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

By Melissa S. Kearney and Phillip B. Levine*

This paper explores the impact of the introduction of the widely viewed MTV reality show 16 and Pregnant *on teen childbearing. Our main analysis relates geographic variation in changes in teen childbearing rates to viewership of the show. We implement an instrumental variables (IV) strategy using local area MTV ratings data from a pre-period to predict local area* 16 and Pregnant *ratings. The results imply that this show led to a 4.3 percent reduction in teen births. An examination of Google Trends and Twitter data suggest that the show led to increased interest in contraceptive use and abortion.* (*JEL* D12, J13, J16, L82, Z13)

It is a longstanding and open question how exposure to media images affects viewers' behavior. Policy advocates and cultural observers worry, in particular, about the impact of exposure to sexual and violent content on adolescent behaviors. In some circles, the idea that teenagers respond to media content is a foregone conclusion, but determining whether media images themselves cause teens to behave in certain ways is a difficult empirical task.[1]

This paper examines the impact of the MTV show, *16 and Pregnant*, on teen childbearing. 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. Could exposure to these media images of pregnant teens and very young new moms have had an impact on how teens behave, leading to changes in the likelihood of becoming teen mothers themselves?

[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.

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 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 US teen birth rate has declined dramatically over the past 20 years, falling from 61.8 births 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 (Kearney and Levine 2014a).

Teen motherhood is generally considered to be an important social problem; there is substantial policy interest in figuring out how to reduce its prevalence. In previous work (Kearney and Levine 2012a, 2014b), we have argued that teen motherhood is more appropriately considered a marker of a social problem rather than a direct cause. We maintain that view. However, children born to teen mothers tend to be immediately at an economic and social disadvantage in life. So while it is true that "solving" the issue of teen motherhood may not have large effects on poverty rates or other social problems for those women, it is also likely to be true that children are off to a better start in life if they are born to older mothers, especially if their mothers get married, acquire more education and earnings potential, and acquire the skills needed to be more effective parents during those additional years.

The question of what types of policies are effective at influencing teen birth rates is still very much an open one. When trying to explain the recent decline, policy observers have offered many competing explanations, often pointing to the presumed success of their favored program or policy, such as sex education programs, abstinence-only education programs, or improved access to contraception. But, previous academic research on the topic offers limited support for the effectiveness of any of these targeted policies.[2] This paper adds a new "policy" into the mix—the role of what might be considered a social media campaign in the guise of a very popular reality TV show depicting the challenges of teen pregnancy and motherhood.

The timing of the introduction of MTV's *16 and Pregnant* is such that it might have contributed 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.[3] 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

---

[2] See Kearney and Levine (2012b) and Lopoo and Raissian (2012) for discussions of this body of research.

[3] For example, see Patrik Jonsson, "A force behind the lower teen birthrate: MTV's '16 and Pregnant,'" *Christian Science Monitor*, December 21, 2010 (http://www.csmonitor.com/USA/Society/2010/1221/A-force-behind-the-lower-teen-birthrate-MTV-s-16-and-Pregnant and Bill Albert); "U.S. Teen Birth Rate Drops a Dramatic 9% in 2010," The National Campaign to Prevent Teen and Unplanned Pregnancy, November 17, 2011 (http://thenationalcampaign.org/blog/us-teen-birth-rate-drops-dramatic-9-2010).

show airs.[4] Interestingly, in a study for the National Campaign to Prevent Teen and Unplanned Pregnancy, Albert (2010, p. 33) 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."

Our main analysis exploits geographic variation in measures of the show's viewership (as captured by Nielsen ratings data) to investigate whether differential exposure to the show led to differential changes in teen birth rates. This analysis is conducted at the level of "designated market areas" (DMAs), which reflect collections of counties representing local television markets for which ratings data are available. We append to these data county-level Vital Statistics birth data aggregated to the DMA level for conceptions occurring between 2005 and 2010. To address the possibility that the appeal of a show about teen pregnancy might be correlated with teen birth rate trends, 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 that MTV ratings in the period before the show aired would be unrelated to subsequent trends in teen childbearing, except for the introduction of the *16 and Pregnant* content. We also introduce event study methods that look for changes in teen births well-timed to the introduction of the show to identify a causal effect.

The results of this analysis imply that the introduction of *16 and Pregnant* led teens to noticeably reduce the rate at which they give birth. Our estimates imply that this show led to a 4.3 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 24 percent of the total decline in teen births over that period.

We supplement this analysis with an exploration of data about online activity to determine whether the MTV show influenced viewers' interest in the topics of birth control and abortion, in so far as it is reflected in their Internet search and social media activity. In particular, we examine the potential influence of *16 and Pregnant* on teens' decision-making process with data on the frequency with which teens search for terms like "how to get birth control" or use the terms "birth control" or "abortion" in their tweets.[5] This analysis provides some insight into the mechanisms that would lead to changes in rates of teen childbearing in the absence of sufficient data on actual levels of sexual activity and contraceptive use.[6]

Specifically, in this analysis we first use high frequency, national data (daily and 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 national data

---

[4] http://www.cnn.com/2011/OPINION/05/04/henson.teen.mom.show/ (accessed June 18, 2013).

[5] We explored other terms attempting to identify interest in sexual activity and adoption, but were unable to identify well-defined terms in this regard. 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, but were about issues unrelated to the concepts we were interested in capturing. 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.

[6] Data sources that specifically include sexual activity and contraceptive use are available, but are not sufficient for use in this project, as we describe below.

from these sources to look for increases in relevant searches and tweets related to birth control and abortion co-timed with the spikes in searches/tweets on the show's title. Third, we take advantage of lower frequency data that varies at the state level to examine 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. Combining these empirical approaches enable us to take advantage of all of the potential sources of statistical identification available in these data.

The full set of analyses we have conducted using Google Search and Twitter data provides supportive evidence that the show had an impact on attitudes toward pregnancy prevention consistent with the show leading to a reduction in teen birth rates. Large spikes in search activity and tweets about the show are evident exactly at the time a new episode was released. In some specifications, 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 these types of searches/tweets when the show was on the air.

I. Previous Research on the Impact of Media Exposure on Social Outcomes

There are now a number of economics papers documenting causal effects of media exposure on social outcomes.[7] One important example is the work of Gentzkow and Shapiro (2008) that exploits variation in the timing of television's introduction to different local markets in the United States to identify the effect of preschool television exposure on standardized test scores during adolescence, finding small positive effects on test scores. Jensen and Oster (2009) investigate the effects of cable television introduction in a set of villages in India, finding that exposure to cable television led to more progressive views about the role of women, in addition to changes in other social outcomes. These studies are examples of papers that focus on exposure to television media in general. A limitation of this type of study is that it does not consider specific media content and thus one must speculate about the mechanism through which cable television is having its effect.

There are examples of studies that focus on specific media content—as opposed to more general media exposure—in political economy. For example, DellaVigna and Kaplan (2007) consider exposure to the politically right-leaning content of Fox News on voting behavior and Gerber et al. (2011) examine the impact of exposure to a specific partisan television political advertising campaign on voter preferences. There are also examples of quasi-experimental analyses looking at outcomes of violence. 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 "incapacitation effect."

Our paper is more closely related to a pair of papers that study the effect of exposure to soap opera content in Brazil. These papers take advantage of variation in timing of when the network that broadcast soap operas entered into different areas of the country. They find that the introduction of this network led to increased rates

---

[7] Price and Dahl (2012) provide a helpful summary of the set of studies focusing on the impact of the media on family outcomes, with an emphasis on describing methodological challenges and approaches.

of divorce (Chong and La Ferrara 2009) and lower levels of fertility (La Ferrara, Chong, and Duryea 2012). Since soap operas portrayed smaller families and family dissolution, there is a strong case to be made that it was exposure to this specific content that led to these changes in social outcomes.

Our paper makes a contribution to this literature by examining how exposure to a very specific type of content—namely, a reality television program that portrays the difficulties of being a new teen mother—affects rates of teen childbearing.[8] In doing so, we provide an additional piece of evidence that content can be a very powerful mechanism driving the effect of media on important social outcomes. Furthermore, 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.[9] As noted in the introduction, there is great policy interest in reducing rates of teen childbearing, with little convincing evidence about effective ways to do so. Yet, the role that the media can play in contributing to further declines in teen childbearing has not been previously addressed.[10]

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 to 7 month period in the life of a teenager as she navigates the bumpy terrain of adolescence, growing pains, rebellion, and 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).[11]

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

---

[8] We know of one other paper that specifically examines the effect of an MTV program–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.

[9] Chandra et al. (2008, p. 1052) found that teens who viewed more sexual content on television were more likely to become pregnant. 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."

[10] In a paper written concurrently with this paper, Trudeau (forthcoming) investigates how teen birth rates changed after the introduction of *16 and Pregnant*. That paper focuses on age differences, state sexual education mandates, and in one section of the paper, uses Nielsen ratings data from a limited set of DMAs. Relative to that paper, our analysis considers the universe of DMAs, uses an instrumental variables strategy, and examines Google Trends and Twitter data, among other differences.

[11] 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.

the population of new teen mothers. Across episodes, 68 percent (32 of 47) of teen moms are white, non-Hispanic. In the United States in 2010, 57 percent of teens are white, non-Hispanic (US Census Bureau 2014), but among teens who gave birth that year, only 39 percent are (Martin et al. 2012). 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 toward 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 girls commonly report that they did not think that they would have sex or become pregnant (36 of 47) and that they were ambivalent about getting pregnant (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 prior to the birth and three led to adoption. There were 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.[12] In addition, in 8 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. *Show Popularity and Nielsen Ratings Data*

Before undertaking an analysis of the effect of *16 and Pregnant* on teens' childbearing outcomes, it is important to document that teens' exposure to the show was substantial.[13] This is, of course, a necessary condition for the show to have had any sizable effect on aggregate birth rates. In addition, it must be the case that some of those who were exposed to the show were teens who might have become teen moms, in other words "marginal teen moms." To state the obvious, if the only girls who watched the show were ones who have a probability of zero of becoming a teen mother, or a probability of one of becoming a teen mother, then regardless of

---

[12] In the United States in 2007, the C-section delivery rate for women under age 20 was also 23 percent. For details, see Menacker and Hamilton (2010).

