CHARLES UNIVERSITY
FACULTY OF SOCIAL SCIENCES
Institute of Economic Studies

The Effect of Face Masks on Covid Transmission: A Meta-Analysis

Master's thesis

Author: Bc. Martina Lušková
Study program: Economics and Finance
Supervisor: doc. PhDr. Zuzana Havránková, Ph.D.
Year of defense: 2023

Declaration of Authorship

The author hereby declares that he or she compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain any other academic title.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis in whole or in part and agrees with the thesis being used for study and scientific purposes.

Prague, May 2, 2023

Martina Lušková

Abstract

The effect of face masks on Covid-19 transmission is crucial for the health of populations. Nevertheless, its economic consequences cannot be overlooked. To perform a quantitative meta-analysis, 258 estimates from 44 primary studies were collected together with more than 30 variables mirroring the differences among the studies. Publication bias was examined by implementing various statistical tests resulting in mild evidence for the phenomenon. We contribute to other meta-analyses on the topic by employing the Bayesian and Frequentist model averaging to identify the drivers behind the heterogeneity of the estimates. The results suggest that temperature, geographical latitude, and panel data structure have a highly statistically significant and positive effect on the risk of transmission associated with mask-wearing. Moreover, a positive effect was identified for healthcare set-up. In contrast, performing an aerosol-generating procedure shifts the risk in the negative direction.

JEL Classification	I1, I11, I19,
Keywords	meta-analysis, Covid-19, face masks, pandemic, Covid-19 transmission, publication bias, Bayesian model averaging
Title	The Effect of Face Masks on Covid Transmission: A Meta-Analysis

Abstrakt

Vplyv rúšok na prenos Covidu-19 je kľúčový pre zdravie populácií, napriek tomu nemožno prehliadať jeho ekonomické dôsledky. Na vykonanie kvantitatívnej meta-analýzy bolo zhromaždených 258 odhadov z 44 primárnych štúdií spolu s viac ako 30 premennými reprezentujúcimi rozdiely medzi štúdiami. Publikačné skreslenie bolo skúmané implementáciou rôznych štatistických testov. Výsledky vykonaných testov dokazujú miernu prítomnosť tohto fenoménu. Pridaná hodnota našej práce leží v použití Bayesovského a Frekventistického priemerovania modelov na identifikáciu faktorov zodpovedných za heterogenitu odhadov. Výsledky naznačujú, že teplota, geografická šírka a panelová štruktúra dát majú štatisticky významne pozitívny vplyv na riziko prenosu spojené s nosením rúšok. Navyše, pozitívny vzťah bol identifikovaný aj pre zdravotnícke prostredie. Naopak, zdravotnícke výkony, pri ktorých dochádza k vzniku aerosólu ovplyvňujú riziko negatívne.

Klasifikace JEL	I1, I11, I19,
Klíčová slova	meta-analýza, Covid-19, rúška, pandémia, prenos Covid-19, publikačná selektivita, Bayesovské modelové priemerovanie
Název práce	Vplyv rúšok na prenos Covidu: Meta-Analýza

Acknowledgments

The author is grateful especially to doc. PhDr. Zuzana Havránková, Ph.D.and prof. doc. PhDr. Tomáš Havránek, Ph.D. for their professional insights, valuable guidance, advice, and support, which they have provided throughout writing this thesis. Without them, the thesis would not be completed.

This thesis is part of a project that has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 870245.

This output was supported by the NPO "Systemic Risk Institute" number LX22NPO5101, funded by European Union - Next Generation EU (Ministry of Education, Youth and Sports, NPO: EXCELES)

Typeset in LaTeXusing the IES Thesis Template.

Bibliographic Record

Lušková, Martina: *The Effect of Face Masks on Covid Transmission: A Meta-Analysis*. Master's thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies, Prague. 2023, pages 90. Advisor: doc. PhDr. Zuzana Havránková, Ph.D.

Contents

List of Tables

List of Figures

Acronyms

AGP	Aerosol Generating Procedures
BE	Between Effects
BMA	Bayesian Model Averageing
EMA	European Medicines Agency
EU	European Union
EUL	Emergency Use Listing
FAT	Funnel Asymmetry Test
FE	Fixed Effects
FMA	Frequentist Model Averageing
HR	Hazard Ratio
IV	Instrumental Variable
JCR	Journal Citation Reports
MERS-CoV	Middle East Respiratory Syndrome Coronavirus
OLS	Ordinary Least Squares
OR	Odds Ratio
PET	Precision Effect Test
PIP	Posterior Inclusion probability
PMP	Posterior Model Probability
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RR	Relative Risk or Risk Ratio
RVIs	Respiratory Viral Illnesses
SARS-CoV-1	Severe Acute Respiratory Syndrome Coronavirus 1
SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2
VIF	Variance Inflation Factor
WAAP	Weighted Average of the Adequately Powered
WHO	World Health Organisation
PPE	Personal Protective Equipment

Master's Thesis Proposal

Author	Bc. Martina Lušková
Supervisor	doc. PhDr. Zuzana Havránková, Ph.D.
Proposed topic	The Effect of Face Masks on Covid Transmission: A Meta-Analysis

Motivation As Covid-19 disease began to spread rapidly at the beginning of 2020 affecting the health and wellbeing of the population all around the world, some protective measures were taken. Considering transmission channels (droplet and airborne particles infected by virus penetrate the body mainly via the respiratory system) of the disease, one of the protective measures was the usage of face masks to prevent the spread of the disease.

The aim of my thesis is to assess the papers published on the relationship between face masks and the spread of Covid. So far, several studies investigated the mentioned relationship using different techniques such as mathematical modelling of Covid infection in a population (Eikenberry et al., 2020), logistic regression and others.

Even though several meta-analyses were conducted on the relationship between face masks and Covid transmission, there are several drawbacks. Firstly, some meta-analyses include only a limited number of studies (Li et al., 2021). Secondly, there is a missing link to economic and econometric reasoning in the interpretation of results and lacking more complex policy recommendations based on the obtained results. And lastly, since Covid is still present and evolving, I expect to find more recent studies, that were not included in the other meta-analyses.

Hypotheses

Hypothesis #1: Publication bias is present in the literature estimating the effect of face masks on Covid transmission.

Hypothesis #2: The publication bias exaggerates the mean value of the estimated effect of face masks on Covid transmission.

Hypothesis #3: The estimated effect of face masks on Covid transmission is driven by the geographical location.

Methodology Firstly, the dataset consisting of primary studies needs to be constructed. I will define a search query and use the Google Scholar database to full-text search the studies. Secondly, I will examine already published meta-analyses and make sure to incorporate the studies included by authors who already performed a meta-analysis on the topic (Chu et al., 2020; Li et al., 2021; Liang et al., 2020). Additionally, I will search for the studies published recently, to include more recent evidence in my dataset. In the process of collection, I will also focus on other characteristics of the studies, such as standard errors, number of observations, standard deviation, confidence intervals and other effects relevant to the analysis of heterogeneity.

Once I will collect the dataset, I will examine the publication bias using the graphical method – a funnel plot (Egger et al., 1997). In, addition, I will also perform the following tests: funnel asymmetry test (FAT) (Stanley, 2005) with different estimators and weighting matrices, statistical power and bias (Ioannidis et al., 2017), selection model (Andrews & Kasy, 2019), stem-based method (Furukawa, 2019), a kinked meta-regression model (Bom & Rachinger, 2019) and p-uniform* method (Aert & Assen, 2018).

In the second part of the thesis, I will focus on the examination of heterogeneity. For this purpose, I will use the following methods. Bayesian (BMA) and frequentist model averaging, which is used to deal with uncertainty by allowing to assign weights to different models taking into account their data fit, specification and parsimony (Steel, 2020; Amini & Parmeter, 2012). Moreover, I will include several robustness checks – using different priors and weights.

Expected Contribution The effect of facemasks on Covid transmission is important to investigate since it provides a base for policy implications and public health perspective. Since there have been several meta-analyses conducted on the topic (Chu et al., 2020; Li et al., 2021; Liang et al., 2020), I will focus on updating the included studies with more recent evidence.

Additionally, the contribution of my thesis lies in the interpretation of the future results to form a policy recommendation and include an economic and econometric rationale behind chosen study design. I will also focus more on the literature review of the relationship between face masks and Covid transmission, which was not included to a larger extent in other meta-analyses. Additionally, my thesis will include the examination of publication bias, which was also not included by other authors.

Outline

1. Introduction – I will introduce the topic and provide my motivation and contribution to the thesis.
2. Literature review – I will describe already published literature on the topic, its methods, and the main results.
3. Data – This section will describe the process of collection of the dataset (search query, inclusion criteria, etc.) The obtained dataset will be described, and summary statistics will be presented.
4. Methodology – I will describe the methods used to perform a meta-analysis. This section will include both methods related to the examination of publication bias and methods of heterogeneity analysis.
5. Results – I will describe the obtained results and provide their interpretation.
6. Conclusion – This section will summarize the thesis, provide the possible policy implication regarding the obligation of face mask usage, and states any potential drawbacks and limitations. Additionally, possible topics for further research will be mentioned.

Core bibliography

Aert, R.C.M. van & Assen, M.A.L.M. van (2018). Correcting for Publication Bias in a Meta-Analysis with the P-uniform* Method. [Online]. Available from: https://osf.io/preprints/metaarxiv/zqjr9/. [Accessed: 25 June 2022].

Amini, S.M. & Parmeter, C.F. (2012). Comparison of Model Averaging Techniques: Assessing Growth Determinants. Journal of Applied Econometrics. [Online]. 27 (5). p.pp. 870–876. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/jae.2288. [Accessed: 25 June 2022].

Andrews, I. & Kasy, M. (2019). Identification of and Correction for Publication Bias. American Economic Review. [Online]. 109 (8). p.pp. 2766–2794. Available from: https://www.aeaweb.org/articles?id=10.1257/aer.20180310&&from=f. [Accessed: 25 June 2022].

Bom, P.R.D. & Rachinger, H. (2019). A kinked meta-regression model for publication bias correction. Research Synthesis Methods. [Online]. 10 (4). p.pp. 497–514. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/jrsm.1352. [Accessed: 25 June 2022]

Chu, D.K., Akl, E.A., Duda, S., Solo, K., Yaacoub, S., Schünemann, H.J., Chu, D.K., Akl, E.A., El-harakeh, A., Bognanni, A., Lotfi, T., Loeb, M., Hajizadeh, A., Bak, A., Izcovich, A., Cuello-Garcia, C.A., Chen, C., Harris, D.J., Borowiack, E., Chamseddine, F., Schünemann, F., Morgano, G.P., Schünemann, G.E.U.M., Chen, G., Zhao, H., Neumann, I., Chan, J., Khabsa, J., Hneiny, L., Harrison, L., Smith, M., Rizk, N., Rossi, P.G., AbiHanna, P., El-khoury, R., Stalteri, R., Baldeh, T., Piggott, T., Zhang, Y., Saad, Z., Khamis, A., Reinap, M., Duda, S., Solo, K., Yaacoub, S. & Schünemann, H.J. (2020). Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. The Lancet. [Online]. 395 (10242). p.pp. 1973–1987. Available from: https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)31142-9/fulltext. [Accessed: 25 June 2022].

Egger, M., Davey Smith, G., Schneider, M. & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. BMJ (Clinical research ed.). 315 (7109). p.pp. 629–634.

Eikenberry, S.E., Mancuso, M., Iboi, E., Phan, T., Eikenberry, K., Kuang, Y., Kostelich, E. & Gumel, A.B. (2020). To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic. Infectious Disease Modelling. [Online]. 5. p.pp. 293–308. Available from: https://www.sciencedirect.com/science/article/pii/S2468042720300117. [Accessed: 25 June 2022].

Furukawa, C. (2019). Publication Bias under Aggregation Frictions: Theory, Evidence, and a New Correction Method. SSRN Electronic Journal. [Online]. Available from: https://www.ssrn.com/abstract=3362053. [Accessed: 25 June 2022].

Ioannidis, J.P.A., Stanley, T.D. & Doucouliagos, H. (2017). The Power of Bias in Economics Research. The Economic Journal. [Online]. 127 (605). p.pp. F236–F265. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/ecoj.12461. [Accessed: 25 June 2022].

Li, Y., Liang, M., Gao, L., Ayaz Ahmed, M., Uy, J.P., Cheng, C., Zhou, Q. & Sun, C. (2021). Face masks to prevent transmission of COVID-19: A systematic review and meta-analysis. American Journal of Infection Control. [Online]. 49 (7). p.pp. 900–906. Available from: https://www.sciencedirect.com/science/article/pii/S0196655320310439. [Accessed: 25 June 2022].

Liang, M., Gao, L., Cheng, C., Zhou, Q., Uy, J.P., Heiner, K. & Sun, C. (2020).

