

IZA DP No. 10317

**Does Reality TV Induce Real Effects?
On the Questionable Association Between
16 and Pregnant and Teenage Childbearing**

David A. Jaeger
Theodore J. Joyce
Robert Kaestner

October 2016

Does Reality TV Induce Real Effects? On the Questionable Association Between 16 and Pregnant and Teenage Childbearing

David A. Jaeger

*CUNY Graduate Center,
University of Cologne, IZA and NBER*

Theodore J. Joyce

Baruch College, CUNY and NBER

Robert Kaestner

University of California, Riverside and NBER

Discussion Paper No. 10317
October 2016

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Does Reality TV Induce Real Effects? On the Questionable Association Between *16 and Pregnant* and Teenage Childbearing*

We reassess recent and widely reported evidence that the MTV program *16 and Pregnant* played a major role in reducing teen birth rates in the U.S. since it began broadcasting in 2009 (Kearney and Levine, *American Economic Review* 2015). We find Kearney and Levine's identification strategy to be problematic. Through a series of placebo and other tests, we show that the exclusion restriction of their instrumental variables approach is not valid and find that the assumption of common trends in birth rates between low and high MTV-watching areas is not met. We also reassess Kearney and Levine's evidence from social media and show that it is fragile and highly sensitive to the choice of included periods and to the use of weights. We conclude that Kearney and Levine's results are uninformative about the effect of *16 and Pregnant* on teen birth rates.

JEL Classification: J13, L82

Keywords: teen childbearing, media, social media, internet

Corresponding author:

David A. Jaeger
Ph.D. Program in Economics
CUNY Graduate Center
365 Fifth Ave
New York, NY 10016
USA
E-mail: djaeger@gc.cuny.edu

* The authors thank Phil Levine for his generous assistance in replicating results, and Bill Evans, Gary Solon, and seminar participants at Columbia University, the CUNY Graduate Center, the CUNY Institute for Demographic Research, the Guttmacher Institute, and the University of Michigan for helpful comments. Onur Altindag provided exemplary research assistance. Joyce and Kaestner acknowledge support from the National Institute of Child Health and Human Development Grant R01 HD082133 and from the Health Economics Program of the National Bureau of Economic Research. Nielsen ratings data are proprietary and cannot be shared. Birth data from the National Center for Health Statistics (NCHS) have confidential elements and must be obtained directly from NCHS. Institutional Review Board (IRB) approval was obtained from the National Bureau of Economic Research.

The birth rate among 15 to 19 year olds in the United States was 61.1 per 1,000 teens in 1991, the highest rate since 1971. The rate among black teens was 114.8 in 1991. Many policy makers viewed such high rates as a national crisis (National Research Council 1987). Since 1991, however, the overall teen birth rate has fallen almost continuously, declining by more than half to 24.2 in 2014. Over the same period, the birth rate among African American teens has fallen by nearly two thirds, to 35.1. While these are substantial changes, the teen birth rate in the U.S. is still higher than that in Canada and much of Western Europe, and a consensus as to the causes of the decline remains elusive.¹ In an effort to understand which policies and interventions might prove effective at helping teens avoid an unintended pregnancy, in 2010 the U.S. Office of Adolescent Health awarded 100 million dollars in grants to fund 102 Teen Pregnancy Prevention (TPP) projects. Each project was required to have a rigorous evaluation plan and over 90 percent used randomized designs. A recently released summary of results from these evaluations revealed that few of the interventions were effective at changing teen behavior with respect to sexual activity and contraceptive use.²

The widely-held belief that teen pregnancy has adverse consequences for young mothers and the failure of the TPP initiative to demonstrate a clearly effective intervention to prevent teen pregnancy may explain why a recent study that claimed a reality TV show caused a significant

¹ There are a variety of possible explanations (Boonstra 2014), including more effective contraception (John S. Santelli, *et al.* 2007; Laura Lindberg, John Santelli and Shelia Desai forthcoming), welfare reform (Robert Kaestner, Sanders Korenman and June O'Neill 2003; Leonard M. Lopoo and Thomas Delaire 2006), labor market conditions (Rajeev Dehejia and Adriana Llearas-Muney 2004, and Elizabeth Olitmans Ananat, Anna Gassman-Pines and Christina Gibson-Davis 2013), and income inequality (Kearney and Levine 2014b). None of these explanations appear to be definitive.

²See <http://www.hhs.gov/ash/oah/oah-initiatives/evaluation/grantee-led-evaluation/summary-researchdemonstration.pdf>; <http://www.hhs.gov/ash/oah/oah-initiatives/evaluation/grantee-led-evaluation/summary-ebps.pdf> (last accessed 15 September 2016)

decline in teen birth rates garnered immediate and widespread attention in the national media.³ The study by Melissa S. Kearney and Philip B. Levine (2015c, henceforth KL) concludes that the MTV reality shows *16 and Pregnant*, *Teen Mom*, and *Teen Mom 2*, by dramatizing the challenges of pregnancy and childrearing, caused a 4.3 percent drop in teen birth rates between July 2009 and December 2010. This effect is so large that it accounts for a quarter of the total reduction in teen childbearing during this period. KL assert that their estimates represent a causal impact and claim that "...a social media campaign in the guise of a very popular reality TV show... adds a new 'policy mix'" (p. 3598) to current interventions designed to reduce teen pregnancy.

In this paper, we reassess the basic hypothesis and empirical foundations of KL's conclusions. At the heart of KL's paper is an identification strategy that is, at best, problematic. *16 and Pregnant* began broadcasting nationally in June 2009, providing a pre- and post-treatment analysis, but no clear counterfactual. To overcome this limitation, KL use the audience share watching MTV in the period before *16 and Pregnant* was aired as an instrument for watching *16 and Pregnant*. Both measures of viewership reflect choices, however, and even KL concede that their instrument does not provide a truly exogenous source of identifying variation in the viewing of *16 and Pregnant*.

³ For the initial reaction to the working paper by Kearney and Levine (2014) see: *CNN*: <http://www.cnn.com/2014/01/13/health/16-pregnant-teens-childbirth/>; *The Washington Post*: <https://www.washingtonpost.com/news/arts-and-entertainment/wp/2014/04/09/how-mtvs-16-and-pregnant-led-to-declining-teen-birth-rates/>; *Time*: <http://time.com/825/does-16-and-pregnant-prevent-or-promote-teen-pregnancy/>; *Newsweek*: <http://www.newsweek.com/why-teen-birth-rate-keeps-dropping-333946>; and *The New York Times*: <http://www.nytimes.com/2014/01/13/business/media/mtvs-16-and-pregnant-derided-by-some-may-resonate-as-a-cautionary-tale.html> (All last accessed 15 January 2016). For the continuing popularity of *16 and Pregnant* as an explanation for the decline in teen birth rates, see <http://www.nytimes.com/2016/07/19/opinion/winning-the-campaign-to-curb-teen-pregnancy.html> (Last accessed 24 August 2016).

We show that the assumption of common trends in birth rates between low and high MTV-watching areas necessary for identifying a causal effect in KL’s research design is not met. Placebo analyses using only the pre-*16 and Pregnant* period yield similar results to those of KL, indicating a violation of the exclusion restriction necessary for identification. We also replicate KL’s results on the association between *16 and Pregnant* and Google searches and Twitter tweets for birth control. When we estimate these regressions without using weights and/or include in the analysis the appropriate counterfactual periods when *16 and Pregnant* was not being broadcast, we find their results to be so fragile that nothing can be inferred from them.

We conclude that causal claims about the effect of *16 and Pregnant* on teen birth rates are unwarranted. More generally, we show the importance of clearly and robustly establishing that the parallel trends assumption holds in difference-in-difference analyses that use geographic variation for identification. Our results also underscore the great and greatly underappreciated importance of replication in the social sciences, especially for results that receive widespread and uncritical attention in the national media, potentially leading unproven strategies becoming part of the “policy mix.”

I. Identification in Previous Studies of the Impact of Television on Behavior

The previous literature has examined the impact of the introduction of television or specific television content on a variety of outcomes, such as voter turnout (Matthew Gentzkow 2006), tests scores (Gentzkow and Jesse M. Shapiro 2008), women’s empowerment (Robert Jensen and Emily Oster 2009); voting behavior (Stefano DellaVigna and Ethan Kaplan 2007); and fertility and divorce (Eliana La Ferrara, Alberto Chong, and Suzanne Duryea 2012). In each of these studies, the authors acknowledge the endogeneity of television viewing and address this

problem using plausibly exogenous changes across both space and time in access to television or specific content in a reduced form model and/or in an instrumental variables approach. Indeed, Kearney and Levine (2015b) themselves follow a similar approach in their recent study on the effect of *Sesame Street* broadcasts on long-term educational outcomes. They use the distance to a UHF or VHF television tower and the introduction of *Sesame Street* in different places at different times as sources of plausibly exogenous variation, arguing that that historical FCC decisions regarding the placement of television towers are plausibly unrelated to the characteristics of children who could potentially watch *Sesame Street*.

Unlike these studies, KL do not exploit a plausibly exogenous source of variation in viewership of *16 and Pregnant* to obtain estimates of the show's effect on teen fertility. *16 and Pregnant* started to be broadcast everywhere in the U.S. at a single point in time, providing no clear control group (i.e. where *16 and Pregnant* was not available) to serve as a counterfactual. To address this issue, KL use exposure to *16 and Pregnant*, measured by the Nielsen ratings for the show, and compare areas with greater and smaller viewership. The authors recognize that *16 and Pregnant* viewership is unlikely to be exogenous and, therefore instrument for it using Nielsen ratings for MTV viewership in the period directly preceding the beginning of *16 and Pregnant*. MTV viewership is also unlikely to be exogenous, however, because it is based on socioeconomic and other differences across places. Although KL control for area fixed effects, their instrumental variables approach is vulnerable to time-varying confounders and lacks a strong theoretical basis for its validity.

II. KL's Empirical Framework

KL use an instrumental variables approach to obtain estimates of the effect of *16 and Pregnant* viewership on teen fertility, as described by the following four equations, which we reproduce here:

$$(1) \quad \ln(BR_{jt}) = \beta_0 + \beta_1(Rate16P_j \times Post_t) + \beta_2 U_{jy} + \mathbf{X}_{jy} \gamma + \theta_t + \delta_{js} + \epsilon_{jt},$$

$$(2) \quad \ln(BR_{jt}) = \beta_0 + \beta_1 \overbrace{(Rate16P_j \times Post_t)} + \beta_2 U_{jy} + \mathbf{X}_{jy} \gamma + \theta_t + \delta_{js} + \epsilon_{jt},$$

$$(3) \quad (Rate16P_j \times Post_t) = \beta_0 + \beta_1(MTV0809_j \times Post_t) + \beta_2 U_{jy} + \mathbf{X}_{jy} \gamma + \theta_t + \delta_{js} + \epsilon_{jt},$$

and

$$(4) \quad \ln(BR_{jt}) = \beta_0 + \beta_1(MTV0809_j \times Post_t) + \beta_2 U_{jy} + \mathbf{X}_{jy} \gamma + \theta_t + \delta_{js} + \epsilon_{jt}.$$

We follow KL's notation, although in actuality the parameters represent different quantities across the different equations, i.e. β_1 in equation (3) is not the same quantity as β_1 in equation (4). In the equations, BR_{jt} is the birth rate by calendar quarter of conception t , in Designated Market Area (DMA) j ; $Rate16P_j$ is the average Nielsen rating in DMA j for the shows *16 and Pregnant*, *Teen Mom*, and *Teen Mom 2* (henceforth *16 and Pregnant*) for 12 to 24 year-old viewers for the seven Nielsen sweep months of July 2009, February 2010, July 2010, and November 2010, May 2011, November 2011, and May 2012; $Post_j$ is a dichotomous indicator for the period after the introduction of *16 and Pregnant*; $MTV0809$ is the average Nielsen rating in for 12 through 24 year olds in DMA_j for all MTV shows in the sweeps months of July 2008, November 2008, February 2009, and May 2009, U_{jy} is the annual unemployment rate in DMA j in year y ; \mathbf{X}_{jy} is vector that includes the percent population in the DMA that is that non-Hispanic black, and the percent that is Hispanic in calendar year y ; θ_t is a set of quarterly dummy

variables; and δ_{js} is a full set of DMA \times season fixed effects, which implicitly defines DMA fixed effects.

Equation (1) represents the equation of interest that yields estimates of the association between *16 and Pregnant* viewership and teen birth rates. Equation (3) is the first stage regression with $(MTV0809_j \times Post_t)$ as an instrument. Equation (2) is the second stage regression and identical to equation (1) except that predicted *16 and Pregnant* viewership derived from equation (2) is used instead of actual *16 and Pregnant* viewership. Equation (4) represents the reduced form effect of MTV viewership on teen birth rates.

KL's study period includes 24 quarters (2005:QI – 2010:QIV) for 205 DMAs for a total (potential) sample of 4,920 observations.⁴ Ratings for *16 and Pregnant* are measured during the time slot from 9:00 to 10:00 pm for 12 to 24 year-old viewers on Tuesdays in the sweep months.⁵ KL average these Tuesday ratings within each month and then average the four months of ratings

⁴ The dependent variable is missing in some DMAs in some quarters because there are no teen births. This occurs more frequently in the analyses stratified by age or race/ethnicity.

⁵ KL use ratings for a broad age range from a single time in the later evening (9:00-10:00 p.m.) when a new episode of *16 and Pregnant* was broadcast. Using this as a measure of the total exposure to *16 and Pregnant* for teenagers has some obvious problems. Previous episodes are broadcast at various times on the same day as the new episode, which is also rebroadcast at different times during the week. KL could have used what Nielsen terms "Gross Rating Percentages," which averages the percent of households watching any broadcast of *16 and Pregnant* at any time during the sweep month. This double counts any individual who watches the same show more than once, but is perhaps a better measure of exposure. Nielsen also collects separate ratings for teens aged 12-17 separately from young women 18 to 24 years old. The younger teens are more likely to watch *16 and Pregnant* earlier in the day than women 18 to 24. By aggregating ratings across age groups, KL potentially reduce noise in the data, but at the expense of accurately measuring exposure. KL also average Nielsen ratings from seven sweep months drawn from 2009-2012 and assign this average to conceptions from 2009:III to 2010:QIV, leading to another source of mis-measurement of exposure to *16 and Pregnant*. Teens who conceived in August of 2009, for example, are assigned a rating that pertains to future broadcasts to which they could not have been exposed. Ratings almost tripled in the first year suggesting important variation by quarter of conception.

and assign that value to the six post-*16 and Pregnant* quarters within each DMA. They follow the same procedure for the MTV ratings during the period 2008:QIII to 2009:QII by using the average for the four quarters within each DMA. The measures that KL use for *16 and Pregnant* viewership and MTV viewership therefore vary only in the cross section. Note that because the models also include DMA fixed effects, which absorb the time-invariant determinants of fertility (in equations (1), (2), and (4)) or of *16 and Pregnant* viewership (in equation (3)), the main effect of *16 and Pregnant* cannot be estimated in equations (1), (2), and (4), nor can the main effect of MTV be included in equation (3). Rather, only the interactions between $Rate16P_j$ and $MTV0809_j$ with $Post_t$ can be included on the right hand side of the relevant regressions.

The identification strategy implied by equations (1) through (4) is a comparison of teen birth rates before and after the introduction of *16 and Pregnant*, stratified by levels MTV viewership in the year before the show began. This is essentially a difference-in-differences approach.⁶ The key identifying assumption is therefore parallel trends in teen birth rates by levels of MTV ratings in the years prior to *16 and Pregnant*.

III. Placebo Tests

For estimates from equation (3) to have a causal interpretation, MTV ratings should be unrelated to trends in birth rates conditional on DMA fixed effects except through its relationship

⁶ Trudeau (2016) also uses a difference-in-differences strategy in her analysis of *16 and Pregnant* and teen sexual behavior. She compares sexual activity of younger teens ages 15-18 to older women ages 19-22 and then birth rates of 15-16 year olds to 19 and 20 years before and after *16 and Pregnant*, i.e. she uses the older women as the comparison group for the younger ones. The strategy is highly questionable because both groups are exposed to *16 and Pregnant* and they differ dramatically in their levels of sexual activity and birth rates. Although Trudeau includes area fixed effects, small variations in trends can yield spurious findings. She finds no effect on birth rates when she contrasts changes between younger and older women, but statistically significant declines among younger teens in a first-difference model interacted with the Nielsen ratings of the show by designated market area (DMA).

with *16 and Pregnant* ratings. Any association with MTV ratings would be a violation of the exclusion restriction required for the validity of the instrumental variables estimator, as MTV should only affect birth rates through its relationship with *16 and Pregnant*. In Table 1 we present results from a series of reduced form and IV regressions in which we artificially start a placebo “show” in sequential quarters. Each regression includes 24 quarters, the same length of analysis used by KL: 18 prior to the beginning of the placebo show and 6 after, including the quarter that the “show” starts. We estimate the reduced form equation (4) and IV equation (5) and use MTV viewing from 2008-2009 as the instrument.

