7-30-2024

Digital Trace Data and Demographic Forecasting: How Well Did Google Predict the US COVID-19 Baby Bust?

Joshua Wilde
Portland State University

Wei Chen
Max Planck Institute for Demographic Research

Sophie Lohmann
Max Planck Institute for Demographic Research

Jasmin Abdel Ghany
University of Oxford

Citation Details

Wilde, J., Chen, W., Lohmann, S., & Abdel Ghany, J. (2024). Digital Trace Data and Demographic Forecasting: How Well Did Google Predict the US COVID-19 Baby Bust? Population and Development Review, 50(S1), 421–446. Portico.

Digital Trace Data and Demographic Forecasting: How Well Did Google Predict the US COVID-19 Baby Bust?

JOSHUA WILDE, WEI CHEN, SOPHIE LOHMANN AND JASMIN ABDEL GHANY

At the onset of the first wave of COVID-19 in the United States, the pandemic's effect on future birthrates was unknown. In this paper, we assess whether digital trace data—often touted as a panacea for traditional data scarcity—held the potential to accurately predict fertility change caused by the COVID-19 pandemic in the United States. Specifically, we produced state-level, dynamic future predictions of the pandemic's effect on birthrates in the United States using pregnancy-related Google search data. Importantly, these predictions were made in October 2020 (and revised in February 2021), well before the birth effect of the pandemic could have possibly been known. Our analysis predicted that between November 2020 and February 2021, monthly United States births would drop sharply by approximately 12 percent, then begin to rebound while remaining depressed through August 2021. While these predictions were generally accurate in terms of the magnitude and timing of the trough, there were important misses regarding the speed at which these reductions materialized and rebounded. This ex post evaluation of an ex ante prediction serves as a powerful demonstration of the "promise and pitfalls" of digital trace data in demographic research.

Joshua Wilde, Nuffield Department of Population Health, Leverhulme Centre for Demographic Science, Demographic Science Unit, University of Oxford, Oxford, UK. E-mail: josh.wilde@demography.ox.ac.uk. Research Scientist, Population Research Center, Portland State University, Portland, OR, USA. Affiliate Researcher, Laboratory of Fertility and Well-Being at the Max Planck Institute for Demographic Research, Rostock, Mecklenburg-Vorpommern, Germany. Research Fellow, IZA Institute for Labor Economics, Bonn, Nordrhein-Westfalen, Germany. Wei Chen, Sophie Lohmann, Department of Digital and Computational Demography, Max Planck Institute for Demographic Research, Rostock, Mecklenburg-Vorpommern, Germany. E-mail: victoriaweichen7@gmail.com; E-mail: s.lohmann@irees.de. Jasmin Abdel Ghany, Nuffield Department of Population Health and Nuffield College, Leverhulme Centre for Demographic Science, Demographic Science Unit, University of Oxford, Oxford, UK. E-mail: jasmin.abdelghany@nuffield.ox.ac.uk

Introduction and background

The COVID-19 pandemic has had significant consequences for human society. As of April 28, 2022, the COVID-19 pandemic has led to approximately 510 million confirmed cases and 6.2 million deaths worldwide (WHO 2021). In the immediate aftermath of the first wave in March and April 2020, most attention focused on the mortality and economic consequences of the pandemic, while its effects on fertility were mainly a matter of speculation (Aassve et al. 2020). Some in the popular media suggested the pandemic would result in a "baby boom" as couples spent more time together. Such pronouncements were viewed with skepticism by many demographers, citing evidence on the short-term fertility effects from other mortality crises, including natural disasters (Finlay 2009; Evans, Hu, and Zhao 2009; Nobles, Frankenberg, and Thomas 2015; Behrman and Weitzman 2016) and previous pandemics (Bertillon 1892; Chandra et al. 2018). These studies generally show that mortality spikes are followed by reductions in births within a year with some evidence of fertility rebounds after several years (Palloni 1988). However, in spite of the broad public and scientific interest in the effect of the COVID-19 pandemic on birthrates, there was a lack of concrete forecasts at national or subnational levels regarding the possible magnitude or timing of these hypothesized birthrate effects. In the United States, the first of these concrete predictions came in July 2020, in a short Brookings Institution report (Kearney and Levine 2020), and focused on aggregate national changes without any prediction regarding the timing of the predicted declines. The second prediction in October 2020—which included specific state-level, dynamic predictions at the month level—is the subject of this research article (Wilde, Chen, and Lohmann 2020).

The intense public interest in the effect of the pandemic on birthrates was partially due to the fact that fertility change has significant economic and social consequences (Ashraf, Weil, and Wilde 2013; Karra, Canning, and Wilde 2017). Birth reductions may accelerate population aging and increase dependency ratios in populations already far below replacement fertility (PRB 2020), potentially lowering economic growth in the face of smaller working-age populations and a higher private and public care burden (Maestas, Mullen, and Powell 2023; Beard and Bloom 2015). However, these social and economic effects are mainly affected by long-run fertility change. If the pandemic leaves lifetime births per woman unchanged and births are only postponed, the long-term economic effects from postponement should be minimal. However, recent postcrisis fertility declines—including the 2008–2009 financial crisis—did not experience a rebound and led to permanently lower fertility rates. Since the exact nature of the effect of COVID-19 on the future of human fertility was unclear, the economic and social effects of the crisis due to demographic change were also unknown.

As human gestation takes on average of 268 days, there was a natural delay from the onset of the COVID-19 pandemic crisis and its effect on full-term births (Jukic et al. 2013). For example, full-term births from conceptions realized during the rapid onset of the pandemic in February or March of 2020 would not appear until November or December. This delay in understanding the effect of the pandemic on fertility was further compounded since natality data do not become available for analysis instantaneously. For example, the US Natality File birth microdata from the National Center for Health Statistics (NCHS) are generally released at least 6–9 months after the end of the year in which those births occurred. As a result, the full data from 2020 was not available for analysis until September 2021. Even in countries that have faster data releases (such as in European countries with registered data), significant delays between the advent of births and the release of data hamper the ability of researchers to make timely analyses of the relationship between COVID-19 and fertility rates.

Digital trace data have long been hoped to fill these gaps through now-casting methods where data are scarce or delayed. An emerging literature suggests that these sources—and Google search data in particular—can be used to monitor a number of social and biological phenomena in the absence of more reliable or timely data (Cesare et al. 2018). Such data have been used for real-time analyses of disease outbreaks such as the seasonal flu and Dengue (Ginsberg et al. 2009; Carneiro and Mylonakis 2009) as well as studies on well-being (Brodeur et al. 2021), tourism (Siliverstovs and Wochner 2018), financial trading behavior (Preis et al. 2013), and demographic processes such as fertility (Billari, D'Amuri, and Marcucci 2016; Ojala et al. 2017; Lin, Cranshaw, and Counts 2019), migration (Zagheni and Weber 2012; Wladyka 2017), sexual behavior (Markey and Markey 2013; Wilde, Apouey, and Jung 2017; Stephens-Davidowitz 2017), mortality (Tamgno, Faye, and Lishou 2013; Ricketts and Silva 2017), and suicide (Solano et al. 2016). Moving beyond now-casting with Google data is generally difficult due to the complexity or uncertainty surrounding the long-term processes which govern such phenomena.

Google search data are particularly appealing for predictions regarding fertility for multiple reasons. First, most phenomena can only be now-casted and not forecasted using these methods (Choi and Varian 2012). However, since behavior and information-seeking surrounding human gestation takes place in predictable phases with well-known time lags, now-casting conceptions can help us to predict birth 7–9 months into the future. Second, inasmuch as sexual behavior and conception as well as their discussion and scientific investigation may be considered a social taboo by many, some individuals may be more willing to search for such information on the Internet than to discuss their behavior in person (Stephens-Davidowitz 2017). For example, there is a well-known disconnect between fertility intentions and fertility behavior (Morgan and Rackin 2010). Therefore, although Google data are only an imperfect reflection of the behaviors that affect fertility, it may more accurately reflect actual behavioral change than a direct self-report of those behaviors. Finally, while most economic and demographic indicators are subject to lengthy time lags in data compilation and release, Google search data are immediately available and free to all, making it an ideal data source for a broad range of now-casting and forecasting applications.

