G²LM|LIC
Gender, Growth, and Labour Markets in Low Income Countries Programme

WP

G²LM|LIC Working Paper No. 64 | May 2022

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)

IZA Institute of Labor Economics
Initiated by Deutsche Post Foundation

UKaid
from the British people

G²LM|LIC Working Paper No. 64 | May 2022

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)

G²LM|LIC
c/o IZA – Institute of Labor Economics
Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-510

Email: glm-review@iza.org

G²LM|LIC Working Paper No. 64 | May 2022

ABSTRACT

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

We measure the digital impact of the initial Indian COVID-19 lockdown using an online survey coupled with consensually provided browser history records from over 1,000 individuals, spanning over 30 million website visits. Both men and women in our sample dramatically increased their internet activity during the lockdown, which reflects the heightened importance of digital access, but men's activity increased by significantly more. Gender differences are present overall and for key categories including leisure, production, video streaming and social media. The exception is for self-investment through online learning websites and educational YouTube videos, where men and women had similar significant increases. Among full-time employed respondents, women's lower browser usage is mainly in leisure browsing, while it is concentrated in productive activities among part-time workers and non-workers. The lockdown also saw a significant reduction in women's online job search, alongside a significant increase in men's, with larger effects among likely job seekers, indicating potentially persistent harm to women's employment. The gender gap is larger among parents, consistent with increased childcare obligations as the driver. Yet in our survey, fathers self-reported significantly larger increases in childcare time than mothers. This relative increase in paternal childcare was not corroborated in partners' reports or in childcare-related browser usage, which we identify leveraging machine learning methods to analyze text from website titles and YouTube video descriptions. The inconsistency within the self-reported data and contrast with the digital trace data underscore the value of accessing objective "digital footprint" records to gain insight into time use and activity.

JEL Classification:
J16, J22, J24, L82, L86, D1

Keywords:
digital activity, data triangulation, data privacy, gender, COVID19 lockdown, time use

Corresponding author:
Amalia R. Miller
Department of Economics
University of Virginia
P.O. Box 400182
237 McCormick Road
Charlottesville, VA 22904
USA
E-mail: armiller@virginia.edu

1. Introduction

Inequality in the adoption, use, and impact of technology has been extensively studied in the Information Systems literature.[1] We extend this literature by leveraging the ubiquity of online information to quantify the gendered impact of the COVID-19 pandemic. We employ a digital approach to data collection that combines a brief online survey with consensual sharing of objective internet browser histories to collect longitudinal time-use data during the initial Indian lockdown, enabling cross-validation. Despite the strictness of the Indian lockdown, which closed schools and blocked non-essential workers from leaving home (MHA 2020), as well as our informed-consent protocol for all data collection, we were able to obtain data from over a thousand people, covering over 30 million website visits before and during the lockdown.

Our IT-enabled data provide a unique perspective on how Indian men and women adapted differently to the pandemic lockdown that was both sudden and severe. This perspective is particularly valuable because of the prohibitive barriers to collecting reliable time-use data during lockdowns, combined with important concerns that lockdowns and school closures enacted to mitigate the COVID-19 pandemic imposed particularly severe time burdens on women (United Nations 2020a; Alon et al. 2020; Burki 2020). Browser histories provide us with rich and objective historical time use information that we collected without access to the extensive in-depth and in-person questioning, involving multiple time windows (e.g., past day/week/month), typically required for verification and validation in time-use surveys, while also avoiding the measurement problems of recall bias and misreporting in subjective reports.

The digital divide is a key issue in India, where internet connectivity is rapidly growing, but significant gender gaps in technology adoption and active internet usage persist.[2] Despite recognition in the Information Systems community that more research focused on developing

[1] See, for example, Acquisti and Fong (2019) in the context of hiring discrimination; Atasoy, Banker and Pavlov (2021) in the context of IT skills and employment; Ahuja and Thatcher (2005) in the context of workplace environment; and Mejia and Parker (2021) in the context of ridesharing platforms.
[2] Recent reports estimate 57% of men and 43% of women in urban India are active online, e.g., https://economictimes.indiatimes.com/tech/technology/india-to-have-900-million-active-internet-users-by-2025-says-report/articleshow/83200683.cms. The gender gap is also reflected in the male dominance in our sample.

countries is needed, and of gender issues in digital access (Walsham et al. 2007), research has been hampered by limited data availability.[3] The gender disparity in digital access in India is also reflective of stark gender disparities across a range of economic, health and social outcomes (Duflo 2012), which motivates policy concerns about women being particularly vulnerable to lockdown disruptions (United Nations 2020a).

Our first findings of significant increases in browser use across a range of activities, including leisure, production, and human capital investment, highlight the greatly amplified value of digital access during the Indian lockdown. However, this increase was unequal, and men's browser time increased by significantly more, both overall and for important categories including leisure and production. This result shows that the pandemic expanded the existing gender gap in active internet usage in India, even among those with digital access. This is particularly concerning in light of the large digital divide by gender in digital access and use, even before the pandemic.

When we split the sample by working status, we find that the gender gap in online leisure time is more prominent among full-time workers, while the gap in online productive activities is larger among those not in full-time jobs. This suggests that relative to men, working women sacrificed online leisure to maintain productive time use, which could lead to stress and burnout, while other women forwent potential earnings opportunities. We also find indications that the gendered economic effects of the lockdown may persist – we see women's online job search activity decline, both in absolute terms and relative to men's, particularly among job seekers.

The gender differential in the impact of the lockdown is larger among parents, consistent with additional household obligations during the lockdown disproportionally consuming mothers' time. We investigated household time use directly with our survey but did not find that women self-reported larger increases in childcare time.[4] Rather, men and women both reported

[3] https://www.epw.in/engage/article/where-data-study-internet-india

[4] Most research quantifying the pandemic's effect on time use relies on surveys, focusing on binary outcomes (Del Boca et al. 2020), examining cross-sectional differences during the pandemic (Giurge, Whillans, and Yemiscigil 2021), or using repeated cross-sections (Teodorovicz et al. 2021) or retrospective questions to obtain pre-pandemic baselines (e.g., Adams-Prassl et al. 2020; United Nations 2020b). Time diaries are more reliable (Hamermesh et al. 2005), and can distinguish between primary and secondary

significant increases in time devoted to housework and childcare during the lockdown, and men reported significantly larger (absolute and relative) increases in childcare time than women. This surprising finding echoes Zhou et al. (2020), a rare pandemic study with longitudinal time-use data, where self-reported housework time increased more for men (3.5 hours) than for women (3 hours) at the onset of the UK lockdown. However, in our survey, the pattern of men's increased time devoted to childcare is not corroborated by responses to questions about spousal time use. Although this inconsistency could come from differences across households, it is notable that men reported greater childcare time while also spending significantly more time online.

To triangulate across another data source, we adopted machine learning and textual analysis methods to identify childcare-related browser usage to examine whether men were spending more time online on childcare-related usage, e.g., watching children's videos. Across various externally validated measures we find no statistically significant increase in men's childcare-related browser time use during the lockdown, relative to women's. Also, consistent with prior research,[5] women in our sample shared their devices with others more than men did, and among those who did not share their devices, the gender gap in usage was even larger. Although it is possible that men increased both their childcare and leisure time by more than women did because they were more likely to have lost their jobs or started working from home during the lockdown, we find that neither of these factors explains the effects in our data. This raises concerns about the quality of simple self-reported time use measures and the possibility that men and women perceive or report their activities differently, as suggested by prior findings that men tend to overreport their household production time (Kan and Pudney 2008) and that fathers devote a higher fraction of their childcare time to secondary or passive care (Folbre and Yoon 2007).

Our use of objective and detailed browser data together with subjective time use reports that are subject to recall bias and misreporting combines aspects of Myers et al. (2020) and Cui et al. (2021) who use either surveys or objective output measures to study of the productivity effects

childcare (Folbre and Yoon 2007), but are more onerous to collect and infeasible during strict pandemic lockdowns. The American Time Use Study was suspended between March 19 and May 11, 2020.

[5] https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2019/03/PI_2019.03.07_Mobile-Connectivity_FINAL.pdf

of the pandemic.[6] The browser data also enable us to examine novel outcomes related to internet activity, not typically included even in detailed time diaries. The exceptional depth and detail in browser histories – in our continuous clickstream data, we observe the timing and title of every web page that is opened – allow us to go beyond usual time allocations to examine a range of specific categories and activities, such as video watching or job search, and to compare differences across and even within days.

By using IT-enabled objective measurements of online time use, this paper contributes to literatures in IS and economics that study time allocations in households (Becker 1965; Blau and Kahn 2017; Hamermesh, 2016) and among information workers (Bhansali and Brynjolfsson 2007). The paper resembles Collopy (1996) in comparing self-reported and IT-enabled objective measures of time use, but we depart from much of the prior literature in our focus on gender differences across a novel set of outcomes related to internet time use and in measuring the shock of pandemic lockdowns.

The internet is a technology that profoundly affects people's lives, yet research on gendered impacts has been limited. By studying gender differences in internet use, we contribute to a growing body of work in IS that harnesses digitally gathered data to study inequality between demographic groups. Previous studies have empirically examined the increased risks of group-based discrimination created by the rapid rise of personal data digitalization (Leidner and Tona 2021) in the context of online hiring (Acquisti and Fong 2020; Chan and Wang forthcoming), hate speech (Ananthakrishnan and Tucker 2021), sharing-economy platforms (Cui, Li and Zhang 2020; Mejia and Parker 2021), and online job advertisements (Lambrecht and Tucker 2019). In the area of IT-related labor market outcomes, studies have measured gender differences in promotion rates for IT service workers (Langer, Gopal and Bapna 2020), in effects of "teleworkability" on employment (Hou et al. 2021), and in labor market returns to IT skills in developing countries (Atasoy, Banker, and Pavlou 2021). While this research includes productivity and human capital investment measures as outcomes, our scope also includes leisure activities and household production. Unlike prior IS studies that have focused on time and activity devoted to specific

[6] Outside of the pandemic, Bandiera et al. (2020) examine objective measures of CEO activities drawn from their calendar entries.

internet applications or social media platforms (Ghose and Han 2011; Han, Park and Oh 2015; Bapna et al. 2016; Rishika and Ramaprasad 2019), we study both aggregate measures of overall browser use, as well as measures by category and for specific pages.

