Quantifying Community Characteristics of Maternal Mortality Using Social Media

Rediet Abebe*
rabebe@fas.harvard.edu
Harvard University

Salvatore Giorgi*
sgiorgi@seas.upenn.edu
University of Pennsylvania

Anna Tedijanto
ajt232@cornell.edu
Cornell University

Anneke Buffone
buffone.anneke@gmail.com
University of Pennsylvania

H. Andrew Schwartz
has@cs.stonybrook.edu
Stony Brook University

ABSTRACT

While most mortality rates have decreased in the US, maternal mortality has increased and is among the highest of any OECD nation. Extensive public health research is ongoing to better understand the characteristics of communities with relatively high or low rates. In this work, we explore the role that social media language can play in providing insights into such community characteristics. Analyzing pregnancy-related tweets generated in US counties, we reveal a diverse set of latent topics including *Morning Sickness*, *Celebrity Pregnancies*, and *Abortion Rights*. We find that rates of mentioning these topics on Twitter predicts maternal mortality rates with higher accuracy than standard socioeconomic and risk variables such as income, race, and access to health-care, holding even after reducing the analysis to six topics chosen for their interpretability and connections to known risk factors. We then investigate psychological dimensions of community language, finding the use of less trustful, more stressed, and more negative affective language is significantly associated with higher mortality rates, while *trust* and *negative affect* also explain a significant portion of racial disparities in maternal mortality. We discuss the potential for these insights to inform actionable health interventions at the community-level.

KEYWORDS

maternal mortality, health disparities, language, topic modeling, community characteristics

1 INTRODUCTION

The United States has one of the highest maternal mortality rates of any country in the Organization for Economic Cooperation and

---

*Both authors contributed equally to this research.

Development group [11, 23]. Approximately 700 individuals die from pregnancy-related causes [3, 15, 68] and an estimated 60% of these deaths are suspected to be preventable [68]. While the international trend has seen a reduction in maternal mortality, despite increased budgets, rates in the US have more than doubled in the past 25 years [3].[1] Black and Latina mothers bear a disproportionate brunt of this burden: Black women are three to four times more likely to die during childbirth, even after controlling for numerous socioeconomic and risk factors [68]. These rates vary by geography: e.g., in New York City, Black women are 12 times more likely to die during childbirth than white women [64, 73].[2]

Public health research has examined potential causes for maternal mortality and disparities, pointing to issues such as access to insurance, bias in health-care, segregated hospitals, and inadequate post-delivery care [8, 27, 49, 50, 54]. While it is understood that each of community, health facility and system, patient, and provider all play a part, there is an overall pervasive concern that the specific causes and mechanisms for maternal mortality and disparities are not adequately understood [68]. The WHO cites a "general lack of good data – and related analysis – on maternal health outcomes" as a bottleneck for gaining insights into this issue [3].

In this work, we seek to partially address this gap, focusing on community-level factors that characterize maternal mortality as revealed through social media language. We examine whether community variables derived from social media language data can predict community maternal mortality rates and its racial disparity. While emotions and language analyzed using social media data are shown to have high-efficacy in tasks ranging from predicting allergies or life satisfaction to depression or heart disease mortality [26, 30, 33, 66, 76], the potential of social media has yet to be examined in this manner to help shed understanding on maternal mortality at the community level.

Our contributions in this work are in three-folds:

- We show that there is a diverse set of pregnancy-related topics ranging from *Morning Sickness*, to *Abortion Rights*, to *Maternal Studies*. We demonstrate that these topics predict maternal mortality rates with higher accuracy than standard socioeconomic (SES), risk factors, and race.

---

[1]Note, on the other hand, US infant mortality is at a historic low [13].

[2]This issue has garnered increases attention in part due to concentrated efforts by policy-makers, advocacy groups, and celebrities, in addition to long-standing work by community organizations [35, 36, 40, 59, 81]. e.g., see collaborations between the Atlanta-based Black Mamas Matter Alliance and the Black Maternal Health Caucus.

- We show that a select set of six topics, chosen for their interpretability and relations to known maternal health factors, hold as much predictive power as all pregnancy-related topics. Specifically, four of these topics – *Maternal Studies, Teen Pregnancy, Abortion Rights*, and *Congratulatory Remarks* – have negative associations with mortality rates.
- We examine variables associated with racial disparities in maternal mortality (i.e. the difference between rates for Black women and other races), finding that language-based scores for *trust* and *affect* hold explanatory power for the county-level relationship between race and maternal mortality, even after controlling for standard SES and risk-factors.

2 BACKGROUND AND RELATED WORK

Maternal Mortality Background. Public health research has sought better measurements of maternal mortality rates and their causes and consequences [3, 23, 68]. There is a long line of work exploring what community, patient, hospital, provider, or systemic-level factors may contribute to high rates of mortality and disparities in the US [34, 48, 56, 57]. At the patient-level, cardiovascular conditions, which are related to stress, cause about one third of all pregnancy-related deaths [68]. At the community and systemic-level, studies have shown that delivery site, segregation, and discrimination in maternity care during visits all play a role [8, 27, 49, 50]. At the systemic-level, sociological and economic research have shown racial disparities in mortality and life-expectancy [18, 55]. In line with such studies, there are numerous calls to use a data-driven approach to better grasp the role and causes of maternal mortality related to each of the above main categories [68].

Social Media Data for Health. Twitter data and more generally social media data has been a popular source for exploring community-level health measurements [67]. Examples include excessive alcohol consumption [26], depression [29, 62], heart disease, [33], and more generally population health and well-being [25, 38, 76]. In addition to measuring community-level insights, these data sources have been used to study health information seeking and sharing [31] and individual-level predictions [30]. In recent years, there has also been interest in understanding the societal and ethical implications and limitations around the use of social media data for health studies and roles for computing as a diagnostic of social problems [1, 4, 16, 17, 20].

Maternal Health. An emerging topic of interest has been the use of language-driven analysis to understand pregnancy and maternal experiences. For instance, De Choudhury et al. [28] studied Twitter posts to understand changes in emotions for mothers; Antoniak et al. [5] looked at narrative paths in individuals sharing childbirth stories on an online forum. Focusing on support, Costa Figueiredo et al. [22], Gui et al. [42], Vydiswaran et al. [80] looked at how online peer support and information exchange for pregnant individuals, their caregivers, and individuals experiencing fertility issues. Abebe et al. [2] looked at information seeking for pregnancy and breastfeeding related to HIV. To our knowledge, ours is the first work to employ a language-driven study to understand maternal mortality in the US.

3 DATA

We used three sets of data sets for this study, described below:

3.1 Twitter Data and Seed-Words

To generate our pregnancy data set, we started with a random 10% sample of the entire Twitter stream collected between 2009 and 2015 [70]. We then used this data set to build two subsets: (1) pregnancy-related tweets and (2) tweets geo-located to US counties.

Pregnancy-Related Tweets. The first data set consisted of tweets related to pregnancy and birth. Tweets were pulled from the main data set if they contained the following seed-words: *pregnancy, pregnant, infant, fetus, miscarriage, prenatal, trimester, complications, pregnant, birth, childbirth, pregnancies, baby, children, pregnancy, mother, newborn, child,* as well as their plural form, hashtags such as *#pregnancy*, and capitalizations such as *Pregnancy*. These seed-words were selected by examining nearest neighbors from word2vec for words related to 'pregnancy' and 'pregnant.'

