	Yes	Yes
Date fixed effects	Yes	Yes	Yes	Yes	Yes

Notes: This table presents the main estimates for daily internet browser activity counts. Standard errors are clustered at the individual level. Dependent variables are the natural log transformation of the outcome of interest plus 1 second. Significance at *** p<0.01, ** p<0.05, * p<0.1.

TABLE 5. HETEROGENEOUS EFFECTS OF THE LOCKDOWN ON GENDER GAPS IN BROWSER TIME USE BY PARENTAL STATUS

	One Child or More			No Children		
	Total	Leisure	Production	Total	Leisure	Production
	(1)	(2)	(3)	(4)	(5)	(6)
Lockdown × *Female*	-0.376**	-0.627***	-0.353**	-0.182	-0.0404	-0.317
	(0.174)	(0.190)	(0.168)	(0.247)	(0.243)	(0.232)
Observations	49,316	49,316	49,316	32,613	32,613	32,613
Number of Individuals	661	661	661	437	437	437
Individual fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Date fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table presents separate estimates for two subsamples: those with at least one child (columns 1-3) and those with no children (columns 4-6). Standard errors are clustered at the individual level. Dependent variables are the natural log transformation of 1 plus the outcome of interest. Significance at *** p<0.01, ** p<0.05, * p<0.1.

TABLE 6. HETEROGENEOUS EFFECTS OF THE LOCKDOWN ON GENDER GAPS IN BROWSER TIME USE BY EMPLOYMENT STATUS

	Full-time Employed			Not Full-time Employed		
	Total Time Use	Leisure Time Use	Production Time Use	Total Time Use	Leisure Time Use	Production Time Use
	(1)	(2)	(3)	(4)	(5)	(6)
Lockdown × *Female*	-0.264	-0.574***	-0.233	-0.469*	-0.0841	-0.675**
	(0.172)	(0.173)	(0.163)	(0.282)	(0.310)	(0.268)
Observations	59,433	59,433	59,433	22,496	22,496	22,496
Number of Individuals	794	794	794	304	304	304
Individual fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Date fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table presents separate estimates for the full-time employed and not full-time employed (including students and part-time employed) samples. Standard errors are clustered at the individual level. Each regression includes an interaction of the Lockdown and female indicators. Dependent variables are the natural log transformation of 1 plus the outcome of interest. Significance at *** p<0.01, ** p<0.05, * p<0.1.

TABLE 7. DIFFERENTIAL EFFECTS OF THE LOCKDOWN ON ONLINE JOB SEARCH BY GENDER

	Full Sample		Not Employed Full-Time	
	Visited a Job Search Page	Job Search Page Time Use	Visited a Job Search Page	Job Search Page Time Use
	(1)	(2)	(3)	(4)
Lockdown × Female	-0.0241**	-0.147**	-0.0402**	-0.274***
	(0.00943)	(0.0577)	(0.0169)	(0.103)
Observations	81,929	81,929	19,987	19,987
Number of Individuals	1,098	1,098	271	271
Individual fixed effects	Yes	Yes	Yes	Yes
Date fixed effects	Yes	Yes	Yes	Yes

Notes: This table presents the results on the job search websites. Outcomes in columns 1 and 3 are indicator variables for whether a participant visited a job search website; outcomes in columns 2 and 4 are measures of time spent on job search websites (with the log transformation to the value plus 1 second). Columns 1 and 2 are from models estimated on the entire sample, while columns 3 and 4 use the subset of participants that are not employed full time at the time of the survey and had no change in employment status over the prior 90 days. Standard errors are clustered at the individual level. Significance at *** p<0.01, ** p<0.05, * p<0.1.