[13] The public controversy regarding the potential "glamorizing" effect of *16 and Pregnant* may include other forms of exposure outside of television exposure, like tabloid coverage. It is possible that our analysis could generate additional insights if we were able to measure exposure through these alternative forms of media.

how compelling the show's content was, the show would not lead to changes in teen birth rates. Ratings data do not allow us to parse out viewership by latent probability of becoming a teen mother, but we will establish that viewership was widespread across the country, and not singularly clustered in areas with either unusually high or low rates of teen childbearing.

The entertainment media, basing its assessments on Nielsen television 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 three 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).

For our analysis, we have obtained ratings data for the show under an agreement with the Nielsen Corporation and these constitute our primary measure of show viewership. Nielsen ratings data have been the gold standard in measuring exposure to television shows for decades.[14] The Nielsen 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.[15]

Geography is defined by DMAs; we have data on 205 DMAs in the continental United States.[16] A DMA is a collection of counties that is generally larger than a metropolitan area. It is based on the geographic region where a signal could be received from a broadcast television tower in a given market. A DMA has an average of 15 counties with 1.5 million people each.

---

[14] 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 makeup 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. 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.

[15] Demographic information is mainly 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).

[16] 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.

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.[17]

The Nielsen data available to us include ratings data during the "sweeps" periods (November, February, May, and July) in each DMA by age group—ages 12 to 17 and ages 18 to 24—and all ages, for each season of *16 and Pregnant*. The success of this show led MTV to also generate spin-offs, called *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.[18] We also were able to obtain comparable ratings data for these sequel shows as well. To ease exposition, we will refer to this collection of three shows simply as *16 and Pregnant*. In our subsequent analysis, we aggregate the data for 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. We do so to reduce the sampling variability that is generated by relatively small samples of young viewers in each DMA during a single sweeps period.[19] Because of the limited nature of the data, we are not able to take advantage of any time series variability in the popularity of the shows or of different episodes.

Ratings data by DMA reveal a considerable amount of geographic variation in show popularity. 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

---

[17] Between 2011:I and 2013:I, those between ages 18 and 24 reduced the hours spent watching traditional television from 26.5 hours per week to 23.4 hours per week (Marketing Charts 2015). Comparably small declines are also apparent among those between ages 12 and 17. This indicates that online viewing may be making some inroads into traditional TV viewing, but the traditional method is still overwhelmingly the main form of exposure to television content. 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 two million viewers, on average, who watched the show "live" (within one day of initial airing—MTV 2012).

[18] The exact dates that each of these shows aired during our sample window is as follows:

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.

[19] To be more specific, we have data from seven 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 seven 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.
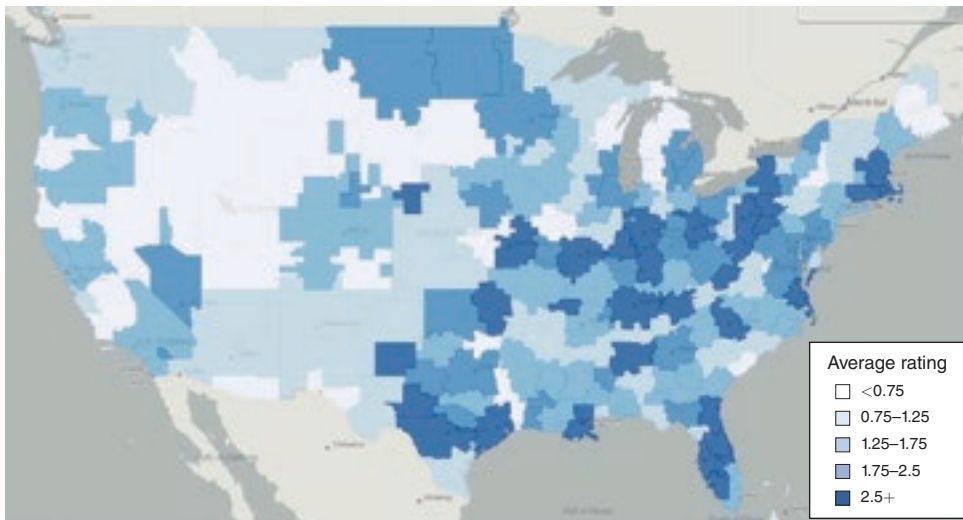
FIGURE 1. NIELSEN RATINGS FOR *16 AND PREGNANT/TEEN MOM/TEEN MOM 2* FOR THOSE AGED 12 TO 24,
BY DESIGNATED MARKET AREA

Source: Authors' calculations based on Nielsen data from 2009 through 2012.

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. These facts are not crucial to our identification strategy, but relevant to building an understanding of who watched the show and how concentrated viewership was. We would be concerned if the show only resonated with teens in one part of the country, for example, and less inclined to believe it could have had an effect on aggregate birth rates across the country.

Figure 2 displays the geographic variation in baseline teen childbearing rates by DMA to compare to the show's popularity.[20] Teen birth rates tend to be highest in DMAs in the South and Midwest, where the show is popular, as well as in the Southwest, where the show is not popular. Teen birth rates are moderate in DMAs in California, where the show is not popular. Teen birth rates are very low in New England, where the show is reasonably popular. Statistically, the population weighted correlation between Nielsen ratings data and teen birth rates is 0.16, which is positive, but small. This correlation is shown as a scatter plot in Figure 3, in which each data point is a DMA.

It is interesting, though not necessary to our identification strategy, that ratings are neither strongly positively correlated nor strongly negatively correlated with teen birth rates. If ratings covaried closely with teen birth rates, it might suggest selection into viewership based on the teen childbearing environment in the area. Our IV strategy is designed to overcome this type of selection, so such a relationship would not be insurmountable to the identification of a causal relationship.

---

[20] 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.
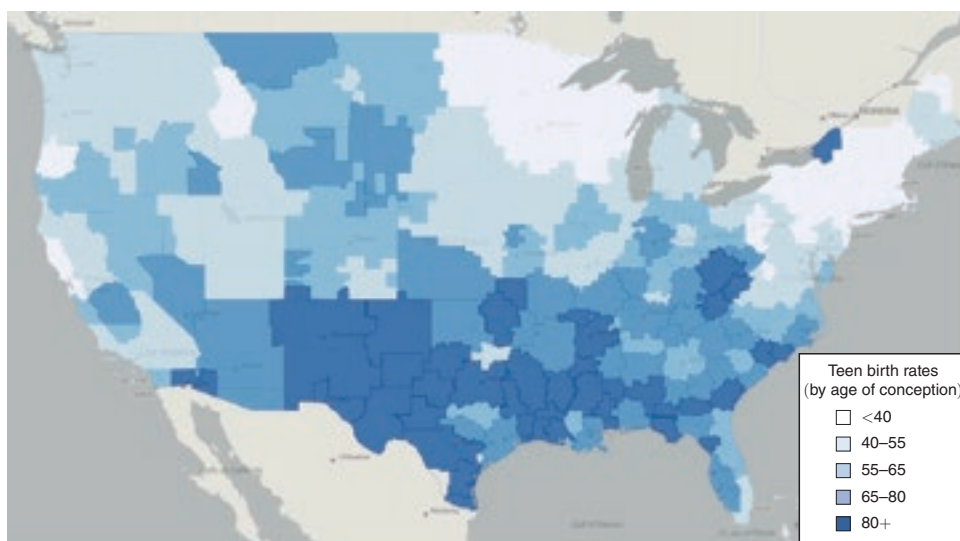
FIGURE 2. TEEN BIRTH RATES IN 2008 BY DESIGNATED MARKET AREA

Note: Teen birth rates by age of conception reflect the number of births conceived while the mother was between age 15 and 19 per 1,000 girls between ages 15 and 19.

Source: Authors' calculations based on Vital Statistics birth data.
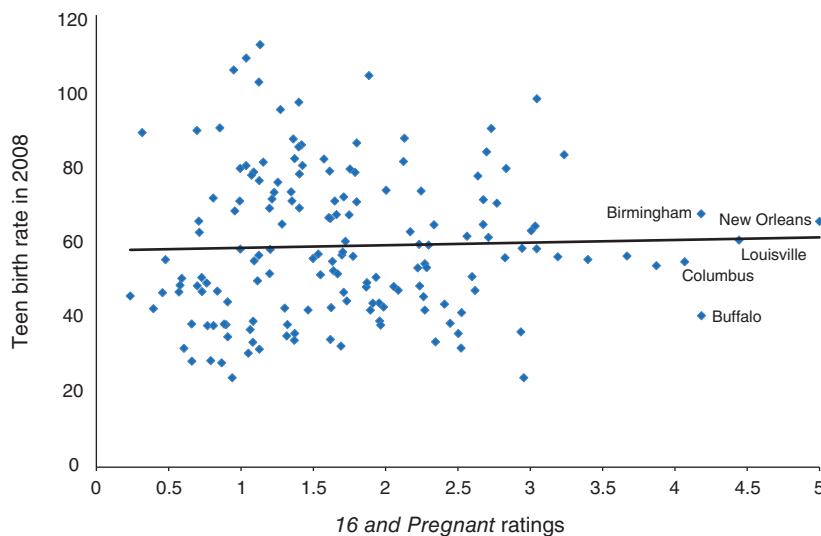
FIGURE 3. RELATIONSHIP BETWEEN *16 AND PREGNANT* RATINGS AND TEEN BIRTH RATES IN 2008

Perhaps more importantly, if viewership had a very strong negative correlation with teen birth rates, we might suspect that the only people tuning into the show were those teens who were not at risk of becoming teen parents, which would make it unlikely that the show had a sizable effect on aggregate teen birth rates. We offer these figures to make the relevant observation that teens across the country

tuned into the show (or not) essentially independently of teen birth rates where they lived.

