

A Service of

ZBW

Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

Dyroff, Philipp; Miller, Robert

## Working Paper A Comment on "Calvert et al. (2023): Changes in Preterm Birth and Stillbirth during COVID-19 Lockdowns in 26 Countries"

I4R Discussion Paper Series, No. 222

**Provided in Cooperation with:** The Institute for Replication (I4R)

*Suggested Citation:* Dyroff, Philipp; Miller, Robert (2025) : A Comment on "Calvert et al. (2023): Changes in Preterm Birth and Stillbirth during COVID-19 Lockdowns in 26 Countries", I4R Discussion Paper Series, No. 222, Institute for Replication (I4R), s.l.

This Version is available at: https://hdl.handle.net/10419/315049

#### Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

#### Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



# WWW.ECONSTOR.EU

# **INSTITUTE** for **REPLICATION**

No. 222 I4R DISCUSSION PAPER SERIES

# A Comment on "Calvert et al. (2023): Changes in Preterm Birth and Stillbirth during COVID-19 Lockdowns in 26 Countries"

Philipp Dyroff Robert Miller

**April 2025** 



# **I4R DISCUSSION PAPER SERIES**

I4R DP No. 222

# A Comment on "Calvert et al. (2023): Changes in Preterm Birth and Stillbirth during COVID-19 Lockdowns in 26 Countries"

Philipp Dyroff<sup>1</sup>, Robert Miller<sup>1</sup>

<sup>1</sup>Berlin University of Psychology, Berlin/Germany

### **APRIL 2025**

Any opinions in this paper are those of the author(s) and not those of the Institute for Replication (I4R). Research published in this series may include views on policy, but I4R takes no institutional policy positions.

I4R Discussion Papers are research papers of the Institute for Replication which are widely circulated to promote replications and metascientific work in the social sciences. Provided in cooperation with EconStor, a service of the <u>ZBW – Leibniz Information Centre for Economics</u>, and <u>RWI – Leibniz Institute for Economic Research</u>, I4R Discussion Papers are among others listed in RePEc (see IDEAS, EconPapers). Complete list of all I4R DPs - downloadable for free at the I4R website.

I4R 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.

#### Editors

Abel Brodeur University of Ottawa Anna Dreber Stockholm School of Economics Jörg Ankel-Peters RWI – Leibniz Institute for Economic Research

E-Mail: joerg.peters@rwi-essen.de RWI – Leibniz Institute for Economic Research Hohenzollernstraße 1-3 45128 Essen/Germany www.i4replication.org

## A comment on "Calvert et al. (2023): Changes in preterm birth and stillbirth during COVID-19 lockdowns in 26 countries"

Philipp Dyroff (Berlin University of Psychology) Robert Miller (Berlin University of Psychology)

Correspondence concerning this report should be addressed to PD (<u>p.dyroff@stud.phb.de</u>) and RM (<u>r.miller@phb.de</u>), Chair of Psychological Methods, Psychologische Hochschule Berlin, Am Köllnischen Park 2, 10179 Berlin, Germany.

#### Abstract

Calvert et al. (2023) meta-analyzed effect estimates from interrupted time series (ITS) analyses of changes in preterm birth- and stillbirth rates following the first four months lockdown in various countries. Evidence for small relative reductions was reported regarding preterm birth rates in high income countries, which was not countered by an increase in stillbirth rates. This comment attempts to (recreate) reproduce these findings in population-based samples. To check robustness, effect size estimates were additionally obtained after accounting for for serial autocorrelation. Due to restricted data access, the data for reproduction were not identical to those analyzed by the original study, but were extracted from the provided time-series plots of birth rates using a Web Plot Digitizer (Rohatgi, 2024). Our results show very similar effect size estimates with ITS analysis conducted using Poisson regression and their subsequent random-effects meta-analysis. Despite of the methodological constraints arising from a lack of data openness, our reproductive analyses provide reasonable indications for the robustness of "Changes in preterm birth rate and stillbirth during COVID-19 lockdowns in 26 countries" by Calvert et al. (2023).

#### 1 Introduction

In "Changes in preterm birth and stillbirth during COVID-19 lockdowns in 26 countries" from 2023, Calvert et al. examined the effects of lockdown on preterm-, very preterm-, spontaneous preterm-, and stillbirth rates (Calvert et al., 2023). The study used data from a total of 52,067,596 births over the period of January 2015 to July 2020, collected from 26 countries. If individual country data sources included more than 90% of all births, these data were considered population-based (this was the case for 18 countries) and non-population-based for a percentage below 90%. There were 51,340,025 from population-based and 727,571 from non-population-based data sources. Based on this data, i.e. birth rates before lockdown, weighted interrupted time series (ITS) models were fitted on the monthly *log(odds)* of birth rates to estimate a predicted value at a specific point in time. The models were created as follows:

"Given that countries could have different trends in perinatal outcomes, we fitted five different potential models for each outcome for each country evaluating the trend as a linear, square, quadratic, logarithmic and second-order polynomial effect. The model with the lowest Akaike Information Criterion was chosen as the best fit model. We assessed the goodness of fit of the best model by examining the standardized residuals. Second, to compare the forecast of the best fit model to the post-lockdown observed values, we refitted the model to the pre-lockdown observations using the same trend effect selected through the Akaike Information Criterion." (cf. ibid. p. 539).

The ITS predicted birth rates based on the months before lockdown as they would be, if lockdown had not occurred. The odds ratios were then calculated by comparing the estimated values with the actual values. For preterm, very preterm, spontaneous preterm and stillbirth, separate data sets were created and within these, prediction estimates were calculated for each country. Pooled *ORs* were calculated for each dataset using a random-effects meta-analysis, with subgroups created within the datasets based on country income (high income, upper-middle income). For any further in-detail information on the methods used in the original study, please check Calvert et al (2023) p. 539. For defining perinatal outcomes and lockdown, the authors used global standard definitions which can be found at Calvert et al. (2023) p. 538.

As different countries provided different sorts of data (some only live births, some all births) questions of comparability arose, which were consecutively addressed by sensitivity analyses. First, an ITS analysis was conducted exclusively for countries whose preterm birth rates were based on all births (as opposed to live births only). Secondly, an ITS analysis was conducted including all preterm births from 28 weeks' gestation onwards. Thirdly, to eliminate the influence of the countries with the highest birth rates, an ITS analysis was carried out excluding the USA and Brazil.

This present report was prepared for the Institute for Replication (Brodeur et al., 2024). It investigated whether the analytical results of Calvert et al. (2023) are recreationally reproducible using an extraction tool to gather the data depicted in the original studies graphs, and Poisson regression for the ITS analysis. Robustness was further investigated by accounting for serial autocorrelation of the dependent variable.

#### **1.1** Results of the original study

The meta-analysis of the population-based data shows a slight reduction in the effects of lockdown on the preterm birth rate.

population bases date indicated small reductions in the first (*OR* 0.96, 95% *Cl* 0.95–0.98, *P* < 0.001), second (*OR* 0.96, 95% *Cl* 0.92–0.99, *P* = 0.03), and third month (*OR* 0.97, 95% *Cl* 0.94–1.00, *P* = 0.09) of lockdown, but none in the fourth month (*OR* 0.99, 95% *Cl* 0.96–1.01, *P* = 0.34) Between-country heterogeneity (*I*) was 0%, 64%, 53% and 34% for the first to fourth month of lockdown, respectively. Stratifying by country income level indicated similar reductions in the odds of preterm birth for both high and upper-middle-income country settings, with higher between-country heterogeneity among upper-middle-income countries. (Calvert et al., 2023, p. 530)

The analysis of very preterm and spontaneous preterm birth rates, on the other hand, shows no or weak effects.

For very preterm births there was no evidence of an impact of lockdown over the four months of lockdown with ORs for all population-based datasets varying between 1.00 and 1.02 and *Cls* spanning the null value. For spontaneous preterm births, in the subset of countries with data available, there were small relative decreases (3–4%) in the first three months following lockdown in HICs, but not in Brazil, the only upper-middle-income country providing these data. There was also evidence for a decrease in the fourth month of lockdown using only the non-population-based data (*OR* 0.88, 95% *Cl* 0.78–0.99, *P* = 0.04, *l*<sup>2</sup>= 0%) (cf. ibid. p. 531).

Following results for the correlations between lockdown and stillbirth rate were revealed:

we found no clear evidence of an impact of lockdown on stillbirth in the first month of lockdown overall (*OR* 1.04, 95% *Cl* 0.99–1.09, *P* = 0.10,  $l^2$ = 0%), but an increase was observed when restricting to HICs (*OR* 1.14, 95% *Cl* 1.02–1.29, *P* = 0.02,  $l^2$  = 0%), driven by Canada (*OR* 1.26, 95% *Cl* 1.04–1.51, *P* = 0.02). There was an increase in the odds of stillbirth across all population based datasets in the second (*OR* 1.07, 95% *Cl* 1.02–1.12, *P* = 0.001,  $l^2$ = 0%), third (*OR* 1.08, 95% *Cl* 1.02–1.13, *P* = 0.004,  $l^2$ = 0%) and possibly fourth month (*OR* 1.07, 95% *Cl* 1.00–1.15, *P* = 0.07,  $l^2$ = 11%) of lockdown.

These ORs were driven largely by Brazil and when we restricted the meta-analysis to HICs only, we found no evidence for an association between lockdown and stillbirth in the second month (*OR* 1.00, 95% *Cl* 0.88–1.12, *P* = 0.98,  $l^2$ = 0%), third month (*OR* 0.99, 95% *Cl* 0.88–1.12, *P* = 0.89,  $l^2$ = 0%) and fourth month (*OR* 1.01, 95% *Cl* 0.87–1.18, *P* = 0.86,  $l^2$ =13%) of lockdown (cf. ibid. p. 531).

Overall, a reduction in the preterm birth rate of 3-4% was observed in the months of the lockdown; at the same time, the results did not indicate an opposite increase in the stillbirth rate (in high-income countries), which could have served as an explanation for the falling preterm birth rates. The sensitivity analyses showed "negligible changes in the country-specific estimates of the impact of lockdown on preterm birth rates" (cf. ibid. p. 533) when excluding countries that only provided live birth data and when restricting preterm births to >28 weeks' gestation. Excluding the USA and Brazil, which together account for more than 70% of the births included in the original study, also resulted in negligible changes in the effect estimates (cf. ibid. p. 536).

#### 1.2 Aims of this study

This study tests the original study for recreate reproducibility and robustness reproducibility (Dreber & Johanneson, 2023). Constitutively, it evaluates, as defined in the above-mentioned study, whether the results of the original study can be reproduced without access to the original data and the original code, using only the information given in the original study, but also tests if the results are robust to alternative analytical decisions (cf. ibid.). The hypothesis is that no significant differences will be found between the results of the original study and the results of this report, both for Poisson regression and when controlling for autocorrelation.

I4R DP No. 222

#### 2 Methods

This report attempts to reproduce the findings of the original study despite the lack of access to the individual patient data (IPD) of the original study. There were several methodological hindrances in the reproduction as will become clear below. The following section presents these hindrances as well as the solutions used.

#### 2.1 Data

One of the main differences between this report and the original study lies in analyzed data. The Calvert et al. (2023) were unable to provide direct IPD access, whereas the data curating SAIL databank refused to provide IPD access without financial compensation of utilizing their infrastructure for the purpose of the present replication endeavor. Consequently, this report had to rely on the aggregated data that were extractable from the original studies graphs. The procedure was as follows.

Fig. 1 and Extended Data Fig. 1-3 in the original study show the crude rates of preterm births, very preterm births, spontaneous preterm births and stillbirths (calculated as relative number of births per month) for each country during the time period from January 2015 to July 2020, consisting of on average 67 data points per country and type of birth. These monthly data are marked and extracted using the Web Plot Digitizer data extraction tool (Rohatgi, 2024). The reliability and validity of this method and the instrument have been proven by numerous studies (Drevon et al., 2017; Van der Mierden et al., 2021). To additionally ensure the highest possible reliability, the data of each graph is digitized three times and the intraclass coefficient is calculated from the given data. Furthermore, Wilcoxon sign rank test (Woolson, 2005) is used to evaluate if differences in the first month of lockdown in preterm birth do significantly deviate in the replication from original data. The mean value from all extracted data finally forms the data set that is used for further analysis.

#### 2.2 Interrupted time series (ITS) analysis

Following the analysis plan of the original study, ITS analyses are also carried out for this report. As reported by Calvert et al. (2023), the tutorial by Bernal, Cummins and Gasparrini (Bernal et al., 2017) was used to specify and implement these analyses. Due to

7

I4R DP No. 222

the unavailability of IPD, however, ITS models are not fitted to monthly *log(odds)* of the births, but on the monthly *log(rate)* using Poisson regression.

#### 2.2.1 Recreate Reproducibility

To enable the utilization of Poisson regression, the extracted data were transformed into integer counts, that is, the crude rates extracted from the graphs were multiplied by a value of 100 (e.g. Australia: preterm birth rate approx. 6.86\*100 = 686). Likewise, the reference population that was originally used to calculate those crude rates, was rescaled to 10,000 births per month and included as regression offset. The obvious disadvantages of this modeling approach are that any monthly fluctuation in total births is disregarded and that the individual contribution of each country to the final meta-analyses becomes dependent on the length of their respective time series. The latter challenge will be addressed in more detail later, when it comes to the calculation of pooled effect estimates.

For subsequent model selection, the analytical procedure proposed by Bernard et. al. (2017) is closely adhered to: ITS modeling regresses the accordingly offset birth counts on the time (scaled in months) relative to lock-down and allows for variability in intercept and slope across countries and birth types alike the original study. Linear, quadratic, cubic, quartic and logarithmic models of time were fitted, and the best fitting model was determined using the Akaike Information Criterion (AIC; Akaike, 1998; Calvert et al., 2023). Fourier transforms were applied to model seasonal trends and capture the periodic patterns in the data (Akula, 2020). Different harmonics (pairs of sine and cosine functions) (1; 2; 3) were evaluated and the best-fitting model was determined with AIC again and used further. Based on the selected models and the respective pre-lockdown data, predictions of birth rates were generated for the first three-, respectively four months of lockdown. Odds ratio (*OR*) was calculated by dividing the observed values by the predicted values.

#### 2.2.2 Robustness Reproducibility

To check the robustness of the results of the original study, autocorrelation was further investigated, as the description of the methodology in the original study does not indicate if this was considered. Autocorrelation occurs when the residual of a data point t of a

14R DP No. 222

time series is correlated with the residual of another data point of the previous period t - 1; t - 2; etc. (Stocker, 2023, p. 2). As a result, the standard errors of these values are also correlated, which in turn can lead to distortion of the standard errors and invalid test statistics (cf. ibid., p. 14). In principle, autocorrelation in time series data is rarely a problem, as it is often caused by seasonal fluctuations and can be managed after controlling for those (Bernal et al., 2017). Nevertheless, autocorrelation plots showed significant autocorrelations on one or more lags in about half of the data sets. To determine whether these depicted autocorrelations have a significant impact on the effect sizes, the used model was extended by a term that temporally lags the dependent variable (Idv). This is a common and robust method for dynamic models to reduce autocorrelation (Keele & Kelly, 2005), which is particularly useful since this is a less complex extension and easily added to the model.

In terms of robustness testing, the aim of this report was therefore not to find the "perfect" model, but to check whether the original model is robust enough to cope with autocorrelation and therefore not to allow the effect estimates to deviate significantly from those of this model. In order to minimize the complexity of the model and not to go beyond the scope of this report, only the value of the numerically largest lag (if significant) of each country was integrated into the corresponding model as a ldv. For similarly large lags, the better fitting model was selected using AIC. However, adding a ldv into the model leads to the exclusion of the same amount of data as the numerical value of the ldv. To minimize loss of information, only a lag that did not lead to the exclusion of more than 10% of data was chosen, otherwise the highest within the 10% range was taken as replacement. As multicollinearity can increase after the inclusion of an ldv, this was also checked.