We then manually examined a random sample of 1,000 tweets from the data set to test for relevance to pregnancy. Tweets that were deemed off-topic, such as those containing phrases like "miscarriage of justice" were used to generate phrases for further data cleaning. We also randomly sampled tweets for specific seed-words and if a substantial (i.e., more than 20%) of the tweets were unrelated to pregnancy, all tweets were removed from the data set, reducing the seed-set. After these cleaning steps, we kept 74.40% of the data set, and validated in fresh sample of 1,000 tweets that over 95% of them are related to pregnancy.

U.S. County Tweets. The second data set consisted of tweets geo-located to U.S. counties. For this we used the County Tweet Lexical Bank [39]. This data set was geo-located using self-reported location information (from the user description field) and latitude / longitude coordinates [76]. The data were then filtered to contain only English tweets [58]. We then limited our data set to Twitter users with at least 30 posts and U.S. counties with at least 100 such users. The final Twitter data set consisted of 2,041 U.S. counties.

3.2 Mortality Rates

The World Health Organization (WHO) defines maternal mortality as "the death of a woman while pregnant or within 42 days of termination of pregnancy, irrespective of the duration and site of the pregnancy, from any cause related to or aggravated by the pregnancy or its management but not from accidental or incidental causes" with the Centers for Disease Control and Prevention (CDC) expanding this time period to 1 year [24, 82]. Data for maternal mortality was collected from the CDC WONDER online database [14]. We collected rates from 2009-2017, so as to match the time-span of our Twitter sample in addition to more recent years (2016 and 2017) since these rates are on the rise [68]. Mortality rates are listed under the following International Classification of Diseases, Tenth Revision (ICD-10) categories: O00-O07 (pregnancy with abortive outcome) and O10-O99 (other complications of pregnancy, childbirth and the puerperium). The CDC suppresses data if a county experiences less than 10 deaths in a given time period for privacy reasons. Of the 2,041 counties in our Twitter set only 197 also had mortality rates (i.e., counties experiencing 10 or more deaths).

Since the CDC does not report age-adjusted rates for counties with low mortality numbers, we took the crude rate as reported and created our own age-adjusted rate. To do this, we built a model using median age of females (American Community Survey, 2014; 5-year estimates) and predicted maternal mortality, taking the residuals as our new "age-adjusted maternal mortality rate." This age-adjusted value is used throughout the paper.

3.3 Socioeconomic Measures and Risk Factors

In addition to mortality, we collected additional county-level variables on *socioeconomics*, *risk factors*, and *race*. *Socioeconomics* included unemployment rate, median income, and education (percentage of people with Bachelor's degrees and High School graduate percentage). For *risk factors*, we included insurance rates and access to health-care (the ratio of population to number primary care providers). Finally, we also explored the relationship between language and maternal mortality with respect to percentage of Black individuals in each county. As discussed previously, the disparity in mortality rates for Black women is large and providing evidence toward the factors at play for such a disparity is a key application for our analyses. Additionally, to account for overall rates of birth, all analysis included a birth rate covariate (the rate per 1,000 women, aged 15-50, with births in the past 12 months).

The birth rate, race, SES variables, and insurance rates were collected from the 2014 American Community Survey (5 year estimates), whereas the primary care providers was collected from the 2017 County Health Rankings (as reported by the Area Health Resource File/American Medical Association, 2014). We were able to obtain these values for each of the counties which met the Twitter and mortality inclusion criteria above.

Overall, we obtained data for 197 U.S. counties and county equivalents that met each of the data requirements above and conducted our study on these counties. The full list of these counties is included in the project page.[5]

4 TOPICS AND THEORETICAL LINGUISTIC FEATURES

We used three sets of features that will characterize maternal mortality through language. First, we created a set of automatically-derived topics built over the pregnancy-related tweets. These topics reveal a diversity of themes in discussions around pregnancy on the platform. Next, we used a small set of theoretically-driven language features – (*affect*, *depression*, *stress*, and *trust*) – in order to access psychological traits of a community and their relations to maternal mortality. Finally, we use a large, general set of topics (non-pregnancy related) to identify broader language patterns.

4.1 Pregnancy-Related Topics

We start with our data set of over 5 million pregnancy-related tweets described in Section 3. We automatically extracted *topics* using Latent Dirichlet Allocation (LDA) [12]. LDA is a generative statistical model which assumes that each document (in our case tweet) contains a distribution of topics, which in turn, are a distribution of words. We use the Mallet software package [60], which estimates the latent variable of the topics using Gibbs sampling [37]. All default Mallet settings were used, except α, which is a prior on

the expected topics per document. We set $\alpha = 2$ since tweets are shorter than the typical length of documents. The number of topics is a free parameter and we chose 50 topics.[3]

Topic Label	Top Weighted Words
Teen Pregnancy (1.34%)	teen, rate, rates, teenage, highest, mortality, low, states, teens, higher, number, 20, country, american, united, education, lowest, population
Morning Sickness (0.54%)	morning, sickness, purpose, symptoms, lives, wanted, williamson, tv, experience, cure, bra, marianne, thinking, signs, oral, teenagers, simon
Celebrity Pregnancies (1.42%)	kim, kardashian, kayne, amber, rose, beyonce, years, west, harry, finish, swear, north, who's, kayne's, taylor, sets, louis, wiz
Abortion Rights (2.15%)	women, abortion, care, health, abortions, bill, mortality, #prolife, rights, law, gift, support, circumstances, crisis, irrelevant, #prochoice, forced
Maternal Studies (2.56%)	risk, defects, study, health, weight, linked, flu, cancer, early, diet, drinking, smoking, blood, safe, alcohol, diabetes, autism, acid, disease, drug
Congratulatory Remarks (3.06%)	congrats, congratulations, :), boy, happy, love, daughter, son, <3, wait, sister, late, healthy, cousin, xx, amazing, :d, meet, proud

Table 1: Sample pregnancy topics with representative words

We find that our data reveals a rich set of themes related to pregnancy and birth. In Table 5, we show a sample of six topics, which are hand-selected to demonstrate the breadth of topics in the data set.[5] The first column provides the topic label, which were hand-generated by the authors, and the frequency with which the topic occurs in the data set.[4] The last column corresponds to the top 10 most representative words for the topic.

These above topics show that pregnancy-related discussions on Twitter can range from personal-health disclosure such as in *Morning Sickness*, to political conversations related to *Abortion Rights*, and light topics such as *Congratulatory Remarks*. Topics that were not included in manuscript due to length constraints include *Royal Baby*, *Food Cravings*, and *Pregnancy Timeline*. Each of these topics shows varying levels of popularity across the counties.

4.2 Theoretical Features

We also explore a set of theoretically-driven language features: *affect*, *depression*, *trust*, and *stress*. We downloaded pre-existing models to derive county-level language features including:

- **affect** – positive and negative emotional valence trained over Facebook posts [71].
- **depression** – degree of depressive personality (a facet of the big five personality test) fit over social media users' language [75].

---

[3]Before running the rest of our analysis, we ran LDA using 10, 20, 50, 100, and 200 topics. We selected 50 topics based on manual inspection of coherence and interpretability of the topics.

[4]Note, since there are 50 topics, the average value is 2%. Furthermore, since some themes, such as celebrity pregnancy, occur in more than one topic, the overall frequency of this theme in the data set is higher than the corresponding value in this table.

- **trust** – degree of trustfulness (how much one tends to trust persons or entities that they do not personally know) fit over social media users' language [83].
- **stress** – amount of stress fit over social media users' language and Cohen's Stress scale [19, 43].

