A CAUTIONARY TALE OF EVALUATING IDENTIFYING ASSUMPTIONS:
DID REALITY TV REALLY CAUSE A DECLINE IN TEENAGE CHILDBEARING?

David A. Jaeger
Theodore J. Joyce
Robert Kaestner

June 2018

A Cautionary Tale of Evaluating Identifying Assumptions:
Did Reality TV Really Cause a Decline in Teenage Childbearing?
June 2018
JEL No. J13, L82

## ABSTRACT

Evaluating policy changes that occur everywhere at the same time is difficult because of the lack of a clear counterfactual. Hoping to address this problem, researchers often proxy for differential exposure using some observed characteristic in the pre-treatment period. As a cautionary tale of how difficult identification is in such settings, we re-examine the results of an influential paper by Melissa Kearney and Philip Levine, who found that the MTV program *16 and Pregnant* had a substantial impact on teen birth rates. In what amounts to a difference-in-differences approach, they use the pre-treatment levels of MTV viewership across media markets as an instrument. We show that controlling for differential time trends in birth rates by a market's pre-treatment racial/ethnic composition or unemployment rate causes Kearney and Levine's results to disappear, invalidating the parallel trends assumption necessary for a causal interpretation. Extending the pre-treatment period and estimating placebo tests, we find evidence of an "effect" long before *16 and Pregnant* started broadcasting. Our results highlight the difficulty of drawing causal inferences from national point-in-time policy changes.

Keywords: Social Media, Difference-in-Differences, Shift-Share Instruments, Teen Pregnancy

David A. Jaeger
Ph.D. Program in Economics
CUNY Graduate Center
365 Fifth Ave
New York, NY 10016
and University of Cologne
and IZA
and NBER
dj@djaeger.org

Robert Kaestner
Institute of Government and Public Affairs
University of Illinois
815 West Van Buren Street, Suite 525
Chicago, IL 60607
and NBER
kaestner@uic.edu

Ted Joyce
Baruch College, CUNY
Department of Economics and Finance
55 Lexington Ave
New York, NY 10010
and NBER
theodore.joyce@baruch.cuny.edu

Evaluating policy changes that occur nationally at one point in time is particularly challenging. Cross-country comparisons are unlikely to be convincing because of institutional and cultural difference and simple before-and-after comparisons within a country are likely to be confounded by secular changes over time in the outcome. Finding an appropriate counterfactual in such cases is daunting and the assumptions necessary for identifying a causal effect should be thoroughly scrutinized.

One common identification strategy when evaluating a point-in-time change in policy is to use pre-treatment characteristics of the units under study as a measure of the exposure to the treatment. These characteristics are often interacted with indicators of time or other variables that vary over time to create instruments that are used to isolate exogenous variation in the endogenous variable of interest. The threats to identification are likely to be substantial, however, as the pre-treatment characteristics must stratify the units in a way that is orthogonal to any subsequent trends in the outcome apart from the effect of treatment.

As a cautionary tale of the difficulty of meeting such assumptions, we reexamine the results of a recent and influential paper by Kearney and Levine (2015b, henceforth KL) that claimed that the MTV reality shows *16 and Pregnant, Teen Mom*, and *Teen Mom 2* (henceforth collectively referred to as *16 and Pregnant*) caused a 4.3 percent drop in teen birth rates between July 2009 and December 2010. This effect is large and would account for a quarter of the total reduction in teen childbearing during this period. KL interpret their estimates as causal and write 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.

KL use the audience share watching MTV at a specific time of day in each media market in the year before *16 and Pregnant* began broadcasting, interacted with a post-*16 and Pregnant*

dummy variable as an instrument for the *16 and Pregnant* viewership. There is little theoretical justification for this instrument, however, because MTV viewership is a choice variable that is likely to be determined by the same demographic and socioeconomic factors that determine *16 and Pregnant* viewership. Because the same demographic and socioeconomic factors are likely to be associated with teen births, it is natural to question the exogeneity of the instrument. KL's identification fundamentally relies on comparing changes in teen birth rates before and after *16 and Pregnant* in media markets with relatively high- and low-MTV viewership in the year before the show.

Identification strategies that use fixed unit characteristics (like MTV viewership in the pre-treatment period) interacted with something that varies over time (like the post-treatment dummy variable) are quite common. In another recent example, Nunn and Qian (2014) examine the impact of U.S. food aid on conflict in recipient countries. They use U.S. wheat production in the period prior to aid receipt interacted with the fraction of years that a country received any food aid in a fixed twenty-five-year period to create an instrument for U.S. food aid that varies by country and time period. Nunn and Qian note explicitly that their identification strategy is a difference-in-differences analysis. In a critique similar to ours about KL, Christian and Barrett (2017) conclude that the parallel trends assumption inherent in Nunn and Qian's analysis does not hold.

As noted by Goldsmith-Pinkham, Sorkin, and Swift (2018), KL's identification strategy is also related to the use of "Bartik" or shift-share instruments common in the trade literature, and the literature estimating the impact of immigration on wages and other outcomes, among others. Goldsmith-Pinkham, Sorkin, and Swift (2018) discuss the general conditions for validity of Bartik instruments in static contexts, while Jaeger, Ruist, and Stuhler (2018) raise concerns about such instruments in presence of dynamic adjustment processes, particularly in the immigration

literature. All of these studies can be viewed as essentially relying on a difference-in-differences research design on the basis of pre-treatment characteristics.

KL's reliance on a variant of the "shift-share" methodology stands in marked contrast to other studies of the impact of media (Gentzkow 2006, Della Vigna and Kaplan 2007, Gentzkow and Shapiro 2008, Jensen and Oster 2009, La Ferrara, Chong, and Duryea 2012, Campante and Hojman 2013, Kearney and Levine 2015a, among others). In all of these studies, the authors exploit plausibly exogenous changes in media exposure across *both* space and time. The addition of spatial variation allows these authors to control more convincingly for trends that may confound the treatments in question.

The assumption necessary for a causal interpretation of KL's results is that trends in teen birth rates in geographic areas with high- and low-MTV viewership would have been the same, conditional on area and time fixed effect and a limited set of covariates, in the absence of *16 and Pregnant*. How likely is it that this identifying assumption holds? Teen birth rates began falling precipitously with the onset of the Great Recession in 2008 and there were profound racial and ethnic differences in the rate of decline. Given the substantial geographic variation in racial/ethnic composition, trends in birth rates in the media markets used by KL are also likely to vary because of different trends in birth rates by race/ethnicity. These differences may confound KL's estimates and violate the parallel trend assumption necessary for a causal interpretation.

This is precisely what we find. Allowing for differential time trends by the racial/ethnic composition of a media market eliminates any association between *16 and Pregnant* and teen birth rates, as well as the birth rates of women ages 20 to 24 and 25-29. We obtain the same null findings when we interact the racial/ethnic composition of an area with the unemployment rate. The importance of differential time trends in teen birth rates prior to *16 and Pregnant* is evident from

KL's own results. They report no association between *16 and Pregnant* and the birth rates of non-Hispanic black and Hispanic teens when their analyses are stratified by race/ethnicity, despite national data indicating that black and Hispanic young women watched the show as much, or even more, than their white counterparts (Kearney and Levine 2014).

To assess whether parallel trends in birth rates between low- and high-MTV viewing areas held before *16 and Pregnant* began broadcasting, we perform a series of placebo tests and find an "effect" of *16 and Pregnant* prior to the advent of the show. Using Kearney and Levine's "IV Event Study" methodology, we also show that including even a few additional periods in the analysis prior to the beginning of *16 and Pregnant* leads to a rejection of the hypothesis that pre-treatment trends were parallel. Combined with our in-sample results, our out-of-sample rejection of the parallel trends assumption leave little question that KL's results do not warrant a causal interpretation.