Each row of Table 1 presents the estimated coefficients and standard errors from the estimation of the reduced form and instrumental variables models from a different 24 quarter-period. For example, in row (1) the 24-quarter period begins in 1999:Q1 and ends in 2004:QIV, with the placebo show beginning in 2003:QIII. Row 25, which is boxed in the table, reproduces KL’s reduced form and instrumental variables results from their Table 1 columns (4) and (3), respectively. The lightly shaded rows include periods that include some of the actual post-*16 and Pregnant* period in the placebo post-“show” period and the darkly shaded rows include some of the actual post-*16 and Pregnant* period in both the pre- and post-“show” periods.⁷

For the reduced form, we find that 10 of the 19 estimated coefficients in rows (1) through (19), which use data exclusively from the pre-*16 and Pregnant* period, are statistically different from zero at conventional levels. That all of the estimates except for row (1) are negative and

⁷ Following KL, we weight all regressions by the female population aged 15-24 in the DMA in the quarter of observation. We present unweighted results for the same placebo tests from Table 1 in Appendix Table 1. In general, fewer of the placebo tests are statistically significant when we do not weight, but the IV estimate for the actual *16 and Pregnant* period that KL analyze becomes implausibly large (-4.504 with standard error of 1.343) suggesting that *16 and Pregnant* can explain over 50 percent of the decline in teen birth rates in the first 18 months of the show (see KL, p. 3616).

roughly equal to KL's reduced form result in magnitude strongly suggests that MTV viewing is not a valid instrument and that the parallel trends assumption required for the difference-in-differences strategy does not hold.⁸ Because the sample period is fixed at 24 quarters, the first stage relationship between MTV and *16 and Pregnant* ratings is quite stable, with the first-stage coefficient on MTV Ratings ranging from 1.457 in row (5) to 1.524 in row (17); the first-stage coefficient is 1.513 in row (25), which corresponds to KL's analysis period. The IV results therefore reflect those from the reduced form.

IV estimates in Table 1 are similar to KL's "IV Event Study" (p. 3614) in the sense that we use data on MTV and *16 and Pregnant Ratings* that post-date the period under study. For example, in row (13) our IV estimate suggests that teen birth rates decrease 1.958 percent from a one rating point increase in our placebo *16 and Pregnant* show, which is very similar to the estimate of -1.88 that KL report in their Figure 6 and Appendix Table B1. Our estimate captures

⁸ Adjusting for multiple comparisons in Table 3 using the Carlo E. Bonferroni (1936) correction or the somewhat more powerful Bonferroni-Holm (Sture Holm 1979) correction leads to rejection of none of the null hypotheses that the coefficients in both the reduced form results and the instrumental variables results are equal to zero, including KL's results in row (25). For the purposes of these tests, we treat the reduced form and instrumental variables results separately. Both the Bonferroni and Bonferroni-Holm multiple comparison procedures fix the so-called *familywise error rate* (the rate of committing *any* Type I error) and tend to have substantially less power than single comparison procedures. By fixing the familywise error rate, the probability of committing a Type II error (not rejecting one of the hypotheses when, in fact, it is not true) can increase quite substantially. An alternative procedure is to focus on the so-called *false discovery rate*, which is less conservative and fixes the share of null hypothesis rejections that are false rather than fixing the probability that *any* null hypothesis rejections are false. We have performed the Benjamini-Hochberg (1995) procedure and do not reject any of the hypotheses implicit in Table 3 when we fix the false discovery rate at 5 or 10 percent. If the false discovery rate is increased to 15.8 percent for the reduced form and 13.1 percent for instrumental variables, we still reject the null hypothesis that the coefficient is zero for 7 of the 19 pre-"show" periods for the reduced form and 8 of the 19 pre-"show" periods for instrumental variables when we consider all 36 multiple comparisons. If we limit our attention only to the tests in which the data are fully from the pre-*16 and Pregnant period*, we reject the null hypothesis of a zero coefficient for 10 of the 19 regressions, with a false discovery rate of 16.9 percent for the reduced form and 8 of 19 with a false discovery rate of 14.7 percent for instrumental variables. These calculations are available from the authors by request.

unobserved and time-varying factors within the DMA that are correlated with teen birth rates and *16 and Pregnant* viewership.

KL's decision to restrict their analysis to 24 quarters period is arguably limited, given the profound trends in teen births rates over the past 25 years and the centrality of the parallel trends assumption to the validity of their difference-in-differences analysis. Figure 1 shows the natural logarithm of teen births for all teens and separately for non-Hispanic whites, non-Hispanic blacks, and Hispanics from 1990 to 2014. Teen births fall continuously from 1991 to 2005 after which they rise briefly between 2005 and 2007. After 2007, however, birth rates begin a rapid decline that precedes the airing of *16 and Pregnant*.

To capture more of the long decline in birth rates, in Table 2 we perform a similar placebo exercise, but now hold the analysis period fixed and use all of the pre-*16 and Pregnant* quarters from 1999:QI to 2009:QII. The quarter in which the placebo “show” starts sweeps through the period such that there are always at least 18 pre-“show” quarters and at least 6 post-“show” quarters. Here the relationship between the reduced form and IV estimates changes across different placebo tests because first stage estimates change as the percentage of zeroes (i.e. the pre-“show” period) in MTV and *16 and Pregnant* viewership vary. The first stage coefficient is largest in row (1) at 1.539 and smallest in row (10) at 0.195. We find that all the reduced form and IV estimates are negative and of approximately the same magnitude as the actual KL estimates, which one would not expect if KL's instrument met the exclusion restriction. For the reduced form, the last 5 coefficients in rows (15) through (19) are statistically significantly different from zero at the 10 percent level while for the IV estimates those in rows (1) through (4) and those in rows (15) through (18) are statistically significant at the 10 percent

level. Notably, *16 and Pregnant* was not on the air during any of the periods in Table 2 and yet we find consistently negative and often statistically significant estimates of its effect.⁹

The results in Tables 1 and 2 cast substantial doubt on the validity of KL’s empirical strategy. If their approach was valid, and MTV viewership “randomly” assigned DMAs into those that are more and less exposed to *16 and Pregnant*, then the estimated coefficients in Tables 1 and 2 should not (nearly) all be of the same sign and few if any should be statistically significantly different from zero. Yet many of the coefficients from placebo tests that occurred wholly before *16 and Pregnant* was broadcast are statistically different from zero, and most are of the same sign and magnitude of KL’s actual estimates.

IV. Replicating and Extending KL’s Parallel Trend Assessments

KL present an event study in their Figure 5 and Appendix Table B1 as a means of testing the parallel trends assumption. The coefficients in their Figure 5 are from the following regression, which is similar to equation (4):

$$(5) \quad \ln(BR_{jt}) = \beta_0 + \sum_{i=1}^{14} \alpha_i (MTV0809_j \times PreQ_i) + \sum_{i=19}^{24} \beta_i (MTV0809_j \times PostQ_i) + \phi U_{jt} + \theta_t + \delta_{js} + \epsilon_{jt}.$$

The coefficients α_i show the average difference in log birth rates for varying levels of MTV viewership by quarter of conception in the pre-*16 and Pregnant* period (quarters 1 to 14, 2005:QI-2008:QII) relative to the reference category (quarters 15 to 18, 2008:QIII-2009:QII).

The β_i coefficients show the same contrast but for the six quarters after the introduction of *16 and Pregnant* (quarters 19 to 24: 2009:III-2010:QIV). KL find that they cannot reject the joint

⁹ We present the unweighted version of Table 2 in Appendix Table 2. As in Table 1, fewer of the individual tests are statistically significantly different from zero when not using weights.

null hypothesis that $\alpha_1 = 0, \alpha_2 = 0, \dots, \alpha_{14} = 0$, but they can reject the null hypothesis that $\beta_{19} = 0, \beta_{20} = 0, \dots, \beta_{24} = 0$. Visually their Figure 5 shows a modest downward trend in teen birth rates in the *pre-16 and Pregnant* period, with a noticeable decrease after *16 and Pregnant* began in quarter 19.

Figure 2, Panel A replicates KL’s event study analysis.¹⁰ Unlike KL, we show the actual four quarters of the reference period.¹¹ This more appropriate spacing makes the “discontinuity” at the introduction of *16 and Pregnant* less noticeable. Like KL, we show least squares fit lines through the coefficients for both the *pre-16 and Pregnant* and *post-16 and Pregnant* periods.¹² The vertical line at 2009:QIII marks the quarter in which *16 and Pregnant* was first broadcast.

KL’s decision to use 2008:QIII–2009:QII as the reference period may obscure the fall in teen birth rates in the year before *16 and Pregnant*. The absence of a noticeable discontinuity at point the show began would undercut what they term the “... strong visual support for the notion of a causal effect” (KL p. 3614). In Panel B of Figure 2 we change the reference period to 2005:QI to 2005:QIV and plot the coefficients from estimation of equation (5), *mutatis mutandis*. There is little apparent discontinuity at 2009:QIII, however, and the 95 percent confidence

¹⁰ The note to KL’s Figure 5 states: “Estimates reflect coefficients on 2008–2009 MTV ratings interacted with quarter from a regression model controlling for DMA \times season fixed effects.” We were not able to replicate their figure with that specification. We *were* able to replicate their figure when we also included the DMA-specific unemployment rate in the model (but not the race/ethnicity variables). All of the models in Figures 2 through 5 therefore include the unemployment rate as a regressor to retain comparability to KL’s results. Our general points hold, however, if we follow KL’s specification as described.

¹¹ KL’s plot of the event study coefficients in Figure 5 is misleading. Each tick mark on the horizontal axis represents a quarter of the year and the vertical axis shows the magnitude of coefficients at each quarter. But only a single tick represents the reference period, which comprises four quarters. This exaggerates the sharpness of the falloff in teen births rates in the *post-16 and Pregnant* period by making a full year appear to be only one quarter.

¹² The relevance of these fitted lines is unclear, as they are not derived from the estimated regression and do not represent formal statistical tests of a break in trend. We therefore omit them from our subsequent event study graphs in Figures 3 through 5.

intervals are noticeably wider than in Panel A. The pattern in Panel B of Figure 2 is also consistent with the unconditional trends in teen birth rates, which began to decline in 2008, a year before *16 and Pregnant* was first broadcast. We cannot, however, reject the null hypothesis that the coefficients in the pre-*16 and Pregnant* period are jointly equal to zero ($F=1.46, p=0.128$).¹³

The conclusions change when we extend the event study analysis temporally in both directions. In Figure 3, we add two years in the pre-*16 and Pregnant* period and one year in the post-*16 and Pregnant* period, so that the period under examination is 2003:Q1-2011:QIV. As in Figure 2, in Panel A we follow KL and use 2008:QIII-2009:QII as the reference and show least squares fit lines for the pre- and post-*16 and Pregnant* periods. In Panel B of Figure 3 we plot estimates from the same model, but with 2003:Q1-2003:QIV as the reference period. The downward trend in teen birth rates in the years prior to *16 and Pregnant* and the lack of a clear discontinuity at 2009:QIII is now quite apparent and we strongly reject the null hypothesis that the pre-*16 and Pregnant* coefficients are jointly equal to zero ($F=2.13, p=0.003$). Figures 4 and 5 repeat this exercise for the periods 2001:Q1-2012:QIV and 1991:Q1-2013:QIV, respectively.¹⁴ In both cases, we again strongly reject the null hypothesis that the coefficients for the period prior to *16 and Pregnant* are equal to zero ($F=2.83, p<0.001$ in Figure 5 and $F=3.41, p<0.001$ in Figure 6). There is also nothing in either figure that suggests a discontinuous break in the birth rate once *16 and Pregnant* began broadcasting.

Overall, we find that there is no evidence to suggest that *16 and Pregnant* altered the existing downward trend in birth rates that began long before its first broadcast. By using only

¹³ Note that the F -test on pre-*16 and Pregnant* coefficients is invariant to the choice of reference period. KL erroneously report the p -value from this test as 0.21 on p. 3613.

¹⁴ *16 and Pregnant* (including *Teen Mom* and *Teen Mom 2*) continued to be a popular shows through 2013.

four and one half years of data in the period before *16 and Pregnant*, KL sectioned off a time period in which teen birth rates were at roughly similar levels between 2005 and 2008 due to the rise and fall in teen birth rates over these years. When we extend the study period, the identifying assumption that teen birth rates are unrelated to MTV viewership no longer holds.¹⁵

V. Direct Tests of Parallel Trends

As KL appropriately note, their event study analysis assumes a linear relationship between MTV ratings and teen birth rates. Any association between birth rates and ratings of *MTV or 16 and Pregnant* in the period before the show began would, however, undermine a causal interpretation. We now allow a more flexible, non-parametric relationship between birth rates and MTV and *16 and Pregnant* ratings than the linear specification of equation (5). In equation (6), we interact the quarter of conception indicators with indicators for quartiles of MTV or *16 and Pregnant* ratings:

$$(6) \quad \ln(BR_{jt}) = \beta_0 + \sum_{i=1}^{17} \alpha_i Quarter_t + \sum_{i=1}^{17} \sum_{k=2}^4 \beta_{ik} (Quarter_t \times Quartile_k) + \mathbf{X}_{jt} \gamma + \beta_2 U_{jt} + \theta_t + \delta_j + \epsilon_{jt}$$

where $Quarter_t$ is the calendar quarter (of conception, for those who give birth) and $Quartile_k$ is an indicator that groups MTV or *16 and Pregnant* ratings by quartile. The other terms are the same as in equations (1) through (4) above.

We initially limit our analysis to the 18 quarters that KL use before *16 and Pregnant* began (2005:QI-2009:QII), omitting one quarter as the reference period. The coefficients $\alpha_1, \dots, \alpha_{17}$ therefore capture the trend in log teen birth rates over time for DMAs in the lowest

¹⁵ In Appendix Figures 1 through 4, we show results from the same event study analyses without using weights in the underlying regressions. Our conclusions remain the same.

quartile of MTV viewership. The β_{ik} coefficients capture differences between the quarterly time effects of the first quartile and the second, third, or fourth quartiles of MTV or *16 and Pregnant* ratings. If MTV viewing is a valid instrument, we should not reject the null hypothesis that the β 's are jointly equal to zero. We also test a simpler model where we examine whether the trends in DMAs that have ratings of MTV or *16 and Pregnant* that are above the respective medians differ from those below the median. As with the event study analysis, we extend the pre-period back to 1999 in increments.

Table 3 shows the results of testing whether the interaction terms between the quarterly fixed effects and quartile indicators for MTV ratings (Panel A) or *16 and Pregnant* ratings (Panel B) are jointly equal to zero. Each column in each panel is from a separate regression and the table shows the F statistics for the relevant joint hypothesis tests.¹⁶ In Panel A, column (1) we examine KL's 2005-2009 period and reject the joint hypothesis that that the trends for the second through fourth quartiles are equal to the main effects ($F=1.95, p<0.001$). We also reject the null hypothesis that the coefficients for the third quartile are equal to the main effects ($F=2.82, p=0.001$) as well that the DMAs above the median have the same effects as for below the median ($F=1.64, p=0.057$). In columns (2), (3), and (4) we extend the time period backwards to 2003, 2001, and 1999, respectively. In each case, we reject the null hypothesis that the quarterly effects are the same across the quartiles of MTV ratings.

In Panel B of Table 3 we perform the same exercise using quartiles of *16 and Pregnant* ratings instead of MTV ratings. Using KL's study period, in column (1) we reject the null

¹⁶ Strictly speaking, the test statistics based on the clustered and heteroskedasticity-consistent variance-covariance matrix are valid only asymptotically. As the number of degrees of freedom in the denominator of the F statistics goes to infinity, the reported test statistics multiplied by the numerator degrees of freedom will approach the limiting χ^2 distribution. The p -values of the tests shown are more conservative (larger) than from the corresponding asymptotically valid χ^2 statistics, however, and so we show these instead.

hypothesis that the period coefficients for the second through fourth quartiles are jointly equal to zero ($F=1.46$, $p=0.034$), as well as the second quartile ($F=1.70$, $p=0.045$) and third quartile ($F=2.35$, $p=0.003$) separately. We cannot, however, reject the null hypothesis that the pre-period trends differ above and below the median ratings of *16 and Pregnant*. In columns (2), (3), and (4), when we extend the sample period, the results become more consistent and similar to those using MTV ratings: we reject the null hypothesis that there is no difference in trends between the second, third, and fourth quartiles relative to the first quartile, as a group and individually. In results not shown in the table, we should also note that the coefficient on the unemployment rate remains stable and robust despite fitting non-parametric trends within quartiles of the ratings.

VI. Including DMA-Specific and MTV-Quartile-Specific Trends

Our previous results clearly indicate that the common trends assumption required for a causal interpretation of KL's results is not met. We now test the sensitivity of KL's IV results to controlling for differential trends by including a quadratic trend for each DMA in the analysis. Sensitivity of the results would indicate that KL's empirical strategy relies heavily on the assumption that *16 and Pregnant* viewing is independent of the disturbance term conditional on the other covariates rather than being an actual quasi-experimental research design. Allowing for differential trends across geographic units assumes knowledge of the correct functional form and is vulnerable to specification searches (see, for example Leora Freidberg 1998; Justin Wolfers 2006; and Kearney and Levine 2015a).

In Table 4, Panels A through D, we present IV results from the same four periods represented in Figures 2 through 5, respectively. Column (1) reproduces KL's specification from equation (3) without DMA-specific trends, so that Panel A, column (1) replicates KL's IV result

from their Table 1, column (3). In column (2) we augment their model with a quadratic trend for each DMA. Comparing columns (1) and (2) between the four panels indicates that controlling for DMA-specific trends leads to estimates of the impact of *16 and Pregnant* that are 32 to 80 percent smaller than when we do not control for those trends. In Panel C, for the period from 2001 to 2012, the estimated coefficient is not statistically significantly different from zero. The coefficients becoming smaller and sometimes statistically not significant underscores the importance of our earlier evidence that the parallel trends assumption does not hold, and shows that a substantial portion of the estimated effect of *16 and Pregnant* in KL's analysis can be accounted for by different trends in birth rates across DMAs.

In column (3) we allow for a more flexible first stage relationship by including indicators for each quartile of MTV viewing in 2008-2009 (as we did in Table 3) and in column (4) add to this specification DMA-specific quadratic trends. In column (3), the estimates are qualitatively similar to (although larger than) those that KL estimate regardless of the period under examination. When we include DMA-specific trends in column (4), however, only the 2005-2010 period in Panel A yields similar results to KL. When we lengthen the period of analysis in Panels B through D, including DMA-specific trends leads to estimates that are substantially smaller (in absolute value) and not statistically significant, again suggesting that DMA-specific trends are responsible for much of the effect of *16 and Pregnant* that KL estimate.