4.3 General Topics

Finally, we use a larger set of LDA topics built over a more general data set. By doing this in tandem with the pregnancy-related topics, we can zoom in on pregnancy-related themes while also exploring a larger set of language correlates, which might help in characterizing communities suffering from higher or lower rates of mortality. To this end, we downloaded a set of 2,000 topic posteriors that were automatically-derived over the MyPersonality data set [77]. These topics have been used over a large class of problems and have been found to be robust both in terms of interpretability and predictive power [33, 51, 65, 69], so they form a point of comparison for our domain-specific topics.

5 METHODS

To understand the relationship between community level language and maternal mortality, we perform three types of statistical analyses: (1) *prediction* — can language be used to predict mortality rates in a cross-sectional cross validation setup? (2) *differential language analysis* – can we gain insights into communities which suffer from higher or lower maternal mortality through language? and (3) *mediating language analysis* — can language be used to understand the mechanisms through which Black communities experience increased rates of maternal mortality? All data processing, feature extraction and statistical analysis are performed using the open source Python package DLATK [78].

5.1 Prediction

We use two types of predictive models, depending on the type of independent variables. All non-language variables (i.e., SES and risk factors) are modeled with an ordinary least squares (OLS) regression, whereas language features use an ℓ_2 regularized (Ridge) regression [47]. In addition to regularization, we also use a feature selection pipeline in all language based models, since the number of features can be larger than the number of observations (N=197 counties). The pipeline first removes all low variance features and then features that were not correlated with our outcome. Finally, we applied Principal Component Analysis (PCA) to further reduce the number of features. All models are evaluated in a 10-fold cross validation setup, with the Ridge regularization parameter α tuned on the training set within each fold. Predictive accuracy is measured in terms of a single Pearson correlation between the actual values and the predicted values, whereas standard errors are calculated across all 10 folds.

5.2 Differential Language Analysis

Differential Language Analysis (DLA) is used to identify language characterizing maternal mortality [52, 77]. Here we individually regress each of our language variables (i.e., pregnancy related topics and theoretical features) using an OLS regression, adding in access to health-care, birth rates, socioeconomics and risk factors

as covariates. We adjust for multiple comparisons by applying a Benjamini—Hochberg false discovery rate correction to the significance threshold ($p < .05$) [10]. For LDA topics we visualize topics significant correlations as word clouds. The word clouds display the top 15 most prevalent words within a topic sized according to their posterior likelihood.

5.3 Mediating Language Analysis

We explore the relationship between maternal mortality and the percentage of Black individuals within a county, as expressed through the county's language. Language based mediation analysis has been used in the past to explore the relationship between socioeconomics and excessive drinking [26]. For this analysis, we residualize the crude maternal mortality rate, as reported by the CDC, on median age of female, birth rates, all socioeconomic variables (income, education and unemployment), insurance rates and rates of primary care providers.

For each language variable, both the pregnancy related LDA topics and theoretical language features, we consider the mediating relationship between the topic (mediator), percentage Black (independent variable) and residualized maternal mortality rates (dependent variable). We follow the standard three-step, Baron and Kenny approach [9]. Step 1: we regress our independent (x) and dependent variables (y; path c) in a standard OLS regression. Step 2: we regress the independent variable (x) and mediator (m; path α). Finally, in Step 3 we create a multi-variate model and regress both the mediator (m; topic) and independent variable (x; percentage Black) with maternal mortality (y; path c'). The three models are as follows:

$$y = cx + \beta_1 + \epsilon_1, \tag{1}$$
$$m = \alpha x + \beta_2 + \epsilon_2, \tag{2}$$
$$y = c'x + +\beta m + \beta_3 + \epsilon_3. \tag{3}$$

The mediation effect size ($c - c'$) is taken as the reduction in the effect size between the direct relationship (i.e., percentage Black and maternal mortality) and the mediated relationship. To test for significance, we use a Sobel p [79] and correct all p values for false discoveries via a Benjamini—Hochberg procedure.

6 RESULTS

We begin by looking at correlations between maternal mortality and various socioeconomics and risk factors. Table 2 shows the set of correlation coefficients. These results state that the percentage of the population that is Black and unemployment rate were positively correlated with maternal mortality rate and insurance access, income, and education were negatively correlated with maternal mortality rate. Additionally, birth rates were not significantly correlated with maternal mortality. Note that, in this paper, we only consider 197 counties in the US due to constraints around Twitter and county-mapped data as discussed in Section 3. While the correlation values do not exactly match correlations for all US counties, the general direction of relationship between maternal mortality rates and these SES and risk-factors was the same, with those the strongest associations – such as percent Black – also matching.

We next look at the predictive accuracy of our 50 topics, the 2000 general topics, and the above SES and risk-factors as well as percent Black values. For this, note that we used linear regression

	Correlation
Birth Rates	
Rate per 1,000 women	.10 [-.04,.24]
Race	
Black (percent)	.49 [.36,.61]***
Risk Factors	
Primary Care Providers	-.23 [-.38,-.09]**
Uninsured (percent)	.27 [.12,.41]***
Socioeconomics	
Income (log median)	-.42 [-.55,-.29]***
High School or more (percent)	-.14 [-.28,.01]
Bachelor's Degree (percent)	-.38 [-.52,-.23]***
Unemployment (percent)	.26 [.12,.39]***

Table 2: Correlations with risk factors, socioeconomics, and race. All non-birth rate correlations controlled for birth rates. Reported standardized β with 95% confidence intervals in square brackets; ***$p < 0.001$, **$p < 0.01$, *$p < .05$, after Benjamini—Hochberg correction
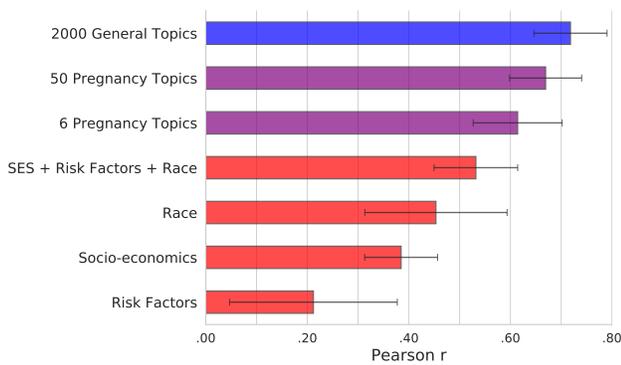

Figure 1: Prediction accuracy for non-language variables (red), pregnancy related LDA topics (purple) and general set of LDA topics (blue). Reported Pearson r from 10-fold cross validation, errors bars are 95% CI.

with maternal mortality values as the outcome variable and the aforementioned language variables as the explanatory variables. Figure 1 shows that the 2000 general Facebook topics had the highest predictive power with a Pearson $r = .72$ [.65, .79]*** while risk factors (PCP access and insurance rate) were the lowest with a Pearson $r = .21$ [.05, .38]**. Overall SES factors, risk factors, and race, had significantly less predictive accuracy (using a paired t-test) than the 50 pregnancy-related topics from the Twitter data ($t = -4.63$, $p < .001$) and the 2000 general topics ($t = -4.74$, $p < .001$).

