Media Influences on Social Outcomes: The Impact of MTV's *16 and Pregnant* on Teen Childbearing

Melissa S. Kearney
Department of Economics
University of Maryland
and National Bureau of Economic Research
kearney@econ.bsos.umd.edu

Phillip B. Levine
Department of Economics
Wellesley College
and National Bureau of Economic Research
plevine@wellesley.edu

January 2014

Abstract: This paper explores how specific media images affect adolescent attitudes and outcomes. The specific context examined is the widely viewed MTV franchise, *16 and Pregnant*, a series of reality TV shows including the *Teen Mom* sequels, which follow the lives of pregnant teenagers during the end of their pregnancy and early days of motherhood. We investigate whether the show influenced teens' interest in contraceptive use or abortion, and whether it ultimately altered teen childbearing outcomes. We use data from Google Trends and Twitter to document changes in searches and tweets resulting from the show, Nielsen ratings data to capture geographic variation in viewership, and Vital Statistics birth data to measure changes in teen birth rates. We find that *16 and Pregnant* led to more searches and tweets regarding birth control and abortion, and ultimately led to a 5.7 percent reduction in teen births in the 18 months following its introduction. This accounts for around one-third of the overall decline in teen births in the United States during that period.

Acknowledgements: We are grateful to Dan Fetter, Craig Garthwaite, Kelleen Kaye, Tim Moore, Diane Schanzenbach, and Doug Staiger for very helpful comments and conversations. We acknowledge the valuable research assistance of Filippos Petroulakis, Alex Roth, and Helen Willis. Seminar participants at Boston College, Brookings, Dartmouth College, George Washington University, the University of Virginia Batten School, and Wellesley College offered helpful comments. We gratefully acknowledge financial support from the University of Maryland Department of Economics and Wellesley College that enabled us to purchase the data used in this paper. Nielsen ratings data used in this project are proprietary and cannot be shared.

I. INTRODUCTION

It is a longstanding and open question how exposure to media images affects the behavior of viewers. Policy advocates and cultural observers worry, in particular, about the impact of exposure to sexual and violent content on the behaviors of adolescents. In some circles, the idea that teenagers respond to media content is a foregone conclusion, but determining whether the media images themselves cause the behavior is a very difficult empirical task.[1] The purpose of this paper is to examine the impact of the widely viewed MTV show, *16 and Pregnant*, on teenage attitudes and outcomes. This show purports to show the difficult reality of becoming a teen mother. As we document below, *16 and Pregnant* has drawn large audiences among relevant subpopulations. One clear indication of the show's success is the spinoffs it has generated: *Teen Mom, Teen Mom 2, and Teen Mom 3*. Could exposure to these media images of pregnant teens and very young new moms have had an impact on how teens think about pregnancy and ultimately on whether they become teen mothers themselves? If so, this would have important implications for thinking about how to effectively communicate with teens and influence their behaviors.

The context of teen childbearing in the United States makes this question an important one to study. In 2012, 29.4 out of every 1,000 girls between the ages of 15 and 19 (2.94 percent) gave birth in the United States. This rate is considerably higher than that in any other developed country, where typical rates of teen childbearing are more often in the range of 5 to 10 births per 1,000 girls in this age group (Kearney and Levine, 2012a). Though still an outlier internationally, the U.S. teen birth rate has declined dramatically over the past 20 years, falling from 61.8 births

[1]In terms of sexual behavior, evidence is cited showing that teens who watch sexual content on television are more likely to experience a teenage pregnancy. For example, popular press pieces by Stein (2008) and Tanner (2008) cite evidence from Chandra, et al. (2008). But evidence like this does not isolate the effect of the exposure from the choice of a particular type of individual to watch such content.

per 1,000 teen girls in 1991. This decline has occurred in two distinct waves. Between 1991 and 2008, it fell largely continuously from 61.8 to 40.2, representing an annual average rate of decline of 2.5 percent per year. Teen birth rates fell far more rapidly in the next four years, dropping from 40.2 to 29.4, or 7.5 percent per year.

The timing of the introduction of MTV's *16 and Pregnant* is such that it might conceivably have contributed in some measure to the most recent, very sharp decline. Ever since its introduction, various observers have made conflicting claims about the show's influence on teens. Some have highlighted the show's focus on the difficulties of raising a child at such a young age and have concluded from this coincident timing that the show is at least partially responsible for the recent decrease in teen childbearing rates.[2] Others argue that the show glamorizes teen pregnancy, with its cast members essentially becoming media "stars," whose lives are followed in the tabloids well after their show airs.[3]

In this paper, we take advantage of data from a number of sources to investigate the impact of the MTV show. Specifically, we investigate the following three questions: (1) Was exposure to the show substantial? (2) Did exposure to the show influence teens' interest in birth control or abortion?[4] (3) Did teen childbearing outcomes change as a result of the show's introduction? We use several measures of exposure, including Nielsen ratings data and the

---

[2]For example, see http://www.csmonitor.com/USA/Society/2010/1221/A-force-behind-the-lower-teen-birthrate-MTV-s-16-and-Pregnant and http://blog.thenationalcampaign.org/pregnant_pause/2011/11/us-teen-birth-rate-drops-a-dra.php.

[3]http://www.cnn.com/2011/OPINION/05/04/henson.teen.mom.show/index.html?_s=PM:OPINION, accessed 6/18/2013.

[4]Ideally, we would also conduct a full analysis of the impact of the show on sexual activity and contraceptive use as behavioral outcomes. We attempted to use data from the Youth Risky Behavior Surveillance (YRBS) system – years 2007, 2009, and 2011, to do that. These data are collected biannually in odd numbered years since 1991 for high school students who respond to the survey at school. A major disadvantage of using these data for this purpose, though, is that the sample sizes of youth in each geographic area can be quite small. In many states in which data are available, surveys were only completed by perhaps a thousand students regarding sexual activity and then perhaps a few hundred of them regarding use of contraception among those sexually active. Sample sizes at this level per state resulted in weak statistical power, leading us to omit a discussion of our analysis in this paper.

frequency with which individuals search for the show on Google and tweet about it on Twitter. We measure the influence on teens' decision making process with data on the frequency with which teens include terms like birth control or abortion in their searches/tweets.[5] Finally, we examine the impact on teen birth rates using Vital Statistics Natality microdata.

We use multiple empirical approaches to answer these questions. First, we present a descriptive analysis using several of our data sources to measure exposure. Nielsen ratings data is a direct indicator, but we can also examine high frequency data (daily or weekly) available from Google Trends and Twitter to look for time series spikes in searches and tweets on the show's title on the day/week that a new episode is released. Second, we conduct a similar analysis using high frequency data from these sources to look for increases in searches and tweets on terms like "birth control" when spikes appear in searches/tweets on the show's title. Third, we take advantage of geographic variation in the data, examining whether locations with relatively higher levels of searches/tweets about *16 and Pregnant* during the period when the show is on the air also experience relatively higher levels of searches/tweets about things like birth control.

Fourth, we exploit geographic variation in measures of the show's viewership (as captured by ratings) to investigate whether differential exposure to the show led to differential changes in teen birth rates. We correct for the fact that interest in a show about teen pregnancy is likely to be higher in locations where the teen birth rate is high. Geographic fixed effects hold constant time invariant factors. Greater interest may also occur in those areas where teen

---

[5]Identifying relevant terms included in searches or tweets is critical in this process. We explored other terms attempting to identify interest in sexual activity and adoption, but were unable to convince ourselves that they were sufficiently narrowly focused on the same concept we wanted to explore. For instance, searches for "how adopt" were dominated over our time period by the 2010 earthquake in Haiti and tweets including the term "adopt" or "adoption" were frequently not about children. In terms of sexual activity, we considered searches for terms like "does sex hurt" or "sex first time," as ones that may represent the reflective consideration about having sex that we want to test. We did not find a relationship between media exposure and these search terms.

childbearing is rising (or falling more slowly). To correct for this, we implement an instrumental variables (IV) strategy, predicting *16 and Pregnant* ratings with a broad measure of MTV ratings from a previous period. The identifying assumption of the IV approach is MTV ratings in the period before the show aired would be unrelated to subsequent trends in teen childbearing.

The results of our analysis indicate that exposure to *16 and Pregnant* was high and that it had an influence on teens' thinking regarding birth control and abortion. Large spikes in search activity and tweets about the show are evident exactly at the time a new episode was released. We also see an associated spike in Google searches and twitter messages containing the terms "birth control" and "abortion." Locations in which the show was more popular experienced greater increases in searches/tweets like this when the show was on the air.

Our most important finding is that the introduction of *16 and Pregnant* along with its partner shows, *Teen Mom* and *Teen Mom 2*, led teens to noticeably reduce the rate at which they give birth.[6] Our estimates imply that these shows led to a 5.7 percent reduction in teen births that would have been conceived between June 2009, when the show began, and the end of 2010. This can explain around one-third of the total decline in teen births over that period. Data limitations preclude us from conducting separate analyses of pregnancies and abortions, but we note that teen abortion rates also fell over this period (Pazol, et al., 2013). This suggests that the shows' impact is attributable to a reduction in pregnancy rather than greater use of abortion.