In addition to uncertainty surrounding the pandemic's effect on aggregate fertility, there could have also been significant heterogeneous effects across subnational regions, or between socioeconomic groups. For example, COVID-19 incidence and mortality have been elevated among the Black or African American community in the United States, and the economic impacts have been particularly acute for workers with lower levels of education (CDC 2020; Finch and Finch 2020). Additionally, while planned births may fall as the economic fallout of the pandemic induces couples to delay child-bearing, reduced contraceptive access may lead to an increase in unintended pregnancies. This effect is particularly acute for areas with historically poor contraceptive access: A 2020 UNFPA report noted that COVID-19 had already exacerbated unmet family planning needs due to a variety of reasons, including decreased demand for health facility visits, unavailability of trained medical staff, and supply chain disruptions for contraceptive commodities (UNFPA 2020). Analyzing differential changes in Google search volumes across regions with heterogeneous populations may yield early insights into the potential mechanisms behind the effects of COVID-19 on birthrates.

While digital trace data have been touted as a possible panacea for data scarcity with its promise of now-casting social, economic, and health phenomena, significant skepticism abounds. Following an explosion of research advocating digital trace data to study social phenomena, there was a significant backlash against their use.[1] The most high-profile of these cases was the supposed failure of Google Flu, which stopped being predictive after just several years due to model overfitting run amok (Lazer et al. 2014). In addition, while most research with digital trace data have focused on either ex post tests of predictive fit or now-casting exercises, tests of digital data's ex ante predictive capacity are rare.

In October 2020, we did precisely this. Using state-of-the-art prediction methodologies, we put the promise of digital trace data to the test by making an ex ante prediction on the future of US fertility (Wilde, Chen, and Lohmann 2020) as a result of the COVID-19 pandemic using digital trace data. Our predictions were very specific: they were by state and month, up to seven months into the future.[2] Our forecasting model was primarily based on current Google search volumes for a set of keywords relating to conception, pregnancy, childbirth, and economic stability. The idea behind this strategy is simple: if one observes a sharp increase in searches for conception-related terms such as "pregnancy test," "missed period," "ovulation," or "abortion," one might expect a corresponding change in births seven to nine months later.

We did this in four steps. First, we showed that before 2019, periods of above-normal search volume for Google keywords relating to conception and pregnancy were associated with higher numbers of births in the following months at the expected time lags. Excess searches for unemployment keywords just before possible conception had the opposite effect. Second, by employing simple machine learning techniques, we demonstrated that including information on keyword search volumes in prediction models significantly improved forecast accuracy over a number of cross-validation criteria. Third, we used data on Google searches during the COVID-19 pandemic to predict changes in aggregate fertility rates in the United States at the state level through February (later August) 2021. Finally, we showed our predictions were heterogeneous in understandable ways across states by sociodemographic characteristics such as income, education, racial or ethnic identity, and COVID-19 caseload.

In this paper, we present our previous process and prediction and provide a postmortem.[3] By so doing, we hope to make a larger statement regarding the "promise and pitfalls" of using digital data in demography (Cesare et al. 2018). Specifically, while many of the predicted effects came to pass almost exactly as forecast, there were a number of significant misses. This rare ex post evaluation of a real-time ex ante prediction serves as a powerful demonstration that digital data can indeed be used in forecasting, significantly improve model forecasts, and be useful indicators of population behavior; yet they are not a panacea for traditional problems of scare data and forecasting error as many would hope.

Data and methods

In the spirit of keeping the main text accessible and compelling, we moved many of the specifics about the methods and data to Supporting Information. However, we include a brief overview here.

Data sources

Our data on keyword search frequency comes from Google Trends (http://trends.google.com). We use monthly searches at the state level in the United States since data for smaller geographic regions are often suppressed due to low search volumes. The Google data began in January 2004 and ended at the time our model was run (either July 2020 or January 2021). The model was fit only using pre-Covid data, defined by us as December 2019.

Data on births by state and month come from the National Vital Statistics System (NVSS) of the National Center for Health Statistics (NCHS).

Since 2004 is the first year Google search data are available, we use monthly birth counts for each US state and DC from 2004 to 2019. This yields 16 years of data across 51 geographic regions or 9792 possible state-month-year observations.

Prediction model

Our prediction model utilizes a time-lagged fixed-effects regression controlling for state-specific time-invariant effects, common seasonal birth variation across states, and state-specific linear growth trends. Specifically, we estimate the following prediction model:

$$Y_{smy} = \alpha_s + \theta_m + \gamma_s * t \ + \sum_{w}^{W} \sum_{l=t_0}^{T} \beta_{w,t-l} \ln I_{smy}^{w} + \epsilon_{smy},$$

where s, m, and y are index state, month, and year, respectively, α_s is a state fixed effect, θ_m is a month fixed effect, and $\gamma_s * t$ are a set of state-specific time trends. The double summation represents a series of β coefficients for the natural log of the normalized search volume for keywords within keyword set W at a number of monthly time lags. The dependent variable Y_{smy} is the natural log of births in a given state for a certain month and year, implying the interpretation of the βs as an elasticity—the percentage change in births from each percentage change in search volume. For our prediction model we use t_0 of 7 and a T of 12, representing monthly time lags from 7 to 12 months before birth. Heteroskedasticity-robust Huber–White standard errors are utilized for the prediction model unless otherwise stated.

Given the state and month fixed effects, this regression controls for national seasonality in both births and keyword search volumes. In essence, it estimates the effect on births of larger-than-normal search volume for a given month, in a given state, compared with that same month across years. It also controls for changes in aggregate births over time specific to the state due to the linear time trends. We selected this specification since it minimized mean-squared prediction error (MSPE) relative to models with nonlinear growth trends and state-specific seasonal effects due to model over-fitting. More information on the different models we tested can be found in the Methodology section of the Supporting Information.

We chose to use the latest time lag of the independent variable at seven months before birth to include information-seeking in early pregnancy, but no later in order to predict births as far in advance as possible. To demonstrate associations between keywords and births during trimesters besides the first, we include monthly search volume lags up to one month before birth in some specifications not used for prediction. Controls for overall search volume—a proxy for general Internet usage—are also included, in addition to monthly unemployment rates in some specifications.

Keyword selection and model cross-validation

Keyword sets were created in a multistep process. We initially consulted the literature for words which may be predictive, specific, and common enough for use in forecasting. Keywords were grouped into seven categories: unplanned pregnancy, pregnancy intention, prenatal services, pregnancy symptoms, pregnancy termination, unemployment, and other. We then sought input from other experts to inform ourselves of any important topics or keywords we might be missing. Due to the necessarily arbitrary initial keyword set, all keywords were preselected before looking at Google data to avoid data mining or hypothesizing after the results are known (commonly known as HARKing). Both topic and keyword searches were used: keywords utilized search data using Boolean logic operators, while topics include searches for the keyword in question but also include related queries without requiring an exact character match. For example, common misspellings, translations into other languages common to the geographic area in question, and other closely related keywords are included under a topic search. Further details and the list of initial keywords or topics are given in Table 1.

We then tested each keyword for goodness of fit and predictive power through three screens. The first screen omitted any keyword for which more than 30 percent of the 9792 state-month observations were missing. This Data (D) screen selected 37 of the 44 initial words. Second, we added all lags of search volumes for months 7–12 before birth for all 37 of these words to our prediction model and omitted any keywords whose 7–12 months lags were not jointly significant at the 5 percent level, or which failed to have any lag between 7 and 12 which was not individually significant at the same threshold. This process was done iteratively, such that the model was reestimated with the smaller keyword set, and words which then did not meet the two significance criteria were again omitted. This Significance (S) screen selected 16 keywords, listed in alphabetical order: ClearBlue, Condom, Divorce, Emergency Contraception, Human chorionic gonadotropin (hCG), Intrauterine Device (IUD), Medical Abortion, Morning Sickness, OBGYN, Online Dating, Ovulation, Porn, Pregnancy, Sexually Transmitted Infection (STI), Ultrasound, and Unemployment. Information on this process and which words were omitted each round is given in Supporting Information Table 3.

The third screen is called the MSPE Reduction (M) screen, which uses a machine learning procedure called forward stepwise selection to select keywords from the Significance screen which most reduced MSPE. To find the

TABLE 1 Keyword selection criteria

		Screen					Screen		
Word	**Topic or keyword**	**D**	**S**	**M**	**Word**	**Topic or Keyword**	**D**	**S**	**M**
Unplanned Pregnancy					**Pregnancy Termination**				
Emergency Contraception	Topic	X	X		Baby Heartbeat	Keyword			
"Morning After Pill"	Keyword				Cytotec	Keyword			
Plan B	Keyword	X			Dilation and Curettage	Topic	X		
STI	Topic	X			Medical Abortion	Topic	X	X	X
"Unprotected Sex"	Keyword				Mifepristone	Topic			
Online Dating[a]	Keyword	X			Miscarriage	Topic	X		
					Misoprostol	Topic	X		
Pregnancy Intention					Pregnancy Bleeding	Keyword	X		
Clearblue	Topic	X	X	X					
Conceive	Keyword	X			**Unemployment**				
HCG - Weight - Diet	Keyword	X	X	X	File Unemployment How	Keyword			
IVF	Topic	X			Furlough	Keyword			
Ovulation	Topic	X	X		Layoff	Topic	X		
Ovulation Test	Keyword	X			Unemployment	Topic	X	X	
Pregnancy Test	Topic	X			"Unemployment Office"	Keyword	X		

/...