TABLE 8. EFFECTS OF THE LOCKDOWN SURVEY-BASED MEASURES OF HOUSEHOLD PRODUCTION TIME

	Childcare time use		Housework time use	
	Own *Married Sample with Children*	Partner's *Married Sample with Children*	Own *Married Sample*	Partner's *Married Sample*
	(1)	(2)	(3)	(4)
Female	0.634***	-1.243***	1.227***	-1.899***
	(0.207)	(0.198)	(0.182)	(0.196)
Lockdown	1.519***	0.733***	1.608***	0.501***
	(0.163)	(0.175)	(0.151)	(0.170)
Lockdown × Female	-0.765**	0.000864	-0.378	0.270
	(0.297)	(0.287)	(0.258)	(0.272)
Constant	2.577***	3.942***	2.619***	4.774***
	(0.107)	(0.122)	(0.107)	(0.126)
Observations	1,154	1,154	1,382	1,382
Number of Individuals	577	577	691	691

Notes: This table presents the estimates for survey-based time use outcomes related to household production. The unit of observation is a person-period (before or after the lockdown is imposed). Married respondents answered questions about their own and their partners' usual daily time spent on childcare (if they had children) and housework activities during the pre-lockdown and the lockdown periods. Daily time use was measured as an interval variable using 2-hour buckets up to 8 or more hours. We converted it to a continuous variable using the mid-point of each bin and assigning 10 hours to participants who selected 8 or more hours. The online survey was conducted during the lockdown period, so only the lockdown values are contemporaneous. Robust standard errors are in parentheses. Significance at *** p<0.01, ** p<0.05, * p<0.1.

APPENDIX FIGURES AND TABLES

FIGURE A1. DAILY BROWSER TIME USE FOR MEN AND WOMEN BY CATEGORY

Notes: This graph depicts the average internet browser time use for male (red) and female (blue) users across various categories and website domains. The pale blue shaded region represents the period before the COVID-19 lockdown in India on March 25, 2020.

TABLE A1. EFFECTS OF THE LOCKDOWN ON GENDER GAPS IN DAILY BROWSER TIME: EXTENSIVE MARGIN

	Any Browser Time	Any Leisure Time	Any Production Time	Any YouTube Time	Any Facebook Time	Any Google Searches
	(1)	(2)	(3)	(4)	(5)	(6)
Lockdown × *Female*	-0.0233	-0.0375**	-0.0308*	-0.040**	-0.0476***	-0.0384**
	(0.0145)	(0.0164)	(0.0157)	(0.0157)	(0.0108)	(0.0174)
Observations	81,929	81,929	81,929	81,929	81,929	81,929
Number of Individuals	1,098	1,098	1,098	1,098	1,098	1,098
Individual fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Date fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table presents the results for the extensive margin of internet browser usage. Standard errors are clustered at the individual level. In each column, the dependent variable is a binary equal to 1 if the user visited the corresponding website category and 0 otherwise. Significance at *** p<0.01, ** p<0.05, * p<0.1.

TABLE A2. EFFECTS OF THE LOCKDOWN ON GENDER GAPS IN DAILY BROWSER TIME: INTENSIVE MARGIN

	Total	Leisure	Production	YouTube	Facebook	Google search
	(1)	(2)	(3)	(4)	(5)	(6)
Lockdown × *Female*	-0.0980**	-0.169**	-0.120***	-0.199**	-0.0720	-0.0411
	(0.0402)	(0.0696)	(0.0453)	(0.0859)	(0.0815)	(0.0605)
Observations	68,677	52,143	65,886	33,978	16,819	48,576
Number of Individuals	1,098	1,088	1,098	1,040	947	1,085
Individual fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Date fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table presents the intensive margin results for daily internet browser time use. Standard errors are clustered at the individual level. Dependent variables are the natural log transformation of the outcome of interest. For each outcome variable, observations with no-usage days are dropped. Significance at *** p<0.01, ** p<0.05, * p<0.1.

TABLE A3. EFFECTS OF THE LOCKDOWN ON GENDER GAPS IN DAILY WEBSITE VISITS: INTENSIVE MARGIN

	Total Unique URLs	Unique Leisure URLs	Unique Production URLs	Unique YouTube Videos	Unique Google Searches
	(1)	(2)	(3)	(4)	(5)
Lockdown × Female	-0.192***	-0.252***	-0.179***	-0.166***	-0.0647*
	(0.0596)	(0.0600)	(0.0566)	(0.0510)	(0.0358)
Observations	68,677	52,143	65,886	32,011	48,903
Number of Individuals	1,098	1,088	1,098	1,037	1,085
Individual fixed effects	Yes	Yes	Yes	Yes	Yes
Date fixed effects	Yes	Yes	Yes	Yes	Yes

Notes: This table presents the intensive margin results for daily internet browser usage counts. Standard errors are clustered at the individual level. Dependent variables are the natural log transformation of the outcome of interest. For each outcome variable, observations with no-usage days are dropped. Significance at *** p<0.01, ** p<0.05, * p<0.1.

