Supplementary Materials for Pregnancy during the pandemic: The impact of COVID-19-related stress on risk for prenatal depression

Lucy S. King[1], Daisy E. Feddoes[1], Jacklyn S. Kirshenbaum[1], Kathryn L. Humphreys[2], & Ian H. Gotlib[1]

[1]Department of Psychology, Stanford University
[2]Department of Psychology and Human Development, Vanderbilt University

Method

Measures

COVID-19 Stress and Adversity. We examined 16 indicators of COVID-19 stress and adversity, listed in Table S1. Distributions for each of these variables are presented in Figure S1.

Community-level Risk Factors. We geocoded women's current addresses using the Census Geocoding API and the "tidygeocoder" package in R (Cambon, 2020). We next used the "tigris" package in R (Walker, 2020) to identify the census tract code corresponding to each participant's geolocation.

Statistical Analysis

Aim 2: Explore the Latent Structure of COVID-19 Stress and Adversity among Pregnant Women. Given that the 16 indicators of stress and adversity (see Table S1) included a mix of continuous normal variables, continuous variables with floor effects (i.e., censored from below), and count variables (modeled with Poisson regression), the EFA was estimated with maximum likelihood robust standard errors using a numerical integration algorithm (MLR; Muthén & Muthén, 2017). Missing data were handled using Full Information Maximum Likelihood and factors were allowed to correlate using an oblique rotation ("geomin"; Muthén & Muthén, 2017). Traditional fit statistics for EFA are not calculated when using MLR. Therefore, we determined the number of factors based on chi-squares tests comparing model fit for a model with one additional factor, the Bayesian Information Criterion (BIC) values, and the interpretability of the solution.

Results

Descriptive Statistics

We present distributions of indicators of COVID-19 stress and adversity in Figure S1, of estimated factor scores for *objective COVID-19 adversity* and *subjective COVID-19 stress* in Figure S2, of community-level risk factors in Figure S3, and individual-level risk and protective factors in Figure S4.

Pre-pandemic Cohort. Participants reported living with 0-4 children (mean[SD]=0.88[1.10]) and 1-7 adults (mean[SD]=2.38[1.12]). Eighty-three percent were married or partnered. Sixty-eight percent had paid jobs, 17% were stay-at-home caregivers, 2% were students, and 10% were unemployed. Sixty-seven percent of participants had previously received treatment for mental health difficulties or substance abuse.

COVID-19 Cohort. Participants reported living with 0-9 children (mean[SD]=0.63[0.93]) and 1-8 adults (mean[SD]=2.30[0.81]). Ninety-seven percent were married or partnered. Seventy-six percent had paid jobs, 10% were stay-at-home caregivers, 2% were students, 3% reported "other" types of employment, and 8% said they were unemployed and looking for work. Forty-one percent of participants had previously received treatment for mental health difficulties or substance abuse.

Matching of Participants in the COVID-19 and Pre-pandemic Cohorts for Aim 1 Analyses

We present density plots for the continuous variables examined in the matching process in Figure S5 and descriptive statistics for these variables in Table S3.

Participants in the pre-pandemic and COVID-19 samples were imbalanced with respect to age, gestational weeks, number of children in the household, parity (primiparous vs. multiparous), marital status (married/partnered vs. single/divorced/separated), history of mental health/substance use treatment (positive history vs. no history), and, to a lesser degree, on education (< vs. ≥ 4-year college degree) and employment status (employed for wages vs. stay-at-home home caregiver/student/unemployed). Notably, although several of the between-cohort mean differences for these variables were statistically significant in the full data and

became non-significant after matching (see Table S3), hypothesis testing should not be used to determine balance between groups given that balance is a within-sample property that does not have reference to a broader population and that hypothesis tests for balance can be misleading due to variation in statistical power (Stuart, 2010).

First, because all participants in the pre-pandemic sample had known addresses in the San Francisco Bay Area, we first removed 390 participants from the COVID-19 sample who did not provide addresses in the Bay Area. Second, we implemented Optimal Nearest Neighbor Matching using the "MatchIt" package in R (Ho, Imai, King, & Stuart, 2011) to identify participants from the COVID-19 cohort that were similar to the pre-pandemic cohort with respect to age, gestational weeks, number of children, parity, marital status, history of mental health/substance use treatment, marital status, and education (employment status was no longer considered after finding that including it in the matching procedure led to greater imbalance). Optimal Nearest Neighbor Matching selects the best participant from one group (e.g., COVID- 19) for each participant in another group (e.g., pre-pandemic) using propensity scores as distance measures, while taking into account the overall set of matches when choosing individual matches in order to minimize the global propensity score. The result is two groups of equal size with similar distributions across the variables considered.

We used the following established numerical guidelines (Rubin, 2006) to determine whether the resulting matched samples were sufficiently balanced: 1) the absolute standardized mean difference (SMD) in propensity scores is < 0.50 and 2) the ratio of the variances (variance ratio) of the propensity scores is close to 1.00. In our matched samples, we achieved a SMD = 0.15 and a variation ratio = 1.66 for the propensity scores. For all variables except gestational age, SMDs ranged from -0.03-0.19 (see Figure S6) and variance ratios ranged from 1.23-1.52. However, as is evident in Figure S5, Panel B, we did not achieve balance for gestational weeks given that participants in the pre-pandemic cohort were recruited to have a central tendency in mid-pregnancy and therefore had more restricted variance than did participants in the COVID-19 cohort. Thus, although we achieved an SMD = 0.10 for gestational weeks, the variance ratio was low at 0.32. We could not remove any participants from the pre-pandemic cohort in order to achieve balance on gestational weeks due to concerns for loss in statistical power. Indeed, an approach of Coarsened Exact Matching resulted in sample sizes < 35 per cohort. Based on findings that we had achieved adequate overall balance and the fact that gestational weeks was not associated with depressive symptoms in either cohort (see Figure S7), the matched samples derived from Optimal Nearest Neighbor Matching were used in Aim 1 analyses. We tested our hypotheses for Aim 1 when adjusting for all the variables examined in the matching process.