Adding DMA-specific trends to the model may be an over-parameterization of the model. To address this issue, in column (5) of Table 4 we reduce the number of trend parameters by including quadratic trends just for the four quartiles of MTV viewing in 2008-2009 while continuing to instrument for *16 and Pregnant* with the quartile dummy variables as in columns (3) and (4). As in column (4), we obtain results similar to KL's only in their sample period in

Panel A. Otherwise, the estimated impact of *16 and Pregnant* is not statistically significant. In addition, in Panels B through D we find that the quartile-specific trends are jointly statistically significant – again indicating that the assumptions for KL’s estimates to be interpreted causally are not met.¹⁷

To assess whether the difference-in-differences approach that includes DMA-specific trends is plausibly valid, in Tables 5 and 6 we repeat the placebo tests from Tables 1 and 2 while adding DMA-specific quadratic trends.¹⁸ As before, we should expect to find no relationship between MTV or *16 and Pregnant* and birth rates in the period before *16 and Pregnant* was actually broadcast. In Table 5, we find 4 of 19 (21 percent) of the IV estimates are statistically significant and several more IV estimates that are of the same approximate magnitude as the KL’s estimate of the impact of *16 and Pregnant* (-1.591). When we focus on a fixed pre-*16 and Pregnant* period in Table 6, we find 5 of 19 (26 percent) of the IV estimates are statistically significant and again several more that are of the same approximate magnitude as KL’s estimate of the effect of *16 and Pregnant*. We conclude from these results that even if one was to ignore the overwhelming evidence that KL’s approach is not quasi-experimental, but highly dependent on the conditional independence assumption of regression methods, the IV estimate of -1.591 from the difference-in-difference specification with DMA-specific trends is still not credible.

¹⁷ Appendix Table 4 presents the analogous estimates to Table 4 without using weights. In general the results that do not employ weights are more stable than those using weights and are always statistically significant. They are, however, often of implausibly large magnitudes. For example, the results in column (2), which is KL’s model with the addition of DMA-specific trends, suggest that *16 and Pregnant* could explain between 29 and 44 percent of the decline in teen births between 2008 and 2010. Moreover, as noted by Haider, Solon, and Wooldridge (2015), large differences between weighted and unweighted results is an indication of specification error.

¹⁸ Appendix Tables 5 and 6 show the unweighted versions of Tables 5 and 6.

VII. Social Media and the Causal Chain

KL use Twitter feeds and Google searches to provide evidence of a potential mechanism linking *16 and Pregnant* to lower teen birth rates. If tweets and Google searches about birth control and abortion allow teen girls to acquire new information that leads to increases in prevention or voluntary termination of pregnancies there would be, presumably, fewer teen births. By establishing a correlation between *16 and Pregnant* and these online activities, KL suggest that it is reasonable to infer that the program lowers teen birth rates. Although KL's results are somewhat mixed, they conclude that there is sufficient evidence to support a causal link.

In Table 7, we present results from time series analyses of national Twitter activity about birth control and abortion related to measures of potential exposure to *16 and Pregnant*. In Panel A we follow KL and present results using the “in season” period, comprising only days in which a new season of *16 and Pregnant* was being broadcast weekly. Columns (1) and (4) of panel A replicate KL's results using Twitter from their Tables 3 and 4 in which they use the log tweet rate about birth control and abortion, respectively, as dependent variables. In part (i) of Panel A we replicate KL's estimate that the tweets about birth control increase 12 percent on the first day an episode is broadcast and 23 percent on the next day relative to the other five days of the week after the broadcast. In columns (1) and (4) of Panel A, part (ii), we replicate KL's results in which they use the log tweet rate about *16 and Pregnant* as a proxy for exposure to the show.¹⁹

¹⁹ About this specification KL write: “The 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 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” (p. 3621)

They also find here a statistically significant relationship between *16 and Pregnant* and Twitter activity about birth control and abortion.

In all of these regressions, KL use the total number of daily tweets as weights. This means that days with increased tweeting will receive more weight. For example, Michael Jackson died on 25 June 2009, the same day that “Amber,” the third episode of season 1 of *16 and Pregnant*, aired. This date would get more weight in KL’s regressions if Michael Jackson’s death caused a spike in Twitter activity. In columns (2) and (5) of Panel A, we re-estimate all four of the regressions in columns (1) and (4) of Panel A without using weights and report standard errors that are calculated using the method of Whitney K. Newey and Kenneth D. West (1987) to allow for two autoregressive lags in the disturbance.²⁰ Merely by not weighting, we find that the relationship between tweets about birth control is no longer statistically significantly related to *16 and Pregnant* using either the broadcast indicators or tweets.

Two thirds of *16 and Pregnant* episodes were first broadcast on Tuesdays with the other episodes distributed between Sundays, Wednesdays, Thursdays and Fridays. There is, therefore, sufficient variation in the broadcast schedule to control day-of-week effects, which possibly arise because, for example, people may tweet less when they are at work. There may also be seasonality in Twitter activity and we therefore explore whether KL’s results are sensitive to controlling for these day of the week and month of year differences. In columns (3) and (6) of Panel A, we repeat the analysis from columns (1) and (4) (including weighting by the number of daily tweets), but add day-of-the week and month-of-year fixed effects. Including these controls

²⁰ We calculated Durbin-Watson statistics for the models in columns (1) and (3) of Panels A and B and in all cases easily rejected the null hypothesis of no serial correlation in the residuals. These results are available from the authors by request. Allowing for additional lags in the Newey-West procedure did not substantively change the estimated standard errors.

causes any association between tweeting about birth control and abortion both measures to disappear.

KL's sample in these results is limited to the 336 days in which *16 and Pregnant* was "in season" out of the 1,455 available days with positive tweets between 1 January 2009 and 31 December 2012 as presented in KL's Figure 7.²¹ Days in which *16 and Pregnant* was not "in-season" arguably provide a better baseline for ongoing tweet activity related to contraception and abortion. In Panel B, columns (1), (2), (4) and (5) of Table 6 we show estimates from the exact same specifications as in in Panel A, but use all of the available data. In part (i) of Panel B, the coefficients on the indicator for days with episodes of *16 and Pregnant*, or just after, now have the wrong sign, are of small magnitude, and/or are statistically insignificant. In part (ii), we find a modest association between tweets about *16 and Pregnant* and that for birth control if we use weights, but the rest of the estimates in part (ii) of Panel B do not reflect any association between *16 and Pregnant* and tweets about birth control or abortion.

In columns (3) and (6) of Panel B part (i), we now include an indicator for the day an episode was broadcast along with six lags for the subsequent days along with the day-of-week and month-of-year fixed effects as in Panel B. In addition to the coefficients on the first and second of these seven indicators, we also show their average and its standard error. The results indicate that daily tweets for birth control and abortion are, on average, between 10 and 13 percent *lower* during a season of *16 and Pregnant* than when a new season is not being broadcast. In Panel B(ii), we repeat the exercise from Panel A(ii) but using the full period and find no association between the log tweet rate for *16 and Pregnant* and the tweet rate for birth control and abortion. Rather than bolstering their claim that *16 and Pregnant* caused teen birth

²¹ There are 1,461 total days in this period, but KL use log tweet rates and 6 days are lost due to zero tweets.

rates to fall, their use of social media raises additional doubts about the credibility of their conclusions.²²

VIII. Discussion and Conclusion

The evidence presented by KL is inadequate to conclude anything about the causal effect of *16 and Pregnant* on teen birth rates. Their analysis fundamentally rests on the assumption that areas in which young adults watched less MTV prior to *16 and Pregnant* were no different with regard to trends in teen birth rates than areas in which MTV was more popular. We find that there *were* differences between these areas. This is seen most clearly in the placebo tests in which we found an “effect” of *16 and Pregnant* before it existed, strongly suggesting that the parallel trends assumption necessary for a causal interpretation of KL’s results does not hold. The sensitivity of the event study estimates and tests of pre-show trends when we extend the study period in short increments reinforce this result. A similar issue arises with their results using Twitter feeds. By only analyzing days during a season of *16 and Pregnant*, they neglect to exploit an arguably more appropriate baseline of tweet activity. When we control for day of week and month fixed effects, KL’s social media results disappear.

What may have affected the acceleration in the decline in birth rates that begin in 2008, if not *16 and Pregnant*? The obvious candidate is the Great Recession and its aftermath. Sobotka, Skirbekk and Philipov (2011) show that younger women’s fertility is more to recessions than responsive to recessions than older women’s. Goldstein, et al. (2013) demonstrate this relationship help in Europe during the Great Recession. Cherlin, et al. (2013) confirm that this

²² KL’s last effort to link *16 and Pregnant* with online activity uses within-state variation between the pre-*16 and Pregnant* and post-*16 and Pregnant* periods. This exercise suffers similar sensitivity to selected samples and weighting as their analysis of teen birth rates. Results are available from the authors by request.

relationship also held in the U.S. during that time and show that that poorer women and Hispanic women experienced especially sharp declines at the Great Recession's onset. Although KL control for the overall unemployment rate, it is possible that using a more refined measure of unemployment for teen woman would alter their results.²³ There also may be meaningful interactions between changes in unemployment, longer schooling durations, and changes in more effective contraception.

Are there general lessons to be learned from this exercise beyond demonstrating that causal claims about *16 and Pregnant's* impact on birth rates are not justified? We believe that there are. Given the importance of the parallel trends assumption in differences-in-differences studies, studies that use geographic variation to identify a causal effect should clearly test the robustness of the parallel trends assumption over long and varying periods. The parallel trend assumption is not specific to particular time periods and difference-in-difference estimates should be invariant to modest extensions to the periods of analysis. Only when the assumption is met over substantial and varying periods can the resulting estimates be given a causal interpretation. Perhaps more importantly, our revisiting KL's analysis adds to the growing evidence in economics and other social sciences that replication is important and necessary. Without replication, the myriad problems with KL's analyses would not have come to light and there would be no opportunity to correct the public record about the effect of reality television on

²³ Data from the Current Population Survey indicate that the unemployment rate for 15 to 24 year-old, non-Hispanic whites, non-Hispanic blacks and Hispanics increased by 8.8, 10.5 and 15.8 percentage points, respectively, between 2007 and 2010. These are substantial increases. Calculations available from the authors by request.

teen reproductive activity.²⁴ Because KL's study attracted extensive media coverage, policymakers may believe that "nudges" like those represented by *16 and Pregnant* are effective when, at least in this case, no causal link has been proven. Getting the answer right, which depends on both revisiting analyses with the original data and replication of the "experiment" in different contexts, should have a higher priority in economics and social sciences journals.

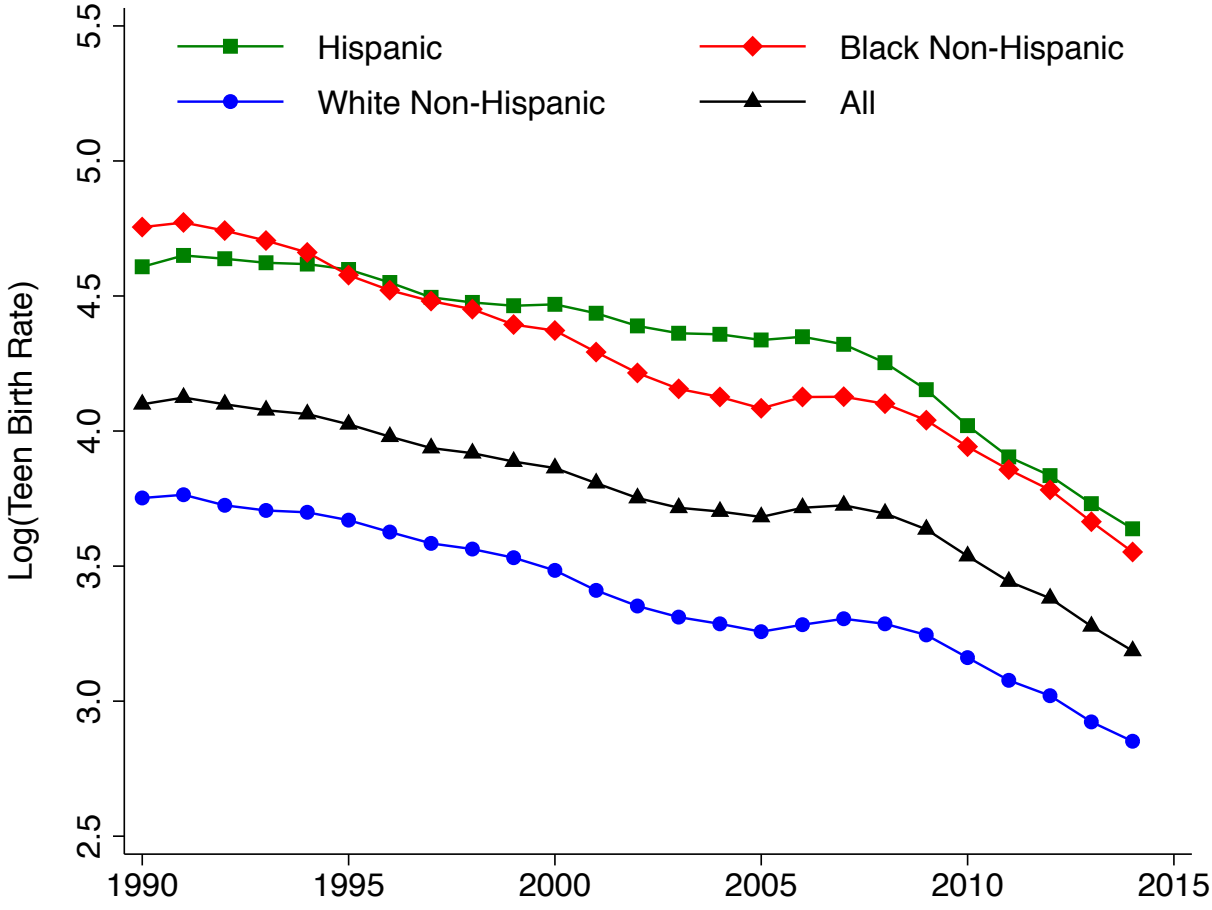
²⁴ We want to emphasize that such replication exercises require the cooperation of the authors of the original paper. Phil Levine was exemplary in providing us with programs and responding to our many requests while we were trying to reproduce KL's results.

References

- Ananat, Elizabeth Oltmans, Anna Gassman-Pines, and Christina Gibson-Davis.** 2013. "Community-Wide Job Loss and Teenage Fertility: Evidence From North Carolina." *Demography* 50 (6): 2151–2171.
- Benjamini, Yoav, and Yosef Hochberg.** 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society. Series B (Methodological)* 57 (1): 289–300.
- Bonferroni Carlo E.** 1936. "Teoria statistica delle classi e calcolo delle probabilità." *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze* 8: 3–62.
- Boonstra, Heather.** 2014. "What Is Behind the Declines in Teen Pregnancy Rates?" *Guttmacher Policy Review* 17 (3): 15–21.
- Cherlin, Andrew, Erin Cumberworth, S. Philip Morgan, and Christopher Wimer.** 2013. "The Effects of the Great Recession on Family Structure and Fertility." *The Annals of the American Academy of Political and Social Science* 650 (1): 214–231.
- Dehejia, Rajeev, and Adriana Lleras-Muney.** 2004. "Booms, Busts, and Babies' Health." *The Quarterly Journal of Economics* 119 (3): 1091–1130.
- DellaVigna, Stefano, and Ethan Kaplan.** 2007. "The Fox News Effect: Media Bias and Voting." *The Quarterly Journal of Economics* 122 (3): 1187–1234.
- Friedberg, Leora.** 1998. "Did Unilateral Divorce Raise Divorce Rates? Evidence from Panel Data." *The American Economic Review* 88 (3): 608–627.
- Gentzkow, Matthew.** 2006. "Television and Voter Turnout." *The Quarterly Journal of Economics* 121 (3): 931–972.
- Gentzkow, Matthew, and Jesse M. Shapiro.** 2008. "Preschool Television Viewing and Adolescent Test Scores: Historical Evidence from the Coleman Study." *The Quarterly Journal of Economics* 123 (1): 279–323.
- Goldstein, Joshua, Michaela Kreyenfeld, Aiva Jasilioniene, and Deniz Dilan Karaman Örsal.** 2013. "Fertility Reactions to the 'Great Recession' in Europe: Recent Evidence from Order-Specific Data." *Demographic Research* 29 (4): 85–104.
- Hamilton, Brady E, Martin, Joyce A., Michelle J.K. Osterman, Sally C. Curtin, and T.J. Matthews.** 2015. *Births: Final Data for 2014. National Vital Statistics Reports.* Volume 64, Number 12. Hyattsville, MD: National Center for Health Statistics.
- Holm, Sture.** 1979. "A Simple Sequentially Rejective Multiple Test Procedure." *Scandinavian Journal of Statistics* 6 (2): 65–70.
- Jensen, Robert, and Emily Oster.** 2009. "The Power of TV: Cable Television and Women's Status in India." *The Quarterly Journal of Economics* 124 (3): 1057–1094.
- Kaestner, Robert, Sanders Korenman, and June O'Neill.** 2003. "Has Welfare Reform Changed Teenage Behaviors?" *Journal of Policy Analysis and Management* 22 (2): 225–248.