TABLE A4. EFFECTS OF THE LOCKDOWN ON GENDER GAPS IN DAILY BROWSER TIME:
NO DEVICE SHARING SUB-SAMPLE

	Total Time Use	Leisure Time Use	Production Time Use	YouTube Time Use	Facebook Time Use	Google search Time Use
	(1)	(2)	(3)	(4)	(5)	(6)
Lockdown × Female	-0.522***	-0.588***	-0.428**	-0.705***	-0.305***	-0.406***
	(0.197)	(0.213)	(0.186)	(0.194)	(0.110)	(0.143)
Observations	42,659	42,659	42,659	42,659	42,659	42,659
Number of Individuals	566	566	566	566	566	566
Individual fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Date fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table presents the results on daily internet browser time use on the subsample who do not share their devices (smartphone, tablet, computer) with others. Standard errors are clustered at the individual level. Dependent variables are the natural log transformation of 1 plus the outcome of interest. Significance at *** p<0.01, ** p<0.05, * p<0.1.

TABLE A5. EFFECTS OF THE LOCKDOWN ON GENDER GAPS IN DAILY WEBSITE VISITS: NO DEVICE SHARING SUB-SAMPLE

	Total URL Count	Leisure URL Count	Production URL Count	Unique YouTube Video Count	Unique Google Search Count
	(1)	(2)	(3)	(4)	(5)
Lockdown × Female	-0.360***	-0.345***	-0.277**	-0.250***	-0.147***
	(0.119)	(0.0959)	(0.115)	(0.0587)	(0.0539)
Observations	42,659	42,659	42,659	42,659	42,659
Number of Individuals	566	566	566	566	566
Individual fixed effects	Yes	Yes	Yes	Yes	Yes
Date fixed effects	Yes	Yes	Yes	Yes	Yes

Notes: This table presents the results on daily internet browser usage counts on the subsample who do not share their devices (smartphone, tablet, computer) with others. Standard errors are clustered at the individual level. Dependent variables are the natural log transformation of 1 plus the outcome of interest. Significance at *** p<0.01, ** p<0.05, * p<0.1.

TABLE A6. INTERNET BROWSER USAGE BY GENDER AND PARENTAL OR EMPLOYMENT STATUS

	Parental Status				Employment Status			
	One Child or More		No Children		Employed Full Time		Not Full-Time Employed	
	Female	Male	Female	Male	Female	Male	Female	Male
Total Time	203.7	221.9	211.7	252.1	200.5	234.2	218.7	231.8
	(211.3)	(228.4)	(228.8)	(246.1)	(209.8)	(237.4)	(233.6)	(230.9)
Leisure Time	94.66	99.75	120.2	139.3	90.35	109.5	132.1	134.2
	(166.7)	(174.7)	(194.5)	(205.4)	(159.0)	(185.3)	(208.1)	(196.6)
Production Time	82.54	91.91	66.55	78.31	83.09	92.70	62.97	66.16
	(104.0)	(110.1)	(86.79)	(97.18)	(105.7)	(110.2)	(79.01)	(84.49)
Observations	17,139	32,177	12,100	20,513	18,823	40,610	10,416	12,080

Notes: Unit of observation is a person-day. Browser time use measured in minutes per day. Standard deviations in parentheses.

TABLE A7. EFFECTS OF THE LOCKDOWN ON GENDER GAPS IN YOUTUBE-PURPOSE-ADJUSTED DAILY BROWSER TIME

	Leisure time use	Production time use	Leisure URL count	Production URL count
	(1)	(2)	(3)	(4)
Lockdown × Female	-0.392***	-0.308**	-0.271***	-0.265***
	(0.143)	(0.140)	(0.0668)	(0.0852)
Sample Mean	70.412	97.510	38.725	176.299
Observations	81,929	81,929	81,929	81,929
Number of Individuals	1,098	1,098	1,098	1,098
Individual fixed effects	Yes	Yes	Yes	Yes
Date fixed effects	Yes	Yes	Yes	Yes

Notes: This table presents the main estimates for daily internet browser time use and activity counts after adjusting for the YouTube leisure and production split. Dependent variables are the natural log transformation of the outcome of interest plus 1 second. The sample mean and standard deviation are at the person-day level and reported in levels (minutes). Standard errors are clustered at the individual level. Significance at *** p<0.01, ** p<0.05, * p<0.1.