#### 2.3 Meta-analysis

In the final step of replication, pooled odds ratios (*OR*) and their confidence intervals (*CI*) were obtained. As in the original study, a random-effects meta-analysis was carried out (Calvert et al., 2023; DerSimonian & Laird, 1986), both for the high-income and upper-middle income subgroups and for all countries in a data set together. The weights assigned to the individual countries were taken from the original study, the reason for

9

I4R DP No. 222

this being unavailability of IPD for the replication. Basically, the pooled *log(OR)* result from the following formula (Borenstein et al., 2009, p. 73):

$$\hat{\theta}_{pooled} = \frac{\sum_{i=1}^{k} \omega_i \hat{\theta}_i}{\sum_{i=1}^{k} \omega_i} \text{ whereby: } \omega_i = \frac{1}{Variance_i + \tau^2}$$

As the variance is in turn dependent on the sample size N, in such a way that a higher N leads to smaller variances and thus to higher weights  $\omega_i$ , this has the effect that individual countries with higher birth rates, that is, higher N are weighted much more heavily in the calculation of  $\theta_{pooled}$ . As stated in paragraph 2.2.1, the ITS regression modeled birth counts assuming a time- and country-invariant reference population of births. This obviously has only a limited connection with the actual birth rate of each country, which complicates comparability, as it leads to similar precision of effect estimates across countries and thus to their much more homogeneous contribution than in the original study. To ensure that the present reproduction and the original study are also comparable on a pooled effects level, the weights of the original study were used as the most economic option. The calculation of the 95% confidence intervals of the pooled effect is as follows (cf. ibid., p.74):

$$\begin{split} LL_{\theta_{pooled}} &= \hat{\theta}_{pooled} - 1.96 * SE_{\hat{\theta}} \\ UL_{\theta_{pooled}} &= \hat{\theta}_{pooled} + 1.96 * SE_{\hat{\theta}} \\ \\ \text{Whereby: } SE_{\hat{\theta}} &= \sqrt{\frac{1}{\Sigma \omega_i}} \end{split}$$

As mentioned in the previous section, the aim of this report is not to produce a "perfect model" or an exact estimate of the change in preterm birth rates, but to probe for differences in the results that might be attributable to model misspecification. With that in mind, we suggest to not solely rely on the pooled effect estimates, but to also consider the differences in the individual *ORs* computed by original study vs. this report.

#### 2.4 Sensitivity analysis

Two sensitivity analyses were conducted. First, to mirror the procedure of the original study, data from USA and Brazil were excluded in the meta-analysis for the association

Institute for Replication

I4R DP No. 222

between preterm-birth rate and lockdown to rule out the influence of the two countries with the highest birthrate by far, contributing over 70% of the births included in the original study (Calvert et al., 2023).

Second, ITS analyses were performed for all countries across all types of births, but instead of choosing the highest significant lag as the ldv in the regression model,  $Y_{t-1}$  was integrated as the ldv in all datasets, regardless of the actual autocorrelations. On one hand it is possible that the most significant lag is only influenced by the structure of the data or by chance, on the other hand lockdown in our data was only present for a maximum of four months, which means the effects of it may not be captured by including the most significant but the most recent event. As literature states, many observations that exhibit autocorrelation tend to be strongest correlated to previous observations in the very recent past anyway (Paoella, 2018). This means that including a  $Y_{t-1}$  ldv as standard, could possibly capture the effects of lockdown better while making the results more robust in case they don't deviate significantly from the original ones. All other analytical decisions in this model remain the same as described in section 2.2.1.

#### 2.5 Comparing the results

Firstly, the results of recreate reproducibility were compared with those of the original study. Then the results after inclusion of the ldv are compared with those of the previous replication study and those of the original study. Effect size estimates were compared at both country-specific and pooled level. For the specific levels as for pooled effects statistical tests such as the log-likelihood ratio test could not be used, due to the unknown code and data, and therefore the *N* of the original study. The comparison of the results was therefore carried out descriptively, looking at overlapping *Cls*. It would have been possible to compare the regression models of recreate reproducibility with those of robustness reproducibility via likelihood-ratio, but since the statistical comparison with the original study, which is the most important one, was not feasible, and as the aim of the robustness model was not to find the best fitting model, statistical comparison was omitted. The replication is considered successful if effect estimates do not differ significantly from those of the original study in terms of effect sizes and the direction of the effects (Muradchanian et al., 2021).

#### 2.6 Availability of Data and Code

Code and data used were stored in an OSF repository and can be viewed at: <u>https://osf.io/q8dct/?view\_only=8c3498e1ce324ff098466878efd85c69</u>

R version 4.4.0 was used to analyze the data. The coding was largely carried out with the aid of relevant literature (Bernal et al., 2017, 2020; Field et al., 2012). ChatGPT v4 was used in the creation of codes to produce figures as similar as possible to those in the original study.

#### **3** Results

As methods, the results part was split into data extraction, recreational reproduction, robustness reproduction, and sensitivity analysis.

#### 3.1 Data extraction

The extraction of the data points of the individual graphs resulted in one data point per month for the period of the data collected in the original study. Per country, this resulted in a maximum of 67 data points per data set (e.g. Brazil) or a minimum of 43 (Iran) with 3-4 being from lockdown. The precision of the extraction was high while not being exact, as the mean deviation of the last four observed data points of the preterm data sets from the reported observed data of the original study resulted in a percentage point value of 0.083 (very preterm: 0.012; spontaneous preterm: 0.046; stillbirth: 0.06). Wilcoxon signed rank test showed no significant difference between original and replicated data (P=0.516 for first month of observed preterm birth rates).

Intraclass correlation indicated that there were very high and significant similarities between the three extraction processes, with ICCs varying between 0.92-1.00, which allowed the mean value of the data to be used for further analysis (ICC outputs can be found in the OSF repositorium).

One irregularity was spotted in the results part of the original study, as the observed preterm data from Peru in Supplementary fig. 13-16 didn't match the corresponding graph depicted in figure 1 of the main study. This of course led to minor effect size differences that were not attributable to the method used in this report.

#### 3.2 Recreational Reproduction

Regarding preterm birth, on country specific level, effect size differences between the original and the present work were small to non-existent, with 95% *Cls* overlapping. The only exceptions being Peru, which was due to the incorrect data depicted in the graph, and Iceland, which was probably due to the high fluctuation of the preterm birth data, which leads to less precise estimates due to fewer data points used in the replication. Pooled effects show a small but significant reduction in preterm birth in the first (*OR* 0.96, 95% *Cl* 0.92-0.99, *P*=0.020), second month (*OR* 0.96, 95% *Cl* 0.94-0.99, *P*=0.004) as they did in the original study, but not in the third due to the influence of Peru, (*OR* 0.98, 95%

*Cl* 0.96-1.01, *P*=0.162) and fourth month (*OR* 0.98, 95% *Cl* 0.96-1.01, *P*=0.286). Figure 1 below shows country-specific and pooled data for preterm birth rate in the first month of lockdown. For all other months and very preterm, spontaneous preterm, and stillbirth tables please check the Supplement.

Country	Observed	Predicted	Odds Ratio	95% CI	Weights	
Australia (New South Wales*)	5.9	6	0.99	10 88 1 111	2 4%	
Belgium	81	89	0.91	[0.83 1.00]	2.6%	
Canada	8.3	9	0.93	[0 85 1 02]	7.0%	
Chile**	8.7	8.7	1.00	[0.91, 1.10]	6.0%	
Denmark	5.3	6.1	0.87	[0.77, 0.97]	0.3%	
Finland	5.6	5.4	1.04	[0.93, 1.17]	1.0%	
Hungary	8.4	8.4	1.00	[0.91, 1, 10]	1.8%	
Iceland	7.6	5.1	1.47	[1.32, 1.65]	0.1%	
Norway	7	6.8	1.03	[0.92, 1.14]	1.6%	_ <b>.</b>
Scotland	8.2	8.2	0.99	[0.90, 1.09]	1.5%	
Sweden	6	6.1	0.99	[0.88, 1.10]	2.4%	
Switzerland	6.7	6.9	0.97	[0.87, 1.08]	2.1%	
Uruguay	9.4	10	0.94	[0.86, 1.03]	1.0%	
USA**	9.8	10.2	0.96	[0.88, 1.05]	38.8%	
Wales**	8.4	8.7	0.97	[0.89, 1.07]	1.2%	
Pooled Effect (High Income)			0.97	[0.92, 1.02]		
Brazil	12.2	12.8	0.96	[0.88, 1.03]	17.9%	
Iran	8.7	9.3	0.93	[0.85, 1.02]	2.3%	
Peru**	5.8	6.7	0.87	[0.78, 0.97]	9.7%	
Pooled Effect (Upper-middle Income	)		0.93	[0.87, 0.98]		
Pooled Effect (Overall)			0.95	[0.92, 0.99]		

#### Preterm birth rate in first month of lockdown

**Figure 1:** Association between lockdown and preterm birth rates in the replication study. Individual and pooled estimates of the association between preterm birth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. The forest plot on the right shows individual ORs (blue dot) with corresponding CIs (black line). \*Births from 24 weeks onwards \*\*Live births only (For further information regarding definitions of preterm birth rates, please check original study)

For very preterm birth (Supplementary Fig. 5-8), there were some differences in forecasted predictions and therefore Odds Ratios but there were no clear significant deviances which could be detected on a descriptive level, as all the 95% *Cls* were overlapping. Backing the finding of the original study, there were no significant changes in pooled effect estimates, with ORs varying between 1.02-1.04 and *Cls* spanning 1. Taking a closer look at both high-income and upper-middle-income there were also no significant changes. For spontaneous preterm birth (Supplementary Fig. 9-12), there were also no significant changes in comparison to the original study regarding *ORs* on individual level. However, the "small relative decreases (3–4%) in the first three months" (Calvert et al., 2023, p. 531) could not be reproduced, probably due to wider *Cls* in the replication, spanning 1 on every month following lockdown.

I4R DP No. 222

Unlike preterm birth datasets, reproduced stillbirth data showed stronger transnational deviations from the original on individual levels. There were significant differences (meaning 95% Confidence Intervals not overlapping) in second and fourth month of lockdown in Iceland and in the third month in Wales. On the contrary, pooled *ORs* turned out to be quite similar to the original, with no clear evidence for an increase in stillbirth in the first (*OR* 1.06, 95% *CI* 1.00-1.13, *P*=0.067), but possibly in second (*OR* 1.07, 95% *CI* 1.00-1.14, *P*=0.051), third (*OR* 1.07, 95% *CI* 1.00-1.14, *P*=0.047) and fourth (*OR* 1.07, 95% *CI* 1.02-1.13, *P*=0.009) month of lockdown. As in the original, Brazil with contributing approx. 70% of stillbirths was the driving factor and restricting the meta-analysis to HICs only, led to effect estimates being non-significant in second (*OR* 1.00, 95% *CI* 0.95-1.05, *P*=0.980), third (*OR* 1.02, 95% *CI* 0.97-1.08, *P*=0.370) and fourth (*OR* 0.99, 95% *CI* 0.94-1.05, *P*=0.717) month of lockdown.

#### **3.3** Robustness Reproduction

Within a total of 50 ITS analyses, 35 were identified in which autocorrelation and or partial autocorrelation plots showed one or more significant lags. On preterm birth rates, only in 3 out of 10 countries did the inclusion of an ldv mean that the plots no longer showed significant lags (Belgium, Finland and Hungary). In the other cases there was no effect; ldv inclusion led to significant lags at a different position or even increased existing autocorrelation. On very preterm birth rates, in only 2/10 of the countries (Iceland and Uruguay) the ldv proofed to be effective, furthermore 2/8 on (Australia and Belgium) spontaneous- and 1/7 (Switzerland) on stillbirth rates. Figure 2 below shows the changes in acf and pact plots before and after the inclusion of a ldv, exemplified on the ITS analyses

of preterm birth Peru. To see all changes in autocorrelation in acf and pacf plots, check Supplementary Figures 33-36.



Figure 2: Changes in autocorrelation and partial autocorrelation plots in Perus preterm birth ITS analyses, before and after inclusion of a ldv. Figures on the left show lags <u>before</u> inclusion of a ldv (In this case  $Y_{t-2}$ ) into the ITS analysis, whereas figures on the right show lags <u>after</u> inclusion. Upper figures show autocorrelation, lower partial autocorrelation. The Blue dotted line indicates significance level

For preterm birth rates, there were none to minor changes on individual levels of predicted rates and therefore ORs. The only exception is Iceland, whose values surprisingly were now much closer to those of the original study (First month: *OR* 1.06, 95% *Cl* 0.96-1.17; second month: *OR* 0.79, 95% *Cl* 0.70-0.88; third month: *OR* 0.76, 95% *Cl* 0.86-0.85; fourth month: *OR* 0.78, 95% *Cl* 0.70-0.87). Brazil, the month with the highest weight among those with significant autocorrelation, showed no change whatsoever compared to the results calculated in the recreate reproduction. Pooled effects also display the same effect size estimates as the recreate reproduction, with indication of a small but significant reduction of preterm birth rates in the first (*OR* 0.96, 95% *Cl* 0.92-1.00, *P*=0.029) and second (*OR* 0.96, 95% *Cl* 0.94-0.99, *P*=0.004) but not in the third (*OR* 0.98, 95% *Cl* 0.95-1.01, *P*=0.114) and fourth month (*OR* 0.99, 95% *Cl* 0.96-1.02, *P*=0.483)

of lockdown. Individual and pooled preterm birth rates are illustrated in Figure 3, as the recreation reproducibility plots, very preterm, spontaneous preterm and stillbirth can be found in the Supplement.

Preterm birth rate in first month of lockdown

Country	Observed	Predicted	Odds Ratio	95% CI	Weights	
Australia (New South Wales*)	5.9	6	0.99	[0.88, 1.11]	2.4%	
Belgium (with Idv)	8.1	8.6	0.94	[0.86, 1.04]	2.6%	
Canada (with ldv)	8.3	9	0.93	[0.85, 1.02]	7.0%	<b>-</b>
Chile**	8.7	8.7	1.00	[0.91, 1.10]	6.0%	
Denmark (with Idv)	5.3	6.1	0.87	[0.77, 0.97]	0.3%	
Finland (with Idv)	5.6	5.5	1.02	[0.91, 1.15]	1.0%	
Hungary (with Idv)	8.4	8.7	0.96	[0.88, 1.06]	1.8%	
Iceland (with Idv)	7.6	7.2	1.06	[0.96, 1.17]	0.1%	
Norway	7	6.8	1.03	[0.92, 1.14]	1.6%	
Scotland	8.2	8.2	0.99	[0.90, 1.09]	1.5%	
Sweden	6	6.1	0.99	[0.88, 1.10]	2.4%	
Switzerland	6.7	6.9	0.97	[0.87, 1.08]	2.1%	
Uruguay (with Idv)	9.4	10.1	0.93	[0.85, 1.02]	1.0%	
USA**	9.8	10.2	0.96	[0.88, 1.05]	38.8%	
Wales** (with Idv)	8.4	8.6	0.98	[0.89, 1.08]	1.2%	
Pooled Effect (High Income)			0.96	[0.92, 1.02]		
Brazil (with ldv)	12.2	12.8	0.96	[0.88, 1.03]	17.9%	
Iran	8.7	9.3	0.93	[0.85, 1.02]	2.3%	
Peru** (with Idv)	5.8	6.5	0.90	[0.80, 1.00]	9.7%	
Pooled Effect (Upper-middle Income)			0.94	[0.88, 1.00]		
Pooled Effect (Overall)			0.96	[0.92, 1.00]		
						0.775 0.8 0.825 0.85 0.875 0.9 0.925 0.95 0.975 1 1.025 1.05 1.075 1.1 1.125 1.15 1.17

**Figure 3:** Association between lockdown and preterm birth rates in the replication study after inclusion of a ldv. Individual and pooled estimates of the association between preterm birth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. Countries, for which a ldv was inserted into the Poisson regression, are marked with (with ldv). The forest plot on the right shows individual ORs (blue dot) with corresponding CIs (black line). \*Births from 24 weeks onwards \*\*Live births only (For further information regarding definitions of preterm birth rates, please check original study)

Similar observations could also be made when inspecting effect estimates for very preterm birth data. At the individual level, Iceland again shows a significant change, although this again consists of moving closer to the effect estimators of the original study. Denmark instead took the opposite approach and showed results that were significantly further away from the original study. Among the countries with high weightings, Brazil, and particularly its second (*OR* 0.97, 95% *Cl* 0.80-1.18 compared to *OR* 1.04, 95% *Cl* 0.85-1.27), third (*OR* 0.93, 95% *Cl* 0.76-1.33 compared to *OR* 1.01, 95% *Cl* 0.85-1.23) and fourth (*OR* 0.99, 95% *Cl* 0.81-1.21 compared to *OR* 1.07, 95% *Cl* 0.87-1.30) month showed a change in direction of effect. The inclusion of a ldv here did not lead to a disappearance of significant lags but merely to a shift (Supplementary Fig. 34), which is why the results should be viewed with caution. At a pooled level, the results of the meta-analysis remained almost unchanged compared to the recreate reproduction. The only exception

Institute for Replication

I4R DP No. 222

here was the upper-middle income data in fourth month (*OR* 0.99, 95% *Cl* 0.86-1.15, *P*=0.928 compared to *OR* 1.04, 95% *Cl* 0.90-1.20, *P*=0.594), as these are subject to a strong influence from Brazil.