We also contribute by exploiting the exogenous shock of a strict national COVID-19 lockdown to study gendered changes in internet activity. Like other pandemic studies, in IS and economics, we leverage digital footprints left by individuals in their ordinary activities. Prior work has used aggregate or de-contextualized data from Google Trends reports (Bacher-Hicks et al. 2021; Brodeur et al. 2021), de-identified email and meeting meta-data (DeFilippis et al. 2020), and smartphone geolocation data (Chiou and Tucker 2020; Chen et al. 2021; Ananthakrishnan et al. 2020). Our approach differs in that we explicitly ask individuals to respond to our survey and to grant us one-time access to obtain a snapshot of their recent browser history. The survey responses provide key context for the digital histories of individual respondents.

Although our data collection approach has clear 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 of our approach is that linking browser and survey information enables us to study gender, family status and employment. More fundamentally, our data collection method represents another way to balance between the competing interests of obtaining rich digital data and protecting the privacy of individual users. Rather than embedding or exploiting trackers on individual computers, we worked in partnership with *PY Insights*, 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 a study (e.g., Athey, Catalini and Tucker 2017; Prince and Wallsten 2021; Lin forthcoming), recruiting costs will be higher in such an approach, and the sample less representative. These biases are offset by the ability to study changes in the activities of individuals over time.

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. Our approach also addresses the methodological concern that forward-looking data collection with informed consent could affect online activity – for example, people who agree to being tracked for a study might alter their online behavior. As concerns about digital privacy increase among regulators and the public (Schwartz 2019; Goldfarb and Tucker 2019; Acquisiti, Taylor and Wagman 2016; Al-Natour et al. 2020), and data privacy laws are increasingly adopted, including in developing countries,[7] the importance of considering alternative models for ethical and privacy-protecting digital data collection will increase as well.

2. Data

2.1. Primary Data Collection

We collaborated with *PY Insights*, an internet-browser analytics platform, and *Dynata*, a global first-party data platform, to field our survey between mid-May and early June 2020. Individuals drawn from Dynata's marketing pools in India were invited to participate in an online survey that ended with a consensual browser data upload using the *PY Insights* software. Participants with valid data were compensated for their effort. *PY Insights'* internet browser extension collects retrospective data stored in each user's browser account history. This is identical to what a participant would observe if they visited the *History* section of their internet browser on their personal computer (see Figure A2 for an illustrative example). The records cover up to 90 days of past activity on the browser account, accumulated across all electronic devices (computer, smartphone, tablet). We observe every website visit, including the URL (uniform resource locator, i.e., web address) and timestamp.[8] Although our browser data can include records from multiple types of electronic devices, most smartphone browser apps do not support internet browser extensions or add-ons, so the *PY Insights* technology only collects data from personal computers. No information is collected from private browsing or Incognito mode, and personal identifiers are removed prior to analysis.

Each URL has an associated title, which conveys 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, *PY Insights* calculates its duration in seconds and

---

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

[8] The software only captures retrospective data. Once the data transfer is over, it automatically deletes itself and redirects participants to the survey platform.

provides a detailed categorization scheme for each website domain.[9] We use these categories to identify websites as being primarily related to leisure (entertainment) or production (non-recreational).[10] Because YouTube represents a sizeable portion of usage and is classified as leisure by *PY Insights*, we also conduct robustness checks in which we re-classify YouTube videos as leisure or production related using Google's YouTube API.

We obtained data that met our quality control standards from 1,094 individuals aged 22 to 54 located in 28 states across India. We prevented individuals using a new browser account or a secondary browser type that is not used regularly from participating by requiring at least 30 days of browser data. We dropped one user who preferred not to state their gender and took two steps to avoid computer bots: we included an attention test question in the survey and manually dropped all users with an average of more than 3,000 URL visits per day.[11]

In total, we collected over 31.5 million webpage visits to 134,123 unique websites. 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 81,462 days of individual browser usage data with 52,509 days coming from 701 men and 28,953 days from 393 women.

2.2. Summary Statistics

Although we targeted equal gender balance, 64% of our respondents were male (Table 1), which may reflect the gender gap in digital access in India.[12] The need for computer access to participate

[9] The categories are based on Google Cloud Platform's natural language processing algorithm. The universe of categories are at https://cloud.google.com/natural-language/docs/categories.
[10] *Leisure* includes Adults, Arts & Entertainment, Games, Online Communities (including social media), and Shopping. *Production* includes Business & Industrial, Computers & Electronics, Finance, Internet & Telecom (including e-mail and search engines), Jobs & Education, Law & Government, News, Science, and Reference. Other Google Cloud categories combined cover 0.8% of our data. Some websites – such as spam webpages – are also labelled as "other". Median "other" category usage on a day covers 7% of total time use.
[11] The 19 users who failed this requirement show browsing that is unlikely to come from a human, such as spending entire days repeatedly visiting the same handful of business websites, refreshing every 5 seconds.
[12] See https://www.oecd.org/going-digital/bridging-the-digital-gender-divide.pdf, and http://rchiips.org/nfhs/factsheet_NFHS-5.shtml

likely also contributes to high educational attainment in our sample, with over 90% of men and women being college graduates. Full-time employment (including self-employment) is higher for men (77%) than for women (64%), but the latter is elevated relative to the Indian population. Men in our sample are also somewhat older and more likely to be married and have children than are women.

That this sample is unrepresentative of India's population does not threaten the internal validity of the within-person changes that we measure (using objective data untainted by personal recollection).[13] However, to the extent that the impact of the lockdown varied across individuals, the sample's composition will affect the average effects we compute, and the differences by gender in those effects. Therefore, the estimates should be interpreted as applying to individuals of the type that would and could complete this survey, a relatively advantaged Indian subpopulation – literate in English and having access to an internet-connected computer – that is rapidly growing and increasingly engaging with information technology.[14]

Average daily browser time use in our sample is 3.7 hours (Table 2), with about two hours devoted to leisure – including 1 hour watching YouTube videos, and about 1.5 hours on production (column 1). Men and women had similar browser use in the pre-lockdown period (columns 2 and 3), with no statistically significant gender difference in total time. Both men (column 6) and women (column 7) significantly increased their time online during the lockdown, by over an hour a day on average (column 5). However, the increase in browser time was smaller for women, a pattern apparent in Figure 1. The figure shows that browser time, and its daily fluctuations, were similar for men and women at the start of the sample. Male usage started increasing following the World Health Organization's March 11, 2020 declaration that the COVID-19 outbreak is a global pandemic. During the lockdown, men and women both show increasing browser time, and the increase for men is noticeably larger.

[13] The difficulty of recruiting a representative sample during theer pandemic is not unique to this study. Low response rates have generated quality concerns even for well-established government surveys in the US, such as the ACS (https://www.census.gov/newsroom/press-releases/2021/changes-2020-acs-1-year.html). The challenges have been immense in lower-income countries (e.g., Egger et al. 2021).

[14] Computer penetration in India is estimated to be about 3 percent and growing about 15 percent a year (IDC, 2020).

When we examine internet use by purpose, we see sizable increases in both leisure and production usage for men and women, but the increases in leisure are larger both absolutely and relative to pre-lockdown mean (Table 2 and Figure 1). The increase in leisure time online likely reflects a shift to online leisure to replace offline social and leisure activities prevented by the lockdown. Some of the increased production time may come from 49.8 percent of our employed sample reporting a shift to working at home during the lockdown (Table 1). Although productive activity naturally also took place offline or outside of the browser, it is reassuring to observe the cyclical pattern of weekly usage, with regular drops on Sundays, unique to production. It is also reassuring that Figure 1 shows a sharp drop in online shopping activity at the outset of the lockdown, which is consistent with the severity of the lockdown, that prevented home deliveries. Online shopping repeats the common pattern of a relative decline in women's time use during the lockdown, but differs in that women's usage exceeded men's before it.

There is one small but important category that stands out as an exception to the pattern of relative reductions in women's time online, which is human capital investment. When measured by time spent on online learning domains, the category shows similar and significant increases for both men and women that are statistically indistinguishable from one another. This feature is also present in data on time spent on YouTube videos in the "educational" category (Table A5) and in subjective reports of frequency of "self-investment" activities (Table A2).

3. Gendered Effects of the Lockdown on Browser Activity

We estimate the differential effect of the lockdown by gender using panel data and two-way fixed effects for individuals and time. We define the pre-lockdown baseline period through March 24, 2020, and the lockdown period as starting on March 25, 2020, the date of the first national COVID-19 lockdown in India. The lockdown 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:

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

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

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, $Female_i$ 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, β, captures the average differential impact of the lockdown on women relative to men. We apply a natural logarithmic transformation on our outcomes, after adding 1 (second or click) to all daily observations to retain zero values.

3.1. Gendered Effects on Internet Browser Activity

Confirming the patterns in the raw data, the regression estimates in Panel A of Table 3 show sizable and significant relative declines in women's time online during the lockdown across a variety of measures. Women's total browser time decreased by 25.3 percent relative to men's – i.e., nearly half an hour less time per day.[16] Women's online time use decreased relative to men's by 27.8 percent for leisure and by 28.6 percent for production websites. We find similar relative declines in our count-based measures of activity in Panel B. Women's daily count of unique URLs visited dropped by 24.4 percent relative to men's, amounting to about 40 fewer URLs per day. We also find significant usage drops for women, relative to men, for video streaming (YouTube, time and clicks), social media (Facebook time) and Google searches.