Aim 1: Examine Prenatal Depressive Symptoms in Matched Samples of Women Pregnant Prior to and During the COVID-19 Pandemic

We repeated the analyses for Aim 1 using all available data from both the pre-pandemic and COVID-19 cohorts instead of the matched sub-cohorts. We instead adjusted for the variables examined in the matching process, including age, gestational weeks, number of children in the household, parity, marital status, history of mental health/substance use treatment, education, employment status, and race/ethnicity. Consistent with the analyses of the matched sub-cohorts, we found that women in the full COVID-19 cohort of 725 women had significantly higher symptoms of depression than did the 88 women in the pre-pandemic cohort (B=2.99, SE=0.61, $t(757)=4.89$, $p<.001$, β[95% CI]=0.57[0.34, 0.81]).

Although the EPDS was not designed to measure anxiety symptoms, there is evidence that three of the questions on the EPDS (*I have blamed myself unnecessarily when things went wrong; I have been anxious or worried for no good reason; I have felt scared or panicky for no very good reason*) correlate to form an anxiety subscale (EPDS-3A; (Tuohy & Mcvey, 2008). To determine whether differences in EPDS scores between the matched pre-pandemic and COVID-19 cohorts were driven by differences in anxiety symptoms, we repeated the analyses for Aim 1 modeling EPDS-3A scores instead of total EPDS scores. Women in the COVID-19 cohort did not differ significantly in EPDS 3-A scores from women in the pre-pandemic cohort (Welch's $t[81]=0.99$, $p=.326$, Cohen's d[95% CI]=0.15[-0.15, 0.45]), suggesting that differences in symptoms between women who were pregnant during the pandemic and women who were pregnant prior to the pandemic were driven by overall depressive symptoms rather than anxiety symptoms specifically.

Although gestational weeks was not associated with women's depressive symptoms in either cohort (see Figure S7), we explored whether the difference in symptoms between the pre-pandemic and COVID-19 cohorts depended on stage of pregnancy. Only 1 participant in the pre-pandemic cohort and 17 participants in the COVID-19 cohort were in the first trimester of pregnancy. Therefore, we conducted analyses testing the difference in depressive symptoms between women in each cohort stratified based on whether women were in

earlier pregnancy (≤20 gestational weeks; N=53) or in later pregnancy (>20 gestational weeks; N=111). Given different numbers of women in early vs. later pregnancy, we focused our interpretation on effect sizes rather than *p*-values. The effect size for the difference in depressive symptoms between the pre-pandemic and COVID-19 cohorts was smaller among women earlier in pregnancy (Welch's $t(49.30)=1.52$, Cohen's *d*[95% CI]=0.42[-0.14, 0.96]) than among women in later pregnancy (Welch's $t(104.74)=3.99$, Cohen's *d*[95% CI]=0.76[0.38, 1.15]). However, as evidenced by a non-significant interaction between cohort and phase of pregnancy (earlier vs. later; B=-2.07, SE=1.82, $t(160)=-1.13$, $p=.259$, β[95% CI]=-0.36[-1.00, 0.27]), these effect sizes were not significantly different from each other. Thus, although these findings suggest that, compared to women who were pregnant prior to the pandemic, women who were pregnant during the pandemic who were in later pregnancy had more severely elevated symptoms of depression than did women who were in earlier pregnancy, we cannot draw strong conclusions

Aim 2: Explore the Latent Structure of COVID-19 Stress and Adversity among Pregnant Women

Based on chi-square tests of model fit and the BIC values, more complex models provided better fits to the data. Specifically, a 2-factor solution (BIC=43,752.83) fit better than a 1-factor solution (BIC=44,252.81.16; $\chi^2(15)=581.11$, $p<.001$), and a 3-factor solution (BIC=43,425.15) fit better than a 2-factor solution ($\chi^2(14)=461.49$, $p<.001$). When items with high oblique rotated cross- loadings (>|.30|) were removed (distress due to current financial/employment impacts, distress due to expected financial/employment impacts, and restrictions on activity), leaving 13 items in the model, we observed a similar pattern. Specifically, the 2-factor solution provided a significantly better fit to the data than the 1-factor solution ($\chi^2(12)=591.18$, $p<.001$), and the BIC was reduced in the 2-factor compared to the 1-factor model and in the 3-factor model compared to the 2-factor model ($BIC_{1\text{-factor}}$=36,788.16; $BIC_{2\text{-factor}}$=36,283.36; $BIC_{3\text{-factor}}$ =36,100.37). However, the 3-factor model remained uninterpretable due high cross-loadings. In contrast, the 2-factor solution contained no high cross-loadings and was interpretable.

Aim 4: Examine the Associations of Dimensions of COVID-19 Stress and Adversity with Prenatal Depressive Symptoms during the COVID-19 Pandemic

We depict the zero-order associations of COVID-19 objectivity adversity and COVID-19 subjective stress with prenatal depressive symptoms in Figure S8.

Gestational weeks was not associated with scores for subjective COVID-19 stress (*r*[723]=.12) or objective COVID-19 adversity (*r*[723]=.02). Among women in the full COVID-19 cohort of 725 pregnant women, we explored whether the association between subjective COVID-19 stress and depressive symptoms and the association between objective COVID-19 adversity and depressive symptoms depended on stage of pregnancy. Specifically, we conducted analyses testing the association between each dimension of COVID-19 stress and adversity and depressive symptoms stratified based on whether women were in earlier pregnancy (≤20 gestational weeks; N=175) or in later pregnancy (>20 gestational weeks; N=550). Findings were highly similar for women in earlier and later pregnancy, although the effect size for the association between subjective COVID-19 adversity and prenatal depressive symptoms was slightly larger in later than in earlier pregnancy. In both earlier and later pregnancy and when covarying for objective COVID-19 adversity, subjective COVID-19 stress was strongly positively associated with depressive symptoms (**earlier pregnancy:** β=0.38, 95% CI[0.22, 0.56], R^2=.12; **later pregnancy:** β=0.45, 95% CI[0.36, 0.54], R^2=.19). In contrast, when covarying for subjective COVID-19 stress, objective COVID-19 adversity was not associated with depressive symptoms (**earlier pregnancy:** β=0.03, 95% CI[-0.24, 0.16], , R^2=.02; **later pregnancy:** β=0.06, 95% CI[-0.02, 0.14], R^2=.04)