- Kearney, Melissa S., and Phillip B. Levine.** 2014a. *Media Influences on Social Outcomes: The Impact of MTV's 16 and Pregnant on Teen Childbearing.* Working Paper 19795. National Bureau of Economic Research. <http://www.nber.org/papers/w19795>.
- 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.** 2015a. "Investigating Recent Trends in the U.S. Teen Birth Rate." *Journal of Health Economics* 41 (May): 15–29.
- Kearney, Melissa S., and Phillip B. Levine.** 2015b. *Early Childhood Education by MOOC: Lessons from Sesame Street.* Working Paper 21229. National Bureau of Economic Research. <http://www.nber.org/papers/w21229>.
- Kearney, Melissa S., and Phillip B. Levine.** 2015c. "Media Influences on Social Outcomes: The Impact of MTV's 16 and Pregnant on Teen Childbearing." *American Economic Review* 105 (12): 3597–3632.
- 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.
- Lindberg, Laura, John Santelli, and Sheila Desai.** forthcoming. "Understanding the Decline in Adolescent Fertility in the United States, 2007-2012." *Journal of Adolescent Health.*
- Lopoo, Leonard M., and Thomas DeLeire.** 2006. "Did Welfare Reform Influence the Fertility of Young Teens?" *Journal of Policy Analysis and Management* 25 (2): 275–298.
- National Research Council.** Panel on Adolescent Pregnancy and Childbearing. 1987. *Risking the Future: Adolescent Sexuality, Pregnancy, and Childbearing.* Edited by Cheryl D. Hayes and Sandra L. Hofferth. 2 vols. Washington, D.C.: National Academy Press.
- Newey, Whitney K., and Kenneth D. West.** 1987. "Hypothesis Testing with Efficient Method of Moments Estimation." *International Economic Review* 28 (3): 777–787.
- Santelli, John S., Laura Duberstein Lindberg, Lawrence B. Finer, and Susheela Singh.** 2007. "Explaining Recent Declines in Adolescent Pregnancy in the United States: The Contribution of Abstinence and Improved Contraceptive Use." *American Journal of Public Health* 97 (1): 150–156.
- Sobotka, Tomáš, Vegard Skirbekk, and Dimiter Philipov.** 2011. "Economic Recession and Fertility in the Developed World." *Population and Development Review* 37 (2): 267–306.
- Solon, Gary, Steven J. Haider, and Jeffrey M. Wooldridge.** 2015. "What Are We Weighting For?" *Journal of Human Resources* 50 (2): 301–316.
- Trudeau, Jennifer.** 2016. "The Role of New Media on Teen Sexual Behaviors and Fertility Outcomes—The Case of 16 and Pregnant." *Southern Economic Journal* 82(3): 975-1003.
- Wolfers, Justin.** 2006. "Did Unilateral Divorce Laws Raise Divorce Rates? A Reconciliation and New Results." *American Economic Review* 96(5): 1802-1827.

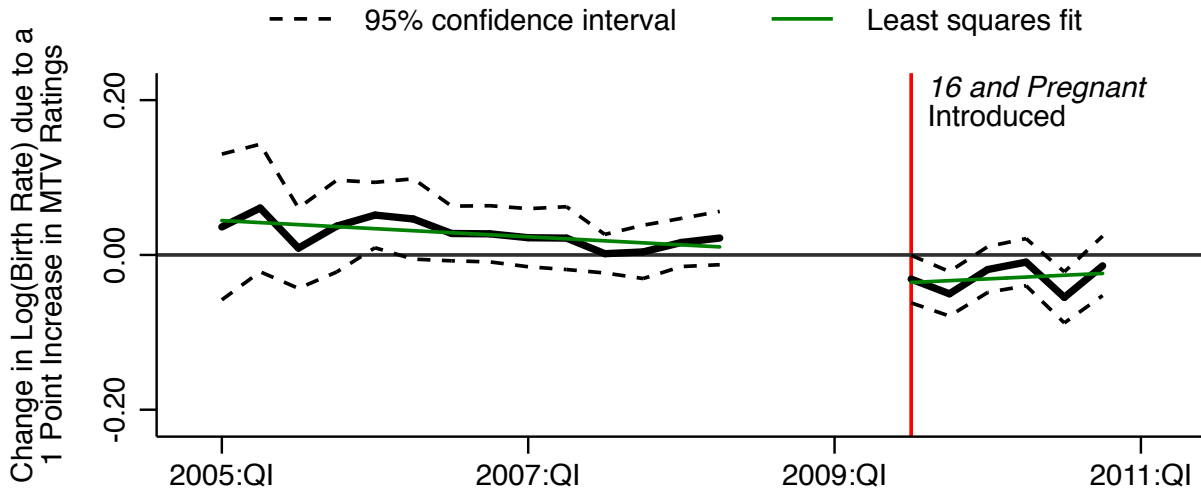
Figure 1
Log(Teen Birth Rate) for Race/Ethnicity Groups



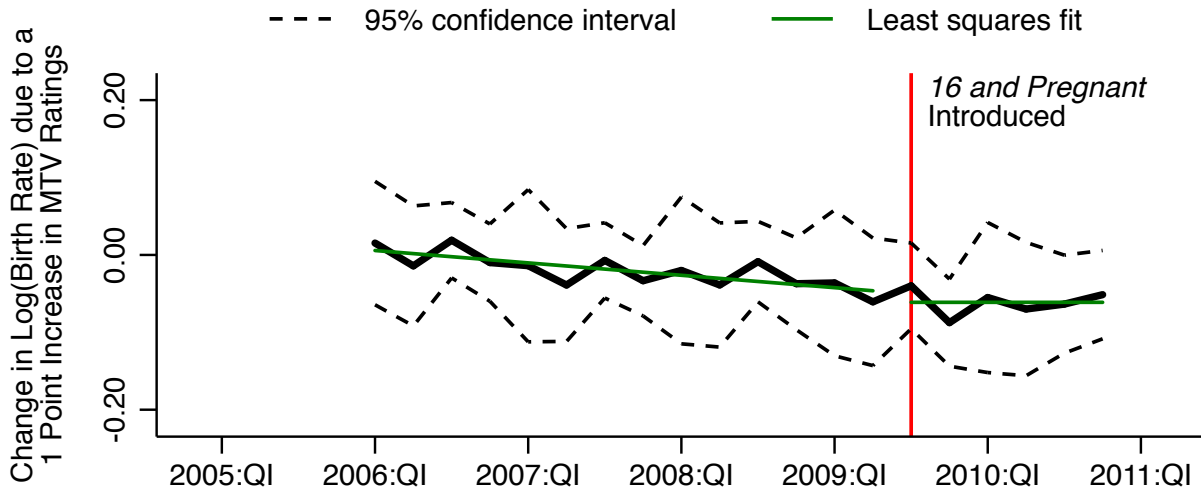
Source: Hamilton, et al. (2015)

Figure 2
Reduced Form Event Study: 2005-2010

Panel A: Reference period is 2008:QIII-2009:QII (Replicating Fig. 5 from KL)



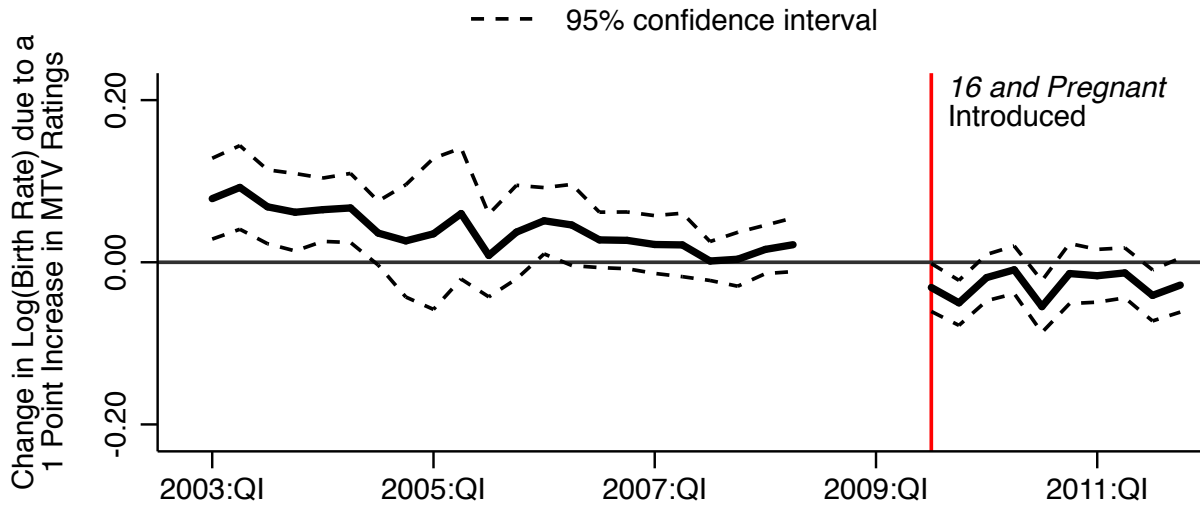
Panel B: Reference Period is 2005:Q1-2005:QIV



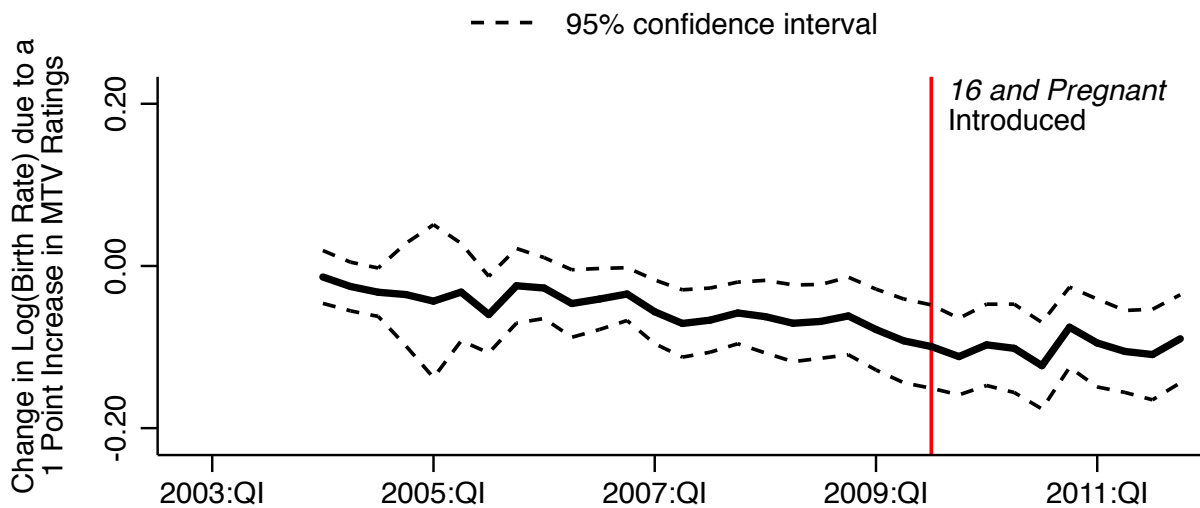
Notes: These figures plot the coefficients of average MTV Ratings in 2008-2009 interacted with period dummy variables. The regressions also include the unemployment rate and a full set of period dummy variables and DMA \times season interactions. Confidence intervals are calculated by clustering by DMA. Regressions are weighted by the female population aged 15-19 in the DMA. Panel A uses 2008:QIII through 2009:QII as the reference period. Panel B uses 2005:Q1 through 2005:QIV as the reference period. Least squares lines in both panels are fitted separately for the pre- and post-16 and Pregnant periods. The F statistic for the test of joint significance of the coefficients in pre- and post-16 and Pregnant period is 1.46 with a p -value of 0.128. Panel A replicates Kearney and Levine's (2015b, KL) Figure 5.

Figure 3
Reduced Form Event Study: 2003-2011

Panel A: Reference period is 2008:QIII-2009:QII



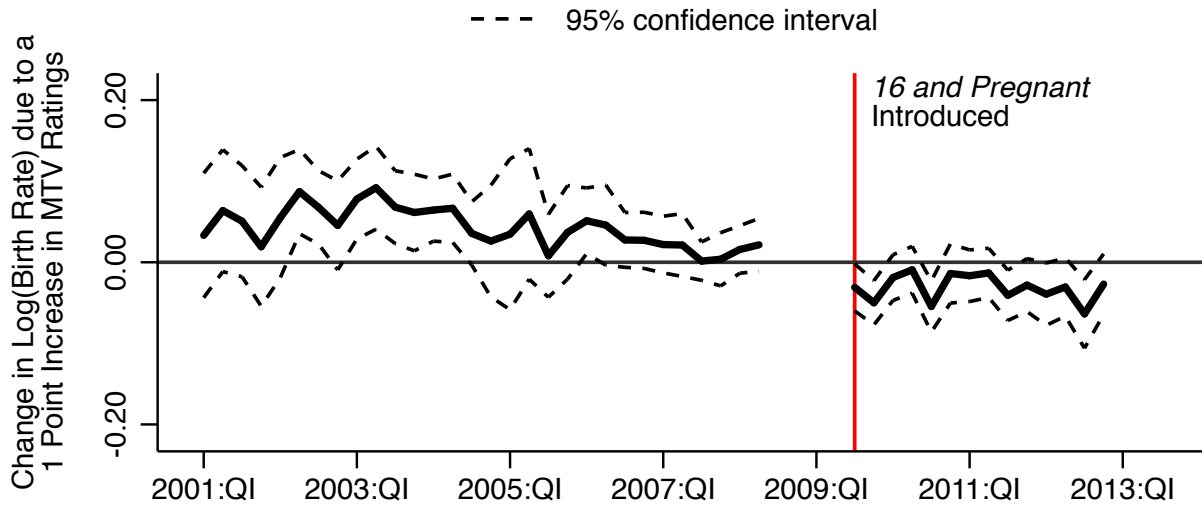
Panel B: Reference period is 2003:Q1-2003:QIV



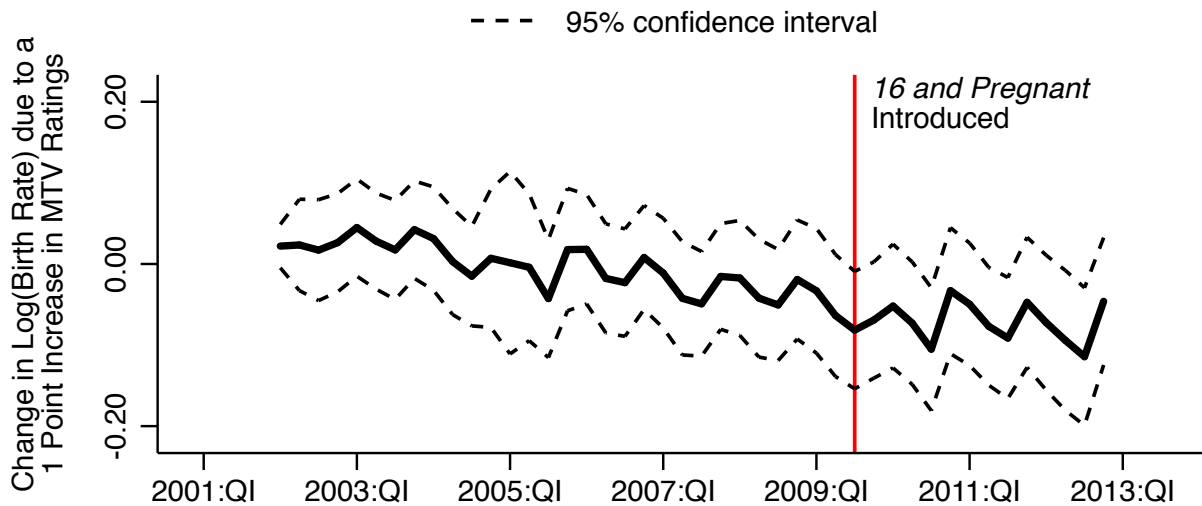
Notes: These figures plot the coefficients of average MTV Ratings in 2008-2009 interacted with period dummy variables. The regressions also include the unemployment rate and a full set of period dummy variables and DMA \times season interactions. Confidence intervals are calculated by clustering by DMA. Regressions are weighted by the female population aged 15-19 in the DMA. Panel A uses 2008:QIII through 2009:QII as the reference period. Panel B uses 2003:Q1 through 2003:QIV as the reference period. Least squares lines in both panels are fitted separately for the pre- and post-*16 and Pregnant* periods. The F statistic for the test of joint significance of the coefficients in the pre-*16 and Pregnant* period is 2.13 with a p -value of 0.003.