II. BACKGROUND

A. *Show Content*

MTV describes *16 and Pregnant* as an "hour-long documentary series focusing on the controversial subject of teen pregnancy. Each episode follows a 5-7 month period in the life of a teenager as she navigates the bumpy terrain of adolescence, growing pains, rebellion, and

[6] *Teen Mom 3* was not on the air yet during our sample period.

coming of age; all while dealing with being pregnant." The show first aired in June of 2009. Through October of 2013, it has been on the air for five "seasons" of 8 to 13 hour-long episodes (47 episodes tracking individual girls plus a handful of specials).[7]

To understand the nature of the messages conveyed in the show, we hired research assistants to view every episode of *16 and Pregnant* and record aspects of its content. Appendix A tabulates the results of that exercise. Here we offer some highlights. The girls on the show are largely from high birth rate states (15 of 47 from Texas, Florida, and Alabama). The racial/ethnic demographics of the girls on the show are fairly representative of the teen population in the United States, but not of the population of new teen mothers. Across episodes, 32 of 47 featured teen moms are white, non-Hispanic (in the U.S. teen population in 2010, 59 percent are white, non-Hispanic; but among teens who gave birth that year, only 39 percent are). Most girls (38 of 47) did not live in two parent households at time of pregnancy, which is consistent with broader statistics of teens giving birth (National Campaign to Prevent Teen and Unplanned Pregnancy, 2009).

Among the girls on the show, ambivalence towards teen childbearing is rampant. Only 18 out of 47 report opposition to their pregnancy when they found out, although none report that they were looking to get pregnant. The most common reasons for getting pregnant include not thinking she would become pregnant or have sex at that time (36 of 47) and ambivalence (28 of 47). Only 5 of 47 report trying to avoid a pregnancy, but failing. Three-quarters of the girls (36 of 47) report not using any form of contraception at the time they got pregnant.

An important emphasis on most episodes is the relationship between the girl and the father of her child, who is typically her boyfriend. Of all the pregnancies, four led to a marriage

[7]MTV only labels the show as having four seasons, but "season 2" ran from February through April and then October through December of 2010, which we label as two seasons. The show has spawned the spin-offs, *Teen Mom*, *Teen Mom 2*, and *Teen Mom 3*, which we will address subsequently.

prior to the birth and three led to adoption. There we no abortions. Almost all (40 of 47) of the boyfriends stick around through the pregnancy. Many fathers (31 of 44) live with the girl and her child afterwards and most of them (26 of 31) are heavily involved in the child's life. Only four of the fathers are completely uninvolved. Just over half (24 out of 44) of the relationships between the girl and her boyfriend either collapsed or were very strained by the end of the episode.

The show also emphasizes the implications of teen childbearing for the teen mother's health and well-being. Consistent with national trends, 11 out of the 47 births (23 percent) occurred via C-section; some occurring after up to 26 hours of labor.[8] In addition, in eight of the 47 pregnancies the mother or her baby experienced a significant health complication. One mother needed to spend a full month in the hospital as a preventative measure. One baby needed to be airlifted to another hospital to receive needed treatment. The show portrays extensive sleep deprivation for the teen mothers. Overall, the realities of the lives of teen mothers are presented in ways that may have been unknown or difficult to imagine for other teens viewing the show.

B. Previous Research on the Impact of Media Exposure

To the best of our knowledge, ours is the first paper to offer a credible estimate of the causal effect of specific media content on teen childbearing rates. There is descriptive survey work suggesting that the images portrayed specifically on the show *16 and Pregnant* might have had an influence on adolescent view. In a study for the National Campaign to Prevent Teen and Unplanned Pregnancy, Albert (2010) reports that 82 percent of teens who report watching *16 and Pregnant* indicate that it "helps teens better understand the challenges of pregnancy and parenthood." Only 17 percent report that it "glamorizes teen pregnancy." Chandra, et al. (2008) found that teens who viewed more sex content on television were more likely to become

---

[8]In the United States in 2007, the C-section delivery rate for women under age 20 was also 23 percent. For details, see http://www.cdc.gov/nchs/data/databriefs/db35.pdf.

pregnant.[9] Making the link to actual behavior that might lead to teen childbearing, however, is complicated by the fact that those who are more likely to give birth at a young age may be the ones who are more likely to view media with greater sexual content. No past research of which we are aware has adequately overcome this limitation.

Related research on the impact of media exposure on other outcomes has implemented methods that plausibly lead to causal estimates. In a true experimental setting, Gerber, et al. (2011) examine the impact of political advertising on television on public opinion polls, finding that greater exposure to the ads has substantial, but short-lived, effects on voter preferences. In quasi-experimental analyses, Dahl and DellaVigna (2009) find that rates of violent crime actually fall the same day that popular violent movies are released, partly attributable to an "incarceration effect." Card and Dahl (2011) find that family violence increases in the aftermath of an upset loss for a home football team. Both of these studies use the plausibly random timing of those events to determine that the effect may be thought of as causal. La Ferrara, Chong, and Duryea (2012) exploit geographic variation in exposure to Brazilian soap operas, finding that fertility rates fall when areas become exposed to soap operas on television that portray small families. Similarly, Chiou and Lopez (2010) find that businesses in Laguna Beach, CA, experienced more break-ins relative to a neighboring city following the introduction of another MTV reality TV show, *Laguna Beach: The Real Orange County*, which portrays the extensive wealth of the community. We employ similar empirical strategies in our subsequent analyses.[10]

---

[9]This analysis links the level of exposure to media with sexual content between the ages of 12 and 17 to subsequent rates of teen childbearing, finding that a positive relationship exists even after controlling for many observable factors. The authors recognize that their findings are not necessarily causal, stating "although our model included a wide range of potentially confounding factors as covariates, there is the possibility that we did not account for all factors that may alternatively explain the relationship we uncovered." On the other hand, they conclude, "despite these limitations, our study clearly suggests that television plays a role in shaping adolescent reproductive health outcomes."

[10]Price and Dahl (2012) describe the advantages of these types of approaches in identifying causal estimates of the impact of media influences on family outcomes.

Our paper offers a number of contributions to research on the influence of the media in affecting social outcomes. First, as noted earlier, most previous work addressing the influence of the media is correlational in nature. Our methods provide an important contribution to this literature. Second, our paper is among the first to provide plausibly causal estimates of the intermediate steps that occur between exposure and behavior. Our use of data from Google Trends and Twitter enable us to provide some gauge of what viewers are thinking about when they watch the show. Third, our use of data from these sources is among the first academic papers in which they are applied to examine social outcomes like teen childbearing.

III. DESCRIPTION OF THE DATA

Our analysis relies on a number of data sources, namely, Nielsen ratings, Google Trends, Twitter records, and Vital Statistics Natality microdata. In this section we provide relevant details on these various sources.

A. Nielsen Data

Nielsen ratings data have been the gold standard in measuring exposure to television shows for decades.[11] The Nielson Corporation collects these data from households either through meters that are attached to television sets or through diaries that are kept by household members. The results of their data acquisition are used to generate "ratings points," which represents the percentage of the population that watched an episode of a show. Ratings by demographic status and geographic location are also available.[12] Geography is defined by media markets, which are

[11] Nielsen is very restrictive in the release of its data. Researchers have access to these data through contractual relationships that allow very limited disclosure.

[12] Demographic information is only available in diary data because that is how Nielsen can determine who is watching the television when it is on. This does introduce a potential source of bias in that different demographic groups may have different propensities to complete the diary (cf. Kanazawa and Funk, 2001).

labeled "Designated Market Areas," or DMAs, which are collections of counties. We have data on 205 DMAs in the continental United States.[13]

Technological change has somewhat hindered Nielsen's ability to completely monitor household viewing habits. For several years, individuals have had the ability to record shows and watch them later and, more recently, individuals have the ability to stream recently released episodes of shows over the internet. Nielsen does measure some time shifting in that they can determine whether a recorded show is watched within one, three, or seven days of its release. The data we have obtained from Nielsen suggests the geographic correlation in ratings of shows watched live versus recorded is very high. This means that our use of ratings within one day of airing is a suitable measure of exposure. Similarly, despite the decline in viewership that has occurred recently as online streaming has become more accessible, watching television shows live is still the predominant manner in which individuals, and even teens, watch shows.[14]

The Nielsen data available to us include ratings data during the "sweeps" periods (November, February, May, and July) in each DMA for those aged 12 to 17, 18 to 24, and all ages for each season of *16 and Pregnant, Teen Mom*, and *Teen Mom 2*. *Teen Mom* (initial airing date, December 2009) follows four of the "stars" from the initial season of *16 and Pregnant* through their first years of motherhood. *Teen Mom 2* (initial airing date, January 2011) follows a similar strategy, focusing on "stars" of *16 and Pregnant*, season 2.[15] We aggregate the data for

[13]There are 210 DMAs in total. We do not have ratings data for the 4 DMAs in Alaska and Hawaii and our crosswalk between counties and DMAs omits the Laredo, Texas DMA, leaving us with usable data for 205 DMAs.
[14] Between 2011 and 2013, those between ages 12 and 17 (18 and 24) reduced the hours spent watching traditional television from 24.3 (26.8) hours per week to 21.4 hours per week (23.4) (Marketing Charts, 2013). This indicates that online viewing may be making some inroads into traditional TV viewing, but the traditional method appears to still be the main form. Moreover, MTV reports that in 2011 around 350,000 individuals per week streamed episodes of *16 and Pregnant* from MTV's website compared to the 2 million viewers, on average, who watched the show "live" (within one data of initial airing – MTV, 2012). Although this is a substantial number of streaming viewers, it also suggests that live TV ratings still represent a useful source of exposure to a TV show.
[15]The exact dates that each of these shows aired during our sample window is as follows:

those 12 to 24 and across all seasons/shows to generate an average measure of viewership in each DMA of the entire *16 and Pregnant* franchise. Our reason for doing this is to reduce the sampling variability that is generated by relatively small samples of young viewers in each DMA during a single sweeps period.[16] Based on this construction, we do not take advantage of any time series variability in the popularity of the shows once they began. Viewership as measured by Nielsen and aggregated by us is one of our primary measures of exposure.[17] To ease exposition, we will refer to this collection of three shows simply as *16 and Pregnant*.