For the Differential Language Analysis (DLA), we selected the 6 topics of interest. We ran a multi-linear regression, treating the maternal mortality rate as an outcome variable and the prevalence of these topics in the counties as the explanatory variable with birth rates, race, risk factors and socioeconomics as covariates. We found that five of the 6 topics, shown in Figure 2 had significant associations with maternal mortality rates. *Maternal studies* had

the most negative association – i.e., counties where there are relatively more tweets related to this topic had lower rates of mortality. Note that each of the four topics in the figure – *Maternal Studies, Teen Pregnancies, Congratulatory Remarks*, and *Abortion Rights* – all show negative associations with maternal mortality rates. *Celebrity Pregnancies*, not shown, is positively associated (.20 [.07, .33]*) with higher mortality.

-.38 [-.49,-.25]***   -.35 [-.47,-.22]***

-.28 [-.40,-.14]***   -.19 [-.32,-.05]*
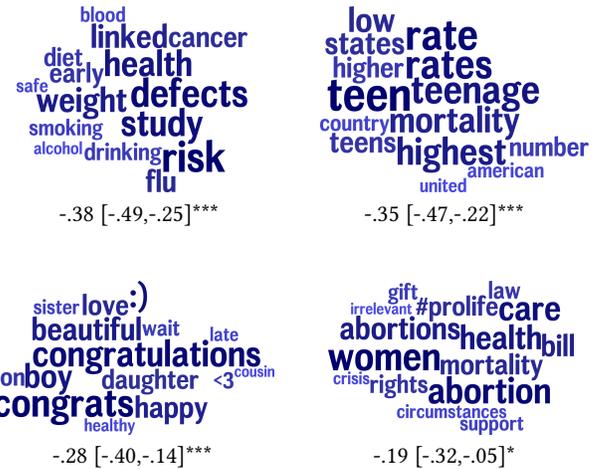
Figure 2: Differential Language Analysis using 6 pregnancy related LDA topics, controlled for race, risk factors and socioeconomics. Reported standardized β with 95% confidence intervals in square brackets; ***$p < 0.001$, **$p < 0.01$, *$p < .05$, after Benjamini—Hochberg correction.

We also used 4 theoretical features within the DLA framework: *affect, depression, stress* and *trust*. Results are presented in Table 3. We see higher rates of maternal mortality associated with higher distrust, higher stress, higher depression, and with less affect.

	Correlation
Affect	-.30 [-.43,-.17]***
Depression	.23 [.10,.36]**
Stress	.24 [.10,.37]**
Trust	-.38 [-.49,-.25]***

Table 3: Differential Language Analysis of theoretically relevant features. Reported standardized β with 95% confidence intervals in square brackets; ***$p < 0.001$, **$p < 0.01$, *$p < .05$, after Benjamini—Hochberg correction.

Finally, we explore disparities by race at the population level. The county-level health disparity itself can be seen simply from the strong correlation between the two variables: communities that are more Black, have greater maternal mortality. We turn to Twitter-based community characteristics as mediators (i.e. explainers) of this race-mortality relationship. The idea behind mediation analysis, is that if included a 3rd variable (i.e. a Twitter measurement) in the linear analysis reduces the relationship of the first 2 (i.e. race and

	$c - c'$	α	β
Affect	.11**	-.40 [-.53,-.27]***	-.27 [-.41,-.13]***
Depression	-.04	-.26 [-.39,-.12]***	.14 [.01,.28]*
Stress	-.01	-.06 [-.20,.08]	.15 [-.02,.28]*
Trust	.14**	-.51 [-.63,-.39]***	-.27 [-.42,-.12]***

Table 4: Mediating Language Analysis: Analysis seeks to explain the correlation, $c = .36$, between *percent Black* and *residualized maternal mortality* through differences in language. α: correlation between the theoretical factor and percent Black; $beta$: correlation between the theoretical factor and residualized maternal mortality. Reported Pearson r with 95% confidence intervals in square brackets; ***$p <$ 0.001, **$p < 0.01$, *$p < .05$, after Benjamini- –Hochberg correction. The $c - c'$ column uses a Sobel p for significance [79].

maternal mortality), then this third variable is accounting for some of the covariance between the first two.

We considered each of the 4 theoretical dimensions as potential mediators. To zero in on explaining what is novel about the race-mortality, we controlled for all previously mentioned socioeconomic and risk factor variables by producing a residual of the variance left over. The correlation between percent Black and maternal mortality was then $c = .36$ Without this step, it could be that any mediators were simply accounting for socioeconomic or risk factor effects. As seen in Table 4, we found two of the theoretical dimensions and 3 of the topics had a significant mediation effect, in part explaining the disparity. For example, trust mediated the relationship – the fact that communities expressing lower trust had greater maternal mortality, partially explained why Black percentage related to greater mortality.

7 DISCUSSION

The results shown in this work demonstrate the efficacy of social media language to shed some light on community characteristics of maternal mortality. While social media data, by itself, is not able to reliably identify causes for high maternal mortality rates and disparities, it can provide supporting evidence for existing conjectures and generate hypotheses for further investigation.

The observation that pregnancy-related topics, as well as the general 2,000 topics, both hold more predictive power than SES, risk factors, and race, combined, shows that such language-based data sets may contain characteristics of communities beyond that captured in standard variables used to study maternal mortality. Furthermore, the diversity of discussion themes in the pregnancy-related data set presents an opportunity to consider how different topics relate with maternal mortality rates and patterns of topic popularity across US counties.

The novel mediation results presented in this work allow us to gain further insights into how *affect, depression, stress,* and *trust* relate to mortality rates and disparities. The results that *trust* and *affect* related significantly with mortality rates mirrors discussions from public health research: for instance, failure by hospitals, providers, and facilities to provide unbiased and nondiscriminatory care has already been shown to result in lower follow-up visits by

Black and Latina women, which is believed to drive higher mortality rates. Trust in physicians and medical institutions has been extensively studied [44–46, 61], with multiple studies focusing on racial and ethnic differences in levels of trust [6, 7, 32, 41]. Findings repeatedly show ethnic and racial differences in trust towards health-care systems, in addition to showing that distrust is associated with racial disparities in use of preventive services [63]. The *affect* result is also related to the *Congratulatory Remarks* topic, indicating that communities with both more positive language and more positive discussions around pregnancy and birth may also be experiencing lower maternal mortality rates and disparities. These observations, along with existing discussions, provide potential actionable insights for policies at the community level.

The results here are not without limitations: as with other studies heavily relying on social media data, there are inherent issues of selection bias in who is on the platform and which users meet the inclusion thresholds we set for the pregnancy-related and county-mapping data sets. There is also selection bias in tweets that are geo-located as well as language use by the individuals on Twitter compared to other platforms. It is imperative to not take these data sets as being representative of the U.S., the counties we study, or even individuals that maybe included in the data sets.

Furthermore, we do not control for linguistic differences across different parts of the U.S. and some topics, as a result, may show significant spatial and geographic associations. Likewise, we set the seed-words for constructing the pregnancy-related data set using word2vec, which may also suffer bias issues: e.g, certain words which may be commonly used to discuss pregnancy and birth by certain groups of under-represented individuals may not pass this analysis. While we attempt to control for this by having a relatively large number of seed-words and instead relying on data cleaning, this remains a notable limitation.

We were hindered by the availability of outcome data: a lot of the relevant data is available only at the county-level and crucial data like disparities by race were entirely unavailable. While we believe that studies like ours will provide additional data-sources, models, and measurements to further our understanding of maternal mortality and disparities, availability of ground truth data presents a significant bottleneck. The availability of ground truth data about mortality and disparities, including data regarding mortality rates for groups of individuals belonging to marginalized communities, as well as disaggregated data by different demographics such as race, age, education, income, and other demographics would allow for more fine-grained analysis.