Figure 4
Reduced Form Event Study: 2001-2012

Panel A: Reference period is 2008:QIII-2009:QII



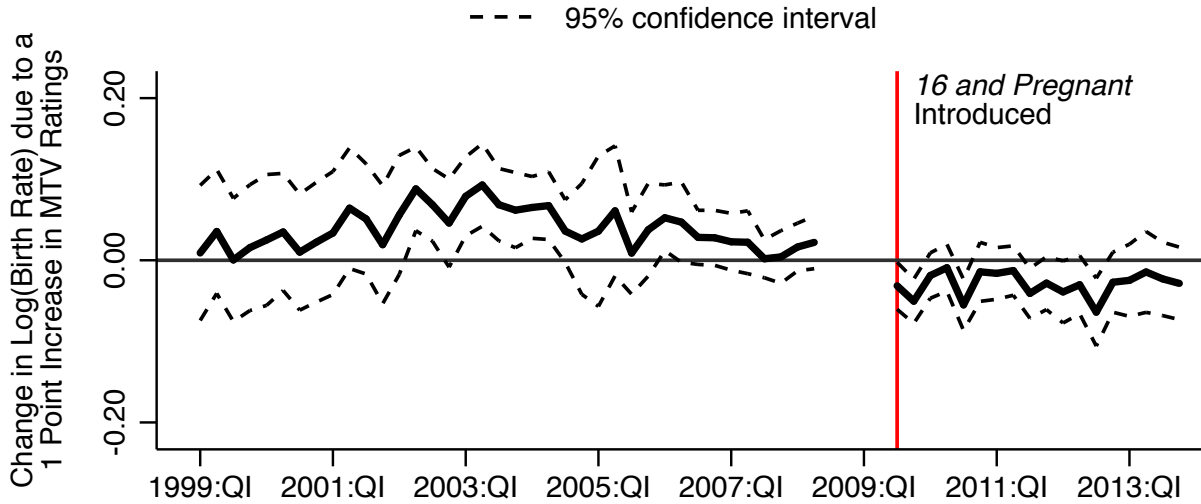
Panel B: Reference period is 2001:Q1-2001:QIV



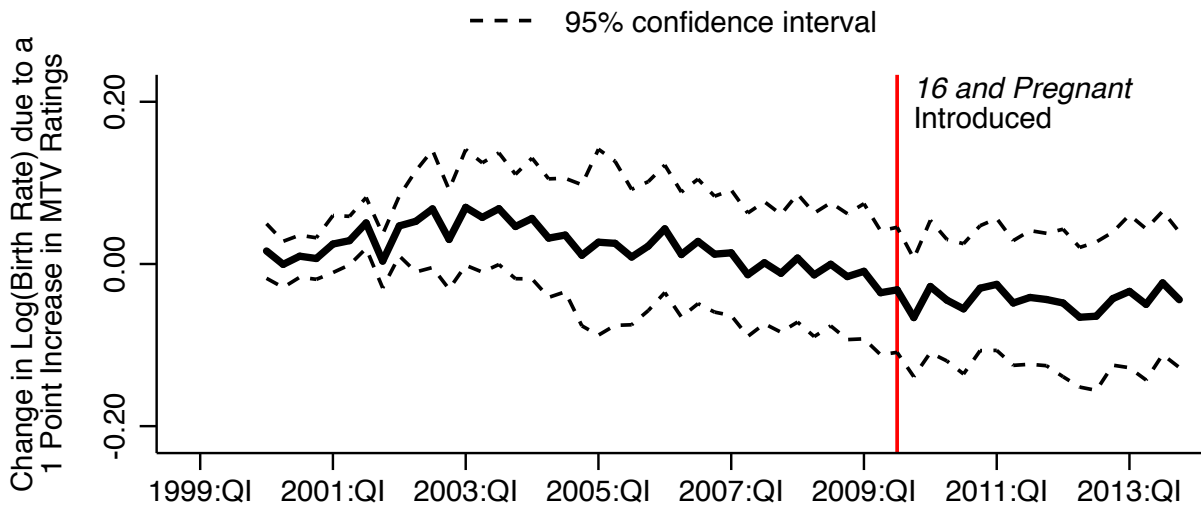
Notes: These figures plot the coefficients of average MTV Ratings in 2008-2009 interacted with period dummy variables. The regressions also include the unemployment rate and a full set of period dummy variables and DMA \times season interactions. Confidence intervals are calculated by clustering by DMA. Regressions are weighted by the female population aged 15-19 in the DMA. Panel A uses 2008:QIII through 2009:QII as the reference period. Panel B uses 2001:Q1 through 2001:QIV as the reference period. Least squares lines in both panels are fitted separately for the pre- and post-*16 and Pregnant* periods. The F statistic for the test of joint significance of the coefficients in the pre-*16 and Pregnant* period is 2.83 with a p -value of <0.001 .

Figure 5
Reduced Form Event Study: 1999-2013

Panel A: Reference period is 2008:QIII-2009:QII



Panel B: Reference period is 1999:Q1-1999:QIV



Notes: These figures plot the coefficients of average MTV Ratings in 2008-2009 interacted with period dummy variables. The regressions also include the unemployment rate and a full set of period dummy variables and DMA \times season interactions. Confidence intervals are calculated by clustering by DMA. Regressions are weighted by the female population aged 15-19 in the DMA. Panel A uses 2008:QIII through 2009:QII as the reference period. Panel B uses 1999:Q1 through 1999:QIV as the reference period. Least squares lines in both panels are fitted separately for the pre- and post-*16 and Pregnant* periods. The F statistic for the test of joint significance of the coefficients in the pre-*16 and Pregnant* period is 3.13 with a p -value of <0.001 .

Table 1
 Placebo Tests of Estimated Reduced Form and Instrumental Variables Impact on Teen Birth Rates
 Rolling 24 Quarter Periods

Row	Dates			Reduced Form		Instrumental Variables	
	Begin	"Show" Start	End	Coefficient	Std. Err.	Coefficient	Std. Err.
(1)	1999:QI	2003:QIII	2004:QIV	0.397	2.061	0.268	1.262
(2)	1999:QII	2003:QIV	2005:QI	-0.683	2.298	-0.463	1.403
(3)	1999:QIII	2004:QI	2005:QII	-1.604	2.588	-1.091	1.554
(4)	1999:QIV	2004:QII	2005:QIII	-2.692	2.813	-1.839	1.661
(5)	2000:QI	2004:QIII	2005:QIV	-2.681	3.042	-1.840	1.806
(6)	2000:QII	2004:QIV	2006:QI	-2.621	2.865	-1.789	1.684
(7)	2000:QIII	2005:QI	2006:QII	-2.903	2.303	-1.977	1.353
(8)	2000:QIV	2005:QII	2006:QIII	-2.675 *	1.570	-1.808 *	0.954
(9)	2001:QI	2005:QIII	2006:QIV	-2.437 *	1.324	-1.639 *	0.846
(10)	2001:QII	2005:QIV	2007:QI	-1.994	1.299	-1.330	0.837
(11)	2001:QIII	2006:QI	2007:QII	-2.554 *	1.441	-1.693 *	0.948
(12)	2001:QIV	2006:QII	2007:QIII	-2.929 **	1.473	-1.936 **	0.977
(13)	2002:QI	2006:QIII	2007:QIV	-2.968 *	1.521	-1.958 **	0.992
(14)	2002:QII	2006:QIV	2008:QI	-3.144 **	1.569	-2.068 **	1.028
(15)	2002:QIII	2007:QI	2008:QII	-3.341 **	1.508	-2.194 **	1.014
(16)	2002:QIV	2007:QII	2008:QIII	-3.079 **	1.458	-2.021 **	0.966
(17)	2003:QI	2007:QIII	2008:QIV	-2.756 *	1.476	-1.809 *	0.967
(18)	2003:QII	2007:QIV	2009:QI	-2.543 *	1.485	-1.673 *	0.969
(19)	2003:QIII	2008:QI	2009:QII	-2.377	1.558	-1.566	1.002
(20)	2003:QIV	2008:QII	2009:QIII	-2.442	1.580	-1.610	1.006
(21)	2004:QI	2008:QIII	2009:QIV	-2.978 *	1.604	-1.964 *	1.029
(22)	2004:QII	2008:QIV	2010:QI	-3.320 **	1.663	-2.192 **	1.074
(23)	2004:QIII	2009:QI	2010:QII	-3.589 **	1.639	-2.372 **	1.040
(24)	2004:QIV	2009:QII	2010:QIII	-3.898 **	1.611	-2.577 **	1.024
(25)	2005:QI	2009:QIII	2010:QIV	-3.581 **	1.517	-2.368 **	0.942
(26)	2005:QII	2009:QIV	2011:QI	-3.496 **	1.444	-2.318 ***	0.888
(27)	2005:QIII	2010:QI	2011:QII	-2.624 *	1.343	-1.743 **	0.802
(28)	2005:QIV	2010:QII	2011:QIII	-2.439 *	1.285	-1.620 **	0.773
(29)	2006:QI	2010:QIII	2011:QIV	-2.173 *	1.227	-1.443 *	0.744
(30)	2006:QII	2010:QIV	2012:QI	-1.765	1.220	-1.174	0.735
(31)	2006:QIII	2011:QI	2012:QII	-1.992	1.270	-1.328 *	0.774
(32)	2006:QIV	2011:QII	2012:QIII	-1.843	1.288	-1.229	0.792
(33)	2007:QI	2011:QIII	2012:QIV	-1.360	1.305	-0.908	0.806
(34)	2007:QII	2011:QIV	2013:QI	-1.041	1.368	-0.695	0.844
(35)	2007:QIII	2012:QI	2013:QII	-0.612	1.353	-0.409	0.831
(36)	2007:QIV	2012:QII	2013:QIII	0.237	1.343	0.158	0.810

Notes: Entries in the table are a) in the reduced form, the estimated coefficient on MTV Ratings in 2008:QIII-2009:QII interacted with a dummy variable for being in the "post" period and b) for instrumental variables, the estimated coefficient on *16 and Pregnant* Ratings interacted with a dummy variable for being in the "post" period where the instrument is the regressor of interest from the reduced form regressions. Standard errors, clustered by DMA, are shown in parentheses. All regressions are weighted by the female population aged 15-19 in the DMA at the time of the observation. In each regression there are 18 pre-"show" quarters and 6 post-"show" quarters. All regressions also include the unemployment rate, the percent of the population that is non-Hispanic black and the percent of the population that is Hispanic, 24 quarter fixed effects as well DMA \times season fixed effects as regressors. The period analyzed by Kearney and Levine's (2015b) is boxed. In the periods that are lightly shaded, the placebo post period partially includes the actual post-*16 and Pregnant* period. In the periods that are darkly shaded, both the placebo pre- and the placebo post-"show" period include the actual post-16 and Pregnant period. Sample size in rows (1) through (17), (19) through (21), and (36) is 4,918, the sample size in row (18) is 4,917, and the sample size in rows (22) through (35) is 4,919. *** indicates significant at the 1 percent level, ** indicates significant at the 5 percent level, * indicates significant at the 10 percent level.

Table 2

Placebo Tests of Reduced Form and Instrumental Variables Estiamtes of Impact on Teen Birth Rates
Fixed for the Entire Pre-16 and Pregnant Period

Row	Placebo "Show" Start	Reduced Form		Instrumental Variables	
		Coefficient	Std. Err.	Coefficient	Std. Err.
(1)	2003:QIII	-1.343	2.680	-2.067 **	0.985
(2)	2003:QIV	-1.731	2.621	-2.106 **	0.981
(3)	2004:QI	-2.056	2.575	-2.001 **	1.002
(4)	2004:QII	-2.337	2.491	-1.846 *	0.987
(5)	2004:QIII	-2.516	2.421	-1.495	0.995
(6)	2004:QIV	-2.622	2.387	-1.359	1.030
(7)	2005:QI	-2.644	2.300	-1.212	1.053
(8)	2005:QII	-2.666	2.191	-0.833	1.006
(9)	2005:QIII	-2.818	2.187	-0.698	0.988
(10)	2005:QIV	-2.685	2.199	-0.524	0.968
(11)	2006:QI	-2.873	2.227	-0.791	0.943
(12)	2006:QII	-3.160	2.219	-0.877	0.859
(13)	2006:QIII	-3.258	2.189	-0.908	0.785
(14)	2006:QIV	-3.453	2.181	-0.942	0.720
(15)	2007:QI	-3.682 *	2.168	-1.179 *	0.678
(16)	2007:QII	-3.752 *	2.138	-1.340 **	0.663
(17)	2007:QIII	-3.629 *	2.113	-1.708 ***	0.636
(18)	2007:QIV	-3.597 *	2.131	-1.909 ***	0.740
(19)	2008:QI	-3.627 *	2.161	-2.052	1.328

Notes: Entries in the table are a) for the reduced form, the estimated coefficient on MTV Ratings in 2008:QIII-2009:QII interacted with a dummy variable for being in the "post" period and b) for instrumental variables, the estimated coefficient on *16 and Pregnant* Ratings interacted with a dummy variable for being in the "post" period where the instrument is the regressor of interest from the reduced form regressions. Standard errors, clustered by DMA, are shown in parentheses. All regressions are weighted by the female population aged 15-19 in the DMA at the time of the observation. Each regression uses the period 1999:QI to 2009:QII. The only difference in each regression is the quarter in which the placebo "show" begins. All regressions also include the unemployment rate, the percent of the population that is non-Hispanic blackk and the percent of the population that is Hispanic, 24 quarter fixed effects as well DMA \times season fixed effects as regressors. The first stage relationship between MTV ratings and *16 and Pregnant* Ratings ranges from 0.273 (row 13) to 1.539 (row 1). Sample size for all regressions is 8,607. *** indicates significant at the 1 percent level, ** indicates significant at the 5 percent level, * indicates significant at the 10 percent level.

Table 3
Differential Trends in Teen Birth Rates during the Pre-16 and Pregnant Period by Quartiles of MTV Ratings and 16 and Pregnant Ratings

Period:	(1) 2005:Q1-2009:QII	(2) 2003:Q1-2009:QII	(3) 2001:Q1-2009:QII	(4) 1999:Q1-2009:QII
<i>Panel A: Quartiles of MTV Ratings</i>				
<i>F</i> : I(Second quartile), I(Third quartile), and I(Fourth quartile) × Time Dummies = 0	1.953 *** [<0.001]	2.954 *** [<0.001]	4.317 *** [<0.001]	6.665 *** [<0.001]
<i>F</i> : I(Second quartile) × Time Dummies = 0	1.286 [0.204]	1.453 * [0.083]	1.580 ** [0.030]	1.544 ** [0.027]
<i>F</i> : I(Third quartile) × Time Dummies = 0	2.817 *** [0.001]	3.544 *** [<0.001]	3.248 *** [<0.001]	4.579 *** [<0.001]
<i>F</i> : I(Fourth quartile) × Time Dummies = 0	1.322 [0.181]	3.300 *** [<0.001]	4.175 *** [<0.001]	4.352 *** [<0.001]
<i>F</i> : I(Median or above) × Time Dummies = 0	1.640 * [0.057]	1.844 ** [0.011]	1.852 *** [0.005]	1.625 ** [0.015]
<i>Panel B: Quartiles of 16 and Pregnant Ratings</i>				
<i>F</i> : I(Second quartile), I(Third quartile), and I(Fourth quartile) × Time Dummies = 0	1.464 ** [0.034]	2.041 *** [<0.001]	3.680 *** [<0.001]	5.355 *** [<0.001]
<i>F</i> : I(Second quartile) × Time Dummies = 0	1.699 ** [0.045]	1.862 *** [0.010]	2.144 *** [0.001]	2.868 *** [<0.001]
<i>F</i> : I(Third quartile) × Time Dummies = 0	2.347 *** [0.003]	2.403 *** [<0.001]	3.124 *** [<0.001]	3.594 *** [<0.001]
<i>F</i> : I(Fourth quartile) × Time Dummies = 0	1.292 [0.200]	2.158 *** [0.002]	3.568 *** [<0.001]	3.891 *** [<0.001]
<i>F</i> : I(Median or above) × Time Dummies = 0	0.971 [0.493]	0.933 [0.559]	1.574 ** [0.031]	1.625 ** [0.015]
Number of time periods	14	26	34	42
Sample Size	3,689	5,327	6,967	8,607

Notes: Entries in the table are *F* statistics. *p*-values in brackets. Variance-covariance matrix of estimates calculated by clustering by DMA. Regressions are weighted by the female population aged 15 to 19 in each DMA. Estimates in each column of each panel are from a separate regression. *F*-statistics are from tests that the coefficients on the interactions between quarterly indicators and indicators for the relevant quartiles of either 16 and Pregnant ratings or MTV ratings are jointly equal to zero. All models also include DMA fixed effects, the percent of black non-Hispanics, the percent of Hispanics in the DMA, and the unemployment rate as regressors. The results in column (1) use the same pre-16 and Pregnant period as Kearney and Levine (2015b). *** indicates significant at the 1 percent level, ** indicates significant at the 5 percent level, * indicates significant at the 10 percent level.

Table 4
Instrumental Variables Estimates of Impact of *16 and Pregnant* on Birth Rates
with DMA-Specific or MTV-Quartile-Specific Trends

	(1)	(2)	(3)	(4)	(5)
<i>Panel A: 2005:Q1-2010:QIV</i>					
<i>16 and Pregnant</i> Ratings	-2.368 ** (0.942)	-1.591 ** (0.756)	-2.741 ** (1.131)	-2.046 ** (0.899)	-2.022 ** (0.914)
$\chi^2(6)$ test: MTV Quartile-Specific Trends=0					10.31 [0.112]
<i>Panel B: 2003:Q1-2011:QIV</i>					
<i>16 and Pregnant</i> Ratings	-3.090 *** (1.064)	-1.736 ** (0.714)	-3.883 *** (1.249)	-1.227 (0.985)	-0.993 (1.030)
$\chi^2(6)$ test: MTV Quartile-Specific Trends=0					11.20 [0.082]
<i>Panel C: 2001:Q1-2012:QIV</i>					
<i>16 and Pregnant</i> Ratings	-3.704 *** (1.314)	-0.755 (0.738)	-4.802 *** (1.560)	0.214 (1.132)	0.243 (1.134)
$\chi^2(6)$ test: MTV Quartile-Specific Trends=0					12.47 [0.052]
<i>Panel D: 1999:Q1-2013:QIV</i>					
<i>16 and Pregnant</i> Ratings	-3.561 ** (1.489)	-1.354 ** (0.683)	-4.957 *** (1.849)	-0.556 (1.146)	-0.608 (1.111)
$\chi^2(6)$ test: MTV Quartile-Specific Trends=0					13.15 [0.041]
<i>Control Variables</i>					
DMA-Specific Quadratic Trends	no	yes	no	yes	no
MTV Quartile-Specific Quadratic Trends	no	no	no	no	yes
<i>Instruments</i>					
MTV \times Post	yes	yes	no	no	no
I(MTV Quartile) \times Post	no	no	yes	yes	yes

Notes: Standard errors, clustered by DMA, are shown in parentheses. p -values of χ^2 are shown in brackets. All regressions are weighted by the female population aged 15-19 at the time of the observation. Estimates in each column of each panel are from separate regressions. Panel A, Column (1) replicates the IV results in Table 1, Column (3) of Kearney and Levine (2015b). All models also include the unemployment rate, the percent of the population that is non-Hispanic black and the percent of the population that is Hispanic, 24 quarter fixed effects as well DMA \times season fixed effects as regressors. Sample sizes are 4,91 in Panel A; 7,377 in Panel B; 9,837 in Panel C; and 12,296 in Panel D. *** indicates significant at the 1 percent level, ** indicates significant at the 5 percent level, * indicates significant at the 10 percent level.