III. Analysis of Teen Births

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. The key to our approach is to investigate whether there was a more rapid decline in rates of teen childbearing starting after the show was introduced in locations in which the show was more widely viewed.

A. *Data and Empirical Approach*

We use data on births between 2005 and 2011 from the Vital Statistics system to create a dataset of all births conceived between January 2005 and December 2010. Vital Statistics data contain birth certificate information, including length of gestation to determine approximate date of conception, for virtually every live birth in the United States. The main strength of these data is their 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 to aggregate births by occurrence in each DMA, the geographic level at which Nielsen ratings data are available.[21] We use a crosswalk procedure to aggregate counties into DMA units. Dating births by approximate month of conception enables us to better link the timing of exposure to the show to the timing of the activities that led to a subsequent birth.

The key to our empirical approach is to relate geographic variation in changes in teen birth rates to geographic variation in rates of exposure to the *16 and Pregnant* content, as captured by television ratings. The ordinary least squares (OLS) representation of this relationship is

$$(1) \quad \ln(B_{jt}) \;=\; \beta_0 \;+\; \beta_1 Rate16P_j \times post_t + \beta_2 U_{jy} + \mathbf{X}_{jy}\boldsymbol{\gamma} + \boldsymbol{\theta}_t + \boldsymbol{\delta}_{js} + \varepsilon_{jt},$$

where t indexes quarters, j indexes media markets (DMAs), and s indexes season of the year. The outcome variable of interest $\ln(B)$ represents the natural logarithm of the teen birth rate, defined as births to girls between ages 15 and 19 in DMA j conceived in quarter t scaled by the population of women age 15–19.[22] We include quarter fixed effects ($\boldsymbol{\theta}_t$) and DMA $\times$ season fixed effects ($\boldsymbol{\delta}_{js}$) to control for the seasonality that exists in teen birth rates (as documented by Buckles and Hungerman 2013). Our sample of births consists of those that were conceived between January

---

[21] Beginning in 2005, Vital Statistics Natality data with county identifiers require permission to use them, which we obtained through a nontransferable researcher agreement. Prior to 2005, county level identifiers exist in public use files, but births in smaller counties are omitted from the files. Because of this, we are unable to construct DMA level birth rates prior to 2005.

[22] The population data used for the denominator is only available annually, not quarterly.

2005 and December 2010, yielding 24 quarters of data. We use quarterly variation in births in our analysis because monthly data is noisy, particularly in smaller locations. Along with data from 205 DMAs, our full sample consists of 4,920 DMA by quarter observations.[23]

The explanatory variable of primary focus is the interaction of *Rate16P*, which represents Nielsen ratings for the show among those between ages 12 and 24, and *post*, which is an indicator variable for calendar quarters after June 2009 when the show began. The variable U_{jy} and the set of variables, $\mathbf{X}_{jy}$, control for the average unemployment rate and the percent non-Hispanic, black and percent Hispanic, respectively, in the DMA in the year (y) in which each quarter (t) falls. The unemployment rate is constructed by taking the population weighted average of annualized county unemployment rates; we do this because local monthly unemployment rates are notoriously noisy. This measure is included in the model as a proxy for economic conditions at the time of conception. Previous work has found that higher unemployment rates are associated with lower rates of teen childbearing (e.g., Colen, Geronimus, and Phipps 2006; Kearney and Levine 2012b). Race and ethnicity measures are constructed from county level population estimates available from the Census Bureau.

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.[24] 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.[25]

As described earlier, the ratings measure is an average taken 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. Again, data limitations preclude us from exploiting variation over time within a market in the show's popularity after its introduction.[26]

---

[23] 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.

[24] 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 many DMAs cross state lines.

[25] We also estimated analogous models without using sample weights and obtained qualitatively similar results. For instance, rather than our reported coefficient in column 1 of Table 1 of -2.368 (0.942), in the unweighted model we get -5.361 (1.342). Although the coefficient is larger, so is the standard error. This result is consistent with the use of population weights to reduce heteroskedasticity. We also experimented with two-way clustering on DMAs and quarters in estimating standard errors, but that had very little effect on our results.

[26] We experimented with splitting up the post period into two nine-month windows, calculating average ratings by DMA in each window, and examining whether the show had differential effects in the two windows. That

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 in terms of teen childbearing rates. To address this possibility (despite the low cross-sectional correlation between ratings and teen birth rates displayed in Figure 3) we account for time-invariant differences across markets by including 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, as 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 toward finding a positive effect of the show on teen birth rates.

To address this form of bias, we first 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 PM and 10:00 PM in the four 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 IV equation takes the form of the OLS equation as represented in equation (1) above, but the variable of interest is replaced by predicted ratings:

$$(2) \quad \ln(B_{jt}) = \beta_0 + \beta_1 \widehat{Rate16P_j} \times post_t + \beta_2 U_{jy} + \mathbf{X}_{jy} \gamma + \theta_t + \delta_{js} + \varepsilon_{jt},$$

where $\widehat{Rate16P_j} \times post_t$ is generated by the first stage regression in the IV framework:

$$(3) \qquad\qquad Rate16P_j \times post_t = \beta_0 + \beta_1 MTV0809_j \times post_t$$

$$+ \beta_2 U_{jy} + \mathbf{X}_{jy} \gamma + \theta_t + \delta_{js} + \varepsilon_{jt}.$$

The variable *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 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 2008–2009 in the following quarters.[28]

---

analysis did not yield economically or statistically meaningful differences between the two periods.

[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 toward teens/young adults, none of them share themes similar to those on *16 and Pregnant*.

[28] We estimated these equations without the time varying control variables (the unemployment rate and percent non-Hispanic black and percent Hispanic) and obtained similar results. We also estimated analogous regression models with these control variables on the left-hand sided and found no relationship between them and the *MTV0809* interaction variable.

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 preexisting 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.[29]

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

$$(4) \qquad \ln(B_{jt}) \ = \ \beta_0 \ + \ \beta_1 MTV0809_j \ \times \ post_t \ + \ \beta_2 U_{jy}$$

$$+ \ \mathbf{X}_{jy}\boldsymbol{\gamma} \ + \ \boldsymbol{\theta}_t \ + \ \boldsymbol{\delta}_{js} \ + \ \varepsilon_{jt}.$$

The explanatory time-invariant variable $MTV0809_j$ captures preexisting differences in the size of the 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 the relationship with teen birth rates, which could plausibly be attributed 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, conditional on the controls in the model, subsequent trends in teen childbearing would not be correlated with 2008–2009 MTV ratings except for a direct effect of the *16 and Pregnant* content. We also need the relationship between the earlier MTV ratings and the later *16 and Pregnant* ratings to be monotonic. Figure 4 presents a

---

[29] 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 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.
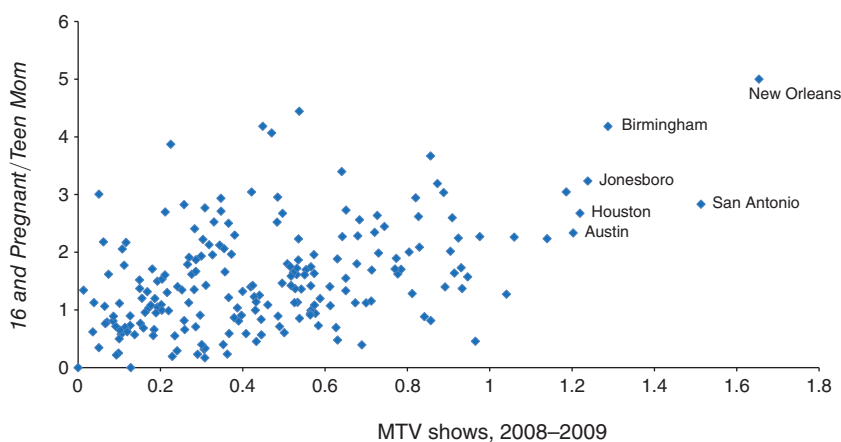
FIGURE 4. RELATIONSHIP BETWEEN MTV RATINGS IN 2008/2009 AND RATINGS FOR *16 AND PREGNANT/TEEN MOM*
(*Verification of IV Monotonicity Assumption*)

Notes: All data points reflect ratings for those between the ages of 12 and 24. Data from 2008–2009 reflect average ratings between July 2008 and May 2009 for all shows airing between 9:00 PM and 10:00 PM on weekdays. Data for *16 and Pregnant/Teen Mom* reflect average ratings over all seasons through May 2012 for all shows in the franchise (including *Teen Mom 2*).

Source: Authors' calculations based on Nielsen data from 2008 through 2012.

simplified bivariate version of the first stage relationship in the IV approach. It presents a simple scatter plot depicting the relationship between MTV ratings before *16 and Pregnant* started and the ratings for the *16 and Pregnant* show in the period after its introduction across DMAs. The pattern is clearly linear, demonstrating that this monotonicity requirement holds. 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 2008–2009. The correlation coefficient is around 0.5 weighted or unweighted. Given the strength of these relationships, it is not surprising that the data exhibit sufficient power in our first stage regression; the *F*-statistic on the omitted instrument is 48.1.

We augment this approach by implementing event-study methods designed to examine whether the time path of the estimated impact on birth rates changes discretely 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. The impact of the show on behavior might accumulate over time, but it should not start before the show began. If it did, it may signal a spurious relationship between the show's initial release and behavioral outcomes.