We aggregated activity to the daily level for most of our analysis, but we also explored variation in the impacts by time of day. We divided each day into twelve 2-hour intervals and ran separate regressions on browser time use (total, leisure and production) for each interval. The results are in Figure 2, with estimates starting at 6 AM on the left. The effects are largest midday and in the late evening. These times coincide with lunch and dinner, which are both typically hot meals in Indian households. Because of gender roles typically assigning South Asian women with responsibility for these tasks (Duflo, Greenstone and Hanna 2008; Dhar, Jain and Jayachandran 2018), we expect that women in our sample are more likely than men to be involved in meal

[16] Because the outcome is logged, the coefficient of -0.292 implies a change of -25.3% = 100*($e^{-0.292}$-1).

preparation, service and clean-up, which could explain the observed gender differences in internet use.

3.2. 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 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 by splitting the sample based on parental status.

Table 4 presents separate estimates for samples of individuals with at least one child and with no children (summary statistics in Table A6). We observe significant drops in total, leisure and production time use for mothers relative to fathers, while among childless adults, we find no significant gender differences in any of these measures. The difference between the two samples is greatest (and statistically significant) for leisure time. Mothers experienced a relative drop in online leisure of 43.3 percent compared to fathers, while childless women experienced an insignificant increase relative to childless men. The disproportionate effect of the lockdown on mothers is primarily manifesting in our data as a relative reduction in leisure time.

We next split our sample by employment status, to test whether effects are stronger for women who have less economic power and autonomy. Consistent with this prediction, our estimates for total time use and production time are smaller and less significant in the full-time employed sample (columns 1 and 3 of Table 4, Panel B) than in the sample of individuals not employed full-time (columns 4 and 6). Nevertheless, relative to full-time employed men, full-time women had a substantial and significant 38.8 percent decrease in leisure time online. Thus, the differential impact of the lockdown was not limited to only those working part time or less. 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 48.5 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 relative to men's and choosing instead to sacrifice leisure time.[17] 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.

3.3. Gendered Effects on Online Job Search

We next consider differences in online job search. Although it accounts for a small share of browser time, job search can have lasting effects on labor market outcomes. Over three-quarters of job applications worldwide are submitted online and India's growing online job market remained active during the lockdown while in-person networking and job applications were strictly disallowed.[18] We created a comprehensive list of job search websites frequented in India and classified website visits as relating to job search if their URL domain is included in this list. Because many observations in our sample have zero time devoted to online job-search, we supplemented our usual log-transformed measure of daily browser time use with a daily indicator for whether the person visited any job-search websites.

The gender differences are striking. The summary statistics show that men's time devoted to online job search increased by about 40 percent during the lockdown, while women's job search time decreased by a similar amount (Table 2). Regression estimates in Table 4 (Panel C) show the significance of the relative drop in women's time spent performing online job search during the lockdown: a 2.2 percentage point drop on the extensive margin (column 1) and a 12.9 percent decrease in duration (column 2).

Because we lack a measure of offline job search, our results for online job seeking may in part reflect a shift in medium rather than amount, yet they are concerning indicators of worsening gender gaps in Indian labor markets following the lockdown. Indian women's labor force participation remains low despite the country's economic growth, declining fertility and rising education levels. The absolute decline in women's online job search during the lockdown is particularly troubling in light of prior findings that women have lower access to social protections

[17] Their observed online leisure time is significantly lower than their self-reported ideal allotment.

[18] https://www.statista.com/statistics/881116/recruitment-share-of-job-applications-by-source-worldwide/

(Cameron 2019), yet often lack information about available jobs and search for jobs less efficiently than men (Fletcher, Pande and Moore 2017).

4. Effects of the Lockdown on Household Production

4.1. Childcare and Housework Time Use

Our focus in this paper is on internet browser activity, where we have the most robust data. Because housework and childcare activities are impossible to capture with browser data alone, we asked survey questions about time use for those categories. Despite caveats about the quality of subjective time-use reports, a relative increase in women's time expended on household production would provide direct evidence of the hypothesized mechanism underlying women's relative decreases in time online. However, that is not what we find.

The results of our analysis of survey-based measures of time use, comparing men and women, before and during the lockdown, are in Figure 3 (see Table A8 for the regression results). We asked married individuals separately about their own and their partner's usual daily time spent on housework and (if they had children) on childcare. We converted the interval responses (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 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 0.7 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 large positive and highly significant lockdown coefficients confirm our expectation that childcare and housework time increased sharply during 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: over 1.5 hours more per day on childcare, double the increase reported by women (0.756 = 1.523 – 0.767; column 1). With this additional 45 minutes a day, men report devoting as much time as women do to childcare during the lockdown. This finding of a 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 (column 2) and the gender difference 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 the increase in housework time during the lockdown for either own or partner's time. This conflict suggests that the gender difference in the impact of the lockdown on self-reported time devoted to children may 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).

4.2. Self-Reported Measures and Textual Analysis of Browser Data

We considered the possibility that men devoted more of their time online to browsing child-targeted content with their children. We found no empirical support for this in our data. Because we are unable to identify browsing activity shared with children, we focus on webpages and videos aimed at children, to enable data triangulation. We identify such activities online by applying textual analysis and a machine learning algorithm to the website 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 childcare-related keywords and used by Indian parents. These keywords were identified by conducting semi-structured interviews with multiple Indian parents, who have internet access. We code each webpage visit as childcare-related if the title of the page contains a word from this dictionary. We resort to dictionary-based methods here because we lack labelled data on childcare-related website categories to use as a training dataset and because topic models (e.g., Latent Dirichlet Allocation (Blei, Ng, and Jordan, 2003)) are unlikely to endogenously form a childcare-related website category (Gentzkow, Kelly and Taddy, 2019).

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 expert knowledge to curate the dictionaries. This makes it difficult for manual dictionaries to comprehensively capture the full range of words that refer to a particular topic. 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 $\vec{w} \in R^K$. We use the skip-gram implementation of Word2Vec. 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, in a fixed window. 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. Mikolov et al. (2013) show that the resulting word vectors capture semantic and syntactic similarities between words in an efficient way. 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 5 words most similar for each seed word, measured by the cosine similarity between word vectors, to form our model-driven dictionary.

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 and can serve as a robustness check to validate our textual analysis results.

We are able to confirm that each of these 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 5).

Table 5 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 found small (ranging from < 10 to 50 seconds) and statistically insignificant gender differences in the effect of the lockdown on child-related internet use. Results from the Word2Vec-based dictionary (columns 3 and 4) and YouTube Kids channels (columns 5 and 6) are similar to the findings from the manual dictionary method (columns 1 and 2).

Another way to reconcile these findings is that men spent more time consuming online content while caring for their children during the lockdown. The content may not have been targeted primarily at children or even consumed together with their children. Men may have multi-tasked – pursuing leisure and productive activities online while also keeping an eye on their children. This could produce the gender differences we observe in self-reports if men are more likely to describe time spent on a device in the vicinity of children as "childcare" time, while women tend to reserve that term for primary childcare, time spent actively caring for children or supervising their activities. In that case, a rise in secondary childcare time among fathers, even if matched with a similar or larger rise among mothers, could generate the pattern in the self-reported data. The browser data are not able to resolve this conflict, but they suggest that simple self-reported data on childcare time use paint an incomplete picture at best.

5. Robustness Checks

5.1. Alternative Model Specifications

Our main results in Table 3 are robust to various alternative specifications. We first examined alternative definitions of the dependent variables, separating the extensive and intensive margins and using a linear model. We studied the extensive margin using an indicator variable for the person visiting any URL (overall or within the category) in Table A1, Panel A, and the intensive margin separately for daily browser time use and daily count of website visits using the log transform of the outcome variable of interest without adding 1 (which drops zero usage days from the sample) in Table A1 Panels B and C. Finally, we repeated our main estimates from Table 3 without employing the log transformation in Table A1, Panels C and D. Across all of these models, the estimates confirm the main results, except for a few cases where the effects have the same direction but lack the statistical significance. Those exceptions are the extensive margin of any browser time in the day, intensive margin on time spent on Facebook and the number of Google searches, and the raw count of production URLs.

We next confirmed the robustness of our results to an alternative clustering structure of standard errors. Table A10 shows the results are unchanged when we use two-way (participant and date) cluster-robust standard errors instead of just clustering on the participant level.

5.2. Alternative Estimation Samples

This section reports results from estimates on sub-samples aimed at addressing concerns about the pre-lockdown changes in behavior coming from the World Health Organization (WHO) declared COVID-19 as a global pandemic on March 11, 2020 and from sharing of electronic devices.

We first address the concern that individuals in our sample responded to the WHO announcement even before the formal lockdowns was imposed. In that case, the dates from the WHO announcement to the start of our treatment period (on March 25) may not represent usual pre-pandemic behaviour. We therefore repeated our main models from Table 3 excluding data between March 11 to 24. The estimated impacts of the lockdown in Table A10 are larger across all categories, increasing by 3 to 21 percent.

We also considered the possibility that the relative increase in browser time attributed to men is due to their greater sharing of devices with others in the household. Because our survey elicits device sharing, we are able to estimate separate effects for the sub-group that does not share their smartphone, computer, or tablet. Consistently across all regressions in our main analysis in Table 3, we find larger effects for this sub-group (Table A3). Women in this sub-sample decrease their total time online by 40.7 percent relative to men (compared to 25.3 percent in the full sample). This difference suggests that women in the full sample shared their devices more intensively than men, and a greater share of their browser activity was consumed by others. Thus, our full-sample results may underestimate the relative decline in women's time online.