## I.    KL's Empirical Framework

KL's empirical model consists of the following four equations

(1)          $\ln(BR_{jt}) = \beta_{10} + \beta_{11}(Rate16P_j \times Post_t) + \beta_{12}U_{jy} + \mathbf{X}_{jt}\gamma_1 + \theta_{1t} + \delta_{1js} + \epsilon_{1jt}$

(2)          $\ln(BR_{jt}) = \beta_{20} + \beta_{21}(\widehat{Rate16P_j} \times Post_t) + \beta_{22}U_{jy} + \mathbf{X}_{jt}\gamma_2 + \theta_{2t} + \delta_{2js} + \epsilon_{2jt}$

(3) $(Rate16P_j \times Post_t) = \beta_{30} + \beta_{31}(MTV0809_j \times Post_t) + \beta_{32}U_{jy} + \mathbf{X}_{jt}\gamma_3 + \theta_{3t} + \delta_{3js} + \epsilon_{3jt}$

and

(4)          $\ln(BR_{jt}) = \beta_{40} + \beta_{41}(MTV0809_j \times Post_t) + \beta_{42}U_{jy} + \mathbf{X}_{jt}\gamma_4 + \theta_{4t} + \delta_{4js} + \epsilon_{4jt}$

where $BR_{jt}$ 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* for 12 to 24 year-old viewers for the seven Nielsen sweep months, when viewership is measured to set advertising rates for the following period, of July 2009, February 2010, July 2010, November 2010, May 2011, November 2011, and May 2012; $Post_t$ is a dichotomous indicator for the period after the introduction of *16 and Pregnant*; $MTV0809_j$ is the average Nielsen rating 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 a vector that includes the percent population in the DMA that is that non-Hispanic black, and the percent that is Hispanic in calendar year $y$; the $\theta_t$'s represent time period (quarter) fixed effects; and the $\delta_{js}$'s are full sets of DMA × season fixed effects, which also implicitly include DMA fixed effects.  Note that we have modified KL's notation to reflect that the parameters and disturbance terms differ across equations.

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.  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. 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. KL average these Tuesday ratings within each month and then average the four months of ratings 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 ratings that KL use for *16 and Pregnant* and MTV therefore vary only in the cross section (by DMA).

KL use the interaction between a time-invariant, DMA-specific measure of MTV viewership and an indicator for the post-*16 and Pregnant* period as the instrument for the viewership of *16 and Pregnant*. For the instrument to be valid, MTV viewership in the pre-*16 and Pregnant Period* must "randomly" assign *16 and Pregnant* viewership, conditional on DMA fixed effects and a small number of covariates. Because of the inclusion of DMA fixed effects, this interaction is necessary to generate time series variation in the instrument, which consists of zeros in all DMAs and all quarters from 2005:QI to 2009:QII with a discrete jump for each DMA in 2009:QIII. The measure of *16 and Pregnant* is also time invariant in the post-treatment period (and obviously zero in the pre-treatment period), generating a strong first-stage association.

## II.    Within-Sample Evaluation of KL's Identifying Assumptions

KL's reduced form compares teen birth rates before and after the introduction of *16 and Pregnant,* stratified by levels of MTV viewership in the year before the show began. The key identification assumption is therefore parallel trends: in the absence of *16 and Pregnant*, birth rates should change in the same way in areas with high and low MTV viewership. Figure 1 shows

6

that overall teen birth rates decline significantly prior to *16 and Pregnant* and that the rate of decline accelerates after 2007 with the onset of the Great Recession. There is substantial variation in the rate of decline across race/ethnicity groups. The average annual change in teen births between 2005 and 2010 was -1.9 percent among non-Hispanic whites, -2.9 percent among non-Hispanic blacks and -6.3 percent among Hispanics (our calculations based on data reported in Hamilton, *et al.* 2015).

While differential trends across race/ethnicity groups in birth rates do not, by themselves, invalidate KL's identification strategy, the likelihood that these differential trends will cause the parallel trends assumption to be violated increases to the extent to which DMAs vary in their ethnic composition. We find that the correlation between MTV viewership (in 2008-9) and the non-Hispanic black share of the female population (in 2005) is quite high (0.413) while substantially lower for the Hispanic share (0.093). This masks some variation, however, for Hispanics. In 2005, the Hispanic share of the female population across the four quartiles (lowest to highest) of MTV viewership was 11.1, 19.8, 16.6 and 18.1, respectively. The corresponding values for the non-Hispanic black share in 2005 across quartiles of MTV viewership were monotonic: 7.6, 11.6, 13.9, and 20.3, respectively. These substantial differences in race/ethnicity across quartiles of the instrument are a threat to KL's identification strategy if there are differential trends by racial groups within high- and low-MTV watching areas. This is what we find in Figure 2, which presents trends in log teen birth rates for each of the race/ethnicity groups that KL examine for groups of DMAs by quartile of MTV viewership in 2008-09. For non-Hispanic blacks and Hispanics, in particular, the trends in log teen birth rates across MTV quartiles do not appear to be parallel.

To assess formally the importance of this potential confounding for KL's estimates, we augment their model with time trends that vary with levels of the exogenous covariates, particularly the pre-*16 and Pregnant* racial/ethnic composition. This approach has been used in several recent articles (Hoynes and Schanzenbach 2009; Hoynes, Schanzenbach, and Almond 2016; and Hjort, Sølvsten, and Wüst, 2017). Large changes in the IV and reduced form coefficients after including these trends would suggest that KL's estimates are biased.

The results are shown in Table 1. In each panel, we show the reduced form estimates from equation (4) and the IV estimates from equation (2), which include DMA and period fixed effects. Panel 1 exactly replicates KL's reduced form and IV estimates for comparison. In panel 2 we show estimates from a specification in which we interact the values of the covariates, **X**, in 2005 (well before the start of *16 and Pregnant*) with the period fixed effects. Both the reduced form relationship between MTV and the IV estimate of the impact of *16 and Pregnant* are reduced by a factor of approximately 10 and are no longer statistically significantly different from zero.

The results in panel 2 strongly suggest that the parallel trends assumption in KL's model does not hold. In panel 3 we reduce the possibility of overfitting by estimating a more parsimonious specification, adding only interactions of the three covariates with a linear trend to the base models from equations (4) and (2). These results are very similar to those in panel 2 and, again, suggest that the parallel trends assumption does not hold. In panel 4 we limit the trend interactions to the percent non-Hispanic black and percent Hispanic in 2005, yielding reduced form and IV coefficients that are still approximately only 1/3 the size of those in panel 1 and are not statistically significantly different from zero. In contrast, in panel 5 we include only the interaction of a trend with the unemployment rate in 2005 and obtain results that are quite similar to KL's in panel 1.

As Figures 1 and 2 suggest, the results in panels 4 and 5 lead to the conclusion that the effect of *16 and Pregnant* that KL estimate is largely driven by differential trends in birth rates across race and ethnic groups that are correlated with the timing of the introduction *16 and Pregnant*. We confirm this in panels 6 through 8, where we simplify the trend interactions further and merely interact the covariate levels in 2005 with *Post$_t$*. The single interaction of the covariate levels in 2005 with the post-period dummy is nested within the specification that interacts the same covariates with the complete set of time fixed effects. A Wald test indicates that we cannot reject the restrictions implied in panel 6 relative to panel 2. The results are essentially the same as in panels 3 through 5. Taken together, the results in Table 1 strongly suggest that the parallel trend assumption underlying the KL analysis does not hold and that a causal interpretation of KL's estimated impact of *16 and Pregnant* is unwarranted and, indeed, that *16 and Pregnant* had no meaningful effect on teen birth rates.

### *Stratification by Age*

In their Table 2, KL present results of the effect of *16 and Pregnant* stratified by age and race/ethnicity. We exactly replicate their estimates for 20 to 24 year olds (panel 1), 25 to 29 year olds (panel 2) and 30 to 34 year olds (panel 3) in the odd-numbered rows in our Table 2. In the even-numbered rows, we modify those specifications by including, as in panel 3 of Table 1, a linear time trend interacted with the level of the three covariates in 2005. We have also estimated these models using both the full set of period × 2005 covariate interactions as well as the 2005 covariates interacted with the *Post$_t$* dummy. The results are qualitatively similar and available from the authors by request. As in Table 1, when we include these interactions, we find no

statistically significant associations between *16 and Pregnant* and birth rates by age. For example, KL report that a one ratings point increase in the viewership of *16 and Pregnant* lowers birth rates of 20-24 year-olds by 2.4 percent (row 3, panel 1). Inclusion of the additional trend terms reduces the effect to a statistically insignificant reduction (*p*-value of 0.19) of 0.9 percent (row 4, panel 1). We find the same pattern for 25-29 and 30-34 year-olds, respectively.