We sought to identify single indicators of COVID-19 stress and adversity that were most strongly associated with prenatal depressive symptoms. Given that factor scores are difficult to interpret, identifying single indicators may aid in characterizing our findings and translating them to clinical settings. The LOOCV for the EN model of depressive symptoms identified an optimal λ=.02, resulting in 9 variables with non-zero coefficients, 7 of which were variables that had loaded onto the *subjective COVID-19 stress* factor and two of which were variables that had loaded onto the *objective COVID-19 adversity* factor. We present the associations between each of these variables and prenatal depressive symptoms in Figure S9. Listed in order of zero-order effect size, these variables were as follows: distress due to social disruptions (EN estimate=0.20), distress due to reduced access to resources (EN estimate=0.18), severity of overall impact of COVID-19 on daily life (EN estimate=-0.06), valance of overall impact of COVID-19 (EN estimate=-0.11), concern about the availability of social support during birth (EN estimate=0.16), distress that family would

contract virus (EN estimate=0.04), distress that participant would contract virus (EN estimate=0.02), current employment/financial impacts (EN estimate=0.09), number of changes to prenatal care due to COVID-19 (EN estimate=0.02).

Table S1. Descriptive statistics for indicators of COVID-19 stress and adversity. For valence of overall impact, lower scores indicate that the impact of the pandemic was viewed as having a more negative impact on life.

Variable	Mean (SD)	Range
Changes to prenatal care	2.29 (1.39)	0-6
Concern for birth medical care	4.06 (2.40)	0-7
Concern for birth social support	4.96 (2.14)	0-7
Concern for caregiving capacity	3.52 (2.85)	0-7
Concern for infant's health	4.04 (2.66)	0-7
Self-quarantine/government restrictions	1.44 (0.62)	0-4
Current employment/financial impacts	4.71 (2.88)	0-19
Expected employment/financial impacts	3.87 (3.69)	0-22
Distress due to current financial impacts	2.69 (2.06)	0-6
Distress due to expected financial impacts	2.98 (2.08)	0-6
Distress due to potential family illness	3.31 (2.02)	0-6
Distress due to own potential illness	1.86 (1.99)	0-6
Distress due to reduced access	1.47 (0.69)	0-3
Distress due to social disruptions	2.76 (1.94)	0-6
Overall impact to daily life	4.14 (1.21)	0-6
Valence of overall impact	-1.74 (1.04)	-3-3

Table S2. Descriptive statistics, elastic net estimates, and zero-order effect sizes for individual- and community-level risk and protective factors. Descriptive statistics are for the 343 participants who provided valid addresses in California. Elastic net and effect size estimates are for the 319 participants included in the elastic net regression. FPL=Federal Poverty Line. Objective adversity and subjective stress are estimated factor scores from solution presented in Figure 2. Zero-order effect sizes are Cohen's *d* values for discrete variables, Pearson's *r* values continuous variables., and Spearman's ρ values for ordinal or count data.

Variable	N (%) or Mean (SD)	Range	Elastic Net Estimate / zero-order effect size [95% CI]	
			Objective Adversity $R^{2=}$.11	***Subjective Stress*** $R^{2=}$.18
Individual-level Factors				
Primiparous	173 (50)		0.10 *.15 [-.07, .37]*	0.13 *.28 [.06, .50]*
Previous miscarriage or abortion	106 (31)			
History of serious medical conditions	80 (23)			
History of mood or anxiety disorder	93 (27)		0.04 *.20 [-.05, .45]*	0.03 *.26 [.01, .51]*
High risk for COVID-19 due to existing medical conditions	83 (24)			0.10 *.44 [.18, .70]*
Person of color	136 (39)			0.11 *.31 [.09, .54]*
Immigrant to U.S.	69 (20)		0.07 *.20 [-.07, .47]*	
Low income according to FPL				
Multi-bedroom housing	279 (81)			
Age			0.04 *.06 [-.05, .17]*	0.03 *.03 [-.08, .14]*
Environmental resources	4.21 (0.85)	0-5	-0.11 *-.11 [-.21, <.01]*	-0.10 *-.16 [-.26, -.05]*
Prenatal medical conditions	0.22 (0.48)	0-2	0.15 *.20 [.09, .30]*	0.13 *.17 [.06, .28]*
Support from prenatal care team	1.60 (0.56)	0-2	-0.12 *-.14 [-.24, -.03]*	-0.10 *-.13 [-.23, -.02]*
Current social support	3.91 (1.58)	0-6		
Change in social support	-0.64 (1.46)	-5-4		-0.11 *-.16 [-.27, -.05]*
Community-level Factors				
Rate of ER visits for asthma	39.13 (28.65)	0.02-99.61	0.05 *.07 [-.04, .18]*	
Rate of ER visits for heart attacks	29.52 (23.85)	0.32-95.72		
Rate of low weight births	43.73 (27.64)	0.00-99.82	0.01 *.05 [-.06, .16]*	
% low educational attainment	32.40 (24.22)	0.18-91.72		0.13 *.15 [.05, .25]*
% housing burdened	39.87 (25.58)	0.27-98.95		
% limited English-peaking	47.94 (26.27)	0.00-99.34	0.07 *.06 [-.05, .17]*	0.01 *.11 [.01, .29]*
% in poverty	30.79 (21.67)	0.24-90.80	-0.09 *-.04 [-.15, .07]*	-0.03 *.03 [-.08, .14]*
% unemployed	29.53 (22.60)	0.04-98.45		

Table S3. Descriptive statistics for variables examined in matching process. Low income is ≤ 200% below the federal poverty line. Prior to matching, 88 participants in pre-pandemic cohort and 725 in COVID-19 cohort. After matching, 82 participants in each cohort. Welch's *t*-statistics do not assume equality of variances.