The augmented regression model takes the following form:

$$(5) \quad \ln(B_{jt}) = \beta_0 + \sum_{i=18}^{5} \alpha_i \big(rate16P_j \times preQ_i\big) + \sum_{i=1}^{6} \beta_i \big(rate16P_j \times postQ_i\big)$$

$$+ \phi U_{jy} + \mathbf{X}_{jy} \boldsymbol{\gamma} + \boldsymbol{\theta}_t + \boldsymbol{\delta}_{js} + \varepsilon_{jt}.$$

TABLE 1—ESTIMATES OF THE IMPACT OF *16 AND PREGNANT* RATINGS ON TEEN BIRTH RATES

	OLS (1)	First stage (2)	IV (3)	Reduced form (4)
Dependent variable:	ln(birth rate)	*16 and Pregnant* ratings	ln(birth rate)	ln(birth rate)
16 and Pregnant ratings	−1.020* (0.552)		−2.368** (0.942)	
MTV ratings 2008–2009		1.511*** (0.204)		−3.581** (1.512)
Unemployment rate	−1.440*** (0.401)	−0.001 (0.026)	−1.487*** (0.375)	−1.485*** (0.432)
F-statistic on omitted instrument		48.1		

Notes: The birth data used for this analysis represents quarterly birth rates by DMA for conceptions leading to live births between 2005 and 2010. The sample size in each model is 4,919 (205 DMAs, 24 quarters, and one observation was dropped because there were no teen births). Coefficients and standard errors (reported in parentheses) in birth rate regressions are multiplied by 100. Each model also includes the percentage of a DMA's female teen that is Hispanic and non-Hispanic black along with quarter and DMA × season fixed effects. Regressions are weighted by the relevant sample sizes for each outcome. Reported standard errors are clustered at the DMA level.

The first summation term counts from the first quarter of 2005 (18 quarters before *16 and Pregnant* began) to the second quarter of 2008 (five quarters beforehand). The second summation term counts from the first quarter after the show started to six quarters afterwards (2009:III through 2010:IV). The respective coefficients map out the quarterly pattern in the birth rates in response to *16 and Pregnant* ratings. Other elements in this equation include the DMA level unemployment rate, quarterly fixed effects, and DMA × season fixed effects. This last set of fixed effects require four base period quarters, which we set as 2008:III through 2009:II, representing the full year before the show began. This specification enables us to treat the values of β as the effect of the show relative to a base year just prior to its introduction. We use an analogous instrumental variables procedure (described earlier) to estimate this model, interacting our single instrumental variable, *MTV0809*, with the relevant quarterly indicators.

B. *Results*

The baseline results of our formal econometric analysis are presented in Table 1; the first column provides OLS estimates as described by equation (1). In that specification, we find that an additional ratings point for *16 and Pregnant* reduced the teen birth rate by around 1 percent (*p*-value = 6.4 percent). As we described earlier, however, OLS estimates are likely to be biased upward (toward a less negative effect). To address this, we estimate an IV model; the first-stage regression results are reported in column 2 and the IV estimates are presented in column 3. The first stage estimates, represented by equation (3), show a strong relationship between *16 and Pregnant* ratings and our instrument, MTV ratings in 2008–2009, the year before the show began. The IV estimates from equation (2) indicate that a one point increase in ratings for *16 and Pregnant* (about 50 percent of its average value) reduces the teen birth rate by 2.37 percent, which is statistically significant at any standard level. Column 4 provides reduced form estimates represented by equation (4).
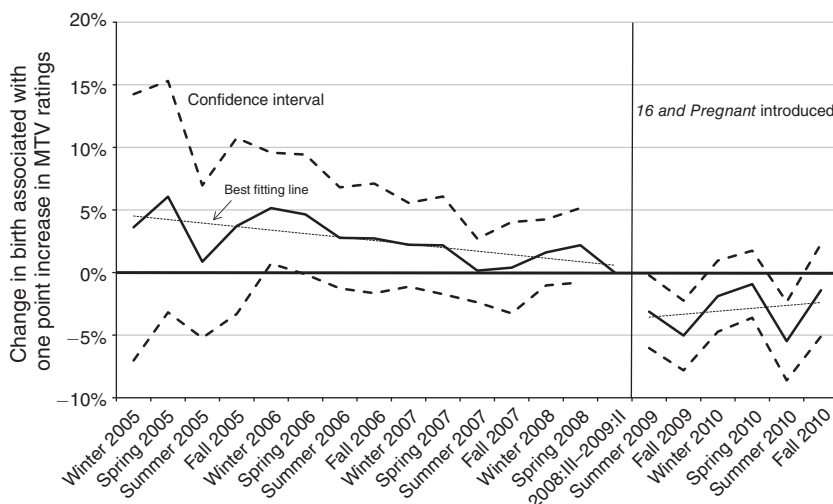
FIGURE 5. REDUCED-FORM EVENT STUDY ESTIMATES OF THE IMPACT OF *16 AND PREGNANT*

Notes: Estimates reflect coefficients on 2008–2009 MTV ratings interacted with quarter from a regression model controlling for DMA × season fixed effects. Dashed lines reflect 95 percent confidence intervals. Dotted lines reflect the best fitting lines through the pre- and post-period coefficients.

We continue with a reduced form approach to visually represent an event study analysis of the effects of *16 and Pregnant*. In Figure 5, we present a simplified, reduced form version of equation (5), where quarterly indicator variables are interacted with 2008–2009 MTV ratings (the instrument in our IV model) in a model that only includes quarter and DMA × season fixed effects to absorb much of the residual variation. Figure 5 simply plots these quarterly interaction coefficient estimates along with their confidence intervals.[30] For the period before the show began, coefficients are not statistically significant (with one exception among 14 coefficients). Jointly, the group is not significantly different from zero (p-value $= 0.21$). After the show began, all the coefficients turn negative and two of the six are statistically significant. A test of joint significance of those six coefficients rejects that they are all equal to zero (p-value $= 0.002$).

Still, visual examination of the pre-period suggests a downward slope that goes against a causal interpretation, despite the fact that the F-tests indicate that the higher values during that period are insignificant. To take this visual analysis one step further, we simply introduce best fitting lines through the pre- and post-period coefficients, which are also shown in the figure. These lines make it clear that there is, indeed, a downward slope before *16 and Pregnant* began, but there is still a sizable downward shift in point estimates precisely at the point when the show was

---

[30] In a model with DMA × season fixed effects, the base period of omitted interactions totals four, one for a particular DMA in each of the four seasons. This means that the base period is one year long, not one quarter long. We have chosen to use July 2008 through June 2009 as the base period to make it easier to see the "treatment effect," which occurs once the show starts in June 2009.
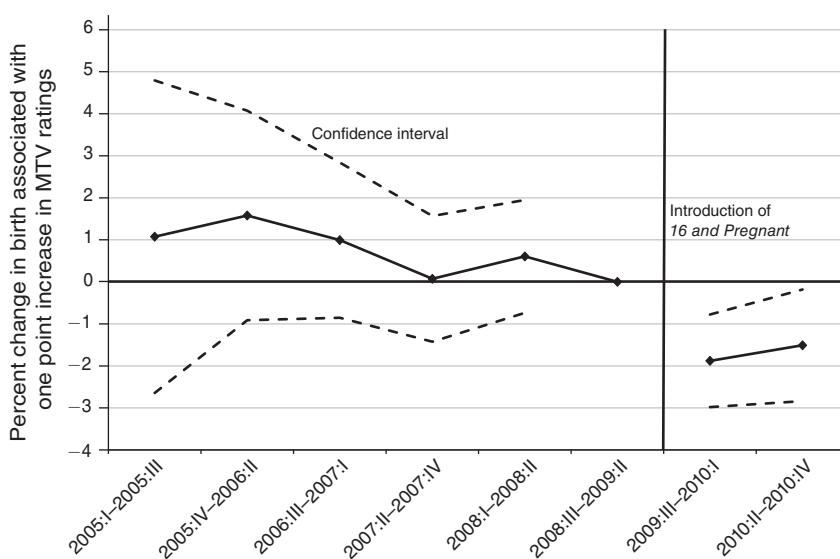
FIGURE 6. IV EVENT STUDY ESTIMATES OF THE IMPACT OF *16 AND PREGNANT*

Notes: Estimates reflect coefficients on *16 and Pregnant* ratings from an IV model where 2008–2009 MTV ratings are the instruments. Other control variables include labor market conditions, age and racial composition, and DMA × season fixed effects. Dotted lines reflect 95 percent confidence intervals.

introduced. We believe this provides strong visual support for the notion of a causal effect.[31]

We extend this event study exercise and in Figure 6, we report the results of our fully specified IV model based on equation (5). Here we report IV estimates of the impact of *16 and Pregnant* in a model including all covariates and aggregating quarters into groups of three to provide greater power. These estimates also strongly suggest a causal interpretation of the show's impact, with a more solid econometric underpinning. In the period before *16 and Pregnant* began, none of the quarterly groups were significantly different from the base period (the year before the show began) and these effects are not jointly significantly different from zero. What appears to be a slow, steady decline in point estimates is much too small to forecast the decline in the coefficients once the show began. From there, the data indicate negative and statistically significant effects on the teen birth rate associated with a one point increase in teen births. Point estimates indicate that a one unit increase in ratings points reduced the teen birth rate by -1.88 percent in the following three quarter period. Once the show began, its estimated impact appears to be roughly constant. Appendix Table B1 provides the exact results from this specification, along with results from OLS and reduced form specifications.

---

[31] This conclusion strongly relies on the assumption of linearity in the response of teen births to MTV ratings.