Table 5
 Placebo Tests of Estimated Reduced Form and Instrumental Variables Impact on Teen Birth Rates
 Rolling 24 Quarter Periods with Quadratic DMA-Specific Trends

Row	Dates			Reduced Form		Instrumental Variables	
	Begin	"Show" Start	End	Coefficient	Std. Err.	Coefficient	Std. Err.
(1)	1999:QI	2003:QIII	2004:QIV	-0.305	1.975	-0.207	1.168
(2)	1999:QII	2003:QIV	2005:QI	-1.165	2.415	-0.788	1.460
(3)	1999:QIII	2004:QI	2005:QII	-2.645	1.739	-1.787 *	1.082
(4)	1999:QIV	2004:QII	2005:QIII	-2.817	1.720	-1.906 **	0.942
(5)	2000:QI	2004:QIII	2005:QIV	-2.329	3.428	-1.576	1.889
(6)	2000:QII	2004:QIV	2006:QI	-1.522	3.676	-1.036	2.084
(7)	2000:QIII	2005:QI	2006:QII	-0.029	2.334	-0.020	1.383
(8)	2000:QIV	2005:QII	2006:QIII	1.030	1.264	0.700	0.705
(9)	2001:QI	2005:QIII	2006:QIV	0.595	2.067	0.404	1.177
(10)	2001:QII	2005:QIV	2007:QI	3.192 *	1.876	2.196 **	1.006
(11)	2001:QIII	2006:QI	2007:QII	1.955	1.412	1.386	0.859
(12)	2001:QIV	2006:QII	2007:QIII	1.130	1.950	0.782	1.150
(13)	2002:QI	2006:QIII	2007:QIV	1.018	1.728	0.700	1.017
(14)	2002:QII	2006:QIV	2008:QI	-0.351	1.673	-0.241	0.994
(15)	2002:QIII	2007:QI	2008:QII	-1.481	1.411	-1.018	0.876
(16)	2002:QIV	2007:QII	2008:QIII	-1.053	0.947	-0.724	0.574
(17)	2003:QI	2007:QIII	2008:QIV	-0.132	1.127	-0.091	0.664
(18)	2003:QII	2007:QIV	2009:QI	0.547	0.900	0.375	0.539
(19)	2003:QIII	2008:QI	2009:QII	1.638	1.015	1.120 *	0.593
(20)	2003:QIV	2008:QII	2009:QIII	1.503	1.316	1.027	0.783
(21)	2004:QI	2008:QIII	2009:QIV	1.575	1.626	1.074	0.981
(22)	2004:QII	2008:QIV	2010:QI	0.225	1.361	0.153	0.801
(23)	2004:QIII	2009:QI	2010:QII	-1.356	1.349	-0.925	0.780
(24)	2004:QIV	2009:QII	2010:QIII	-1.412	1.099	-0.962	0.638
(25)	2005:QI	2009:QIII	2010:QIV	-2.333 *	1.293	-1.591 **	0.756
(26)	2005:QII	2009:QIV	2011:QI	-2.815 **	1.222	-1.925 **	0.756
(27)	2005:QIII	2010:QI	2011:QII	0.358	1.257	0.245	0.749
(28)	2005:QIV	2010:QII	2011:QIII	0.774	1.136	0.528	0.743
(29)	2006:QI	2010:QIII	2011:QIV	0.442	1.065	0.301	0.686
(30)	2006:QII	2010:QIV	2012:QI	2.823 **	1.171	1.927 ***	0.624
(31)	2006:QIII	2011:QI	2012:QII	0.679	1.436	0.460	0.738
(32)	2006:QIV	2011:QII	2012:QIII	0.632	1.423	0.431	0.835
(33)	2007:QI	2011:QIII	2012:QIV	-0.100	1.347	-0.068	0.827
(34)	2007:QII	2011:QIV	2013:QI	-0.864	1.501	-0.590	0.793
(35)	2007:QIII	2012:QI	2013:QII	-2.005	1.573	-1.377	0.892
(36)	2007:QIV	2012:QII	2013:QIII	-1.204	1.427	-0.822	0.960

Notes: Entries in the table are a) in the reduced form, the estimated coefficient on MTV Ratings in 2008:QIII-2009:QII interacted with a dummy variable for being in the "post" period and b) for instrumental variables, the estimated coefficient on 16 and Pregnant Ratings interacted with a dummy variable for being in the "post" period where the instrument is the regressor of interest from the reduced form regressions. Standard errors, clustered by DMA, are shown in parentheses. All regressions are weighted by the female population aged 15-19 in the DMA at the time of the observation. In each regression there are 18 pre-"show" quarters and 6 post-"show" quarters. All regressions also include DMA-specific quadratic trends, the unemployment rate, the percent of the population that is non-Hispanic black and the percent of the population that is Hispanic, 24 quarter fixed effects as well DMA × season fixed effects and DMA-specific quadratic time trends as regressors. The period analyzed by Kearney and Levine's (2015b) is boxed. In the periods that are lightly shaded, the placebo post period partially includes the actual post-16 and Pregnant period. In the periods that are darkly shaded, both the placebo pre- and the placebo post-"show" period include the actual post-16 and Pregnant period. Sample size in rows (1) through (17), (19) through (21), and (36) is 4,918, the sample size in row (18) is 4,917, and the sample size in rows (22) through (35) is 4,919. *** indicates significant at the 1 percent level, ** indicates significant at the 5 percent level, * indicates significant at the 10 percent level.

Table 6
 Placebo Tests of Instrumental Variables Estimates of Impact on Teen Birth Rates
 Fixed for the Entire Pre-16 and Pregnant Period with Quadratic DMA-specific Trends

Row	Placebo "Show" Start	Reduced Form		Instrumental Variables	
		Coefficient	Std. Err.	Coefficient	Std. Err.
(1)	2003:QIII	-1.405	1.282	-0.958	0.732
(2)	2003:QIV	-2.217	1.469	-1.513 *	0.802
(3)	2004:QI	-2.778	1.897	-1.899 *	1.039
(4)	2004:QII	-3.054	2.269	-2.086 *	1.265
(5)	2004:QIII	-3.004	2.683	-2.053	1.535
(6)	2004:QIV	-2.777	2.820	-1.901	1.640
(7)	2005:QI	-2.258	2.344	-1.550	1.386
(8)	2005:QII	-1.594	1.322	-1.091	0.795
(9)	2005:QIII	-1.517	1.133	-1.038	0.733
(10)	2005:QIV	-0.231	1.281	-0.158	0.812
(11)	2006:QI	-0.411	1.530	-0.283	0.976
(12)	2006:QII	-0.685	1.833	-0.471	1.172
(13)	2006:QIII	-0.051	1.772	-0.035	1.120
(14)	2006:QIV	0.025	1.807	0.017	1.137
(15)	2007:QI	-0.106	1.778	-0.072	1.120
(16)	2007:QII	0.561	1.500	0.383	0.919
(17)	2007:QIII	1.905	1.550	1.300	0.907
(18)	2007:QIV	2.505 *	1.378	1.712 **	0.783
(19)	2008:QI	2.600 *	1.425	1.778 **	0.799

Notes: Entries in the table are a) for the reduced form, the estimated coefficient on MTV Ratings in 2008:QIII-2009:QII interacted with a dummy variable for being in the "post" period and b) for instrumental variables, the estimated coefficient on 16 and Pregnant Ratings interacted with a dummy variable for being in the "post" period where the instrument is the regressor of interest from the reduced form regressions. Standard errors, clustered by DMA, are shown in parentheses. All regressions are weighted by the female population aged 15-19 in the DMA at the time of the observation. Each regression uses the period 1999:QI to 2009:QII. The only difference in each regression is when the placebo "show" begins. All regressions also include DMA-specific quadratic trends, the unemployment rate, the percent of the population that is non-Hispanic black and the percent of the population that is Hispanic, 24 quarter fixed effects as well DMA \times season fixed effects and DMA-specific quadratic time trends as regressors. The first stage relationship between MTV Ratings and 16 and Pregnant Ratings ranges from 1.451 (row 19) to 1.467 (row 1). Sample size for all regressions is 8,607. *** indicates significant at the 1 percent level, ** indicates significant at the 5 percent level, * indicates significant at the 10 percent level.

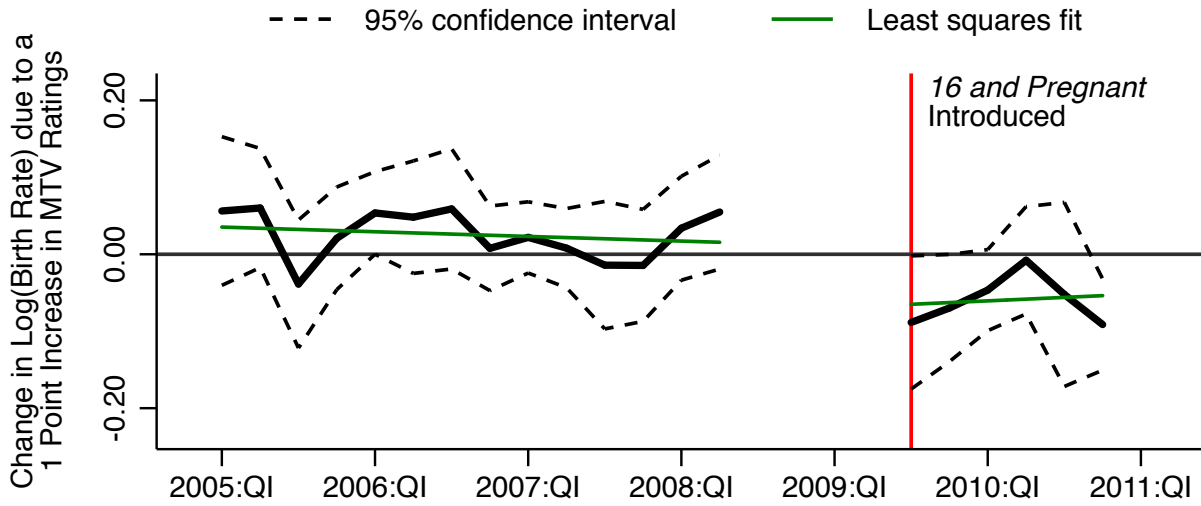
Table 7
Estimated Impact of *16 and Pregnant* on Tweets about Birth Control and Abortion

	Birth Control			Abortion		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A: National Tweet Rate Using Only "In Season" Daily Data between 1 January 2009 and 31 December 2012</i>						
<i>(i) Impact of 16 and Pregnant Broadcasts (336 days)</i>						
Day new <i>16 and Pregnant</i> episode released	0.120 *** (0.046)	0.065 (0.052)	0.015 (0.064)	0.142 *** (0.046)	0.176 *** (0.061)	0.027 (0.051)
Day after new <i>16 and Pregnant</i> episode released	0.229 *** (0.047)	0.157 (0.052)	0.035 (0.055)	0.212 *** (0.046)	0.180 *** (0.061)	0.024 (0.063)
Replicates Kearney and Levine (2015b) results	Table 3, Column 4	no	no	Table 3, Column 5	no	no
<i>(ii) Impact of ln(tweet rate) of 16 and Pregnant (336 days)</i>						
ln(tweet rate), <i>16 and Pregnant</i>	0.077 *** (0.027)	0.029 (0.027)	0.004 (0.027)	0.064 ** (0.027)	0.007 (0.026)	0.019 (0.030)
Replicates Kearney and Levine (2015b) results	Table 4, Panel A, Column 4	no	no	Table 4, Panel A, Column 5	no	no
<i>Panel B: National Tweet Rate Using All Available Daily Data between 1 January 2009 and 31 December 2012</i>						
<i>(i) Impact of 16 and Pregnant Broadcasts (1,455 days)</i>						
Day new <i>16 and Pregnant</i> episode released	-0.045 (0.044)	-0.045 (0.039)	-0.171 *** (0.051)	-0.015 (0.064)	0.052 (0.043)	-0.112 ** (0.055)
Day after new <i>16 and Pregnant</i> episode released	0.046 (0.055)	0.041 (0.049)	-0.124 ** (0.052)	0.037 (0.063)	0.046 (0.052)	-0.092 (0.056)
Sum of Day 1 through Day 7 indicators "in season"			-0.126 *** (0.031)			-0.108 *** (0.028)
<i>(ii) Impact of ln(tweet rate) of 16 and Pregnant (1,322 days)</i>						
ln(tweet rate), <i>16 and Pregnant</i>	0.035 ** (0.017)	0.011 (0.012)	0.013 (0.016)	-0.020 (0.016)	-0.009 (0.019)	-0.016 (0.017)
<i>All Panels: Weighting and Covariates</i>						
Includes day of week and month indicators	no	no	yes	no	no	yes
Weighted by number of tweets	yes	no	yes	yes	no	yes

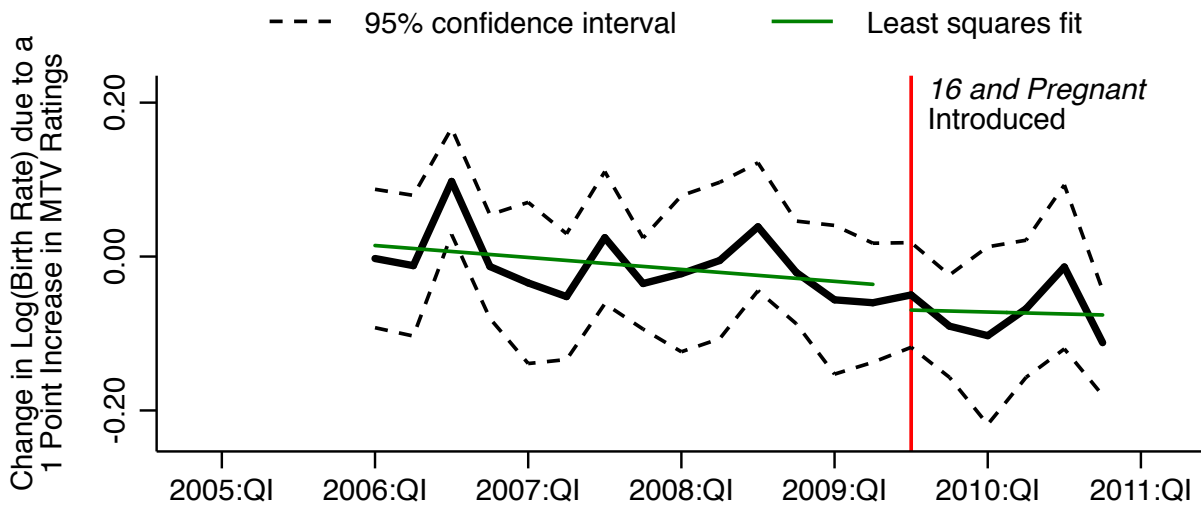
Notes: Dependent variable is ln(tweet rate). Panel A, columns (1) and (4) use heteroskedasticity-consistent standard errors (in parentheses) to be consistent with Kearney and Levine (2015b). All other regressions use Newey-West (1987) standard errors (in parentheses) with two autoregressive lags, which were not sensitive to including additional lags. Panel A uses national time series that includes only the day that a new episode of *16 and Pregnant* was broadcast and the 6 subsequent days between 1 January 2009 and 31 December 2012. Panel B uses the same specifications as in Panel A but the sample includes all tweets between 1 January 2009 and 31 December 2012. Sample sizes in Panel B(ii) are different from those in Panel B(i) because days with no tweets about *16 and Pregnant* are dropped. Estimates in columns (1), (3) and (4), and (6) are weighted by the number of daily tweets. Estimates in columns (2) and (5) use the same sample and specification as columns (1) and (4), respectively, but regressions are not weighted. Regressions in Panels A(i) and B(i) include 7 indicators for the day that a new episode of *16 and Pregnant* was broadcast and the subsequent 6 days. Following Kearney and Levine (2015b), we show the first two of these estimated coefficients. In Panel B(ii), columns (3) and (6) we also show the average of all 7 of these indicators. Data used are the same as in Figure 7 of Kearney and Levine (2015b), but at a daily frequency. *** indicates significant at the 1 percent level, ** indicates significant at the 5 percent level, * indicates significant at the 10 percent level.