Variable	Prior to Matching			After Matching		
	Pre-pandemic Cohort	COVID-19 Cohort		Pre-pandemic Cohort	COVID-19 Cohort	
	N (%)		χ^2	N (%)		χ^2
Primiparous	24 (27)	358 (49)	14.52, p<.001	20 (24)	23 (28)	0.13, p=.723
Married/partnered	73 (82)	706 (97)	37.23, p<.001	70 (85)	74 (90)	0.51, p=.474
Person of color	35 (40)	267 (37)	0.11, p=.736	33 (40)	38 (46)	0.40, p=.528
Low income	7 (8)	50 (7)	0.08, p=.773	7 (9)	8 (10)	<0.00, p=1.00
≥4-year college degree	69 (78)	611 (85)	1.20, p=.274	66 (80)	58 (71)	1.62, p=.203
Employed for wages	60 (68)	553 (76)	2.35, p=.125	55 (67)	55 (67)	<0.00, p=1.00
Past mental health/ substance treatment	59 (67)	296 (41)	20.05, p<.001	57 (69)	56 (68)	<0.00, p=1.00
	Mean (SD) Range		t	Mean (SD) Range		t
Age (years)	32.55 (5.04) 20-44	33.69 (4.38) 19-50	t=2.01, p=.047	32.65 (5.12) 20.52-44.42	32.62 (4.79) 19.73-41.90	-0.46, p=0.964
Gestational weeks	24.44 (5.48) 12-37	26.79 (8.81) 4-41	t=3.51, p<.001	24.35 (5.64) 12-37	23.92 (10.06) 5-40	-0.34, p=.736
Number of children	0.88 (1.10) 0-4	0.63 (0.93) 0-9	t=1.97, p=.052	0.89 (1.10) 0-4	0.88 (0.88) 0-5	-0.08, p=.938

Figure S1. Distributions of indicators of COVID-19 stress and adversity examined in exploratory factor analysis. Variables with floor effects were censored from below (Tobin, 1958) in the EFA model.

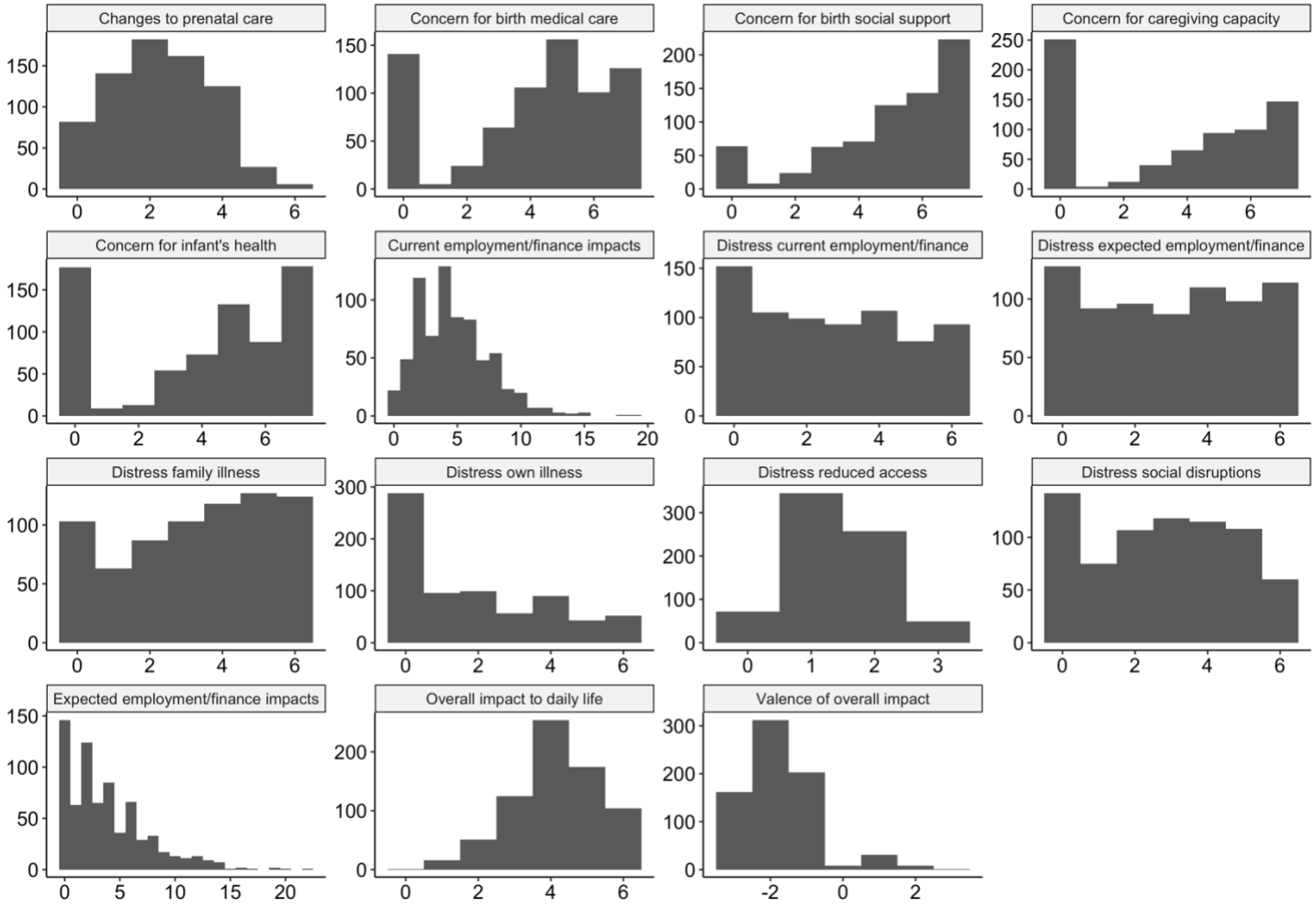


Figure S2. Distributions of estimated factor scores used in analyses for Aims 2-3.