TABLE A8. EFFECTS OF THE LOCKDOWN ON YOUTUBE USAGE

	Leisure time use	Production time use	Leisure URL count	Production URL count
	(1)	(2)	(3)	(4)
Lockdown × Female	-0.358***	-0.316***	-0.115***	-0.101***
	(0.115)	(0.0922)	(0.0334)	(0.0238)
Sample Mean	32.453	14.711	3.294	1.478
Observations	81,929	81,929	81,929	81,929
Number of Individuals	1,098	1,098	1,098	1,098
Individual fixed effects	Yes	Yes	Yes	Yes
Date fixed effects	Yes	Yes	Yes	Yes

Notes: This table presents the estimates for YouTube usage. `Leisure' and `Production' are aggregate measures based on the YouTube categories collected via the YouTube API. `Leisure' includes Autos & Vehicles, Comedy, Entertainment, Film & Animation, Gaming, Movies, Music, People & Blogs, Pets & Animals, Sports, Trailers, and Travel & Events. `Production' includes Education, How to & Style, News & Politics, and Science & Technology. The daily average YouTube time use is 73.63 minutes (Table 2). This includes the time spent on non-video URLs such as the YouTube home and search pages. It also includes time spent on videos whose category could not be determined via the YouTube API. Dependent variables are the natural log transformation of the outcome of interest plus 1 second. Columns 1 and 2 are duration, and columns 3 and 4 are daily URL count measures. Standard errors are clustered at the individual level. The sample mean and standard deviation are at the person-day level and reported in levels (minutes or counts). Significance at *** p<0.01, ** p<0.05, * p<0.1.

TABLE A9. EFFECT OF THE LOCKDOWN ON FINER YOUTUBE CATEGORY USAGE

	Leisure Usage						Production Usage		
	Movies	Music	Games	People & Blogs	Entertainment	Other Leisure	News & Media	Education	Other Production
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Lockdown × Female	-0.05	-0.094	-0.034	-0.277***	-0.233***	-0.147**	-0.469*	-0.0841	-0.675**
	(0.047)	(0.0664)	(0.0405)	(0.070)	(0.0805)	(0.0728)	(0.282)	(0.310)	(0.268)
Sample Mean	3.374	7.534	2.655	5.947	9.575	9.396	3.871	6.320	4.520
Observations	81,929	81,929	81,929	81,929	81,929	81,929	81,929	81,929	81,929
Number of Individuals	1,098	1,098	1,098	1,098	1,098	1,098	1,098	1,098	1,098
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table presents the estimates for YouTube category usage. We collect the category information of YouTube videos by feeding our URL data into the YouTube API. The "Movies" category consists of Film & Animation, Movies, and Trailers. "Other Leisure" includes Autos & Vehicles, Comedy, Pets & Animals, Sports, and Travel & Events. "Other Production" includes How to & Style, and Science & Technology. Dependent variables are the natural log transformation of the outcome of interest plus 1 second. All outcome variables are time-use measures. Standard errors are clustered at the individual level. The sample mean and standard deviation are at the person-day level and reported in levels (minutes). Significance at *** p<0.01, ** p<0.05, * p<0.1.