The results in Tables 1 and 2 suggest that the estimated association between *16 and Pregnant* and teen birth rates reported by KL is due to differential trends in birth rates across areas in ways that are associated with MTV viewing. The more challenging task is to explain why. As Figure 1 shows, the differential rate of decline in teen birth rates by race/ethnicity appears to have increased with the onset of the Great Recession. Several recent studies demonstrate that the Great Recession had particularly large effects on the fertility of young women (Sobotka, Skirbekk, and Philipov, 2011; Cherlin, *et al*. 2013). Autor, Dorn, and Hanson (2017) show that recent trade shocks that affect the employment opportunities of less educated men and women decrease both marriage and birth rates. Kearney and Wilson (2017) have documented that improved economic conditions from the fracking boom may lead to increased non-marital and marital fertility. Most relevant for our analysis, Villarreal (2014) finds that areas with relatively large Hispanic populations may have been more affected by the Great Recession through decreased employment and flow of immigrants.

To explore the differential response of birth rates to the Great Recession, in Table 3 we interact each of the three covariate levels in 2005 with the unemployment rate. We find that for teens, the IV coefficient on *16 and Pregnant* falls from -2.37 (Table 2, row 3) to -0.562 (Table 3, row 2). For this group, the interaction between the share of the DMA that is black in 2005 and the unemployment rate is negative and statistically significant – consistent with the hypothesis that

10

worse labor market conditions lowered birth rates differentially for minority populations. We find similar results for Hispanics in the other age groups. For all age groups, the interaction between the unemployment rate in 2005 and the contemporaneous unemployment rate are also statistically significant, suggesting a non-linear relationship between birth rates and labor market conditions. As in Tables 2 and 3, allowing for the effect of the racial/ethnic composition on birth rates to vary over time reduces the reduced form and IV coefficients such that they are no longer statistically significant.

*Stratification by Race/Ethnicity*

KL stratify by race/ethnicity and find that the effects of *16 and Pregnant* are statistically significant for non-Hispanic whites and marginally so for Hispanics. In the odd-numbered rows of Table 4 we replicate these results for all three race/ethnicity groups. KL do not include race/ethnicity variables in these regressions. To test the parallel trends assumption, therefore, we now include a full set of DMA-specific time trends. KL find that an increase in *16 and Pregnant* viewership lowers the birth rates of non-Hispanic whites by 2.4 percent (panel 1, row 3). When we include DMA-specific trends, this coefficient turns positive and is statistically insignificantly different from zero (panel 1, row 4). We find no statistically significant reduced form or IV effects for non-Hispanic blacks (panel 2) or Hispanics (panel 3). As with the full sample, we find little in these results to suggest that *16 and Pregnant* affected birth rates in any of the race/ethnicity groups.

## III.    Out-of-Sample Evaluation of KL's Identifying Assumptions


As Kahn-Lang and Lang (forthcoming) note, differences in mean outcomes between the treatment and control groups at some point prior to the beginning of treatment suggest that parallel trends between the two groups do not hold unless there is some identifiable explanation.  One way to assess these differences is to extend the pre-treatment period.  The choice of how long into the past to extend the pre-treatment period is, of course, arbitrary.  This is a practical not a conceptual issue, however, as the beginning of the pre-treatment period is usually determined arbitrarily, by data availability, or by path dependence in the conduct of the analysis.  Even if there are level differences at some point in the past, small changes in the length of the pre-treatment period should not greatly affect estimated treatment effects if the parallel trends assumption holds.

In practice, there are few guidelines for what constitutes a sufficiently long (or short) pre-treatment period, which may lead to "cherry picking" results.  Robustness of results to pre-treatment periods of different lengths would increase confidence in those results, particularly when the pre-treatment period includes geographic and over-time variation in factors that are known predictors of the outcome. Periods, like the Great Recession, in which the covariates differ in their relationship to the outcome relative to other periods, may be ill-suited for inclusion in the analysis.

We conduct two types of analyses to assess the validity of the parallel trends assumption. In the first, we do placebo tests to assess the "impact" of *16 and Pregnant* in periods before the show was broadcast.  In the second, we replicate KL's "Reduced Form Event Study" and incrementally extend the pre-treatment period further into the past.

*Placebo Tests*

As we have noted, MTV ratings should be unrelated to trends in birth rates conditional on time and DMA fixed effects except through their relationship with *16 and Pregnant* ratings. Any association between MTV ratings and birth rates in the pre-*16 and Pregnant* period would suggest a violation of the exclusion restriction (i.e., parallel trends) necessary for identification. In Table 5, 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 interacted with a post-"show" indicator as the instrument.

Each row of Table 5 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 2001:Q1 and ends in 2006:QIV, with the placebo show beginning in 2005:QIII. Row 12, which is shaded in the table, reproduces KL's reduced form and instrumental variables results from their Table 1 columns (4) and (3), respectively. For the reduced form, we find that 9 of the 11 estimated coefficients in rows (1) through (11), which use data exclusively from the pre-*16 and Pregnant* period, are statistically different from zero at conventional levels. That all of the estimates are negative and roughly equal to KL's reduced form result in magnitude strongly suggests that that the parallel trends assumption required for the difference-in-differences strategy does not hold. 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 (1) to 1.524 in

13

row (9); the first-stage coefficient is 1.513 in row (12), which corresponds to KL's analysis period. The IV results therefore reflect those from the reduced form.

Adjusting for multiple comparisons in Table 5 using the Bonferroni (1936) correction or the somewhat more powerful Bonferroni-Holm (Holm 1979) correction leads to rejection of none of the null hypotheses that the coefficients in both the reduced form and instrumental variables placebo results equal to zero. 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. An alternative procedure is to focus on the so-called *false discovery rate* (FDR), 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. When we perform the Benjamini-Hochberg (1995) on the 11 placebo reduced form tests, we do not reject any of the hypotheses implicit in Table 5 when we fix the FDR at 5 or 10 percent. Increasing the FDR to 11 percent leads to the rejection of the null in rows 7 through 9, and increasing it to twenty percent allows us to reject the null in all but row 1. If we fix the FDR at 5 percent for the placebo instrumental variables results, we reject the null that the coefficient in row 7 is equal to zero. Increasing the FDR to 10 percent, leads to the rejection of the null hypotheses that the coefficients in rows 6 through 10 are separately equal to zero, and if the FDR is set at 20 percent we reject the nulls in rows 4 through 11.

The IV estimates in Table 5 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 (7) our IV estimate suggests that teen birth rates decrease 2.194 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 unobserved and time-varying factors within the DMA that are correlated with teen birth rates and *16 and Pregnant* viewership. Finding any statistically significant coefficients in a period before *16 and Pregnant* began broadcasting strongly suggests that KL's instrument is invalid.

Although we have presented a series of placebo results in order to keep from "cherry picking" those that are particularly favorable to our argument, we recognize that the results in Table 5 are not independent of one another when the "treatment" periods overlap. We have therefore boxed rows 1 and 7 to illustrate results where the "post" periods are disjoint (and as indicated above, all placebo "treatment" periods are disjoint from the actual period when *16 and Pregnant* was being broadcast). Focusing only on rows 1, 7, and 12, on the basis of the coefficients and their statistical significance, it would be virtually impossible to distinguish which of these results was generated by the actual *16 and Pregnant* and which were placebo tests.