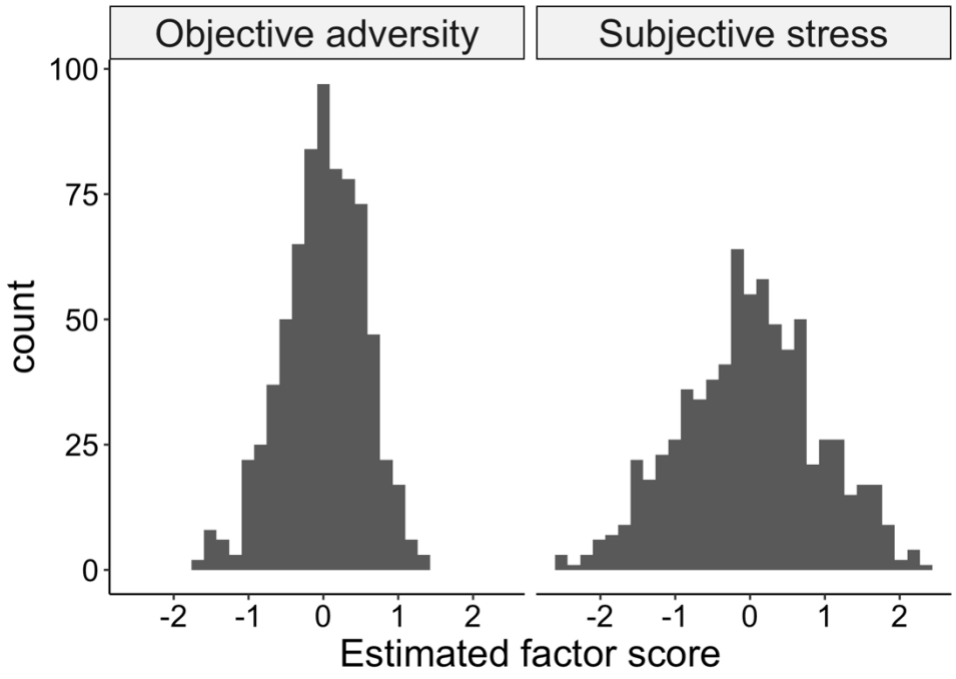


Figure S3. Distributions of community-level risk factors. Based on data from the CalEnviroScreen, the indicators were quantified by the census tract percentile relative to the distribution across all 8,057 tracts in California.

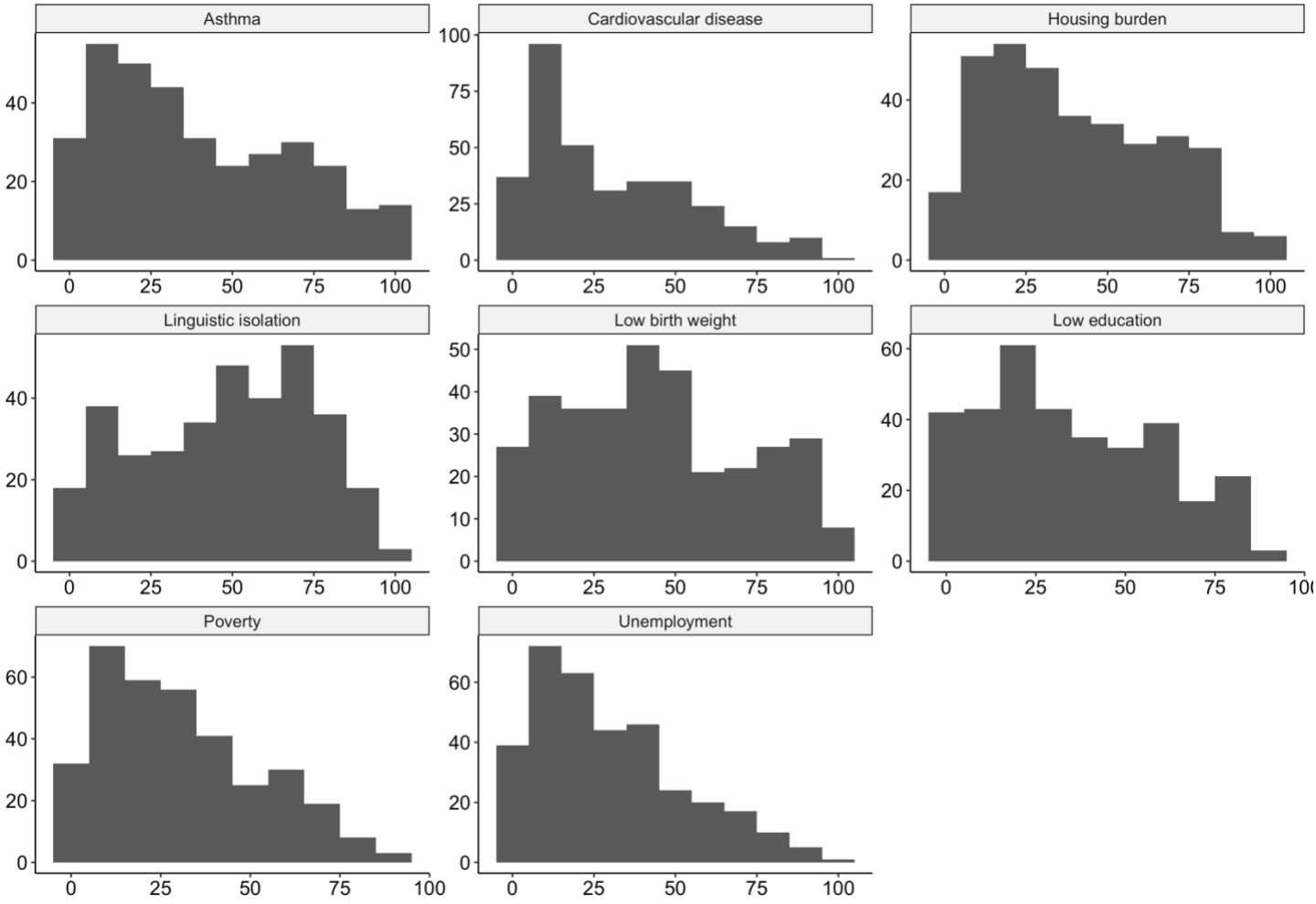


Figure S4. Distributions of continuous individual-level risk and protective factors.