5.3. Job Changes Do Not Explain the Gendered Effects of the Lockdown

Because the lockdown is associated with greater job loss for men (4.3 percent) than for women (3.6 percent) in our sample, the relative increase in men's time devoted to job search may come from their greater need for search rather than from women's increased household obligations. We address this concern by identifying individuals who are more likely to be job seekers throughout the 90-day lookback window: those that 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 decreases in both the extensive margin (3.9 percentage points) and in overall duration (24 percent) measures of job search activity for women relative to men on this sub-sample (Table 4 Panel C, columns 3 and 4).

We also consider the possibility that higher rates of male job loss, or more shifting to work from home, can explain the apparent puzzle that men self-reported relative increases in time devoted to childcare while also increasing in their relative browser time use. These two things could both be true if lockdowns reduced external demands on men's time by more. However, the main results for both browser time use and for self-reported childcare time use are unchanged when we drop from our sample the 44 people who reported losing a job during the pandemic (Table A7, Panel A). Furthermore, the gender differences remain significant after we expand our models to control for differential impacts of the lockdown on people experiencing job loss (separately for themselves and their spouses) or starting to work from home (self or spouse) in Table A7, Panel B.

5.4. Separating Leisure and Production Activities in YouTube

Finally, because YouTube accounts for almost 20 percent of total browser time in our sample, we further parsed the video content of 308,497 unique YouTube URLs using Google Cloud's

YouTube Data API. For each URL, the YouTube API provides an array of information about the associated video, such as its title, category, description, and channel name.[19]

We used YouTube video categories to identify videos that are more related to leisure or production. Two-thirds of YouTube time is devoted to leisure in this scheme. The results of our main analysis are unchanged if we revise our category-level usage measures by moving productive YouTube content into the production category (Table A4, Panel A). We confirmed that the pattern of results from the full browsing data is present within YouTube videos as well: women's time devoted to both leisure and production videos drops considerably relative to men's during the lockdown (Tables A5, Panel B).

6. Discussion and Conclusions

Around the world, the curtailment of face-to-face activities during the COVID-19 pandemic made the internet a vital avenue for leisure, production, and human capital investment. This paper provides a unique view into how pandemic lockdowns changed digital activity and time use, drawing on novel data from an online survey data and continuous clickstream data tracing internet browser histories, collected during the initial Indian lockdown.

This exercise provides a novel demonstration of the value of information that can be extracted by analyzing "digital footprints" left by people going through their normal online activities, even in a setting with strong privacy protections and fully informed consent for all data collection. Collecting data in a consensual manner was more costly than using de-contextualized or anonymized data sources, which limited the scale of the collection, but we managed to obtain data in a short time frame from over a thousand people. The higher cost was more than offset by the availability of supplemental information on demographic, contextual and subjective factors necessary for this analysis. This model can be applied and extended to other settings in which

[19] Details at https://developers.google.com/youtube/v3. This information was not available for videos that had been removed by the time we collected YouTube API data.

organizations and researchers want to benefit from "big data" but are constrained by legal, ethical, and practical considerations.

Our browser data provide a rich and objective record that enable us to measure changes in online activity and time use around the time of the lockdown. By capturing the substantial increases in browser use among both men and women, across a range of activity domains, our data illustrate the heightened importance of digital access during times of disruption and physical danger. This increased value of digital access during the pandemic lockdown supports greater public and private investment in expanding such access more broadly, and particularly of reducing existing disparities in access between demographic groups.

The benefits of increased internet use that we find on our relatively privileged sample of highly-educated Indians with personal computers and internet access were not available to people without such access. This has implications for gender equality because of the digital divide by gender, which is in part reflected in the male-dominated composition of our sample. Within this sample, our findings of relative increases in internet activity for men, overall, and across a range of activities, further suggest widening gaps in wellbeing from uneven digital use. Access to a device and an internet connection are not enough to ensure full use when other factors interfere. These results have implications for policymakers concerned with IT diffusion and with its uneven distribution.

By combining browser and survey data, we are also able to measure gender differences in the impact of the lockdown for different sub-groups. We find the relative decline for women particularly in the leisure domain, is concentrated among parents. This suggests that a source may be that the lockdown disproportionately exacerbated the caretaking burdens on women. However, this was not detected in our time-use survey on childcare time, where men self-reported relatively larger increases in time spent caring for children than women did. The pattern in self-reports is also not echoed in reports from spouses or in objective data on child-related internet browsing, suggesting that self-reports may be unreliable because of the subjective aspects of responses to simple time use questions about childcare time. With increasing availability of objective digital trace data and development of machine learning methods that enable highly granular measures, similar cross-validation of survey data may become common.

In addition to providing evidence on how the immediate effects of the initial COVID-19 lockdown in India differed by gender, our results also have implications for employers and organizations that seek to attract and retain female talent. Two of our findings suggest additional challenges coming from lower labor force attachment among women without full-time jobs. The first is the relative decline in online production time use for those women compared to similar men. The second is the decrease in women's online job seeking activity, both absolutely and relative to men. These outcomes may be directly observable to employers who observe work time and job applications. Our third finding, for full time working women, is less directly visible. These women maintained their productive time online, relative to men, but they experienced significant relative drops in online leisure time. This occurred while in-person leisure activities were largely proscribed and may have long term consequences, such as burnout, that drive some women to leave their jobs. The finding suggests that employers could benefit from investing proactively in inquiring about and supporting the mental health and work-life balance challenges of their workers, in addition to efforts and programs developed in response to the pandemic to expand opportunities for remote and flexible work. That women did not fall behind in human capital development suggests that they continued to aspire towards career advancement despite the setbacks the pandemic caused, a positive outlook that employers and governments could benefit from nurturing.

Acknowledgements

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 PY Insight and Dynata for assistance with data collection; and Ritu Agarwal, Ashish Arora, Oriana Bandiera, Sanjay Jain, Selin Kesebir, Chris Parker, Jeff Prince, Carmit Segal, Michael Smith, Bruce Weber and conference participants at the 2020 IZA G2LM|LIC annual research conference and the 2021 NBER Summer Institute on IT and Digitization 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.

References

Acquisti A, Fong C (2020) An Experiment in Hiring Discrimination via Online Social Networks. *Management Science* 66(3): 1005-1024.

Adams-Prassl A, Teodora B, Marta G, Rauh C (2020) Inequality in the Impact of the Coronavirus Shock: Evidence from Real Time Surveys. *Journal of Public Economics* 189: 104245.

Ahuja MK, Thatcher JB (2005) Moving Beyond Intentions and Toward the Theory of Trying: Effects of Work Environment and Gender on Post-adoption Information Technology Use. *Management Information Systems Quarterly* 29(3): 427-459.

Al-Natour S, Cavusoglu H, Benbasat I, Aleem U (2020) An Empirical Investigation of the Antecedents and Consequences of Privacy Uncertainty in the Context of Mobile Apps. *Information Systems Research* 31(4), 1037-1063.

Alon T, Doepke M, Olmstead-Rumsey J, Tertilt M (2020) The Impact of COVID-19 on Gender Equality. NBER Working Paper Number 26947.

Ananthakrishnan UM, Tucker C (2021) The Drivers and Virality of Hate Speech Online. Available at SSRN 3793801.

Ananthakrishnan UM, Chen J, Susarla A (2020) No Pain, No Gain: Examining the Digital Resilience of the Fitness Sector During the COVID-19 Pandemic. Available at SSRN 3740143.

Atasoy H, Banker RD, Pavlou PA (2021) Information Technology Skills and Labor Market Outcomes for Workers. *Information Systems Research* 32(2):437-461

Athey S, Catalini C, Tucker C (2017) The Digital Privacy Paradox: Small Money, Small Costs, Small Talk. NBER Working Paper Number 23488.

Bacher-Hicks A, Goodman J, Mulhern C (2021) Inequality in Household Adaptation to Schooling Shocks: Covid-Induced Online Learning Engagement in Real Time. *Journal of Public Economics* 193: 104345.

Baker SR, Bloom M, Davis SJ (2016) Measuring Economic Policy Uncertainty. *Quarterly Journal of Economics*, 131 (4): 1593–1636.

Bandiera O,Prat A, Hansen S, Sadun R (2020) CEO Behavior and Firm Performance. *Journal of Political Economy* 128(4): 1325-1369.

Bapna R, Ramaprasad J, Shmueli G, Umyarov A (2016) One-way Mirrors in Online Dating: A Randomized Field Experiment. *Management Science* 62(11), 3100-3122.

Becker GS (1965) A Theory of the Allocation of Time. *Economic Journal* 75 (299): 493-517.

Bhansali S, Brynjolfsson E (2007) Digitizing Work: Driving and Measuring Changes in Information Worker Time Use and Performance Via a Longitudinal Quasi-Experiment. *International Conference on Information Systems Proceedings* 26.

Blau FD, Kahn LM (2017) The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature*, 55 (3): 789-865.

Blei DM, Ng AY, Jordan MI (2003) Latent Dirichlet Allocation. *The Journal of Machine Learning Research* 3: 993-1022.

Brodeur A, Clark AE, Fleche S, Powdthavee N (2021) COVID-19, Lockdowns and Well-being: Evidence from Google Trends. *Journal of Public Economics* 193: 104346.

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

Cameron L (2019) Social Protection Programs for Women in Developing Countries. *IZA World of Labor.* https://wol.iza.org/uploads/articles/471/pdfs/social-protection-programs-for-women-in-developing-countries.pdf?v=1

Chan J, Wang J (Forthcoming) Hiring Preferences in Online Labor Markets: Evidence of a Female Hiring Bias. *Management Science*

Chen MK, Chevalier JA, Long EF (2021) Nursing Home Staff Networks and COVID-19. *Proceedings of the National Academy of Sciences* 118 (1): e2015455118.

Chiou L, Tucker C (2020) Social Distancing, Internet Access and Inequality. NBER Working Paper Number 26982.

Collopy F (1996) Biases in Retrospective Self-reports of Time Use: An Empirical Study of Computer Users. *Management Science* 42(5): 758-767.