TABLE 1 (Continued)

Word	Topic or keyword	Screen D	Screen S	Screen M	Word	Topic or Keyword	Screen D	Screen S	Screen M
Prenatal Services					**Other**				
BabyCenter	Topic	X			Birth Control	Topic	X		
Folate	Topic	X			Condom	Topic	X	X	
Midwife	Topic	X			Divorce	Topic	X	X	
Obgyn	Keyword	X	X		Gestation	Topic	X		
Prenatal Vitamins	Topic	X			IUD	Topic	X	X	
Ultrasound	Topic	X	X	X	Menstruation	Topic	X		
					Porn	Keyword	X	X	X
Pregnancy Symptoms					Pregnancy	Topic	X	X	X
Missed Period	Keyword	X			Pregnant	Keyword	X		
Morning Sickness	Topic	X	X						
Pregnancy Symptoms	Keyword	X			**Controls**				
					Control Set[b]	Keyword	X	X	X

NOTES: This table provides information on the specific keywords used in the analysis. Quotations around a keyword indicate that the returned search volume reflects only exact character matches, including word order. A minus sign indicates the turned search volume for the first word, but only if the search did not also include the words after the minuses. A plus sign indicates the Boolean OR operator, meaning the search volume contained searches which contained any of the listed words. A space between words indicates the Boolean AND operator, meaning searches were only included if they contained all listed words. Both Topic and Keywords searches were used: Keywords utilized search data using the Boolean logic operators outlined above, while Topics include searches for the keyword in question but also include related queries without requiring an exact character match. For example, common misspellings, translations into other languages common to the geographic area in question, and other closely related keywords are included under a topic search. D = Data screen (data for that keyword or topic was not missing for more than 30 percent of the sample); S = Significance screen (at least one monthly lag coefficient between 7 and 12 months before birth which was significant at the 5 percent level using standard Huber–White errors AND where all lags between 7 and 12 months were jointly significant at the 5 percent level, while also including all other Screen S keywords); M = Mean-squared prediction error reduction screen (selected by MSPE-minimizing forward stepwise machine learning keyword selection algorithm).

[a] Online Dating refers to a composite keyword search for the seven most popular online dating platforms used over our sample period. The query in Google Trends is equivalent to the Boolean logic query: match.com OR tinder OR okcupid OR bumble OR zoosk OR eHarmony OR POF OR plentyoffish.

[b] Control refers to a composite of the 10 most searched general words in the English language in Google. (See the Google Data section in the Supporting Information for details on the theoretical rationale behind this set of controls.) The query in Google Trends is equivalent to the Boolean logic query: the OR of OR how OR an OR is OR what OR and OR me OR you OR do.

MSPE reduction, we use eightfold time series split cross-validation, which splits the sample into a training set and test set eight times, and calculates the prediction error as the absolute value—in percentage terms—of the difference between these predictions and the actual births for every observation in our test set.

The forward stepwise selection learning procedure first estimates the base model without any keywords. Then, we add each keyword one at a time and employ the MSPE calculation procedure described above. The word which minimizes the MSPE is selected and becomes part of the base model. Then a second round begins, where each remaining word is added one at a time, and the word which minimizes MSPE is selected. This procedure continues until the additional cumulative reduction in MSPE by adding an additional keyword is less than one percentage point. This screen selected seven keywords, listed in order of selection: Unemployment, Clearblue, Ultrasound, Medical Abortion, hCG, Pregnancy, and Porn. Information on this process is given in Supporting Information Table 4

Once the model is estimated for each keyword set using data from January 2004 to December 2019, the coefficients from this model are applied to current Google Search volume data to provide a prediction seven months in advance.[4] Due to low search volumes, data for some important keywords are missing for 14 states, leading to missing predictions. These are low-population states, whose combined population comprises only 5 percent of the US population.[5]

Results (February 2021 update)

To illustrate the association between specific keywords and later births, we estimate the month-specific effect of relevant keywords individually on births for each month between 1 and 12 before the observed state-level birth count and plot these coefficients for a subset of keywords in Figure 1. Specifically, we plot these associations for seven of the most predictive keywords (ClearBlue, Medical Abortion, Misoprostol, Morning Sickness, Ovulation, Pregnancy, Unemployment) and the unemployment rate. We show two types of 95 percent confidence intervals—one with heteroskedasticity-robust Huber–White standard errors, and another for the implied 95 percent confidence intervals adjusting for multiple hypotheses using sharpened false discovery rate q-values.[6]

In general, these figures largely confirm intuitive associations between these keywords and future births. For example, a doubling of excess searches for ClearBlue—the name of a popular pregnancy test brand—is associated with 0.39 percent (95 percent CI: 0.11–0.67 percent) more births nine months later, 0.53 percent (95 percent CI: 0.25–0.81 percent) more births eight months later, 0.50 percent (95 percent CI: 0.23–0.78 percent) more births seven months later, and 0.46 percent (95 percent CI:

FIGURE 1 Fertility keyword searches and later births

NOTES: These figures show regression coefficients between births in a state and Google keyword search volume for the preceding 12 months as a solid black line. Coefficients are elasticities and can be interpreted as the percentage change in births due to a doubling of excess keyword search volume in a given month. Dashed lines represent 95 percent confidence intervals using Huber–White standard errors, while dotted lines represent 95 percent confidence intervals from inferred standard errors when correcting for multiple hypotheses using sharpened false discovery rate q-values

0.18–0.73 percent) more births six months later. This corresponds to the first trimester when most pregnant women experience their first pregnancy symptoms. Similarly, excess searches for "Morning Sickness" are associated with more births seven months later (0.46 percent, 95 percent CI: 0.17–0.75 percent), roughly consistent with when morning sickness most often occurs.

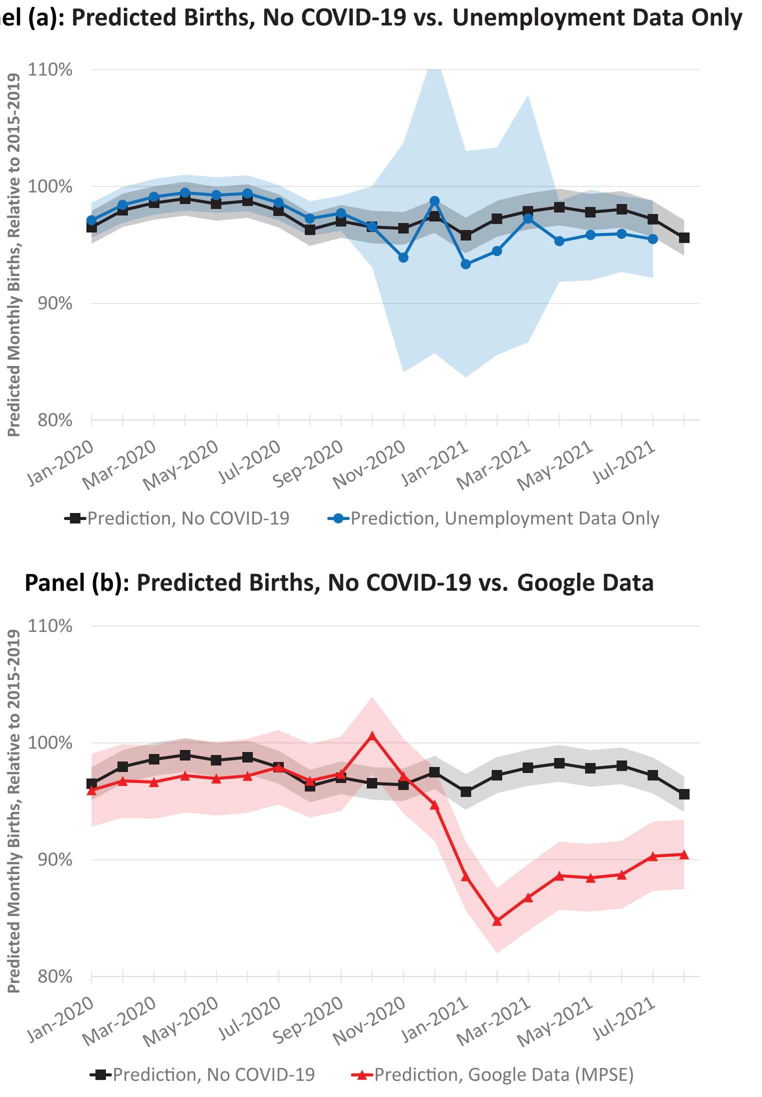
Searches for Unemployment immediately preceding conception are strongly negatively associated with births. A doubling of excess searches for Unemployment nine months before birth corresponds to 0.82 percent (95 percent CI: 0.38–1.26 percent) fewer births, while at months 10 and 11 correspond with 1.14 percent (95 percent CI: 0.70–1.58 percent) and 0.77 percent (95 percent CI: 0.31–1.23 percent) fewer births, respectively. Interestingly, the actual unemployment rate is uncorrelated with later births for any month between 1 and 12 months before birth.