Appendix Figure 1
 Reduced Form Event Study: 2005-2010
 Unweighted

Panel A: Reference period is 2008:QIII-2009:QII



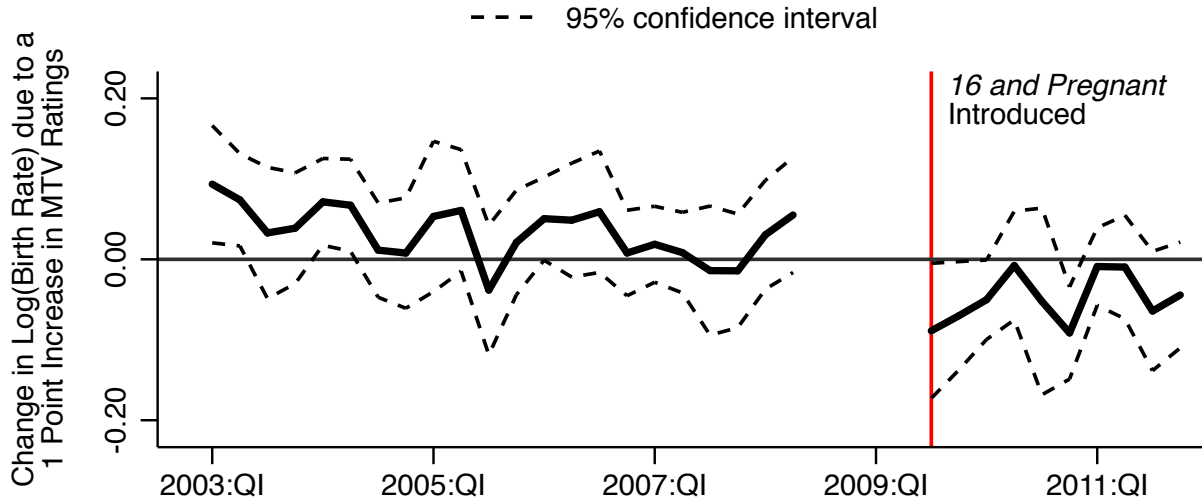
Panel B: Reference period is 2005:Q1-2005:QIV



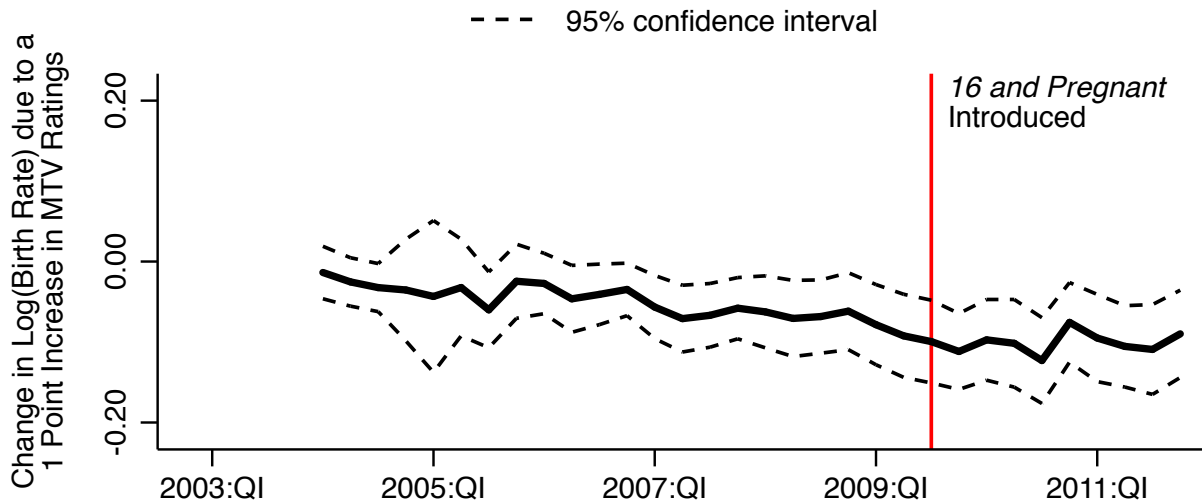
Notes: These figures plot the coefficients of average MTV Ratings in 2008-2009 interacted with period dummy variables. The regressions also include the unemployment rate and a full set of period dummy variables and DMA \times season interactions. Confidence intervals are calculated by clustering by DMA. Panel A uses 2008:QIII through 2009:QII as the reference period. Panel B uses 2005:Q1 through 2005:QIV as the reference period. Least squares lines in both panels are fitted separately for the pre- and post-16 and Pregnant periods. The F statistic for the test of joint significance of the coefficients in the pre-16 and Pregnant period is 1.88 with a p -value of 0.030.

Appendix Figure 2
 Reduced Form Event Study: 2003-2011
 Unweighted

Panel A: Reference period is 2008:QIII-2009:QII



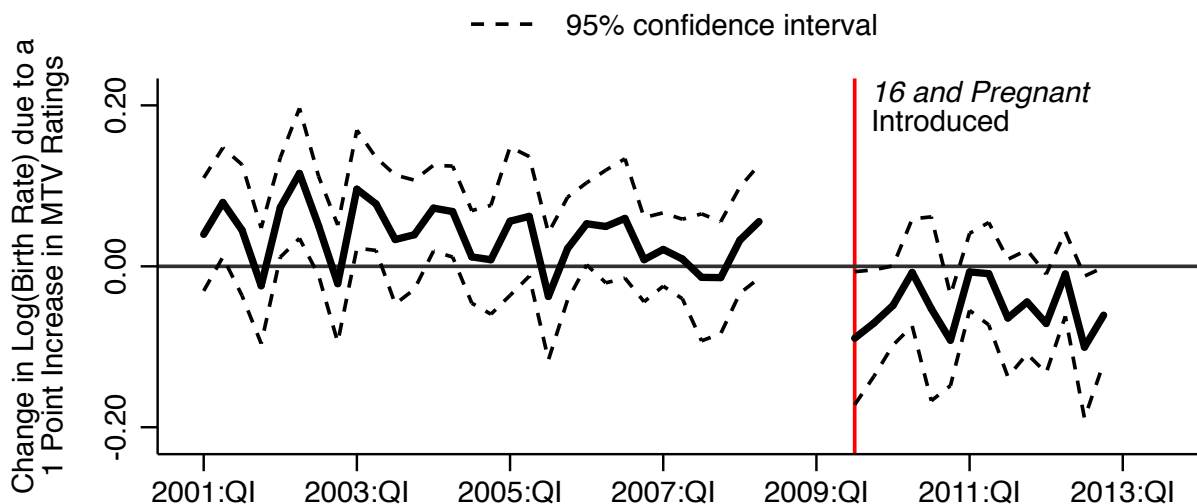
Panel B: Reference period is 2003:Q1-2003:QIV



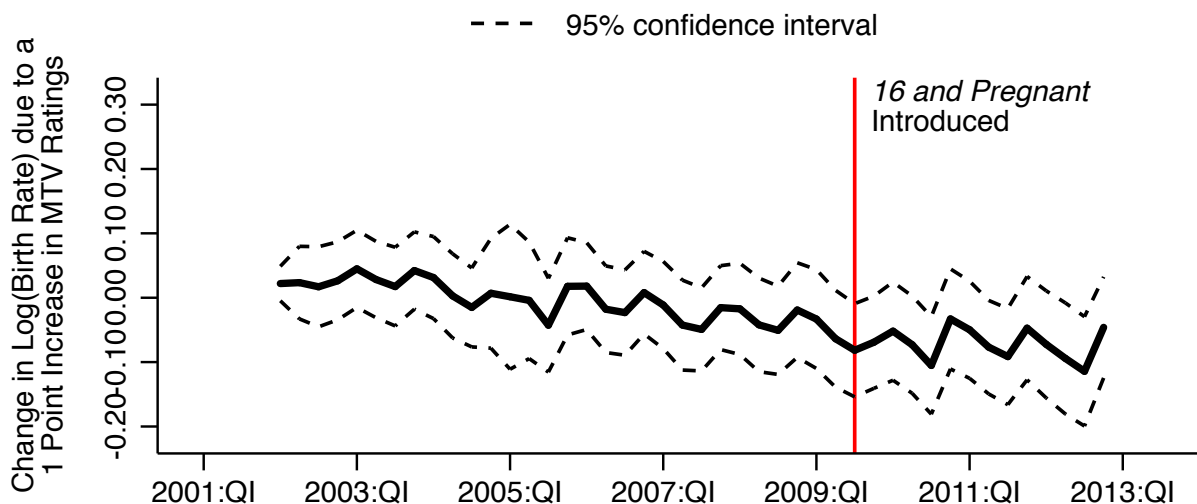
Notes: These figures plot the coefficients of average MTV Ratings in 2008-2009 interacted with period dummy variables. The regressions also include the unemployment rate and a full set of period dummy variables and DMA \times season interactions. Confidence intervals are calculated by clustering by DMA. Panel A uses 2008:QIII through 2009:QII as the reference period. Panel B uses 2003:Q1 through 2003:QIV as the reference period. Least squares lines in both panels are fitted separately for the pre- and post-16 and Pregnant periods. The F statistic for the test of joint significance of the coefficients in the pre-16 and Pregnant period is 2.13 with a p -value of 0.003.

Appendix Figure 3
 Reduced Form Event Study: 2001-2012
 Unweighted

Panel A: Reference period is 2008:QIII-2009:QII



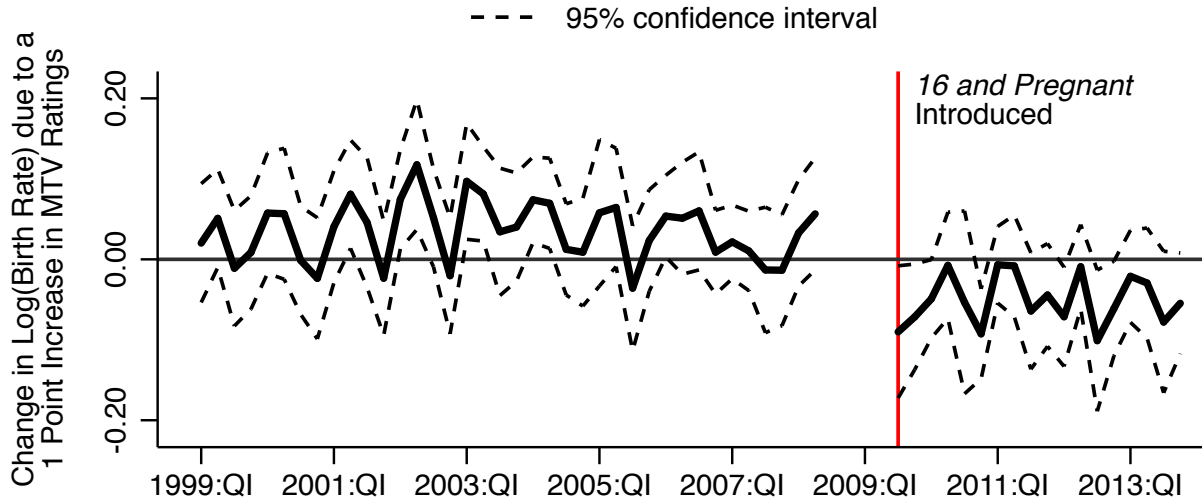
Panel B: Reference period is 2001:Q1-2001:QIV



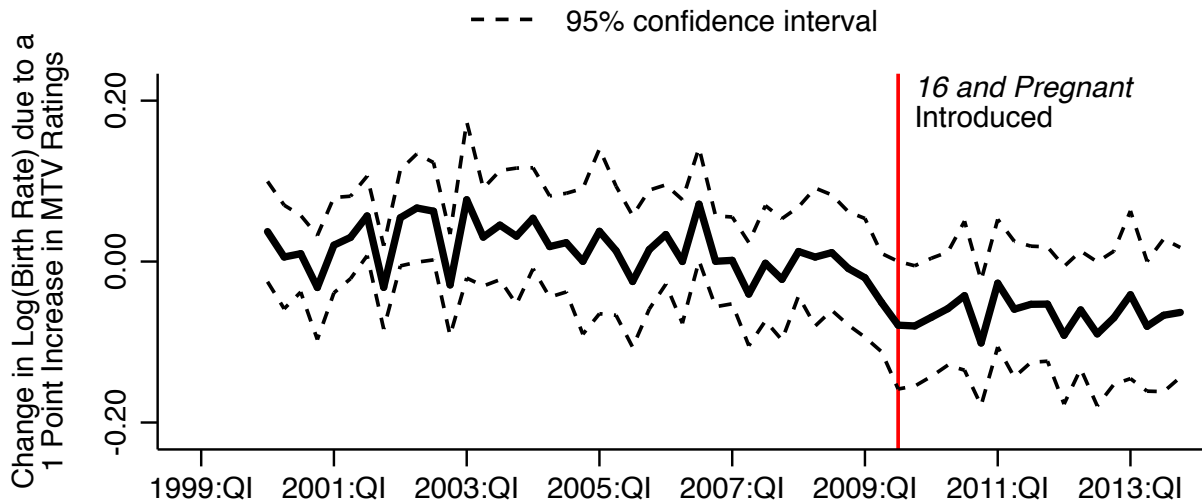
Notes: These figures plot the coefficients of average MTV Ratings in 2008-2009 interacted with period dummy variables. The regressions also include the unemployment rate and a full set of period dummy variables and DMA \times season interactions. Confidence intervals are calculated by clustering by DMA. Panel A uses 2008:QIII through 2009:QII as the reference period. Panel B uses 2001:Q1 through 2001:QIV as the reference period. Least squares lines in both panels are fitted separately for the pre- and post-*16 and Pregnant* periods. The F statistic for the test of joint significance of the coefficients in the pre-*16 and Pregnant* period is 2.83 with a p -value of <0.001 .

Appendix Figure 4
 Reduced Form Event Study: 1999-2013
 Unweighted

Panel A: Reference period is 2008:QIII-2009:QII



Panel B: Reference period is 1999:Q1-1999:QIV



Notes: These figures plot the coefficients of average MTV Ratings in 2008-2009 interacted with period dummy variables. The regressions also include the unemployment rate and a full set of period dummy variables and DMA \times season interactions. Confidence intervals are calculated by clustering by DMA. Panel A uses 2008:QIII through 2009:QII as the reference period. Panel B uses 1999:Q1 through 1999:QIV as the reference period. Least squares lines in Panel A are fitted separately for the before and after *16 and Pregnant* periods. The F statistic for the test of joint significance of the coefficients in the pre-*16 and Pregnant* period is 1.55 with a p -value of 0.030.

Appendix Table 1
 Placebo Tests of Estimated Reduced Form and Instrumental Variables Impact on Teen Birth Rates
 Rolling 24 Quarter Periods
 Unweighted

Row	Dates			Reduced Form		Instrumental Variables	
	Begin	"Show" Start	End	Coefficient	Std. Err.	Coefficient	Std. Err.
(1)	1999:QI	2003:QIII	2004:QIV	1.074	1.810	0.781	1.232
(2)	1999:QII	2003:QIV	2005:QI	0.577	2.100	0.420	1.417
(3)	1999:QIII	2004:QI	2005:QII	-0.717	2.402	-0.523	1.546
(4)	1999:QIV	2004:QII	2005:QIII	-1.849	2.705	-1.347	1.678
(5)	2000:QI	2004:QIII	2005:QIV	-1.374	2.911	-1.003	1.840
(6)	2000:QII	2004:QIV	2006:QI	-1.515	2.797	-1.089	1.720
(7)	2000:QIII	2005:QI	2006:QII	-2.261	2.412	-1.610	1.406
(8)	2000:QIV	2005:QII	2006:QIII	-1.578	1.833	-1.113	1.079
(9)	2001:QI	2005:QIII	2006:QIV	-0.855	1.479	-0.599	0.896
(10)	2001:QII	2005:QIV	2007:QI	-0.070	1.306	-0.049	0.824
(11)	2001:QIII	2006:QI	2007:QII	-1.160	1.422	-0.806	0.873
(12)	2001:QIV	2006:QII	2007:QIII	-1.628	1.571	-1.132	0.987
(13)	2002:QI	2006:QIII	2007:QIV	-1.516	1.387	-1.055	0.867
(14)	2002:QII	2006:QIV	2008:QI	-2.797 *	1.609	-1.948 *	1.040
(15)	2002:QIII	2007:QI	2008:QII	-2.500	1.602	-1.744 *	1.046
(16)	2002:QIV	2007:QII	2008:QIII	-1.876	1.463	-1.309	0.955
(17)	2003:QI	2007:QIII	2008:QIV	-1.107	1.496	-0.773	0.963
(18)	2003:QII	2007:QIV	2009:QI	-1.147	1.409	-0.801	0.921
(19)	2003:QIII	2008:QI	2009:QII	-1.125	1.433	-0.784	0.938
(20)	2003:QIV	2008:QII	2009:QIII	-2.120	1.394	-1.474	0.944
(21)	2004:QI	2008:QIII	2009:QIV	-3.661 **	1.507	-2.537 **	1.074
(22)	2004:QII	2008:QIV	2010:QI	-4.961 ***	1.589	-3.436 ***	1.191
(23)	2004:QIII	2009:QI	2010:QII	-5.747 ***	1.877	-3.979 ***	1.393
(24)	2004:QIV	2009:QII	2010:QIII	-5.791 ***	1.696	-4.024 ***	1.303
(25)	2005:QI	2009:QIII	2010:QIV	-6.455 ***	1.695	-4.504 ***	1.343
(26)	2005:QII	2009:QIV	2011:QI	-5.000 ***	1.535	-3.521 ***	1.178
(27)	2005:QIII	2010:QI	2011:QII	-4.308 ***	1.360	-3.054 ***	1.018
(28)	2005:QIV	2010:QII	2011:QIII	-3.757 ***	1.382	-2.671 ***	1.001
(29)	2006:QI	2010:QIII	2011:QIV	-3.351 **	1.571	-2.388 **	1.076
(30)	2006:QII	2010:QIV	2012:QI	-3.626 **	1.410	-2.588 ***	0.970
(31)	2006:QIII	2011:QI	2012:QII	-2.460 *	1.378	-1.760 *	0.921
(32)	2006:QIV	2011:QII	2012:QIII	-2.712 *	1.533	-1.946 *	1.011
(33)	2007:QI	2011:QIII	2012:QIV	-2.195	1.533	-1.582	1.033
(34)	2007:QII	2011:QIV	2013:QI	-1.722	1.579	-1.249	1.064
(35)	2007:QIII	2012:QI	2013:QII	-2.147	1.645	-1.567	1.141
(36)	2007:QIV	2012:QII	2013:QIII	-1.132	1.598	-0.825	1.080

Notes: Entries in the table are a) in the reduced form, the estimated coefficient on MTV Ratings in 2008:QIII-2009:QII interacted with a dummy variable for being in the "post" period and b) for instrumental variables, the estimated coefficient on *16 and Pregnant* Ratings interacted with a dummy variable for being in the "post" period where the instrument is the regressor of interest from the reduced form regressions. Standard errors, clustered by DMA, are shown in parentheses. In each regression there are 18 pre-"show" quarters and 6 post-"show" quarters. All regressions also include the unemployment rate, the percent of the population that is non-Hispanic black and the percent of the population that is Hispanic, 24 quarter fixed effects as well DMA × season fixed effects as regressors. The period analyzed by Kearney and Levine's (2015b) is boxed. In the periods that are lightly shaded, the placebo post period partially includes the actual post-*16 and Pregnant* period. In the periods that are darkly shaded, both the placebo pre- and the placebo post-"show" period include the actual post-*16 and Pregnant* period. Sample size in rows (1) through (17), (19) through (21), and (36) is 4,918, the sample size in row (18) is 4,917, and the sample size in rows (22) through (35) is 4,919. *** indicates significant at the 1 percent level, ** indicates significant at the 5 percent level, * indicates significant at the 10 percent level.