Spontaneous preterm birth rates showed effect estimate differences at an individual level in Wales in first (*OR* 0.99, 95% *Cl* 0.86-1.14 compared to *OR* 1.15, 95% *Cl* 0.99-1.33), second (*OR* 0.89, 95% *Cl* 0.76-1.02 compared to *OR* 1.13, 95% *Cl* 0.96-1.31), third (*OR* 0.96, 95% *Cl* 0.82-1.11 compared to *OR* 1.14, 95% *Cl* 0.97-1.33) and fourth (*OR* 1.27, 95% *Cl* 1.11-1.46 compared to *OR* 1.49, 95% *Cl* 1.28-1.72) month following lockdown and in Brazil in first month (*OR* 0.93, 95% *Cl* 0.85-1.01 compared to *OR* 1.02, 95% *Cl* 0.93-1.11). The minor individual changes in most countries and larger ones in Wales came closer to the effect size estimates of the original study. However, Brazil once again did not seem to respond to the inclusion of a ldv in the desired way (Expended Data Figure 35), as was again demonstrated by significant lags even after the inclusion. Due to minor changes in the effect size estimates of other countries, particularly the USA, pooled effects after inclusion of the ldv showed results on HIC subgroup data in the first month, that were also closer to those of the original study (*OR* 0.96, 95% *Cl* 0.90-1.02, *P*=0.069). Beyond this, there were no changes in the meta-analysis.

Stillbirth rates showed no clear trend in terms of whether they moved closer to or further away from the values of the original study. At the same time, only Sweden and Switzerland showed a reduction in significant lags to below the significance level. For all others, there was only a shift or no response after inclusion of the ldv. Brazil, the country that had by far the highest weighting of all countries, differed from the original study in the second (*OR* 1.03, 95% *Cl* 0.94-1.16), third (*OR* 1.03, 95% *Cl* 0.97-1.12) and fourth (*OR* 1.03, 95% *Cl* 0.94-1.12) months more clearly than the effect size estimates of the ITS analysis without ldv. However, there was no satisfactory reduction in autocorrelation after inclusion of the ldv in Brazil once more. Due to Brazil's high influence and lesser effect size estimates while remaining the same direction, pooled effect size estimates overall turned insignificant in second (*OR* 1.02, 95% *Cl* 0.94-1.10, *P*=0.579), third (*OR* 1.03, 95% *Cl* 0.97-1.10, *P*=0.367) and fourth (*OR* 1.02, 95% *Cl* 0.96-1.07, *P*=0.563) month following lockdown, while remaining insignificant in first (*OR* 1.05, 95% *Cl* 0.99-1.12, *P*=0.128). Contrary to this, second (*OR* 0.95, 95% *Cl* 0.90-1.00, *P*=0.035) and fourth (*OR* 

I4R DP No. 222

0.92, 95% *CI* 0.87-0.97, *P*=0.004) month of pooled HIC subgroup data now showed a significant decrease in stillbirth. This was due to lesser effect estimates in Finland and Belgium in second month, and in Belgium, Norway, and Switzerland in fourth month. Following table shows the comparison of original data's-, recreate reproducibility's and robustness reproducibility's effect sizes. For very preterm, spontaneous and stillbirth comparisons, please check the Supplementary Tables 1-4.

Effect size estimates	Original Study (1)	Recreate Reproduction (2)	Robustness Reproduction (with ldv) (3)
Month 1			
High-Income	OR 0.96	OR 0.97	OR 0.96
	95% CI 0.94-0.98	95% CI 0.92-1.02	95% CI 0.92-1.02
		P=0.175	P=0.168
Upper-middle Income	OR 0.96	OR 0.93	OR 0.94
	95% CI 0.90-1.02	95% CI 0.87-0.98	95% CI 0.88-1.00
		P=0.013	P=0.035
Total	OR 0.96	OR 0.95	OR 0.96
	95% CI 0.95-0.98	95% CI 0.92-0.99	95% CI 0.92-1.00
	P<0.001	P=0.020	P=0.029
Month 2			
High-Income	OR 0.97	OR 0.97	OR 0.97
	95% CI 0.95-0.99	95% CI 0.94-0.99	95% CI 0.94-0.99
		P=0.016	P=0.017
Upper-middle Income	OR 0.93	OR 0.96	OR 0.98
	95% CI 0.79-1.08	95% CI 0.91-1.01	95% CI 0.90-1.01
		P=0.126	P=0.095
Total	OR 0.96	OR 0.96	OR 0.96
	95% CI 0.92-0.99	95% CI 0.94-0.99	95% CI 0.94-0.99
	P=0.03	P=0.004	P=0.004

Table 1 – Comparison of estimated effects gained through meta-analysis in preterm birth rates

Month 3

High-Income	OR 0.97	OR 0.98	OR 0.98
	95% CI 0.95-0.99	95% CI 0.95-1.01	95% CI 0.98-1.01
		P=0.161	P=0.123
Upper-middle Income	OR 0.95	OR 0.99	OR 0.98
	95% CI 0.83-1.09	95% CI 0.94-1.04	95% CI 0.93-1.04
		P=0.676	P=0.587
Total	OR 0.97	OR 0.98	OR 0.98
	95% CI 0.94-1.00	95% CI 0.96-1.01	95% CI 0.95-1.01
	P=0.09	P=0.162	P=0.114
Month 4			
High-Income	OR 0.98	OR 0.99	OR 0.99
	95% CI 0.96-1.01	95% CI 0.96-1.02	95% CI 0.95-1.02
		P=0.551	P=0.538
Upper-middle Income	OR 0.98	OR 0.97	OR 0.99
	95% CI 0.90-1.07	95% CI 0.92-1.03	95% CI 0.94-1.05
		P=0.323	P=0.737
Total	OR 0.99	OR 0.98	OR 0.99
	95% CI 0.96-1.01	95% CI 0.98-1.01	95% CI 0.96-1.02
	P=0.34	P=0.286	P=0.483

*Note:* P-Values of the original study could not be found for every subgroup effect size estimate, please check Calvert et. al p.530ff. for details. P-Values of the reproduction were calculated using Wald-test

#### 3.4 Sensitivity analysis

As in the original study, excluding Brazil and the United States in the meta-analysis calculating ORs for the association between lockdown and preterm birth led to imperceptible changes in effect estimates (Supplementary Table 5). Likewise, the inclusion of the ldv  $Y_{t-1}$  for every ITS analysis conducted also led to very similar results on country-specific and pooled levels on preterm and stillbirth data to the recreate reproduction with negligible differences in detail (Supplementary Table 6).

I4R DP No. 222

#### 4 Discussion

The aim of this study was to reproduce the work "Changes in preterm birth and stillbirth during COVID-19 lockdowns in 26 countries" and to check its robustness, whereby only the population-based data were of interest. A successful replication was predefined as one that does not differ significantly from the original in terms of direction of the effect and effect size. While the main study found evidence for "small (3-4%) relative reduction in the overall preterm births following lockdown" (Calvert et al., 2023, p. 536), this study replicated these small findings in 3 of 4 months, with a reduction in preterm birth rates of 4-5%. For very preterm birth rates, there was also no evidence for an impact of lockdown found. Spontaneous preterm birth rates showed similar ORs than the original, while having wider CIs and no significant P-values in all four months in HIC countries. All these results also appeared to be robust to the inclusion of a ldv at pooled levels, with only minor deviations from the original and recreated results. Stillbirth, on the other hand, showed changes in the effect size in months 2-4, which was first and foremost due to the change in effect sizes in Brazil as a result of the ldv's influence.

To check whether the reproduction can be assessed as successful, how the individual significant deviations presumably occurred must be understood. To this end, we will first look at the strengths and limitations of this present work.

The main limitations are the lack of original data and the lack of access to the original code. So, although the extraction of data from graphs has proven to be relatively accurate and reliable, deviations in results cannot be solely attributed to the method only but can also be a consequence of minor discrepancies in data. Another problem arises from the economical decision of setting the offset of the regression model generally to 10000, as this does not resemble the original birth rates and leads a homogenous distribution of births between countries in reproduction. As described in Section 2.3, an attempt was made to counter this problem by using the weightings of the original study for calculating the effect size estimators in the reproduction. However, this has the consequence that variances of the original study are also included in the calculation, which weakens the influence of the reproduction variances and thus possible effects of the inclusion of a ldv on the pooled results. Utilizing the weightings of the original study

21

I4R DP No. 222

also prevents individual monthly weightings in the reproduction, which leads to a very important problem itself. In a regression model that uses the same offset in each month, the residuals of a certain point in time may show up differently in relation to the residuals of another point in time than they would with non-homogeneous offsets. And as the weightings of the individual months are missing, the variability is somewhat smoothed, which can lead to an insufficient explanation of fluctuations by the regression model. This can lead to an amplification or emergence of autocorrelation. In addition, inconsistencies between the data implied in the preterm mapping of Peru and the observed rates, made comparability between both works harder. Knowledge of the original code would have enabled a statistical comparison (log-likelihood ratio test), which would've been more insightful than the descriptive comparison.

Another limitation is, that there was neither checking nor controlling for overdispersion done. Especially graphs in stillbirth indicate high variance in monthly stillbirth occurrence and therefore make overdispersion quite likely, which, if it really was the case, meant that a quasi-Poisson model would have been the better choice.

One last important limitation is the method used to control for autocorrelation itself. The method in this study removes e.g. three data points of the dependent variable from the data set when a ldv of  $Y_{t-3}$  (four for  $Y_{t-4}$  etc.) is included. A better model fit after inclusion of the ldv could therefore also merely be the result of a reduced amount of data, which is why the sensitivity analysis with a general ldv of  $Y_{t-1}$  was also of interest. Additionally, the inclusion of a single ldv proved to be insufficient in most cases to really eliminate autocorrelation, but in many cases just led to a shift in lag structure. Other methods such as ARIMA or SARIMA (Ghysels et al., 2006; Nelson, 1998), could have maybe done a better job getting rid of autocorrelation.

One strength of the reproduction is, that although the original code is not available, the tutorial used by the authors is freely accessible. So, although a different regression approach was chosen, within these regression models most of analytical decisions such as evaluating the trend as a linear, square, quadratic, logarithmic and second-order polynomial and choosing the best fit via AIC, or allowance for different within-period trend and intercept for the pre-lockdown and lockdown periods, presumably remain the same. Another strength is that extracted data is very close to the Institute for Replication

I4R DP No. 222

original while showing high reliability scores. These strengths both allow the results of robustness reproduction to be meaningful compared and interpreted.

With all this in mind, we can now check if the reproduction can be assessed as successful. In recreational reproduction, the results at individual and pooled level were very similar to those of the original study except for most of Iceland and stillbirth rates. While the pooled effect sizes are also similar because they use the variances of the original study for their calculation, there are also mainly marginal deviations at the country-specific level. Minor differences can probably be attributed to the method being used to extract data from graphs. Larger differences can probably be explained by the high variability within the original data caused by the comparatively lower event occurrence which might have led to overdispersion, this is particularly visible in stillbirths. Robustness reproduction provides no clear evidence of significant changes in effect estimates following the inclusion of a ldv. In preterm birth data, only a few countries showed changes compared to the recreational reproduction, as e.g. Brazil in very preterm birth or Wales in spontaneous preterm birth. But since the deviations from Wales, like other smaller deviations, were closer to the original data, and the deviations from Brazil are presumably explained by an inadequate model fit (because of remaining significant lags after adding a ldv), none of this provides a sufficient evidence base for the assumption that the method of the model is not robust against the inclusion of a lagged dependent variable.

Results of the sensitivity analysis further support these findings. As in the original study, effect size estimates did not change after exclusion of USA and Brazil in preterm birth rates, which speaks for the robustness of the original method. Inclusion of a general ldv of  $Y_{t-1}$  also negligible effects on preterm and stillbirth rates, indicating that regression models created seem to be robust against changes in autocorrelation structure.

In summary, this study succeeded in replicating most of the findings of the original study. The main deviations between are presumably due to methodological weaknesses in reproduction, such as small differences in used data, homogenous offsets, missing control for overdispersion and insufficient ability of the robustness reproduction model to counter autocorrelation. The recreational reproduction can therefore be viewed as a success and the results of the original work as reproducible. However, checking the autocorrelation plots revealed significant autocorrelation and partial autocorrelation in most of the regressions conducted (which could also be a consequence of overdispersion). While adding a specific or general ldv did not change much of calculated predicted birth rates, especially in preterm birth datasets, there might be small differences possible using a more elaborate method than just adding one ldv. Nevertheless, the fact that most effect size estimates do not change by adding an ldv, although it changes the autocorrelation structure, is an indication of robustness.

#### **5** References

- Akaike, H. (1998). Information Theory and an Extension of the Maximum Likelihood Principle. In
  E. Parzen, K. Tanabe, & G. Kitagawa (Eds.), *Selected Papers of Hirotugu Akaike* (pp. 199–213). Springer. https://doi.org/10.1007/978-1-4612-1694-0\_15
- Akula, L. (2020, October 16). *How to add Fourier terms to your regression & seasonality analysis (using Python & SciPy)*. Medium. https://towardsdatascience.com/how-to-add-fourierterms-to-your-regression-seasonality-analysis-using-python-scipy-99a94d3ae51
- Bernal, J. L., Cummins, S., & Gasparrini, A. (2017). Interrupted time series regression for the evaluation of public health interventions: A tutorial. *International Journal of Epidemiology*, 46(1), 348–355. https://doi.org/10.1093/ije/dyw098
- Bernal, J. L., Cummins, S., & Gasparrini, A. (2020). Corrigendum to: Interrupted time series regression for the evaluation of public health interventions: a tutorial. *International Journal of Epidemiology*, *50*(3), 1045. https://doi.org/10.1093/ije/dyaa118
- Borenstein, M., Hedges, L. V., Higgins, J., & Rothstein, H. (2009). Random-Effects Model. In Introduction to Meta-Analysis (pp. 69–75). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470743386.ch12
- Brodeur, A., Dreber, A., Hoces de la Guardia, F. *et al.* Reproduction and replication at scale. *Nat Hum Behav* **8**, 2–3 (2024). https://doi.org/10.1038/s41562-023-01807-2
- Calvert, C., Brockway, M. (Merilee), Zoega, H., Miller, J. E., Been, J. V., Amegah, A. K., Racine-Poon, A., Oskoui, S. E., Abok, I. I., Aghaeepour, N., Akwaowo, C. D., Alshaikh, B. N., Ayede, A. I., Bacchini, F., Barekatain, B., Barnes, R., Bebak, K., Berard, A., Bhutta, Z. A., ... Azad, M. B. (2023). Changes in preterm birth and stillbirth during COVID-19 lockdowns in 26 countries. *Nature Human Behaviour*, *7*(4), 529–544. https://doi.org/10.1038/s41562-023-01522-y

- DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, 7(3), 177–188. https://doi.org/10.1016/0197-2456(86)90046-2
- Dreber, A., & Johanneson, M. (2023). A Framework for Evaluating Reproducibility and Replicability in Economics. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4458153
- Drevon, D., Fursa, S. R., & Malcolm, A. L. (2017). Intercoder Reliability and Validity of WebPlotDigitizer in Extracting Graphed Data. *Behavior Modification*, *41*(2), 323–339. https://doi.org/10.1177/0145445516673998

Field, A., Miles, J., & Field, Z. (2012). *Discovering Statistics Using R* (1st ed.). SAGE Publications Ltd.

Ghysels, E., Osborn, D. R., & Rodrigues, P. M. M. (2006). Chapter 13 Forecasting Seasonal Time
Series. In G. Elliott, C. W. J. Granger, & A. Timmermann (Eds.), *Handbook of Economic Forecasting* (Vol. 1, pp. 659–711). Elsevier. https://doi.org/10.1016/S15740706(05)01013-X

Keele, L., & Kelly, N. (2005). Dynamic Models for Dynamic Theories: The Ins and Outs of Lagged Dependent Variables. https://web.archive.org/web/20160604130444/https://www.nuffield.ox.ac.uk/politics/ papers/2005/Keele%20Kelly%20LDV.pdf

- Muradchanian, J., Hoekstra, R., Kiers, H., & van Ravenzwaaij, D. (2021). How best to quantify replication success? A simulation study on the comparison of replication success metrics. *Royal Society Open Science*, *8*(5), 201697. https://doi.org/10.1098/rsos.201697
- Nelson, B. K. (1998). Statistical methodology: V. Time series analysis using autoregressive integrated moving average (ARIMA) models. Academic Emergency Medicine: Official Journal of the Society for Academic Emergency Medicine, 5(7), 739–744. https://doi.org/10.1111/j.1553-2712.1998.tb02493.x
- Paoella, M. (2018). The AR(1) Model. In *Linear Models and Time-Series Analysis* (pp. 185–221). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119432036.ch4

Rohatgi, A. (2024). *Web Plot Digitizer* [Computer software]. https://automeris.io/WebPlotDigitizer.html

Stocker, H. (2023). *Methoden der empirischen Wirtschaftsforschung*. Universität Innsbruck.