Efficacy of face mask in preventing respiratory virus transmission: A systematic review and meta-analysis. Travel Medicine and Infectious Disease. [Online]. 36. p.p. 101751. Available from: https://www.sciencedirect.com/science/article/pii/S1477893920302301. [Accessed: 26 June 2022].

Stanley, T.D. (2005). Beyond Publication Bias. Journal of Economic Surveys. [Online]. 19 (3). p.pp. 309–345. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.0950-0804.2005.00250.x. [Accessed: 25 June 2022].

Steel, M.F.J. (2020). Model Averaging and Its Use in Economics. Journal of Economic Literature. [Online]. 58 (3). p.pp. 644–719. Available from: https://www.aeaweb.org/articles?id=10.1257/jel.20191385. [Accessed: 25 June 2022]

Chapter 1

Introduction

As Covid-19 disease began to spread rapidly, the lives of populations all around the world were influenced. To slow down the spread of the contagious disease, various measures were implemented. Since Covid-19 is transmitted mainly by the droplets spread by an infected individual, all the interventions were centred around social distancing. Social distancing can take various forms: stay-at-home orders, restrictions on opening hours, indoor and/or outdoor gatherings, travelling restrictions, school closures and more. Nevertheless, the most popular measure was ordering populations to wear face masks. Knowing the true unbiased effect of face masks on Covid-19 transmission is essential not only for the well-being and health of populations but also for proper policy setting during the pandemic. Apart from the health-related reasons for the evaluation of the mentioned effect, we need to consider the economic consequences of the Covid-19 pandemic as well. As an outcome of social distancing measures, economic activity experienced a major decline. As a result, according to The World Bank, the world's GDP annual growth experienced a drop to -3.1% in 2020. A cost-effectiveness analysis of face masks was performed with the following results: According to Bagepally *et al.* (2021) the additional incurred costs associated with mask-wearing amounted to almost 1 billion USD with the additional 1,121 prevented Covid-19 cases per million subjects with 328 quality-adjusted life years gained. These results are however sensitive to the effectiveness of face masks in preventing Covid-19.

The objective of this thesis was to assess the literature published on the effect of face masks on Covid transmission and perform a quantitative meta-analysis. To do so, we collected 258 estimates from 44 studies, their standard errors, and the variables representing the differences among the studies. We intend to estimate the true value of the effect corrected for publication bias. Publication bias is a serious issue present in the majority of published literature (Stanley 2005). Since the publication of a paper is often determined by the statistical significance of its results, the authors engage in the manipulation of sample sizes and specification of models to achieve significance (Gerber *et al.* 2008; Rothstein *et al.* 2005; Brodeur *et al.* 2018). To examine whether publication bias is present in the collected literature on the mentioned effect, we implemented several modern statistical tests. Firstly, the FAT-PET with different specifications (OLS, Fixed effects, Between effects) and weights were performed. Secondly, we applied a variety of current techniques such as the Endogenous kink model by Bom & Rachinger (2019), the Stem-based method as suggested by Furukawa (2019), the Selection model as in Andrews & Kasy (2019), and more. Thirdly, methods allowing for endogeneity such as FAT-PET with instrumental variable, p-uniform* method as proposed by van Aert & Van Assen (2021) and Caliper tests (Gerber *et al.* 2008) were employed. Based on the results of performed tests, we concluded that there is only mild evidence for publication bias.

Apart from publication bias detection the majority of enumerated methods can be used to estimate the effect beyond bias. The significant estimates of risk associated with face mask-wearing were raging from -0.187 to -0.440. These values can be interpreted as follows: Wearing a face mask is associated with a reduced risk of Covid-19 infection by 18.7% to 44%. Such results suggest a significant protective ability against Covid-19. As a consequent implication in the case of another wave of Covid-19 or a variant resistant to available vaccines, we recommend face masks be used. This thesis also aims to determine the potential drivers behind the heterogeneity of estimates of the effect of face masks on Covid-19 transmission. It is not unlikely that the estimated effects of primary studies vary not only because of the publication bias but because of different settings of the studies, methodology and many other factors including the geographical location and the temperature. Despite several meta-analyses already published on the mentioned effect, they all contain several drawbacks. Firstly, the meta-analysis by Chu *et al.* (2020) published in the Lancet evaluates the effect of face masks, however, the studies included in the meta-analysis are focused on various respiratory diseases, not on Covid-19 specifically. The number of included studies on Covid-19 regarding mask use is as low as four. Including other respiratory illnesses in the meta-analysis can be seen in papers by Jefferson *et al.* (2023); Liang *et al.* (2020); Chaabna *et al.*

(2021). Moreover, the findings of mentioned meta-analyses are contrasting. While Chu *et al.* (2020) reports huge protective abilities of face masks, Jefferson *et al.* (2023) found little to no difference in wearing a mask compared to not wearing one. In addition, the contribution of this thesis lies in performing a quantitative meta-analysis of studies on Covid-19 only. Furthermore, we aim to focus on the examination of heterogeneity and determining its drivers as this was not included in greater detail in the mentioned meta-analyses. Since many variables reflecting the differences among the studies were collected, the model uncertainty needs to be addressed. The solution we applied is the Bayesian and Frequentist model averaging. We found the temperature, geographical latitude, panel data structure, risk ratio estimate type, healthcare set-up, standard error and age to have a positive effect on the risk of Covid-19 infection associated with mask-wearing. The positive effect means that for these variables masks provide lower protection. On the other hand, we found performing an aerosol-generating procedure to have a negative effect. The interpretation of such a result is that mask-wearing is essential during these procedures. Moreover, as a robustness check, the Bayesian model averaging was estimated with different model priors and g-priors yielding highly comparable posterior inclusion probabilities for the variables.

Lastly, we would like to emphasize that the contribution of this thesis lies in its relevance to policy-makers. Moreover, this thesis improves other meta-analyses on the topic by including 44 studies specifically on Covid-19. As compared to other authors, we implemented a wide spectrum of modern meta-regression methods. In addition, we go beyond just estimating the true value of the effect of face masks on Covid transmission and determine the drivers behind the heterogeneity of the estimates.

The thesis is structured as follows. Chapter 2 provides a literature review of not only the primary studies but also the meta-analyses on the effect of face masks on Covid transmission. In order to understand the differences among the studies, we also focus on their approaches used to estimate the effect. Chapter 3 describes in detail the procedure used to obtain the data and the recalculation of both effects and standard errors to achieve comparability of the estimates. Chapter 4 focuses on the examination of publication bias by various modern methods. Chapter 5 implements the model averaging methods to explain the drivers of heterogeneity. Chapter 6 presents the derived best practice estimates and Chapter 7 summarises the thesis.

The data and code are available upon request.

Chapter 2

Literature review

The effect of face masks on Covid transmission has been subject to many debates not only of the general public but especially the policymakers. The following chapter provides a literature review of both meta-analyses and primary studies on the effect of face masks on Covid or other respiratory viral illnesses transmission.

2.1 The Covid-19 pandemic

This section is intended to provide a better understanding of the Covid-19 pandemic and the role of face masks and, why and how they are able to prevent the Covid infection. The pathogen causing Covid-19, officially named Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) was first discovered in China, Hubei province's city of Wuhan in December 2019. The coronavirus has firstly been given the working name 2019 novel coronavirus (2019-nCoV) (Lu *et al.* 2020). According to the World Health Organisation (WHO), 651 918 402 cumulative cases of SARS-CoV-2 infection and 6 656 601 cumulative deaths due to SARS-CoV-2 were reported. The provided numbers were reported as of 23rd December of 2022 (World Health Organisation 2022c). According to Guo *et al.* (2020b) SARS-CoV-2 could come from bats that are believed to be the virus's native host or other unidentified intermediary hosts before infecting humans. Ciotti *et al.* (2020) suggests SARS-CoV-2 originating by mutation of RaTG13 virus that is infecting the horseshoe bat is supported by the substantial similarity of SARS-CoV-2 and RaTG13 of approximately 96%.

Covid-19 is an acute respiratory infection that causes several symptoms. WHO provides the list of the most common symptoms including fever, cough, tiredness, loss of taste and/or smell. The following symptoms: sore throat, headache, aches and pains, diarrhoea, a rash on the skin, or discolouration of fingers or toes, and red or irritated eyes are less common (World Health Organisation 2022a). The majority of infected patients are able to overcome the disease at home. However, some SARS-CoV-2 infected patients, who develop more severe symptoms such as difficulty breathing or shortness of breath, loss of speech or mobility, or confusion, and chest pain might require hospitalisation (World Health Organisation 2022a).

The transmission of Covid-19 can be done via two main channels. Direct infection can be acquired by droplet particles infected by the virus that penetrates the body mainly via the respiratory system. The direct infection is transferred human-to-human when the infected person coughs, sneezes, talks, or sings. In addition, it is also possible to become infected indirectly. The indirect infection can be acquired by contact with contaminated objects and airborne contagion (Lotfi *et al.* 2020; Liu *et al.* 2020b). When talking about transmission, it is essential to mention the reproductive number. According to Chaudhry (2022) the basic reproduction rate (R0) is the most important metric for predicting a new pathogen's capacity to spread. It is defined as the average number of secondary transmissions from a single sick individual. The R0 exceeding one means that the epidemic is progressing. For Covid-19 the estimates by WHO are raging from 1.4 to 2.5, however, the review that examined fourteen studies found the average reproduction number to be 3.28 and median R0 to be 2.79 (Liu *et al.* 2020a).

Before the availability of vaccines, the main preventive measures were using face masks and other Personal Protective Equipment (PPE) such as protective shields, goggles, gowns, and gloves to block pathogen transmission. These were used especially by medical workers. Social distancing, home quarantine, and disinfection of both home surfaces and hands are other common protective measures (Lotfi *et al.* 2020; Ciotti *et al.* 2020; World Health Organisation 2022b). According to Covid19 vaccine tracker (2022), there are eleven vaccines against Covid-19 that have obtained the Emergency Use Listing (EUL) by WHO as of 2nd of December 2022. In European Union (EU) there are seven vaccines approved by European Medicines Agency (EMA), the last one was granted approval as of October 2022 (European Medicines Agency 2022).

From an economic point of view, the pandemic of Covid-19 is associated with increased costs related to the treatment of the infected. Dôvera, a Slovak health insurance company that provides health insurance for more than 30%

of the Slovak population reports that the average cost of hospitalization for a patient with Covid-19 during the second wave of the pandemic was 2,155 euros. Whereas, if the patient additionally required artificial pulmonary ventilation, the average costs increased up to 11,245 euros (Dôvera 2022). Additionally, Brodeur *et al.* (2021) suggests the pandemic has also a negative impact on labour markets, mental well-being, racial disparity, and gender-related outcomes.

2.2 Meta-analyses on the effect of face masks on Covid transmission

So far, some meta-analyses on the mentioned effect have already been published. The most complex meta-analysis so far is the one published by Chu *et al.* (2020). The authors included 44 comparative studies from 9 countries. Included primary studies estimated the effect of face masks not only on SARS-CoV-2 transmission (SARS-CoV-2 is the virus causing the infectious disease Covid-19) but also on Severe Acute Respiratory Syndrome Coronavirus 1 (SARS-CoV-1) and Middle East Respiratory Syndrome Coronavirus (MERS-CoV). In terms of methodology, Bayesian and frequentist averaging and random effects meta-regression was used. Chu *et al.* (2020) suggest that face masks usage decreased the potential odds of infection by 85% (adjusted odds ratio = 0.15) The authors argue the effect is more substantial for N95 type of face masks (or a face mask of comparable quality) than it is for single-use surgical masks.

The meta-analysis by Li *et al.* (2021b) included only 11 primary studies from 4 countries. In contrast to the first mentioned meta-analysis, the authors only included primary studies on SARS-CoV-2. In terms of methodology, the random effects meta-regression model was used. Authors suggest that using a face mask is associated with lower odds of getting infected by SARS-CoV-2 compared to not wearing a face mask (odds ratio = 0.38). Additionally, the effect was more apparent for a group consisting of healthcare workers (odds ratio = 0.29).

Liang *et al.* (2020) meta-analysed 21 primary studies from 8 countries. These studies focused on 5 different respiratory viruses: SARS-CoV-1, Influenza virus, H1N1 (causing swine flu), SARS-CoV-2, and respiratory virus not specified in greater detail. The analysis was conducted using the fixed-effects model (for pooled odds ratio) and the random-effects model. To assess the potential publication bias Begg's and Egger's tests were performed. Based on these tests, the authors concluded that there seems to be no publication bias. Liang *et al.* (2020) suggest that face masks provide increased protection against Respiratory Viral Illnesses (RVIs) (odds ratio = 0.35) with the effect being stronger in the healthcare setting.