***"Reduced Form Event Study"***

KL present results from an "event study" of the reduced-form relationship between MTV viewership and teen birth rates (KL Figure 5 and Appendix Table B1). Their primary goal is to assess the validity of the parallel trends assumption. They estimate the following regression:

(5)
$$\ln(BR_{jt}) = \beta_0 + \sum_{i=1}^{14} \alpha_i (MTV0809_j \times PreQ_i) +$$
$$\sum_{i=19}^{24} \beta_i (MTV0809 \times PostQ_i) + \phi U_{jy} + \theta_t + \delta_{js} + \epsilon_{jt}.$$

The coefficients $\alpha_i$ 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-2009:QII) relative to the reference category (quarters 15 to 18, 2008:QIII-2009:QII). The $\beta_i$

15

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 reported that the joint null hypothesis that $\alpha_1 = 0, \alpha_2 = 0, \dots, \alpha_{14} = 0$ cannot be rejected, but the null hypothesis that $\beta_{19} = 0, \beta_{20} = 0, \dots, \beta_{24} = 0$ is rejected ($p<0.01$).

KL take the lack of a statistically significant relationship between MTV and birth rates in the pre-*16 and Pregnant* period as evidence that their instrument meets the exclusion restriction. We test this assumption further by extending the event study analysis backwards in increments of two years. In Figure 3, Panels 1, 2, and 3, we present results from estimating equation (5) for periods with the starting year in 2005 (KL's period), 2003, and 2001, respectively. We present the $\chi^2$ statistic for the joint test of the coefficients in the pre-*16 and Pregnant* period. Unlike KL, we use the first four quarters in each panel as the reference period. The $\chi^2$ tests for the coefficients in the pre-*16 and Pregnant* period are invariant to the choice of a reference category (in the pre-treatment period). KL's choice of using the four quarters prior to the start of *16 and Pregnant* as the reference category is arbitrary, tends to obscure the smooth trends in the birth rates, and gives the appearance of a discontinuity when, in fact, a break in trend is difficult to discern.

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 quite apparent in all three panels and we strongly reject the null hypothesis that the pre-*16 and Pregnant* coefficients are jointly equal to zero in Panels 2 and 3 with *p*-values of 0.036 and <0.001, respectively. These results are not artifacts of greatly extending the period of analysis. Extending the pre-*16 and Pregnant* period even by only two years (Panel A of Table 2) results in rejection of the null that pre-treatment coefficients are jointly equal to zero. The sensitivity of KL's results to extending the pre-treatment period by even a small amount suggests that the assumptions needed to identify a causal effect are unlikely to be met.

Although there is no formal and global test of the parallel trends assumption, even a lack of rejection of differential trends cannot disprove a change in trends for the counterfactual group in the post-treatment period.  In the case of *16 and Pregnant*, a number of factors suggest that underlying trends in teen fertility may be confounding KL's estimate of its impact.   As we show in Figure 2, there are clearly differential trends in teen birth rates by race/ethnicity and MTV viewership prior to the beginning of *16 and Pregnant*.  There is also a noticeable dip in the reduced-form relationship between MTV viewership and birth rates beginning in 2008, as shown in panels B and C of Figure 3, that is coincident with the onset of the Great Recession.  A point-in-time change at the national level is more vulnerable to coincident shocks than interventions that occur across time and space.

**IV.    Conclusion**

Identification strategies that interact fixed pre-treatment characteristics with variables that vary over time are quite common and may seem particularly enticing for evaluating policies or treatments that begin at a single point in time for an entire population.  In such cases, the lack of an obvious counterfactual for the treated group in the post-treatment period makes identification of causal effects difficult, and studies claiming to estimate a causal effect in in the face of such challenges should be viewed with appropriate skepticism.

As an example of the difficulties associated with identification in these situations, we have reexamined causal claims that the MTV program *16 and Pregnant* was responsible for lowering

birth rates substantially in the United States.  Our results strongly suggest that the association between *16 and Pregnant* and teen birth rates is due to differential trends in birth rates across areas with different demographic characteristics in the wake of Great Recession. The fertility rates of less advantaged teens, as proxied by race/ethnicity, fell much faster with the onset of the recession than those of white non-Hispanics. We find a relationship between MTV and birth rates even in the period prior to the existence of *16 and Pregnant*, casting substantial doubt on the validity of the parallel trends assumption required for identification. While is possible that *16 and Pregnant* may have affected teen birth rates, our results point to the difficulty of making a causal claim that the show, which debuted nationally at the nadir of the Great Recession, had a significant impact on teen birth rates.

Our results add to a growing literature that highlights the difficulties associated with identification strategies that employ instruments derived from interacting pre-treatment characteristics with a time varying variable.  Future researchers wishing to identify a causal effect of a policy that was introduced everywhere at a single point in time, who may be tempted to use similar instruments, should recognize that "shift-share" instruments are not a panacea and be careful to assess thoroughly whether the necessary assumptions inherent in such research designs are credible.

# Acknowledgements

## References

Autor, D., D. Dorn, and G. Hanson (2017) "When Work Disappears: Manufacturing Decline and the Falling Marriage-Market Value of Men," Mimeo. MIT Department of Economics. Available at https://economics.mit.edu/files/12736.

Campante, Filipe R. and Daniel A. Hojman (2013) "Media and Polarization: Evidence from the Inroduction of Broadcast TV in the United States," *Journal of Public Economics* 100, 79–92.

Cherlin, A., E. Cumberworth, S. P. Morgan, and C. Wimer (2013) "The Effects of the Great Recession on Family Structure and Fertility," *Annals of the American Academy of Political and Social Science* 650, 214–231.

Christian, Paul and Christopher B. Barrett (2017) "Revisiting the Effect of Food Aid on Conflict: A Methodological Caution," World Bank Policy Research Working Paper 817.

Della Vigna, S., and E. Kaplan (2007) "The Fox News Effect: Media Bias and Voting," *Quarterly Journal of Economics* 122, 1187–1234.

Dobkin, C., A. Finkelstein, R. Kluender, and M. J. Notowidigdo (2016) "The Economic Consequences of Hospital Admissions." NBER Working Paper 22288.

Gentzkow, M. (2006) "Television and Voter Turnout," *Quarterly Journal of Economics* 121, 931–972.

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

Goldsmith-Pinkham, Isaac Sorkin, and Henry Swift (2018) "Bartik Instruments: What, When, Why, and How," NBER Working Paper 24408.

Hamilton, B. E, J.A. Martin, M. J. K. Osterman, S. 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.

Hjort, J., M. Sølvsten, and M. Wüst. (2017) "Universal Investment in Infants and Long-Run Health: Evidence from Denmark's 1937 Home Visiting Program," *American Economic Journal: Applied Economics* 9, 78–104.

Hoynes, H. W. and D. W. Schanzenbach (2009) "Consumption Responses to In-Kind Transfers: Evidence from the Introduction of the Food Stamp Program," *American Economic Journal: Applied Economics* 1, 109–139.

Hoynes, H. W., D. W. Schanzenbach, and Douglas Almond (2016) "Long-Run Impacts of Childhood Access to the Safety Net," *American Economic Review* 106 (4): 96-34.

Jensen, R., and E. Oster (2009) "The Power of TV: Cable Television and Women's Status in India," *Quarterly Journal of Economics* 124, 1057–1094.

Kahn-Lane, Ariella and Kevin Lang (forthcoming) "The Promise and Pitfalls of Differences-in-Differences: Reflections on 'Sixteen and Pregnant' and Other Applications," *Journal of Business and Economic Statistics*.

Kearney, M. S., and P. B. Levine (2014) "Media Influences on Social Outcomes: The Impact of MTV's 16 and Pregnant on Teen Childbearing," National Bureau of Economic Research Working Paper 19795. Available at http://www.nber.org/papers/w19795.

_____ (2015a) "Early Childhood Education by MOOC: Lessons from Sesame Street," National Bureau of Economic Research Working Paper 21229. Available at http://www.nber.org/papers/w21229.

_____ (2015b) "Media Influences on Social Outcomes: The Impact of MTV's *16 and Pregnant* on Teen Childbearing," *American Economic Review* 105, 3597–3632.

Kearney, M. S., and R. Wilson (2017) "Male Earnings, Marriageable Men, and Nonmarital Fertility: Evidence from the Fracking Boom.," National Bureau of Economic Research Working Paper 23408. Available at http://www.nber.org/papers/w23408.