A number of previous economics papers have made use of Nielsen ratings data, mainly to investigate issues of racial preference or discrimination. Kanazawa and Funk (2001) examine Nielson Ratings for professional basketball games as a function of white player participation; Aldrich, Arcidiacono, and Vigdor (2005) examine Nielsen ratings for ABC's Monday Night Football as a function of quarterback race; and Myers (2008) examines how the racial make-up of a local television news staff affects ratings. These papers examine Nielson ratings as the variable of interest and investigate determinants of ratings. We have a different conceptual goal, which is to use Nielsen ratings as a measure of exposure to content.[18]

---

16 and Pregnant: 6/11/2009-7/31/2009, 2/16/2010-4/20/2010, 10/26/2010-1/5/2011, 4/19/2011-6/28/2011, 3/27/2012-6/6/2012.

Teen Mom: 12/8/2009-1/26/2010, 7/20/2010-10/12/2010, 7/5/2011-9/20/2011, 6/12/2012-8/28/2012.

Teen Mom 2: 1/11/2011-3/29/2011, 12/6/2011-2/14/2012, 11/12/2012-2/12/2013.

All shows are ongoing in new seasons beyond those dates. A new series, *Teen Mom 3*, began in August of 2013 and follows the lives of the girls from season 4 of *16 and Pregnant*. Its later starting date pushes it beyond the window of analysis in this paper.

[16]To be more specific, we have data from 7 months in which one of these shows was on the air during a sweeps period. The average number of viewers tracked for those 12-24 is 192 per DMA/month. This leaves us with 1,344 viewers in all 7 months, on average for a DMA. With an average rating of 1.8, the average number of viewers per DMA who watched one of these shows is 24. We have chosen not to conduct this analysis on a monthly basis, because it would leave us with just 3.5 viewers per DMA per month, on average.

[17]The public controversy regarding the potential "glamorizing" effect of *16 and Pregnant* may include other forms of exposure, like tabloid coverage, which we cannot capture. It is possible that our analysis could generate different results if we focused on exposure through these alternative forms of media.

[18]Some studies in marketing and advertising have used Nielsen ratings to measure exposure to advertising. An example from economics includes a study by Saffer, Wakefield, and Terry-McElrath (2007) on the effect of exposure to nicotine replacement therapy advertising on youth smoking.

B. Google Trends

Data from Google Trends provides indexed values of the relative frequency with which individuals search Google for a particular term. Historical data going back to 2004 are available for countries, states, and cities; we focus on data from the United States collectively and by state in different parts of the analysis. We focus on searches specifically for *16 and Pregnant*, ignoring searches for *Teen Mom* or *Teen Mom 2* because of the ambiguity of those words (particularly teen mom) in a search. The index is designed to vary between 1 and 100, where a 100 is assigned to the period or location in which the search frequency is the highest relative to all searches conducted. Index values in other periods or locations are determined proportionally to that maximum value. Data are available beginning in 2004. Geographical searches are possible because Google is able to identify a user's location by the IP address. Since IP addresses are usually assigned to internet service providers within region-based blocks, an IP address can often be used to identify the region or country from which a computer is connecting to the Internet.

Presumably because of the extent of search activity, the results provided by Google Trends are based on a sample of searches rather than the full universe. If a data cell is defined narrowly enough (i.e. too short a period of time over too small a population in the area), in a particular sample search, activity may be sufficiently limited that Google chooses not to report it (indicated by a reported index value of 0). This approach by Google generates two potential problems. The first is sample selection bias, which would occur if, in small samples, only those random draws that generate unusually high search activity are reported. The second is just ordinary sampling variability, which is a problem largely for standard error calculations if we treat the reported data as fixed constants rather than random variables. To overcome these obstacles, we repeat our searches on Google Trends multiple times and select only those states for which data are available in each period. We then take the average of the index values, which substantially reduces the sampling variability. Another shortcoming of these data is that they do not include any demographic information, including age, about the individual conducting the search. We maintain the assumption, which we recognize to be imperfect, that the types of searches that interest us for this project are likely, but certainly not exclusively, made by adolescents and young adults.

Google Trends search data has now been used by a number of previous authors, mostly in studies interested in forecasting economic activity and by studies in the field of finance. Papers by Choi and Varian (2009a and 2009b) showed that Google Trends could help predict economic indexes, including predictions of tourism activity and unemployment insurance claims.[19] Google Trends data have also been used in a number of finance studies as a measure of investor attention and search activity (cf. Da, Gao and Engelberg, 2011; and Vlastakis and Markellos, 2012). It has also been used to examine labor market questions. For example, Baker and Fradkin (2013) use these data combined with administrative unemployment insurance data from 2006-2011 in Texas; they show that individuals receiving UI search less than individuals who are unemployed and who are not receiving UI. Garthwaite, Gross, and Notowidigo (forthcoming) use Google Search data to examine job search activity in their analysis of the relationship between public health insurance provision and employment; they find that searches for the term “job openings” peaked in months when public health insurance access was limited.

Perhaps the papers that use Google data that are closest in intent to our analysis are those by Stephens-Davidoff (2013a and 2013b). The first paper uses Google Trends data to explore the

---

[19]Choi and Liu (2011) cite nearly 20 papers that have used Google Search data in a variety of market contexts: capital markets, entertainment markets, labor markets, real estate markets, and healthcare markets.

impact of racism on votes for Barak Obama. To do this, the author compares changes in election results between 2004 and 2008 as a function of the rate of Google searches for racial epithets by state. The second paper uses Google Trends data to better track child maltreatment over the course of the business cycle based on searches one might conduct if they were suspicious of this form of activity (like "child abuse" or "child neglect"). Still, our paper is among the first to use these data as a lens to interpret the intermediate steps between exposure to an external stimulus and social outcomes.

C. Twitter

Twitter records all "tweets" – defined as messages to the online Twitter message service, where messages are limited to 140 characters – made by individuals. Accessing those data is not as straightforward as using Google Trends. One can conduct a search on twitter.com/search and receive a list of recent tweets that contain a search term. There is no way to access historical data nor is there a way to count the frequency of tweets on this public website. The Twitter "fire hose" (a library of past tweets) can be obtained, but it is extremely difficult to work with because of the format and amount of data available. Obtaining these data requires the use of a third party vendor that has a contract with Twitter to process searches, aggregating the frequency of tweets that contain specific search terms and providing geographic variation in their frequency. We used Topsy Labs.[20] As with our use of Google Trends data, we restrict our attention to tweets including *16 and Pregnant*, ignoring *Teen Mom* and *Teen Mom 2* because of the ambiguity of those terms in a tweet.

[20]In December of 2013, Apple Inc. purchased Topsy Labs. Although the service's future is unclear, at that time they ceased accepting new user contracts. The data we have extracted from Topsy Labs are available upon request from us, but they can no longer be accessed at this time from Topsy Labs itself. Other providers of Twitter data are available, however.

One limitation in these data is that geographic detail is difficult to obtain. Only a very small share of users identify where they live when they open their Twitter account and those locations are not updated if the account holder moves. Twitter also does not release IP addresses for tweets coming from the Twitter fire hose (largely because of the extent of tweeting from cell phones), so geography cannot be identified that way. To determine the location of the person sending the tweet, Topsy uses a probabilistic model that assigns an individual to a place based on a large number of factors. Examples include some information provided in the user's profile, check-ins at events/locations (i.e. foursquare), times tweets are made (to help get time zone), specific tweet content ("if you don't drive at least 80 on the mass pike, get off the mass pike"), and geographic detail contained in hashtags ("Excited for our upcoming kickball tournament to benefit NCCF! Hope you can join us on 9/14! #chevychase #bethesda"). They can then use the small sample of account holders who report geographic information to validate their methods. They report: "our methods allow us to geoinfer the origin of tweets with over 90% coverage by country, 80% accuracy by state/province, 40% accuracy by city, at a 90-95% confidence level."[21]

In our analysis, we use all available tweets between January 1, 2009 and December 31, 2012; 38 billion tweets from the United States in English are available over this time period.[22] Using these data, we can tabulate the total number of tweets made from each state, conditional on the availability of location. We then calculate a "tweet rate" by determining the number of

---

[21]This statement was provided in a personal communication between us and Topsy in an email dated May 21, 2013. We should also indicate that the process of generating these geographic data appears to be a work in progress. In working with these data we have identified data abnormalities that seem to occur in the geographic data, but not the national data. Over time, some of these issues have been resolved and new issues introduced as the algorithm for creating these data has been modified. The basic patterns of our results have never been substantively affected, but we believe it is prudent to interpret our reported results using geographic Twitter data with some caution.

[22]The data available to us do not include all tweets made during this period. We do not have access to tweets that have been deleted at the request of a user or tweets from users who have closed their account. This also raises another issue with respect to the use of these data – even the historical data is dynamic. As users delete tweets or close accounts, the available historical data changes. This means that subsequent investigators who attempt to replicate our results will not be able to do so without obtaining the exact data extract that we obtained. We are able to provide these data upon request.

tweets per one million total tweets that contain a specific search term (in the United States in English). As with Google Trends data, no demographic detail, including age, is available on the person sending the tweet. We simply assume that those tweeting about *16 and Pregnant* are more likely to be younger and a potential viewer of the show. Appendix B describes specific Twitter content related to the entry "*16 and Pregnant"* and other associated terms.