We use these estimated associations between future births and pregnancy-related keyword searches to forecast state-level births through August 2021 using Google search volumes up to January 2021 for seven different models, some of which utilize information on Google keywords. When incorporating these keywords, we generally utilize two different keyword sets: the Significance set and the MSPE set. The list of words and methodologies used to select these keyword sets are given in Section Data and Methods and in Table 3 and 4 in the Supporting Information. Briefly, the Significance set is a set of keywords which were significantly associated with future births, and the MSPE set is the keyword set selected by a machine-learning algorithm based on their ability to reduce MSPE.

The full results for all seven models are displayed in Figure 7 in the Supporting Information, but the most important subset is shown in Figure 2. For each of these models, aggregated state-specific predictions are shown relative to the average level of births between 2015 and 2019 for each month.[7] The seven models are (1) no Covid, which estimates a model using only monthly seasonal variation in births and state-specific linear growth trends; (2) unemployment rate only, which adds information on actual unemployment rates to the no Covid model; (3) significance, which adds information on keywords from the Significance screen to the no Covid model; (4) MSPE, which adds information on keywords from the MSPE screen; (5) Early, which removes keywords related to clinical pregnancy services from the significance model which may have been closed during lockdown or avoided due to fears surrounding infection; (6) unemployment searches only, which adds information on searches for unemployment and a control to the no Covid model; and (7) MSPE, no unemployment, which is the same as the MSPE model but omits searches for unemployment.

Figure 2 clearly demonstrates the usefulness of Google data in prediction. In panel a, two models which do not incorporate Google data are shown. Predictions which utilize information using only unemployment

FIGURE 2 Predicted US births by month



NOTES: National predicted births, relative to average monthly births from 2015 to 2019 for that month, for various prediction models. The three models are No COVID-19—a baseline model in which births follow normal seasonal patterns and remain on state-specific annual trends: Unemployment Data Only—which adds information on monthly unemployment rates to the No COVID-19 model; and Google Data (MPSE)–which uses information on search volumes for Google keywords selected from a mean-squared prediction error minimizing forward stepwise machine learning selection method. Shaded areas represent 95 percent standard error bands

rates show almost no change in births as a result of the pandemic and are highly uncertain, since while annual changes in unemployment rates are strong predictors of fertility (Sobotka, Skirbekk, and Philipov 2011), short-run monthly deviations are not, as shown in Figure 1.

However, when incorporating information on Google keywords from the MSPE screen (Figure 2, panel b), we predict a large and highly significant decline in birthrates beginning in December 2020, reaching a decline of 15.2 percent (95 percent CI: 12.4–18.0 percent) from their 2015–2019 average by February 2021. This is equivalent to a 12.3 percent decline from 2019 alone. After February, a tepid rebound is predicted to begin, and by August 2021 predicted births are still 9.5 percent (95 percent CI: 6.6–12.5 percent) below the average 2015–2019 level. These predictions imply that aggregate births between December 2020 and August 2021 will be 240,072 (95 percent CI: 161,946–326,212) fewer than their 2015–2019 average, an annualized reduction of 8.7 percent (95 percent CI: 5.8–11.6 percent). This is similar yet smaller than the June 2020 prediction by Kearney and Levine (2020). Importantly, these results are robust to different keyword sets. In Figure 7 in the Supporting Information, we show all seven model predictions simultaneously and find very little difference in the predictions between the three models which utilize both unemployment and pregnancy-related keywords. Further, we show that models which only utilize search terms on unemployment, or only search terms on pregnancy, reduce MSPE significantly less and predict smaller declines in fertility than those which utilize both, highlighting the importance of including both keyword sets in prediction models.

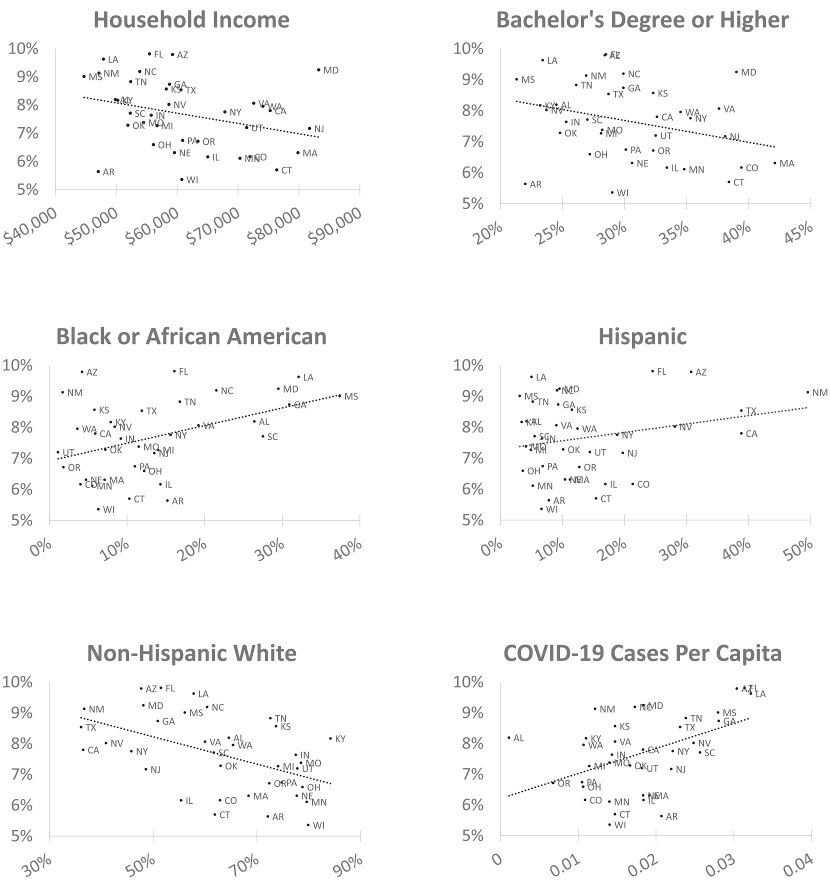
Finally, we explore whether our predictions vary systematically across states. We consider four sources of heterogeneity—education, income, race/ethnicity, and COVID-19 caseload. We find sizable differences across US states: the largest predicted annualized decline between December 2020 and August 2021 using the MSPE reduction set in Florida (9.8 percent), while the smallest decline is found in Wisconsin (5.4 percent). As shown in Supporting Information Figure 8, the largest predicted declines are generally in the Southern US and the Sun Belt. In Figure 3, we present state-level scatter plots of predicted annualized birth decline and six sociodemographic characteristics: median household income, a fraction of adults with a bachelor's degree or higher, a fraction of Black or African American, a fraction of Hispanic, a fraction of non-Hispanic White, and per capita COVID-19 cases.[8] States with characteristics associated with lower socioeconomic status, larger minority populations, and more COVID-19 cases per capita show larger predicted birth declines.[9]

Discussion of predictions

Ex ante prediction credibility

A double-digit collapse in births over a four-month period would have been the largest experienced in the United States in over a century. Therefore, it may be useful to explore the ex ante plausibility of such a bold prediction.

FIGURE 3 Predicted state fertility decline and various mother characteristics



NOTES: These panels show scatter plots relating the aggregate predicted percentage reduction in statewide births between November 1, 2020, and August 31, 2021 (vertical axis) with various state-level characteristics (horizontal axis). COVID-19 cases per capita only reflect cases before October 31, 2020, nine months before our final estimate.
SOURCE: American Community Survey 2013–2017, CDC, authors calculations.