Cui R, Hao D, Zhu F (2021) Gender Inequality in Research Productivity during the COVID-19 Pandemic. *Manufacturing & Service Operations Management.*

Cui R, Li J, Zhang DJ (2020) Reducing discrimination with reviews in the sharing economy: Evidence from field experiments on Airbnb. *Management Science* 66(3), 1071-1094.

DeFilippis E, Impink SM, Singell M, Polzer JY, Sadun R (2020) Collaborating During Coronavirus: The Impact of COVID-19 on the Nature of Work. NBER Working Paper Number 27612.

Del Boca D, Oggero N, Profeta P, Rossi M (2020) Women's and Men's Work, Housework and Childcare, before and during COVID-19. *Review of Economics of the Household* 18(4): 1001-1017.

Dhar D, Jain T, Jayachandran S (2018) Reshaping Adolescents' Gender Attitudes: Evidence from a School-based Experiment in India. NBER Working Paper Number 25331.

Duflo E, Greenstone M, Hanna R (2008) Cooking Stoves, Indoor Air Pollution and Respiratory Health in Rural Orissa. *Economic and Political Weekly*: 43(32): 71-76.

Duflo E (2012) Women Empowerment and Economic Development. *Journal of Economic Literature* 50(4): 1051-79.

Egger D, Miguel E, Warren SS, Shenoy A, Collins E, Karlan D, Parkerson D, et al. (2021) Falling Living Standards during the COVID-19 Crisis: Quantitative Evidence from Nine Developing Countries. *Science Advances* 7 (6): eabe0997.

Enke B (2020) Moral Values and Voting. *Journal of Political Economy* 128 (10): 3679-3729.

Fletcher EK, Pande R, Moore CT (2017) Women and Work in India: Descriptive Evidence and a Review of Potential Policies. CID Working Paper Series 2017.339, Harvard University, Cambridge, MA.

Folbre N, Yoon J (2007) What is Child Care? Lessons from Time-use Surveys of Major English-speaking Countries. *Review of Economics of the Household* 5(3): 223-248.

Gentzkow M, Kelly B, Taddy M (2019) Text as data. *Journal of Economic Literature* 57 (3): 535-74.

Ghose A, Han SP (2011) An Empirical Analysis of User Content Generation and Usage Behavior on the Mobile Internet. *Management Science* 57(9): 1671-1691.

Giurge LM, Whillans AV, Yemiscigil A (2021) A Multicountry Perspective on Gender Differences in Time Use During COVID-19. *Proceedings of the National Academy of Sciences* 118 (12): e2018494118.

Goldfarb A, Tucker C (2019) Digital Economics. *Journal of Economic Literature* 57 (1): 3-43.

Han SP, Park S, Oh W (2015) Mobile App Analytics: A Multiple Discrete-continuous Choice Framework. *Management Information Systems Quarterly* 40(4): 983-1008.

Hamermesh DS (2016) What's to Know About Time Use? *Journal of Economic Surveys*, 30 (1): 198–203.

Hamermesh DS, Frazis H, Stewart J (2005) Data Watch: The American Time Use Survey. *Journal of Economic Perspectives*, 19 (1): 221-232.

Hou J, Liang C, Chen P, Gu B (2021). Gender Heterogeneity in the Effect of Telework on Labor Market Outcomes during the COVID-19 Pandemic . *ICIS 2021 Proceedings 21.*

IDC (2020) Indian Traditional PC Market Finishes the Year at a Six-year High with 11 Million Shipments in 2019. https://www.idc.com/getdoc.jsp?containerId=prAP46039420.

Kan MY, Pudney S (2008) Measurement Error in Stylized and Diary Data on Time Use. *Sociological Methodology* 38(1):101-132.

Langer N, Gopal RD, Bapna R (2020) Onward and Upward? An Empirical Investigation of Gender and Promotions in Information Technology Services. *Information Systems Research* 31(2): 383-398.

Lambrecht A, Tucker C (2019) Algorithmic Bias? An Empirical Study of Apparent Gender-based Discrimination in the Display of STEM Career Ads. *Management science* 65(7): 2966-2981.

Leidner DE, Tona O (2021) The CARE Theory of Dignity Amid Personal Data Digitalization. *Management Information Systems Quarterly* 45(1): 343-370.

Lin T (Forthcoming) Valuing Intrinsic and Instrumental Preferences for Privacy. *Marketing Science*.

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

Mikolov T, Sutskever I, Chen K, Corrado G, Dean J (2013) Distributed Representations of Words and Phrases and Their Compositionality. *Proceedings of the 26th International Conference on Neural Information Processing Systems* pp. 3111-3119.

Mejia J, Parker C (2021) When Transparency Fails: Bias and Financial Incentives in Ridesharing Platforms. *Management Science* 67(1): 166-184

Myers KR, Tham WY, Yin Y, Cohodes N, Thursby JG, Thursby MC, Schiffer P, Walsh JT, Lakhani KR, Wang D (2020) Unequal Effects of the COVID-19 Pandemic on Scientists. *Nature Human Behaviour* 4(9): 880-883.

Prince J, Wallsten S (2021) Empirical Evidence of the Value of Privacy. *Journal of European Competition Law & Practice* 12(8): 648-654.

Rishika R, Ramaprasad J (2019) The Effects of Asymmetric Social Ties, Structural Embeddedness, and Tie Strength on Online Content Contribution Behavior. *Management Science* 65(7): 3398-3422.

Schwartz PM (2019) Global Data Privacy: The EU Way. *New York University Law Review* 94 (4): 771.

Teodorovicz T, Sadun R, Kun AL, Shaer O (2021) Working from Home during COVID-19: Evidence from Time-Use Studies. Harvard Business School Working Paper Number 21-094.

United Nations (2020a) Policy Brief: The Impact of COVID-19 on Women. https://www.unwomen.org/en/digital-library/publications/2020/04/policy-brief-the-impact-of-covid-19-on-women

United Nations (2020b) Whose time to care: Unpaid care and domestic work during COVID-19. https://data.unwomen.org/publications/whose-time-care-unpaid-care-and-domestic-work-during-covid-19

Walsham G, Robey D, Sahay S (2007) Foreword: Special issue on information systems in developing countries. *Management Information Systems Quarterly* 31(2): 317-326.

Zhou M, Hertog E, Kolpashnikova K, Kan MY (2020) Gender Inequalities: Changes in Income, Time Use and Well-being Before and During the UK COVID-19 Lockdown. SocArXiv. 2 June. doi:10.31235/osf.io/u8ytc.

FIGURES AND TABLES

Figure 1 Average Daily Internet Browser Time Use for Men and Women

Notes. The COVID-19 lockdown in India started on March 25, 2020, and continued through the end of the sample period. The pale blue shaded region represents the pre-lockdown period. The WHO officially declared COVID-19 as a global pandemic on March 11, 2020.

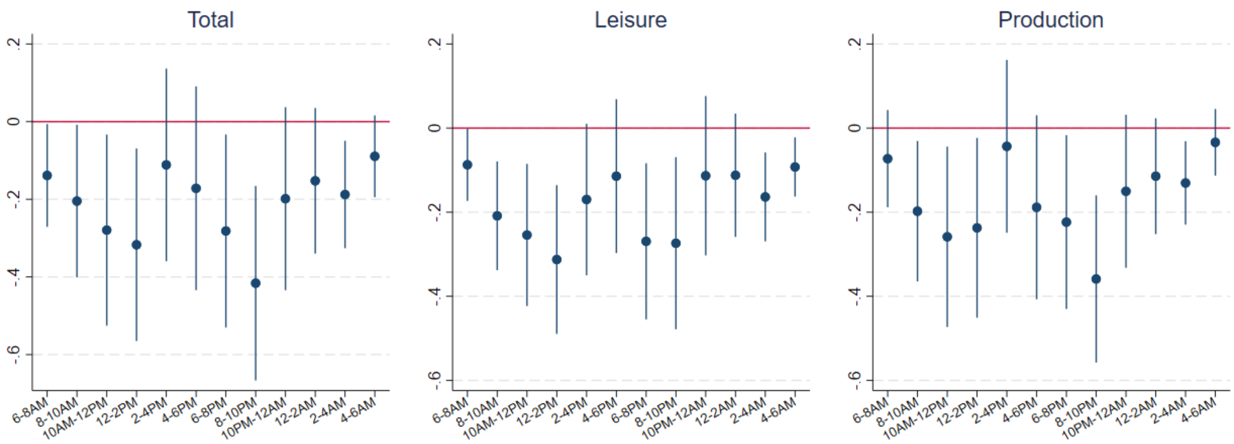

Figure 2 Within-Day Changes in Internet Browser Time Use

Notes. This figure presents 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 the daily browser time plus 1 second. We divided each day into twelve 2-hour intervals and ran a separate regression 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		Men		Female—Male	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Error
Age	30.71	7.369	33.124	7.864	-2.414	0.485***
Any Children	0.58	0.494	0.612	0.488	-0.032	0.031
Any Children Under 8	0.425	0.495	0.485	0.5	-0.060	0.031*
Married	0.603	0.49	0.642	0.48	-0.039	0.031
College Graduate	0.921	0.27	0.916	0.278	0.005	0.017
Employed Full Time	0.639	0.481	0.772	0.42	-0.133	0.029***
White-Collar Occupation	0.214	0.41	0.27	0.444	-0.058	0.022**
Self-Employed	0.122	0.328	0.18	0.384	-0.057	0.022***
Started Working from Home	0.438	0.497	0.449	0.498	-0.012	0.031
Number of Individuals	393		701		1,094	

Notes. Survey responses from 1,094 individuals in India, between 10 May and 4 June, 2020.