TABLE 2—ESTIMATES OF THE IMPACT OF *16 AND PREGNANT* RATINGS ON BIRTH RATES, BY DEMOGRAPHIC GROUP

	Age 15–19	Age 20–24	Age 25–29	Age 30–34	White, non-Hispanic Age 15–19	Black, non-Hispanic Age 15–19	Hispanic Age 15–19
Panel A. IV							
Rating	−2.368**	−2.422**	−1.606**	−0.328	−2.413**	−0.140	−3.782
	(0.942)	(0.993)	(0.817)	(0.689)	(1.075)	(1.380)	(2.063)
Unemployment rate	−1.487***	−1.810***	−0.800***	−0.910***	−1.399***	−0.243	−2.512***
	(0.375)	(0.308)	(0.229)	(0.146)	(0.329)	(1.049)	(0.784)
Panel B. Reduced form							
Rating	−3.581**	−3.498***	−2.344**	−0.455	−3.370**	−0.194	−6.510
	(1.517)	(1.312)	(1.151)	(1.016)	(1.581)	(2.109)	(4.432)
Unemployment rate	−1.485***	−1.726***	−0.721***	−0.897***	−1.354***	−0.240	−2.418***
	(0.432)	(0.325)	(0.252)	(0.165)	(0.375)	(1.170)	(0.944)
Sample size	4,919	4,920	4,920	4,918	4,911	4,513	4,750

Notes: This data used for this analysis represents quarterly birth rates by DMA for conceptions leading to live births between 2005 and 2010. The sample size in each model starts at 4,920 (205 DMAs, 24 quarters, and then drops DMA/quarter cells in which 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 × season fixed effects and age-specific models include the percentage of the relevant population that is Hispanic and non-Hispanic black. Regressions are weighted by the relevant sample sizes for each outcome. Reported standard errors are clustered at the DMA level.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

In Table 2, we report estimates from IV and reduced form models for different subpopulations. Because we found no evidence of differential effects by quarter after show introduction, we focus on equations (2) and (4), which have just one coefficient estimated for the effect of the show on birth rates. First we consider births to somewhat older women, including those 20 to 24, 25 to 29, and 30 to 34. The results for women between 20 and 24 are comparable to the estimated effect for teens; we do not consider this surprising since ratings for those who are between the ages of 12 and 17 are similar to those who are between the ages of 18 and 24. The point estimate of the effect declines for the older age groups and is statistically insignificant for those who are 30 to 34, as we might expect.[32]

In the final three columns of Table 2, we distinguish teens by their race and ethnicity. The point estimate of the effect for an additional one point in ratings is −2.413 (standard error of 1.075) for whites; −0.140 (standard error of 1.380) for blacks; and −3.782 (standard error of 2.063) for Hispanics. None of these are statistically different from one another. In thinking just about the point estimates, however, it is perhaps not too surprising that the effect on teen births to white, non-Hispanic mothers is estimated to be relatively large. As described earlier, mothers on *16 and Pregnant* were disproportionately white, non-Hispanic, so perhaps there was a greater role model effect for them. This conclusion, though, should be considered more as conjecture since the differences in results across race/ethnicity are not

---

[32] Tests of statistical significance for differences in pairwise comparisons of coefficients between the two younger age groups and the two older age groups have *p*-values between 0.03 and 0.12.

statistically significant and we are unable to adequately test the race/ethnic-specific role model hypothesis.[33]

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 those results indicate a negative and significant 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 greater numbers (up to age 24), but dropped off at older ages. All of these findings support a causal interpretation of our results.

To interpret the magnitude of our results, we focus on the estimate from Table 1, column 1, which reports the results from an IV model in which we treat the impact of the show as constant after it began.[34] Again, this estimate suggests that a one point increase in Nielsen ratings for *16 and Pregnant* and its companion series reduced teen birth rates by about 2.37 percent. We also note that ratings for these shows averaged 1.8 ratings points. This means that the shows would have contributed to a $2.37 \times 1.8 = 4.27$ percent reduction in teen births over our sample period after they began airing in the middle of June 2009. According to our data, between the fourth quarter of 2008 and the fourth quarter of 2010, the quarterly teen birth rate (defined by age of conception) fell from 14.62 to 12.05, representing a fall of 17.6 percent.[35] The predicted drop attributable to *16 and Pregnant* can explain 4.27/17.6 or 24.3 percent of this decline.

The coefficient estimates presented in Table 1 also reveal the importance of labor market conditions in determining teen childbearing outcomes. The estimated effect suggests that when the labor market is weak, teens respond by having fewer children. In other words, teen births—like non-teen births, as documented in numerous previous studies—appear to be procyclical. The magnitude of the effect is rather large: a 1 percentage point increase in the unemployment rate leads to about a 1.5 percent reduction in the teen childbearing rate.[36] This means that the five point increase in the unemployment rate that the United States experienced in the Great Recession would

---

[33] We attempted to examine this hypothesis by exploring responses based on the timing of appearances of mothers by race/ethnicity. We distinguish mothers by whether they are white or non-white; more precise racial/ethnic breakdowns are not readily determined by visual inspection and the frequency of what appears to be mixed race/ethnicity. Although each season of *16 and Pregnant* included both white and non-white mothers, *Teen Mom*, and *Teen Mom 2* only included white, non-Hispanic mothers. We broke up the post period into those in which these shows were on the air and examined whether the birth effect for white, non-Hispanic teens was larger during those quarters. The data do not suggest such a pattern. We also found no evidence that teen birth rates for black, non-Hispanics or Hispanics fell by more during the periods in which non-white mothers were on the show. Whether we could reasonably expect the timing of these effects to be as precise as that embedded in this exercise, though, is unclear and that limits its value.

[34] Data limitations restricted us from estimating models with time variation in ratings. The proper interpretation of this set of results is that markets in which *16 and Pregnant* was generally more popular experienced a reduction in teen childbearing that did not noticeably change over time. It does not tell us that the impact was insensitive to changing market exposure once the show began.

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

[36] The magnitude of this effect is larger than that estimated in Kearney and Levine (2012b), which found that a 1 percent increase in the unemployment rate reduces teen births by 0.6 percent. That analysis used data from 1980 through 2010, though, whereas the present study focuses on cyclical variation from the recent Great Recession. It is conceivable that the severity of the recent recession led to changes in behavior that were not just linear extrapolations of that observed in past, less severe, recessions.

generate a 7.5 percent reduction in teen childbearing. This amounts to 42.6 percent of the overall decline, which is a larger effect than that which we attribute to *16 and Pregnant*. The fact that a TV show can have an impact that even approaches the magnitude of the effect of a major swing in economic conditions is a testament to the practical significance of its impact. Overall, our findings suggest that much of the dramatic decline in teen childbearing since *16 and Pregnant* first aired in June of 2009 can be attributed to the MTV show and to the weak labor market.

IV. Exploring Mechanisms Using Google Search and Twitter Activity

The results presented above indicate that teenagers responded to the content on *16 and Pregnant* by reducing the rate at which they gave birth. Presumably this result is driven by the show's content influencing the way teens think about teen motherhood—namely, showing in a compelling and accessible way that being a teen mom (and dad) is challenging. The link must be that this change in attitude toward teen parenthood led to an increased commitment to avoid a pregnancy and birth. Ideally we would have data on teenage attitudes over time, preferably by DMA, combined with data on rates of unprotected sex over time, also preferably by DMA. Data on sexual activity and contraceptive use from available data sources, however, are not available at the DMA level and sample sizes are sufficiently small that they are not conducive to conducting an analysis comparable to what we have reported using Vital Statistics birth data.[37] Instead, we look to data from Google Trends and Twitter to measure changing levels of interest in birth control and abortion, as captured by online search and social media activity. These data also are not available at the DMA level, forcing us to adopt alternative identification strategies that take advantage of the massive amounts of data available from these sources. This section begins by describing these data sources, and then detailing these methodological approaches before presenting our results.

A. *Data from Google Trends and Twitter*

Data from Google Trends provides indexed values of the relative frequency with which individuals search Google for a particular term over time in a particular location. Google is able to identify geographical information for search activity through the user's IP address. The index is created by determining the number of searches for a particular search term as a share of the total number of searches conducted in each time and place. Google Trends then assigns an index value of 100 for the time/place in which this "relative search rate" is maximized. Other index values are set by taking the ratio of the search rate in a particular time/place to the maximum search rate. In other words, a time/place where the relative search rate is half the maximum value would be assigned an index value of 50. Because of the massive amount of searches conducted, Google creates this index by sampling from its underlying database. Users can request an analysis of search patterns over any time

---

[37] Data limitations also 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.

period between 2004 and the present and across countries, states, and cities. If the search volume resulting from the sample in a particular time and place is too low, Google Trends will report an index value equal to zero, which is the equivalent of a missing value.[38]

Google's approach in constructing this index generates results that are strictly ordinal; no cardinal interpretation can be applied to the values. Locations/places with higher index values are those where searches for a particular term relative to total searches are higher. This could occur because of more absolute searches for that term or fewer other searches. The process of assigning a value of 100 to the maximum level of relative searches also hinders cardinal interpretations. Two separate searches for different terms (i.e., "how get birth control" and "how get abortion") in a place/time with values of 100 almost certainly do not have equal relative search rates. They are just the highest relative search rates. All of this discussion indicates that any analysis of Google Trends data can only focus on the direction of any effect. Specific values cannot be compared or specifically interpreted.

Despite these limitations, Google Trends search data has now been used by a number of previous authors, mostly in studies focused on forecasting economic activity and by studies in the field of finance,[39] but also by analysts addressing social issues like we do in this paper. For instance, Stephens-Davidowitz (2013a,b) uses Google Trends data to explore the impact of racism on votes for Barak Obama (comparing changes in election results between 2004 and 2008 as a function of the rate of Google searches for racial epithets by state) and to better track child maltreatment over the course of the business cycle (when searches for terms like "child abuse" or "child neglect" may rise by those suspicious of this form of activity).

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

---

[38] 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. States included in this analysis are CA, FL, GA, IL, MA, MI, NJ, NY, OH, PA, TX, and VA for "how get abortion" and the same states along with MO, NC, and WA for "how get birth control."

[39] Choi and Varian (2009a and 2009b) showed that Google Trends could help predict economic indexes, including predictions of tourism activity and unemployment insurance claims. Choi and Liu (2011) cite nearly 20 papers that have used Google Search data to study outcomes in a variety of market contexts: capital markets, entertainment markets, labor markets, real estate markets, and healthcare markets. Google Trends data have also been used in a number of finance studies as a measure of investor attention and search activity, for example, in Da, Engelberg, and Gao (2011); and Vlastakis and Markellos (2012). It has also been used as a measure of job search activity by Baker and Fradkin (2015) and Garthwaite, Gross, and Notowidigdo (2014).

variation in their frequency. We used Topsy Labs.[40] 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.