Schoberer *et al.* (2022) focused on the effect of several protective measures on Covid transmission. The meta-analysis of the effect of face masks was performed based on 7 observational studies in a healthcare setting. The authors used a (pooled) random effects model. Results suggest that wearing a face mask was associated with a decreased probability of getting infected by Covid for healthcare workers (odds ratio = 0.16). However, the authors additionally pointed out that based on the magnitude of the effect, the certainty of the results was classified as moderate.

In addition, there are more meta-analyses on the topic, however, there are several drawbacks present. Firstly, Chaabna *et al.* (2021) included 13 primary studies, but only one of these studies was focused on Covid. Secondly, Talic *et al.* (2021) included only 6 primary studies focused directly on Covid. And thirdly, Tabatabaeizadeh (2021) focused directly on Covid, still, the number of included primary studies was only 4. As expected, the results of these meta-analyses suggest that the usage of masks was associated with decreased risk of infection. The odds ratios ranged from 0.12 to 0.66. Fixed and random effects models were used.

2.3 Primary literature on the effect of face masks on Covid transmission

Naturally, the research regarding the effect of face masks on Covid transmission started as Covid-19 disease began to spread rapidly at the beginning of 2020 affecting the health and well-being of the population all around the world. There is quite a lot of evidence published on the topic of face masks and SARS-CoV-1 or MERS-CoV transmission. On the other hand, the literature that specifically examines the effect of face masks on Covid-19 transmission is quite limited. There are several study designs and parameters of the primary literature. The main differentiation feature of the papers is whether they focus on healthcare, non-healthcare setting, or on both of them. The papers focusing on healthcare settings include Wang *et al.* (2020b), Heinzerling *et al.* (2020), Guo *et al.*

(2020a), Khalil *et al.* (2020), Chen *et al.* (2020), Wang *et al.* (2020a), and Wang *et al.* (2020c). The effect of face masks on Covid transmission was examined in the non-healthcare environment by Cheng *et al.* (2020), Catching *et al.* (2021), Doung-Ngern *et al.* (2020), Eikenberry *et al.* (2020) and Mittal *et al.* (2020). Moreover, Burke *et al.* (2020), and Chatterjee *et al.* (2020) focused on both healthcare and non-healthcare setting.

As mentioned above, the availability of primary literature on the SARS-CoV-2 transmission and the effect of face masks is not ideal. There are, therefore, many different methodologies used to estimate the effect. Wang *et al.* (2020b) focused on laboratory-confirmed cases and their close contacts. These were used to examine the secondary clinical attack rate. To do so, only symptomatic cases were taken exclusively. The secondary attack rate was compared in different settings. In one of the settings, PPE used by healthcare workers were included. Similarly, the close contact with the index patient was analysed by Heinzerling *et al.* (2020). The authors used a medical records review of the healthcare workers of two hospitals, where one was specifically intended to treat Covid-19 patients, and thus face masks were used by the hospital staff. After acquiring the number of medical workers who became infected following close contact with the index patient, the comparison was carried out. Another paper by Wang *et al.* (2020c) focusing on Covid infection in healthcare staff collected data from six hospital departments in Wuhan, China. Half of these departments were treating Covid patients. As a result, the personnel was protected by face masks and frequently sanitized their hands. Once again, the authors compared the results to the medical workers from the other three departments, where no masks were used and hands were sanitized only sporadically.

Wang *et al.* (2020a) used the data from 107 hospitals located in Hubei province, China. The number of Covid-19 cases among the neurosurgical healthcare personnel was calculated. The effect of face masks was evaluated based on the relative risk. The authors concluded that inadequate protection was associated with an increased risk of contracting the infection. Guo *et al.* (2020a) also focused on a specific medical speciality: orthopaedic surgeons and trainees in Wuhan's metropolitan region. The study was performed on hospital level, 24 hospitals were studied. Data were collected using an online self-administered questionnaire. The results were reported using the odds ratio. Additionally, authors focused not only on face masks but also on other factors such as hand sanitation, and availability of PPE. Another paper using a questionnaire to collect data from almost 200 physicians at several health institutions in Bangladesh (Khalil *et al.* 2020). The results likewise in the previous study reported odds ratio supporting the protective ability of face masks. Furthermore, Chen *et al.* (2020) also reported the odds ratio. However, the methodological approach was different to the two papers described above. The authors used nasopharyngeal swab samples and serum samples to determine the prevalence of SARS-CoV-2 among 105 healthcare workers who were in close contact with infected patients.

Papers by Burke *et al.* (2020) and Chatterjee *et al.* (2020) focused on both healthcare and non-healthcare setting. Chatterjee *et al.* (2020) used the data from the national database of performed Covid-19 tests in India. The database included both medical workers and the general population. To gather information on face mask use, the authors distributed a questionnaire among healthcare workers. To report the results authors presented an adjusted odds ratio. On the other hand, the paper by Burke *et al.* (2020) used data acquired by contact investigation of nine early travel-related Covid-19 cases in the United States. The results were provided as a percentage of infected patients out of all observed subjects for different categories of PPE.

Lastly, two papers focused on non-healthcare setting: Cheng *et al.* (2020) and Doung-Ngern *et al.* (2020). Firstly, Doung-Ngern *et al.* (2020) reported odds ratios and adjusted odds ratios. The study evaluated the effect of protective measures, masks included, in the general public in Thailand on Covid-19 transmission. Secondly, Cheng *et al.* (2020) focused on close contacts of Covid-19 infected individuals in Taiwan. The authors reported risk ratios.

2.3.1 Types of study designs in healthcare

In healthcare, there are several types of study designs. According to the article by Röhrig *et al.* (2009), primary research can be divided into three main categories: basic research, clinical research, and epidemiological research.

- Basic research focuses especially on theoretical, experimental (animal experiments), and other applied study designs such as genetic studies, or cell studies.

- Clinical research can be divided into experimental (phase I up to IV clinical trials) and observational studies. Examples of observational studies include therapy studies, prognostic, and diagnostic studies.

- Similarly to clinical research, epidemiological research contains experimental (interventional studies) and observational study designs (cohort study, cross-sectional study, case-control study, and others).

Probably the most common type of study in estimating the effect of face masks on Covid transmission is the observational study.

2.3.2 Approaches to estimating the effect of face masks on Covid transmission

To report the effectiveness of face masks in reducing Covid transmission authors can estimate a Hazard Ratio (HR). The hazard ratio is defined by National Cancer Institute (2022a) as *"A measure of how often a particular event happens in one group compared to how often it happens in another group, over time."* The interpretation of HR lower than one would be that the hazard of infection in the first group is lower as compared to the second group by $(1 - HR) * 100$ %. The interpretation of HR greater than one, on the other hand, would be that the hazard of infection of the first group is higher by $(HR - 1) * 100$ % as compared to the second group. The HR that is equal to one means the hazard of infection is the same for both groups.

As an alternative to the HR, researchers report the Relative Risk or Risk Ratio (RR) which is defined by National Cancer Institute (2022b) as *"A measure of the risk of a certain event happening in one group compared to the risk of the same event happening in another group."* The risk ratio of 1 means that there is no difference in the risk of Covid infection in one group compared to the second group. For RR lower (greater) than one, the risk of Covid infection is lower (greater) in the first group compared to the second one. If the protective face masks were used by the first group, the interpretation of the RR lower than 1 would be as follows. The usage of face masks decreases the risk of being infected by Covid.

The third ratio that is often reported by researchers is the Odds Ratio (OR). The definition of OR is the following. *"A measure of the odds of an event happening in one group compared to the odds of the same event happening in another group* (National Cancer Institute 2022c)*."* The two groups differ in exposure to a certain factor (face masks, for example). The OR = 1 indicates that the odds of exposure were the same for both groups, and as a result, the exposure to the factor would most likely not increase the risk of infection. If the OR > 1, the exposure to the factor would probably raise the chance of developing infection, and if the OR < 1, it is likely that the exposure would lower the risk of infection. The OR can be estimated using the logit regression model as well as the method described by Altman (1990).

In this sub-chapter, we will focus on approaches that were specifically used to estimate the effect of face mask usage on Covid transmission.

Cox proportional hazards regression model

Health economics is often focused on survival analysis. The survival analysis examines and models the duration of occurrences of an event. The title "survival analysis" and most of its language are derived from the paradigmatic such event, death, but the range of applications for survival analysis is far wider (Fox & Weisberg 2002). One of the applications of the survival analysis is its utilization in estimating the effect of face masks on Covid transmission and corresponding hazard ratios.

Cox (1972) proposed a regression model on how to estimate a hazard ratio. Firstly we will introduce basic concepts that are later used in the estimation procedure. Equation 2.1 describes the hazard function h(t) - the instantaneous risk of an event, given that no event occurred up till now.

$$h(t) = \lim_{\Delta t \to 0} \frac{Pr[(t \leq T < t + \Delta t) \mid T \geq t]}{\Delta t} = \frac{f(t)}{S(t)} \tag{2.1}$$

Where survival time is represented by T. We assume non-negativity, $T \geq 0$. Survival time is viewed as a continuous random variable with the following cumulative distribution function: $P(t) = Pr(T \leq t)$. The probability density function can be derived as the following derivative: $p(t) = \partial P(t)/\partial t$. Additionally, the distribution of survival times can be represented by the survival function illustrated by Equation 2.2.

$$S(t) = Pr(T > t) = 1 - P(t) \tag{2.2}$$

Survival function provides the probability that a person will live through time t (Fox & Weisberg 2002; Cameron & Trivedi 2005).

To estimate the mentioned hazard ratio, we need the Cox proportional hazards regression model specified by Equation 2.3.

$$\log h_i(t) = \alpha(t) + \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_k x_{ik} \tag{2.3}$$

The model is said to be semi-parametric, because baseline hazard function $\alpha(t) = \log h_0(t)$ is left unspecified, without any distributional assumptions. It does not assume any distribution such as exponential, Weibull, log-normal or any other distribution frequently used in survival analysis. While the covariates are represented in the model linearly. Now, the following two equations are assumed.

$$\xi_i = \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_k x_{ik} \tag{2.4}$$

and

$$\xi_{i'} = \beta_1 x_{i'1} + \beta_2 x_{i'2} + ... + \beta_k x_{i'k} \tag{2.5}$$

Equations 2.4 and 2.5 are the predictions of the linear part of the Equation 2.3 for two observations i and i' respectively. The two observations are different in terms of their values of x-es. Using the Equations 2.4 and 2.5, and the baseline hazard function the hazard ratio is defined as the following equation.

$$\frac{h_i(t)}{h_{i'}(t)} = \frac{h_0(t)e^{\xi_i}}{h_0(t)e^{\xi_{i'}}} = \frac{e^{\xi_i}}{e^{\xi_{i'}}} \tag{2.6}$$

The resulting hazard ratio is independent of time t, meaning that the model does not estimate the hazard rate's variation over time. On the other hand, the model estimates how covariates affect a baseline hazard rate (Fox & Weisberg 2002). The described model needs several specific assumptions to be fulfilled. Firstly, the proportional hazards assumption needs to hold. The assumption can be tested graphically by plotting the instantaneous hazard for the two subgroups of the analysis. If the assumption holds, the curves for both subgroups should follow the same trend. The proportional hazard assumption can be also tested by plotting the log cumulative cause-specific hazard function for both subgroups of the analysis. In this case, the assumption holds if the two curves do not cross. In a case when one does not want to rely on graphical methods only, there are tests based on scaled Schoenfeld residuals available. The residuals are calculated from the estimated model specifically for each of the covariates and for the model as a whole, and correlated with time. The assumption is violated if at least one of the covariates or the entire model exhibits a significant relationship between the residuals and time (Grambsch & Therneau 1994; Grambsch 1995; Fox & Weisberg 2002).

The second assumption that needs to hold is the linearity of functional form in the parametric part of the model. Similarly, as with linear and extended linear models, the martingale residuals may be used to create component-plusresidual (or partial-residual) plots and to visualize nonlinearity by plotting them against variables.

Relative risk method

Altman (1990) suggests that in prospective research, groups of participants with various characteristics are monitored to see if the desired outcome materializes. This is true of much clinical research as well as observational ones in which the trait of interest, cannot be randomised. The proportions of each group that have the result are simple to compute, and the ratio of these two proportions indicates how much higher the risk is in one group than in the other. This ratio is known as relative risk. Additionally, in the healthcare setting, relative risk is often reported to provide information about the risk of an event occurring in an exposed group relative to the control group. Generally, the results of a study are usually presented in the following way (Table 2.1). Regarding the situation of the effect of face masks on Covid transmission, the

Table 2.1: General presentation of study results

	Outcome present	Outcome not present	Total
Group 1	a	c	a + c
Group 2	b	d	b + d
Total	a + b	c + d	n

Source: Altman (1990)

event is considered the Covid infection. The two groups are individuals exposed to Covid infection who were or were not using protective face masks. Using the Table 2.1, the relative risk or risk ratio can be calculated as shown in Equation 2.7.

$$RR = \frac{a/(a+b)}{c/(c+d)} \tag{2.7}$$

The following Equation 2.8 can be used to calculate the confidence intervals of the relative risk/risk ratio. The equation uses the standard error of the logarithm of relative risk.

$$\text{SE (ln RR)} = \sqrt{\frac{1}{a} - \frac{1}{a+c} + \frac{1}{b} - \frac{1}{b+d}} \tag{2.8}$$

Finally, the confidence intervals are calculated as shown in Equation 2.9

$$CI_{95\%} = \exp(\ln \text{RR} \pm N_{0.975} * \text{SE}(\ln \text{RR})) \tag{2.9}$$

Since Altman (1990) assumed that the sampling distribution of the logarithm of relative risk is the normal distribution, $N_{0.975}$ is the appropriate value from the normal distribution.