Appendix Table 2

Placebo Tests of Reduced Form and Instrumental Variables Estiamtes of Impact on Teen Birth Rates
 Fixed for the Entire Pre-16 and Pregnant Period
 Unweighted

Row	Placebo "Show" Start	Reduced Form		Instrumental Variables	
		Coefficient	Std. Err.	Coefficient	Std. Err.
(1)	2003:QIII	-0.271	2.067	-0.859 **	0.411
(2)	2003:QIV	-0.453	2.019	-0.923 **	0.408
(3)	2004:QI	-0.826	2.001	-0.665	0.430
(4)	2004:QII	-1.083	1.971	-0.593	0.443
(5)	2004:QIII	-1.197	1.944	-0.181	0.499
(6)	2004:QIV	-1.198	1.919	0.023	0.528
(7)	2005:QI	-1.299	1.828	0.276	0.547
(8)	2005:QII	-1.424	1.667	0.588	0.515
(9)	2005:QIII	-1.518	1.577	0.723	0.495
(10)	2005:QIV	-1.031	1.536	0.886 *	0.511
(11)	2006:QI	-1.331	1.502	0.591	0.462
(12)	2006:QII	-1.472	1.514	0.474	0.433
(13)	2006:QIII	-1.473	1.458	0.472	0.401
(14)	2006:QIV	-2.194	1.465	0.250	0.386
(15)	2007:QI	-2.489	1.544	-0.063	0.409
(16)	2007:QII	-2.351	1.531	-0.256	0.429
(17)	2007:QIII	-1.855	1.493	-0.821 *	0.490
(18)	2007:QIV	-1.683	1.401	-0.786	0.530
(19)	2008:QI	-1.730	1.434	-1.191	0.945

Notes: Entries in the table are a) for the reduced form, the estimated coefficient on MTV Ratings in 2008:QIII-2009:QII interacted with a dummy variable for being in the "post" period and b) for instrumental variables, the estimated coefficient on *16 and Pregnant* Ratings interacted with a dummy variable for being in the "post" period where the instrument is the regressor of interest from the reduced form regressions. Standard errors, clustered by DMA, are shown in parentheses. Each regression uses the period 1999:QI to 2009:QII. The only difference in each regression is when the placebo "show" begins. All regressions also include the unemployment rate, the percent of the population that is non-Hispanic black and the percent of the population that is Hispanic, 24 quarter fixed effects as well DMA \times season fixed effects as regressors. The first stage relationship between MTV Ratings and *16 and Pregnant* Ratings ranges from -0.859 (row 10) to 3.164 (row 1). Sample size for all regressions is 8,607. *** indicates significant at the 1 percent level, ** indicates significant at the 5 percent level, * indicates significant at the 10 percent level.

Appendix Table 3
Differential Trends in Teen Birth Rates during the Pre-16 and Pregnant Period by Quartiles of MTV Ratings and 16 and Pregnant Ratings
Unweighted

Period:	(1)	(2)	(3)	(4)
	2005:Q1-2009:QII	2003:Q1-2009:QII	2001:Q1-2009:QII	1999:Q1-2009:QII
<i>Panel A: Quartiles of MTV Ratings</i>				
<i>F:</i> I(Second quartile), I(Third quartile) and I(Fourth quartile) × Time Dummies = 0	2.006 *** [<0.001]	2.384 *** [<0.001]	2.249 *** [<0.001]	2.909 *** [<0.001]
<i>F:</i> I(Second quartile) × Time Dummies = 0	1.260 [0.222]	1.433 * [0.091]	1.242 [0.184]	1.267 [0.146]
<i>F:</i> I(Third quartile) × Time Dummies = 0	2.337 *** [0.003]	2.560 *** [<0.001]	2.376 *** [<0.001]	2.933 *** [<0.001]
<i>F:</i> I(Fourth quartile) × Time Dummies = 0	2.481 *** [0.001]	1.903 *** [0.008]	1.939 *** [0.003]	1.854 *** [0.003]
<i>F:</i> I(Median or above) × Time Dummies = 0	2.465 *** [0.002]	1.888 *** [0.009]	1.560 [0.034]	1.602 [0.018]
<i>Panel B: Quartiles of 16 and Pregnant Ratings</i>				
<i>F:</i> I(Second quartile), I(Third quartile) and I(Fourth quartile) × Time Dummies = 0	1.441 ** [0.040]	1.554 *** [0.008]	2.188 *** [<0.001]	2.464 *** [<0.001]
<i>F:</i> I(Second quartile) × Time Dummies = 0	2.053 *** [0.010]	2.015 *** [0.004]	2.295 *** [0.001]	2.466 *** [<0.001]
<i>F:</i> I(Third quartile) × Time Dummies = 0	2.314 *** [0.003]	1.746 ** [0.019]	1.715 ** [0.013]	1.93 *** [0.002]
<i>F:</i> I(Fourth quartile) × Time Dummies = 0	1.797 ** [0.030]	1.769 ** [0.017]	1.651 ** [0.020]	1.839 *** [0.003]
<i>F:</i> I(Median or above) × Time Dummies = 0	1.312 [0.187]	1.053 [0.401]	1.090 [0.347]	1.284 [0.133]
Number of time periods	14	26	34	42
Sample Size	3,689	5,327	6,967	8,607

Notes: *p*-values in [brackets], standard errors in (parentheses). Variance-covariance matrix of estimates calculated by clustering by DMA. Regressions are unweighted. Estimates in each column of each panel are from a separate regression. *F*-statistics are from tests that the coefficients on the interactions between quarterly indicators and indicators for the relevant quartiles of either 16 and Pregnant ratings or MTV ratings are jointly equal to zero. All models also include DMA fixed effects, the percent of non-Hispanic blacks and percent of Hispanics in the DMA as regressors. The results in column (1) use the same pre-16 and Pregnant period as Kearney and Levine (2015b). For the estimated coefficients, *** indicates significant at the 1 percent level, ** indicates significant at the 5 percent level, * indicates significant at the 10 percent level.

Appendix Table 4
Instrumental Variables Estimates of Impact of *16 and Pregnant* on Birth Rates
with DMA-Specific or MTV-Quartile-Specific Trends
Unweighted

	(1)	(2)	(3)	(4)	(5)
<i>Panel A: 2005:Q1-2010:QIV</i>					
<i>16 and Pregnant</i> Ratings	-4.504 *** (1.343)	-3.620 ** (1.845)	-5.349 *** (1.538)	-4.962 ** (1.954)	-4.800 ** (1.949)
$\chi^2(6)$ test: MTV Quartile-Specific Trends=0					4.99 [0.545]
<i>Panel B: 2003:Q1-2011:QIV</i>					
<i>16 and Pregnant</i> Ratings	-4.453 *** (1.256)	-4.302 *** (1.266)	-5.377 *** (1.550)	-5.031 *** (1.489)	-4.911 *** (1.471)
$\chi^2(6)$ test: MTV Quartile-Specific Trends=0					11.17 [0.083]
<i>Panel C: 2001:Q1-2012:QIV</i>					
<i>16 and Pregnant</i> Ratings	-4.816 *** (1.352)	-2.855 *** (0.974)	-5.573 *** (1.624)	-3.313 *** (1.156)	-3.406 *** (1.149)
$\chi^2(6)$ test: MTV Quartile-Specific Trends=0					11.30 [0.080]
<i>Panel D: 1999:Q1-2013:QIV</i>					
<i>16 and Pregnant</i> Ratings	-4.670 *** (1.435)	-3.029 *** (1.013)	-5.282 *** (1.669)	-3.835 *** (1.247)	-4.196 *** (1.357)
$\chi^2(6)$ test: MTV Quartile-Specific Trends=0					8.34 [0.214]
<i>Control Variables</i>					
DMA-Specific Quadratic Trends	no	yes	no	yes	no
MTV Quartile-Specific Quadratic Trends	no	no	no	no	yes
<i>Instruments</i>					
MTV \times Post	yes	yes	no	no	no
I(MTV Quartile) \times Post	no	no	yes	yes	yes

Notes: Standard errors, clustered by DMA, are shown in parentheses. p -values of χ^2 -statistics are shown in brackets. Estimates in each column of each panel are from separate regressions. All models also include the unemployment rate, the percent of the population that is non-Hispanic black and the percent of the population that is Hispanic, 24 quarter fixed effects as well DMA \times season fixed effects as regressors. Sample sizes are 4,91 in Panel A; 7,377 in Panel B; 9,837 in Panel C; and 12,296 in Panel D. *** indicates significant at the 1 percent level, ** indicates significant at the 5 percent level, * indicates significant at the 10 percent level.

Appendix Table 5
 Placebo Tests of Estimated Reduced Form and Instrumental Variables Impact on Teen Birth Rates
 Rolling 24 Quarter Periods with Quadratic DMA-Specific Trends
 Unweighted

Row	Dates			Reduced Form		Instrumental Variables	
	Begin	"Show" Start	End	Coefficient	Std. Err.	Coefficient	Std. Err.
(1)	1999:QI	2003:QIII	2004:QIV	0.219	2.666	0.159	1.662
(2)	1999:QII	2003:QIV	2005:QI	0.576	2.685	0.418	1.658
(3)	1999:QIII	2004:QI	2005:QII	-2.748	2.238	-1.995	1.474
(4)	1999:QIV	2004:QII	2005:QIII	-2.092	2.555	-1.517	1.606
(5)	2000:QI	2004:QIII	2005:QIV	-1.397	3.351	-1.012	2.047
(6)	2000:QII	2004:QIV	2006:QI	-0.090	3.698	-0.065	2.307
(7)	2000:QIII	2005:QI	2006:QII	-0.609	3.180	-0.444	1.972
(8)	2000:QIV	2005:QII	2006:QIII	-1.866	3.667	-1.351	2.298
(9)	2001:QI	2005:QIII	2006:QIV	-1.897	3.023	-1.374	1.940
(10)	2001:QII	2005:QIV	2007:QI	5.155 **	2.495	3.774 **	1.567
(11)	2001:QIII	2006:QI	2007:QII	3.287	2.330	2.469	1.531
(12)	2001:QIV	2006:QII	2007:QIII	4.027	3.142	2.946	1.928
(13)	2002:QI	2006:QIII	2007:QIV	3.407	2.941	2.469	1.799
(14)	2002:QII	2006:QIV	2008:QI	-3.530	2.380	-2.553	1.639
(15)	2002:QIII	2007:QI	2008:QII	-5.315	3.252	-3.862 *	2.191
(16)	2002:QIV	2007:QII	2008:QIII	-2.590	2.437	-1.881	1.554
(17)	2003:QI	2007:QIII	2008:QIV	-0.113	2.740	-0.082	1.713
(18)	2003:QII	2007:QIV	2009:QI	2.170	2.686	1.579	1.718
(19)	2003:QIII	2008:QI	2009:QII	4.360	3.259	3.169	2.050
(20)	2003:QIV	2008:QII	2009:QIII	4.690	4.115	3.398	2.617
(21)	2004:QI	2008:QIII	2009:QIV	1.568	3.113	1.135	1.964
(22)	2004:QII	2008:QIV	2010:QI	-1.014	2.726	-0.734	1.688
(23)	2004:QIII	2009:QI	2010:QII	-4.693	3.160	-3.382 *	1.977
(24)	2004:QIV	2009:QII	2010:QIII	-5.620	3.422	-4.073 *	2.147
(25)	2005:QI	2009:QIII	2010:QIV	-4.979 *	2.835	-3.620 **	1.845
(26)	2005:QII	2009:QIV	2011:QI	-1.251	2.791	-0.909	1.751
(27)	2005:QIII	2010:QI	2011:QII	0.558	2.148	0.405	1.336
(28)	2005:QIV	2010:QII	2011:QIII	2.834	2.438	2.048	1.511
(29)	2006:QI	2010:QIII	2011:QIV	1.447	4.734	1.042	2.938
(30)	2006:QII	2010:QIV	2012:QI	2.044	3.182	1.473	1.995
(31)	2006:QIII	2011:QI	2012:QII	4.671 *	2.719	3.342 *	1.753
(32)	2006:QIV	2011:QII	2012:QIII	2.679	2.195	1.936	1.441
(33)	2007:QI	2011:QIII	2012:QIV	1.545	2.749	1.117	1.719
(34)	2007:QII	2011:QIV	2013:QI	-0.170	3.053	-0.123	1.898
(35)	2007:QIII	2012:QI	2013:QII	-3.312	2.882	-2.407	1.829
(36)	2007:QIV	2012:QII	2013:QIII	0.148	1.992	0.106	1.235

Notes: Entries in the table are a) in the reduced form, the estimated coefficient on MTV Ratings in 2008:QIII-2009:QII interacted with a dummy variable for being in the "post" period and b) for instrumental variables, the estimated coefficient on *16 and Pregnant* Ratings interacted with a dummy variable for being in the "post" period where the instrument is the regressor of interest from the reduced form regressions. Standard errors, clustered by DMA, are shown in parentheses. In each regression there are 18 pre-"show" quarters and 6 post-"show" quarters. All regressions also include DMA-specific quadratic trends, the unemployment rate, the percent of the population that is non-Hispanic black and the percent of the population that is Hispanic, 24 quarter fixed effects as well DMA × season fixed effects and DMA-specific quadratic time trends as regressors. The period analyzed by Kearney and Levine's (2015b) is boxed. In the periods that are lightly shaded, the placebo post period partially includes the actual post-*16 and Pregnant* period. In the periods that are darkly shaded, both the placebo pre- and the placebo post-"show" period include the actual post-*16 and Pregnant* period. Sample size in rows (1) through (17), (19) through (21), and (36) is 4,918, the sample size in row (18) is 4,917, and the sample size in rows (22) through (35) is 4,919. *** indicates significant at the 1 percent level, ** indicates significant at the 5 percent level, * indicates significant at the 10 percent level.

Appendix Table 6
 Placebo Tests of Instrumental Variables Estimates of Impact on Teen Birth Rates
 Fixed for the Entire Pre-16 and Pregnant Period with Quadratic DMA-specific Trends
 Unweighted

Row	Placebo "Show" Start	Reduced Form		Instrumental Variables	
		Coefficient	Std. Err.	Coefficient	Std. Err.
(1)	2003:QIII	-0.818	2.023	-0.597	1.325
(2)	2003:QIV	-0.981	1.874	-0.715	1.212
(3)	2004:QI	-1.898	2.105	-1.383	1.312
(4)	2004:QII	-2.292	2.309	-1.666	1.427
(5)	2004:QIII	-2.166	2.559	-1.570	1.589
(6)	2004:QIV	-1.699	2.614	-1.229	1.645
(7)	2005:QI	-1.658	2.434	-1.197	1.527
(8)	2005:QII	-1.548	1.981	-1.118	1.267
(9)	2005:QIII	-1.312	1.647	-0.951	1.085
(10)	2005:QIV	1.461	1.538	1.065	1.043
(11)	2006:QI	0.768	1.812	0.566	1.231
(12)	2006:QII	0.890	2.124	0.650	1.413
(13)	2006:QIII	1.760	1.907	1.279	1.226
(14)	2006:QIV	-1.085	2.014	-0.786	1.400
(15)	2007:QI	-1.757	2.142	-1.272	1.520
(16)	2007:QII	-0.396	1.801	-0.286	1.220
(17)	2007:QIII	2.316	2.018	1.675	1.263
(18)	2007:QIV	3.133	2.030	2.266 *	1.269
(19)	2008:QI	2.909	2.419	2.104	1.528

Notes: Entries in the table are a) for the reduced form, the estimated coefficient on MTV Ratings in 2008:QIII-2009:QII interacted with a dummy variable for being in the "post" period and b) for instrumental variables, the estimated coefficient on 16 and Pregnant Ratings interacted with a dummy variable for being in the "post" period where the instrument is the regressor of interest from the reduced form regressions. Standard errors, clustered by DMA, are shown in parentheses. All regressions are weighted by the female population aged 15-19 in the DMA at the time of the observation. Each regression uses the period 1999:QI to 2009:QII. The only difference in each regression is when the placebo "show" begins. All regressions also include DMA-specific quadratic trends, the unemployment rate, the percent of the population that is non-Hispanic black and the percent of the population that is Hispanic, 24 quarter fixed effects as well DMA \times season fixed effects and DMA-specific quadratic time trends as regressors. The first stage relationship between MTV Ratings and 16 and Pregnant Ratings ranges from 1.358 (row 11) to 1.385 (row 7). Sample size for all regressions is 8,607. *** indicates significant at the 1 percent level, ** indicates significant at the 5 percent level, * indicates significant at the 10 percent level.