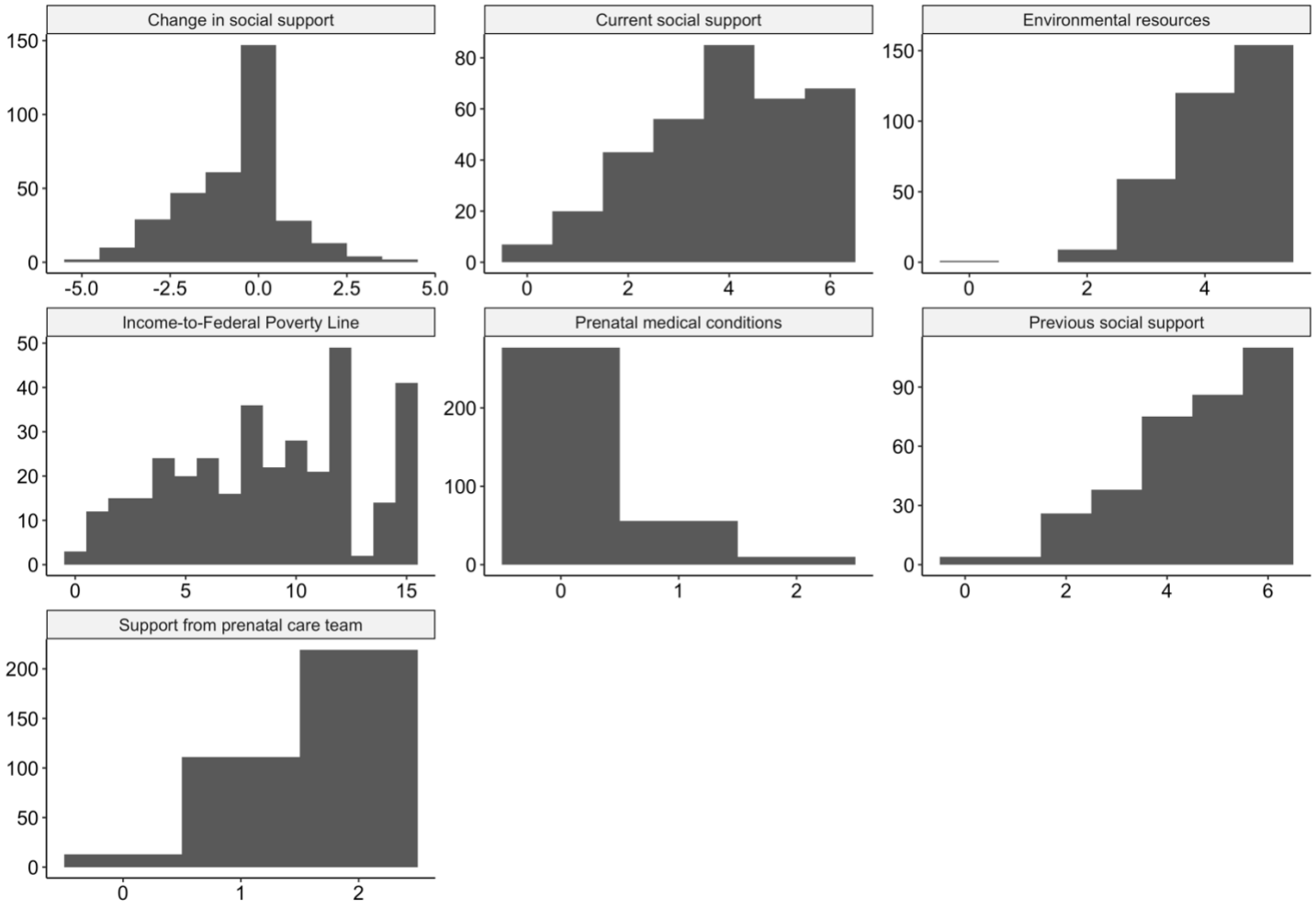
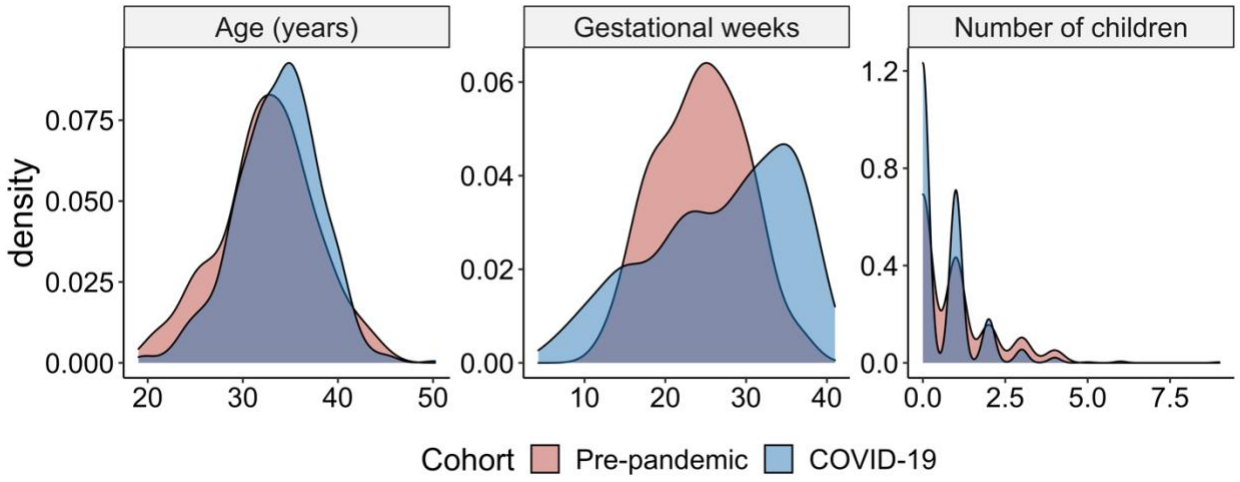


Figure S5. Distributions of continuous variables examined in process of matching a subsample of participants in the COVID-19 cohort to the pre-pandemic cohort.