8 ETHICS STATEMENT

This study was reviewed by the University of Pennsylvania institutional review board and, due to lack of individual human subjects, found to be exempt. All data used in this study are publicly available. While the county-level language estimates are publicly available and will be posted on the project page[5], the original tweets, which are also publicly available, are unable to be redistributed by the authors due to Twitter's Terms of Service. For additional privacy protection, we automatically replace any Twitter user names with <user> in our analysis and presentation in this paper.

---

[5]All data available at: https://github.com/wwbp/maternal_mortality

REFERENCES

[1] Rediet Abebe, Solon Barocas, Jon Kleinberg, Karen Levy, Manish Raghavan, and David G Robinson. 2020. Roles for computing in social change. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. 252–260.

[2] Rediet Abebe, Shawndra Hill, Jennifer Wortman Vaughan, Peter M Small, and H Andrew Schwartz. 2019. Using search queries to understand health information needs in africa. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 13. 3–14.

[3] Priya Agrawal. 2015. Maternal mortality and morbidity in the United States of America.

[4] Tim Althoff. 2017. Population-scale pervasive health. *IEEE pervasive computing* 16, 4 (2017), 75–79.

[5] Maria Antoniak, David Mimno, and Karen Levy. 2019. Narrative Paths and Negotiation of Power in Birth Stories. In *Proc. ACM Human Computer Interaction*. CSCW.

[6] Katrina Armstrong, Mary Putt, Chanita Hughes Halbert, David Grande, J Sanford Schwartz, Kaijun Liao, Noora Marcus, Mirar Bristol Demeter, and Judy A Shea. 2013. Prior experiences of racial discrimination and racial differences in health care system distrust. *Medical Care* 51, 2 (2013), 144.

[7] Katrina Armstrong, Karima L Ravenell, Suzanne McMurphy, and Mary Putt. 2007. Racial/ethnic differences in physician distrust in the United States. *American journal of public health* 97, 7 (2007), 1283–1289.

[8] Laura Attanasio and Katy B Kozhimannil. 2017. Health care engagement and follow-up after perceived discrimination in maternity care. *Medical care* 55, 9 (2017), 830–833.

[9] Reuben M Baron and David A Kenny. 1986. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology* 51, 6 (1986), 1173.

[10] Yoav Benjamini and Yosef Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)* 57, 1 (1995), 289–300.

[11] Cynthia J Berg, William M Callaghan, Carla Syverson, and Zsakeba Henderson. 2010. Pregnancy-related mortality in the United States, 1998 to 2005. *Obstetrics & Gynecology* 116, 6 (2010), 1302–1309.

[12] David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *Journal of machine Learning research* 3, Jan (2003), 993–1022.

[13] CDC. 2019. Center for Disease Control and Prevention. https://www.cdc.gov/nchs/data/databriefs/db229.

[14] CDC. 2019. Center for Disease Control and Prevention: CDC Wonder. https://wonder.cdc.gov/.

[15] CDC. 2019. Center for Disease Control and Prevention: Pregnancy Mortality Surveillance System. https://www.cdc.gov/reproductivehealth/maternalinfanthealth/pregnancy-mortality-surveillance-system.htm.

[16] Stevie Chancellor, Michael L Birnbaum, Eric D Caine, Vincent Silenzio, and Munmun De Choudhury. 2019. A taxonomy of ethical tensions in inferring mental health states from social media. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*. ACM, 79–88.

[17] Irene Y Chen, Peter Szolovits, and Marzyeh Ghassemi. 2019. Can AI help reduce disparities in general medical and mental health care? *AMA journal of ethics* 21, 2 (2019), 167–179.

[18] Raj Chetty, Michael Stepner, Sarah Abraham, Shelby Lin, Benjamin Scuderi, Nicholas Turner, Augustin Bergeron, and David Cutler. 2016. The association between income and life expectancy in the United States, 2001-2014. *Jama* 315, 16 (2016), 1750–1766.

[19] Sheldon Cohen, Ronald C Kessler, and Lynn Underwood Gordon. 1997. *Measuring stress: A guide for health and social scientists*. Oxford University Press on Demand.

[20] Mike Conway and Daniel O'Connor. 2016. Social media, big data, and mental health: current advances and ethical implications. *Current opinion in psychology* 9 (2016), 77–82.

[21] Paul T Costa and Robert R McCrae. 1992. Normal personality assessment in clinical practice: The NEO Personality Inventory. *Psychological assessment* 4, 1 (1992), 5.

[22] Mayara Costa Figueiredo, Clara Caldeira, Tera L Reynolds, Sean Victory, Kai Zheng, and Yunan Chen. 2017. Self-tracking for fertility care: collaborative support for a highly personalized problem. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017), 36.

[23] Andreea A Creanga and William M Callaghan. 2017. Recent increases in the US maternal mortality rate: disentangling trends from measurement issues. *Obstetrics & Gynecology* 129, 1 (2017), 206–207.

[24] Andreea A Creanga, Carla Syverson, Kristi Seed, and William M Callaghan. 2017. Pregnancy-related mortality in the United States, 2011–2013. *Obstetrics and gynecology* 130, 2 (2017), 366.

[25] Aron Culotta. 2014. Estimating county health statistics with twitter. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*. ACM, 1335–1344.

[26] Brenda Curtis, Salvatore Giorgi, Anneke EK Buffone, Lyle H Ungar, Robert D Ashford, Jessie Hemmons, Dan Summers, Casey Hamilton, and H Andrew Schwartz.

[27] 2018. Can Twitter be used to predict county excessive alcohol consumption rates? *PloS one* 13, 4 (2018), e0194290.

[27] Heike Thiel de Bocanegra, Monica Braughton, Mary Bradsberry, Mike Howell, Julia Logan, and Eleanor Bimla Schwarz. 2017. Racial and ethnic disparities in postpartum care and contraception in CaliforniaâĂŹs Medicaid program. *American journal of obstetrics and gynecology* 217, 1 (2017), 47–e1.

[28] Munmun De Choudhury, Scott Counts, and Eric Horvitz. 2013. Major life changes and behavioral markers in social media: case of childbirth. In *Proceedings of the 2013 conference on Computer supported cooperative work*. ACM, 1431–1442.

[29] Munmun De Choudhury, Scott Counts, and Eric Horvitz. 2013. Social media as a measurement tool of depression in populations. In *Proceedings of the 5th Annual ACM Web Science Conference*. ACM, 47–56.

[30] Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. 2013. Predicting depression via social media. In *Seventh international AAAI conference on weblogs and social media*.

[31] Munmun De Choudhury, Meredith Ringel Morris, and Ryen W White. 2014. Seeking and sharing health information online: comparing search engines and social media. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*. ACM, 1365–1376.

[32] Mark P Doescher, Barry G Saver, Peter Franks, and Kevin Fiscella. 2000. Racial and ethnic disparities in perceptions of physician style and trust. (2000).

[33] Johannes C Eichstaedt, H Andrew Schwartz, Margaret L Kern, Gregory Park, Darwin R Labarthe, Raina M Merchant, Sneha Jha, Megha Agrawal, Lukasz A Dziurzynski, Maarten Sap, Christopher Weeg, Emily E Larson, Lyle H Ungar, and Martin EP Seligman. 2015. Psychological language on Twitter predicts county-level heart disease mortality. *Psychological Science* 26 (2015), 159–169. Issue 2.