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 percent coverage by country, 80 percent accuracy by state/province, 40 percent accuracy by city, at a 90–95 percent confidence level."[41]

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.[42] 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 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.[43]

---

[40] 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.

[41] This statement was provided in a personal communication between us and Topsy in an e-mail 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.

[42] 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.

[43] Few examples exist of papers that use Twitter data in an economic 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. Barberá 2015). Another set of papers in the field of finance are also available that examine the relationship between Twitter data and financial activity (for example, Blankespoor, Miller, and White 2013). Rui, Lui, and Whinston (2013) find that positive tweets increase movie sales and negative tweets

B. *Empirical Approaches*

The nature of the Google and Twitter data requires us to implement alternative approaches in this analysis relative to those which we reported earlier using Vital Statistics birth data. It seems the best way to take advantage of the strength of these data, which rests in their sheer volume, is to focus on the time series nature and the frequency with which outcomes can be measured. We can establish causal relationships if the exact timing of an outcome matches perfectly with the introduction of or exposure to the show. For instance, if searches/tweets regarding *16 and Pregnant* spike in the next 24 hours after the show is aired and then reverts back to some lower rate immediately afterwards, we are inclined to conclude that the show itself is responsible for that spike. This analysis is usefully presented graphically.

In addressing other outcomes, like searches looking for information on birth control or tweets that include mentions of birth control, a similar strategy may be applied, although regression-based methods appear more appropriate. We can certainly consider, for instance, daily trends in the number of tweets mentioning birth control and look for spikes on the days when the show was on the air. That may be a more difficult task, though, because many tweets are likely to mention birth control regardless of the show. Observing a spike, therefore, would be rather unlikely; an upward blip from what is otherwise a large and noisy number of tweets containing that term is a more plausible outcome. Graphical approaches, therefore, would be unlikely to yield relevant insights.

Instead, we estimate regression models that are designed to capture this form of identification. Regressions have the property of averaging out some of the statistical noise so they may be more conducive to identifying correlated spikes when the spikes themselves are not sufficiently large to clearly reveal themselves in a figure. We estimate the following OLS regression model:

$$(6) \qquad searchBC_t = \beta_0 + \beta_1 NewRelease_t + \mathbf{X}_t \gamma + \varepsilon_t,$$

where *searchBC* is shorthand for a search or tweet for a term like "birth control," *NewRelease* indicates that period in which a new episode is released, and *X* represents quadratic trends and, when we use weekly data from Google Trends, month fixed effects as described below.

Because of the way the data are tracked, we use different units of time depending on whether we are looking at Google Trends data or Twitter data. When we use Google Trends 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. Because of the potential for seasonal variation in this approach, we also include seasonal fixed effects when we use these data. When we use Twitter data, we restrict our attention just to those weeks in which the show is "in season" (which we listed previously) and take advantage of daily variation in outcomes. In our analysis of these data, we examine the response

---

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.

the day the show aired as well as the following day to approximate the 24 hour period following the airing of a new episode. This would modestly alter equation (6) to take the following form:

$$(7) \quad searchBC_t = \beta_0 + \beta_1 NewRelease_t + \beta_2 NewRelease_{t-1} + \mathbf{X}_t\gamma + \varepsilon_t.$$

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

$$(8) \qquad\qquad searchBC_t = \beta_0 + \beta_1 search16P_t + \mathbf{X}_t\gamma + \varepsilon_t.$$

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. We estimate these models using data in the same units and time frame as we indicated for equation (6).

In all of these approaches using high frequency data, we believe that the results plausibly provide causal estimates of the impact of the show. For the results to be spurious, there would need to be some other nationally trending variable whose time series pattern is correlated with the specific day or week that *16 and Pregnant* aired over the course of a few years. Although there may be something that satisfies this, we are at a loss to imagine what that might be.

In one final identification approach, we 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 above captures the methodology we introduce here, as represented by equation (9):

$$(9) \qquad\qquad searchBC_{jt} = \beta_0 + \beta_1 search16P_{jt} + \beta_2 U_{jt}$$

$$+ \gamma_j + \delta_t + \varepsilon_{jt}.$$

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.

This approach is the most closely associated with our earlier analysis of Vital Statistics birth data, with some important exceptions. First, we cannot estimate this model using quarterly data in the period before and after the show began. Google Trends data are too thin to do so; instead we aggregate over the January 2005–May 2009 and June 2009–December 2010 periods in all states with available data in both

periods in our analysis. 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). Twitter was not used extensively until around the time when *16 and Pregnant* was introduced. In those data, we use state variation 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. Second, the key explanatory variable is not the level of Nielsen ratings; we are unable to include those ratings in this analysis because Google Trends and Twitter data are not available at the county or DMA level. Third, for similar reasons we are also unable to estimate instrumental variables models using MTV ratings before the show began. Despite these limitations, we believe that this analysis can provide supportive, if perhaps not dispositive, evidence.[44]

C. *Results*

We begin by presenting graphical evidence indicating that young people clearly increased their Internet activities related to the show exactly at the times when new shows aired. Figure 7 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.

We next plot daily trends in searches and spikes between November 1, 2010 and December 21, 2010, as shown in Figure 8. We chose a period this short because Google Trends daily rate data are only available when the search is restricted to a relatively short period of time. We chose these particular days because they are 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.

---

[44] We have experimented with aggregating Nielsen ratings to the state level by assuming that ratings across counties within a DMA are constant and applying county population estimates. We then combined these state level ratings data along with Google Trends data on search activity at the state level for specific terms. Note that this is not possible using Twitter data because tweeting was only common during the "post" period. Using these Google Trends data, we estimate models comparable to those reported previously except using ratings as the key explanatory variable in both OLS and IV specifications. From this analysis we learned: (i) Google searches for *16 and Pregnant* and Nielsen ratings are positively correlated; (ii) if we replace *16 and Pregnant* Google search index values with *16 and Pregnant* ratings, we also find statistically significant effects of ratings on our other Google search terms (how to get birth control or how to get an abortion) in comparable OLS specifications; and (iii) these data are not sufficient to estimate IV models based on the subset of states for which we have relevant search data because first stage regressions have very limited power in this context. We also tried using search volume for MTV shows popular in 2008–2009 (including "The Hills," "Real World," and "The Challenge") as alternative instruments, but it became apparent that only searches for "The Hills" can be identified from Google Trends as focusing specifically on the TV show. Search data for this one show was insufficient to provide sufficient identification for the IV approach.
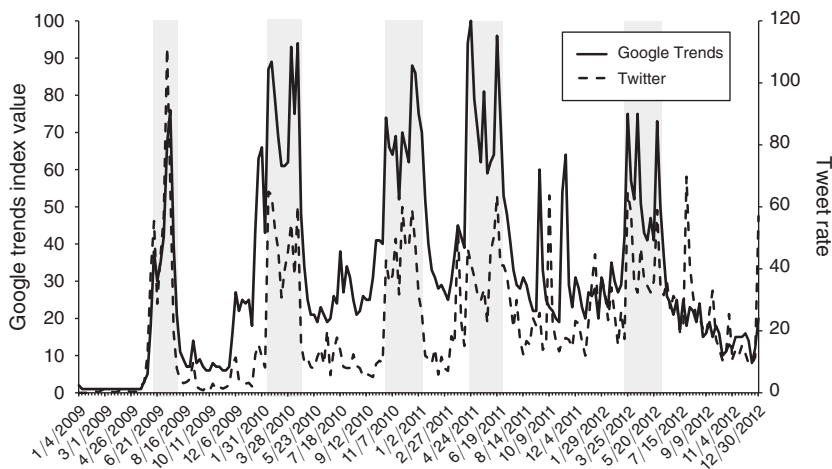
FIGURE 7. WEEKLY GOOGLE SEARCHES AND TWEETS ABOUT *16 AND PREGNANT*

Note: Shaded areas reflect weeks that new episodes of *16 and Pregnant* were broadcast.

Sources: Search index values are available directly from Google Trends. Tweet rates are computed based on the frequency of tweets obtained from Topsy Labs.
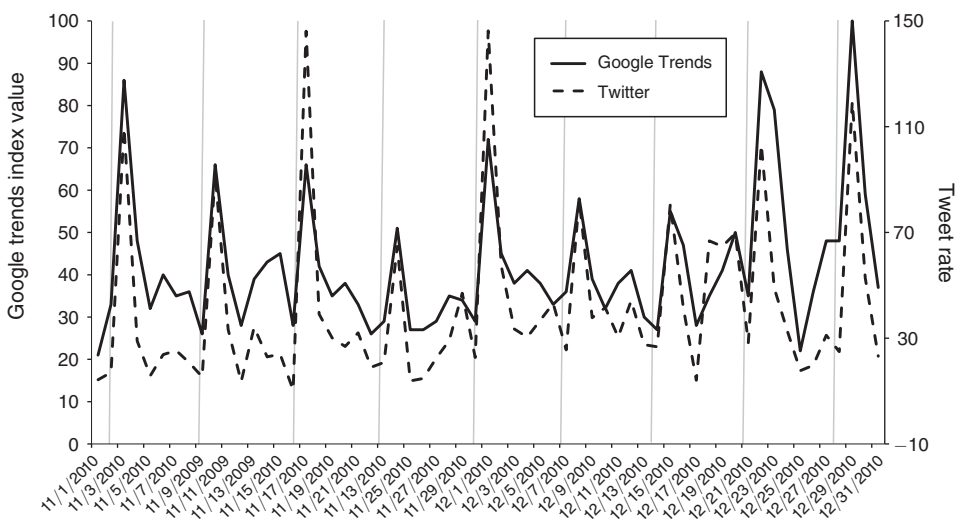
FIGURE 8. DAILY GOOGLE SEARCHES AND TWEETS ABOUT *16 AND PREGNANT*

Note: Vertical lines reflect days that new episodes of *16 and Pregnant* were broadcast.