Odds ratio method

According to Tenny & Hoffman (2017), an indicator of how closely an incident is linked to exposure is the odds ratio. The odds ratio uses the two odds. Firstly, the odds of an event occurring in a group that has been exposed to the disease. Secondly, the odds that the event will occur in a group that has not been exposed to the disease. The odds ratio aids in determining how likely it is for an exposure to cause a certain occurrence. Similarly, as in the relative risk method, the results can be presented in a general form of a Table 2.1. If we have a look at the formula for relative risk presented in Equation 2.7, we can derive the formula for the odds ratio. Assuming the a to be small and c to be small as well, the odds ratio can be calculated as shown in Equation 2.10. Altman (1990) suggest that in case-control studies the case-defining outcome of interest is usually rare, which is the basis for assuming a and c to be small enough.

$$OR = \frac{ad}{bc} \tag{2.10}$$

Once again, to calculate the confidence intervals we need the equation using the logarithm of odds ratio as in Equation 2.11.

$$\text{SE } (\ln \text{OR}) = \sqrt{\frac{1}{a} + \frac{1}{c} + \frac{1}{b} + \frac{1}{d}} \tag{2.11}$$

And the corresponding 95% confidence interval can be calculated according to the formula presented in Equation 2.12

$$CI_{95\%} = \exp(\ln \text{OR} \pm N_{0.975} * \text{SE}(\ln \text{OR})) \tag{2.12}$$

(Altman 1990). Additionally, it is important to mention, that if the outcome of interest is rare, the risk ratio and odds ratio are not very different from each other. However, if the outcome of interest is more common to observe, the relative risk and odds ratio are not directly comparable. To deal with this issue, Holland (1989); Zhang & Kai (1998) proposed a method for transitioning between the relative risk and odds ratio that was later summarised by Schmidt & Kohlmann (2008). The relationship between risk ratio and odds ratio can be seen in Equation 2.13.

$$RR = \frac{OR}{1 - I_0 + I_0 * OR}$$
$$\text{alternatively } RR = OR * \frac{1 - I_1}{1 - I_0} \tag{2.13}$$

Where I_1 stands for the fraction $a/(a+b)$ and I_0 represents the ratio of $c/(c+d)$. The fractions represent the prevalence (when cross-sectional data were used) or incidence (when longitudinal data were used) between exposed (I_1) and unexposed (I_0) individuals (Schmidt & Kohlmann 2008).

Logistic regression

Another method for estimating the odds ratio is using the logistic regression model. Firstly, the dependent variable is binary, where 1 usually represents the feature of interest. According to the Altman (1990), logistic regression is implemented in order to predict the proportion of individuals with the mentioned feature of interest for a given combination of the explanatory variables used in the model. In the case of Covid, the feature of interest would be the infection by Covid. To keep the predicted proportion of individuals in the interval bounded by 0 and 1, transformation needs to be used. Equation 2.14 shows the logit transformation.

$$\text{logit}(p) = ln\left(\frac{p}{1-p}\right) \tag{2.14}$$

Where p represents the proportion of subjects with the given feature of interest (to give an example, it could be the proportion of individuals infected by Covid.) Whereas, the expression $1 - p$ stands for the proportion of individuals who do not have the given feature (individuals not infected by Covid). Additionally, the two predictions, one for the group with some characteristics and one without can be compared. Hence, the two groups, one that has been using the face masks and one that has not can be compared by taking the difference of their log odds as can be seen in Equation 2.15.

$$\begin{aligned} l_1 - l_2 &= \text{logit}(p_1) - \text{logit}(p_2) = \\ &= ln\left(\frac{p_1}{1-p_1}\right) - ln\left(\frac{p_2}{1-p_2}\right) = ln\left[\frac{p_1(1-p_2}{p_2(1-p_1)}\right] \end{aligned} \tag{2.15}$$

(Altman 1990).

When it comes to the estimation, the logit is a result of a general latent regression framework $y^* = x'\beta + \epsilon$, where $y = 1$ for $y^* > 0$ and $y = 0$ for $y^* \leq 0$. Taking into account that estimated probabilities need to be bounded by 0 and 1, we need the probability $P(y = 1|x) = 1$ for $x'\beta$ approaching plus infinity and probability $P(y = 1|x) = 0$ for $x'\beta$ approaching minus infinity. Logit uses the logistic distribution as can be seen in the following Equation 2.16.

$$P(Y = 1|x) = \frac{exp(x'\beta)}{1 + exp(x'\beta)} = \Lambda(x'\beta) \tag{2.16}$$

Where $\Lambda(.)$ represents the logistic distribution function. The estimation framework used to estimate the logit model is the maximum likelihood (Greene 2018).

Many studies in healthcare that are based on logit models report the odds ratio. Firstly, the odds in favour of $Y = 1$ for the logit model are expressed as in Equation 2.17.

$$Odds = \frac{P(Y = 1|x)}{P(Y = 0|x)} = \frac{exp(x'\beta)/[1 + exp(x'\beta)]}{1/[1 + exp(x'\beta)]} = exp(x'\beta) \tag{2.17}$$

Secondly, when a change of a dummy variable is considered, the odds ratio can be written as in 2.18.

$$\begin{aligned} OddsRatio &= \frac{Odds(x, d = 1)}{Odds(x, d = 0)} = \\ &= \frac{\left[\frac{exp(x'\beta + \delta * 1)/[1 + exp(x'\beta + \delta * 1)]}{1/[1 + exp(x'\beta + \delta * 1)]}\right]}{\left[\frac{exp(x'\beta + \delta * 0)/[1 + exp(x'\beta + \delta * 0)]}{1/[1 + exp(x'\beta + \delta * 0)]}\right]} = exp(\delta) \end{aligned} \tag{2.18}$$

As a result, although it is not a derivative, the odds' change when a variable changes by one unit is somewhat similar to a partial effect (Greene 2018).

Chapter 3

Data

In this chapter, we will describe the process of obtaining the data that will be further used to examine potential publication bias and heterogeneity. To construct the dataset, we first started by reading the primary studies included in the meta-analyses already conducted on the effect of face masks on Covid transmission by Chu *et al.* (2020); Li *et al.* (2021b); Liang *et al.* (2020) and others. Out of these studies, we constructed a list of crucial primary studies, that was used to adjust the search query until a sufficient number of these crucial studies were included in the top results. Different combinations of the keywords such as "mask", "face mask", "respirator", "Covid-19", "coronavirus", "SARS-CoV-2", "transmission" and similar terms were used. The following query was used to search the studies in the Google Scholar database.

("SARS-CoV-2" OR "2019-nCoV" OR "coronavirus" OR "COVID-19") respirator transmission (observational OR descriptive OR case-control) face mask respirator epidemiological -meta

Google Scholar is considered superior to other databases because of its ability to search through the full text of studies. In this way, we were able to include studies that do not have all the desirable keywords combined in the title or abstract (Gechert *et al.* 2022). Additionally, including only one query for one database allows the search process to be replicated. The search was performed on the 2nd of February and returned more than 8,300 studies. Out of these 8,300 studies, the first 250 were examined. The studies were examined based on the abstract, brief overview of the study and/or quick inspection of the methods and results section (if methods were not described in the abstracts). Sometimes the examination of the abstract was not sufficient because of the following reason. A number of studies from the list of crucial studies did not include the required keywords (such as "mask", "face mask", and "respirator") in the title. Accounting for this, the search query was designed to include also studies that do not have required keywords in the title. As a consequence, the results of the search included numerous narrative reviews, opinions of experts, and papers of non-empirical nature. To quickly identify a study that can be used in the meta-analysis, a deeper examination was needed even in this initial step. The search was restricted to include only studies since 2019. The year when Covid-19 has first been discovered in China. The search query was once again repeated with a restriction including only the last 3 years (2021, 2022, 2023). Since the original search was quite restricted, because of the nature of the topic of this thesis, no new eligible studies were identified this way.

Another source of primary studies is a technique called snowballing. Generally, snowballing is using the list of references of a study to identify additional studies to be included in the meta-analysis (Wohlin 2014). To identify the highest number of studies, we performed the snowballing on two levels. Firstly, when initially screening the studies, we not only downloaded the studies relevant for data collection but also the reviews and other types of literature that included information on research already conducted on the topic. These sources provided us with primary studies that were not captured by the search query. Secondly, after reviewing the papers identified by the search query and the first snowballing, we exported their reference lists. Scopus database was used to download all the references. The references were sorted by the frequency of appearance of the studies. All the studies whose frequency was at least 3 or more were examined. The number of examined studies was 47. In addition, we also reviewed the meta-analyses on the effect of face masks on Covid transmission and identified 7 new studies.

3.1 Inclusion Criteria

In order to perform the quantitative analysis, the following inclusion criteria were set.

- The effect of face masks on Covid transmission is reported as OR, RR, hazard ratio, increase in the number of identified Covid cases for both treatment and control groups or relative change in Covid cases.

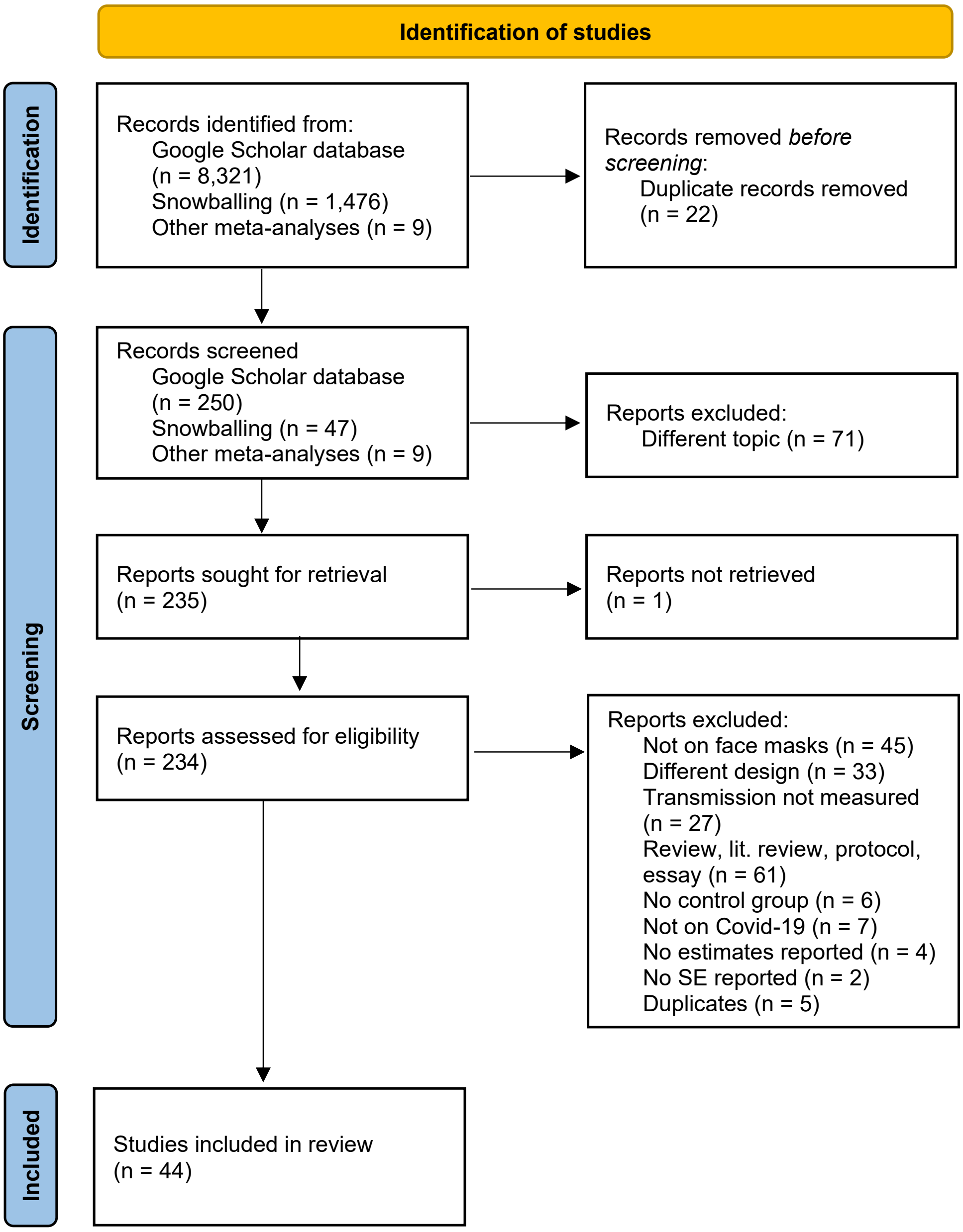
- The study reports standard errors, confidence intervals or p-values.