- Van der Mierden, S., Spineli, L. M., Talbot, S. R., Yiannakou, C., Zentrich, E., Weegh, N., Struve, B., Zur Brügge, T. F., Bleich, A., & Leenaars, C. H. C. (2021). Extracting data from graphs: A case-study on animal research with implications for meta-analyses. *Research Synthesis Methods*, 12(6), 701–710. https://doi.org/10.1002/jrsm.1481
- Woolson, R. F. (2005). Wilcoxon Signed-Rank Test. In *Encyclopedia of Biostatistics*. John Wiley & Sons, Ltd. https://doi.org/10.1002/0470011815.b2a15177

### **Supplement**

#### **Supplementary Figures**

Preterm birth rate in first month of lockdown								
Country	Observed	Predicted	Odds Ratio	95% CI	Weights			
Australia (New South Wales*)	5.9	6	0.99	[0.88, 1.11]	2.4%			
Belgium	8.1	8.9	0.91	[0.83, 1.00]	2.6%			
Canada	8.3	9	0.93	[0.85, 1.02]	7.0%			
Chile**	8.7	8.7	1.00	[0.91, 1.10]	6.0%			
Denmark	5.3	6.1	0.87	[0.77, 0.97]	0.3%			
Finland	5.6	5.4	1.04	[0.93, 1.17]	1.0%			
Hungary	8.4	8.4	1.00	[0.91, 1.10]	1.8%			
Iceland	7.6	5.1	1.47	[1.32, 1.65]	0.1%			
Norway	7	6.8	1.03	[0.92, 1.14]	1.6%			
Scotland	8.2	8.2	0.99	[0.90, 1.09]	1.5%			
Sweden	6	6.1	0.99	[0.88, 1.10]	2.4%			
Switzerland	6.7	6.9	0.97	[0.87, 1.08]	2.1%			
Uruguay	9.4	10	0.94	[0.86, 1.03]	1.0%			
USA**	9.8	10.2	0.96	[0.88, 1.05]	38.8%			
Wales**	8.4	8.7	0.97	[0.89, 1.07]	1.2%			
Pooled Effect (High Income)			0.97	[0.92, 1.02]				
Brazil	12.2	12.8	0.96	[0.88, 1.03]	17.9%			
Iran	8.7	9.3	0.93	[0.85, 1.02]	2.3%			
Peru**	5.8	6.7	0.87	[0.78, 0.97]	9.7%			
Pooled Effect (Upper-middle Income)			0.93	[0.87, 0.98]				
Pooled Effect (Overall)			0.95	[0.92, 0.99]				
						1		

Supplementary Figure 1: Association between lockdown and preterm birth rates in the replication study. Individual and pooled estimates of the association between preterm birth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. The forest plot on the right shows individual ORs (blue dot) with corresponding Cis (black line). \*Births from 24 weeks onwards \*\*Live births only (Please check original study for further information)

#### Preterm birth rate in second month of lockdown

Country	Observed	Predicted	Odds Ratio	95% CI	Weights	
Australia (New South Wales*)	6.3	6	1.04	[0.93, 1.17]	5.9%	
Belgium	8.2	8.9	0.92	[0.83, 1.01]	5.8%	
Canada	8.4	9.4	0.90	[0.82, 0.99]	8.7%	
Chile**	9	8.7	1.04	[0.95, 1.14]	8.2%	
Denmark	5.9	5.5	1.07	[0.95, 1.20]	1.3%	
Finland	5.9	5.4	1.09	[0.97, 1.23]	3.3%	
Hungary	8.7	8.6	1.02	[0.93, 1.12]	4.8%	
Iceland	5.6	5	1.11	[0.98, 1.25]	0.4%	
Norway	6.4	6.8	0.95	[0.85, 1.06]	4.2%	
Scotland	7.9	8.2	0.96	[0.87, 1.06]	4.2%	
Sweden	5.5	6	0.93	[0.83, 1.05]	5.5%	
Switzerland	6.3	7	0.91	[0.82, 1.02]	5.2%	
Uruguay	9.4	10	0.93	[0.86, 1.03]	3.2%	
USA**	10	10.4	0.96	[0.88, 1.05]	11.1%	
Wales**	7.5	8.4	0.90	[0.82, 0.99]	3.4%	
Pooled Effect (High Income)			0.97	[0.94, 0.99]		
Brazil	12	12	1.00	[0.92, 1.08]	10.2%	
Iran	9.2	9.6	0.96	[0.88, 1.05]	5.7%	
Peru**	6	6.6	0.91	[0.81, 1.01]	8.9%	+
Pooled Effect (Upper-middle Income)			0.96	[0.91, 1.01]		
Pooled Effect (Overall)			0.96	[0.94, 0.99]		
						0.8 0.825 0.85 0.875 0.9 0.925 0.95 0.975 1 1.025 1.05 1.075 1.1 1.125 1.15 1.175 1.2 1.221 Odds Ratio

Supplementary Figure 2: Association between lockdown and preterm birth rates in the replication study. Individual and pooled estimates of the association between preterm birth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. The forest plot on the right shows individual ORs (blue dot) with corresponding Cis (black line). \*Births from 24 weeks onwards \*\*Live births only (Please check original study for further information)

Country	Observed	Predicted	Odds Ratio	95% CI	Weights	
Australia (New South Wales*)	6.1	6.1	1.01	[0.90, 1.13]	5.4%	
Belgium	8.4	8.9	0.94	[0.86, 1.03]	5.5%	
Canada	8.4	8.8	0.95	[0.86, 1.04]	9.0%	
Chile**	8.5	8.8	0.96	[0.88, 1.06]	8.2%	
Denmark	6.6	5.6	1.17	[1.04, 1.31]	1.2%	· · · · · · · · · · · · · · · · · · ·
Finland	5.5	5.3	1.05	[0.93, 1.18]	2.7%	
Hungary	8.6	8.7	0.98	[0.90, 1.08]	4.4%	
Iceland	5.6	4.7	1.18	[1.05, 1.34]	0.3%	· · · · · · · · · · · · · · · · · · ·
Norway	5.8	6.9	0.85	[0.76, 0.95]	3.6%	
Scotland	9	8	1.11	[1.01, 1.23]	4.1%	
Sweden	5.6	5.8	0.96	[0.86, 1.08]	5.1%	
Switzerland	7.3	6.9	1.06	[0.95, 1.17]	5.0%	_ <b>+</b> •
Uruguay	9.8	10.4	0.94	[0.86, 1.03]	2.8%	
USA**	10.1	10.4	0.97	[0.89, 1.05]	12.8%	
Wales**	8	8.2	0.98	[0.89, 1.08]	3.6%	
Pooled Effect (High Income)			0.98	[0.95, 1.01]		
Brazil	12.2	11.8	1.03	[0.95, 1.11]	11.3%	
Iran	9.8	9.9	0.99	[0.91, 1.08]	5.3%	
Peru**	6.3	6.7	0.94	[0.84, 1.04]	9.4%	
Pooled Effect (Upper-middle Income)			0.99	[0.94, 1.04]		
Pooled Effect (Overall)			0.98	[0.96, 1.01]		
						0.75 0.80.825 0.8750.90.925 0.975 1 1.025 1.0751.11.125 1.1751.21.225 1.2751.31.325 Odds Ratio

#### Preterm birth rate in third month of lockdown

Supplementary Figure 3: Association between lockdown and preterm birth rates in the replication study. Individual and pooled estimates of the association between preterm birth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. The forest plot on the right shows individual ORs (blue dot) with corresponding Cis (black line). \*Births from 24 weeks onwards \*\*Live births only (Please check original study for further information)

Country	Observed	Predicted	Odds Ratio	95% CI	Weights	
Belgium	9.1	8.8	1.03	[0.94, 1.13]	5.5%	<b>-</b>
Canada	8.7	8.7	1.00	[0.91, 1.10]	10.8%	
Chile**	8.3	8.8	0.93	[0.85, 1.03]	8.8%	
Denmark	5.5	5.8	0.95	[0.85, 1.07]	0.8%	
Finland	5.4	5.3	1.03	[0.91, 1.16]	2.4%	
Hungary	8.8	8.6	1.02	[0.93, 1.12]	4.2%	<b>_</b>
Iceland	5.5	4.1	1.35	[1.19, 1.54]	0.2%	
Norway	6.4	6.9	0.93	[0.84, 1.04]	3.3%	
Scotland	8.3	7.7	1.08	[0.98, 1.19]	3.5%	
Switzerland	7	6.7	1.05	[0.95, 1.17]	4.5%	
Uruguay	9.6	11.1	0.86	[0.79, 0.94]	2.2%	
USA**	10	10.3	0.98	[0.90, 1.07]	19.7%	
Wales**	8.5	8.3	1.02	[0.93, 1.12]	2.8%	
Pooled Effect (High Income)			0.99	[0.96, 1.02]		
Brazil	12	11.9	1.01	[0.93, 1.09]	14.8%	
Iran	10.1	10.1	1.00	[0.91, 1.09]	5.3%	
Peru**	6.2	6.9	0.91	[0.81, 1.01]	10.9%	
Pooled Effect (Upper-middle Income)			0.97	[0.92, 1.03]		
Pooled Effect (Overall)			0.98	[0.96, 1.01]		
						1 0d

#### Preterm birth rate in fourth month of lockdown

Supplementary Figure 4: Association between lockdown and preterm birth rates in the replication study. Individual and pooled estimates of the association between preterm birth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. The forest plot on the right shows individual ORs (blue dot) with corresponding Cis (black line). \*Births from 24 weeks onwards \*\*Live births only (Please check original study for further information)

Country	Observed	Predicted	Odds Ratio	95% CI	Weights	
Australia (New South Wales*)	1.1	0.9	1.21	[0.91, 1.60]	5.4%	
Belgium	1.3	1.5	0.84	[0.66, 1.06]	3.6%	
Canada	1.3	1.3	1.00	[0.78, 1.27]	6.9%	
Chile**	1.4	1.3	1.03	[0.81, 1.31]	7.3%	
Denmark	1.2	0.6	1.95	[1.43, 2.65]	0.7%	· · · · · · · · · · · · · · · · · · ·
Finland	0.6	0.6	1.02	[0.72, 1.44]	0.7%	
Hungary	1.5	1.3	1.10	[0.87, 1.39]	2.9%	_ <b>-</b>
Iceland	0.9	1.5	0.60	[0.46, 0.78]	0.2%	
Norway	1.2	1.2	1.02	[0.80, 1.32]	2.9%	
Scotland	1.4	1	1.39	[1.08, 1.79]	1.8%	
Sweden	0.9	1	0.91	[0.69, 1.21]	4.1%	
Switzerland	1.4	1.1	1.25	[0.98, 1.61]	4.2%	
Uruguay	1.3	1.2	1.09	[0.86, 1.40]	2.1%	_ <b>_</b> +•
USA**	1.5	1.5	0.98	[0.78, 1.23]	19.3%	<b>_</b>
Wales*	2.1	1.4	1.51	[1.21, 1.87]	1.8%	
Pooled Effect (High Income)			1.05	[0.96, 1.15]		
Brazil	2	2	1.02	[0.84, 1.24]	19.1%	<b>_</b>
Iran	1.7	1.8	0.96	[0.77, 1.18]	7.8%	<b>-</b>
Peru**	0.8	0.9	0.94	[0.70, 1.26]	9.0%	
Pooled Effect (Upper-middle Income)			0.99	[0.86, 1.13]		
Pooled Effect (Overall)			1.03	[0.95, 1.11]		<del>_</del>
						0.5 1 1.5 2 2.5 Odds Ratio

#### Very preterm birth rate in first month of lockdown

Supplementary Figure 5: Association between lockdown and very preterm birth rates in the replication study. Individual and pooled estimates of the association between preterm birth and lockdown. Observed rates are based on original studies Supplementary Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. The forest plot on the right shows individual ORs (blue dot) with corresponding Cis (black line). \*Births from 24 weeks onwards \*\*Live births only (Please check original study for further information)

Country	Observed	Predicted	Odds Ratio	95% CI	Weights	
Australia (New South Wales*)	1	0.9	1.14	[0.85, 1.52]	5.7%	
Belgium	1.4	1.5	0.94	[0.75, 1.18]	4.3%	
Canada	1.2	1.3	0.92	[0.71, 1.18]	7.2%	
Chile**	1.3	1.4	0.97	[0.76, 1.23]	7.7%	<b>-</b>
Denmark	1	0.6	1.63	[1.19, 2.23]	0.7%	
Finland	1.1	0.6	1.86	[1.35, 2.57]	1.5%	
Hungary	1.7	1.5	1.15	[0.93, 1.44]	3.8%	
Iceland	0.8	1.6	0.49	[0.37, 0.64]	0.2%	- <b>-</b>
Norway	1.3	1.2	1.04	[0.82, 1.34]	3.5%	
Scotland	1.5	0.9	1.59	[1.23, 2.06]	2.3%	
Sweden	1.1	0.9	1.16	[0.88, 1.53]	5.4%	
Switzerland	1.3	1	1.24	[0.95, 1.61]	4.6%	
Uruguay	1.7	1.2	1.43	[1.13, 1.81]	2.8%	
USA**	1.5	1.5	0.98	[0.78, 1.23]	16.2%	
Wales*	1.3	1.3	0.99	[0.78, 1.26]	1.5%	
Pooled Effect (High Income)			1.08	[0.99, 1.17]		
Brazil	2	1.9	1.04	[0.85, 1.27]	15.9%	
Iran	1.8	1.8	1.01	[0.82, 1.24]	8.4%	
Peru**	0.7	0.9	0.80	[0.59, 1.09]	8.2%	
Pooled Effect (Upper-middle Income)			0.97	[0.84, 1.11]		
Pooled Effect (Overall)			1.04	[0.97, 1.12]		<b>—</b>
						0.5 1 1.5 2 2 Odds Ratio

Very preterm birth rate in second month of lockdown

Supplementary Figure 6: Association between lockdown and very preterm birth rates in the replication study. Individual and pooled estimates of the association between preterm birth and lockdown. Observed rates are based on original studies Supplementary Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. The forest plot on the right shows individual ORs (blue dot) with corresponding Cis (black line). \*Births from 24 weeks onwards \*\*Live births only (Please check original study for further information)

Country	Observed	Predicted	Odds Ratio	95% CI	Weights	
Australia (New South Wales*)	0.9	0.9	1.02	[0.76, 1.38]	5.5%	_ <b>-</b>
Belgium	1.6	1.5	1.08	[0.86, 1.32]	4.8%	_ <del></del>
Canada	1.1	1.2	0.90	[0.69, 1.16]	7.2%	_ <b>+</b> _
Chile**	1.3	1.4	0.94	[0.74, 1.19]	7.9%	_ <b>-</b> +_
Denmark	2	0.7	2.86	[2.17, 3.76]	1.2%	
Finland	0.7	0.5	1.45	[1.01, 2.09]	1.1%	
Hungary	1.5	1.5	1.00	[0.80, 1.26]	3.8%	
Iceland	0.6	1.5	0.39	[0.29, 0.52]	0.2%	
Norway	1	1.2	0.79	[0.61, 1.04]	3.1%	
Scotland	1.6	0.9	1.88	[1.45, 2.43]	2.5%	
Sweden	0.8	0.9	0.90	[0.67, 1.21]	4.6%	
Switzerland	1.2	0.9	1.24	[0.95, 1.63]	4.6%	+•
Uruguay	1.5	1.2	1.24	[0.98, 1.58]	2.7%	
USA**	1.5	1.5	0.99	[0.79, 1.24]	16.3%	_ <del></del>
Wales*	1.4	1.4	0.97	[0.76, 1.23]	1.9%	<b>_</b> _
Pooled Effect (High Income)			1.04	[0.96, 1.12]		
Brazil	1.9	1.9	1.01	[0.82, 1.23]	15.8%	_ <b>_</b>
Iran	1.9	1.8	1.08	[0.88, 1.32]	8.4%	
Peru**	0.8	0.9	0.86	[0.63, 1.17]	8.6%	
Pooled Effect (Upper-middle Income)			0.99	[0.86, 1.13]		
Pooled Effect (Overall)			1.02	[0.95, 1.09]		<del></del>
						0.5 1 1.5 2 2.5 3 3.5 4 Odds Ratio

#### Very preterm birth rate in third month of lockdown

Supplementary Figure 7: Association between lockdown and very preterm birth rates in the replication study. Individual and pooled estimates of the association between preterm birth and lockdown. Observed rates are based on original studies Supplementary Fig. 1, whereas Predicted rates were conducted using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using metaanalysis with weights from the original study. The forest plot on the right shows individual ORs (blue dot) with corresponding Cis (black line). \*Births from 24 weeks onwards \*\*Live births only (Please check original study for further information)