La Ferrara, E., A. Chong, and S. Duryea (2012) "Soap Operas and Fertility: Evidence from Brazil," *American Economic Journal: Applied Economics* 4, 1–31.

Nunn, Nathan and Nancy Qian (2014) "US Food Aid and Civil Conflict," *American Economic Review* 104, 1630-1666.

Sobotka, T., V. Skirbekk, and D. Philipov (2011) "Economic Recession and Fertility in the Developed World," *Population and Development Review* 37, 267–306.

**Figure 1**
Log(Teen Birth Rate) for Race/Ethnicity Groups
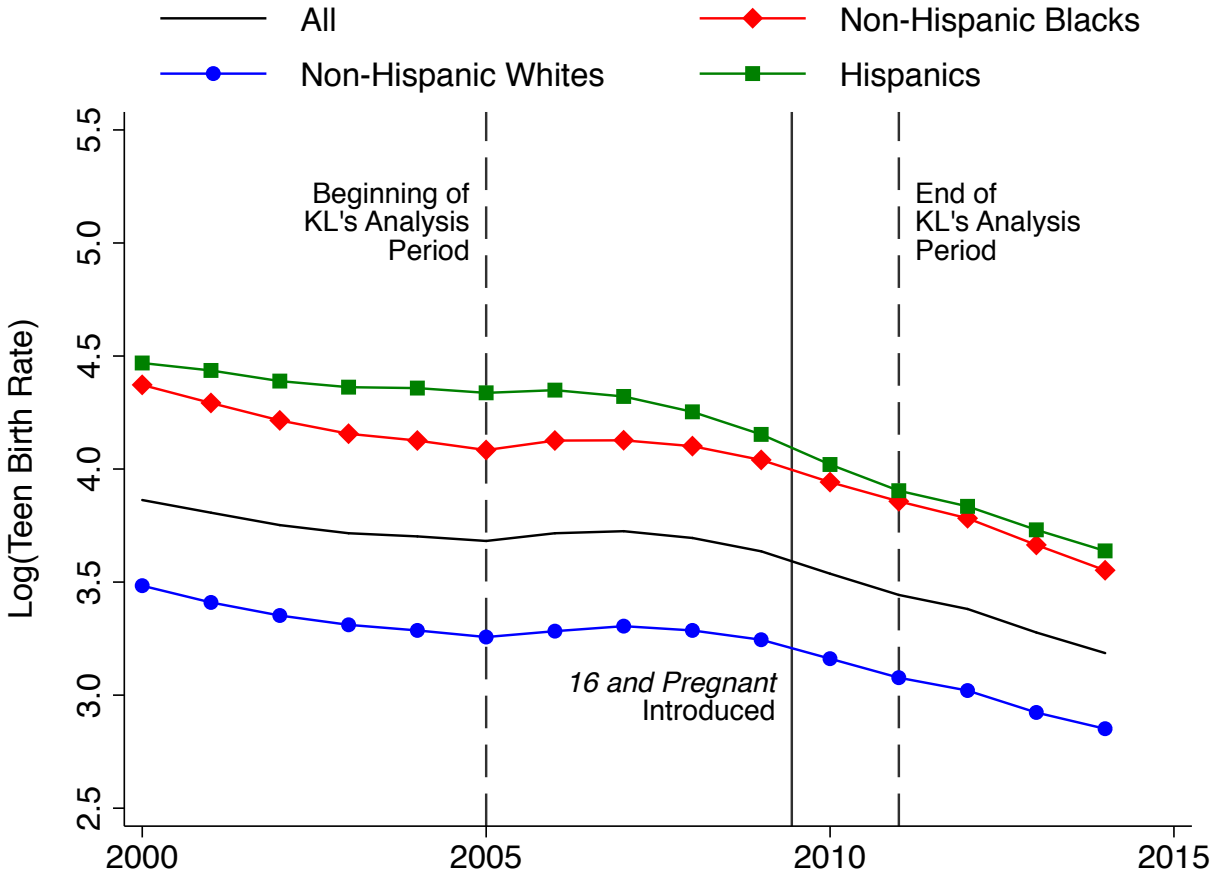
*Source:* Hamilton, et al. (2015)

## Figure 2
### Log(Teen Birth Rate) by MTV Ratings Quartile and Race

*Panel A:* Non-Hispanic Whites
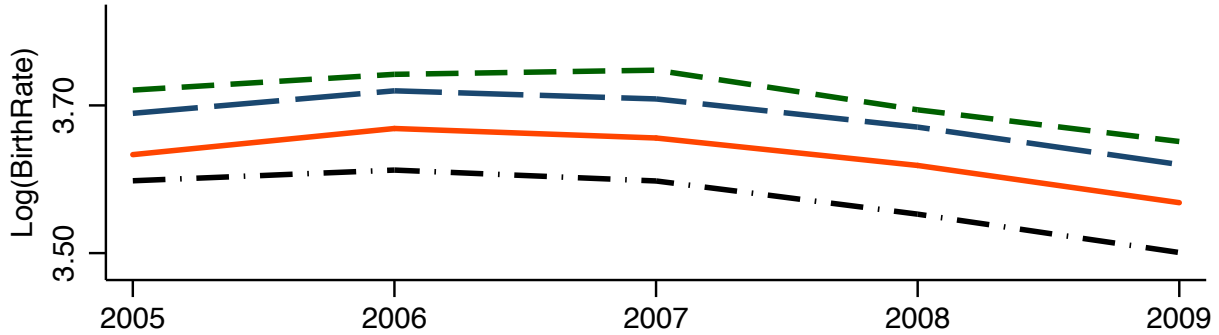
*Panel B:* Non-Hispanic Blacks

*Panel C:* Hispanics

1st    2nd    3rd    4th
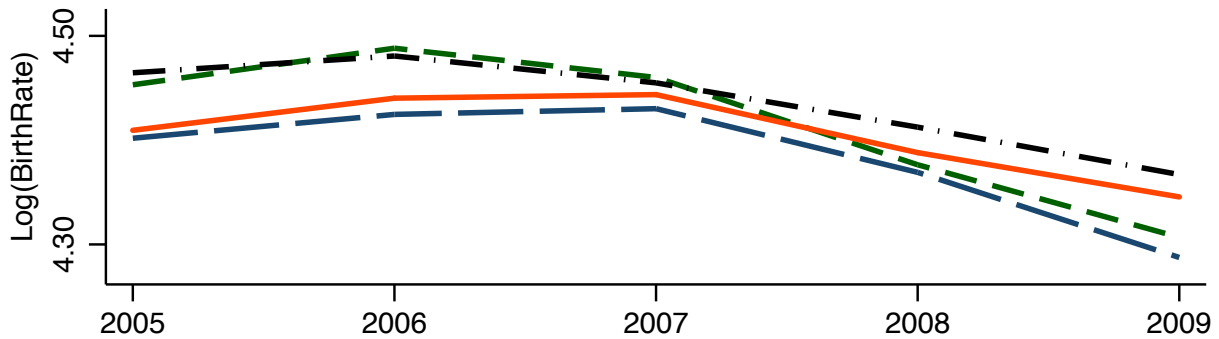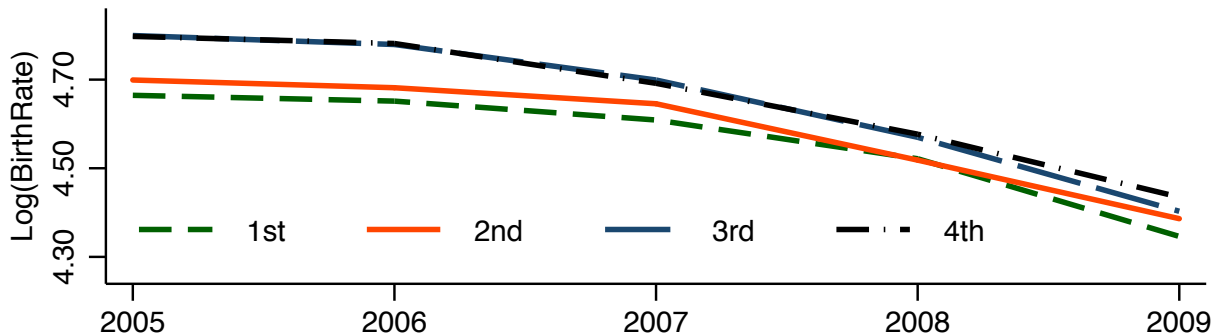
*Source:* Authors' calculations using Bureau of the Census Census County Characteristics, National Vital Statistics System births, and Nielsen ratings data. The latter two data sources are confidential.