- The study reports the sample size or enough data to calculate it.
- The explanatory variable needs to be expressed as a number of transmissions/number of identified Covid cases or any other similar measure. Primary studies that use the number of identified Covid-19 cases in numerical units were excluded, if there was no information on the number of transmissions in the absence of the intervention (face mask usage). Using ratios allows us to express the risk in the treatment group relative to the risk in the control group. Reporting a coefficient in numerical terms only would not provide comparable effects among the studies.
- The study needs to include exact information on the intervention.
- The study provides a piece of sufficient information on the control group and its definition.

In addition to the estimated effect reported as a RR, OR or change in Covid cases for control and treatment groups, we included the studies that only report the number of transmissions/Covid cases for treatment and control groups. From these data, OR or RR can be calculated according to the Equation 2.7 and Equation 2.10 described in Subsection 2.3.2. The main reason for including these studies in the meta-analysis is that they were included also in the most comprehensive meta-analysis on the effect of face masks on Covid transmission by Chu *et al.* (2020). More analysis of these estimates is discussed in the following sections.

The process of identification of studies can be seen on the Figure 3.1 together with the reasons for exclusion. Common reasons for exclusion included the study not being oriented on face masks but on different protective measures such as face shields, googles, gowns (for medical personnel especially) or even other social distancing measures such as stay-at-home orders, restriction on opening hours, indoor and/or outdoor gatherings, travelling restrictions, school closures and others. Studies with different designs such as mathematical modelling, prediction models and models performed on artificial datasets were excluded. 17 studies oriented on the different diseases (SARS-CoV-1 or MERS-CoV, seasonal coronaviruses, influenza) were ruled out. The highest number of studies were not included due to their non-empirical nature. These include reviews, narrative reviews, literature reviews, protocols, essays and opinion letters. The number of studies included in the meta-analysis is 44. The list of these studies can be seen in Table 3.1. Apart from the effects and their standard errors,

Figure 3.1: PRISMA flow diagram



Note: The figure shows a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram constructed in line with the identification of studies. The diagram was created based on the template by Page *et al.* (2021).

we collected variables on the estimation methods, sample size, the data type used in the primary studies, variables on publication, relevant control variables included in the models, variables on study setting and country-level variables. 258 estimates were collected from 44 studies. Together with corresponding variables more than 9,300 data points were collected.

Table 3.1: Studies identified for analysis

Author (year)	
Abaluck *et al.* (2022)	Khalil *et al.* (2020)
Akinbami *et al.* (2020)	Li *et al.* (2021a)
Andrejko *et al.* (2022)	Lio *et al.* (2021)
Budzyn *et al.* (2021)	Loeb *et al.* (2022)
Bundgaard *et al.* (2021)	Maltezou *et al.* (2020)
Davido *et al.* (2021)	Martischang *et al.* (2022)
Dörr *et al.* (2022)	Mitze *et al.* (2020)
Doung-Ngern *et al.* (2020)	Nelson *et al.* (2021)
Fawcett *et al.* (2023)	Nguyen *et al.* (2020)
Fletcher *et al.* (2022)	Payne *et al.* (2020)
Gonçalves *et al.* (2021)	Piapan *et al.* (2020)
Guo *et al.* (2020a)	Pienthong *et al.* (2022)
Guy Jr *et al.* (2021)	Rebmann *et al.* (2021)
Haller *et al.* (2022)	Sharif *et al.* (2021)
Heinzerling *et al.* (2020)	Sugimura *et al.* (2021)
Chatterjee *et al.* (2020)	Toyokawa *et al.* (2022)
Chen *et al.* (2020)	van den Broek-Altenburg *et al.* (2021)
Chernozhukov *et al.* (2021)	Van Dyke *et al.* (2020)
Jehn *et al.* (2021)	Venugopal *et al.* (2021)
Joo *et al.* (2021)	Wang *et al.* (2020a)
Kahlert *et al.* (2021)	Wang *et al.* (2020c)
Karaivanov *et al.* (2021)	Wang *et al.* (2020d))

3.2 Recalculating effects

To perform a meta-analysis one needs the effect from the studies to be directly comparable. Since we identified estimates in different forms, they needed to be recalculated. The highest number of estimates were expressed as OR and RR subsequently. Nevertheless, all the effects were recalculated to the risk of Covid-19 infection. There are several reasons for this decision. Firstly, the risk of infection is centred around zero. This means that if there would be zero risk of infection, the corresponding estimate would be $= 0$ as well. On the other hand, estimates expressed in OR and RR are centred around one, meaning that if there is no effect found, the corresponding estimate would be $= 1$. As a result, the tests performed on these estimates and the computation of standard errors would not be straightforward and would require additional adjustments. Secondly, RR, OR and relative change in Covid cases can be easily recalculated to the risk of Covid-19 infection. On the other hand, the recalculation of the effect expressed as a relative change in Covid cases to OR would require more complex computations. The third reason for not choosing OR as a common measure of effects is interpretation difficulties. Moreover, Higgins *et al.* (2019) suggest that the OR is the hardest measure in terms of understanding, application, and is often misinterpreted by researchers. Throughout the data collection, 7 estimate types were identified. The methods for recalculating each type of estimate can be found below.

Risk Ratio For studies, that report their estimates as risk ratio, we can use Equation 3.1 to express RR as 1 plus risk.

$$RR = \frac{risk_{treated}}{risk_{control}} = \frac{risk_{control} + risk_{change}}{risk_{control}} = 1 + \frac{risk_{change}}{risk_{control}} = 1 + risk \quad (3.1)$$

Thus, to recalculate RR as the risk of Covid-19 infection, we subtract 1 from the estimate.

$$risk = RR - 1 \quad (3.2)$$

Apart from the risk ratio, we can find terms relative risk or rate ratio in the literature. The use of these measures is however inconsistent. The main difference is that the risk ratio and relative risk compare the incidence of an event between treatment and control groups. Whereas, the rate ratio uses the incidence rate in two time intervals. In the studies included in the meta-analysis, the time intervals are implemented in order to differentiate the treatment and control period. As a result, we can treat all of the mentioned ratios similarly.

Prevalence Ratio Estimates reported as prevalence ratios can be considered equivalent to the RR. The only recalculation needed is subtracting 1 from the estimate.

Hazard ratio Hazard ratio is different from RR because it takes into account not only the number of events occurring during the observation period but also the timing. Despite the two ratios not being identical, their interpretation is the same. Spruance *et al.* (2004) suggest that the hazard ratio is an approximation of RR. To standardise the hazard ratio, we subtract 1 from the estimate.

Odds Ratio If authors report their estimates as an odds ratio, we can use the following formula described by Zhang & Yu (1998) to recalculate them to the risk of Covid-19 infection.

$$risk = \frac{OR}{1 - p_0 + p_0 * OR} - 1 \tag{3.3}$$

Where p_0 represents the Covid-19 incidence of the control group. As already mentioned, the OR tends to be misinterpreted as RR. This practice is however troubling. If $p_o < 10\%$ the odds ratio estimated by logistic regression can approximate the risk ratio. On the other hand, the higher the incidence, the less precise the approximation is (Zhang & Yu 1998).

Percentage Increase For studies reporting the estimates as a percentage increase, we implement the following standardisation.

$$risk = \frac{percentage\_increase}{100} \tag{3.4}$$

Change Studies that report their estimates as a change to the absolute number of Covid-19 cases need the following standardisation.

$$risk = \frac{risk_{change}}{risk_{base}} \tag{3.5}$$

Regression Coefficient Studies that report the estimates of the effect of masks on the logarithm weekly case growth rate were standardised according to the following equation based on the interpretation of results of the study by Karaivanov *et al.* (2021).

$$risk = exp(estimate) - 1 \tag{3.6}$$

3.3 Standard error calculation

Standard errors were not always reported in primary studies. Some studies only reported confidence intervals or p-values. In this subsection we describe the process used to calculate standard errors.

Delta Method Firstly, if the standard error was reported, but the estimate needed to be standardised, we employ the Delta Method. The form of the Delta Method always depends on the standardisation applied to the estimate. We were able to use the Delta Method only for the studies that report their estimates as the change to the absolute number of Covid-19 cases. Thus, the Delta Method has the following form.

$$\begin{aligned} se(risk) = var\left(\frac{risk_{change}}{risk_{base}}\right)^{\frac{1}{2}} &= \left(\left(\frac{1}{risk_{base}}\right)^{2} var(risk_{change})\right)^{\frac{1}{2}} = \\ &= \frac{se(risk_{change})}{risk_{base}} \end{aligned} \tag{3.7}$$

Budzyn *et al.* (2021) reported estimates as a change to the absolute number of Covid-19 cases per 100,000 persons but did not report the standard errors. In that case, standard errors of the original estimates were calculated from the confidence intervals and then the Delta Method as in Equation 3.7 was used.

Calculation using p-value If a study reported p-values only. We determined the corresponding t-statistic and calculated the standard error for a recalculated estimate using the t-statistic. This approach was used in the study by Karaivanov *et al.* (2021).

Calculation using confidence intervals For studies reporting only confidence intervals, we calculated the standard error according to the following equation for 95% confidence intervals.

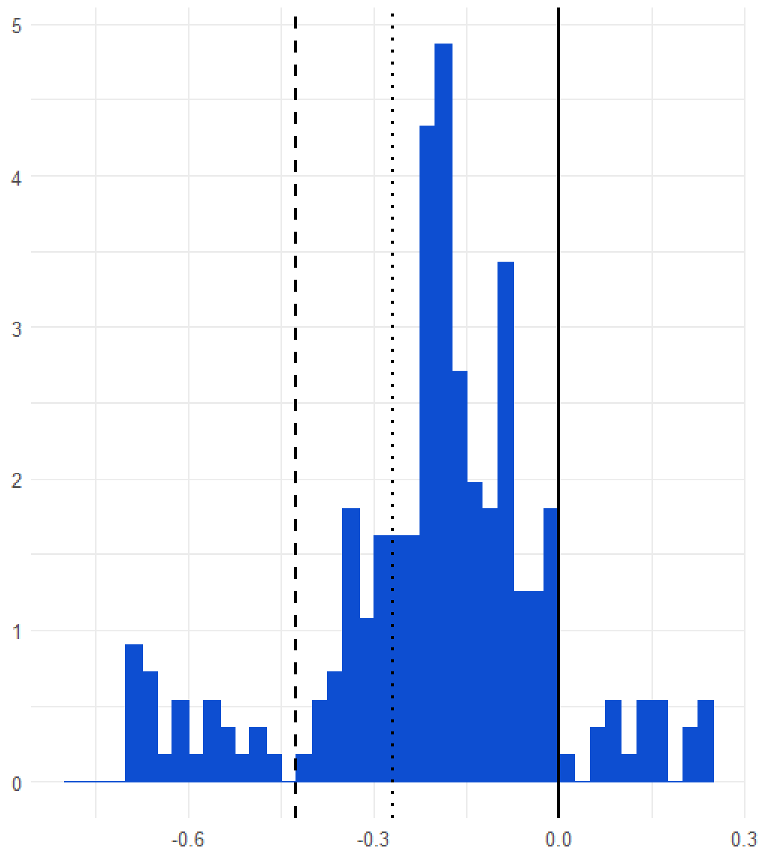
$$se(risk) = \frac{(CI_{upper} - CI_{lower})}{3.92} \tag{3.8}$$

The upper and lower bounds of confidence first needed to be adjusted. If the method was used for calculating standard errors of a study reporting for example a percentage increase, the confidence interval bounds needed to be adjusted at first according to the Equation 3.4. Similarly for other transformations.