Few examples exist of papers that use Twitter data in an economics analysis of individual behavior. There are a handful of studies in political science that use Twitter data to identify political ideology and study information and influence networks in this context (cf. Barbara, 2013). Another set of papers in the field of finance are also available that examine the relationship between Twitter data and financial activity. For instance, Blakespoor, Miller, and White (2012) find that firms' use of Twitter to send hyperlinked information to market participants is associated with relevant market outcomes in a way that is consistent with a reduction in information asymmetry. Liu, Rui, and Whinston (2013) examine how Twitter activity affects movie sales. They find that positive tweets increase movie sales and negative tweets decrease sales. We believe we are the first to use Twitter data in an empirical application to measure what those tweeting are thinking as an intermediate step in a broader investigation of an important social outcome.

D. Vital Statistics Natality Microdata

Our final source of data is from the Vital Statistics system and includes much of the information on birth certificates for virtually every live birth in the United States between January 2006 and December 2010. The main strength of these data is its universal nature and large sample size (several hundred thousand teen births per year). For the purposes of this project, the ability to identify the exact age of the mother for every birth provides us with the ability to generate counts of births for very precise age groups, including younger and older teens. We also take advantage of the county identifiers that are available. Access to county identifiers is particularly useful because it provides us with a way to aggregate births by occurrence in each DMA, which is the only source of geographic information available in the Nielsen data that we use to measure exposure. We are also able to approximate the month of conception in these data because month of birth and gestational age are available. This enables us to better link the timing of exposure to the show to the timing of the activities that led to a subsequent birth. All birth timing reported in the remainder of this analysis is based on the month of conception.

IV. DESCRIPTIVE ANALYSIS OF EXPOSURE

Before we can undertake an analysis of the influence of *16 and Pregnant* on teens' thinking regarding birth control and abortion or on the rate of teen childbearing, it is important to document that teens' exposure to the show was substantial. That is a necessary condition for it to have had much influence. In this section, we use Nielsen ratings along with data on tweets and Google searches to provide that documentation.

The entertainment media, basing its assessments on Nielsen ratings, generally references *16 and Pregnant* as a hit show (see, for instance, Adalian, 2010; and Bricklin, 2012). Just after it was released, MTV reported that the show "has garnered amazing ratings … and consistently taking the #1 spot in its time period among (women between ages 12 and 34) and #1 across all of television among (women between ages 18 and 24)" (Seidman, 2009). More popular episodes attracted over 3 million viewers in total and received ratings up to 8.0 among women between the ages of 18 and 24 (Seidman, 2010). As a point of comparison, 6.4 million viewers, on average, watched weekly episodes of the hit ABC comedy, *Modern Family*, in the 2010-2011 season and the show scored an average rating of 4.9 among adults between the ages of 18 and 49 (Gorman, 2011). Although the target population of *16 and Pregnant* is considerably smaller, among that narrow group it is a very highly rated show. Indeed, a nationally representative survey of around 1,000 teens between the ages of 12 and 19 found that 71 percent had seen an episode of *16 and Pregnant* (Albert, 2010).

Considerable geographic variation is also evident in the extent of exposure. In Figure 1, we report the average rating by DMA for all three shows in the *16 and Pregnant* franchise for those between ages 12 and 24 in all sweeps months during our 2009-2012 sample period in which one of these shows was on the air. The results indicate that the show is the most popular among those in the South and in the Midwest, with some pockets of popularity elsewhere.

As a point of comparison, it is useful to present these results with a similar analysis of the geographic variation in teen childbearing by DMA; we report this in Figure 2. This figure defines teen births by age of conception rather than the more traditional age of birth for consistency with the remainder of our analysis. In this figure, we also see that teen births are more common in the South and Midwest. However, they are also very high in the Southwest where the show is not popular and very low in New England where the show is reasonably popular. Ratings are very low in California, where teen birth rates are moderate. Statistically, the population weighted correlation between Nielson ratings data and teen birth rates is 0.16, which is positive, but small.

We also use data from Google Trends and Twitter on the likelihood of searching for or tweeting about the show as an alternative measure of exposure. The presence and magnitude of spikes in searches/tweets for the show around the time that it is on the air suggests greater exposure. These data indicate that variability in searches/tweets over time is extensive.

Figure 3 displays weekly data between the beginning of 2009 and the end of 2012, highlighting the weeks during that period in which new episodes of the show were released. These data show clear signs of spikes in searches and tweets precisely in the weeks of this period in which new episodes were available. Searches peaked in the spring of 2011 and tweets peaked in the summer of 2009, both during periods in which the show was on the air. Other variability is present in the data, but the predominant pattern is the presence of these spikes.

In Figure 4, we restrict our attention to plotting daily trends in searches and spikes between November 1, 2010 and December 21, 2010. We chose a period this short because Google Trends data are only available at a daily rate when the search is restricted to a relatively short period of time. We chose these particular days because it is in the middle of *16 and Pregnant's* run since it first aired through the end of our sample period. As with the weekly data, we see the existence of spikes in the data that are well-timed to the release of a new show. Interestingly, the spike does not occur on the day the episode is released, but the following day. This makes sense if individuals watch the show at night and want to talk about it with their friends the next day. We view these spikes as supportive of extensive exposure to the show.

V. EMPIRICAL METHODOLOGY

A. Analysis of Google Search and Twitter Activity

1. Use of High Frequency Data on Searches and Tweets

What do people search for and tweet about in the immediate period following the release of a new episode? Data from Twitter and Google Trends is available at very high frequencies, either daily or weekly, providing us with the ability to look for spikes in search terms just after a new episode is aired. We took advantage of these data earlier in examining spikes for searches/tweets about the show itself; we formalize those analyses within an econometric framework here. We can also investigate other terms that individuals search for or tweet about, including terms related to birth control or abortion, which may also spike when a new episode is released. The methods for conducting these analyses are identical.

More specifically, we estimate the following Ordinary Least Squares regression models:

$$16P_t = \beta_0 + \beta_1 NewRelease_t + \beta_2 X_t + \varepsilon_t \qquad [1]$$

$$searchBC_t = \beta_0 + \beta_1 NewRelease_t + \beta_2 X_t + \varepsilon_t \qquad [2]$$

where *16P* represents the term *16 and Pregnant* in a search or tweet, *searchBC* is shorthand for a search or tweet for a term like "birth control," which would represent a form of pregnancy-related behavior, *NewRelease* indicates that period in which a new episode is released, and *X* represents seasonal fixed effects (when we use weekly data from Google) and quadratic trends.

We use two different units of time in our analysis. First, when we use Twitter data, we restrict our attention to those weeks in which the show is "in season," and take advantage of daily variation in outcomes. We modify Equations [1] and [2] modestly in these models to include a lag of the day of new release since "instantaneous" responses may include the following day for a show that airs the prior evening. Second, when we use Google data, we consider the entire time period between January 2009, the beginning of the year in which the show began, and December 2012, and focus on weekly variation, distinguishing between the weeks in which a new episode was in season relative to other weeks of the year. In either case, we would conclude that the introduction of the show has an impact if β_1 is positive. We consider any spike in searches or tweets about the show or any other search term occurring immediately following the release of a new episode to be causally related to the release of that episode.

We extend this analysis by directly treating searches/tweets about *16 and Pregnant* as a measure of exposure and examining whether increased exposure (as captured by this measure) alters the likelihood of searching for other terms related to pregnancy determinants. Formally, we estimate the model:

$$searchBC_t = \beta_0 + \beta_1 16P_t + \beta_2 X_t + \varepsilon_t \qquad [3]$$

This model has the advantage of being able to account for variation in the popularity of particular episodes of the show. In essence, placing an indicator for new release on the right hand side captures exposure to an episode of average popularity. Instead, by using a specific measure of the popularity of each individual episode taken from the same medium (i.e. tweets about birth control as a function of tweets about *16 and Pregnant*), our model is more powerful.

2. Use of Low Frequency Data on Searches and Tweets with Geographic Variation

We also consider the relationship between searches/tweets for pregnancy-related behavior and exposure to *16 and Pregnant* using lower frequency data that varies geographically. Searches/Tweets for *16 and Pregnant* prior to the show's introduction will be minimal, but those searches increase differentially in some states compared to others. This approach simply asks whether those differentials are linked to changes in search/tweet activity for our measures of pregnancy-related behaviors. A straightforward extension of the model represented in equation [3] captures the methodology we introduce here, as represented by equation [4]:

$$searchBC_{jt} = \beta_0 + \beta_1 16P_{jt} + \beta_2 U_{jt} + \gamma_j + \delta_t + \varepsilon_{jt} \qquad [4]$$

Much of the notation is as previously defined. The additional subscript j represents geography, states in this case. State and period fixed effects are also added along with U, the state unemployment rate.

B. Analysis of Teen Births: Vital Statistics Natality Data with Geographic Variation

To investigate whether exposure to the show, *16 and Pregnant,* led to a change in rates of teen childbearing, we exploit the timing of the show's introduction combined with geographic variation across media markets in ratings. We investigate whether there was a more rapid decline in rates of teen childbearing in locations in which the show was more widely viewed. We use quarterly variation in births in our analysis because, as a conceptual matter, changes in sexual behavior and birth outcomes are likely to adjust more slowly (if at all) in response to the airing of new episodes. The next sexual encounter may not be impending. A trip to the drug store, clinic, or doctor is required to obtain contraception. Pregnancies can occur only monthly and only with some probability even if the teen engages in unprotected sex. In these circumstances, lower frequency data makes more sense.