Table 2 Daily Browser Use by Gender and Time Period

	Before Lockdown				Increase During Lockdown		
	Full Sample Mean [Std. Dev.]	Male Sample Mean [Std. Dev.]	Female Sample Mean [Std. Dev.]	Difference Mean (Std. Err.)	Full Sample Mean (Std. Err.)	Male Sample Mean (Std. Err.)	Female Sample Mean (Std. Err.)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Total Time	221.3	184.5	171.0	-13.55	66.05***	75.48***	49.59***
	(231.5)	[216.9]	[207.6]	(10.28)	(4.270)	(5.503)	(6.615)
Total Unique URLs	246.5	193.2	177.9	-15.34	93.19***	107.7***	67.65***
	(599.3)	[379.8]	[360.0]	(15.72)	(10.43)	(10.40)	(22.33)
Leisure Time	111.3	91.08	87.82	-3.269	33.91***	37.98***	26.74***
	(185.5)	[172.3]	[167.7]	(8.398)	(3.240)	(4.096)	(5.276)
Production Time	80.78	68.52	61.01	-7.503*	23.67***	27.03***	17.88***
	(102.7)	[94.30]	[83.43]	(4.038)	(1.849)	(2.287)	(3.113)
YouTube Time	73.48	62.30	51.78	-10.52	23.64***	27.51***	17.01***
	(155.2)	[148.4]	[133.5]	(6.848)	(2.637)	(3.461)	(3.961)
Unique YouTube Videos	6.361	5.183	4.061	-1.122*	2.497***	3.230***	1.220***
	(16.43)	[14.85]	[11.32]	(0.595)	(0.264)	(0.374)	(0.306)
Unique Google Searches	5.050	4.220	4.019	-0.201	1.432***	1.652***	1.046***
	(10.19)	[9.506]	[8.911]	(0.379)	(0.159)	(0.207)	(0.242)
Facebook Time	4.905	3.899	4.166	0.267	1.451***	1.973***	0.512
	(20.59)	[18.78]	[19.65]	(1.077)	(0.336)	(0.479)	(0.359)
Job Search Time	1.561	1.323	1.740	0.416	0.147	0.525***	-0.537
	(10.14)	[8.205]	[15.38]	(0.687)	(0.254)	(0.185)	(0.642)
Online Learning Time	2.758	1.695	2.088	0.393	1.472***	1.536***	1.347**
	(15.49)	[9.867]	[14.29]	(0.560)	(0.310)	(0.346)	(0.610)
Observations	81,462	19,675	10,565	30,240	81,462	52,509	28,953

Notes. Outcomes are at the person-day level and reported here in levels (minutes or counts). Column (4) reports the estimated gender difference (female – male) for each browser use outcome in the pre-lockdown period. Columns (5)-(7) report average increases in browser use (lockdown – pre-lockdown) for the full sample (7) and then for male (6) and female (7) sub-samples. Standard errors are clustered at the individual level. Significance at *** p<0.01, ** p<0.05, * p<0.1

Table 3 Effects of the Lockdown on Gender Gaps in Browser Activity

	Total	Leisure	Production	YouTube	Facebook	Google Search
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Daily Browser Time						
Lockdown × Female	-0.292**	-0.326**	-0.337**	-0.344**	-0.295***	
	(0.148)	(0.151)	(0.140)	(0.138)	(0.0720)	
Panel B. Daily Website Visits						
Lockdown × Female	-0.280***	-0.254***	-0.265***	-0.160***		-0.085**
	(0.0897)	(0.0682)	(0.0862)	(0.0421)		(0.0390)
Observations	81,462	81,462	81,462	81,462	81,462	81,462
Number of Individuals	1,094	1,094	1,094	1,094	1,094	1,094
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 and activity counts. Panel A presents results for browser time use outcomes, and Panel B presents the activity counts, measured as unique URLs generated. Column (1) shows total browser use while subsequent columns are for categories: leisure (2), production (3), YouTube videos (4), Facebook (5) – URLs not examined because extensive activity occurs within the main URL, and Google searches (6) – time on search pages not examined because people typically follow links to results quickly. 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 4 Effects by Parental Status and Employment Status and on Job Search

Panel A. Parental Status

	One Child or More			No Children		
	Total	Leisure	Production	Total	Leisure	Production
	(1)	(2)	(3)	(4)	(5)	(6)
Lockdown × Female	-0.395**	-0.567***	-0.370**	-0.153	0.0141	-0.298
	(0.181)	(0.190)	(0.173)	(0.249)	(0.243)	(0.232)
Observations	48,879	48,879	48,879	32,583	32,583	32,583
Number of Individuals	657	657	657	437	437	437

Panel B. Employment Status

	Full-time Employed			Not Full-time Employed		
	Total	Leisure	Production	Total	Leisure	Production
	(1)	(2)	(3)	(4)	(5)	(6)
Lockdown × Female	-0.275	-0.491***	-0.244	-0.448	-0.0715	-0.664**
	(0.176)	(0.171)	(0.165)	(0.290)	(0.312)	(0.271)
Observations	59,140	59,140	59,140	22,322	22,322	22,322
Number of Individuals	792	792	792	302	302	302

Panel C. Job Search

	Full Sample		Not FT Employed and No Change	
	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	-0.0220**	-0.138**	-0.0404**	-0.275***
	(0.00941)	(0.0580)	(0.0171)	(0.104)
Observations	81,462	81,462	19,824	19,824
Number of Individuals	1,094	1,094	269	269

Notes. This table reports the main estimates for the differential effect of the lockdown on women for various samples. All regressions include individual and date fixed effects. Panel A splits the sample by parental status: adults with at least one child are in columns (1)-(3) and those with no children are in columns 4-6. Panel B presents separate estimates for the full-time employed sample in columns (1)-(3) and for others (including students and part-time employed) in columns (4)-(6). The dependent variables in Panels A and B are the natural log transformation plus 1 second of the outcome of interest. Panel C presents the results on the job search websites. In that panel, the outcome in columns (1) and (3) is an indicator variable for whether the person visited a job search website that day, and the outcome in columns (2) and (4) is the 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 were 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.

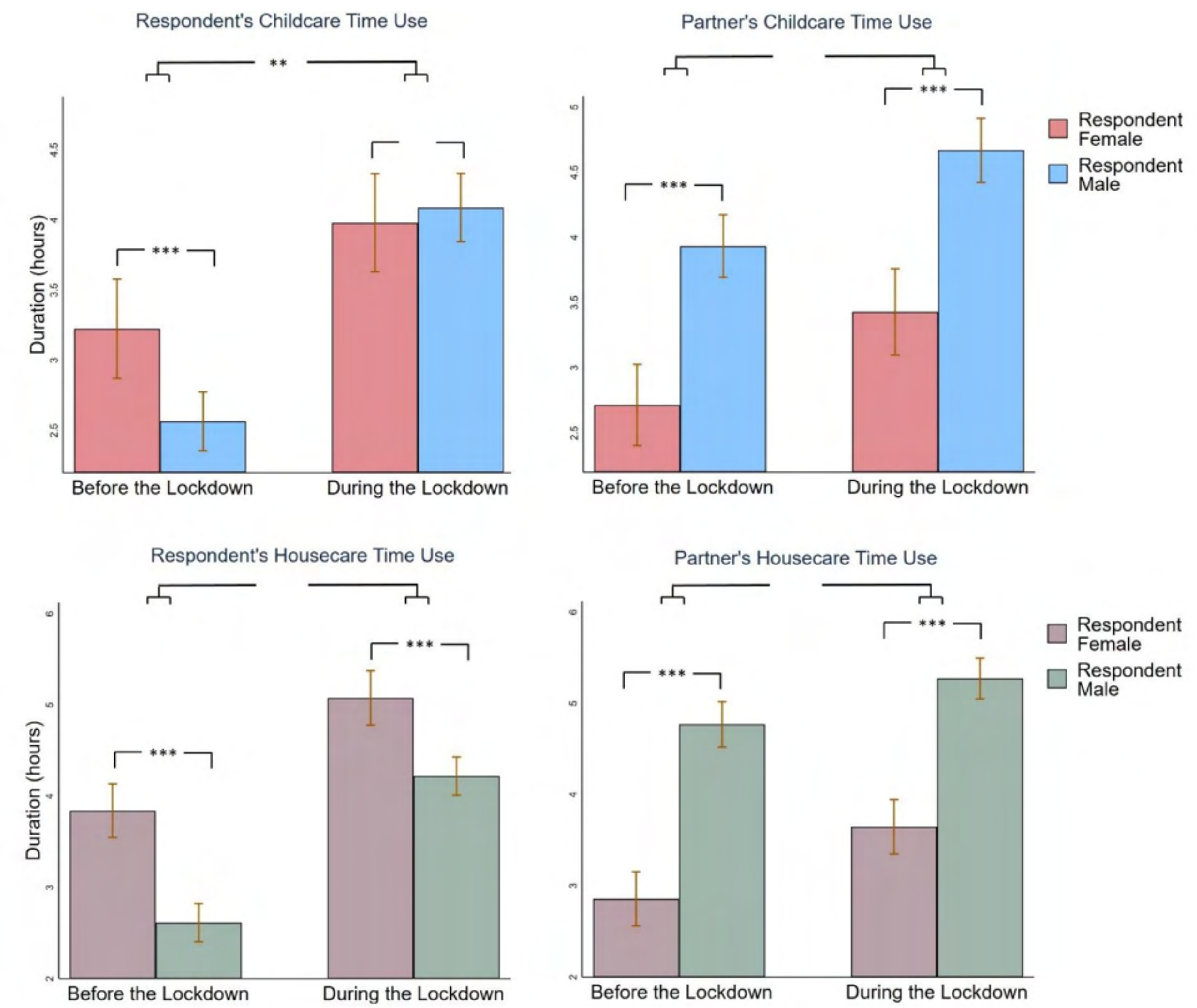

Figure 3 Summary Statistics for Survey-based Measures of Household Production Time

Notes. This figure presents the sample mean and 95% confidence intervals 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. Significance stars report the test for the equality of means. The significance chart on the top panel reports the results for the differential impact of the lockdown on women relative to men. The online survey was conducted during the lockdown period, so only the lockdown values are contemporaneous. Significance at *** p<0.01, ** p<0.05, * p<0.1.