3.4 Data description

After the data from studies presented in Table 3.1 were collected, we carefully inspected the data-set and paid specific attention to the outliers. We excluded two observations from the analysis. Both of these observations were collected from the study by Piapan *et al.* (2020). Since Piapan *et al.* (2020) reported only two estimates, the number of studies was reduced to 43. After careful inspection, we winsorized the effects and their standard errors at 1% level. Figure 3.2 shows the distribution by effect magnitude. The estimates of the

Figure 3.2: Effect distribution

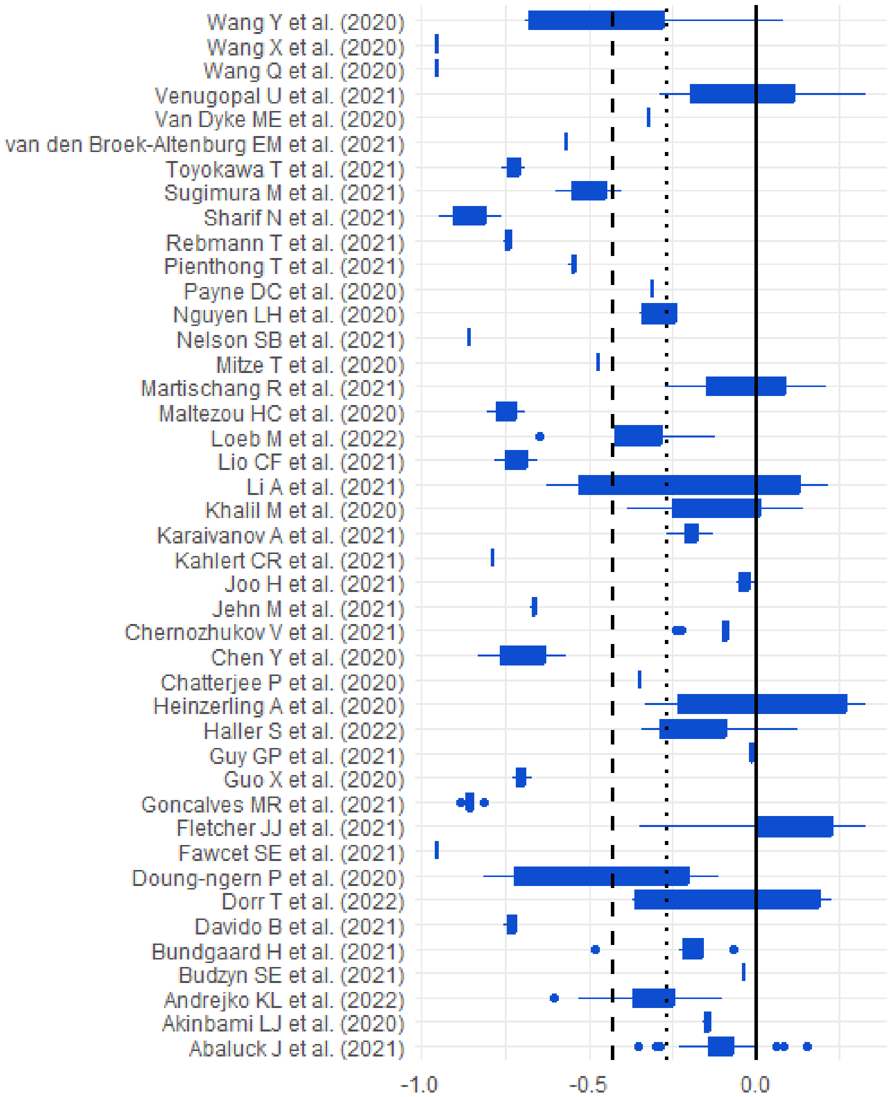


Note: The figure shows the distribution by effect magnitude using winsorized data. The outliers were excluded from the figure but were included in the calculation. The solid vertical line represents 0 intercept. The dotted vertical line is a simple mean and the dashed vertical line represents the weighted mean.

effect of face masks on covid transmission range from -0.956 to 0.33 with a mean value of -0.268 and median value of -0.2. Additionally, we calculated the mean weighted by the inverse number of observations per study which equals to -0.425. The simple mean is higher because studies presenting a higher number of estimates of higher values drive the mean closer to zero. Additionally, the mean value is higher than the median which would suggest a skewed dataset. Chu *et al.* (2020) report the effect as $RR = 0.34$ which can be expressed as $risk = 0.66$. This estimate is much higher that the mean and median values for our collected effects. If we take into account, that even though Chu *et al.* (2020)

included 44 studies in their meta-analysis, only 4 of them were focused on face masks and on Covid-19, their estimates might not be accurate. It is important to mention, that these are just initial remarks based on Figure 3.2 observation and we cannot draw any conclusions yet. Figure 3.3 or forest plot displays the

Figure 3.3: Risk of Covid-19 infection across included studies



Note: The figure shows the effect box plot for every included study calculated using winsorized data. The solid vertical line represents 0 intercept. The dotted vertical line is a simple mean and the dashed vertical line represents the weighted mean.

estimates across studies. Each row represents the individual study included in the meta-analysis. For each study, we present the box plot. Where boxes represent the inter-quartile range (from 25% to 75%). The dots are outliers.

It is apparent from the figure, that estimates vary not only across studies but also within individual studies.

Table 3.2 shows the mean effect of face masks on Covid-19 transmission for selected sub-samples. Some studies or some estimates of studies had their control group protected by lower levels of face masks (respirators for the treatment group and surgical masks for the control group). For these estimates, the mean suggests that wearing a mask might reduce the risk of infection by a lower amount compared to the estimates, where the control group was not protected at all. That would be reasonable since masks might have reduced the risk of transmission in the control group as well. For respirators, we can observe a lower conditional mean risk of infection compared to the surgical masks. For panel data, we can see a higher mean likewise. This could be caused by controlling for other social distancing policies. Additionally the policy control variable and panel data variable are highly correlated.

Interestingly, for estimates with an average minimum temperature during the study period higher or equal to 15°C (warm areas) the mean effect of masks on Covid-19 transmission is higher compared to the mean of estimates where the average maximum temperature was lower or equal to 15°C (cold areas)

For the estimates computing their effect from data only we calculated their simple and weighted mean in order to check whether the effect was not overestimated (masks would be too effective) in these cases. Both means were closer to zero than the means for the rest of the sample. Additionally, these studies were included in the meta-analysis by Chu *et al.* (2020). The estimates also include results from studies with double zero events, which are highly suggested to be included in a meta-analysis Xiao *et al.* (2021). What is more, they represent only around 5% of all observations, thus we decided to include them in the data-set.

Lastly, with available vaccination, the mean is lower. Which is caused by the majority of studies not controlling for vaccination. Thus the seemingly more protective effect of face masks might be probably caused by omitting the vaccination variables from models of primary studies. These are again only observations based on simple descriptive statistics, which cannot be used to draw any conclusions.

Table 3.2: Conditional means

	Mean	95% CI	n
Full sample	-0.268	(-0.805, 0.269)	256
Methodology and effect type			
RR	-0.165	(-0.681, 0.351)	56
OR	-0.425	(-1.062, 0.211)	96
change	-0.158	(-0.320, 0.003)	82
effect from data	-0.160	(-0.938, 0.618)	15
regression	-0.266	(-0.767, 0.235)	238
logit	-0.426	(-1.058, 0.207)	93
cox	-0.240	(-0.524, 0.043)	25
Study set-up			
personal controls	-0.286	(-0.844, 0.273)	104
policy controls	-0.191	(-0.481, 0.098)	94
healthcare	-0.306	(-0.966, 0.354)	68
AGP	-0.379	(-1.044, 0.287)	34
vaccination available	-0.384	(-0.910, 0.143)	31
Mask variables			
mask frequency = all	-0.312	(-0.864, 0.239)	68
mask frequency = some	-0.126	(-0.468, 0.216)	41
respirator	-0.294	(-0.909, 0.321)	46
surgical mask	-0.213	(-0.698, 0.272)	32
control masked = 1	-0.174	(-0.601, 0.253)	35
control masked = 0	-0.283	(-0.830, 0.265)	221
Data characteristics			
panel data	-0.160	(-0.411, 0.091)	164
individual level	-0.315	(-0.932, 0.301)	172
random trial	-0.148	(-0.428, 0.131)	42
data year = 2020	-0.276	(-0.830, 0.279)	189
data year = 2021	-0.245	(-0.730, 0.240)	67
Country characteristics			
China	-0.637	(-1.164, -0.110)	15
Bangladesh	-0.151	(-0.562, 0.261)	36
Switzerland	-0.175	(-0.549, 0.199)	39
USA	-0.186	(-0.735, 0.363)	68
temperature min $\geq$ 15°C	-0.276	(-0.786, 0.234)	108
temperature max $\leq$ 15°C	-0.491	(-0.989, 0.007)	16

Note: The table displays conditional means of the effect of face masks on Covid-19 transmission and corresponding confidence intervals for selected sub-samples. n = sub-sample size

Chapter 4

Publication Bias

In this chapter, we are going to focus on the examination of publication bias. Publication bias is a phenomenon occurring with a preference of researchers for significant effects (Stanley 2005). To describe in greater detail, publication bias occurs when the results of published papers are not a representative sample of all the research conducted on a certain topic. With the increasing number of empirical research, the bias present in the published literature is becoming more severe (Rothstein *et al.* 2005). According to Gerber *et al.* (2008) there are several reasons for publication bias to occur. Firstly, the probability of a paper being published is often determined by whether or not the results of a paper are significant. As a result, researchers would be less likely to submit a paper with non-significant results. This is also known as the file drawer problem: the paper remains in the researchers' drawer.

As a result, researchers might be intentionally adjusting the data-sets or creating sub-samples in order to achieve the desired statistical significance. Additionally, researchers might be prone to using different specifications, which might result in misspecified models with biased results. What is more, the subsequent bias is present not only in individual studies but in the literature as a whole. These practices are known as p-hacking (Brodeur *et al.* 2018). The practice is indeed present in the published literature. Brodeur *et al.* (2016) found evidence for missing published papers with p-values just above the significance threshold of 0.05.

With the use of meta-regression analysis, publication bias and p-hacking have been discovered in the literature on different spheres, among others in economics, social sciences, and medical research (Stanley 2005). Taking into account, that publication bias in medical and related research might have serious consequences for the health of individuals. Some studies, especially at the beginning of the Covid-19 pandemic reported a huge protective ability of face masks (Doung-Ngern *et al.* 2020; Chen *et al.* 2020; Maltezou *et al.* 2020; Wang *et al.* 2020a;c). Combined with uncertainty about the reproduction number of the different variants of Covid-19, populations relying too much on the protective abilities of face masks could have had fatal consequences.

Since some meta-analyses were already published on the effect of face masks on Covid transmission, we can first have a look at their results regarding the publication bias. In the meta-analysis published in *The Lancet*, researchers identified no indication of strong publication bias (Chu *et al.* 2020). On the other hand, the main method used was a graphical Funnel plot approach. No results of more rigorous methods for publication bias detection were reported. The meta-analysis conducted by Li *et al.* (2021b) reported no signs of publication bias. The publication bias was examined using the Funnel plot and two additional statistical tests. Nevertheless, with only six included studies, the credibility of these results should be in question. Similarly, Liang *et al.* (2020) applied the same methods with analogous results. Schoberer *et al.* (2022) addressed the risk of bias in individual primary studies by employing Newcastle Ottawa Scale only. No other tests whether graphical or not were performed.

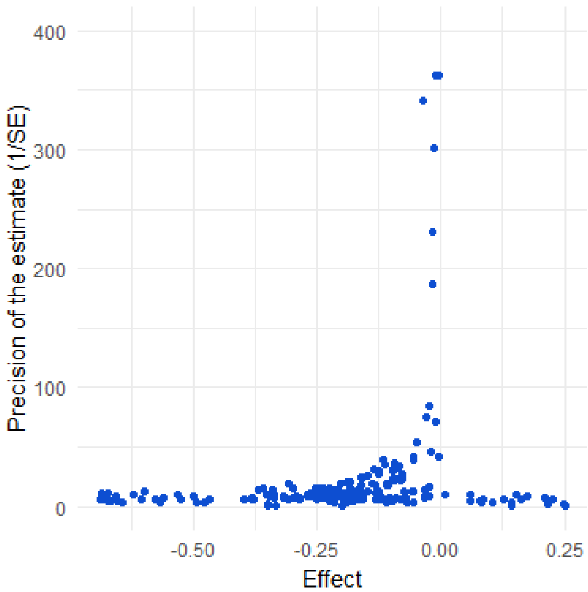
The importance of searching for publication bias on the effect of face masks on Covid-19 transmission has already been outlined. Additionally, from the cited literature, it is obvious that the evaluation of publication bias would benefit from more rigorous methods. The methods that were used in this thesis take inspiration from the ones used by Gechert *et al.* (2022); Havranek *et al.* (2021) in their meta-analyses.

4.1 Graphical method: Funnel plot

The first method we employ is the graphical method for publication bias detection called Funnel plot. The funnel plot was initially described by Egger *et al.* (1997). The graphical test is highly popular among researchers. The majority of meta-analyses include the Funnel plot in the examination of publication bias. The plot is constructed as follows. The horizontal axis plots the estimates of the risk of Covid-19 infection associated with mask-wearing versus their accuracy (the inverse of standard errors) on the vertical axis. The estimates with higher precision should be located around the true value of the risk. On the other hand, the lower the precision of the estimates, the wider the distribution.