The OLS representation of the relationship takes the following form:

$$ln(B_{jt}) = \beta_0 + \beta_1 Rate16P_j{*}post_t + \beta_2 U_{jt} + \gamma_t + \delta_j + \varepsilon_{jt} \qquad [5a]$$

where j indexes media markets (DMAs) and t indexes quarters. The outcome variable of interest ln(B) represents the natural logarithm of the teen birth rate. The explanatory variable of primary focus is the interaction of *Rate16P,* which represents Nielsen ratings among those between ages 12 and 24 for *16 and Pregnant*, and *post*, which is an indicator variable for calendar quarters after June 2009 when the show began. The rest of the notation is the same as defined above.[23] Our sample of births consists of births that were conceived between January 2006 and December 2010, yielding 20 quarters of data. Along with data from 205 DMAs, our full sample consists of 4,100 DMA by quarter observations.[24]

23 When we estimate this OLS model and the IV version presented below, the regression is weighted by female teen population in the media market and standard errors are adjusted for clustering at the level of media market.

[24] In a log birth specification, a handful of observations get dropped because no births occurred in that DMA/quarter, explaining why listed sample sizes are often slightly less than this value.

The model includes a measure of economic conditions in the state, captured by U, the DMA unemployment rate.[25] Previous work has found that weaker economic conditions, are associated with lower rates of teen childbearing (e.g., Colen, Geronimus, and Phipps, 2006; Kearney and Levine, 2012b). Because we are only using data over a relatively short time period, we do not include the usual list of other policy or demographic variables in our analysis because they do not have sufficient variation over this time period to be important.[26]

As described earlier, the ratings measure is an average ratings taking from aggregating all periods after *16 and Pregnant* began and include *Teen Mom* and *Teen Mom* 2. The interaction of this constant measure of ratings with a *post* indicator clarifies that ratings prior to the show's introduction are naturally set to zero and then take on the average value of the show's ratings afterwards. The variation exploited by this empirical specification is the uniform shock that hit all markets when the show aired (which is captured by quarter fixed effects), plus the variation across places in how widely viewed the program was. As we described earlier, data limitations preclude us from exploiting variation over time within a market in the show's popularity after its introduction.

A critical issue in implementing this approach is accounting for the possibility that locations in which the show is more popular are not randomly selected. Perhaps the show is more popular in locations with elevated rates of teen childbearing. If so, OLS estimates of the

[25] To construct DMA level unemployment rates, we applied county level unemployment rates available from the Bureau of Labor Statistics to a cross-walk from counties to DMAs.

[26] For example, none of the three main types of abortion restrictions, mandatory delay periods, state level restrictions on Medicaid funding of abortion, and parental notification laws changed across states during the 2006 to 2010 period. Only three states changed welfare family cap policies. A number of states implemented Medicaid Family Planning waivers, but all in 2006 and 2007. The correlation at the state level between the real value of welfare benefits in 2006 and in 2010 is 0.97. For a discussion of targeted state level policies potentially relevant to teen childbearing, see the discussions in Kearney and Levine (2012b) and Lopoo and Raissian (2012). In addition to not having sufficient variation over the short time period analyzed, the fact that these policies are at the state level provides a complication for the potential inclusion of these policy controls in our DMA level analysis, since some DMAs cross state lines.

relationship between ratings and teen childbearing would include a positive bias, incorrectly suggesting that higher ratings lead to more teen births. To account for time-invariant differences across markets, we include market level fixed effects in our regression models.

Yet we might still worry about unobserved time-varying differences across markets that correlate with both the popularity of the show and the subsequent rate of teen childbearing. Perhaps the show is particularly appealing to individuals living in locations where teen childbearing is rising (or, in the recent past, falling more slowly). If this were the case, the estimated relationship of interest in an OLS model even with fixed effects would have a bias towards finding a positive effect of the show on teen birth rates.

To address this form of bias, we utilize an instrumental variables (IV) approach. We instrument for the show's ratings using ratings among those between ages 12 and 24 for all shows that aired on MTV on weekday evenings between 9:00 and 10:00 in the 4 sweeps months preceding the introduction of *16 and Pregnant* (July 2008 through May 2009). Since those ratings are determined prior to June 2009, our instrumental variable strips out the variation in *16 and Pregnant* ratings that are specifically attributable to that particular show, and potentially reflective of a time-varying latent preference for a show about teen mothers. Before *16 and Pregnant* was introduced, MTV programming contained a range of other reality TV shows, but none were specifically related to teen childbearing.[27]

The first stage regression in the IV framework takes the following form:

$$Rate16P_{j}{*}post_t = \beta_0 + \beta_1 MTV0809_j{*}post_t + \beta_2 U_{jt} + \gamma_t + \delta_j + \varepsilon_{jt} \qquad [5b]$$

where *MTV0809* represents the ratings among those between ages 12 and 24 for shows that aired between 9:00 PM to 10:00 PM on MTV between July of 2008 and May of 2009 in each media

[27]A number of shows aired on MTV during this time period, including *Real World*, *The Hills*, and *The Challenge*. Although all of the shows on MTV are geared towards teens/young adults, none of them share themes similar to those on *16 and Pregnant*.

market. It is time invariant. Its interaction with *post* is consistent with that between *Rate16P* and the *post* indicator, creating an instrumental variable (labeled here as the interaction of two variables) that takes on the value of zero in the quarters before the show was introduced and the value of ratings for shows from 2008/09 in the following quarters.

To facilitate interpretation of the IV estimates, it is useful to consider the experimental analog. First, assume some degree of inertia exists in TV viewing, so that MTV has developed a brand loyalty leading certain individuals to be more inclined to tune in regardless of what is on. Now suppose MTV randomly changed content across markets such that a "treatment" group of DMAs would have seen the MTV programming content switch to *16 and Pregnant* and a control group of DMAs would have continued with existing MTV programming, which is not about teen motherhood. We then would compare the subsequent change in teen births, asking whether there was a significant change between treatment and control DMAs.

The experiment implicit in our IV analysis is related to this sudden change of content to *16 and Pregnant*, but the exposure is not a binary treatment variable. In our IV setting, the content of what was on the air changed across the country at the same time, but "control" and "treatment" groups are differentiated by the size of the MTV audience in the immediately preceding period. Those locations in which the MTV audience is relatively small better approximates the control group and locations in which the audience is relatively large better approximates the treatment group. The treatment effect is identified off the relative change in outcomes in relation to the relative size of the pre-existing MTV audience. If the show is effective at changing attitudes and behaviors, then we would expect to see a relatively larger decline in rates of teen childbearing in places with a larger MTV viewership.[28]

---

[28]To be clear, this approach is *not* designed to identify the effect of assigning a random teenager or young adult to watch *16 and Pregnant*. Individuals select into MTV viewership in all periods. We are identifying the effect of the

The concept of the IV approach is apparent in the reduced form specification:

$$ln(B_{jt}) = \beta_0 + \beta_1 MTV0809_j{*}post_t + \beta_2 U_{jt} + \gamma_t + \delta_j + \varepsilon_{jt} \qquad [5c]$$

The explanatory variable $MTV0809_j$, which is constant over time within a media market, captures longstanding differences in teen birth rates in relation to the characteristics of the typical MTV audience. This term is not included alone because it is captured by market fixed effects. The interaction of this variable with the "post" indicator captures the change in that relationship, which could plausibly be assigned to the change in content of what is being aired for that same audience. This idea mimics the underlying experiment that our IV strategy is seeking to simulate.

For this IV approach to be valid, we need to maintain an exclusion restriction that subsequent trends in teen childbearing would not be correlated with earlier MTV ratings except for a direct effect of *16 and Pregnant.* We also need the relationship between the earlier MTV ratings and the later *16 and Pregnant* ratings to be monotonic. Figure 5 demonstrates that this relationship holds. It presents a simple scatter plot depicting the relationship between these MTV ratings before *16 and Pregnant* started and the ratings for these shows in the period after its introduction across DMAs. The pattern is clearly linear. A bivariate regression between these two variables yields a t-statistic of around 8.0, both weighted and unweighted. The coefficient itself suggest that ratings for *16 and Pregnant* are 1.38 points higher (more than double) than they were for a typical MTV show that aired on weekday evenings in 2008-09. The correlation coefficient is around 0.5 weighted or unweighted. Given the strength of these relationships, it is

---

show on teen births given the realized MTV audience. This approach allows us to say what happened because the show was on the air. The effect might have occurred entirely among MTV viewers, or perhaps there were spillover effects, say in the form of changing peer group norms, to non-MTV watching individuals. The effect that we observe should be thought of as the effect of the show's content on attitudes and behaviors, combined with MTV's ability to draw in an audience of teens who are susceptible to being influenced by MTV content and for whom becoming a teen mother is a possibility, or who have influence on another individual who might become a teen mother. To state the obvious, if no one watched MTV, or if the only ones who watched MTV were individuals who would never become teen mothers (or influence someone who might), then regardless of how compelling the content of this show was, it would not have affected outcomes.

not surprising that we satisfy the need for sufficient power in our first stage regression; the F-statistic on the omitted instrument is 46.