*Notes:* All panels shows weighted average annual Log(Birth Rate) for teenagers by quartile of MTV ratings where the weights are the female population aged 15-19 in the DMA. 2009 includes only the 2 quarters prior to the start of *16 and Pregnant*. Note that the *y*-axis scale is different in each panel.

**Figure 3**
Reduced Form Event Study: Extending the Pre-*16 and Pregnant* Period
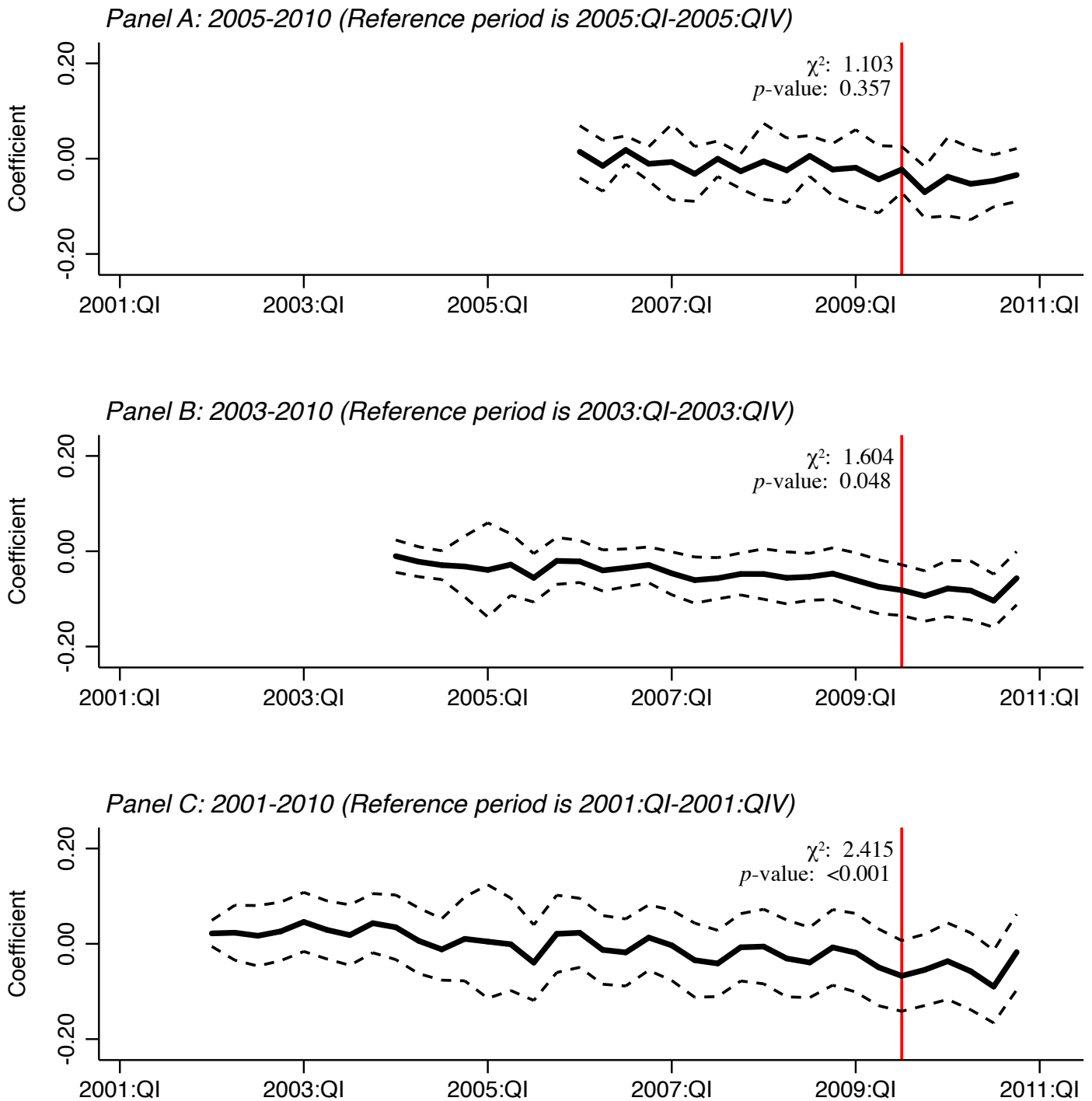
*Panel A: 2005-2010 (Reference period is 2005:QI-2005:QIV)*

χ²: 1.103
*p*-value: 0.357

*Panel B: 2003-2010 (Reference period is 2003:QI-2003:QIV)*

χ²: 1.604
*p*-value: 0.048

*Panel C: 2001-2010 (Reference period is 2001:QI-2001:QIV)*

χ²: 2.415
*p*-value: <0.001

*Source:* Authors' calculations using Bureau of Labor Statistics Local Area Unemployment Statistics, Bureau of the Census Census County Characteristics, National Vital Statistics System births, and Nielsen ratings data. The latter two data sources are confidential.

*Notes:* These figures plot the coefficients of average MTV Ratings interacted with period dummy variables. The regressions also include the unemployment rate, the percent non-Hispanic black, percent Hispanic, and a full set of period dummy variables and DMA × season interactions. Variance-covariance matrices are calculated by clustering by DMA. Ninety-five percent confidence intervals shown. The χ² tests shown are for the joint significance of the coefficients shown in the pre-*16 and Pregnant* period. Panel A shows the period used in Kearney and Levine (2015b). Regressions are weighted by the female population aged 15-19 in the DMA.

