

**Note to readers with disabilities:** *EHP* strives to ensure that all journal content is accessible to all readers. However, some figures and Supplemental Material published in *EHP* articles may not conform to [508 standards](#) due to the complexity of the information being presented. If you need assistance accessing journal content, please contact [ehp508@niehs.nih.gov](mailto:ehp508@niehs.nih.gov). Our staff will work with you to assess and meet your accessibility needs within 3 working days.

### **Supplemental Material**

#### **Residential Proximity to Oil and Gas Development and Birth Outcomes in California: A Retrospective Cohort Study of 2006–2015 Births**

Kathy V. Tran, Joan A. Casey, Lara J. Cushing, and Rachel Morello-Frosch

#### **Table of Contents**

**Figure S1.** Active well density by air basin across California (2005–2015). Map created in ArcGIS 10.6 (ESRI, Redlands, CA). The well density was calculated via the point density tool, which calculates density based on the number of neighboring wells within a 1 km x 1 km cell around each well by air basin.

**Figure S2.** Distribution of total production volume (BOE) per day by non-cases of LBW (A) and cases of LBW (B). The distribution was generated for birth outcome cases and non-cases in order to select a cut-off around the median for each category of exposure to active wells. The distribution of normalized BOE/day was similar across cases of LBW, PTB and SGA. One cut-off was selected for comparability and the value of the cut-off was selected to ensure sufficient overall sample size and number of cases in each exposure group. Figures generated in SAS 9.4 (SAS Institute Inc., Cary, NC).

**Figure S3.** Semi-variogram generated to assess residual spatial dependence. Distances between observations were grouped into 30 distance classes to generate corresponding averages of the semi-variances. The nearly straight line indicates that there is no spatial autocorrelation between the observations. Figure generated in SAS 9.4 (SAS Institute Inc., Cary, NC).

**Figure S4.** Distribution of the average total production volume (BOE) and birth outcomes by year of birth. Total production volume reflects exposure that occurred during pregnancy, most of which occurred the prior year of birth. The annual rates of LBW, PTB and SGA are plotted for rural (A) and urban (B) areas. The birth weight plots include only term births and the mean term birth weight is plotted for rural (C) and urban (D) areas.

**Table S1.** Adjusted odds ratio for binary birth outcomes associated with exposure to inactive wells by rural/urban status. Data for Figure 3A-3C.

**Table S2.** Adjusted odds ratio for term birth weight (grams) associated with exposure to inactive wells by rural/urban status. Data for Figure 3D.

**Table S3.** Unstratified adjusted odds ratios and mean difference (grams) for adverse birth outcomes associated with exposure to oil and gas production.

**Table S4.** Adjusted odds ratios for low birth weight associated with oil and gas production volume by urban/rural status. Data for Figure 4A. Effect modification p-values were derived from two-sample z-tests using strata-specific estimates and variances; significance at  $\alpha=0.05$ .

**Table S5.** Adjusted odds ratios for preterm birth associated with oil and gas production volume by urban/rural status. Data for Figure 4B. Effect modification p-values were derived from two-sample z-tests using strata-specific estimates and variances; significance at  $\alpha=0.05$ .

**Table S6.** Adjusted odds ratios for small for gestational age associated with oil and gas production volume by urban/rural status. Data for Figure 4C. Effect modification p-values were derived from two-sample z-tests using strata-specific estimates and variances; significance at  $\alpha=0.05$ .

**Table S7.** Adjusted mean difference of term birth weight (grams) associated with oil and gas production volume by urban/rural status. Data for Figure 4D. Effect modification p-values were derived from two-sample z-tests using strata-specific estimates and variances; significance at  $\alpha=0.05$ .

**Table S8.** Rural adjusted odds ratios and mean difference (grams) for adverse birth outcomes associated with exposure to oil and gas production during the entire pregnancy by maternal race/ethnicity. Effect modification p-values were derived from two-sample z-tests using strata-specific estimates and variances; significance at  $\alpha=0.05$ . Non-Hispanic Whites were used as the reference in z-tests. EM p-values were not reported for categories with no observations.

**Table S9.** Urban adjusted odds ratios and mean difference (grams) for adverse birth outcomes associated with exposure to oil and gas production during the entire pregnancy by maternal race/ethnicity.

**Table S10.** Rural adjusted odds ratios and mean difference (grams) for adverse birth outcomes associated with exposure to oil and gas production during the entire pregnancy by air basin. Effect modification p-values were derived from two-sample z-tests using strata-specific estimates and variances; significance at  $\alpha=0.05$ . Sacramento Valley was used the reference for all z-tests.

**Table S11.** Urban adjusted odds ratios and mean difference (grams) for adverse birth outcomes associated with exposure to oil and gas production during the entire pregnancy by air basin.

**Table S12.** Adjusted odds ratios and mean difference (grams) for adverse birth outcomes associated with exposure to oil and gas production during the entire pregnancy by urban/rural status for sensitivity analysis models including maternal pre-pregnancy BMI and smoking during pregnancy (2007-2015). BMI and smoking were not available for 2006 births. Effect estimates did not change by >10% compared to main models. Main models excluded these two covariates in order to maximize sample size.

**Table S13.** Adjusted odds ratios and mean difference (grams) for adverse birth outcomes associated with exposure to oil and gas production during the entire pregnancy by urban/rural status for sensitivity analysis models including an indicator for exposure to TRI facilities within 1 km. The variable was missing for 79,371 observations (3%). Effect estimates did not change by >10%, compared to main models.