A. Distributions prior to matching



B. Distributions after matching

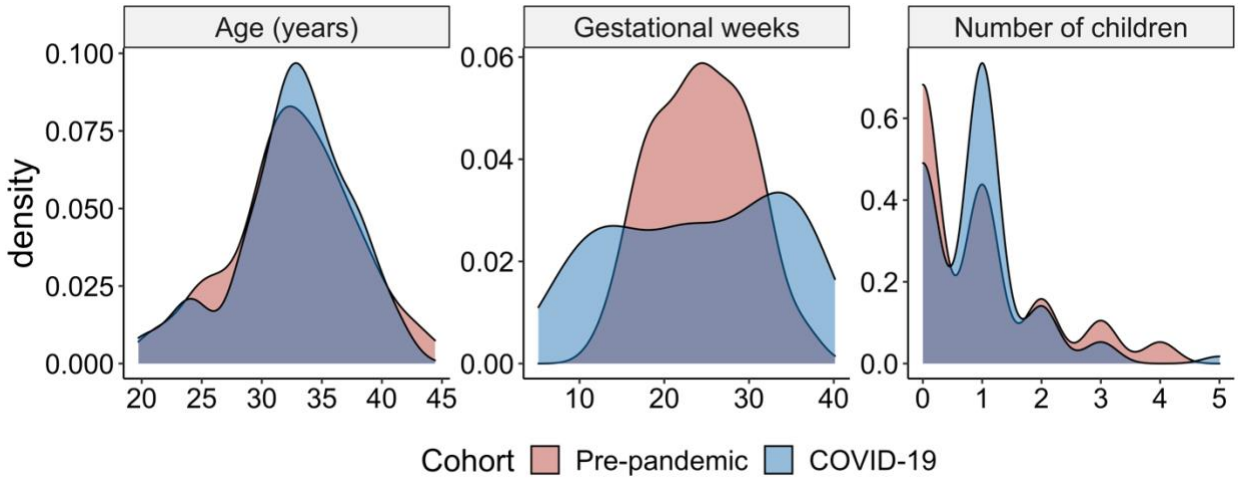


Figure S6. Standardized mean differences between the pre-pandemic and COVID-19 cohorts in all data (prior to matching) and matched data (after matching). Past treatment is whether the participant reported a history of treatment for mental health difficulties or substance use. Absolute standardized mean differences scores < 0.50 are considered acceptable (Rubin, 2006).

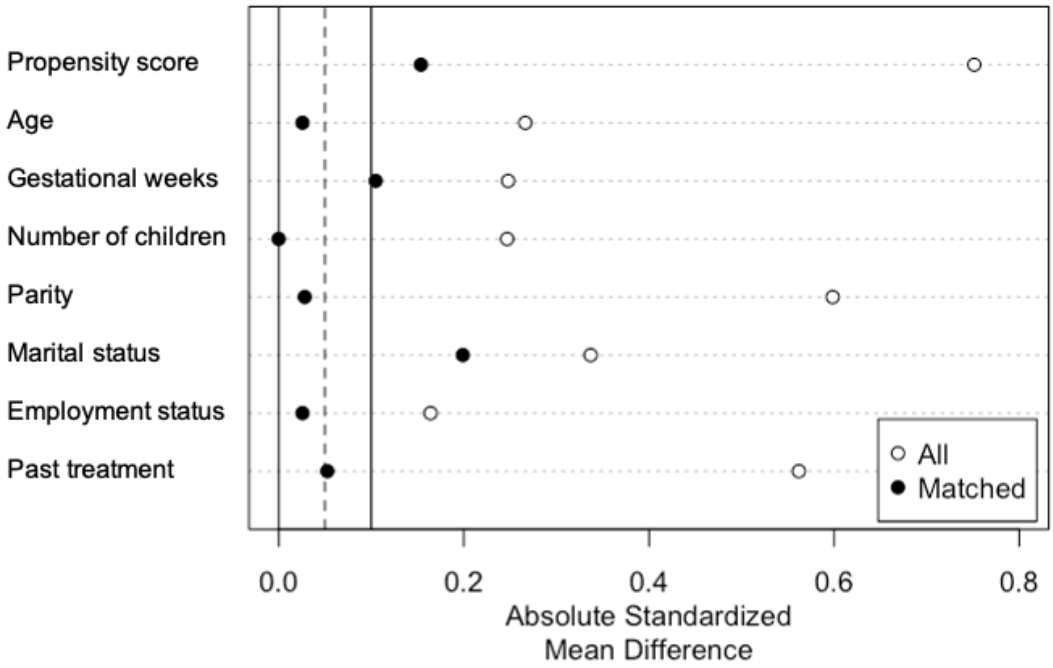


Figure S7. Associations between gestational weeks in pregnancy at time of assessment of prenatal depressive symptoms in the pre-pandemic and COVID-19 cohorts.