### Very preterm birth rate in fourth month of lockdown

Country	Observed	Predicted	Odds Ratio	95% CI	Weights	
Belgium	1.6	1.4	1.09	[0.87, 1.37]	3.0%	_ <b>-</b>
Canada	1.2	1.2	1.01	[0.78, 1.29]	6.1%	
Chile**	1.3	1.3	0.94	[0.74, 1.20]	5.9%	<b>-</b> _
Denmark	1.4	0.7	2.02	[1.51, 2.71]	0.5%	
Finland	0.8	0.4	1.81	[1.25, 2.61]	0.7%	
Hungary	1.2	1.3	0.98	[0.77, 1.26]	2.1%	<b>-</b>
Iceland	1.2	1.9	0.61	[0.48, 0.76]	0.2%	- <b>-</b> -
Norway	0.9	1.1	0.84	[0.64, 1.11]	1.8%	_ <b></b> +
Scotland	1.2	0.8	1.47	[1.11, 1.95]	1.2%	· · · · · · · · · · · · · · · · · · ·
Switzerland	1	0.9	1.09	[0.81, 1.45]	2.6%	
Uruguay	1.8	1.2	1.48	[1.17, 1.87]	1.7%	·
USA**	1.4	1.5	0.96	[0.76, 1.21]	31.3%	<b>-</b> _
Wales*	1.6	1.4	1.10	[0.88, 1.39]	1.0%	_ <b>+•</b>
Pooled Effect (High Income)			1.01	[0.88, 1.15]		
Brazil	2	1.9	1.07	[0.87, 1.30]	27.8%	- <b>-</b>
Iran	1.9	1.9	0.98	[0.80, 1.20]	6.8%	<b>-</b> _
Peru**	0.9	0.9	0.99	[0.74, 1.33]	7.3%	<b>_</b>
Pooled Effect (Upper-middle Income)			1.04	[0.90, 1.20]		
Pooled Effect (Overall)			1.02	[0.93, 1.13]		·+-
						0.5 1 1.5 2 2.5 Odds Ratio

Supplementary Figure 8: Association between lockdown and very preterm birth rates in the replication study. Individual and pooled estimates of the association between very preterm birth and lockdown. Observed rates are based on original studies Supplementary Fig., whereas predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. The forest plot on the right shows individual ORs (blue dot) with corresponding Cis (black line). \*Births from 24 weeks onwards \*\*Live births only (Please check original study for further information)

Country	Observed	Predicted	Odds Ratio	95% CI	Weights	
Australia (New South Wales*)	2.5	2.7	0.91	[0.77, 1.09]	2.8%	
Belgium	5.4	6.1	0.89	[0.79, 1.00]	4.3%	
Canada	4.9	5	0.97	[0.86, 1.10]	13.0%	
Chile**	8.7	8.6	1.01	[0.92, 1.11]	11.6%	-
Finland	3.2	3.6	0.88	[0.76, 1.03]	1.1%	
Iceland	4.1	2.7	1.53	[1.31, 1.78]	0.2%	
Norway	4.1	3.7	1.11	[0.96, 1.27]	2.0%	
Scotland	4.4	4	1.09	[0.95, 1.25]	1.7%	
Sweden	4.7	5	0.95	[0.84, 1.07]	4.1%	
USA**	4.5	4.7	0.97	[0.86, 1.11]	32.5%	
Wales*	3.8	3.3	1.15	[0.99, 1.33]	0.7%	-
Pooled Effect (High Income)			0.98	[0.92, 1.04]		
Brazil	9.8	9.6	1.02	[0.93, 1.11]	25.9%	
Pooled Effect (Upper-middle Income)			1.02	[0.93, 1.11]		
Pooled Effect (Overall)			0.99	[0.94, 1.04]		-
						1

#### Spontaneous preterm birth rate in first month of lockdown

Supplementary Figure 9: Association between lockdown and spontaneous preterm birth rates in the replication study. Individual and pooled estimates of the association between spontaneous preterm birth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. The forest plot on the right shows individual ORs (blue dot) with corresponding Cis (black line). \*Births from 24 weeks onwards \*\*Live births only (Please check original study for further information)

Country	Observed	Predicted	Odds Ratio	95% CI	Weights	
Australia (New South Wales*)	2.6	2.8	0.94	[0.79, 1.11]	3.3%	
Belgium	5.6	6	0.93	[0.83, 1.04]	4.8%	
Canada	5	5	1.00	[0.88, 1.13]	14.1%	
Chile**	9	8.6	1.04	[0.95, 1.14]	12.5%	
Finland	3.2	3.7	0.85	[0.73, 0.99]	1.3%	
Iceland	3.1	2.9	1.07	[0.92, 1.26]	0.2%	
Norway	3.2	3.7	0.86	[0.74, 1.00]	1.9%	
Scotland	3.9	3.9	0.99	[0.86, 1.14]	1.7%	
Sweden	4.5	4.9	0.91	[0.80, 1.03]	4.6%	
USA**	4.6	4.7	0.96	[0.85, 1.10]	30.5%	
Wales*	3.4	3	1.13	[0.96, 1.31]	0.7%	
Pooled Effect (High Income)			0.97	[0.91, 1.03]		
Brazil	9.5	9.6	0.99	[0.91, 1.09]	24.4%	
Pooled Effect (Upper-middle Income)			0.99	[0.90, 1.08]		
Pooled Effect (Overall)			0.98	[0.93, 1.03]		
						0.725 0.7750.80.825 0.8750.90.925 0.975 1 1.025 1.0751.11.125 1.1751.21.225 1.2751.31.3 Odds Ratio

#### Spontaneous preterm birth rate in second month of lockdown

Supplementary Figure 10: Association between lockdown and spontaneous preterm birth rates in the replication study. Individual and pooled estimates of the association between spontaneous preterm birth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. The forest plot on the right shows individual ORs (blue dot) with corresponding Cis (black line). \*Births from 24 weeks onwards \*\*Live births only (Please check original study for further information)

Country	Observed	Predicted	Odds Ratio	95% CI	Weights			
Australia (New South Wales*)	2.8	2.9	0.98	[0.84, 1.16]	3.8%			
Belgium	6	6	0.99	[0.89, 1.11]	5.6%		-+	
Canada	5	5	1.00	[0.88, 1.13]	14.4%			
Chile**	8.4	8.7	0.96	[0.87, 1.05]	12.5%		-+-	
Finland	3.1	3.7	0.84	[0.72, 0.97]	1.4%			
Iceland	1.5	3.2	0.46	[0.38, 0.56]	0.1%			
Norway	3.4	3.7	0.91	[0.79, 1.06]	2.3%			
Scotland	5.1	3.8	1.34	[1.17, 1.53]	2.5%			
Sweden	4.6	4.9	0.94	[0.83, 1.07]	5.0%		-+-	
USA**	4.6	4.8	0.96	[0.85, 1.10]	28.5%		-+	
Wales*	3.3	2.9	1.14	[0.97, 1.33]	1.0%		+-•	
Pooled Effect (High Income)			0.98	[0.92, 1.03]				
Brazil	9.8	9.7	1.01	[0.92, 1.10]	22.9%			
Pooled Effect (Upper-middle Income)			1.01	[0.92, 1.10]				
Pooled Effect (Overall)			0.98	[0.94, 1.03]			+	
						0.5	Odds Ratio	1.5

#### Spontaneous preterm birth rate in third month of lockdown

Supplementary Figure 11: Association between lockdown and spontaneous preterm birth rates in the replication study. Individual and pooled estimates of the association between spontaneous preterm birth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. The forest plot on the right shows individual ORs (blue dot) with corresponding Cis (black line). \*Births from 24 weeks onwards \*\*Live births only (Please check original study for further information)

Country	Observed	Predicted	Odds Ratio	95% CI	Weights			
Belgium	6.4	5.9	1.08	[0.96, 1.21]	8.0%		+	
Canada	5.1	5.1	1.00	[0.89, 1.13]	17.5%		_	
Chile**	8.1	8.8	0.93	[0.84, 1.02]	14.5%			
Finland	3.4	3.8	0.90	[0.78, 1.04]	2.4%			
Iceland	1.5	3	0.50	[0.41, 0.61]	0.2%			
Norway	3.6	3.7	0.96	[0.83, 1.11]	3.4%			
Scotland	4.4	3.5	1.27	[1.10, 1.46]	3.4%			
USA**	4.6	4.7	0.98	[0.86, 1.11]	26.7%			
Wales*	4.5	3	1.49	[1.28, 1.72]	1.7%			
Pooled Effect (High Income)			1.00	[0.94, 1.06]				
Brazil	9.4	9.6	0.98	[0.90, 1.08]	22.5%			
Pooled Effect (Upper-middle Income)			0.98	[0.89, 1.07]				
Pooled Effect (Overall)			1.00	[0.95, 1.04]			+	
						0.5	1 1.5 Odde Patio	

#### Spontaneous preterm birth rate in fourth month of lockdown

Supplementary Figure 12: Association between lockdown and spontaneous preterm birth rates in the replication study. Individual and pooled estimates of the association between spontaneous preterm birth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. The forest plot on the right shows individual ORs (blue dot) with corresponding Cis (black line). \*Births from 24 weeks onwards \*\*Live births only (Please check original study for further information)

Country	Observed	Predicted	Odds Ratio	95% CI	Weights		
Australia (New South Wales*)	5.1	4.4	1.16	[1.02, 1.32]	1.7%		
Belgium	5.3	5.5	0.96	[0.85, 1.08]	1.8%		
Canada	8	7.5	1.06	[0.96, 1.17]	7.1%	<b>+•</b>	
Finland	1.6	1.7	0.96	[0.78, 1.19]	0.3%		
Hungary	4.7	4.8	1.00	[0.88, 1.13]	1.9%		
Norway	2.7	3.2	0.85	[0.72, 1.00]	0.5%		
Scotland	4.7	3.6	1.33	[1.16, 1.53]	0.5%		
Sweden	3.3	2.4	1.35	[1.14, 1.59]	1.2%	_ <b></b> -	
Switzerland	4.5	3.1	1.45	[1.25, 1.67]	1.4%		
Wales*	5.6	2.4	2.33	[2.00, 2.71]	1.1%		
Pooled Effect (High Income)			1.15	[1.09, 1.21]			
Brazil	10.2	9.8	1.04	[0.95, 1.14]	69.0%	+	
Iran	8.9	8.3	1.07	[0.97, 1.18]	13.5%		
Pooled Effect (Upper-middle Income)			1.04	[0.97, 1.13]			
Pooled Effect (Overall)			1.06	[1.00, 1.13]		-	
						0.5 1 1.5 Odds Ratio	2 2.5

#### Stillbirth rate in first month of lockdown

Supplementary Figure 13: Association between lockdown and stillbirth rates in the replication study. Individual and pooled estimates of the association between stillbirth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. The forest plot on the right shows individual ORs (blue dot) with corresponding Cis (black line). \*Births from 24 weeks onwards

Country	Observed	Predicted	Odds Ratio	95% CI	Weights	
Australia (New South Wales*)	3.8	4.5	0.85	[0.74, 0.97]	1.4%	-
Belgium	6	5	1.21	[1.08, 1.37]	1.9%	•
Canada	6.6	7.7	0.86	[0.77, 0.95]	6.4%	-
Finland	2.1	1.3	1.68	[1.35, 2.10]	0.4%	
Hungary	3.8	4.9	0.77	[0.68, 0.89]	1.5%	•
Iceland	7.7	0.8	9.24	[7.37, 11.59]	0.3%	
Norway	3.3	2.9	1.13	[0.96, 1.32]	0.6%	+
Scotland	4.9	3.8	1.28	[1.12, 1.47]	0.5%	-
Sweden	2.1	2	1.05	[0.87, 1.28]	0.8%	+
Switzerland	3.1	3.2	0.94	[0.81, 1.10]	1.0%	+
Wales*	3.3	2.4	1.36	[1.16, 1.61]	0.6%	-
Pooled Effect (High Income)			1.00	[0.95, 1.05]		
Brazil	10.5	9.7	1.08	[0.99, 1.18]	70.9%	+
Iran	8.9	8.4	1.07	[0.97, 1.17]	13.6%	+
Pooled Effect (Upper-middle Income)			1.08	[1.00, 1.16]		
Pooled Effect (Overall)			1.07	[1.00, 1.14]		•
						1 2

#### Stillbirth rate in second month of lockdown

Supplementary Figure 14: Association between lockdown and stillbirth rates in the replication study. Individual and pooled estimates of the association between stillbirth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. The forest plot on the right shows individual ORs (blue dot) with corresponding Cis (black line). \*Births from 24 weeks onwards

Country	Observed	Predicted	Odds Ratio	95% CI	Weights	
Australia (New South Wales*)	3.1	4.4	0.71	[0.62, 0.83]	1.3%	-
Belgium	5.3	4.9	1.07	[0.95, 1.21]	1.8%	-
Canada	6	7.9	0.77	[0.69, 0.85]	6.3%	•
Finland	3.1	1.3	2.28	[1.86, 2.79]	0.6%	
Hungary	4.1	5.4	0.77	[0.68, 0.87]	1.8%	•
Iceland	8.3	1	7.91	[6.45, 9.69]	0.2%	
Norway	2.6	2.8	0.94	[0.79, 1.11]	0.6%	-
Scotland	3.8	3.8	1.01	[0.88, 1.17]	0.4%	+
Sweden	3.6	2	1.84	[1.54, 2.19]	1.4%	-
Switzerland	4.4	3.1	1.42	[1.23, 1.65]	1.5%	+
Wales*	3.9	2.2	1.78	[1.51, 2.11]	1.0%	-
Pooled Effect (High Income)			1.02	[0.97, 1.08]		
Brazil	10.4	9.7	1.07	[0.98, 1.17]	69.2%	+
Iran	9	8.2	1.10	[1.00, 1.21]	14.0%	-
Pooled Effect (Upper-middle Income)			1.07	[1.00, 1.16]		
Pooled Effect (Overall)			1.07	[1.00, 1.14]		•
						1 2

#### Stillbirth rate in third month of lockdown

Supplementary Figure 15: Association between lockdown and stillbirth rates in the replication study. Individual and pooled estimates of the association between stillbirth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. The forest plot on the right shows individual ORs (blue dot) with corresponding Cis (black line). \*Births from 24 weeks onwards

Country	Observed	Predicted	Odds Ratio	95% CI	Weights				
Belgium	7	5.3	1.32	[1.17, 1.48]	4.4%				
Canada	6.3	8.2	0.77	[0.69, 0.85]	11.9%	-			
Finland	2.9	1.6	1.74	[1.44, 2.11]	1.3%			-	
Hungary	4.3	5.9	0.72	[0.64, 0.82]	4.0%	-			
Iceland	2.8	1	2.75	[2.29, 3.45]	0.4%				
Norway	2.1	2.8	0.75	[0.63, 0.90]	1.0%				
Scotland	6	3.4	1.76	[1.54, 2.01]	1.3%				
Switzerland	3.8	2.6	1.46	[1.25, 1.71]	2.7%				
Wales*	2.4	1.8	1.34	[1.10, 1.63]	1.1%				
Pooled Effect (High Income)			0.99	[0.94, 1.05]					
Brazil	10.7	9.7	1.10	[1.01, 1.20]	50.3%				
Iran	8.9	8	1.11	[1.00, 1.22]	21.7%		-		
Pooled Effect (Upper-middle Income)			1.10	[1.03, 1.18]					
Pooled Effect (Overall)			1.07	[1.02, 1.13]			-		
						0.5	1 1.5 2 Odde P	2.5 atio	з з

#### Stillbirth rate in fourth month of lockdown

Supplementary Figure 16: Association between lockdown and stillbirth rates in the replication study. Individual and pooled estimates of the association between stillbirth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. The forest plot on the right shows individual ORs (blue dot) with corresponding Cis (black line). \*Births from 24 weeks onwards

Country	Observed	Predicted	Odds Ratio	95% CI	Weights	
Australia (New South Wales*)	5.9	6	0.99	[0.88, 1.11]	2.4%	
Belgium (with ldv)	8.1	8.6	0.94	[0.86, 1.04]	2.6%	
Canada (with ldv)	8.3	9	0.93	[0.85, 1.02]	7.0%	
Chile**	8.7	8.7	1.00	[0.91, 1.10]	6.0%	
Denmark (with Idv)	5.3	6.1	0.87	[0.77, 0.97]	0.3%	•
Finland (with Idv)	5.6	5.5	1.02	[0.91, 1.15]	1.0%	
Hungary (with Idv)	8.4	8.7	0.96	[0.88, 1.06]	1.8%	
Iceland (with Idv)	7.6	7.2	1.06	[0.96, 1.17]	0.1%	
Norway	7	6.8	1.03	[0.92, 1.14]	1.6%	
Scotland	8.2	8.2	0.99	[0.90, 1.09]	1.5%	
Sweden	6	6.1	0.99	[0.88, 1.10]	2.4%	
Switzerland	6.7	6.9	0.97	[0.87, 1.08]	2.1%	
Uruguay (with Idv)	9.4	10.1	0.93	[0.85, 1.02]	1.0%	
USA**	9.8	10.2	0.96	[0.88, 1.05]	38.8%	
Wales** (with Idv)	8.4	8.6	0.98	[0.89, 1.08]	1.2%	
Pooled Effect (High Income)			0.96	[0.92, 1.02]		
Brazil (with Idv)	12.2	12.8	0.96	[0.88, 1.03]	17.9%	
Iran	8.7	9.3	0.93	[0.85, 1.02]	2.3%	
Peru** (with Idv)	5.8	6.5	0.90	[0.80, 1.00]	9.7%	
Pooled Effect (Upper-middle Income)			0.94	[0.88, 1.00]		
Pooled Effect (Overall)			0.96	[0.92, 1.00]		
						0.775 0.8 0.825 0.85 0.875 0.9 0.925 0.95 0.975 1 1.025 1.05 1.075 1.1 1.125 1.15 1.175 Odds Ratio