Sources: Search index values are available directly from Google Trends. Tweet rates are computed based on the frequency of tweets obtained from Topsy Lab.

TABLE 3—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	0.825 (1.216)	2.227 (2.010)	−1.627 (1.730)	—	—
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: 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, monthly dummy variables. Twitter regressions are weighted by the total number of tweets made on each day.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

Next, we turn to estimates of equations (6) and (7) above, investigating whether airing a new episode increased searches/tweets related to birth control or abortion. In our analysis of Google Trends data, the specific search terms that we considered were "How Get Birth Control;" "How Get Birth Control Pill;" and "How Get Abortion."[45] Table 3 presents the results of this analysis with results for our Google Trends examination reported on the left side of the table. We are unable to detect statistically significant increases for any of these search terms using these data. We report the results of 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.[46]

We report the results of estimating equation (8) in the top panel of Table 4. This analysis estimates regression models where the dependent variable reflects search/tweet activity for terms related to unprotected sexual activity and the key independent variable reflects search/tweet activity mentioning *16 and Pregnant*. These results indicate that increases in Google searches for *16 and Pregnant* are

---

[45] We also experimented with terms that might capture increased interest in getting pregnant or becoming a teen mom, but that exploration was not fruitful.

[46] When we estimate this OLS relationship putting the search index value or the tweet rate for the phrase "*16 and Pregnant*" on the left-hand side, the estimated coefficient indicates that 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 and the tweet rate rises 28 percent the day a new episode airs and 109 percent ($100 \times (e^{\beta} - 1)$ 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 6 and 7. But as described above, we have no meaningful way to interpret these magnitudes; we interpret the fact that there are statistically significant increases in these measures of search and twitter activity as merely consistent with a response to the show, but unfortunately not one that can be readily translated into an absolute number of searches or tweets.

Table 4—Impact of Searches/Tweets about *16 and Pregnant* on Searches/Tweets about 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"
Panel A. National, high frequency data					
Search index/ln(tweet rate)	0.012	0.069	−0.074**	0.077**	0.064***
16 and Pregnant	(0.026)	(0.042)	(0.036)	(0.034)	(0.025)
Number of weeks (searches)/days (tweets)	209	209	209	336	336
Panel B. 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 4. 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 are 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.

  ***Significant at the 1 percent level.
   **Significant at the 5 percent level.
    *Significant at the 10 percent level.

related to 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.

Estimates from difference-in-difference models characterized by equation (9), using lower frequency data that includes state variation, are reported in the lower panel of Table 4. The results of this analysis show that in states in which search activity for *16 and Pregnant* rose the most following the introduction of the show, searches related to getting birth control and abortion also rose the most.[47] 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 teens (presumably) regarding abortion and contraception, providing perhaps a glimpse into their thinking regarding these intermediary steps leading to giving birth.

V. Discussion

Using data from numerous sources, we have examined the impact that the MTV show, *16 and Pregnant*, has had on rates of teen childbearing in the United States.

---

[47] 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.

We find that exposure to the *16 and Pregnant* franchise of shows had a sizable impact on the rate at which teens give birth in the United States, generating a 4.3 percent reduction in teen births that would have been conceived between June 2009, when the show began, and the end of 2010. That reduction can account for 24 percent of the decline over that period. Data from Google Trends and Twitter offer supportive evidence suggesting that the show influenced attitudes, in particular, raising teens' interest in birth control and abortion. We do not have sufficient data to carefully evaluate the role that more contraceptive use or 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.

These findings have important policy implications. Most teen pregnancy prevention efforts are focused on giving teenagers information about how to avoid pregnancies—either through comprehensive sex education or an abstinence-only focused curriculum—or by providing expanded access to contraception. This show produced by MTV was not specifically designed as an anti-teen childbearing campaign, but it seems to have had that effect by showing that being a pregnant teen and a new mother is hard—it strains relationships with friends, parents, and the baby's father, and means physical discomfort, potential health problems, and sleep deprivation. Apparently those images affected teenage viewers' motivation to avoid that outcome. This implies that addressing teens' motivation for avoiding teen parenthood can be an effective tool and, furthermore, that compelling social media can be used as a policy lever to do so.

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.

APPENDIX A: CONTENTS OF *16 AND PREGNANT* EPISODES

1. Show caters to teens in states with high teen birth rate
   a. (15/47 in TX, FL, AL)

2. Ambivalence toward teen pregnancy is rampant
   a. Only 18 out of 47 report opposition to their pregnancy when they found out, although none report that they were looking to get pregnant
   b. Once pregnant, 43 of 47 report ambivalence toward giving birth
   c. Most common reasons for getting pregnant include inexperience (36/47) and ambivalence (28/47). Only 5/47 report trying to avoid a pregnancy but failed

3. Abortion is rarely discussed and the decision not to have one is more about ambivalence than constraints
    a. Most (35/47) did not discuss considering an abortion
    b. Of those the 12 that did, only 3 didn't abort because of a philosophical objection
    c. 1 did so because they were afraid to tell their parents in a parental consent state

4. Contraceptive mistakes not major cause of pregnancies
    a. Most (36/47) report not using contraception at time of pregnancy (can't tell for 2 others)
    b. Of those 9 who report using contraception, 7 report using pill
    c. Suggests either not telling the truth or not taking pill every day

5. Relationships were monogamous
    a. No girl reported having sex with other partners at time of pregnancy

6. Limited information about girls' family background
    a. Most girls (38/47) did not live in two parent households at time of pregnancy
    b. Very little specific information provided about the families of the girls (parents' employment, occupation, education)

7. Demographics of the girls largely match demographics of the population, but do not match demographics of teen girls giving birth
    a. 32/47 of girls are non-Hispanic white, eight are Hispanic, five are black, non-Hispanic, and two are other race, non-Hispanic

8. The vast majority of the boyfriends stick around, although relationships are strained
    a. Of all the pregnancies, only four led to a marriage prior to the birth
    b. Almost all (40/47) of the boyfriends stick around through the pregnancy
    c. By the end of the episode (on average, about a month after birth), most (32/44) are still in relationship after the birth (note that there were three adoptions)
    d. Many, but not all fathers (31/44) are living with their children
    e. For a large minority (12/32) of those still in relationship, that relationship is very strained
    f. Of all the pregnancies, only four led to a marriage prior to the birth
    g. The majority of fathers (26/44) are still heavily involved in their child's life, and most of the others (14/44) have limited involvement. Only 4/44 are totally out of picture

9. After the birth, parents of teen often heavily involved
    a. Most parents (42/47) involved in teen's life during pregnancy and 32 of them provide active support during and after pregnancy
    b. Majority of girls (32/44) still live with parents after giving birth

10. Girls on the show are doing well in their educational attainment under the circumstances
   a. Only 11 of 44 of teen mothers have dropped out. This is not much higher than the national average for all students.

11. Health issues for the mother and the child were prevalent at "normal" rates
   a. 11 out of the 47 births took place via C-section
   b. 6 out of 47 births involved some health complication for the child
   c. 5 out of the 47 mothers experienced some health problems themselves during pregnancy (early contractions, bed rest, preeclampsia)
   d. Some level of sleep deprivation was rampant
   e. Little indication that giving birth led to satisfaction or joy from the experience

Appendix B: Specific Tweet Content

We consider it beyond the scope of our research to systematically review the specific content of the millions of tweets that address *16 and Pregnant*. But we did run a basic search and scan a handful of the specific tweets using data available from search.twitter.com.[48] We found that instructive, and share some of our findings here.

When we have done this for the search term, *16 and Pregnant*, a small number of the resulting tweets indicate that the person was making a largely innocuous tweet:

• *When I go places with Tessa people give me rude looks because they think she is my baby and I was 16 and Pregnant.*

In many other instances, the tweet is clearly about the show, although still largely content free:

• *I just saw a fat kid wearing a shirt that said 16 and Pregnant.*

Others are clearly more pointed and give an indicating of what the person tweeting is actually thinking:

• *16 and Pregnant just makes me not ever want to have a kid hahaha.*
• *still not overly impressed with my decision to take a spring class. I've seen better choices on 16 and Pregnant.*

Examples of tweets that capture the potential glamorizing effect are also present, but far less common:

• *I wanna be 16 and Pregnant. Who wanna be my baby daddy?*

---

[48] These exact tweets may not be reproducible from that source in the future because it is based on searches of recent tweets. We have screen shots capturing what was returned when we conducted these searches.

APPENDIX TABLE B1—ESTIMATES OF THE IMPACT OF *16 AND PREGNANT* RATINGS ON TEEN BIRTH RATES

	OLS (1)	IV (2)	Reduced form (3)
Rating × (2005:I–2005:III)	−0.043 (1.148)	1.075 (1.897)	1.782 (3.077)
Rating × (2005:IV–2006:II)	0.246 (0.874)	1.579 (1.272)	2.464 (2.069)
Rating × (2006:III–2007:I)	−0.120 (0.580)	0.993 (0.942)	1.551 (1.508)
Rating × (2007:II–2007:IV)	−0.234 (0.433)	0.072 (0.763)	0.171 (1.266)
Rating × (2008:I–2008:II)	−0.272 (0.494)	0.605 (0.685)	0.933 (1.121)
Base period: (2008:III–2009:II)	—	—	—
Rating × (2009:III–2010:I)	−1.035 (0.334)	−1.880 (0.562)	−2.816 (0.957)
Rating × (2010:II–2010:IV)	−1.127 (0.400)	−1.510 (0.677)	−2.328 (1.178)
Unemployment Rate	−1.435 (0.405)	−1.490 (0.384)	−1.524 (0.442)

Notes: The data used for this analysis represents quarterly birth rates by DMA for conceptions leading to live births between 2005 and 2010. The sample size in each model is 4,919 (205 DMAs, 24 quarters, and one observation was dropped because there were no teen births). The dependent variable, the birth rate, is measured in natural logs. The rating variable represents ratings for *16 and Pregnant* in OLS and IV models and MTV ratings in 2008–2009 in the reduced form model. Coefficients and standard errors (reported in parentheses) are multiplied by 100. Each model also includes the percentage of a DMA's female teen that is Hispanic and non-Hispanic black along with quarter and DMA × season fixed effects. Regressions are weighted by the relevant sample sizes for each outcome. Reported standard errors are clustered at the DMA level. Bolded results are statistically significant at the 5 percent level.