To do so, we consider three major crises in the United States with similarities to the current pandemic: the Spanish Flu pandemic of 1918–1919, the Great Depression of 1929–1933, and the financial crisis of 2008–2009. In doing so, it is important to keep in mind that the epidemiology of the SARS-CoV-2 virus, the dynamics of the economic fallout, and the social context are fundamentally different for the current pandemic relative to these three crises. Access to modern contraception during the financial crisis and the COVID-19 pandemic also enabled many couples to avoid pregnancy. Comparisons should, therefore, be made with caution. However, as the coronavirus pandemic itself is in many ways unprecedented, these three crises arguably serve as the best comparisons at our disposal.

We first consider the 1918–1919 H1N1 influenza A pandemic—commonly known as the Spanish Flu. Birth rates fell sharply over the first few months of each of the three waves of the pandemic: 18.2 percent in the first wave (August–December 1918), 15.0 percent in the second (March–July 1919), and 12.1 percent in the third (July–November 1920) (Linder and Grove 1947). These declines are very similar to our model prediction for the first COVID-19 wave. The fertility effects of the Great Depression are also very similar in magnitude to the Spanish Flu, although they materialize over a much larger time frame: between 1929 and 1933, birthrates fell by 12.2 percent (Linder and Grove 1947). Finally, after the financial crisis of 2008–2009, births fell by 9.3 percent by 2013 and failed to rebound thereafter.

The evolution of births over these three crises suggests that a short-term 12.3 percent decline in births in response to the COVID-19 pandemic is not unreasonable, and an annualized decline of 8.7 percent is well within the historical range of birth decline. Calculating the elasticity of births to unemployment—defined as the ratio between the percentage change in births and the percentage change in unemployment—yields an elasticity of −0.027 for the current pandemic, precisely equal to the elasticity during the Great Depression (−0.027) and four times smaller than that of the financial crisis (−0.109). Importantly, a further rebound in births past August would reduce the annualized fertility decline to less than 8.7 percent, further lowering this elasticity.

Therefore, the distinguishing characteristic of the predictions in this study is not the magnitude—but rather the speed—of the decline. And since the current pandemic has led to historically fast increases in unemployment, rapid decreases in births should be expected.

Our finding that states with higher proportions of individuals identifying as Hispanic, Black, or African American, or with lower income or educational attainment have larger predicted declines in fertility can be interpreted in two ways. First, areas with high concentrations of individuals from racial or ethnic minorities have been more impacted by COVID-19 and, therefore, will suffer larger birthrate effects. For example, COVID-19 incidence among Black individuals was 2.6 times higher than for non-Hispanic whites and mortality was 2.1 times higher (CDC 2020). In addition, individuals of lower socioeconomic status are disproportionately affected by the virus, and this could manifest in higher predicted fertility declines (Finch and Finch 2020). Second, it may be that the heterogeneity in the predicted birth effect is not caused by the differential impact of the virus itself, but rather by the uneven economic fallout of the pandemic. Historically, during recessions, employment losses are concentrated among low-income and minority groups, and the current economic downturn is no exception. Between February and April 2020, the unemployment rate among non-Hispanic White Americans rose from 3.1 percent to 14.2 percent, while for Black or African Americans it rose from 5.8 percent to 16.7 percent. While these initial increases were similar, the recovery has been much slower for Black or African Americans: by August of 2020 non-Hispanic White unemployment had fallen back to 7.3 percent, yet Black unemployment remained elevated at 13.0 percent. The patterns in unemployment between educated and uneducated workers were even more striking, peaking at only 8.4 percent for those with a bachelor's degree or higher, compared with 17.3 percent and 21.2 percent for noncollege graduates with and without a high school diploma, respectively (BLS 2020).

Ex ante lessons for demographic prediction and digital data

Beyond providing the first dynamic birth forecast for the COVID-19 pandemic, our prediction provided three important insights regarding the potential of using Google Data in demographic forecasting. First, we demonstrated that certain keywords are indeed associated with and highly predictive of births. As shown in Table 4 in the Supporting Information, adding information on searches for the topic unemployment and a control to the model based only on seasonality and time effects reduces MSPE by 72.7 percent, while additionally adding keywords from the MSPE screen further reduces MSPE by an additional 41.6 percent.

Second, while data on unemployment are commonly used to predict birth rates at the annual level, we demonstrate that (1) short-run variation in unemployment rates is uncorrelated with short-run changes in birthrates (Figure 1) and (2) predictions utilizing such data are highly uncertain (Figure 2, panel a). This poor predictive power of short-run unemployment rates stems from the fact that, historically, changes in conceptions due to macroeconomic disruptions begin to fall several quarters before the unemployment rate registers such disruptions (Buckles, Hungerman, and Lugauer 2021). However, Google searches for the topic of unemployment are highly correlated with and predictive of birthrates, as shown in Figure 1. One possible explanation of this phenomenon is that if individuals perceive that a period of economic uncertainty may occur, they may both alter their fertility plans and increase information seeking regarding unemployment before the economic crisis actually occurs. These findings contribute to a broader literature which contrasts the effects of perceived versus objective economic factors, but which has so far rarely distinguished the perceived risk of unemployment from actual unemployment in a causally credible way (Tan et al. 2020).

Third, Google data are particularly effective in improving prediction accuracy during crises. One possible concern with using Google searches to predict the effect of the pandemic on birthrates is that search behavior during the pandemic may be fundamentally different from that in the past. However, we can test how the inclusion of Google keywords affected prediction accuracy during the last fertility crisis in the United States—the Great Recession of 2008–2009. Figure 6 in the Supporting Information demonstrates that prediction error using a model including Google keywords rose less during the Great Recession than one without. Additionally, the same figure shows that the model using Google keywords was able to accurately predict the trend change in fertility which occurred in 2018–2019, whereas the model without did not.

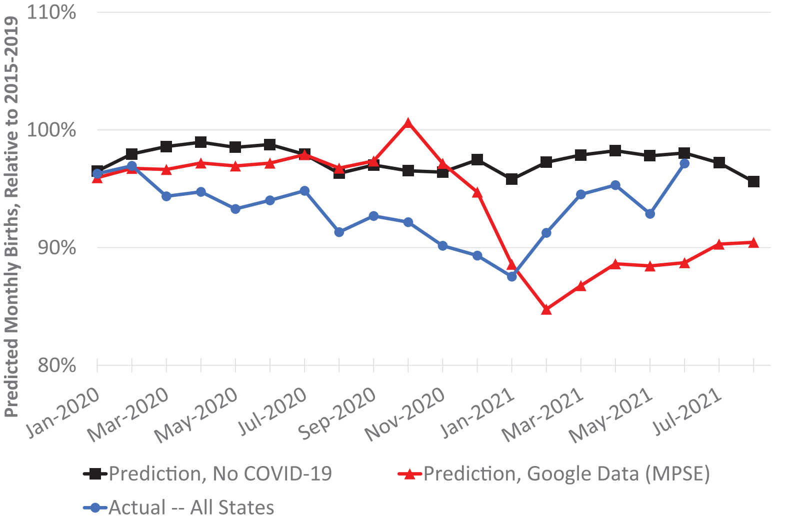
An important limitation of using pregnancy-related Google search data for fertility prediction is that—by construction—it can only predict fertility change at most 7–10 months in the future, leaving the long-run effects unknown. Even if pandemic birth reductions were simply driven by birth postponement and these birth postponements were only of short-term duration under unchanged fertility intentions, worries about permanent changes in quantum persist for several reasons. First, since the 1970s there has been a shift towards births at later ages as women postpone childbearing for educational or career motives (GBD 2017 Population and Fertility Collaborators 2018). As fecundity declines rapidly towards the end of a woman's childbearing years, women who delay childbearing due to the pandemic may face unexpected difficulties trying to conceive, leading to unintentionally lower completed lifetime fertility (Attali and Yogev 2021). Second, reduced access to health care and assisted reproductive technologies during the pandemic may have prevented couples from conceiving at all (Beaujouan 2021). Finally, restrictions on social activities may hinder dating and partnership formation, precluding some women from forming these partnerships during their fertile years (Guetto, Vignoli, and Bazzani 2021).

Prediction postmortem: Hits and misses

We believe predicting the future of US fertility, during a crisis, in real time and ex ante was an important step in assessing the power of digital demography, upon which many researchers have placed high hopes. To our knowledge, ours was the first attempt to do so and provides an important data point in evaluating the usefulness of such data in prediction. In this section, we evaluate how our model prediction performed ex post and discuss its successes and failures. More importantly, however, we try to understand not just where, but why our model missed.