In an alternative specification, we examine the time path of the impact on birth rates as a function of exposure to the content of *16 and Pregnant* episodes. The time path of any change in a behavioral outcome is determined by two phenomena. First, the impact of the show on behavior might accumulate over time, so the change in outcomes might accumulate. Second, outcomes such as getting pregnant or meeting a suitable sexual partner only occur with some probability in any given month, so the impact on outcomes will occur gradually even if the change in attitudes or behaviors occurs immediately. We can also examine the time path leading up to the introduction of the show to determine if any systematic changes in birth rates were occurring prior to the show's introduction. If so, it may signal a spurious relationship between the show's initial release and behavioral outcomes.

The augmented regression model takes the following form:

$$\begin{aligned} Ln(B_{jt}) = \beta_0 &+ \alpha_1(RATE16P_j{*}preQ3_{jt}) + \alpha_2(RATE16P_j{*}preQ2_{jt}) + \alpha_3(RATE16P_j{*}preQ1_{jt}) \\ &+ \beta_1(RATE16P_j{*}postQ1_{jt}) + \beta_2(RATE16P_j{*}postQ2_{jt}) + \beta_3(RATE16P_j{*}postQ3_{jt}) \\ &+ \beta_4(RATE16PE_j{*}postQ4{+}_{jt}) + \beta_4 U_{jt} + \gamma_t + \delta_j + \varepsilon_j \end{aligned} \qquad [6]$$

The "pre" terms represent quarters prior to the introduction of the show. The "post" terms capture the differential change in the behavioral outcome as a function of show exposure for the first, second, third, or subsequent ("4+") quarters after initial airing on June 2009. We use an analogous instrumental variables procedure to estimate this model, interacting our single instrumental variable, *MTV0809*, with the relevant quarterly indicators.

VI. ECONOMETRIC RESULTS

A. Impact on Searches/Tweets

We begin reporting our econometric results in Table 1, which includes a high frequency analysis of the relationship between the release of a new episode and tweets/searches about the show, as expressed in Equation [1]. Using weekly data from Google Trends and daily data from Twitter, which are the highest frequency data available in both sources, we find that search and tweet activity spikes right at the moment that a new show is released. In Google, the search index for *16 and Pregnant* jumps 39 points (out of a maximum of 100) in the weeks in which a new episode is released. In Twitter, tweets rise 28 percent the day a new episode airs and 109 percent ($100*(e^{\beta} - 1)$ where $\beta = 0.249$ and 0.738) the following day. The magnitude of these results is not surprising based on the visual representation of these relationships in Figures 3 and 4, which display very large spikes in search activity and tweets right around the time that new episodes are released. It would be difficult to question the conclusion that individuals respond to this show online in quite meaningful ways. We interpret these responses as an indication that exposure to the show was quite strong.

Table 2 reports the results of a similar analysis focusing on the impact of the release of a new episode on searches and tweets about activities that would alter childbearing outcomes (see Equation [2]). The left side of the table reports results based on our analysis of Google searches. We focus on search terms related to contraceptive use ("how get birth control" and "how get birth control pill") and abortion demand ("how get abortion"), but we are unable to detect statistically significant effects here. We conduct a similar exercise using Twitter data in the right side of the table, focusing on tweets that mention birth control or abortion. Using these data, we see that tweets including the term, "abortion," rise by about 14 percent on the day a new episode is released and an additional 21 percent the following day. Tweets including the term "birth control" rise by 12 percent the day a new episode is released and 23 percent the following day.

In Table 3, we continue to focus on the same outcome measures as in Table 2, but we related them to the level of searches/tweets about *16 and Pregnant,* rather than indicators for when the show is on the air. The top panel of the table reports results using national, high frequency data, as represented in equation [3]. Indeed, when we modify our specification in this way, we see that increases in Google searches for *16 and Pregnant* are estimated to generate a statistically significant increase in searches for "how get birth control pills." Similarly, when we turn to Twitter data on the right side of the table, we see that the elasticity between tweets containing "birth control"/"abortion" and tweets containing *16 and Pregnant* is 0.077/0.064; both are statistically significant.

This table also presents the results from estimating equation [4], which exploits geographic variation in search activity. In our analysis of Google Trends data, we restrict our attention to the changes that take place before and after the introduction of the show in June of 2009 because of limited geographic data availability over relatively short periods. Geographic units are states and changes over time are measured over the intervals, 2005 through May 2009 and June 2009 through 2010. The longer interval in the pre-period is attributable to lower search volume over that period and the desire to obtain the largest feasible sample of states (recall that states with too low a search volume for a particular term are not reported in Google Trends). In our analysis of Twitter data, we focus on the period between January 2009 and December 2012, breaking up that larger period into 11 intervals in which the show was and was not on the air.

The results of this analysis further suggest that *16 and Pregnant* influenced individuals' searches and tweets about birth control and abortion. In states in which search activity for *16 and Pregnant* rose following the introduction of the show, searches related to getting birth control and abortion also rose.[29] Tweets including the term, "birth control," also rose by a statistically significant 13.7 percent in response to increased rates of Twitter activity regarding the show. Overall, the evidence reported here strongly suggests that exposure to *16 and Pregnant* altered searches and tweets by (presumably) teens regarding abortion and contraception, providing perhaps a glimpse into their thinking regarding these intermediary steps leading to giving birth.

B. Impact on Teen Birth Rates

The remainder of our analysis focuses on the relationship between exposure to *16 and Pregnant* and teen childbearing outcomes directly. The clear advantage that we have in this analysis is the virtual universe of birth outcomes available from the Vital Statistics system. The fact that county identifiers are also available in these data to construct births by DMA is also a huge advantage in that it enables us to merge these birth data with the Nielsen ratings data that we use to measure exposure. Again, we aggregate those ratings data from all the *16 and Pregnant* family of shows (including *Teen Mom* and *Teen Mom 2*) in the period after they began in June of 2009 through the end of 2012 to overcome excess noise in the data attributable to small samples within the 12 to 24 year old demographic group in each DMA at a point in time. We continue to use the label, *16 and Pregnant*, to facilitate the subsequent discussion.

We begin the presentation of these results in Table 4, which provides estimates of the impact on birth rates for women between the ages of 15 and 19 at the time of conception. The first stage results, not reported in this table, yield a coefficient on the earlier MTV ratings of 1.44 with a standard error of 0.21. The first column provides IV estimates as described by equations [5a] and [5b]. The results indicate that a one point increase in ratings for *16 and Pregnant* (about

[29] The more specific search regarding how to get birth control pills resulted in too little search activity at the state level to be useful for this analysis as most states had no reported index value in the earlier period.

50 percent of its average value) reduces the teen birth rate by 3.16 percent, which is statistically significant at any standard level.

Column 2 provides evidence that the estimated timing of an impact of the show on teen childbearing is consistent with a causal effect. It represents estimates obtained from the analogous IV version of equation [6], which distinguishes the impact of the show on teen childbearing by quarters before and after its introduction. In this model, we see a marginally significant impact (p-value = .097) of the show the quarter immediately preceding its introduction. This could be because the show started airing in June of 2009 and because dating of conceptions is not exact, so some of the treatment may have spilled over into the quarter preceding the third quarter of 2009. Nevertheless, the estimated impact of the show beginning in the 3rd quarter of 2009 (i.e. Q1Post) are mainly negative and significant. They are also fairly uniform for the remainder of the sample period. Although the estimate for the third quarter after the show was introduced is smaller and not significant, we are not surprised that there is some noise in the quarterly patterns of the estimated effect.[30] Our conclusion from these specifications is that they provide evidence that the show had a statistically significant, causal, negative impact on teen childbearing. Time patterns of these effects are sufficiently stable that it is appropriate to consider the effect to be reasonably uniform.

The remainder of Table 4 reports the results of the main reduced form specification, like in equation [5c], the analogous reduced form model distinguishing effects by time period relative to the introduction of the show (the reduced form version of equation [6]), and OLS specifications constraining the effects to be constant and allowing them to change over time (equations [5a] and [6]). The reduced form specifications provide results that are directly

---

[30] Tests of differences in these coefficients indicate that we cannot reject the hypothesis that all coefficients after the introduction of the show are equal.

comparable to those reported in Columns 1 and 2 using instrumental variables. The OLS results are more positive than those observed using instrumental variables, which is what we would predict if locations experiencing increasing relative rates of teen childbearing are also those where teens want to watch a show about it.

In Table 5, we report estimates from IV and reduced form models for different subpopulations, imposing uniformity in the estimated impact based on the aggregated results reported in Table 4. First we split teens into those who are 15 to 17 and those who are 18 or 19. We also consider births to somewhat older women, including those 20 to 24 and 25 to 29. The results for women in the first three of these age groups are comparable to the overall pattern for teens. Our preferred IV estimates are uniformly negative for all groups under age 25 with roughly similar magnitudes. We do not consider it surprising that those 20 to 24 respond as strongly as younger women. Nielsen data suggest that ratings among those who are between the ages of 12 and 17 are similar to those who are between the ages of 18 and 24. For those who are 25 to 29, however, we see smaller and statistically insignificant estimates, as we might expect.