In our case, it is obvious from Figure 4.1, that the less precise estimates are located close to the horizontal axis. The funnel plot's ability to graphically detect publication bias lies in the following: If the publication bias is not present in the sample, the funnel should appear symmetrical. In a case of publication bias, the funnel will no longer be symmetrical and skewness and asymmetry will be introduced. Observing the Funnel plot of the effects of face masks on

Figure 4.1: Funnel plot



Note: The figure shows the funnel plot as presented by Egger *et al.* (1997). Outliers were excluded from the figure.

Covid-19 transmission and their corresponding precision - Figure 4.1, one can notice the estimates with the highest precision are centred around a negative value relatively close to zero. On the right side of the Funnel plot, we can observe missing values, compared to the left side of the funnel, which only has estimates with low precision. Such a pattern could suggest possible publication bias. The interesting fact, that we consider important to mention, is that the mean and weighted mean values of the effect are both negative and noticeably different from zero, which is caused by a large number of studies with negative effects of higher magnitude. Generally, the true effect being negative would be in line with the existing theory about mask usage (Ueki *et al.* 2020; Wilson *et al.* 2021) and why populations were advised for their use in the first place.

4.2 Linear methods

Apart from the Funnel plot, there exist more precise tests for publication bias. We first apply the numerical methods of Funnel Asymmetry Test (FAT)-Precision Effect Test (PET) according to methodology suggested by Stanley (2005). FAT-PET is designed in the following way. If the publication bias is not present in the collected estimates, the risk of Covid-19 infection associated with mask-wearing should not be correlated with the standard errors of the risk estimates. The relationship could be induced by studies with less precise estimates adjusting their specifications and/or sample sizes in order to achieve significant results. The dependency of the effects on their standard errors can be formally described according to the Equation 4.1.

$$risk_{ij} = \beta_0 + \beta_1 * (SE_{risk})_{ij} + u_{ij} \tag{4.1}$$

Where $risk_{ij}$ represents the i-th estimate of risk from the j-th study. The β_0 stands for the effect beyond bias, hence the effect corrected for the publication bias. $(SE_{risk})_{ij}$ is the standard error of the i-th estimate of risk from the j-th study. The β_1 represents the estimate of the magnitude of publication bias and u_{ij} is the error term. As presented in Table 4.1 we used five different

Table 4.1: Publication bias: linear methods

	OLS	**FE**	**BE**	**Study**	**Precision**
SE	0.074*	-0.436***	-0.306	0.040	-0.436
Publication bias	(0.038)	(0.068)	(0.222)	(0.038)	(2.104)
Constant	-0.282***	-0.187***	-0.243***	-0.197***	-0.187
Effect beyond bias	(0.018)	(0.001)	(0.027)	(0.008)	(0.158)
Studies	43	43	43	43	43
Observations	256	256	256	256	256

Note: The table displays linear methods for publication bias. OLS = Ordinary Least Squares, FE = Fixed Effects, BE = Between Effects, Study = estimates were weighted by the inverse number of observations reported per study, Precision = estimates were weighted by the inverse of standard errors. Standard errors are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

methods to estimate the Equation 4.1. The estimates from the Ordinary Least Squares (OLS) are presented in the first column of the table. The estimated publication bias seems to be quite small and significant only at 10%. The Fixed Effects (FE) model, which was implemented in order to account for the different characteristics on the study level, is the only model showing a highly significant presence of publication bias. The other three methods do not yield a significant estimate of publication bias. The Between Effects (BE) model accounted for between-study variance. The last two columns of the table present models weighted by the inverse number of estimates reported per study and the inverse of the variance as in Ioannidis *et al.* (2017). On the other hand, the estimates of the effect beyond bias are all negative and highly statistically significant in four out of five presented models. The estimation of all models was repeated using the sub-sample of data where effects calculated from data were excluded. The significance of estimates did not change. The table is presented in Appendix A.

4.3 Non-linear methods

In this section, we employ non-linear methods for the examination of publication bias. The non-linear methods are allowing for a non-linear relationship between effects and standard errors. We implemented six methods for estimating the effect beyond bias and two of these methods to estimate the publication bias. Firstly, we estimate the endogenous kink model as proposed by Bom & Rachinger (2019). The authors developed a meta-regression technique for publication bias correction that locates a kink in the distribution of standard errors. The non-linear method features a horizontal part and a sloped line that together create the kink. The kink in the standard errors' distribution is chosen to that publication bias is not probable beneath the distinguishing value. The publication bias estimate is again not significant. Next, we estimate a Hierarchical Bayes model according to Allenby & Rossi (2006). As we already established, the estimates vary both within and between the studies. Thus, with the use of Bayesian statistics, the model utilises the variability of estimates within individual studies and based on these differences determines the weights assigned to each estimate. Similar to the previous method, the publication bias estimate is not significant. Regarding the effect beyond bias, the estimate is yielding a magnitude similar to the weighted mean of the effects.

Ioannidis *et al.* (2017) proposed a method of the Weighted Average of the Adequately Powered (WAAP). The author suggests that a high portion of the research especially in economics is from the statistical point of view not powered enough. Based on the paper, the usual power is as low as 18%. As a result of low power, the non-existent effects are being detected, which can lead to faulty policy implications. To avoid this trap the author suggests, as apparent from the name of the method, including only the adequately powered observations and running a weighted meta-regression only on this sub-sample. As in the original paper, we chose the statistical level of 5% and the bower bias level of 80%. As a result, only 75 observations remained in the sample. The effect beyond bias coefficient is significant and similar to the estimates from the linear methods.

Table 4.2: Publication bias: non-linear methods

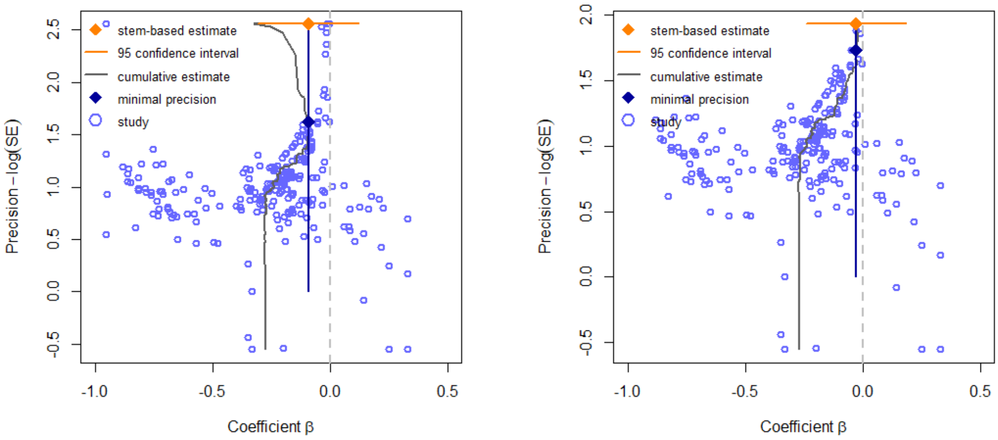
	Effect beyond bias	*Publication bias*
Endogenous Kink	-0.187***	-0.436
(Bom & Rachinger 2019)	(0.024)	(1.315)
Hierarchical Bayes	-0.440***	0.180
(Allenby & Rossi 2006)	(0.097)	(0.304)
WAAP	-0.223***	
(Ioannidis *et al.* 2017)	(0.030)	
Stem-based method	-0.092	
(Furukawa 2019)	(0.110)	
Stem-based method sub-sample	-0.017	
(Furukawa 2019)	(0.041)	
TOP10	-0.094*	
(Stanley *et al.* 2010)	(0.035)	
Selection model	-0.240***	
(Andrews & Kasy 2019)	(0.032)	

Note: The table displays non-linear methods for publication bias. WAAP = Weighted Average of the Adequately Powered, TOP10 = Top 10 Method, Stem-based method sub-sample = Stem-based method performed on data where observations with high negative effects (effect < -0.9) and high precision were removed, *p<0.1; **p<0.05; ***p<0.01

The Stem-based method focuses only on the most precise estimates, the stem of the funnel. The technique as proposed by Furukawa (2019) uses the Equation 4.2 to determine which observations to use in the estimation. The idea is that the bias would decrease as the variance increases (given the higher number of observations).

$$\min_n V\hat{ar}(\hat{b_0^2}, \sigma_0) + B\tilde{ias}^2(\hat{b_0^n}, \hat{b_0}) \text{ subject to } V\hat{ar}(b_i \mid \hat{b_0^n}) = \sigma_0^2 \tag{4.2}$$

The presented equation is the analogue to the similar equation for the mean square error minimization, which would require the knowledge of the unknown true mean b_0. More details are presented in the original paper. We estimated the stem method for the full sample, but as apparent from Figure 4.2, the



(a) Full sample (b) Sub-sample

Figure 4.2: Stem-based method

cumulative estimate at the top of the stem was driven by an effect with high negative magnitude and precision. Additionally, we performed the estimation again to see the behaviour of the estimate. The stem-based method was estimated again but excluded the mentioned observation from the sample. After this procedure, the estimate got closer to zero, but its significance did not change. Next, we employ the TOP10 method as discussed by Stanley *et al.* (2010). It uses a similar principle to the last described method. The author suggests that using only the best 10% of the data can improve the statistical estimation and reduces the publication selection bias, however contradictory to the statistical theory this might be. Nevertheless, the reason for this is that 90% of the data are not representative because of the publication bias. Hence, the remaining 10% of the data should be a better base for efficiently estimating the true effect. The estimated effect beyond bias is significant only at the 10% level and slightly lower than the estimates produced by the other methods. The trustworthiness of the estimate should be in question because of the low number of observations.

The last method for non-linear approaches is the Selection model (Andrews & Kasy 2019). As presented by the authors, the non-parametrically determined probability of a study being published is a function of its results. The probability can be applied in the correction of publication bias. For the purpose of this thesis, we used a 5% significance level and the t-distribution. Results of all methods are presented in Table 4.2.

4.4 Methods allowing for endogeneity

In the last section of this chapter, we discuss the methods allowing for endogeneity. Until now, the methods we discussed assumed that the standard errors are exogenous. The issue with this assumption is the following. The standard errors and the effects could be correlated not only because of the presence of publication bias but as a result of unobserved heterogeneity or measurement errors. We suggest this would be the case for the effects and the standard errors in the collected data-set. As a result of the different methodological approaches used in the primary studies, we expect that some of the methods yield systematically higher standard errors. Table 4.3 shows the results of two methods: Instrumental Variable (IV) estimation and p-uniform* method. The inverse of the square root of the number of observations in primary studies was used as an IV for standard errors (Gechert *et al.* 2022). Secondly, the p-uniform* method developed by van Aert & Van Assen (2021) found the presence of publication bias significant. The effect beyond bias is closer to the weighted mean of the studies compared to the other methods. P-uniform* was estimated using the method of moments. The idea behind the method is the following. The p-values should be distributed uniformly. However, the publication bias is affecting their distribution. Under publication bias, the significant estimates just below the threshold are over-represented, on the other hand, the estimates with p-values just above the 5% level are under-represented. The goal of the p-uniform* method is to find a value around which the p-values follow a uniform distribution.

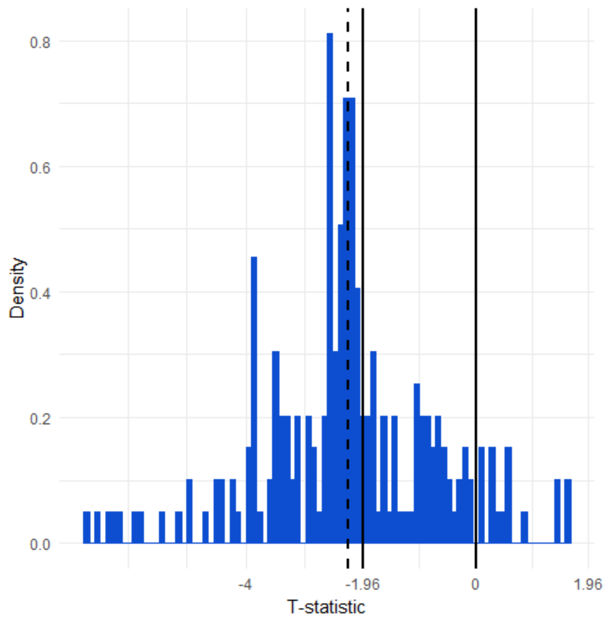
Table 4.3: Publication bias: methods allowing for endogeneity

	IV	**p-uniform***
Publication bias	-0.249	0.148***
	(0.191)	(0.068)
Effect beyond bias	-0.221***	-0.422***
	(0.028)	(0.111)
studies	43	43
number of observations	256	256

Note: The table displays methods for publication bias that allow for endogeneity. IV = regression taking the inverse of the square root of the number of observations as an instrument for the standard error, p-uniform* was estimated using the method of moments, *p<0.1; **p<0.05; ***p<0.01

Let us present the last method utilised in this thesis, allowing for endogeneity in standard errors. Caliper test as described by Gerber *et al.* (2008). Authors suggest that publication bias is the cause of possible jumps in the distribution of t-statistics at significant thresholds of 1.96 and -1.96. Additionally, it is possible to evaluate the behaviour at the 0 threshold. Figure 4.3 shows the distribution of t-statistics for collected effects of face masks on Covid-19 transmission. Looking at the -1.96 threshold we can see a jump in the distribution, with more observations just above the threshold. At 1.96 we cannot observe any t-statistics. Since the majority of our effects are negative, so are the corresponding t-statistics, resulting in no values at this threshold. At 0 we do not observe any major jumps in the distribution. However, only a simple glance at the figure suggests that there are more effects in the $(-1.96, 0)$ interval than in the $(0, 1.96)$ one.