**Table 1**
Estimates of the Impact of *16 and Pregnant* on Teen Birth Rates

| | *16 & Preg.* Ratings or MTV ratings 2008-2009 | Unemp. Rate | Percent Non-Hisp. Black | Percent Hispanic | DMA Fixed Effects, Period Fixed Effects, and | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Period Fixed Eff. × | | | Linear Trend × | | | *Post_t* × | | |
| | | | | | % Non-Hisp. Black 2005 | % Hisp. 2005 | Un-emp. Rate 2005 | % Non-Hisp. Black 2005 | % Hisp. 2005 | Un-emp. Rate 2005 | % Non-Hisp. Black 2005 | % Hisp. 2005 | Un-emp. Rate 2005 |
| **Panel 1: Base model: Replication** | | | | | | | | | | | | | |
| Reduced form (KL) | -3.581 ** | -1.485 *** | -2.305 * | -2.812 *** | | | | | | | | | |
| | (1.517) | (0.043) | (1.267) | (0.636) | | | | | | | | | |
| Instrumental variables (KL) | -2.366 ** | -1.487 *** | -2.197 * | -3.103 *** | | | | | | | | | |
| | (0.941) | (0.375) | (1.197) | (0.553) | | | | | | | | | |
| **Panel 2: Base model plus covariates $_{2005}$ × Period Fixed Effects** | | | | | | | | | | | | | |
| Reduced form | -0.404 | -1.473 *** | -3.312 *** | -2.442 *** | X | X | X | | | | | | |
| | (1.504) | (0.360) | (1.012) | (0.553) | | | | | | | | | |
| Instrumental variables | -0.238 | -1.467 *** | -3.314 *** | -2.459 *** | X | X | X | | | | | | |
| | (0.800) | (0.318) | (0.914) | (0.489) | | | | | | | | | |
| **Panel 3: Base model plus covariates $_{2005}$ × Linear Trend** | | | | | | | | | | | | | |
| Reduced form | -0.437 | -1.414 *** | -3.295 *** | -2.470 *** | | | | X | X | X | | | |
| | -(1.309) | (0.349) | -(0.963) | -(0.536) | | | | | | | | | |
| Instrumental variables | 0.273 | -1.410 *** | -3.302 *** | -2.484 *** | | | | X | X | X | | | |
| | (0.749) | (0.314) | (0.877) | (0.481) | | | | | | | | | |
| **Panel 4: Base model plus Race/Ethnicity $_{2005}$ × Linear Trend** | | | | | | | | | | | | | |
| Reduced form | -1.237 | -1.300 *** | -3.091 *** | -3.032 *** | | | | X | X | | | | |
| | (1.494) | (0.429) | (1.150) | (0.631) | | | | | | | | | |
| Instrumental variables | -0.776 | -1.292 *** | -3.104 *** | -3.071 *** | | | | X | X | | | | |
| | (0.847) | (0.384) | (1.046) | (0.561) | | | | | | | | | |
| **Panel 5: Base model plus Unemployment $_{2005}$ × Linear Trend** | | | | | | | | | | | | | |
| Reduced form | -3.374 *** | -1.632 *** | -2.211 *** | -2.547 *** | | | | | | X | | | |
| | (1.303) | (0.354) | (1.155) | (0.516) | | | | | | | | | |
| Instrumental variables | -2.238 *** | -1.624 *** | -2.115 *** | -2.840 *** | | | | | | X | | | |
| | (0.828) | (0.316) | (1.095) | (0.463) | | | | | | | | | |
| **Panel 6: Base model plus Covariates $_{2005}$ × Post $_t$** | | | | | | | | | | | | | |
| Reduced form | -0.449 | -1.446 *** | -2.870 *** | -2.826 *** | | | | | | | X | X | X |
| | (1.554) | (0.364) | (1.163) | (0.525) | | | | | | | | | |
| Instrumental variables | -0.259 | -1.437 *** | -2.877 *** | -2.840 *** | | | | | | | X | X | X |
| | (0.835) | (0.325) | (1.056) | (0.459) | | | | | | | | | |
| **Panel 7: Base model plus Race/Ethnicity $_{2005}$ × Post $_t$** | | | | | | | | | | | | | |
| Reduced form | -1.441 | -1.348 *** | -2.763 *** | -3.053 *** | | | | | | | X | X | |
| | (1.742) | (0.421) | (1.228) | (0.593) | | | | | | | | | |
| Instrumental variables | -0.838 | -1.325 *** | -2.789 *** | -3.097 *** | | | | | | | X | X | |
| | (0.930) | (0.374) | (1.115) | (0.518) | | | | | | | | | |
| **Panel 8: Base model plus Unemployment $_{2005}$ × Post $_t$** | | | | | | | | | | | | | |
| Reduced form | -3.319 ** | 1.624 *** | -2.247 * | -2.647 *** | | | | | | | | | X |
| | (1.327) | (0.377) | (1.241) | (0.557) | | | | | | | | | |
| Instrumental variables | -2.209 *** | -1.612 *** | -2.152 *** | -2.935 *** | | | | | | | | | X |
| | (0.841) | (0.335) | (1.166) | (0.494) | | | | | | | | | |

*Source:* Authors' calculations using Bureau of Labor Statistics Local Area Unemployment Statistics, Bureau of the Census Census County Characteristics, National Vital Statistics System births, and Nielsen ratings data. The latter two data sources are confidential.

*Note:* Each row shows results from a separate regression. The dependent variable in all regressions is the natural logarithm of the birth rate for 15-19 year olds. Standard errors, clustered by DMA, are shown in parentheses. All regressions are weighted by the female population aged 15 to 19 in the DMA at the time of the observation. All regressions include 24 quarter fixed effects as well DMA × season fixed effects as regressors. The sample size in all regressions 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**

Estimates of the Impact of *16 and Pregnant* on Birth Rates, by Age Groups

| | | 16 & Preg. Ratings or MTV ratings 2008-2009 | Unemp. Rate | Percent Non-Hisp. Black | Percent Hispanic | Linear Trend × | | |
| | | | | | | % Non-Hisp. Black 2005 | % Hisp. 2005 | Un-emp. Rate 2005 |
|---|---|---|---|---|---|---|---|---|
| *Panel 1: Ages 20-24* | | | | | | | | |
| (1) | Reduced form (KL) | -3.498 *** | -1.726 *** | -0.820 | -1.057 | | | |
| | | (1.312) | (0.325) | (0.791) | (1.683) | | | |
| (2) | Reduced form | -1.389 | -1.459 *** | 0.319 | -1.965 | X | X | X |
| | | (1.245) | (0.027) | (0.788) | (1.434) | | | |
| (3) | Instrumental variables (KL) | -2.422 ** | -1.810 *** | -0.923 | 0.845 | | | |
| | | (0.993) | (0.308) | (0.706) | (1.622) | | | |
| (4) | Instrumental variables | -0.897 | -1.459 *** | 0.313 | -1.935 | X | X | X |
| | | (0.755) | (0.247) | (0.717) | (1.322) | | | |
| *Panel 2: Ages 25-29* | | | | | | | | |
| (5) | Reduced form (KL) | -2.344 *** | -0.721 *** | 1.394 | -2.029 | | | |
| | | (1.151) | (0.252) | (0.904) | (1.313) | | | |
| (6) | Reduced form | -1.789 | -0.623 *** | 1.392 | -1.962 | X | X | X |
| | | (1.140) | (0.256) | (0.945) | (1.318) | | | |
| (7) | Instrumental variables (KL) | -1.606 *** | -0.800 *** | 1.504 * | -1.976 | | | |
| | | (0.817) | (0.229) | (0.825) | (1.202) | | | |
| (8) | Instrumental variables | -1.136 | -0.629 *** | 1.406 | -1.941 | X | X | X |
| | | (0.700) | (0.235) | (0.877) | (1.199) | | | |
| *Panel 3: Ages 30-34* | | | | | | | | |
| (9) | Reduced form (KL) | -0.455 | -0.897 *** | 0.538 ** | -2.463 *** | | | |
| | | (1.016) | (0.165) | (0.241) | (0.843) | | | |
| (10) | Reduced form | -0.091 | -0.730 *** | -0.539 ** | -0.652 *** | X | X | X |
| | | (0.496) | (0.139) | (0.252) | (0.813) | | | |
| (11) | Instrumental variables (KL) | -0.328 | -0.910 *** | -0.504 ** | -2.449 *** | | | |
| | | (0.689) | (0.146) | (0.243) | (0.781) | | | |
| (12) | Instrumental variables | -0.141 | -0.730 *** | -0.529 ** | -2.651 *** | X | X | X |
| | | (0.838) | (0.153) | (0.246) | (0.739) | | | |

*Source:* Authors' calculations using Bureau of Labor Statistics Local Area Unemployment Statistics, Bureau of the Census Census County Characteristics, National Vital Statistics System births, and Nielsen ratings data. The latter two data sources are confidential.

*Note:* Each row shows results from a separate regression. The dependent variable is the natural logarithm of the age-specific quarterly birth rate (multipled by 100) in the DMA. Odd-numbered rows reproduce Kearney and Levine's (2015) results from their Table 2. All regressions are weighted by the relevant population at the time of the observation. All regressions include 24 quarter fixed effects as well DMA × season fixed effects as regressors. The sample sizes are 4,920, 4,920, and 4,918 in panels 1, 2, and 3, respectively. *** indicates significant at the 1 percent level, ** indicates significant at the 5 percent level, * indicates significant at the 10 percent level.