TABLE A10. CHILDCARE-RELATED BROWSER USAGE

	Childcare Duration *Full Sample*	Childcare Duration *Sample with Children*	Childcare Duration Word2Vec *Full Sample*	Childcare Duration Word2Vec *Sample with Children*	YouTube Kids Channels *Full Sample*	YouTube Kids Channels *Sample with Children*
	(1)	(2)	(3)	(4)	(5)	(6)
Lockdown × Female	-0.796	-0.804	-0.814	-0.190	-0.149	-0.229
	(1.452)	(1.986)	(0.846)	(1.248)	(0.141)	(0.237)
Sample Mean (Parents)	7.148	-	4.771	-	0.149	-
Sample Mean (Non-Parents)	5.138	-	2.822	-	0.015	-
p-value for test: Parents = Non-Parents	0.054	-	0.002	-	0.001	-
Observations	81,929	49,316	81,929	49,316	81,929	49,316
Number of Individuals	1,098	661	1,098	661	1,098	661
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table presents the results for childcare-related internet browser usage. Outcome variables are measured in minutes. Standard errors are clustered at the individual level. The subsample means are at the person-day level and reported in levels (minutes). P-values report the test for the equality of means, after clustering the standard errors at the individual level. Significance at *** p<0.01, ** p<0.05, * p<0.1

Appendix B. Text Analysis for Childcare-related Browser Usage

Our website category data do not include specific categories like childcare-related browser usage. We identify those activities online by applying textual analysis to the title data and YouTube video descriptions. We define three alternative measures of childcare-related browser usage.

The first approach applies a manually created a dictionary of 165 keywords that are childcare-related and used by Indian parents. We code each webpage visit as childcare-related if the title of the page contains a word from this dictionary.[20]

Although manual dictionary-based methods are common in the literature (e.g., Baker, Bloom, and Davis 2016; Enke 2020), a shortcoming of these techniques is that their performance depends heavily on the quality of the predefined information (in the dictionary). To circumvent this drawback, we also use a natural language processing method to create a model-based dictionary. We first fit a Word2Vec model (Mikolov et al. 2013) to our website title and YouTube description data. Word2Vec is a widely adopted word-embedding technique, where each word w is represented by a K-dimensional vector $\vec{w} \in R^K$. For a given sequence of words $w_1, w_2, .. w_N$, (in a title or video description) the model takes each word as input and aims to predict the surrounding words that come before and after. Therefore, the objective of the model is to choose word vectors so as to maximize the following likelihood function $\sum_n \sum_{i \in S_n} \log p(w_i | w_n)$, where S_n is the set of words surrounding w_n. The vector representation of words allows for measuring semantic or syntactic similarities between words. We leverage this feature to minimize the dependency on prior human information in creating a dictionary. First, we select 8 childcare-related seed words: cartoon, child, infant, kid, nursery, school, toddler, and toy. Then, we pick the

[20] We resort to dictionary-based methods because we do not have any labelled data on childcare-related website categories to use as a training dataset. Also, the title information that identifies a website as childcare-related might be weak. Therefore, topic models are unlikely to endogeneously form a childcare-related website category (see Gentzkow, Kelly and Taddy (2019) for a detailed discussion).

5 words most similar for each seed word, measured by the cosine similarity between word vectors, to form our model-driven dictionary.[21]

Our third approach is to identify 26 YouTube channels that exclusively produce child-targeted content. Capturing usage through these YouTube channels does not provide complete information on the broader childcare-related browser usage. However, as a predictor of child-targeted content usage, it would have minimal type 1 error. Therefore, it provides reliable information on a specific type of childcare-related website usage.

Despite the limitations, we are able to confirm that each of the three measures of child-related content are related to parental status in the expected way. Parents spent significantly more time on the childcare-related content than did childless adults.

Table A10 presents the estimates for each of these measures on the full sample and on the subsample of respondents with children. Unlike the estimates for other online activities, we find no significant drop in childcare-related browser time use for women, relative to men. Results from the Word2Vec-based dictionary (columns 3 and 4) and YouTube Kids channels (columns 5 and 6) are similar to the findings from our manual dictionary method (columns 1 and 2).

[21] Cosine similarity measures the cosine angle between two *K*-dimensional vectors.

Online Appendix References

Enke, Benjamin. 2020. "Moral values and voting." *Journal of Political Economy* 128 (10): 3679-3729.

Gentzkow, Matthew, Bryan Kelly, and Matt Taddy. 2019. "Text as data." *Journal of Economic Literature* 57 (3), 535-74.

Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. "Distributed Representations of Words and Phrases and Their Compositionality." *In Proceedings of the 26th International Conference on Neural Information Processing Systems*.

Scott R. Baker, Nicholas Bloom, Steven J. Davis. 2016. "Measuring Economic Policy Uncertainty." The Quarterly Journal of Economics, 131 (4): 1593–1636.