#### Preterm birth rate in first month of lockdown

Supplementary Figure 17: Association between lockdown and preterm birth rates in the replication study after inclusion of a ldv. Individual and pooled estimates of the association between preterm birth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. Countries, for which a ldv was inserted into the Poisson regression, are marked with (with Idv). The forest plot on the right shows individual ORs (blue dot) with corresponding CIs (black line). \*Births from 24 weeks onwards \*\*Live births only (For further information regarding definitions of preterm birth rates, please check original study)

#### Preterm birth rate in second month of lockdown

Country	Observed	Predicted	Odds Ratio	95% CI	Weights	
Australia (New South Wales*)	6.3	6	1.04	[0.93, 1.17]	5.9%	
Belgium (with Idv)	8.2	8.6	0.95	[0.87, 1.05]	5.8%	
Canada (with ldv)	8.4	9.4	0.90	[0.82, 0.99]	8.7%	
Chile**	9	8.7	1.04	[0.95, 1.14]	8.2%	
Denmark (with Idv)	5.9	5.5	1.07	[0.95, 1.20]	1.3%	
Finland (with Idv)	5.9	6	0.98	[0.95, 1.20]	3.3%	
Hungary (with Idv)	8.7	8.2	1.07	[0.97, 1.18]	4.8%	
Iceland (with Idv)	5.6	7.1	0.79	[0.70, 0.88]	0.4%	
Norway	6.4	6.8	0.95	[0.85, 1.06]	4.2%	
Scotland	7.9	8.2	0.96	[0.87, 1.06]	4.2%	
Sweden	5.5	6	0.93	[0.83, 1.05]	5.5%	
Switzerland	6.3	7	0.91	[0.82, 1.02]	5.2%	
Uruguay (with ldv)	9.4	9.7	0.97	[0.88, 1.06]	3.2%	
USA**	10	10.4	0.96	[0.88, 1.05]	11.1%	
Wales** (with Idv)	7.5	8.5	0.89	[0.80, 0.98]	3.4%	
Pooled Effect (High Income)			0.97	[0.94, 0.99]		
Brazil (with Idv)	12	12	1.00	[0.92, 1.08]	10.2%	
Iran	9.2	9.6	0.96	[0.88, 1.05]	5.7%	
Peru** (with Idv)	6	6.7	0.90	[0.81, 1.01]	8.9%	
Pooled Effect (Upper-middle Income)			0.95	[0.90, 1.01]		
Pooled Effect (Overall)			0.96	[0.94, 0.99]		
						0.7 0.7250.750.775 0.8 0.8250.850.875 0.9 0.9250.950.975 1 1.0251.051.075 1.1 1.1251.151.175 1.2 Odds Ratio

Supplementary Figure 18: Association between lockdown and preterm birth rates in the replication study after inclusion of a ldv. Individual and pooled estimates of the association between preterm birth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. Countries, for which a ldv was inserted into the Poisson regression, are marked with (with ldv). The forest plot on the right shows individual ORs (blue dot) with corresponding CIs (black line). \*Births from 24 weeks onwards \*\*Live births only (For further information regarding definitions of preterm birth rates, please check original study)

Country	Observed	Predicted	Odds Ratio	95% CI	Weights	
Australia (New South Wales*)	6.1	6.1	1.01	[0.90, 1.13]	5.4%	
Belgium (with ldv)	8.4	8.9	0.94	[0.85, 1.03]	5.5%	
Canada (with ldv)	8.4	8.8	0.95	[0.86, 1.04]	9.0%	
Chile**	8.5	8.8	0.96	[0.88, 1.06]	8.2%	
Denmark (with Idv)	6.6	5.6	1.17	[1.04, 1.31]	1.2%	
Finland (with Idv)	5.5	5.8	0.95	[0.85, 1.07]	2.7%	
Hungary (with Idv)	8.6	8.4	1.01	[0.92, 1.12]	4.4%	
Iceland (with Idv)	5.6	7.3	0.76	[0.68, 0.85]	0.3%	
Norway	5.8	6.9	0.85	[0.76, 0.95]	3.6%	
Scotland	9	8	1.11	[1.01, 1.23]	4.1%	
Sweden	5.6	5.8	0.96	[0.86, 1.08]	5.1%	
Switzerland	7.3	6.9	1.06	[0.95, 1.17]	5.0%	
Uruguay (with Idv)	9.8	9.9	0.99	[0.90, 1.08]	2.8%	
USA**	10.1	10.4	0.97	[0.89, 1.05]	12.8%	
Wales** (with Idv)	8	8.3	0.97	[0.88, 1.07]	3.6%	
Pooled Effect (High Income)			0.98	[0.95, 1.01]		
Brazil (with Idv)	12.2	11.8	1.03	[0.95, 1.11]	11.3%	
Iran	9.8	9.9	0.99	[0.91, 1.08]	5.3%	
Peru** (with Idv)	6.3	6.7	0.93	[0.83, 1.04]	9.4%	
Pooled Effect (Upper-middle Income)			0.98	[0.93, 1.04]		
Pooled Effect (Overall)			0.98	[0.95, 1.01]		
						0.675 0.725 0.775 0.825 0.875 0.925 0.975 11.025 1.075 1.125 1.175 1.225 1.275 1.325 Odds Ratio

#### Preterm birth rate in third month of lockdown

Supplementary Figure 19: Association between lockdown and preterm birth rates in the replication study after inclusion of a ldv. Individual and pooled estimates of the association between preterm birth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. Countries, for which a ldv was inserted into the Poisson regression, are marked with (with ldv). The forest plot on the right shows individual ORs (blue dot) with corresponding CIs (black line). \*Births from 24 weeks onwards \*\*Live births only (For further information regarding definitions of preterm birth rates, please check original study)

#### Preterm birth rate in fourth month of lockdown

Country	Observed	Predicted	Odds Ratio	95% CI	Weights	
Belgium (with Idv)	9.1	9	1.01	[0.92, 1.11]	5.5%	
Canada (with ldv)	8.7	8.7	1.00	[0.91, 1.10]	10.8%	
Chile**	8.3	8.8	0.93	[0.85, 1.03]	8.8%	
Denmark (with ldv)	5.5	5.8	0.95	[0.85, 1.07]	0.8%	
Finland (with Idv)	5.4	5.7	0.95	[0.84, 1.07]	2.4%	
Hungary (with Idv)	8.8	8.2	1.08	[0.98, 1.19]	4.2%	
Iceland (with Idv)	5.5	7.1	0.78	[0.70, 0.87]	0.2%	
Norway	6.4	6.9	0.93	[0.84, 1.04]	3.3%	
Scotland	8.3	7.7	1.08	[0.98, 1.19]	3.5%	
Switzerland	7	6.7	1.05	[0.95, 1.17]	4.5%	
Uruguay (with Idv)	9.6	10.1	0.95	[0.87, 1.04]	2.2%	
USA**	10	10.3	0.98	[0.90, 1.07]	19.7%	
Wales** (with Idv)	8.5	8.6	0.99	[0.90, 1.09]	2.8%	
Pooled Effect (High Income)			0.99	[0.95, 1.02]		
Brazil (with ldv)	12	11.9	1.01	[0.93, 1.09]	14.8%	
Iran	10.1	10.1	1.00	[0.91, 1.09]	5.3%	
Peru** (with ldv)	6.2	6.5	0.96	[0.86, 1.07]	10.9%	
Pooled Effect (Upper-middle Income)			0.99	[0.94, 1.05]		
Pooled Effect (Overall)			0.99	[0.96, 1.02]		0.7 0.7250.750.775 0.8 0.8250.850.875 0.9 0.9250.950.975 1 1.025

Supplementary Figure 20: Association between lockdown and preterm birth rates in the replication study after inclusion of a ldv. Individual and pooled estimates of the association between preterm birth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. Countries, for which a ldv was inserted into the Poisson regression, are marked with (with ldv). The forest plot on the right shows individual ORs (blue dot) with corresponding CIs (black line). \*Births from 24 weeks onwards \*\*Live births only (For further information regarding definitions of preterm birth rates, please check original study)

Country	Observed	Predicted	Odds Ratio	95% CI	Weights	
Australia (New South Wales*) (with lag)	1.1	0.9	1.14	[0.86, 1.50]	5.4%	
Belgium	1.3	1.5	0.84	[0.66, 1.06]	3.6%	
Canada	1.3	1.3	1.00	[0.78, 1.27]	6.9%	
Chile**	1.4	1.3	1.03	[0.81, 1.31]	7.3%	<b>-</b>
Denmark (with lag)	1.2	0.7	1.63	[1.22, 2.18]	0.7%	· · · · · · · · · · · · · · · · · · ·
Finland	0.6	0.6	1.02	[0.72, 1.44]	0.7%	
Hungary (with lag)	1.5	1.4	1.03	[0.82, 1.30]	2.9%	
Iceland (with lag)	0.9	0.9	0.97	[0.73, 1.30]	0.2%	
Norway	1.2	1.2	1.02	[0.80, 1.32]	2.9%	<b>_</b>
Scotland	1.4	1	1.39	[1.08, 1.79]	1.8%	·
Sweden	0.9	1	0.91	[0.69, 1.21]	4.1%	
Switzerland	1.4	1.1	1.25	[0.98, 1.61]	4.2%	+
Uruguay (with lag)	1.3	1.2	1.11	[0.86, 1.42]	2.1%	
USA** (with lag)	1.5	1.5	1.02	[0.81, 1.28]	19.3%	<b>_</b>
Wales (with lag)	2.1	1.4	1.44	[1.17, 1.79]	1.8%	
Pooled Effect (High Income)			1.05	[0.96, 1.15]		
Brazil (with lag)	2	1.9	1.02	[0.84, 1.25]	19.1%	<b>_</b>
Iran (with lag)	1.7	1.7	1.01	[0.82, 1.26]	7.8%	
Peru** (with lag)	0.8	0.9	0.95	[0.71, 1.28]	9.0%	
Pooled Effect (Upper-middle Income)			1.00	[0.87, 1.15]		
Pooled Effect (Overall)			1.03	[0.96, 1.11]		<del>_</del>
						0.5 1 1.5 2 Ordris Ratio

#### Very preterm birth rate in first month of lockdown

Supplementary Figure 21: Association between lockdown and very preterm birth rates in the replication study after inclusion of a ldv. Individual and pooled estimates of the association between very preterm birth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. Countries, for which a ldv was inserted into the Poisson regression, are marked with (with ldv). The forest plot on the right shows individual ORs (blue dot) with corresponding CIs (black line). \*Births from 24 weeks onwards \*\*Live births only (For further information regarding definitions of preterm birth rates, please check original study)

#### Very preterm birth rate in second month of lockdown

Country	Observed	Predicted	Odds Ratio	95% CI	Weights	
Australia (New South Wales*) (with lag)	1	0.9	1.10	[0.82, 1.46]	5.7%	
Belgium	1.4	1.5	0.94	[0.75, 1.18]	4.3%	<b>-</b> _
Canada	1.2	1.3	0.92	[0.71, 1.18]	7.2%	
Chile**	1.3	1.4	0.97	[0.76, 1.23]	7.7%	
Denmark (with lag)	1	0.5	2.00	[1.43, 2.80]	0.7%	
Finland	1.1	0.6	1.86	[1.35, 2.57]	1.5%	·
Hungary (with lag)	1.7	1.4	1.18	[0.95, 1.48]	3.8%	+
Iceland (with lag)	0.8	1	0.75	[0.56, 1.01]	0.2%	
Norway	1.3	1.2	1.04	[0.82, 1.34]	3.5%	<del></del>
Scotland	1.5	0.9	1.59	[1.23, 2.06]	2.3%	
Sweden	1.1	0.9	1.16	[0.88, 1.53]	5.4%	<b>+</b> •
Switzerland	1.3	1	1.24	[0.95, 1.61]	4.6%	+
Uruguay (with lag)	1.7	1.4	1.22	[0.97, 1.53]	2.8%	<b>↓</b>
USA** (with lag)	1.5	1.5	1.01	[0.80, 1.26]	16.2%	_ <del></del>
Wales (with lag)	1.3	1.3	1.01	[0.79, 1.28]	1.5%	
Pooled Effect (High Income)			1.08	[0.99, 1.17]		
Brazil (with lag)	2	2.1	0.97	[0.80, 1.18]	15.9%	
Iran (with lag)	1.8	1.7	1.04	[0.56, 1.28]	8.4%	<del>_</del>
Peru** (with lag)	0.7	0.9	0.81	[0.60, 1.11]	8.2%	
Pooled Effect (Upper-middle Income)			0.94	[0.80, 1.11]		
Pooled Effect (Overall)			1.03	[0.96, 1.12]		<del>_</del>
						0.5 1 1.5 2 2.5 Odds Ratio

Supplementary Figure 22: Association between lockdown and very preterm birth rates in the replication study after inclusion of a ldv. Individual and pooled estimates of the association between very preterm birth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. Countries, for which a ldv was inserted into the Poisson regression, are marked with (with ldv). The forest plot on the right shows individual ORs (blue dot) with corresponding CIs (black line). \*Births from 24 weeks onwards \*\*Live births only (For further information regarding definitions of preterm birth rates, please check original study)

Country	Observed	Predicted	Odds Ratio	95% CI	Weights			
Australia (New South Wales*) (with lag)	0.9	0.9	0.96	[0.71, 1.29]	5.5%			
Belgium	1.6	1.5	1.08	[0.86, 1.32]	4.8%			
Canada	1.1	1.2	0.90	[0.69, 1.16]	7.2%			
Chile**	1.3	1.4	0.94	[0.74, 1.19]	7.9%			
Denmark (with lag)	2	0.5	3.60	[2.62, 4.86]	1.2%			·
Finland	0.7	0.5	1.45	[1.01, 2.09]	1.1%			
Hungary (with lag)	1.5	1.4	1.02	[0.81, 1.28]	3.8%	-		
Iceland (with lag)	0.6	1	0.59	[0.43, 0.82]	0.2%			
Norway	1	1.2	0.79	[0.61, 1.04]	3.1%			
Scotland	1.6	0.9	1.88	[1.45, 2.43]	2.5%		-	
Sweden	0.8	0.9	0.90	[0.67, 1.21]	4.6%			
Switzerland	1.2	0.9	1.24	[0.95, 1.63]	4.6%			
Uruguay (with lag)	1.5	1.4	1.04	[0.83, 1.31]	2.7%			
USA** (with lag)	1.5	1.5	1.00	[0.79, 1.25]	16.3%	_ <del></del>		
Wales (with lag)	1.4	1.3	1.05	[0.82, 1.33]	1.9%			
Pooled Effect (High Income)			1.04	[0.95, 1.12]				
Brazil (with lag)	1.9	2.1	0.93	[0.76, 1.33]	15.8%			
Iran (with lag)	1.9	1.8	1.06	[0.87, 1.30]	8.4%			
Peru** (with lag)	0.8	0.9	0.88	[0.65, 1.19]	8.6%			
Pooled Effect (Upper-middle Income)			0.95	[0.80, 1.12]				
Pooled Effect (Overall)			1.01	[0.93, 1.09]		+		
						0.5 1 1.5 2	2.5 3 3.5 Odds Ratio	4 4.5

#### Very preterm birth rate in third month of lockdown

Supplementary Figure 23: Association between lockdown and very preterm birth rates in the replication study after inclusion of a ldv. Individual and pooled estimates of the association between very preterm birth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. Countries, for which a ldv was inserted into the Poisson regression, are marked with (with ldv). The forest plot on the right shows individual ORs (blue dot) with corresponding CIs (black line). \*Births from 24 weeks onwards \*\*Live births only (For further information regarding definitions of preterm birth rates, please check original study)