**Table 3**

Estimates of the Impact of *16 and Pregnant* on Birth Rates, by Age Groups

with Covariate Interactions

| | *16 & Preg.* **Ratings or MTV ratings 2008-2009** | **Unemp. Rate** | **Percent Non-Hisp. Black** | **Percent Hispanic** | **Unemployment Rate ×** | | |
| | | | | | **Unemp. 2005** | **Percent Non-Hisp. Black 2005** | **Percent Hispanic 2005** |
|---|---|---|---|---|---|---|---|
| *Panel 1: Ages 15-19* | | | | | | | |
| (1) Reduced form | -0.885 | -2.836 *** | -2.952 ** | -2.864 *** | 0.371 *** | -0.029 *** | -0.008 |
| | (1.402) | (0.649) | (1.155) | (0.515) | (0.117) | (0.010) | (0.007) |
| (2) Instrumental variables | -0.562 | -2.765 *** | -2.960 *** | -2.899 *** | 0.367 *** | -0.030 *** | -0.009 |
| | (0.821) | (0.606) | (1.052) | (0.459) | (0.107) | (0.009) | (0.006) |
| *Panel 2: Ages 20-24* | | | | | | | |
| (3) Reduced form | -2.236 | -2.263 *** | 0.071 | -1.622 | 0.238 ** | -0.008 | -0.020 *** |
| | (1.426) | (0.511) | (0.735) | (1.621) | (0.089) | (0.012) | (0.007) |
| (4) Instrumental variables | -1.457 | 0.075 | -1.573 | -1.573 | 0.228 *** | -0.009 | -0.024 *** |
| | (0.919) | (0.670) | (1.509) | (1.509) | (0.084) | (0.011) | (0.006) |
| *Panel 3: Ages 25-29* | | | | | | | |
| (5) Reduced form | -1.558 | -0.510 | 1.461 | -2.094 | 0.047 | -0.009 | -0.010 ** |
| | (1.234) | (0.359) | (0.929) | (1.352) | (0.050) | (0.010) | (0.005) |
| (6) Instrumental variables | -1.003 | -0.412 | 1.516 * | -2.084 * | 0.042 | -0.010 | -0.013 *** |
| | (0.770) | (0.349) | (0.849) | (1.230) | (0.047) | (0.009) | (0.004) |
| *Panel 4: Ages 30-34* | | | | | | | |
| (7) Reduced form | -0.306 | -0.643 ** | -0.699 *** | -2.573 *** | 0.011 | 0.003 | -0.009 *** |
| | (0.923) | (0.317) | (0.261) | (0.864) | (0.054) | (0.006) | (0.003) |
| (8) Instrumental variables | -0.202 | -0.624 *** | -0.686 *** | -2.569 *** | 0.010 | 0.002 | -0.010 *** |
| | (0.561) | (0.295) | (0.247) | (0.790) | (0.049) | (0.006) | (0.003) |

*Source:* Authors' calculations using Bureau of Labor Statistics Local Area Unemployment Statistics, Bureau of the Census Census County Characteristics, National Vital Statistics System births, and Nielsen ratings data. The latter two data sources are confidential.

*Note:* Each row shows results from a separate regression. The dependent variable is the natural logarithm of the age-specific quarterly birth rate (multiplied by 100) in the DMA. All regressions are weighted by the relevant population at the time of the observation. All regressions include 24 quarter fixed effects as well DMA × season fixed effects as regressors. The sample sizes are 4,919, 4,920, 4,920, and 4,918 in panels 1 through 4, respectively. *** indicates significant at the 1 percent level, ** indicates significant at the 5 percent level, * indicates significant at the 10 percent level.

**Table 4**

Estimates of the Impact of *16 and Pregnant* on Teen Birth Rates, by Race/Ethnicity

| | 16 & Preg. Ratings or MTV ratings 2008-2009 | Unemp. Rate | Linear Trend × DMA |
|---|---|---|---|
| **Panel 1: Non-Hispanic Whites** | | | |
| (1) Reduced form (KL) | -3.370 ** | -1.354 *** | |
| | (1.581) | (0.375) | |
| (2) Reduced form | 1.425 | -1.011 *** | X |
| | (2.413) | (0.387) | |
| (3) Instrumental variables (KL) | -2.413 ** | -1.399 *** | |
| | (1.075) | -(0.329) | |
| (4) Instrumental variables | 1.013 | -0.960 *** | X |
| | (1.591) | (0.365) | |
| **Panel 2: Non-Hispanic Blacks** | | | |
| (5) Reduced form (KL) | -0.194 | -0.240 | |
| | (2.109) | (1.170) | |
| (6) Reduced form | -3.322 | 1.068 | X |
| | (5.989) | (1.611) | |
| (7) Instrumental variables (KL) | -0.140 | -0.243 | |
| | (1.380) | (1.049) | |
| (8) Instrumental variables | -2.360 | -1.153 | X |
| | (3.382) | (1.471) | |
| **Panel 3: Hispanics** | | | |
| (9) Reduced form (KL) | -6.510 | -2.418 *** | |
| | (4.433) | (0.944) | |
| (10) Reduced form | -1.857 | -1.972 *** | X |
| | (1.960) | (0.589) | |
| (11) Instrumental variables (KL) | -3.782 * | -2.512 *** | |
| | (2.063) | (0.784) | |
| (12) Instrumental variables | -1.066 | -2.055 *** | X |
| | (1.072) | (0.535) | |

*Source:* Authors' calculations using Bureau of Labor Statistics Local Area Unemployment Statistics, Bureau of the Census Census County Characteristics, National Vital Statistics System births, and Nielsen ratings data. The latter two data sources are confidential.

*Note:* Each row shows results from a separate regression. The dependent variable is the natural logarithm of the race/ethnicity-specific quarterly teen birth rate (multiplied by 100) in the DMA. Odd-numbered rows reproduce Kearney and Levine's (2015b) race-specific results from their Table 2. All regressions are weighted by the relevant population at the time of the observation. All regressions include 24 quarter fixed effects as well DMA × season fixed effects as regressors. The sample sizes are 4,911, 4,513, and 4,750 in panels 1, 2, and 3, respectively. *** 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

|  | Dates | | | Reduced Form | | Instrumental Variables | |
| | | "Show" | | | | | |
| Row | Begin | Start | End | Coefficient | Std. Err. | Coefficient | Std. Err. |
|------|----------|-----------|----------|-------------|-----------|------------|-----------|
| (1) | 2001:QI | 2005:QIII | 2006:QIV | -2.437 * | 1.324 | -1.639 * | 0.846 |
| (2) | 2001:QII | 2005:QIV | 2007:QI | -1.994 | 1.299 | -1.330 | 0.837 |
| (3) | 2001:QIII | 2006:QI | 2007:QII | -2.554 * | 1.441 | -1.693 * | 0.948 |
| (4) | 2001:QIV | 2006:QII | 2007:QIII | -2.929 ** | 1.473 | -1.936 ** | 0.977 |
| (5) | 2002:QI | 2006:QIII | 2007:QIV | -2.968 * | 1.521 | -1.958 ** | 0.992 |
| (6) | 2002:QII | 2006:QIV | 2008:QI | -3.144 ** | 1.569 | -2.068 ** | 1.028 |
| (7) | 2002:QIII | 2007:QI | 2008:QII | -3.341 ** | 1.508 | -2.194 ** | 1.014 |
| (8) | 2002:QIV | 2007:QII | 2008:QIII | -3.079 ** | 1.458 | -2.021 ** | 0.966 |
| (9) | 2003:QI | 2007:QIII | 2008:QIV | -2.756 * | 1.476 | -1.809 * | 0.967 |
| (10) | 2003:QII | 2007:QIV | 2009:QI | -2.543 * | 1.485 | -1.673 * | 0.969 |
| (11) | 2003:QIII | 2008:QI | 2009:QII | -2.377 | 1.558 | -1.566 | 1.002 |
| (12) | 2005:QI | 2009:QIII | 2010:QIV | -3.581 ** | 1.517 | -2.368 ** | 0.942 |

*Source:* Authors' calculations using Bureau of Labor Statistics Local Area Unemployment Statistics, Bureau of the Census Census County Characteristics, National Vital Statistics System births, and Nielsen ratings data. The latter two data sources are confidential.

*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" quarers 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. Row (12), the period analyzed by Kearney and Levine (2015b), is shaded. The two boxed rows, (1) and (7), are an example of results that do not share any "post" period in common. Sample size in rows (1) through (9), and (11) is 4,918, the sample size in row (10) is 4,917, and the sample size in row (12) is 4,919. *** indicates significant at the 1 percent level, ** indicates significant at the 5 percent level, * indicates significant at the 10 percent level.