Perhaps more revealing are the results one obtains from a similar exercise where the search terms include "*16 and Pregnant*" and "birth control." The types of tweets that routinely appear include the following:

- *16 and Pregnant is a great way of birth control.*
- *seriously, watching #16 and Pregnant is birth control itself*
- *watching 16 and Pregnant is a great refresher on why NOT to get pregnant. Perfect birth control*
- *Teen Mom and 16 and Pregnant are birth control for me … That shit makes me never want kids.*

Tweets that mention 16 and Pregnant and abortion largely take two forms. Some are purely political:

- *Romney is against abortion so basically that means 3,676 new seasons of teen mom and 16 and Pregnant.*
- *If abortion was running for public office, "16 and Pregnant" would be its most effective campaign ad,"*
- *this week on 16 and Pregnant the virgin mary has to deal with joseph trying to pressure her into an abortion.*

Others clarify that the individuals are thinking about abortion more broadly:

- *Watching 16 and Pregnant, I'd rather have an abortion than give my baby up for adoption.*
- *What would happen if you were 16 and Pregnant?—i dont know to be honest, im against abortion, hopefully im not.*

REFERENCES

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

Albert, Bill. 2010. *With One Voice 2010: America's Adults and Teens Sound Off About Teen Pregnancy.* Washington, DC: 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." *B. E. Journal of Economic Analysis and Policy: Topics in Economic Analysis and Policy* 5 (1): 1–22.

Baker, Scott R., and Andrey Fradkin. 2015. "The Impact of Unemployment Insurance on Job Search: Evidence from Google Search Data." http://andreyfradkin.com/assets/FullTexasJobSearch.pdf (accessed September 29, 2015).

Barberá, Paolo. 2015. "Birds of the Same Feather Tweet Together: Bayesian Ideal Point Estimation Using Twitter Data." *Political Analysis* 23 (1): 76–91.

Blankespoor, Elizabeth, Gregory S. Miller, and Hal D. White. 2013. "The Role of Dissemination in Market Liquidity: Evidence From Firms' Use of Twitter." http://ssrn.com/abstract=1657169.

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

Buckles, Kasey S., and Daniel M. Hungerman. 2013. "Season of Birth and Later Outcomes: Old Questions, New Answers." *Review of Economics and Statistics* 95 (3): 711–24.

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): 1047–54.

Chiou, Lesley, and Mary Lopez. 2010. "The Reality of Reality Television: Does Reality TV Influence Local Crime Rates?" *Economics Letters* 108 (3): 330–33.

Choi, Hyunyoung, and Paul Liu. 2011. "Reading Tea Leaves in the Tourism Industry: A Case Study in the Gulf Oil Spill, 2011." http://google.com/googleblogs/pdfs/google_gulf_tourism_march2011.pdf (accessed September 29, 2015).

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 (accessed September 29, 2015).

Choi, Hyunyoung, and Hal Varian. 2009b. "Predicting Initial Claims for Unemployment Benefits." Google Technical Report. http://research.google.com/archive/papers/initialclaimsUS.pdf (accessed September 29, 2015).

Chong, Alberto, and Eliana La Ferrara. 2009. "Television and Divorce: Evidence from Brazilian Novelas." *Journal of the European Economic Association* 7 (2–3): 458–68.

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): 1531–45.

Da, Zhi, Joseph Engelberg, and Pengjie Gao. 2011. "In Search of Attention." *Journal of Finance* 66 (5): 1461–99.

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

DellaVigna, Stefano, and Ethan Kaplan. 2007. "The Fox News Effect: Media Bias and Voting." *Quarterly Journal of Economics* 122 (3): 1187–1234.

Garthwaite, Craig, Tal Gross, and Matthew J. Notowidigdo. 2014. "Public Health Insurance, Labor Supply, and Employment Lock." *Quarterly Journal of Economics* 129 (2): 653–96.

Gentzkow, Matthew, and Jesse M. Shapiro. 2008. "Preschool Television Viewing and Adolescent Test Scores: Historical Evidence from the Coleman Study." *Quarterly Journal of Economics* 123 (1): 279–323.

Gerber, Alan S., James G. Gimpel, 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): 135–50.

Gorman, Bill. 2011. "2010–11 Season Broadcast Primetime Show 18–49 Ratings Averages." *zap2it TV by the Numbers*, June 1. http://tvbythenumbers.zap2it.com/2011/06/01/2010-11-season-broadcast-primetime-show-18-49-ratings-averages/94402 (accessed October 10, 2013).

Jensen, Robert, and Emily Oster. 2009. "The Power of TV: Cable Television and Women's Status in India." *Quarterly Journal of Economics* 124 (3): 1057–94.

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): 141–63.

Kearney, Melissa S., and Phillip B. Levine. 2012b. "Explaining Recent Trends in the U.S. Teen Birth Rate." National Bureau of Economic Research Working Paper 17964.

Kearney, Melissa S., and Phillip B. Levine. 2014a. *Teen Births Are Falling: What's Going On?* Brookings Policy Brief. Washington, DC: Brookings Institution.

Kearney, Melissa S., and Phillip B. Levine. 2014b. "Income Inequality and Early Nonmarital Childbearing." *Journal of Human Resources* 49 (1): 1–31.

Kearney, Melissa S., and Phillip B. Levine. 2015. "Media Influences on Social Outcomes: The Impact of MTV's *16 and Pregnant* on Teen Childbearing: Dataset." *American Economic Review*. http://dx.doi.org/10.1257/aer.20140012.

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.

Lopoo, Leonard M., and Kerri M. Raissian. 2012. "Natalist Policies in the United States." *Journal of Policy Analysis and Management* 31 (4): 905–46.

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

Martin, Joyce A., Brady E. Hamilton, Stephanie J. Ventura, Michelle J. K. Osterman, Elizabeth C. Wilson, and T. J. Mathews. 2012. "Births: Final Data for 2010." *National Vital Statistics Reports* 61 (1): 1–71.

Menacker, Fay, and Brady E. Hamilton. 2010. "Recent Trends in Cesarean Delivery in the United States." US Department of Health and Human Services, NCHS Data Brief 35. http://www.cdc.gov/nchs/data/databriefs/db35.pdf.

MTV. 2012. "MTV's Groundbreaking Docu-Series '16 and Pregnant' Returns on Tuesday, March 27 at 10pm ET/PT." *MTV Press*, July 12. http://mtvpress.com/press/release/mtvs_groundbreaking_docu_series_16_and_pregnant_returns_on_tuesday_march_27 (accessed October 10, 2013).

Myers, Caitlin Knowles. 2008. "Discrimination as a Competitive Device: The Case of Local Television News." *B. E. Journal of Economic Analysis and Policy* 8 (1). http://dx.doi.org/10.2202/1935-1682.1995.

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 (8): 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): 363–73.

Rui, Huaxia, Yizao Liu, and Andrew Whinston. 2013. "Whose and What Chatter Matters? The Effect of Tweets on Movie Sales." *Decision Support Systems* 55 (4): 863–70.

Saffer, Henry, Melanie Wakefield, and Yvonne Terry-McElrath. 2007. "The Effect of Nicotine Replacement Therapy Advertising on Youth Smoking." National Bureau of Economic Research Working Paper 12964.

Seidman, Robert. 2009. "MTV Greenlights Season Two of Hit Series '16 & Pregnant.'" *zap2it TV by the Numbers*, June 30. http://tvbythenumbers.zap2it.com/2009/06/30/mtv-greenlights-season-two-of-hit-series-16-pregnant/21672/ (accessed October 10, 2013).

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

Stein, Rob. 2008. "Study First to Link TV Sex To Real Teen Pregnancies." *Washington Post*, November 3. http://www.washingtonpost.com/wp-dyn/content/article/2008/11/02/AR2008110202592.html (accessed September 29, 2015).

Stephens-Davidowitz, Seth I. 2013a. "The Cost of Racial Animus on a Black Presidential Candidate: Using Google Search Data to Find What Surveys Miss." http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2238851 (accessed September 29, 2015).

Stephens-Davidowitz, Seth. 2013b. "Unreported Victims of an Economic Downturn." http://static1.squarespace.com/static/51d894bee4b01caf88ccb4f3/t/51e22f38e4b0502fe211fab7/1373777720363/childabusepaper13.pdf (accessed September 29, 2015).

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

Trudeau, Jennifer. Forthcoming. "The Role of New Media on Teen Sexual Behaviors and Fertility Outcomes: The Case of *16 and Pregnant.*" *Southern Economic Journal.*

US Census Bureau. 2014. "Annual Estimates of the Resident Population by Sex, Age, Race Alone or in Combination, and Hispanic Origin for the United States and States: April 1, 2010 to July 1, 2013." http://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk (accessed September 29, 2015).

Vlastakis, Nikolaos, and Raphael N. Markellos. 2012. "Information Demand and Stock Market Volatility." *Journal of Banking and Finance* 36 (6): 1808–21.

This article has been cited by:

1. Vittorio Bassi, Imran Rasul. 2017. Persuasion: A Case Study of Papal Influences on Fertility-Related Beliefs and Behavior. *American Economic Journal: Applied Economics* **9**:4, 250-302. [Abstract] [View PDF article] [PDF with links]
2. Anna Aizer. 2017. A Review Essay on Isabel Sawhill's Generation Unbound: Drifting into Sex and Parenting without Marriage and Laurence Steinberg's Age of Opportunity: Lessons from the New Science of Adolescence. *Journal of Economic Literature* **55**:2, 592-608. [Abstract] [View PDF article] [PDF with links]