In the final three columns of Table 5, we distinguish teens by their race and ethnicity. These results suggest that the impact of *16 and Pregnant* on teen fertility is greatest for black, non-Hispanic teens. The difference in estimated effects between black, non-Hispanic teens and other teens is statistically significant. The estimated effect for Hispanic teens is greater than the aggregate effect, but imprecision in that estimate makes it statistically indistinguishable from zero or from either of the other two groups. We are unable to find much of an effect among white teens. This pattern in the results is consistent with average ratings among those between ages 12 and 24 for all three shows in the *16 and Pregnant* family. National ratings data indicate that black, non-Hispanics are around 30 percent more likely to watch these shows than white, non-Hispanics or Hispanics.[31]

Taking these results as a whole, we conclude that *16 and Pregnant* had a sizable, causal impact on teen birth rates. We implemented an IV methodology designed to overcome the problem of endogenous viewership for *16 and Pregnant*, and our results indicate a negative and significant negative effect on teen births. We also find that the timing of the impact of the show coincided exactly with its introduction. Finally, we observe that the impact of the show was pronounced for those age groups who actually watched it in great numbers (up to age 24), but dropped off considerably for those in the 25-29 age group. All of these findings support a causal interpretation of our results.

The specification we prefer is in Table 4, Column 1, which reports the results from an IV model in which we estimate a constant impact of the show after it began.[32] This estimate suggests that a one point increase in Nielsen ratings for *16 and Pregnant* and its companion series reduced teen birth rates by 3.16 percent. To interpret the magnitude of this estimate, we note that ratings for these shows averaged 1.8 ratings points. This means that the shows would have contributed to a 3.16*1.8 = 5.7 percent reduction in teen births over our sample period after they began in the middle of June 2009. According to our data, between the 4th quarter of 2008 and the 4th quarter of 2010, the quarterly teen birth rate (defined by age of conception) fell from

[31]Ratings data by race/ethnicity for these shows were provided to us by Nielsen. We are unable to conduct a full-scale analysis by race and ethnicity, because sample sizes by DMA are too small to construct race/ethnicity-specific ratings estimates, as we have detailed earlier.

[32]That constant impact may be attributable to the fact that data limitations restricted us from allowing the ratings to vary after it started. The proper interpretation is that markets in which *16 and Pregnant* was generally more popular experienced a reduction in teen childbearing that did not change over time. It does not tell us that the impact was insensitive to changing market exposure once the show began.

14.62 to 12.05, representing a fall of 17.6 percent.[33] The predicted drop attributable to *16 and Pregnant* can explain 5.7/17.6 or 32.4 percent of this decline.

One thing that is clear based on the results presented in Tables 4 and 5 is the importance of labor market conditions in determining teen childbearing outcomes. The estimated effect suggests that when the labor market is weak, teens (like older women) respond by having fewer children. The effect is rather large as well; a one percentage point increase in the unemployment rate reduces the teen childbearing rate by about 2 percent. This means that the five point increase in the unemployment rate that the U.S. experienced in the Great Recession would generate a 10 percentage point reduction in teen childbearing. This finding is consistent with the results in Kearney and Levine (2012b), which addresses this issue in more detail. Overall, our findings suggest that the vast majority of the decline in teen childbearing since *16 and Pregnant* first aired in June of 2009 can be attributed to the show and to the weak labor market.

VII. DISCUSSION

Using data from numerous sources, we examined the impact that the MTV show, *16 and Pregnant*, has had on on-line search and Twitter activity and, ultimately, on rates of teen childbearing in the United States. Our results suggest the introduction of the show led young women to search and tweet about birth control and abortion, indicating that it had some influence on them in a way that could potentially change their behavior. We also find that exposure to the *16 and Pregnant* shows had a sizable impact on the rate at which teens give birth in the United States, generating a 5.7 percent reduction in teen births that would have been conceived between June 2009, when the show began, and the end of 2010. That can account for roughly one-third of the decline over that period. We do not have sufficient data to carefully evaluate the role that

[33]We use a Q4 to Q4 comparison because of the clear seasonality in teen birth rates. Because detrended teen birth rates probably fell somewhat between Q4 of 2008 and Q2 of 2009, our estimates of the impact of *16 and Pregnant* are likely to be slightly understated.

more frequent use of abortion played in contributing to this effect. We do know, however, that aggregate abortion rates for teens were also declining over this period (Pazol, et al., 2013), suggesting that a reduction in pregnancy is the likely mechanism.

The finding that *16 and Pregnant* had an impact suggests that MTV drew in teens who actually were at risk of teen childbearing and conveyed to them information that led them to change their behavior, preventing them from giving birth at such a young age. The fact that MTV knows how to make shows that teens like to watch, which speak to them in ways that resonate, presumably is critical to the show's impact. Apparently, this approach has the potential to yield large results with important social consequences. Typically, the public concern addresses potential negative influences of media exposure, but this study finds it may have positive influences as well. Presumably the effect on the attitudes or behaviors of teens and young adults could be positive or negative, depending on the specific media content and context. We find that media has the potential to be a powerful driver of social outcomes.

REFERENCES

Bricklin, Julia (2012). "Why We Need MTV's '16 And Pregnant'." *Forbes.* Available online at: http://www.forbes.com/sites/juliabricklin/2012/03/30/540/ (accessed 10/10/2013).

Adalian, Josef (2010). "It's a Hit! MTV's '16 and Pregnant' Delivers Huge Ratings." *The Wrap: Covering Hollywood*. February 18. Available at: http://www.thewrap.com/tv/column-post/its-hit-mtvs-16-and-pregnant-delivers-huge-ratings-14381 (accessed 10/10/2013).

Albert, Bill (2010). *With One Voice 2010: America's Adults and Teens Sound Off about Teen Pregnancy*. Washington, DC: The National Campaign to Prevent Teen and Unplanned Pregnancy.

Aldrich, Eric M, Peter S. Arcidiacono, and Jacob L. Vigdor, 2005. "Do People Value Racial Diversity? Evidence from Nielsen Ratings," *The B.E. Journal of Economic Analysis & Policy* 0(1): 4.

Baker, Scott and Andrey Fradkin (2013). "The Impact of Unemployment Insurance on Job Search: Evidence from Google Search Data." Stanford Working Paper.

Barbera, Paolo (2013). "Birds of the Same Feather Tweet Together: Bayesian Ideal Point Estimation Using Twitter Data." NYU Working Paper, available at: http://papers.ssrn.com/abstract=2108098.

Blakespoor, Elizabeth, Gregory S. Miller and Hal D. White (2012). "Dissemination, Direct-Access Information Technology and Information Asymmetry," Research Papers 2106, Stanford University, Graduate School of Business. Available at: http://papers.ssrn.com/abstract=1657169

Card, David and Gordon B. Dahl (2011). "Family Violence and Football: The Effect of Unexpected Emotional Cues on Violent Behavior." *Quarterly Journal of Economics* Volume 126, Issue 1Pp. 103-143.

Chandra, Anita, Steven C. Martino, Rebecca L. Collins, Marc N. Elliott, Sandra H. Berry, David E. Kanouse, and Angela Miu (2008). "Does Watching Sex on Television Predict Teen Pregnancy? Findings from a National Longitudinal Survey of Youth," *Pediatrics* 122(5): 1047-1054.

Chiou, Lesley, and Mary Lopez (2010). "The Reality of Reality Television: Does Reality TV Influence Local Crime Rates? *Economics Letters* 108(3): 330–333.

Choi, Hyunyoung and Paul Liu (2011). "Reading Tea Leaves in the Tourism Industry: a Case Study in the Gulf Oil Spill, 2011." Google Technical Report, http://google.com/googleblogs/pdfs/google_gulf_tourism_march2011.pdf

Choi, Hyunyoung, and Hal Varian (2009a). "Predicting the Present with Google Trends." Google Technical Report, http://google.com/googleblogs/pdfs/google_predicting_the_present.pdf

Choi, Hyunyoung and Hal Varian (2009b). "Predicting Initial Claims for Unemployment Benefits." Google Technical Report, http://research.google.com/archive/papers/initialclaimsUS.pdf

Colen, Cynthia G., Arline T. Geronimus, and Maureen G. Phipps (2006). "Getting a Piece of the Pie? The Economic Boom of the 1990s and Declining Teen Birth Rates in the United States." Social Science and Medicine, 63(6): 1531-1545.

Da, Zhi., Joseph Engelberg, and Pengji Gao (2011). "In Search of Attention" *The Journal of Finance.* 66(5): 1461–1499.

Dahl, Gordon and Stefano DellaVigna (2009). "Does Movie Violence Increase Violent Crime?" *Quarterly Journal of Economics* 124(2): 677-734.

Garthwaite, Craig, Tal Gross, and Matthew J. Notowidigdo (forthcoming). "Public Health Insurance, Labor Supply, and Employment Lock. *Quarterly Journal of Economics.*

Gerber, Alan S., James G. Gimbel, Donald P. Green, and Daron R. Shaw (2011). "How Large and Long-lasting Are the Persuasive Effects of Televised Campaign Ads? Results from a Randomized Field Experiment." *American Political Science Review*. 105(1): 135-150.

Gorman, Bill (2011). "2010-11 Season Broadcast Primetime Show 18-49 Ratings Averages." Available at: http://tvbythenumbers.zap2it.com/2011/06/01/2010-11-season-broadcast-primetime-show-18-49-ratings-averages/94402/ (accessed 10/10/2013).

Kanazawa, Mark. T. and Jonas P. Funk (2001). "Racial Discrimination in Professional Basketball: Evidence from Nielsen Ratings." *Economic Inquiry* 39(4): 599–608.

Kearney, Melissa S. and Phillip B. Levine (2012a). "Why is the Teen Birth Rate in the United States so High and Why Does it Matter?" *Journal of Economic Perspectives* 26(2): 141-166.