#### Very preterm birth rate in fourth month of lockdown

Country	Observed	Predicted	Odds Ratio	95% CI	Weights														
Belgium	1.6	1.4	1.09	[0.87, 1.37]	3.0%														
Canada	1.2	1.2	1.01	[0.78, 1.29]	6.1%														
Chile**	1.3	1.3	0.94	[0.74, 1.20]	5.9%														
Denmark (with ldv)	1.4	0.4	3.47	[2.43, 4.94]	0.5%														
Finland	0.8	0.4	1.81	[1.25, 2.61]	0.7%		-												
Hungary (with Idv)	1.2	1.3	0.97	[0.76, 1.24]	2.1%	-													
Iceland (with Idv)	1.2	1.5	0.79	[0.62, 1.01]	0.2%														
Norway	0.9	1.1	0.84	[0.64, 1.11]	1.8%	-+													
Scotland	1.2	0.8	1.47	[1.11, 1.95]	1.2%	<b> </b> →													
Switzerland	1	0.9	1.09	[0.81, 1.45]	2.6%														
Uruguay (with Idv)	1.8	1.4	1.30	[1.04, 1.62]	1.7%														
USA** (with Idv)	1.4	1.5	0.95	[0.75, 1.15]	31.3%	-													
Wales* (with Idv)	1.6	1.4	1.15	[0.91, 1.45]	1.0%	+													
Pooled Effect (High Income)			1.00	[0.89, 1.14]															
Brazil (with Idv)	2	2	0.99	[0.81, 1.20]	27.8%	-													
Iran (with Idv)	1.9	1.9	1.00	[0.82, 1.23]	6.8%	+													
Peru** (with Idv)	0.9	0.9	1.00	[0.75, 1.35]	7.3%	-													
Pooled Effect (Upper-middle Income)			0.99	[0.86, 1.15]															
Pooled Effect (Overall)			1.00	[0.91, 1.10]		+													
						0.5 1 1.5 2		2.5 3 Odds Ratio	2.5 3 3.5 Odds Ratio	2.5 3 3.5 4 Odds Ratio	2.5 3 3.5 4 Odds Ratio	2.5 3 3.5 4 4. Odds Ratio	2.5 3 3.5 4 4.5 Odds Ratio						

Supplementary Figure 24: Association between lockdown and very preterm birth rates in the replication study after inclusion of a ldv. Individual and pooled estimates of the association between very preterm birth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. Countries, for which a ldv was inserted into the Poisson regression, are marked with (with ldv). The forest plot on the right shows individual ORs (blue dot) with corresponding CIs (black line). \*Births from 24 weeks onwards \*\*Live births only (For further information regarding definitions of preterm birth rates, please check original study)

Country	Observed	Predicted	Odds Ratio	95% CI	Weights	
Australia (New South Wales*) (with lag)	2.5	2.8	0.89	[0.75, 1.05]	2.8%	+
Belgium (with lag)	5.4	6.2	0.88	[0.78, 0.98]	4.3%	_ <b>-</b> -
Canada	4.9	5	0.97	[0.86, 1.10]	13.0%	
Chile**	8.7	8.6	1.01	[0.92, 1.11]	11.6%	<b>—</b>
Finland	3.2	3.6	0.88	[0.76, 1.03]	1.1%	
Iceland (with lag)	4.1	3	1.38	[1.19, 1.81]	0.2%	
Norway (with lag)	4.1	3.8	1.06	[0.93, 1.22]	2.0%	<b></b>
Scotland	4.4	4	1.09	[0.95, 1.25]	1.7%	+
Sweden (with lag)	4.7	5.1	0.92	[0.81, 1.04]	4.1%	
USA** (with lag)	4.5	4.8	0.95	[0.84, 1.09]	32.5%	
Wales* (with lag)	3.8	3.8	0.99	[0.86, 1.14]	0.7%	_ <b>-</b>
Pooled Effect (High Income)			0.96	[0.90, 1.02]		
Brazil (with lag)	9.8	10.6	0.93	[0.85, 1.01]	25.9%	
Pooled Effect (Upper-middle Income)			0.93	[0.85, 1.01]		
Pooled Effect (Overall)			0.95	[0.90, 1.00]		
						1 Odde Patio

#### Sontaneous preterm birth rate in first month of lockdown

Supplementary Figure 25: Association between lockdown and spontaneous preterm birth rates in the replication study after inclusion of a ldv. Individual and pooled estimates of the association spontaneous between preterm birth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. Countries, for which a ldv was inserted into the Poisson regression, are marked with (with ldv). The forest plot on the right shows individual ORs (blue dot) with corresponding CIs (black line). \*Births from 24 weeks onwards \*\*Live births only (For further information regarding definitions of preterm birth rates, please check original study)

Sontaneous preterm birth rate in second month of lockdown
---

Country	Observed	Predicted	Odds Ratio	95% CI	Weights	
Australia (New South Wales*) (with lag)	2.6	2.8	0.95	[0.80, 1.13]	3.3%	
Belgium (with lag)	5.6	6.2	0.89	[0.80, 1.00]	4.8%	
Canada	5	5	1.00	[0.88, 1.13]	14.1%	
Chile**	9	8.6	1.04	[0.95, 1.14]	12.5%	
Finland	3.2	3.7	0.85	[0.73, 0.99]	1.3%	
Iceland (with lag)	3.1	2.8	1.10	[0.94, 1.30]	0.2%	
Norway (with lag)	3.2	3.5	0.90	[0.77, 1.05]	1.9%	
Scotland	3.9	3.9	0.99	[0.86, 1.14]	1.7%	
Sweden (with lag)	4.5	5.3	0.85	[0.75, 0.97]	4.6%	
USA** (with lag)	4.6	4.7	0.97	[0.85, 1.10]	30.5%	
Wales* (with lag)	3.4	3.8	0.89	[0.76, 1.02]	0.7%	
Pooled Effect (High Income)			0.97	[0.91, 1.03]		
Brazil (with lag)	9.5	98	0.97	[0.89, 1.06]	24.4%	
Pooled Effect (Upper-middle Income)			0.97	[0.89, 1.06]		
Pooled Effect (Overall)			0.97	[0.92, 1.02]		
						0.725 0.7750.80.825 0.8750.90.925 0.975 11.025 1.0751.11.125 1.1751.21.225 1.2751.3 Odds Ratio

Supplementary Figure 26: Association between lockdown and spontaneous preterm birth rates in the replication study after inclusion of a ldv. Individual and pooled estimates of the association between spontaneous preterm birth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. Countries, for which a ldv was inserted into the Poisson regression, are marked with (with ldv). The forest plot on the right shows individual ORs (blue dot) with corresponding CIs (black line). \*Births from 24 weeks onwards \*\*Live births only (For further information regarding definitions of preterm birth rates, please check original study)

Country	Observed	Predicted	Odds Ratio	95% CI	Weights	
Australia (New South Wales*) (with lag)	2.8	2.7	1.06	[0.90, 1.25]	3.8%	
Belgium (with lag)	6	6.3	0.94	[0.84, 1.06]	5.6%	
Canada	5	5	1.00	[0.88, 1.13]	14.4%	<b>_</b>
Chile**	8.4	8.7	0.96	[0.87, 1.05]	12.5%	
Finland	3.1	3.7	0.84	[0.72, 0.97]	1.4%	
Iceland (with lag)	1.5	2.3	0.63	[0.51, 0.78]	0.1%	
Norway (with lag)	3.4	3.9	0.87	[0.75, 1.01]	2.3%	
Scotland	5.1	3.8	1.34	[1.17, 1.53]	2.5%	
Sweden (with lag)	4.6	5.1	0.90	[0.79, 1.02]	5.0%	
USA** (with lag)	4.6	4.7	0.98	[0.87, 1.12]	28.5%	
Wales* (with lag)	3.3	3.5	0.96	[0.82, 1.11]	1.0%	
Pooled Effect (High Income)			0.98	[0.92, 1.04]		
Brazil (with lag)	9.8	9.6	1.02	[0.93, 1.06]	22.9%	_ <b>-</b>
Pooled Effect (Upper-middle Income)			1.02	[0.96, 1.09]		
Pooled Effect (Overall)			0.99	[0.94, 1.03]		<del>_</del>
						0.5 1 1.5 Odds Ratio

#### Sontaneous preterm birth rate in third month of lockdown

Supplementary Figure 27: Association between lockdown and spontaneous preterm birth rates in the replication study after inclusion of a ldv. Individual and pooled estimates of the association between spontaneous preterm birth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. Countries, for which a ldv was inserted into the Poisson regression, are marked with (with ldv). The forest plot on the right shows individual ORs (blue dot) with corresponding CIs (black line). \*Births from 24 weeks onwards \*\*Live births only (For further information regarding definitions of preterm birth rates, please check original study)

#### Spontaneous preterm birth rate in fourth month of lockdown

Country	Observed	Predicted	Odds Ratio	95% CI	Weights			
Belgium (with lag)	6.4	6.3	1.02	[0.91, 1.14]	8.0%			
Canada	5.1	5.1	1.00	[0.89, 1.13]	17.5%			
Chile**	8.1	8.8	0.93	[0.84, 1.02]	14.5%			
Finland	3.4	3.8	0.90	[0.78, 1.04]	2.4%			
Iceland (with lag)	1.5	2.9	0.51	[0.42, 0.63]	0.2%			
Norway (with lag)	3.6	3.9	0.94	[0.81, 1.08]	3.4%			
Scotland	4.4	3.5	1.27	[1.10, 1.46]	3.4%			
USA** (with lag)	4.6	4.6	1.00	[0.88, 1.14]	26.7%			
Wales* (with lag)	4.5	3.5	1.27	[1.11, 1.46]	1.7%			
Pooled Effect (High Income)			1.00	[0.94, 1.06]				
Brazil (with lag)	9.4	9.6	0.98	[0.90, 1.08]	22.5%			
Pooled Effect (Upper-middle Income)			0.98	[0.89, 1.07]				
Pooled Effect (Overall)			0.99	[0.95, 1.04]			-	
						0.5	Odds Ratio	1.5

Supplementary Figure 28: Association between lockdown and spontaneous preterm birth rates in the replication study after inclusion of a ldv. Individual and pooled estimates of the association between spontaneous preterm birth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. Countries, for which a ldv was inserted into the Poisson regression, are marked with (with ldv). The forest plot on the right shows individual ORs (blue dot) with corresponding CIs (black line). \*Births from 24 weeks onwards \*\*Live births only (For further information regarding definitions of preterm birth rates, please check original study)

Country	Observed	Predicted	Odds Ratio	95% CI	Weights	
Australia (New South Wales*) (with lag)	5.1	3.6	1.45	[1.26, 1.66]	1.7%	_ <b></b> -
Belgium (with lag)	5.3	6.3	0.85	[0.75, 0.95]	1.8%	
Canada	8	7.5	1.06	[0.96, 1.17]	7.1%	- <b>-</b>
Finland (with lag)	1.6	2.2	0.74	[0.61, 0.91]	0.3%	
Hungary	4.7	4.8	1.00	[0.88, 1.13]	1.9%	
Norway (with lag)	2.7	3.3	0.82	[0.70, 0.97]	0.5%	
Scotland	4.7	3.6	1.33	[1.16, 1.53]	0.5%	
Sweden (with lag)	3.3	2.8	1.18	[1.00, 1.38]	1.2%	
Switzerland (with lag)	4.5	3.7	1.23	[1.07, 1.41]	1.4%	
Wales*	5.6	2.4	2.33	[2.00, 2.71]	1.1%	
Pooled Effect (High Income)			1.13	[1.07, 1.19]		
Brazil (with lag)	10.2	9.9	1.03	[0.94, 1.13]	69.0%	-
Iran	8.9	8.3	1.07	[0.97, 1.18]	13.5%	
Pooled Effect (Upper-middle Income)			1.04	[0.96, 1.12]		
Pooled Effect (Overall)			1.05	[0.99, 1.12]		
						0.5 1 1.5 2 2.5 Odda Datia

#### Stillbirth rate in first month of lockdown

Supplementary Figure 29: Association between lockdown and stillbirth rates in the replication study after inclusion of a ldv. Individual and pooled estimates of the association between stillbirth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. Countries, for which a ldv was inserted into the Poisson regression, are marked with (with ldv). The forest plot on the right shows individual ORs (blue dot) with corresponding Cls (black line). \*Births from 24 weeks onwards \*\*Live births only (For further information regarding definitions of preterm birth rates, please check original study)

#### Stillbirth rate in second month of lockdown

Country	Observed	Predicted	Odds Ratio	95% CI	Weights	
Australia (New South Wales*) (with lag)	3.8	3.7	1.04	[0.90, 1.20]	1.4%	+
Belgium (with lag)	6	6.9	0.87	[0.78, 0.97]	1.9%	-
Canada	6.6	7.7	0.86	[0.77, 0.95]	6.4%	-
Finland (with lag)	2.1	2.3	0.92	[0.77, 1.11]	0.4%	-
Hungary	3.8	4.9	0.77	[0.68, 0.89]	1.5%	•
Iceland	7.7	1	7.66	[6.22, 9.44]	0.3%	
Norway (with lag)	3.3	3	1.09	[0.93, 1.28]	0.6%	+
Scotland	4.9	3.8	1.28	[1.12, 1.47]	0.5%	-
Sweden (with lag)	2.1	2.5	0.86	[0.72, 1.03]	0.8%	-
Switzerland (with lag)	3.1	3.5	0.89	[0.76, 1.03]	1.0%	•
Wales*	3.3	2.4	1.36	[1.16, 1.61]	0.6%	
Pooled Effect (High Income)			0.95	[0.90, 1.00]		
Brazil (with lag)	10.5	10.2	1.03	[0.94, 1.16]	70.9%	÷
Iran	8.9	8.4	1.07	[0.97, 1.17]	13.6%	+
Pooled Effect (Upper-middle Income)			1.04	[0.95, 1.13]		
Pooled Effect (Overall)			1.02	[0.95, 1.10]		ł
						t

Supplementary Figure 30: Association between lockdown and stillbirth rates in the replication study after inclusion of a ldv. Individual and pooled estimates of the association between stillbirth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. Countries, for which a ldv was inserted into the Poisson regression, are marked with (with ldv). The forest plot on the right shows individual ORs (blue dot) with corresponding CIs (black line). \*Births from 24 weeks onwards \*\*Live births only (For further information regarding definitions of preterm birth rates, please check original study)

Country	Observed	Predicted	Odds Ratio	95% CI	Weights	
Australia (New South Wales*) (with lag)	3.1	3.5	0.88	[0.76, 1.03]	1.3%	1
Belgium (with lag)	5.3	6.7	0.78	[0.70, 0.88]	1.8%	•
Canada	6	7.9	0.77	[0.69, 0.85]	6.3%	-
Finland (with lag)	3.1	2.7	1.15	[0.98, 1.36]	0.6%	-
Hungary	4.1	5.4	0.77	[0.68, 0.87]	1.8%	-
Iceland	8.3	0.5	17.73	[13.20, 23.81]	0.2%	
Norway (with lag)	2.6	3.2	0.84	[0.71, 0.98]	0.6%	
Scotland	3.8	3.8	1.01	[0.88, 1.17]	0.4%	
Sweden (with lag)	3.6	2.2	1.64	[1.39, 1.94]	1.4%	•
Switzerland (with lag)	4.4	3.4	1.31	[1.14, 1.51]	1.5%	•
Wales*	3.9	2.2	1.78	[1.51, 2.11]	1.0%	-
Pooled Effect (High Income)			0.97	[0.92, 1.02]		
Brazil (with lag)	10.4	10.1	1.03	[0.94, 1.12]	69.2%	ł
Iran	9	8.2	1.10	[1.00, 1.21]	14.0%	•
Pooled Effect (Upper-middle Income)			1.04	[0.97, 1.12]		
Pooled Effect (Overall)			1.03	[0.97, 1.10]		
						0 2.5

#### Stillbirth rate in third month of lockdown

Supplementary Figure 31: Association between lockdown and stillbirth rates in the replication study after inclusion of a ldv. Individual and pooled estimates of the association between stillbirth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. Countries, for which a ldv was inserted into the Poisson regression, are marked with (with ldv). The forest plot on the right shows individual ORs (blue dot) with corresponding CIs (black line). \*Births from 24 weeks onwards \*\*Live births only (For further information regarding definitions of preterm birth rates, please check original study)

#### Stillbirth rate in fourth month of lockdown

ountry	Observed	Predicted	Odds Ratio	95% CI	Weights	
Belgium (with lag)	7	6.7	1.05	[0.95, 1.17]	4.4%	-
Canada	6.3	8.2	0.77	[0.69, 0.85]	11.9%	-
Finland (with lag)	2.9	1.8	1.61	[1.33, 1.94]	1.3%	
Hungary	4.3	5.9	0.72	[0.64, 0.82]	4.0%	-
Iceland (with lag)	2.8	0.5	5.08	[3.80, 6.79]	0.4%	
Norway (with lag)	2.1	3.9	0.53	[0.45, 0.62]	1.0%	-
Scotland	6	3.4	1.76	[1.54, 2.01]	1.3%	-
Switzerland (with lag)	3.8	3.5	1.11	[0.96, 1.28]	2.7%	•-
Wales*	2.4	1.8	1.34	[1.10, 1.63]	1.1%	
Pooled Effect (High Income)			0.92	[0.87, 0.97]		
Brazil (with lag)	10.7	10.4	1.03	[0.94, 1.12]	50.3%	+
Iran	8.9	8	1.11	[1.00, 1.22]	21.7%	+
Pooled Effect (Upper-middle Income)			1.05	[0.98, 1.13]		
Pooled Effect (Overall)			1.02	[0.96, 1.07]		-
						0 1

Supplementary Figure 32: Association between lockdown and stillbirth rates in the replication study after inclusion of a ldv. Individual and pooled estimates of the association between stillbirth and lockdown. Observed rates are based on the original studies Fig. 1, whereas Predicted rates were obtained using ITS-analysis. Individual Odds were calculated by comparing observed and predicted rates. Pooled effect estimates were calculated using meta-analysis with weights from the original study. Countries, for which a ldv was inserted into the Poisson regression, are marked with (with ldv). The forest plot on the right shows individual ORs (blue dot) with corresponding CIs (black line). \*Births from 24 weeks onwards \*\*Live births only (For further information regarding definitions of preterm birth rates, please check original study)





Supplementary Figure 33: Autocorrelation and partial autocorrelation plots, before and after inclusion of a ldv for <u>preterm birth rates</u>. Acf and pacf plots before ldv were created after adjusting for seasonality. Blue dotted line resembles significance level.