Table 5 Child-Related Browser Usage

	Manual Dictionary *Full Sample*	Manual Dictionary *Sample with Children*	Word Embedding *Full Sample*	Word Embedding *Sample with Children*	YouTube Kids Channels *Full Sample*	YouTube Kids Channels *Sample with Children*
	(1)	(2)	(3)	(4)	(5)	(6)
Lockdown × Female	-0.773	-0.729	-0.882	-0.286	-0.151	-0.233
	(1.463)	(2.011)	(0.852)	(1.265)	(0.143)	(0.241)
Sample Mean (Parents)	7.059		4.600		0.150	
Sample Mean (Non-Parents)	4.947		2.601		0.015	
p-value for t-test: Parents = Non-Parents	0.043		0.002		0.001	
Observations	81,462	48,879	81,462	48,879	81,462	48,879
Number of Individuals	1,094	657	1,094	657	1,094	657
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

Online Appendix. Supplemental Figures and Tables

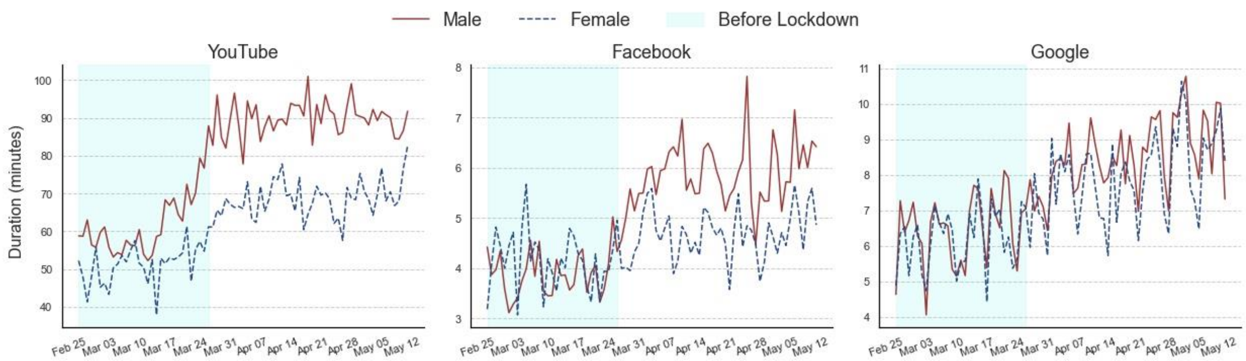

FIGURE A1. DAILY BROWSER TIME USE FOR MEN AND WOMEN BY CATEGORY

Notes. The pale blue shaded region represents the period before the COVID-19 lockdown in India on 25 March, 2020.

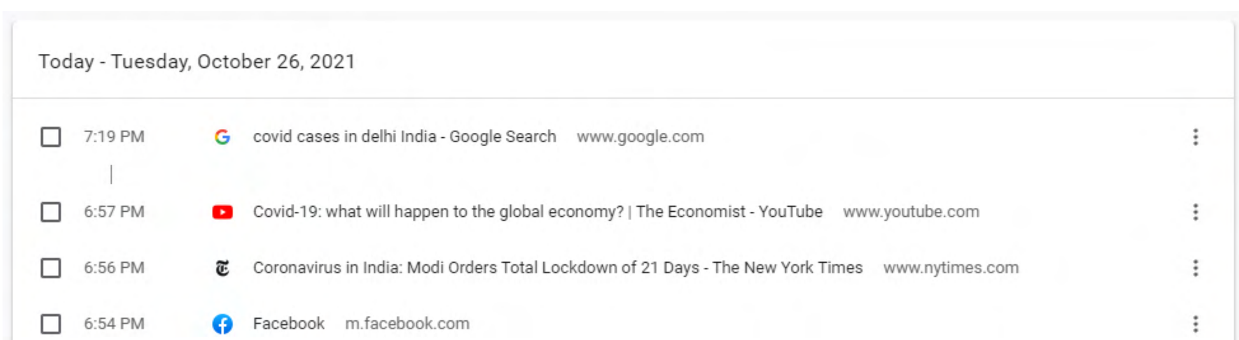

FIGURE A2. SAMPLE BROWSER HISTORY

Notes. This figure presents an illustrative example of a web browser history. Each time a user visits a URL, the date-time information, website title and domain information are saved in the browser history. The title information conveys meaningful information about the website that has been visited (e.g., the google search phrase in the top panel). If a user is logged into the browser app on their mobile devices, web browser history includes the usage under the mobile browser app (e.g., Facebook at the bottom of the panel. The m.facebook.com is the mobile version of Facebook, which is only accessible via a mobile device). Beyond what is available in the browser history, our dataset also includes the category information for each website domain and duration of each URL visit, provided by our industry partner PY Insights.

TABLE A1. ALTERNATIVE SPECIFICATIONS

	Total (1)	Leisure (2)	Production (3)	YouTube (4)	Facebook (5)	Google Search (6)
Panel A. Extensive Margin						
Lockdown × Female	-0.0240	-0.0324**	-0.0316**	-0.0361**	-0.0463***	-0.0378**
	(0.0149)	(0.0165)	(0.0160)	(0.0156)	(0.0108)	(0.0172)
Sample Mean	0.814	0.625	0.783	0.411	0.202	0.584
Observations	81,462	81,462	81,462	81,462	81,462	81,462
Number of Individuals	1,094	1,094	1,094	1,094	1,094	1,094
Panel B. Intensive Margin: Daily Browser Time						
Lockdown × Female	-0.0889**	-0.130*	-0.114**	-0.149*	-0.0805	
	(0.0386)	(0.0688)	(0.0465)	(0.0845)	(0.0849)	
Observations	66,266	50,889	63,820	33,453	16,425	
Number of Individuals	1,094	1,084	1,094	1,032	934	
Panel C. Intensive Margin: Daily Website Visits						
Lockdown × Female	-0.186***	-0.231***	-0.173***	-0.157***		-0.0565
	(0.0570)	(0.0589)	(0.0570)	(0.0518)		(0.0369)
Observations	66,266	50,889	63,820	31,596		47,837
Number of Individuals	1,094	1,084	1,094	1,029		1,082
Panel D. Levels: Daily Browser Time						
Lockdown × Female	-25.49***	-11.29*	-8.062**	-10.66**	-1.359**	
	(8.414)	(6.360)	(3.748)	(4.882)	(0.593)	
Observations	81,462	81,462	81,462	81,462	81,462	
Number of Individuals	1,094	1,094	1,094	1,094	1,094	
Panel E. Levels: Daily Website Visits						
Lockdown × Female	-41.77*	-12.37***	-17.34	-2.026***		-0.509*
	(21.64)	(3.650)	(20.18)	(0.464)		(0.304)
Observations	81,462	81,462	81,462	81,462		81,462
Number of Individuals	1,094	1,094	1,094	1,094		1,094
Individual fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Date fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes. Panel A focuses on the extensive margin and the outcome is an indicator variable for if the person visited any pages within the category. The next two panels focus on the intensive margin and use log-transformed measures of usage (time in Panel B and visits in Panel C) that drop zero values from the sample. The next two panels estimate models in levels and include zero values for usage (time in Panel D and visits in Panel E). Standard errors are clustered at the individual level. Significance at *** p<0.01, ** p<0.05, * p<0.1. All regressions include individual and date fixed effects.

TABLE A2. EFFECTS OF THE LOCKDOWN ON GENDER GAPS IN ONLINE LEARNING

	Browser Time (1)	Any Browser Time (2)	High Frequency (3)	High Frequency (4)	Frequency Category (5)
Lockdown × *Female*	-0.453	0.00424	0.0190	0.0518	0.0164
	(0.591)	(0.00921)	(0.0430)	(0.114)	(0.0966)
Lockdown			0.177***	0.460***	0.397***
			(0.0258)	(0.0684)	(0.0569)
Female			-0.0128	-0.0361	-0.0358
			(0.0292)	(0.0826)	(0.0678)
Sample Mean	2.76	0.123			
Observations	81,462	81,462	2,188	2,188	2,188
Number of Individuals	1,094	1,094	1,094	1,094	1,094
Individual fixed effects	Yes	Yes	No	No	No
Date fixed effects	Yes	Yes	No	No	No

Notes. Columns (1) and (2) present estimated gender differences in the effects of the lockdown on daily internet browser activity related to online learning websites using the person-day unit of observation and controlling for individual and day fixed effects. Browser time is measured continuously in minutes in column (1) and as an indicator for any time during the day in column (2). Columns (3), (4) and (5) examine changes in survey reports of frequency of "self-investment" activities (e.g., taking a course, teaching yourself a new skill, etc.) between the lockdown and period before the pandemic. The questions offered 4 options for response: almost never (16.7%), sometimes (42.8%), frequently (26.9%) and very frequently (13.7%). The outcome in columns (3) and (4) is a binary indicator for reporting high self-investment (frequently or very frequently), using a linear probability (3) or Probit (4) model. Column (5) reports estimates from an ordered Probit using all four categories. Robust standard errors reported in in parentheses in all columns, with clustering at the individual level in Columns (1) and (2). Significance at *** p<0.01, ** p<0.05, * p<0.1.