Kearney, Melissa S. and Phillip Levine (2012b). "Investigating Recent Trends in the U.S. Teen Birth Rates," NBER working paper 17964, March.

La Ferrara, Eliana, Alberto Chong, and Suzanne Duryea (2012). "Soap Operas and Fertility: Evidence from Brazil." *American Economic Journal: Applied Economics*. 4(4): 1-31.

Liu, Yizao, Huaxia Riu, and Andrew Whinston (2013). "Whose and What Chatter Matters? The Impact of Tweets on Movie Sales. *Decision Support Systems*. 55(4): 863–870

Lopoo, Leonard M. and Kerri M. Raissian (2012). "Policy Retrospective: Natalist Policies in the United States." *Journal of Policy Analysis and Management*. 31(4): 905-946.

Marketing Charts (2013). "Are Young People Watching Less TV? (Updated – Q2 2013 Data)." Available at: http://www.marketingcharts.com/wp/television/are-young-people-watching-less-tv-24817/, accessed 6/24/2013.

Myers, Caitlin Knowles (2008). "Discrimination as a Competitive Device: The Case of Local Television News." *The B.E. Journal of Economic Analysis & Policy* 8(1): 28.

MTV (2012). MTV's Groundbreaking Docu-Series "16 And Pregnant" Returns on Tuesday, March 27 At 10pm Et/Pt. Available at: http://mtvpress.com/press/release/mtvs_groundbreaking_docu_series_16_and_pregnant_returns_on_tuesday_march_27 (accessed 10/10/2013).

National Campaign to Prevent Teen and Unplanned Pregnancy (2010). "Evaluating the Impact of MTV's *16 and Pregnant* on Teen Viewers' Attitudes about Teen Pregnancy." *Science Says 45*.

National Campaign to Prevent Teen and Unplanned Pregnancy (2009). "Socio-Economic and Family Characteristics of Teen Childbearing." *Science Says 41*.

Pazol, Karen, Andreea A. Creanga, Kim D. Burley, Brenda Hayes, and Denise J. Jamieson (2013). "Abortion Surveillance – United States, 2010." *Morbidity and Mortality Weekly Report: Surveillance Summaries*. 62(08):1-44.

Price, Joseph and Gordon B. Dahl (2012). "Using Natural Experiments to Study the Impact of Media on the Family." *Family Relations*. 61(3): 363-373.

Saffer, Henry, Melanie Wakefield and Yvonne Terry-McElrath (2007). "The Effect of Nicotine Replacement Therapy Advertising on Youth Smoking." NBER Working Paper 12964.

Seidman, Robert (2010). "Season Premiere of MTV's *16 and Pregnant* Scores Series High 3.4 Million Viewers." Available at: http://tvbythenumbers.zap2it.com/2010/02/18/season-premiere-of-mtvs-16-and-pregnant-scores-series-high-3-4-million-viewers/42357/ (accessed 10/10/2013).

Seidman, Robert (2009). MTV Greenlights Season Two of Hit Series "16 & Pregnant". Available at: http://tvbythenumbers.zap2it.com/2009/06/30/mtv-greenlights-season-two-of-hit-series-16-pregnant/21672/ (accessed 10/10/2013).

Stein, Rob (2008). "Study First to Link TV Sex To Real Teen Pregnancies." *Washington Post*. Available at: http://www.washingtonpost.com/wp-dyn/content/story/2008/11/03/ST2008110300038.html (accessed 10/5/2013).

Stephens-Davidoff, Seth (2013a). "The Cost of Racial Animus on a Black Presidential Candidate: Using Google Search Data to Find What Surveys Miss." Harvard University, unpublished manuscript.

Stephens-Davidoff, Seth (2013b). “Unreported Victims of an Economic Downturn.” Harvard University, unpublished manuscript.

Tanner, Lindsey (2008). “Teen Pregnancies Tied to Tastes for Sexy TV Shows.” *USA Today*. Available at: http://usatoday30.usatoday.com/news/health/2008-11-03-458217694_x.htm (accessed 10/5/2013).

Vlastakis, Nikolaos & Raphael N. Markellos (2012). "Information Demand and Stock Market Volatility," *Journal of Banking & Finance* 36(6): 1808-1821.

Table 1: Impact of Episode Release on National Google Searches and Tweets about *16 and Pregnant*

	Google Trends Index Value, *16 and Pregnant*	ln(Tweet Rate), *16 and Pregnant*
Week New *16 and Pregnant* Episode Released	**38.99**	
	(1.95)	
Day New *16 and Pregnant* Episode Released		**0.249**
		(0.134)
Lagged Day New *16 and Pregnant* Episode Released		**0.738**
		(0.125)
Number of Observations	209 weeks	336 days

Notes: Google Trends data are weekly observations between January 4, 2009 and December 30, 2012. Twitter data are daily observations between January 1, 2009 and December 31, 2012, but just "in-season" days. Standard errors reported in parentheses. These models also include quadratic trends and, in Google Trends analysis, seasonal dummy variables. Twitter regressions are weighted by the total number of tweets made on each day.

Table 2: Impact of Episode Release on National Google Searches and Tweets about Sexual Activity, Contraception, and Abortion

	Google Trends: Search Index			Twitter: ln(Tweet Rate)	
	"How Get Birth Control"	"How Get Birth Control Pill"	"How Get Abortion"	"Birth Control"	"Abortion"
Week New *16 and Pregnant* Episode Released	0.813 (1.209)	3.134 (1.930)	-1.555 (1.742)	---	---
Day New *16 and Pregnant* Episode Released	---	---	---	**0.120** **(0.047)**	**0.142** **(0.036)**
Day New *16 and Pregnant* Episode Released - Lagged	---	---	---	**0.230** **(0.058)**	**0.212** **(0.046)**
Number of Weeks (Searches)/ Days (Tweets)	209	209	209	336	336

Notes: See notes to Table 1.

Table 3: Impact of Searches/Tweets about *16 and Pregnant* on Searches/Tweets about Sexual Activity, Birth Control, and Abortion

	Google Trends: Search Index			Twitter: ln(Tweet Rate)	
	“How Get Birth Control”	“How Get Birth Control Pill”	“How Get Abortion”	“Birth Control”	“Abortion”
			National, High Frequency Data		
Search Index/ln(Tweet Rate)	0.019	**0.093**	-0.054	**0.077**	**0.064**
16 and Pregnant	(0.025)	**(0.040)**	(0.036)	**(0.034)**	**(0.025)**
Number of Weeks (Searches)/ Days (Tweets)	209	209	209	336	336
			State-Level, Lower Frequency Data		
Search Index/ln(Tweet Rate)	**0.751**	**---**	**0.505**	**0.137**	-0.087
16 and Pregnant	**(0.127)**	**---**	**(0.239)**	**(0.054)**	(0.075)
Number of States/Periods	30	**---**	24	537	537

Notes: For national, high frequency specifications, see notes to Table 1. For state-level, lower frequency data, Google Trends data represent two periods, January 2005 through May 2009 and June 2009 through December 2010. These data are only available for states with enough searches for data to be regularly reported (see the text for a discussion of this issue). Twitter data represent 51 states and 11 time periods between January 2009 and December 2012 when *16 and Pregnant* was on and off the air. A small number of observations are dropped because no tweets were reported in that state and period (small states in early periods) and this is a log specification. Standard errors reported in parentheses. All regressions include state and period fixed effects and estimates are obtained from models weighted by the total number of tweets made in the state/period.

Table 4: Estimates of the Impact of *16 and Pregnant* Ratings on Teen Birth Rates

	IV		Reduced Form		OLS	
	(1)	(2)	(3)	(4)	(5)	(6)
Rating	**-3.160**		**-4.567**		-0.838	
	(0.951)		**(1.280)**		(0.523)	
Rating*Q3Pre		-0.974		-1.387		0.078
		(1.117)		(1.653)		(0.531)
Rating*Q2Pre		-0.179		-0.179		-0.507
		(1.138)		(1.667)		(0.758)
Rating*Q1Pre		-2.099		-2.989		0.039
		(1.259)		(1.761)		(0.632)
Rating*Q1Post		**-4.489**		**-6.487**		-1.249
		(1.373)		**(1.904)**		(0.727)
Rating*Q2Post		**-4.287**		**-6.192**		**-1.235**
		(1.189)		**(1.607)**		**(0.622)**
Rating*Q3Post		-1.366		-1.921		0.135
		(1.271)		(1.876)		(0.690)
Rating*Q4-6Post		**-3.434**		**-4.960**		-0.804
		(1.244)		**(1.689)**		(0.693)
Unemployment	**-2.232**	**-2.241**	**-2.090**	**-2.105**	**-2.074**	**-2.060**
Rate	**(0.363)**	**(0.374)**	**(0.365)**	**(0.371)**	**(0.361)**	**(0.362)**

Notes: The data used for this analysis represents quarterly birth rates by DMA for conceptions leading to live births between 2006 and 2010. The sample size in each model is 4099 (205 DMAs, 20 quarters, and one observation was dropped because there were no teen births). The dependent variable, the birth rate, is measured in natural logs. Coefficients and standard errors (reported in parentheses) are multiplied by 100. Each model also includes quarter and DMA fixed effects. As described in the text, the relevant first stage regression for Column 1 yields a coefficient on the earlier MTV ratings of 1.44 with a standard error of 0.21. Regressions are weighted by the relevant sample sizes for each outcome. Reported standard errors are clustered at the DMA level.