Supplementary Figure 34: Autocorrelation and partial autocorrelation plots, before and after inclusion of a ldv for <u>very preterm birth rates</u>. Acf and pacf plots before ldv were created after adjusting for seasonality. Blue dotted line resembles significance level.



Supplementary Figure 35: Autocorrelation and partial autocorrelation plots, before and after inclusion of a ldv for <u>spontaneous preterm birth rates</u>. Acf and pacf plots before ldv were created after adjusting for seasonality. Blue dotted line resembles significance level.



Supplementary Figure 36: Autocorrelation and partial autocorrelation plots, before and after inclusion of a ldv for <u>stillbirth rates</u>. Acf and pacf plots before ldv were created after adjusting for seasonality. Blue dotted line resembles significance level.

#### **Supplementary Tables**

Supplementary Table 1 – Comparison of pooled estimated effects gained through meta-analysis in preterm birth rates

Month following Lockdown	Original Study (1)	Robustness Reproduction (2)	Robustness Reproduction (with Idv) (3)
Month 1			
High-Income	OR 0.96	OR 0.97	OR 0.96
	95% CI 0.94-0.98	95% CI 0.92-1.02	95% CI 0.92-1.02
		P=0.175	P=0.168
Upper-middle Income	OR 0.96	OR 0.93	OR 0.94
	95% CI 0.90-1.02	95% CI 0.87-0.98	95% CI 0.88-1.00
		P=0.013	P=0.035
Total	OR 0.96	OR 0.95	OR 0.96
	95% CI 0.95-0.98	95% CI 0.92-0.99	95% CI 0.92-1.00
	P<0.001	P=0.020	P=0.029
Month 2			
High-Income	OR 0.97	OR 0.97	OR 0.97
	95% CI 0.95-0.99	95% CI 0.94-0.99	95% CI 0.94-0.99
		P=0.016	P=0.017
Upper-middle Income	OR 0.93	OR 0.96	OR 0.98
	95% CI 0.79-1.08	95% CI 0.91-1.01	95% CI 0.90-1.01
		P=0.126	P=0.095
Total	OR 0.96	OR 0.96	OR 0.96
	95% CI 0.92-0.99	95% CI 0.94-0.99	95% CI 0.94-0.99
	P=0.03	P=0.004	P=0.004
Month 3			
High-Income	OR 0.97	OR 0.98	OR 0.98
-	95% CI 0.95-0.99	95% CI 0.95-1.01	95% CI 0.98-1.01
		P=0.161	P=0.123
Upper-middle Income	OR 0.95	OR 0.99	OR 0.98
	95% CI 0.83-1.09	95% CI 0.94-1.04	95% CI 0.93-1.04
		P=0.676	P=0.587

Total	OR 0.97	OR 0.98	OR 0.98
	95% CI 0.94-1.00	95% CI 0.96-1.01	95% CI 0.95-1.01
	P=0.09	P=0.162	P=0.114
Month 4			
High-Income	OR 0.98	OR 0.99	OR 0.99
	95% CI 0.96-1.01	95% CI 0.96-1.02	95% CI 0.95-1.02
		P=0.551	P=0.538
Upper-middle Income	OR 0.98	OR 0.97	OR 0.99
	95% CI 0.90-1.07	95% CI 0.92-1.03	95% CI 0.94-1.05
		P=0.323	P=0.737
Total	OR 0.99	OR 0.98	OR 0.99
	95% CI 0.96-1.01	95% CI 0.98-1.01	95% CI 0.96-1.02
	P=0.34	P=0.286	P=0.483

Note: P-Values of the original study could not be found for every subgroup effect size estimate, please check Calvert et. al p.530ff. for details. P-Values in the reproduction were calculated using Wald-test

# Supplementary Table 2 – Comparison of pooled estimated effects gained through meta-analysis in very preterm birth rates

Month	Original Study (1)	Recreate Reproduction (2)	Robustness Reproduction (with Idv) (3)
1			
High-Income	OR 1.02 95% CI 0.94-1.10	OR 1.05 95% CI 0.96-1.15 P=0.301	OR 1.05 95% CI 0.96-1.15 P=0.273
Upper-middle Income	OR 1.00 95% CI 0.93-1.06	OR 0.99 95% CI 0.86-1.13 P=0.84	OR 1.00 95% CI 0.87-1.15 P=0.998
Total	OR 1.00 95% CI 0.95-1.06	OR 1.03 95% CI 0.95-1.11	OR 1.03 95% CI 0.96-1.11

		P=0.511	P=0.408
2			
High-Income	OR 1.05	OR 1.08	OR 1.08
8	95% CI 0.98-1.13	95% CI 0.99-1.17	95% CI 0.99-1.17
		P=0.078	P=0.0.066
Upper-middle Income	OR 0.95	OR 0.97	OR 0.94
	95% CI 0.80-1.11	95% CI 0.84-1.11	95% CI 0.80-1.11
		P=0.619	P=0.485
Total	OR 1.02	OR 1.04	OR 1.03
	95% CI 0.96-1.09	95% CI 0.97-1.12	95% CI 0.96-1.12
		P=0.287	P=0.400
3			
High-Income	OR 1.02	OR 1.04	OR 1.04
	95% CI 0.93-1.12	95% CI 0.96-1.12	95% CI 0.95-1.12
		P=0.389	P=0.405
Upper-middle Income	OR 0.99	OR 0.99	OR 0.95
	95% CI 0.93-1.05	95% CI 0.86-1.13	95% CI 0.93-1.12
		P=0.829	P=0.524
Total	OR 1.00	OR 1.02	OR 1.01
	95% CI 0.94-1.06	95% CI 0.95-1.09	95% CI 0.93-1.09
		P=0.595	P=0.878
4			
High-Income	OR 0.98	OR 1.01	OR 1.00
8	95% CI 0.94-1.10	95% CI 0.88-1.15	95% CI 0.89-1.14
		P=0.912	P=0.949
Upper-middle Income	OR 1.06	OR 1.04	OR 0.99
	95% CI 1.01-1.12	95% CI 0.90-1.20	95% CI 0.86-1.15
		P=0.594	P=0.928
Total	OR 1.02	OR 1.02	OR 1.00
	95% CI 0.97-1.06	95% CI 0.93-1.13	95% CI 0.91-1.10
		P=0.675	P=0.992

Note: There were no P-Values given for very preterm birth effect size estimates in the original work. P-Values in the reproduction were calculated using Wald-test

Supplementary Table 3 – Comparison of pooled estimated effects gained through meta-analysis in spontaneou
preterm birth rates

Month	Original Study (1)	Recreate Reproduction (2)	Robustness Reproduction (with ldv) (3)
1			
High-Income	OR 0.96	OR 0.98	OR 0.96
-	95% CI 0.93-0.99	95% CI 0.92-1.04	95% CI 0.90-1.02
		P=0.438	P=0.218
Upper-middle Income	OR 0.99	OR 1.02	OR 0.93
	95% CI 0.95-1.03	95% CI 0.93-1.11	95% CI 0.85-1.01
		P=0.661	P=0.099
Total	OR 0.97	OR 0.99	OR 0.95
	95% CI 0.94-0.99	95% CI 0.94-1.04	95% CI 0.90-1.00
		P=0.615	P=0.069
2			
High-Income	OR 0.97	OR 0.97	OR 0.97
	95% CI 0.94-0.99	95% CI 0.91-1.03	95% CI 0.91-1.03
		P=0.356	P=0.304
Upper-middle Income	OR 1.02	OR 0.99	OR 0.97
	95% CI 0.98-1.07	95% CI 0.90-1.08	95% CI 0.89-1.06
		P=0.827	P=0.494
Total	OR 0.98	OR 0.98	OR 0.97
	95% CI 0.95-1.01	95% CI 0.93-1.03	95% CI 0.92-1.02
		P=0.354	P=0.222
3			
High-Income	OR 0.97	OR 0.98	OR 0.98
	95% CI 0.95-1.00	95% CI 0.92-1.03	95% CI 0.92-1.04
		P=0.427	P=0.497
Upper-middle Income	OR 1.03	OR 1.01	OR 1.02
	95% CI 0.98-1.07	95% CI 0.92-1.10	95% CI 0.96-1.09

P=0.553

P=0.827

Total	OR 0.98	OR 0.98	OR 0.99
	95% CI 0.96-1.02	95% CI 0.94-1.03	95% CI 0.94-1.03
		P=0.529	P=0.604
4			
High-Income	OR 0.97	OR 1.00	OR 1.00
	95% CI 0.93-1.02	95% CI 0.94-1.06	95% CI 0.94-1.06
		P=0.994	P=0.904
Upper-middle Income	OR 1.01	OR 0.98	OR 0.98
	95% CI 0.97-1.06	95% CI 0.89-1.07	95% CI 0.89-1.07
		P=0.664	P=0.664
Total	OR 0.98	OR 1.00	OR 0.99
	95% CI 0.95-1.02	95% CI 0.95-1.04	95% CI 0.95-1.04
		P=0.849	P=0.771

Note: There were no P-Values given for spontaneous preterm birth effect size estimates in the original work. P-Values in the reproduction were calculated using Wald-test

# Supplementary Table 4 – Comparison of pooled estimated effects gained through meta-analysis in stillbirth rates

Month	Original Study (1)	Recreate Reproduction (2)	Robustness Reproduction (with ldv) (3)
1			
High-Income	OR 1.14	OR 1.15	OR 1.13
	95% CI 1.02-1.29	95% CI 1.09-1.21	95% CI 1.07-1.19
	P=0.02	P<0.001	P<0.001
Upper-middle Income	OR 1.02	OR 1.04	OR 1.04
	95% CI 0.97-1.08	95% CI 0.97-1.13	95% CI 0.96-1.12
		P=0.269	P=0.372
Total	OR 1.04	OR 1.06	OR 1.05

	95% CI 0.99-1.09	95% CI 1.00-1.13	95% CI 0.99-1.12
	P=0.10	P=0.067	P=0.128
2			
High-Income	OR 1.00	OR 1.00	OR 0.95
	95% CI 0.88-1.12	95% CI 0.95-1.05	95% CI 0.90-1.00
	P=0.98	P=0.98	P=0.035
Upper-middle Income	OR 1.08 95% CI 1.02-1.14	OR 1.08 95% CI 1.00-1.16 P=0.049	OR 1.04 95% CI 0.95-1.13 P=0.434
Total	OR 1.07	OR 1.07	OR 1.02
	95% CI 1.02-1.12	95% CI 1.00-1.14	95% CI 0.96-1.10
	P=0.001	P=0.051	P=0.579
3			
High-Income	OR 0.99	OR 1.02	OR 0.97
	95% CI 0.88-1.12	95% CI 0.97-1.08	95% CI 0.92-1.02
	P=0.89	P=0.37	P=0.248
Upper-middle Income	OR 1.10 95% CI 1.04-1.16	OR 1.07 95% CI 1.00-1.16 P=0.06	OR 1.04 95% CI 0.97-1.12 P=0.286
Total	OR 1.08	OR 1.07	OR 1.03
	95% CI 1.02-1.13	95% CI 1.00-1.14	95% CI 0.97-1.10
	P=0.004	P=0.047	P=0.367
4			
High-Income	OR 1.01	OR 1.02	OR 0.92
	95% CI 0.87-1.18	95% CI 0.97-1.08	95% CI 0.87-0.97
	P=0.86	P=0.717	P=0.004
Upper-middle Income	OR 1.10 95% CI 1.04-1.17	OR 1.07 95% CI 1.00-1.16 P=0.004	OR 1.05 95% CI 0.98-1.13 P=0.134
Total	OR 1.07	OR 1.07	OR 1.02
	95% CI 1.00-1.15	95% CI 1.00-1.14	95% CI 0.96-1.07
	P=0.07	P=0.009	P=0.563

Note: P-Values of the original study could not be found for every subgroup effect size estimate, please check Calvert et. al p.530ff. for details. P-Values in the reproduction were calculated using Wald-test

# Supplementary Table 5 – Comparison of pooled estimated effects gained through meta-analysis in preterm birth rates between primary reproduction analysis and sensitivity analysis excluding Brazil an USA

Month following Lockdown	Primary reproduction analysis (including all countries)	Sensitivity analysis (excluding Brazil and USA)
Month 1		
High-Income	OR 0.97	OR 0.97
	95% CI 0.92-1.02	95% CI 0.94-1.01
Upper-middle Income	OR 0.93	OR 0.88
	95% CI 0.87-0.98	95% CI 0.81-0.96
Total	OR 0.95	OR 0.95
	95% CI 0.92-0.99	95% CI 0.91-0.98
Month 2		
High-Income	OR 0.97	OR 0.97
	95% CI 0.94-0.99	95% CI 0.94-1.00
Upper-middle Income	OR 0.96	OR 0.93
	95% CI 0.91-1.01	95% CI 0.86-1.00
Total	OR 0.96	OR 0.96
	95% CI 0.94-0.99	95% CI 0.93-0.99
Month 3		
High-Income	OR 0.98	OR 0.98
	95% CI 0.95-1.01	95% CI 0.95-1.01

Upper-middle Income	OR 0.99	OR 0.96
	95% CI 0.94-1.04	95% CI 0.89-1.03
Total	OR 0.98 95% CI 0.96-1.01	OR 0.98 95% CI 0.95-1.00
Month 4		
High-Income	OR 0.99 95% CI 0.96-1.02	OR 0.99 95% CI 0.96-1.03
Upper-middle Income	OR 0.97 95% CI 0.92-1.03	OR 0.94 95% CI 0.87-1.02
Total	OR 0.98 95% CI 0.98-1.01	OR 0.98 95% CI 0.95-1.01

	1 Month	2 Month	3 Month	1 Month
Preterm birth	1.19101101	2.10101101	5.1vi0litii	4.10101111
High Income	OR 0.97	OR 0.98	OR 0.98	OR 0.98
	95% CI 0.93-1.02	95% CI 0.95-1.00	95% CI 0.95-1.01	95% CI 0.95-1.02
Upper-middle	OR 0.92	OR 0.95	OR 0.99	OR 0.99
	95% CI 0.87-0.98	95% CI 0.90-1.01	95% CI 0.94-1.05	95% CI 0.94-1.04
Total	OR 0.96	OR 0.97	OR 0.98	OR 0.98
	95% CI 0.92-1.00	95% CI 0.95-0.99	95% CI 0.96-1.01	95% CI 0.96-1.01
Very preterm birth				
High Income	OR 1.05 95% CI 0.96-1.14	OR 1.08 95% CI 0.98-1.17	OR 1.04 95% CI 0.96-1.12	OR 0.99 95% CI: 0.87– 1.13
Upper-middle	OR 0.97	OR 0.96	OR 0.98	OR 1.03
	95% CI 0.85-1.11	95% CI 0.84-1.10	95% CI 0.86-1.13	95% CI: 0.90-1.20
Total	OR 1.02	OR 1.04	OR 1.02	OR 1.01
	95% CI 0.95-1.10	95% CI 0.97-1.11	95% CI 0.95-1.09	95% CI: 0.92-1.11
Spontaneous preterm birth				
High Income	OR 0.98	OR 0.98	OR 0.98	OR 0.99
	95% CI 0.92-1.04	95% CI 0.92-1.04	95% CI 0.93-1.04	95% CI 0.94-1.05
Upper-middle	OR 0.93	OR 0.97	OR 1.02	OR 0.98
	95% CI 0.85-1.01	95% CI 0.89-1.06	95% CI 0.93-1.11	95% CI 0.89-1.07
Total	OR 0.97	OR 0.98	OR 0.99	OR 1.01
	95% CI 0.92-1.02	95% CI 0.93-1.03	95% CI 0.93-1.04	95% CI 0.94-1.04
Stillbirth				
High Income	OR 1.11	OR 0.99	OR 1.02	OR 1.01
	95% CI 1.05-1.17	95% CI 0.93-1.04	95% CI 0.97-1.07	95% CI 0.98-1.09
Upper-middle	OR 1.04	OR 1.07	OR 1.07	OR 1.09
	95% CI 0.96-1.12	95% CI 0.99-1.15	95% CI 1.00-1.15	95% CI 1.02-1.16
Total	OR 1.05	OR 1.06	OR 1.06	OR 1.07
	95% CI 0.98-1.12	95% CI 0.99-1.13	95% CI 1.00-1.13	95% CI 1.02-1.13

### Supplementary Table 6 – Pooled effect estimates after including ldv $Y_{t-1}$