In Figure 4, we represent our main predictions from Figure 2 (panel b), but now include the actual data for reference. Our data come from the Short-run Fertility Fluctuations data series from the Human Fertility Database. While some aspects of the actual and predicted series seem to correspond well, there are significant and important differences. Our prediction generally forecasts the timing and depth of the trough in births

FIGURE 4 US births by month: predicted versus actual

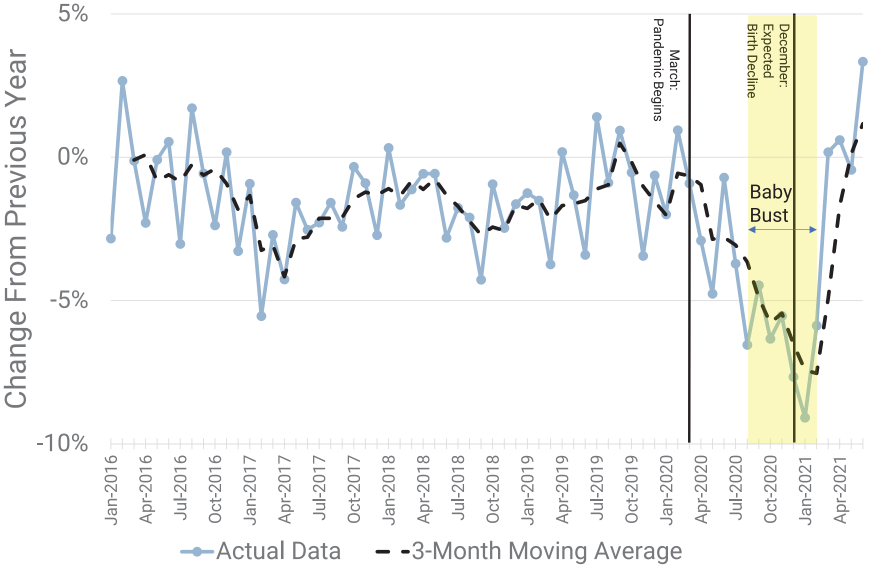


NOTES: National predicted vs. actual births, relative to average monthly births from 2015 to 2019 for that month, for two various prediction models. The three models are No COVID-19—a baseline model in which births follow normal seasonal patterns and remain on state-specific annual trends; and Google Data (MPSE)—which uses information on search volumes for Google keywords selected from a mean-squared prediction error minimizing forward stepwise machine learning selection method. Actual data comes from the Short Term Fertility Fluctuations data series within the Human Fertility Database

correctly. Specifically, while our model predicted a February trough, the actual nadir occurred in January—a difference of just one month. In contrast to our predicted 12.3 percent decline in births at the trough, the true 12-month decline at the trough was slightly less at 9.7 percent.

However, our prediction failed in two main ways. First, as shown in Figure 4, births began deviating significantly from our prediction—and from what was expected in the absence of COVID —almost immediately after the beginning of the pandemic. To show this more clearly, in Figure 5, we show 12-month changes in births from the Short-run Fertility Fluctuations data series for the United States between January 2016 and June 2021. Precisely at the onset of the pandemic in March 2020, a sharp change in trend emerged. According to a simple time-series Autoregressive Integrated Moving Average (ARIMA) model of births over this period, declines in May and August (and each month thereafter through 2020) were already well below the 95 percent confidence interval of expected births. Had these births been full term, they would have been conceived in the early-to-mid fall of 2019, well before the initial lockdown of March 2020. Not only were these declines significant in a statistical sense, but they were also historic in magnitude. The three-month moving average of the 12-month percent change in births has never fallen below −5 percent outside of a recession in the history of the NVSS natality data. However, this happened for the first time

FIGURE 5 12-month percentage change in births, 2016–2020



during the period August–October 2020, several months before births were expected to begin falling in December.

Most explanations for this sharp and historic birth collapse are either unlikely or impossible. One hypothesis states that conceptions may have already been experiencing historically large declines unrelated to the pandemic. However, this seems unlikely given the magnitude of birth reductions, which would have been unprecedentedly absent due to some sort of national crisis. It is also unlikely that births fell in anticipation of the pandemic: the first coronavirus case in China was not confirmed until late December 2019, and US Google searches for the coronavirus—an indicator of public awareness or interest—did not begin to rise until late January 2020. Another possibility is that preexisting conceptions experienced historically high termination rates as a result of the pandemic—either through abortion, miscarriage, or stillbirths. This is similarly unlikely. There is currently no evidence that the COVID-19 pandemic significantly changed miscarriage (Rotshenker-Olshinka et al. 2021) stillbirths (Stowe 2021), prematurity (CDC 2021), or abortion rates (Andersen, Bryan, and Slusky 2021). While the evidence in this area is still evolving, the required change in these variables also makes them unlikely as singular causes of the decline.[10]

Therefore, since our model only utilized time lags around conception (months 7–12 before birth), it effectively precluded the sharp changes in searches at the onset of the first wave to affect births earlier than November. This is an example of the difficulty of forecasting generally, as our modeling choice made theoretical sense ex ante, but not ex post, and should not necessarily be seen as an indictment of the use of Google data—or digital trace data more generally. However, in the case that the early, unexpected, and historic fall in births during the summer of 2020 was a result of a collapse in prepandemic conceptions, then digital trace data can be to blame—precisely because there was no corresponding collapse in conception-related searches in the summer of 2020.

Another possibly damning miss of our prediction is the faster-than-predicted birth rebound in the spring of 2021. While we predicted births to remain depressed through August 2021, Figure 4 shows actual births were no longer statistically different from our no-Covid prediction from as early as March 2021. Fortunately, this miss is also easy to understand ex post. Through our machine learning selection procedure, we demonstrated that historically, searches for the term "unemployment" were more predictive of future births than any other term (see Table 4 in the Supporting Information). Since searches for "unemployment" remained stubbornly high well into the summer of 2021, our model significantly underpredicted the speed of the rebound (see Figure 7 in the Supporting Information). Simply put, a few months after the pandemic began, there was a break from the historical relationship between searches for unemployment and future births.

Conclusions and lessons learned for digital demography

Digital trace data are both alluring and compelling for many applications in the population sciences. It promises a solution to data scarcity, both temporal and geographic in nature. Its sheer volume pairs well with atheoretical data science techniques, promising to uncover hidden relationships without needing to understand the underlying data-generating processes. In addition, it provides data collected passively, possibly reducing misreporting, such as by eliminating cognitive biases from self-reports or respondent priming—or in other words, "no one lies to Google" (Stephens-Davidowitz 2017).

However, such data may have many pitfalls which have only recently been understood (Cesare et al. 2018). In this paper, we outlined a rare test of digital data beyond its usual uses as just another alternative data source or in now-casting applications. Specifically, we used digital trace data to make ex ante future predictions regarding a question of primary interest to both researchers and the general public—to what extent the COVID-19 pandemic would cause a baby boom or baby bust.

Many of our findings and predictions confirmed the usefulness of digital trace data—and Google search data in particular—in demographic research. We demonstrated that periods of excess search volume for keywords relating to conception and pregnancy were indeed associated with higher numbers of births in the following months at the expected time lags, meaning that these data contained true signals regarding actual, fundamental demographic processes. We also showed that including this data significantly improved forecast accuracy, providing a powerful validation of their predictive capacity, and these predictions made sense as the state-level forecasts correlated in understandable ways with states' sociodemographic characteristics. Finally, our predictions were generally correct in terms of the timing and depth of the eventually realized birth declines, which was not the case using traditional data on unemployment alone.

However, we also showed that digital data are no panacea. They do not solve traditional modeling problems with forecasting, and in fact may introduce new ones. Specifically, one of the biggest "misses" of our model—the inability of our forecast to match the early decline of births in the summer of 2020—can likely be traced to modeling choices which were logical ex ante, but incorrect ex post. Nevertheless, if conceptions truly began falling before the onset of the pandemic, and digital data accurately measured missing data on intentions, this decline should have been anticipated. It was not. The other miss—the failure of the prediction to accurately forecast the quick rebound of births in the spring of 2021—is even more damning. The fundamental relationship between searches for unemployment and births broke down after just a few months of the pandemic. While the same can be said for the relationship between traditional data on actual unemployment, employing digital trace data was no better.

Acknowledgments

Joshua Wilde acknowledges funding from the UK Research and Innovation Grant number EP/Y031172/1. Joshua Wilde and Jasmin Abdel Ghany acknowledge support from the Leverhulme Trust (Grant RC-2018-003) for the Leverhulme Centre for Demographic Science. The authors declare that they have no competing interests. All data needed to evaluate the conclusions in the paper are publicly available and fully described in the paper and/or the Technical Appendix.