[34] Centers for Disease Control, Prevention, et al. 2019. Building US capacity to review and prevent maternal deaths. Report from nine maternal mortality review committees.

[35] Kirsten Gillibrand: United States Senator for New York. [n.d.]. With Maternal Mortality Rates On The Rise In The United States, Gillibrand Announces New Legislation To Help Reduce Maternal Deaths, Help Hospitals Implement Best Practices To Prevent Women From Dying Before, During And After Childbirth. https://www.gillibrand.senate.gov/news/press/release/with-maternal-mortality-rates-on-the-rise-in-the-united-states-gillibrand-announces-new-legislation-to-help-reduce-maternal-deaths-help-hospitals-implement-best-practices-to-prevent-women-from-dying-before-during-and-after-childbirth.

[36] Abby Gardner. [n.d.]. Black Women Are Dying During Childbirth. Sen. Kamala Harris Is Working to Change That. https://www.glamour.com/story/senator-kamala-harris-bill-maternal-mortality-crisis.

[37] Alan E Gelfand and Adrian FM Smith. 1990. Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association* 85, 410 (1990), 398–409.

[38] Joseph Gibbons, Robert Malouf, Brian Spitzberg, Lourdes Martinez, Bruce Appleyard, Caroline Thompson, Atsushi Nara, and Ming-Hsiang Tsou. 2019. Twitter-based measures of neighborhood sentiment as predictors of residential population health. *PloS one* 14, 7 (2019), e0219550.

[39] Salvatore Giorgi, Daniel Preotiuc-Pietro, Anneke Buffone, Daniel Rieman, Lyle H. Ungar, and H. Andrew Schwartz. 2018. The Remarkable Benefit of User-Level Aggregation for Lexical-based Population-Level Predictions. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*.

[40] Amanda Michelle Gomez. [n.d.]. There's finally a group of lawmakers focused on one of the widest racial disparities in health care. https://thinkprogress.org/house-forms-first-ever-black-maternal-health-caucus-alma-adams-lauren-underwood-32791417ffd7/.

[41] Howard S Gordon, Richard L Street Jr, Barbara F Sharf, P Adam Kelly, and Julianne Souchek. 2006. Racial differences in trust and lung cancer patients' perceptions of physician communication. *Journal of clinical oncology* 24, 6 (2006), 904–909.

[42] Xinning Gui, Yu Chen, Yubo Kou, Katie Pine, and Yunan Chen. 2017. Investigating Support Seeking from Peers for Pregnancy in Online Health Communities. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017), 50.

[43] Sharath Chandra Guntuku, Anneke Buffone, Kokil Jaidka, Johannes C Eichstaedt, and Lyle H Ungar. 2019. Understanding and measuring psychological stress using social media. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 13. 214–225.

[44] Mark A Hall, Fabian Camacho, Elizabeth Dugan, and Rajesh Balkrishnan. 2002. Trust in the medical profession: conceptual and measurement issues. *Health services research* 37, 5 (2002), 1419–1439.

[45] Mark A Hall, Elizabeth Dugan, Beiyao Zheng, and Aneil K Mishra. 2001. Trust in physicians and medical institutions: what is it, can it be measured, and does it matter? *The milbank quarterly* 79, 4 (2001), 613–639.

[46] Mark A Hall, Beiyao Zheng, Elizabeth Dugan, Fabian Camacho, Kristin E Kidd, Aneil Mishra, and Rajesh Balkrishnan. 2002. Measuring patientsâĂŹ trust in their primary care providers. *Medical care research and review* 59, 3 (2002), 293–318.

[47] Arthur E Hoerl and Robert W Kennard. 1970. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* 12, 1 (1970), 55–67.

[48] Elizabeth A Howell. 2018. Reducing Disparities in Severe Maternal Morbidity and Mortality. *Clinical obstetrics and gynecology* 61, 2 (2018), 387–399.

[49] Elizabeth A Howell, Natalia Egorova, Amy Balbierz, Jennifer Zeitlin, and Paul L Hebert. 2016. Black-white differences in severe maternal morbidity and site of care. *American journal of obstetrics and gynecology* 214, 1 (2016), 122–e1.

[50] Elizabeth A Howell, Natalia N Egorova, Amy Balbierz, Jennifer Zeitlin, and Paul L Hebert. 2016. Site of delivery contribution to black-white severe maternal morbidity disparity. *American journal of obstetrics and gynecology* 215, 2 (2016), 143–152.

[51] Kokil Jaidka, Sharath Chandra Guntuku, Anneke Buffone, H. Andrew Schwartz, and Lyle Ungar. 2018. Facebook versus Twitter: Cross-Platform Differences in Self-Disclosure and Trait Prediction. In *Proceedings of the International AAAI Conference on Web and Social Media*.

[52] Margaret L Kern, Gregory Park, Johannes C Eichstaedt, H Andrew Schwartz, Maarten Sap, Laura K Smith, and Lyle H Ungar. 2016. Gaining insights from social media language: Methodologies and challenges. *Psychological methods* 21, 4 (2016), 507.

[53] Michal Kosinski, Sandra C Matz, Samuel D Gosling, Vesselin Popov, and David Stillwell. 2015. Facebook as a research tool for the social sciences: Opportunities, challenges, ethical considerations, and practical guidelines. *American Psychologist* 70, 6 (2015), 543.

[54] Katy Backes Kozhimannil, Connie Mah Trinacty, Alisa B Busch, Haiden A Huskamp, and Alyce S Adams. 2011. Racial and ethnic disparities in postpartum depression care among low-income women. *Psychiatric Services* 62, 6 (2011), 619–625.

[55] Robert S Levine, James E Foster, Robert E Fullilove, Mindy T Fullilove, Nathaniel C Briggs, Pamela C Hull, Baqar A Husaini, and Charles H Hennekens. 2016. Black-white inequalities in mortality and life expectancy, 1933–1999: implications for healthy people 2010. *Public health reports* (2016).

[56] Judette M Louis, M Kathryn Menard, and Rebekah E Gee. 2015. Racial and ethnic disparities in maternal morbidity and mortality. *Obstetrics & Gynecology* 125, 3 (2015), 690–694.

[57] Michael C Lu. 2018. Reducing maternal mortality in the United States. *Jama* 320, 12 (2018), 1237–1238.

[58] Marco Lui and Timothy Baldwin. 2012. langid. py: An off-the-shelf language identification tool. In *Proceedings of the ACL 2012 system demonstrations (ACL)*. 25–30.

[59] N Martin, E Cillekens, and A Freitas. 2017. Lost mothers. ProPublica.

[60] Andrew Kachites McCallum. 2002. Mallet: A machine learning for language toolkit. *http://mallet. cs. umass. edu* (2002).

[61] David Mechanic. 1996. Changing medical organization and the erosion of trust. *The Milbank Quarterly* (1996), 171–189.

[62] Danielle Mowery, Albert Park, Mike Conway, and Craig Bryan. 2016. Towards automatically classifying depressive symptoms from Twitter data for population health.

[63] Donald Musa, Richard Schulz, Roderick Harris, Myrna Silverman, and Stephen B Thomas. 2009. Trust in the health care system and the use of preventive health services by older black and white adults. *American journal of public health* 99, 7 (2009), 1293–1299.

[64] NYC Health. [n.d.]. Severe Maternal Morbidity: New York City, 2008-2012. https://www1.nyc.gov/assets/doh/downloads/pdf/data/maternal-morbidity-report-08-12.pdf.