TABLE A3. EFFECTS OF THE LOCKDOWN ON GENDER GAPS: NO DEVICE SHARING SUB-SAMPLE

	Total (1)	Leisure (2)	Production (3)	YouTube (4)	Facebook (5)	Google Search (6)
Panel A. Daily Browser Time						
Lockdown × Female	-0.513**	-0.529**	-0.435**	-0.645***	-0.298***	
	(0.204)	(0.212)	(0.190)	(0.192)	(0.109)	
Sample Mean	229.9	116.7	103.5	155.7	6.1	
Panel B. Daily Website Visits						
Lockdown × Female	-0.356***	-0.322***	-0.275**	-0.235***		-0.143***
	(0.122)	(0.0957)	(0.117)	(0.0585)		(0.0540)
Sample Mean	250.7	45.5	168.6	6.3		5.6
Observations	42,419	42,419	42,419	42,419	42,419	42,419
Number of Individuals	564	564	564	564	564	564
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 and activity counts on the subsample of individuals who reported not sharing their devices (smartphone, tablet, computer) with others. Panel A presents the time use outcomes while Panel B presents the activity counts, measured as unique URLs. Dependent variables are the natural log transformation of 1 plus the outcome of interest (seconds or counts). Standard errors are clustered at the individual level. Significance at *** p<0.01, ** p<0.05, * p<0.1.

TABLE A4. EXPLORATION OF YOUTUBE VIDEO CATEGORIES

	Leisure Time use (1)	Production Time use (2)	Leisure URL count (3)	Production URL count (4)
Panel A. Daily YouTube Time				
Lockdown × Female	-0.337***	-0.280***	-0.112***	-0.0931***
	(0.115)	(0.0919)	(0.0336)	(0.0237)
Sample Mean (minutes)	32.469	14.678	3.299	1.473
Panel B. YouTube-Purpose-Adjusted Daily Browser Time				
Lockdown × Female	-0.346**	-0.308**	-0.246***	-0.264***
	-0.143	-0.142	-0.0667	-0.086
Sample Mean (minutes)	70.251	95.458	38.526	168.674
Observations	81,462	81,462	81,462	81,462
Number of Individuals	1,094	1,094	1,094	1,094
Individual fixed effects	Yes	Yes	Yes	Yes
Date fixed effects	Yes	Yes	Yes	Yes

Notes. This table presents the estimates and summary statistics that use information on content categories for YouTube videos. Outcomes in Panel A measure viewing of YouTube videos. Leisure and Production categories are aggregated based on 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. Daily average YouTube time use is 73.48 minutes (Table 2), which includes 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. Panel B presents the main estimates after reclassifying YouTube production videos to the overall production category. Columns 1 and 2 use duration-based measures and columns 3 and 4 use URL counts. Dependent variables are the natural log transformation of 1 plus the outcome of interest (seconds or counts). Standard errors are clustered at the individual level. The sample mean is at the person-day level and reported in levels (minutes or counts). Significance at *** p<0.01, ** p<0.05, * p<0.1.

TABLE A5. BREAKDOWN INTO DETAILED YOUTUBE VIDEO CATEGORIES

	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 ×	-0.0525	-0.0933	-0.0295	-0.261***	-0.237***	-0.141*	-0.259***	-0.0814	-0.207***
Female	(0.0470)	(0.0668)	(0.0409)	(0.0696)	(0.0811)	(0.0731)	(0.0688)	(0.058)	(0.0580)
Sample Mean (minutes)	3.372	7.550	2.663	5.937	9.582	9.399	6.329	3.840	4.510
Observations	81,462	81,462	81,462	81,462	81,462	81,462	81,462	81,462	81,462
Number of Individuals	1,094	1,094	1,094	1,094	1,094	1,094	1,094	1,094	1,094
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 is at the person-day level and reported in levels (minutes). Significance at *** p<0.01, ** p<0.05, * p<0.1.

TABLE A6. INTERNET BROWSER USE 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	198.0	220.3	208.7	249.5	192.9	231.7	219.9	231.7
	(214.9)	(229.2)	(229.2)	(246.3)	(212.3)	(238.0)	(235.1)	(231.1)
Leisure Time	94.49	99.81	119.2	138.3	89.19	109.1	133.3	134.1
	(167.7)	(175.0)	(194.7)	(205.3)	(159.3)	(185.4)	(209.4)	(196.7)
Production Time	77.74	90.67	64.86	77.22	77.50	91.19	62.99	66.11
	(103.7)	(110.3)	(86.01)	(96.78)	(105.1)	(110.3)	(79.18)	(84.55)
Observations	16,878	32,001	12,075	20,508	18,708	40,432	10,245	12,077

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

TABLE A7. ACCOUNTING FOR JOB LOSS AND WORKING FROM HOME

	Browser: Total Time use (1)	Browser: Leisure Time use (2)	Browser: Production Time use (3)	Survey: Childcare Time use (4)
Panel A. Exclude Job Loss				
Lockdown × Female	-0.293*	-0.357**	-0.324**	-0.762**
	(0.150)	(0.153)	(0.142)	(0.305)
Observations	78,233	78,233	78,233	1,104
Number of Individuals	1,050	1,050	1,050	552
Panel B. Browser Usage by Gender and Parental or Employment Status				
Lockdown × Female	-0.257*	-0.301**	-0.307**	-0.796***
	(0.150)	(0.153)	(0.141)	(0.308)
Lockdown × Job Loss	-0.617	-0.349	-0.381	-0.180
	(0.453)	(0.436)	(0.381)	(0.646)
Lockdown × Work from Home	0.124	-0.0510	0.158	-0.401
	(0.153)	(0.159)	(0.142)	(0.302)
Lockdown × Partner Job Loss	-0.0216	0.447	0.109	-0.502
	(0.440)	(0.362)	(0.452)	(0.751)
Lockdown × Partner Work from Home	-0.386**	-0.230	-0.294*	0.102
	(0.180)	(0.183)	(0.170)	(0.311)
Observations	81,462	81,462	81,462	1,146
Number of Individuals	1,094	1,094	1,094	573
Individual fixed effects	Yes	Yes	Yes	No
Date fixed effects	Yes	Yes	Yes	No

Notes. This table presents the results on daily internet browser time use and self-reported childcare time. Browser time is log-transformed adding 1 second, as described in Table 3. Childcare time is in hours, as described in Figure 3. Childcare time is only for married individuals with children. The regression model also includes un-interacted indicators for Female and Lockdown. Panel A reports estimates on a subsample that excludes people who reported losing a job in the prior 90 days. Panel B uses the full sample and adds separate controls for own and spousal job loss and starting to work from home in the past 90 days, interacted with the Lockdown indicator. Standard errors clustered at the individual level (for browser data) and robust standard errors (for survey data) are in parentheses. Significance at *** p<0.01, ** p<0.05, * p<0.1.

TABLE A8. ANALYSIS OF SURVEY-BASED MEASURES OF HOUSEHOLD PRODUCTION TIME

	Childcare time use		Housework time use	
	Own	Partner's	Own	Partner's
	Married Sample with Children	*Married Sample with Children*	*Married Sample*	*Married Sample*
	(1)	(2)	(3)	(4)
Female	0.660***	-1.219***	1.226***	-1.912***
	(0.208)	(0.199)	(0.183)	(0.197)
Lockdown	1.523***	0.735***	1.607***	0.502***
	(0.163)	(0.175)	(0.151)	(0.170)
Lockdown × Female	-0.767**	-0.0205	-0.370	0.287
	(0.300)	(0.289)	(0.260)	(0.273)
Constant	2.565***	3.934***	2.613***	4.769***
	(0.107)	(0.122)	(0.107)	(0.127)
Observations	1,146	1,146	1,374	1,374
Number of Individuals	573	573	687	687

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

TABLE A9. ROBUSTNESS FOR THE RESULTS OF TABLE 3: DROPPING THE PERIOD FROM THE WHO'S PANDEMIC DECLARATION TO THE INITIAL INDIAN LOCKDOWN

	Total (1)	Leisure (2)	Production (3)	YouTube (4)	Facebook (5)	Google Search (6)
Panel A. Daily Browser Time						
Lockdown × Female	-0.360**	-0.370**	-0.357**	-0.433***	-0.334***	
	(0.178)	(0.182)	(0.166)	(0.163)	(0.0828)	
Panel B. Daily Website Visits						
Lockdown × Female	-0.303***	-0.263***	-0.275***	-0.191***		-0.0938**
	(0.107)	(0.0813)	(0.102)	(0.0496)		(0.0460)
Observations	67,128	67,128	67,128	67,128	67,128	67,128
Number of Individuals	1,094	1,094	1,094	1,094	1,094	1,094
Individual fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Date fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes. This table presents the robustness checks on the main estimates in Table 3 for daily internet browser time use and activity counts. We dropped the period from 11 March, when the WHO declared COVID-19 as a pandemic, to 24 March, the date before the initial Indian lockdown. Panel A presents results for browser time use outcomes, and Panel B presents the activity counts, measured as unique URLs generated. Column (1) shows total browser use while subsequent columns are for categories: leisure (2), production (3), YouTube videos (4), Facebook (5; URLs not examined because extensive activity occurs within the main URL), and Google searches (6; time on search pages not examined because people typically follow links to results quickly). 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 A10. ROBUSTNESS FOR THE RESULT OF TABLE 3: ALTERNATIVE STANDARD ERROR CLUSTERING

	Total (1)	Leisure (2)	Production (3)	YouTube (4)	Facebook (5)	Google Search (6)
Panel A. Daily Browser Time						
Lockdown × Female	-0.292**	-0.326**	-0.337**	-0.344**	-0.295***	
	(0.147)	(0.150)	(0.139)	(0.137)	(0.0718)	
Panel B. Daily Website Visits						
Lockdown × Female	-0.280***	-0.254***	-0.265***	-0.160***		-0.0846**
	(0.0892)	(0.0677)	(0.0858)	(0.0419)		(0.0390)
Observations	81,462	81,462	81,462	81,462	81,462	81,462
Number of Individuals	1,094	1,094	1,094	1,094	1,094	1,094
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 and activity counts. Two-way standard errors are clustered at the participant and date level. See Table A9. for variable descriptions. 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.