Notes

1 See Cesare et al. (2018) for an extensive review of this literature.

2 The October 2020 release used search data through July 2020 to predict to February 2021. A February update used search data though January 2021 to predict to August 2021.

3 All methods and results refer to the February 2021 update rather than the original October 2020 release since (1) the original can be found online (Wilde, Chen, and Lohmann 2020) and (2) the February 2021 update allows us to make predictions regarding the rebound as it extends the original prediction until August 2021.

4 This paper focuses on the results of the February 2021 update, which utilized search data though January 31, 2021 and predicted through August 2021. The original October 2020 prediction utilized data through July 31, 2020, and predicted through February 2021.

5 The states are Alaska, Delaware, Hawaii, Iowa, Idaho, Maine, Montana, North

Dakota, New Hampshire, Rhode Island, South Dakota, Vermont, West Virginia, and Wyoming.

6 There are 37 words, plus a control set and the unemployment rate, meaning there are 39 tested variables, each of which has 12 months of data. This makes 39*12 = 468 tested hypotheses.

7 Predicted changes in birthrates for the 14 missing states are set to be equal to the population-weighted average prediction of the remaining 36 states.

8 Covid caseloads only include cases before October 31, 2020, nine months before our final prediction.

9 While we do not provide a formal comparison to the realized birth declines due to the state of birth being suppressed in the US birth data, we note the Nobles et al. (2023) find a complex relationship between socioeconomic status (SES), race and ethnicity, and education using data from California. Specifically, they find that foreign-born individuals contributed the majority of the decline in births. Once foreign status is controlled for, differences in SES seem to play little role in predicting birth declines.

10 Miscarriage rates are 15-20 percent of established pregnancies, and abortions constituted 18.4 percent of established pregnancies which did not involuntarily terminate in 2017 (Jones, Witwer, and Jerman 2019), implying a 6.8 percent drop in August births would necessitate an unprecedentedly large 30–40 percent increase in either of these rates. Stillbirths only affect approximately 1 percent if all established pregnancies, making this channel the most implausible of all.

References

Aassve, Arnstein, Nico Cavalli, Letizia Mencarini, Samuel Plach, and Massimo Livi Bacci. 2020. "The COVID-19 Pandemic and Human Fertility." *Science* 369 (6502): 370–371. https://doi.org/10.1126/science.abc9520

Andersen, Martin, Sylvia Bryan, and David Slusky. 2021. "Covid-19 Surgical Abortion Restriction Did Not Reduce Visits to Abortion Clinics." IZA Discussion Paper No. 13832. Bonn, Germany: IZA Institute of Labor Economics. https://doi.org/10.2139/ssrn.3726434

Ashraf, Quamrul H., David N. Weil, and Joshua Wilde. 2013. "The Effect of Fertility Reduction on Economic Growth." *Population and Development Review* 39 (1):97–130. https://doi.org/10.1111/j.1728-4457.2013.00575.x

Attali, Emmanuel, and Yariv Yogev. 2021. "The Impact of Advanced Maternal Age on Pregnancy Outcome." *Best Practice Research Clinical Obstetrics Gynaecology* 70: 2–9. https://doi.org/10.1016/j.bpobgyn.2020.06.006

Beard, John R., and David E. Bloom. 2015. "Towards a Comprehensive Public Health Response to Population Ageing." *The Lancet* 385 (9968): 658–661. https://doi.org/10.1016/S0140-6736(14)61461-6

Beaujouan, Eva. 2021. "Covid-19 Global Demographic Research Needs? Replacing Speculative Commentaries with Robust Cross-national Comparisons." In *Covid-19 and the Global Demographic Research Agenda*, edited by Landis MacKellar and Rachel Friedman, 8–14. New York: Population Council.

Behrman, Julia Andrea, and Abigail Weitzman. 2016. "Effects of the 2010 Haiti Earthquake on Women's Reproductive Health." *Studies in Family Planning* 47 (1): 3–17. https://doi.org/10.1111/j.1728-4465.2016.00045.x

Bertillon, J.. 1892. La grippe a Paris et dans quelques autres villes de France et de l'étranger en 1889-1890. Imprimerie Municipale.

Billari, Francesco, Francesco D'Amuri, and Juri Marcucci. 2016. "Forecasting Births Using Google." In *Proceedings of the 1st International Conference on Advanced Research Methods and Analytics*. Valencia: Universitat Politecnica Valencia. pp 119. https://doi.org/10.4995/CARMA2016.2015.4301

BLS (Bureau of Labor Statistics). 2020. Unemployment Rate - [time series] lns14000003, lns14000006, lns14027662, lns14027660, and lns14027659. Washington, DC: Bureau of Labor Statistics. https://fred.stlouisfed.org/series/on, October 2, 2020.

Brodeur, Abel, Andrew E Clark, Sarah Fleche, and Nattavudh Powdthavee. 2021. "COVID-19, Lockdowns and Well-Being: Evidence from Google Trends." *Journal of Public Economics* 193: 104346. https://doi.org/10.1016/j.jpubeco.2020.104346

Buckles, Kasey, Daniel Hungerman, and Steven Lugauer. 2021. "Is Fertility a Leading Economic Indicator?" *The Economic Journal* 131 (634): 541–565. https://doi.org/10.1093/ej/ueaa068

Carneiro, Herman Anthony, and Eleftherios Mylonakis. 2009. "Google Trends: A Web-Based Tool for Real-Time Surveillance of Disease Outbreaks." *Clinical Infectious Diseases* 49 (10): 1557–1564. https://doi.org/10.1086/630200

CDC (Centers for Disease Control). 2020. "Covid-19 Hospitalization and Death by Race/Ethnicity, 2020." Atlanta, GA: CDC. Accessed October 2, 2020. https://stacks.cdc.gov/view/cdc/97937.

CDC (Centers for Disease Control and Prevention). 2021. *Provisional estimates for selected maternal and infant outcomes by month, 2018–2020*. 2021. Atlanta, GA: CDC.

Cesare, Nina, Hedwig Lee, Tyler McCormick, Emma Spiro, and Emilio Zagheni. 2018. "Promises and Pitfalls of Using Digital Traces for Demographic Research." *Demography* 55 (5): 1979–1999. https://doi.org/10.1007/s13524-018-0715-2

Chandra, Siddharth, Julia Christensen, Svenn-Erik Mamelund, and Nigel Paneth. 2018. "Short-Term Birth Sequelae of the 1918–1920 Influenza Pandemic in the United States: State-Level Analysis." *American Journal of Epidemiology* 187 (12): 2585–2595. https://doi.org/10.1093/aje/kwy153

Choi, Hyunyoung, and Hal Varian. 2012. "Predicting the present with Google Trends." *Economic Record* 88 (suppl.1): 2–9. https://doi.org/10.1111/j.1475-4932.2012.00809.x

Evans, Richard W., Yingyao Hu, and Zhong Zhao. 2009. "The Fertility Effect of Catastrophe: U.S. Hurricane Births." *Journal of Population Economics* 23 (1): 1–36. https://doi.org/10.1007/s00148-008-0219-2

Finch, W. Holmes, and Maria E. Hernandez Finch. 2020. "Poverty and Covid-19: Rates of Incidence and Deaths in the United States during the First 10 Weeks of the Pandemic." *Frontiers in Sociology* 5: 47. https://doi.org/10.3389/fsoc.2020.00047

Finlay, Jocelyn E. 2009. "Fertility Response To Natural Disasters: The Case Of Three High Mortality Earthquakes." World Bank Policy Research Working Papers. Washington, DC: World Bank. https://doi.org/10.1596/1813-9450-4883.