[65] Gregory Park, H Andrew Schwartz, Johannes C Eichstaedt, Margaret L Kern, Michal Kosinski, David J Stillwell, Lyle H Ungar, and Martin EP Seligman. 2015. Automatic personality assessment through social media language. *Journal of personality and social psychology* 108, 6 (2015), 934.

[66] Michael J. Paul and Mark Dredze. 2011. You Are What You Tweet: Analyzing Twitter for Public Health. In *International Conference on Weblogs and Social Media (ICWSM)*. 265–272.

[67] Michael J Paul and Mark Dredze. 2017. Social monitoring for public health. *Synthesis Lectures on Information Concepts, Retrieval, and Services* 9, 5 (2017), 1–183.

[68] Emily E Petersen, Nicole L Davis, David Goodman, Shanna Cox, Nikki Mayes, Emily Johnston, Carla Syverson, Kristi Seed, Carrie K Shapiro-Mendoza, William M Callaghan, et al. 2019. Vital Signs: Pregnancy-Related Deaths, United States, 2011–2015, and Strategies for Prevention, 13 States, 2013–2017. *Morbidity and Mortality Weekly Report* 68, 18 (2019), 423.

[69] Daniel Preotiuc-Pietro, Jordan Carpenter, Salvatore Giorgi, and Lyle Ungar. 2016. Studying the Dark Triad of personality through Twitter behavior. In *Proceedings of the 25th ACM international on conference on information and knowledge management*. ACM, 761–770.

[70] Daniel Preotiuc-Pietro, Sina Samangooei, Trevor Cohn, Nicholas Gibbins, and Mahesan Niranjan. 2012. Trendminer: An architecture for real time analysis of social media text. In *In Proceedings of the 6th International AAAI Conference on Weblogs and Social Media, Workshop on Real-Time Analysis and Mining of Social Streams, ICWSM*.

[71] Daniel Preoţiuc-Pietro, H Andrew Schwartz, Gregory Park, Johannes Eichstaedt, Margaret Kern, Lyle Ungar, and Elisabeth Shulman. 2016. Modelling valence and arousal in facebook posts. In *Proceedings of the 7th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*. 9–15.

[72] Daniel Rieman, Kokil Jaidka, H Andrew Schwartz, and Lyle Ungar. 2017. Domain adaptation from user-level facebook models to county-level twitter predictions. In *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. 764–773.

[73] Robin Fields. [n.d.]. New York City Launches Committee to Review Maternal Deaths. https://www.propublica.org/article/new-york-city-launches-committee-to-review-maternal-deaths.

[74] James A Russell. 1980. A circumplex model of affect. *Journal of personality and social psychology* 39, 6 (1980), 1161.

[75] H Andrew Schwartz, Johannes Eichstaedt, Margaret L Kern, Gregory Park, Maarten Sap, David Stillwell, Michal Kosinski, and Lyle Ungar. 2014. Towards assessing changes in degree of depression through facebook. In *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*. 118–125.

[76] H Andrew Schwartz, Johannes C Eichstaedt, Margaret L Kern, Lukasz Dziurzynski, Richard E Lucas, Megha Agrawal, Gregory J Park, Shrinidhi K Lakshmikanth, Sneha Jha, Martin E P Seligman, and Lyle H Ungar. 2013. Characterizing geographic variation in well-being using tweets. In *Proceedings of the 7th International AAAI Conference on Weblogs and Social Media (ICWSM)*.

[77] H Andrew Schwartz, Johannes C Eichstaedt, Margaret L Kern, Lukasz Dziurzynski, Stephanie M Ramones, Megha Agrawal, Achal Shah, Michal Kosinski, David Stillwell, Martin EP Seligman, and Lyle H Ungar. 2013. Personality, gender, and age in the language of social media: The Open-Vocabulary approach. *PLoS ONE* (2013).

[78] H Andrew Schwartz, Salvatore Giorgi, Maarten Sap, Patrick Crutchley, Lyle Ungar, and Johannes Eichstaedt. 2017. DLATK: Differential language analysis ToolKit. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*. 55–60.

[79] Michael E Sobel. 1982. Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological methodology* 13 (1982), 290–312.

[80] VG Vinod Vydiswaran, Yang Liu, Kai Zheng, David A Hanauer, and Qiaozhu Mei. 2014. User-created groups in health forums: What makes them special?. In *Eighth International AAAI Conference on Weblogs and Social Media*.

[81] Sen. Elizabeth Warren. [n.d.]. Sen. Elizabeth Warren On Black Women Maternal Mortality: 'Hold Health Systems Accountable For Protecting Black Moms'. https://www.essence.com/feature/sen-elizabeth-warren-black-women-mortality-essence/.

[82] Carla Abou Zahr, Tessa M Wardlaw, and Yoonjoung Choi. 2004. *Maternal mortality in 2000: estimates developed by WHO, UNICEF and UNFPA*. World Health Organization.

[83] Mohammadzaman Zamani, Anneke Buffone, and H Andrew Schwartz. 2018. Predicting Human Trustfulness from Facebook Language. *arXiv preprint arXiv:1808.05668* (2018).

APPENDIX

We include further details on results and discussions from the main text below:

A SAMPLE TWEETS

The last column shows a sample of three tweets for the topic. To find these representative tweets, we extract topic loadings over a random set of 500,000 pregnancy-related tweets. We then order the tweets by topic loadings and hand-select three tweets (out of the top ten) that best describe the topic, ignoring noisy or uninformative tweets. For example, a tweet "teen rates!!!" would load extremely high in our first topic, but it doesn't capture any additional information over the list of the highest-weighted words within the topic. Note that all typos and emoticons in the tweets are included unchanged.

Topic Label	Sample Tweets
Teen Pregnancy (1.34%)	teenage pregnancy #iblamedavidcameron
	Decreasing infant mortality around the world <URL>
	#BecauseOfYolo teenage pregnancy rate has risen
Morning Sickness (0.54%)	The purpose of our lives is to give birth to the best which is within us Marianne Williamson #spirituality
	Ecotopic pregnancy diagnosis symptoms and complications <URL>
	Getting a sickness that isn't morning sickness while #pregnant #sucks #cough #throathurts #stuffynose #blah
Celebrity Pregnancies (1.42%)	Amber rose is pregnant ? #damnwiz
	hopefully kim k's pregnancy doesnt last 72 days
	Taylor swift pregnant by harry
Abortion Rights (2.15%)	Lawmakers ban shackling of pregnant inmates <URL>
	#SouthAfrica to care for all #HIV positive infants <URL> #worldaidsday #womensrights #children
	Nebraska governor rejects prenatal care funding for illegal immigrants <URL>
Maternal Studies (2.56%)	Lower autism risk with folic acid supplements in pregnancy <URL>
	Postpartum cardiovascular risk linked to glucose intolerance during pregnancy <URL>
	Increased autism risk linked to hospital-diagnosed maternal infections <URL>
Congratulatory Remarks (3.06%)	Congrats to <USER> and <USER> on the birth of their baby
	#5yearsago i gave birth to my wonderful daughter <3 <3 <3
	Awwwwww my nephew's wife is pregnant <3 congrats!

Table 5: Sample topics with sample tweets

B THEORETICAL MODELS

We present high-level details for each of the four models used in this paper. Detailed descriptions and evaluations can be found in the corresponding papers. Note that none of the models described below were developed for this paper.