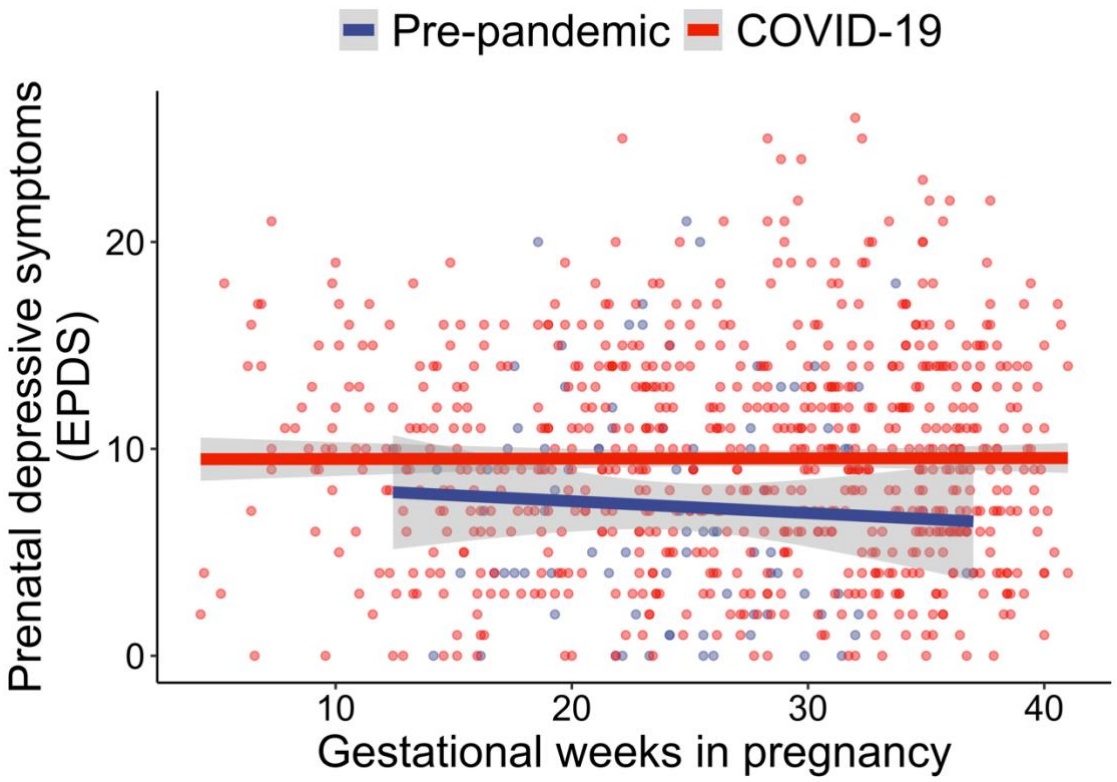
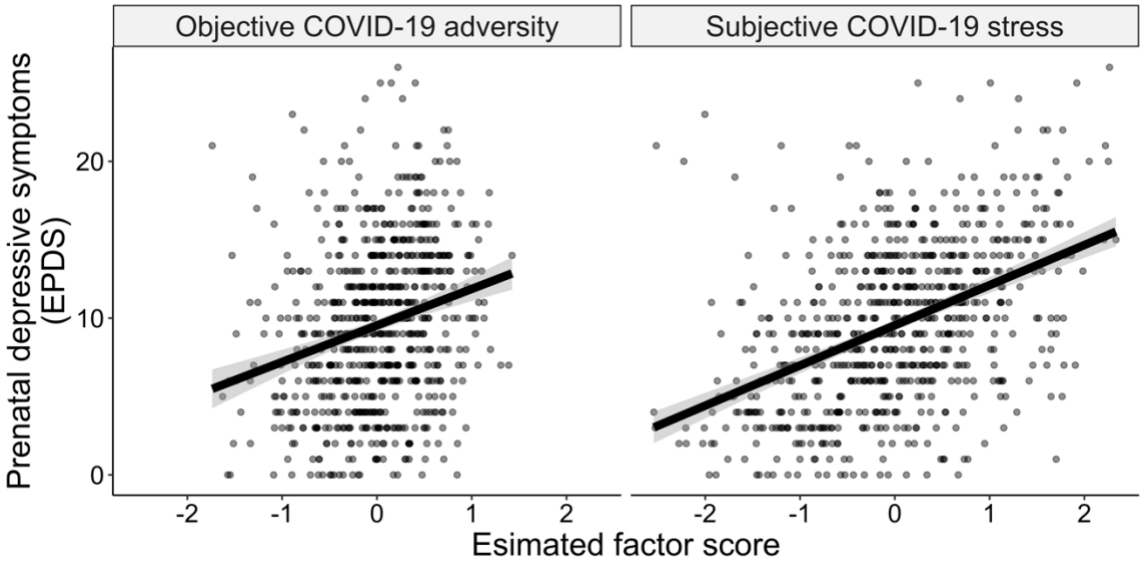


Figure S8. Zero-order associations between estimated factor scores for COVID-19 stress and adversity and prenatal depressive symptoms. When estimated scores for both factors were entered together, only *Subjective COVID-19* stress remained significantly associated with prenatal depressive symptoms.



References

Cambon, J. (2020). *tidygeocoder: Geocoding made easy*. Retrieved from https://jessecambon.github.io/tidygeocoder/

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: Nonparametric preprocessing for parametric causal inference. *Journal of Stastical Software*, *42*(8), 1–28. https://doi.org/10.18637/jss.v042.i08

Muthén, L. K., & Muthén, B. O. (2017). *Mplus User's Guide.* (8th ed.). Los Angeles, CA: Muthén & Muthén.

Rubin, D. B. (2006). Using propensity scores to help design observational studies: Application to the tobacco litigation. *Matched Sampling for Causal Effects*, 365–382. https://doi.org/10.1017/CBO9780511810725.030

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science*, *25*(1), 1–21. https://doi.org/10.1214/09-STS313

Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica*, *26*, 24–36.

Tuohy, A., & Mcvey, C. (2008). Subscales measuring symptoms of non-specific depression , anhedonia , and anxiety in the Edinburgh Postnatal Depression Scale. *British Journal of Clinical Psychology*, *47*, 153–169. https://doi.org/10.1348/014466507X238608

Walker, K. (2020). tigris: Load census TIGER/line shapefiles. Retrieved from https://cran.r-project.org/package=tigris