GBD 2017 Population and Fertility Collaborators. 2018. "Population and Fertility by Age and Sex for 195 Countries and Territories, 1950–2017: A Systematic Analysis for the Global Burden of Disease Study 2017". *The Lancet* 392 (10159): 1995–2051. https://doi.org/10.1016/S0140-6736(18)32278-5

Ginsberg, Jeremy, Matthew Mohebbi, Rajan Patel, Lynnette Brammer, Mark Smolinski, and Larry Brilliant. 2009. "Detecting Influenza Epidemics Using Search Engine Query Data." *Nature* 457 1012–1014. https://doi.org/10.1038/nature07634

Guetto, Raffaele, Daniele Vignoli, and Giacomo Bazzani. 2021. "Marriage and Cohabitation under Uncertainty: The Role of Narratives of the Future during the COVID-19 Pandemic." *European Societies* 23 (Suppl. 1): S674–S688, https://doi.org/10.1080/14616696.2020.1833359

Jones, Rachel K., Elizabeth Witwer, and Jenna Jerman. 2019. *Abortion Incidence and Service Availability in the United States, 2017*. New York: Guttmacher Institute. https://doi.org/10.1363/2019.30760

Jukic, Anne Marie, Donna D. Baird, Clarice R. Weinberg, D. Robert McConnaughey, and Allen J. Wilcox. 2013. "Length of Human Pregnancy and Contributors to Its Natural Variation." *Human Reproduction* 28 (10): 2848–2855. https://doi.org/10.1093/humrep/det297

Karra, Mahesh, David Canning, and Joshua Wilde. 2017. "The Effect of Fertility Decline on Economic Growth in Africa: A Macrosimulation Model." *Population and Development Review* 43: 237–263. https://doi.org/10.1111/padr.12009

Kearney, Melissa S., and Phillip B. Levine. 2020. "*Half a Million Fewer Children? The Coming COVID Baby Bust*.", Washington, DC: Brookings Institution https://www.brookings.edu/articles/half-a-million-fewer-children-the-coming-covid-baby-bust/.

Lazer, David, Ryan Kennedy, Gary King, and Alessandro Vespignani. 2014. "The Parable of Google Flu: Traps in Big Data Analysis." *Science* 343 (6176): 1203–1205. https://doi.org/10.1126/science.1248506

Lin, Allen Yilun, Justin Cranshaw, and Scott Counts. 2019. "Forecasting U.S. Domestic Migration Using Internet Search Queries." In *The World Wide Web Conference (WWW '19)*, 1061–1072. New York: Association for Computing Machinery. https://doi.org/10.1145/3308558.3313667

Linder, Forrest Edward, and Robert D. Grove. 1947. *Vital Statistics Rates in the United States, 1900–1940*. Federal Security Agency, U.S. Public Health Service,Washington DC: National Office of Vital Statistics.

Maestas, Nicole, Kathleen J. Mullen, and David Powell. 2023. "The Effect of Population Aging on Economic Growth, the Labor Force, and Productivity." *American Economic Journal: Macroeconomics* 15 (2): 306–332. https://doi.org/10.1257/mac.20190196

Markey, Patrick M., and Charlotte N. Markey. 2013. "Seasonal Variation in Internet Keyword Searches: A Proxy Assessment of Sex Mating Behaviors." *Archives of Sexual Behavior* 42 (4): 515–521. https://doi.org/10.1007/s10508-012-9996-5

Morgan, S. Philip, and Heather Rackin. 2010. "The Correspondence between Fertility Intentions and Behavior in the United States." *Population and Development Review* 36 (1): 91–118. https://doi.org/10.1111/j.1728-4457.2010.00319.x

Nobles, Jenna, Elizabeth Frankenberg, and Duncan Thomas. 2015. "The Effects of Mortality on Fertility: Population Dynamics After a Natural Disaster." *Demography* 52 (1): 15–38. https://doi.org/10.1007/s13524-014-0362-1

Nobles, Jenna, Alison Gemmill, Sungsik Hwang, and Florencia Torche. 2023. "Fertility in a Pandemic: Evidence from California." *Population and Development Review* padr.12591. https://doi.org/10.1111/padr.12591

Ojala, Jussi, Emilio Zagheni, Francesco Billari, and Ingmar Weber. 2017. "Fertility and Its Meaning: Evidence from Search Behavior." *Proceedings of the International AAAI Conference on Web and Social Media* 11 (1): 640–643. https://doi.org/10.1609/icwsm.v11i1.14915

Palloni, Alberto. 1988. "On the Role of Crises in Historical Perspective: An Exchange: Comment." *Population and Development Review* 14 (1): 145–158. https://doi.org/10.2307/1972503

PRB (Population Reference Bureau). 2020. World Population Data Sheet, 2020. Washington, DC: Population Reference Bureau. https://www.prb.org/wp-content/uploads/2020/07/letter-booklet-2020-world-population.pdf

Preis, Tobias, Helen Susannah Moat, and H. Eugene Stanley. 2013. "Quantifying Trading Behavior in Financial Markets Using Google Trends." *Scientific Reports* 3 (1): 1684. https://doi.org/10.1038/srep01684

Ricketts, Comfort F., and Carlos G. Silva. 2017. "An Analysis of Morbidity and Mortality Using Google Trends." *Journal of Human Behavior in the Social Environment* 27 (6): 559–570. https://doi.org/10.1080/10911359.2017.1299658.

Rotshenker-Olshinka, Keren, Alexander Volodarsky-Perel, Naama Steiner, Eryn Rubenfeld, and Michael H. Dahah. 2021. "Covid-19 Pandemic Effect on Early Pregnancy: Are Miscarriage Rates Altered, in Asymptomatic Women?" *Archives of Gynecology and Obstetrics* 303: 839–845. https://doi.org/10.1007/s00404-020-05848-0

Siliverstovs, Boriss, and Daniel S. Wochner. 2018. "Google Trends and Reality: Do the Proportions Match?" *Journal of Economic Behavior & Organization* 145: 1–23. https://doi.org/10.1016/j.jebo.2017.10.011

Sobotka, Tomáš, Vegard Skirbekk, and Dimiter Philipov. 2011. "Economic Recession and Fertility in the Developed World." *Population and Development Review* 37 (2): 267–306. https://doi.org/10.1111/j.1728-4457.2011.00411.x

Solano, Paola, Morena Ustulin, Enrico Pizzorno, Monica Vichi, Maurizio Pompili, Gianluca Serafini, and Mario Amore. 2016. "A Google-Based Approach for Monitoring Suicide Risk." *Psychiatry Research* 246: 581–586. https://doi.org/10.1016/j.psychres.2016.10.030

Stephens-Davidowitz, Seth. 2017. *Everybody Lies: Big Data, New Data, and What the Internet Can Tell Us About Who We Really Are*. New York: Harper-Collins.

Stowe, Julia, Helen Smith, Kate Thurland, Mary Ramsey, Nick Andrews, and Shamez Ladhani. 2021. "Stillbirths during the Covid-19 Pandemic in England, April-June 2020." *JAMA* 325 (1): 86–87. https://doi.org/10.1001/jama.2020.21369

Tamgno, James K., Roger M. Faye, and Claude Lishou. 2013. "Verbal Autopsies, Mobile Data Collection for Monitoring and Warning Causes of Deaths." In 15th International Conference on Advanced Communications Technology (ICACT), PyeongChang, Korea (South), 495–501.

Tan, Jacinth J. X., Michael W. Kraus, Nichelle C. Carpenter, and Nancy E. Adler. 2020. "The Association between Objective and Subjective Socioeconomic Status and Subjective Wellbeing: A Meta-Analytic Review." *Psychological Bulletin* 146 (11): 970–1020. https://doi.org/10.1037/bul0000258

UNFPA (United Nations Population Fund). 2020. *Impact of the COVID-19 Pandemic on Family Planning and Ending Gender-based Violence, Female Genital Mutilation and Child Marriage*. Technical report. New York: UNFPA. https://www.unfpa.org/sites/default/files/resource-pdf/COVID-19_impact_brief_for_UNFPA_24_April_2020_1.pdf

WHO (World Health Organization). 2021. WHO Coronavirus Disease (COVID-19) Dashboard. Accessed 29 April, 2021. https://covid19.who.int/.

Wilde, Joshua, Bénédicte H. Apouey, and Toni Jung. 2017. "The effect of ambient temperature shocks during conception and early pregnancy on later life outcomes." *European Economic Review* 97: 87–107. https://doi.org/10.1016/j.euroecorev.2017.05.003

Wilde, Joshua, Wei Chen, and Sophie Lohmann. 2020. "Covid-19 and the future of US fertility: What can we learn from Google?" IZA Discussion Paper No. 13776. Bonn, Germany: IZA Institute of Labor Economics. https://docs.iza.org/dp13776.pdf

Wladyka, David. K. 2017. "Queries to Google Search as Predictors of Migration Flows from Latin America to Spain." *Journal of Population and Social Studies* 25 (4): 312–327. https://so03.tci-thaijo.org/index.php/jpss/article/view/102268

Zagheni, Emilio, and Ingmar Weber. 2012. "You are where you e-mail: using e-mail data to estimate international migration rates." In *Proceedings of the 4th Annual ACM Web Science Conference (WebSci '12)*, 348–351. New York: Association for Computing Machinery. https://doi.org/10.1145/2380718.2380764