Affect. An affect model was built using a set of 2,895 annotated Facebook posts. Each post was rated by two psychologists on a nine-point ordinal scale, based on the affective circumplex model introduced by Russell [74]. A ℓ_2 penalized linear (ridge) regression was built using 1—2grams extracted from each message. Using a 10-fold cross-validation setup, the ngram model resulted in a prediction accuracy (Pearson r) of 0.65. Full details can be found in Preoţiuc-Pietro et al. [71].

Depression. The MyPersonality data set [53], which consisted of approximately 154,000 consenting users who shared Facebook statuses and completed a 100-item personality questionnaire was used. The personality questionnaire is based on the International Personality Item Pool proxy for the NEO Personality Inventory [21]. This work then takes the average response to the seven depression-facet items (located within the larger Neuroticism scale) to estimate user-level degree of depression. A ridge-penalized regression model was built [47] using a set of 2,000 LDA topics and 1-3grams extracted over 27,749 individuals and tested on 1,000 random individuals who used at least 1,000 words across all of their statuses. This resulted in a final prediction accuracy (Pearson r) of 0.39. Full details can be found in Schwartz et al. [75].

Trust. Similar to the depression model, the trust model was built using the MyPersonality Facebook data set [53]. Consenting individuals were asked to share their Facebook statuses and answer a Big-Five personality questionnaire. The average of three of the ten trust-facet items from the agreeableness domain – (1) "I believe that others have good intentions," (2) "I trust what people say," and (3) "I suspect hidden motives in others" (reverse-coded) – was used as a measure of trust. A predictive model was built on 26,243 users who answered the above question and also shared Facebook statuses (with at least 1,000 words across all statuses) and evaluated on a smaller set of users (N =621) who answered the ten trust-facet items. Using a set of 2,000 LDA topics and 1-3grams, this resulted in a prediction accuracy (Pearson r) of 0.35. Full details can be found in Zamani et al. [83].

Stress. Participants were recruited through Qualtrics (an online survey platform, similar to Amazon Mechanical Turk), where each participant answered a series of demographic questions, the 10-item Cohen's Stress scale [19] and consented to share their Facebook statuses. The analysis was then limited to those who self-reported age and gender (female/male) and who posted at least 500 words across all Facebook statuses, resulting in a final set of 2,749 participants. A set of 2,000 Facebook topics were used as features in a ridge penalized regression model [47]. This resulted in a prediction accuracy of 0.32 (Pearson r), using a 10-fold cross validation setup. Full details can be found in Guntuku et al. [43].

C   DOMAIN TRANSFER: APPLYING FACEBOOK MODELS TO TWITTER DATA

All four of our theoretical models were trained and evaluated on Facebook data in their original papers, whereas we applied the models to Twitter data. Some of the models have been shown to work in other domains (i.e., stress on Facebook vs Twitter; Guntuku 2019). Additionally, previous work has found is that effect sizes tend to vanish without correcting for the domain transfer [72], which we argue makes our prediction task harder. Additionally, Rieman et al. [72] showed that user-level Facebook models applied to county-level Twitter data are stable in terms of direction of effect sizes.

D   SPATIAL DISTRIBUTIONS

Figure 3 shows the relationship between maternal mortality rates (residualized on race, median age of females, socioeconomics and risk factors) and the topic loadings for the *Congratulatory Remarks* topic. Markers in the scatter plot are colored according to U.S. Census regions (Midwest, Northeast, South and West). We see that lower usage of this topic is associated with high mortality rates. We also see spatial clustering across the regions. For example, the West tends to have lower rates of mortality but large variance in topic usage. The South has the most variation in mortality in addition to the largest outliers in topic usage. Figure **??** includes a similar set of plots for the theoretically-relevant features, showing significant associations between *affect* and *trust* and maternal mortality.
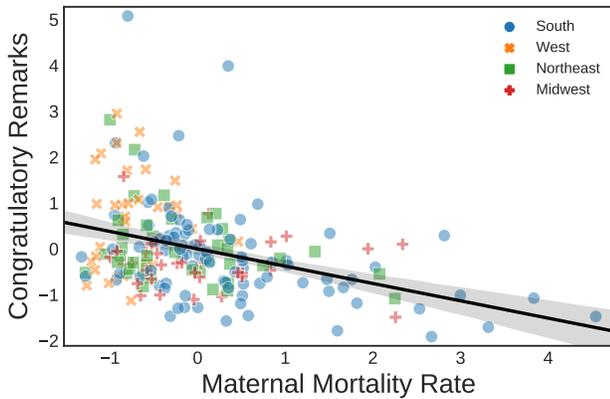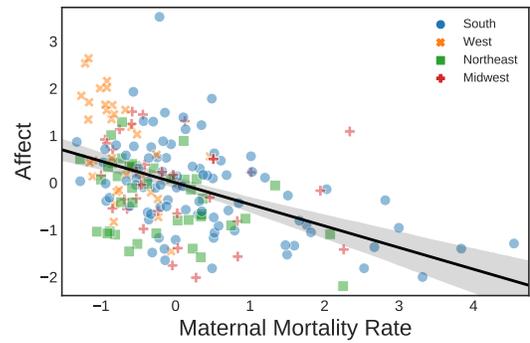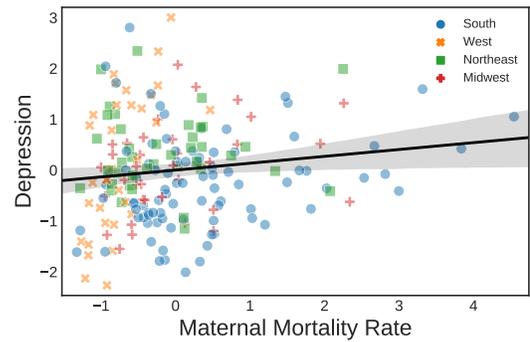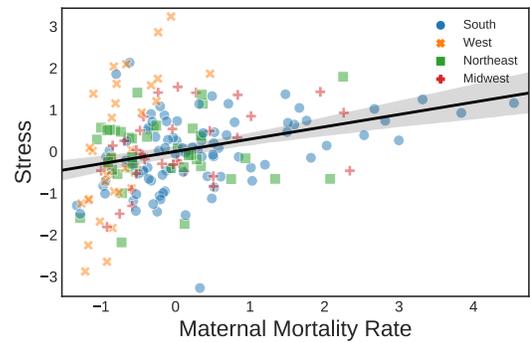
Figure 3: Maternal mortality rate (residualized) vs the *Congratulatory Remarks* topic loading. Dots are colored by which U.S. Census Region the county resides in: Midwest, Northeast, South and West.
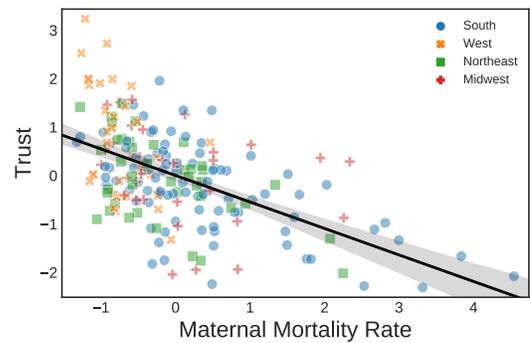
(a) Affect

(b) Depression

(c) Stress

(d) Trust

Figure 4: Maternal mortality rate (residualized) vs theoretically relevant features. Dots are colored by which U.S. Census Region the county resides in: Midwest, Northeast, South and West.