Figure 4.3: t-statistics distribution



Note: The figure shows the distribution of t-statistics. The solid vertical lines display -1.96 and 0 thresholds. The dashed vertical line represents the simple mean of t-statistics. The outliers were excluded from the figure but remained in the calculations.

The Caliper test as compared to the previous methods does not assume any relationship between effects and standard errors. The idea is to compare the number of t-statistics above and below the significance threshold to detect whether publication bias is present. Table 4.4 shows the performed Caliper tests for -1.96 and 0 thresholds for presented Caliper widths. We would like to note that as a result of the sample size the Caliper widths were set quite wide to have enough observations in the given Calipers. Since the Caliper of the width 0.2 contains only 9 t-statistics we would rather not interpret the results of the test. For Caliper widths 0.5, 0.6, 0.7 and 0.8 significant results were obtained. The value 0.684 for 0.5 Caliper width can be interpreted as follows. For interval (-2.21, -1,71) there are 38 t-statistics and 68.4% of them are below the -1.96 threshold. The percentage is even higher for wider Calipers. For the 0 threshold, we do not detect any significant results.

Table 4.4: Publication bias: Caliper tests

	Threshold = -1.96	n	**Threshold = 0**	n
Caliper width = 0.2	0.778*	9		
	(0.147)			
Caliper width = 0.3	0.632	19		
	(0.114)			
Caliper width = 0.4	0.583	24		
	(0.103)			
Caliper width = 0.5	0.684**	38		
	(0.076)			
Caliper width = 0.6	0.745***	51	0.615	13
	(0.062)		(0.140)	
Caliper width = 0.7	0.724***	58	0.556	18
	(0.059)		(0.121)	
Caliper width = 0.8	0.730***	63	0.579	19
	(0.056)		(0.116)	

Note: The table displays the results of the Caliper test as described by Gerber *et al.* (2008) for presented Caliper widths. Caliper width of 0.1 did not contain enough observations even for the -1.96 threshold, standard errors are presented in parentheses, *p<0.1; **p<0.05; ***p<0.01

To summarise, we found significant evidence for publication bias only in some of the performed tests. For tests that identified the significant presence of publication bias in the literature, its magnitude was considered mild. According to Doucouliagos & Stanley (2013) the estimate of $|\hat{\beta}_1| < 1$ is considered to be mild evidence of publication bias. These findings are in line with the ones by Chu *et al.* (2020). The effect beyond bias was estimated to be negative and statistically significant for almost all of the methods. In addition, we did not detect any positive significant estimates of the effect beyond bias. As a result, we believe that the true effect is negative, but its magnitude is varying. The reasons why the effects are different not only among studies but within individual studies too need to be evaluated. We will focus on these reasons in the following chapter.

Chapter 5

Heterogeneity

As already established in the previous chapters, the estimates of the effect of face masks on Covid transmission vary. The effect is very likely driven by different factors. As one was able to observe during the pandemic of Covid-19, in different countries the virus was spreading at varying paces. Thus we expect to find country variables affecting the risk of infection. Additionally, among others, the set-up, quality, and estimation procedures of the included were not the same. In this chapter, we focus our attention on the factors influencing the effect of face masks on Covid transmission.

5.1 Coding of variables

During the data collection process, many variables were collected. To provide the reader with a better understanding of how the studies differ, we explain the rationale behind the individual variables. The variables were divided into categories: Methodology and effect type, Study set-up, Data characteristics, Country and individual characteristics, and Publication characteristics.

Methodology and effect type The methods in primary studies used to estimate the effect of face masks on Covid-19 transmission differ. Nevertheless, as already discussed in Chapter 3 collected estimates were recalculated to the risk of infection. To account for the possibility of different methods producing systematically higher or lower estimates, we collected the corresponding dummy variables. The majority of estimates were obtained by the implementation of a certain type of regression. These estimates account for almost 95% of all collected effects. Logistic regression was used in 22 studies. Additionally, regression with a logit link was the most common method for estimating OR as seen in Andrejko *et al.* (2022); Bundgaard *et al.* (2021); Davido *et al.* (2021); Doung-Ngern *et al.* (2020); Gonçalves *et al.* (2021); Haller *et al.* (2022) and more. The last author additionally used Cox regression to estimate the effects. Cox regression was also the method implemented by Loeb *et al.* (2022); Nguyen *et al.* (2020). As already apparent from previous chapters, another commonly reported effect type is RR. In the majority of studies, the RR was estimated using regression (Loeb *et al.* 2022; Martischang *et al.* 2022; Sugimura *et al.* 2021). For instance, Abaluck *et al.* (2022) estimated weighted OLS. On the other hand, some of the effects (around 5%) were not estimated by regression but calculated from the data according to the equations in Subsection 2.3.2 (Fawcett *et al.* 2023; Fletcher *et al.* 2022; Heinzerling *et al.* 2020; Pienthong *et al.* 2022; van den Broek-Altenburg *et al.* 2021). Lastly, we included a dummy variable to distinguish the studies reporting the absolute and relative change in Covid-19 cases (studies by Budzyn *et al.* (2021) and Van Dyke *et al.* (2020)). The variable was later joined with a dummy variable for percentage increase (studies by Guy Jr *et al.* (2021) and Joo *et al.* (2021)) and a dummy for regression coefficient and corresponding recalculation as seen Equation 3.6 (studies by Karaivanov *et al.* (2021) and Chernozhukov *et al.* (2021)). The reason for joining the three variables is the similarity in the approach of the primary studies which estimated the types of effects.

Study set-up Regarding the estimation procedure and models of the primary studies, we would consider it beneficial to include a set of dummy variables to code the control variables included in the models. Unfortunately, the vast majority of primary studies included in the meta-analysis as well as other studies that we encountered during the identification procedure are of low transparency. The studies do not include the full list of variables included in their models. Because of this reason, we decided to collect at least two dummy variables representing controlling for personal and policy characteristics. Personal characteristics represent for example age, education, number of children, number of household members and occupation. The policy controls variable is designed to represent the estimates, that were obtained from models controlling for other social distancing policies. These policies might include stay-at-home orders, restrictions on public gatherings, school closures, *et cetera*. Brooks-Pollock *et al.* (2021) suggest that these variables influence the transmission of Covid-19. Thus, their omission from the models would result in biased estimates of the effect of face masks. What is more, controlling for vaccination in the models of primary studies was almost non-existent. The first reason could be that the studies were conducted before the vaccination was available in the given region. However, even the studies carried out during the periods when vaccination was already available failed to control for the vaccination. To account for possibly different values of the estimates, a dummy variable indicating the availability of vaccination in a given region and time period was coded.

The primary studies could be divided into two categories: the ones performed in a healthcare environment (among others Fletcher *et al.* (2022); Guo *et al.* (2020a); Wang *et al.* (2020a); Heinzerling *et al.* (2020)) and the ones in a non-healthcare environment such as Karaivanov *et al.* (2021); Toyokawa *et al.* (2022); Wang *et al.* (2020d) and others. Additionally, the healthcare personnel performing Aerosol Generating Procedures (AGP) could be at higher risk of Covid-19 infection given the transmission route of Covid-19 as discussed in Section 2.1. The corresponding dummy variables were coded. We also coded a dummy variable for studies that used a lower grade of protection as their control or base group. Such a practice would likely produce different estimates compared to having non-masked individuals as control.

Some studies were of randomised clinical trial design (Abaluck *et al.* 2022; Bundgaard *et al.* 2021; Loeb *et al.* 2022). The difference is that the random trial studies were properly randomized and the control and treatment groups should be comparable in terms of the characteristics of included subjects (National Cancer Institute 2022d). Consequently, estimates from these studies should be close to the true effect. Lastly, the studies included in the meta-analysis were performed either on an individual level where the outcome for every individual was determined, or on the macro level reporting estimates per 100,000 subjects. Macro estimates were collected from the studies by Chernozhukov *et al.* (2021); Budzyn *et al.* (2021); Guy Jr *et al.* (2021); Jehn *et al.* (2021); Joo *et al.* (2021); Karaivanov *et al.* (2021); Mitze *et al.* (2020); Van Dyke *et al.* (2020).

Data characteristics The collected effect could be influenced by the data structure. We coded a dummy variable panel data, to distinguish the estimates obtained from studies whose dataset was of the panel structure, for example, Abaluck *et al.* (2022); Budzyn *et al.* (2021); Bundgaard *et al.* (2021); Dörr *et al.* (2022); Guy Jr *et al.* (2021); Nguyen *et al.* (2020) and 8 more. Following the reporting guidelines by Havránek *et al.* (2020) we collected the variables for the sample size of a study and the average year in which the study was performed.

Since Covid-19 is relatively a new disease, we included 32 studies conducted in 2020 and 11 in 2021.

Country and individual characteristics One of the advantages of the meta-analysis is that we are able to collect country-level variables and estimate their effect on the risk of Covid-19 infection. According to World Health Organisation (2022c) the number of Covid-19 cases vary for different countries. Thus, we consider it essential to collect also the country-level variables. The variables we collected were the geographical latitude of the region where the study was conducted, and the minimum and maximum average temperatures. The temperature variables were determined based on the area and time period of the study. As suggested by Shi *et al.* (2020) and Notari (2021) the temperature is a fundamental factor in the dynamics of Covid-19 transmission and thus determines the effect of face masks on Covid-19 transmission as well. In addition to country-level variables, we included a variable representing the average age of the subjects of a study. It is important to mention, that the majority of studies reported the average age of subjects to be around 40 years. Abaluck *et al.* (2022) and Joo *et al.* (2021) estimated the effect also for sub-samples where the average age was more than 65 years. Six such estimates were collected. The average age of 12 years was reported by tree studies (Budzyn *et al.* 2021; Jehn *et al.* 2021; Nelson *et al.* 2021). From these studies, three estimates were collected. Lastly, we collected eleven estimates with the average age of subjects around 30 (Abaluck *et al.* 2022; Chen *et al.* 2020; Pienthong *et al.* 2022; Payne *et al.* 2020; Sharif *et al.* 2021).

Publication characteristics The publication characteristics were collected in line with Havránek *et al.* (2020). Namely, we collected a dummy variable for studies published in a peer-reviewed journal. The variable could possibly have an influence on the effects. Since peer-reviewed journals publish studies with higher quality and validity (Kelly *et al.* 2014). We also coded a variable reflecting the impact factor of a journal in which a study was published. Unfortunately, we were not able to use the RePEc factor, since the majority of journals were not of an economic nature. As a substitute, we used Journal Citation Reports (JCR) database which also includes medical journals. Next, we collected a variable on the year of the publication of a study and the number of citations in Google Scholar. The number of citations could be a factor determining the quality of a study.

Table 5.1: Description of variables

Variable	Description	Mean	SD
effect	the risk of Covid-19 infection	-0.268	0.274
standard error	standard error of the risk of Covid-19 infection	0.187	0.470
Methodology and effect type			
RR	=1 if a study reports the estimates as RR	0.219	0.414
OR	=1 if a study reports the estimates as OR	0.375	0.485
change	=1 if a study reports the estimates as a change to identified Covid-19 cases	0.320	0.468
effect from data	=1 if the effect is calculated from data	0.059	0.235
regression	=1 if the effect is estimated using any kind of regression	0.930	0.256
logit	=1 if the effect is estimated using regression with logit link	0.363	0.482
cox	=1 if the effect is estimated using Cox regression	0.098	0.297
Study set-up			
personal controls	=1 if a study controlled for personal characteristics in its model	0.406	0.492
policy controls	=1 if a study controlled for other social distancing policies in its model	0.367	0.483
healthcare	=1 if a study was conducted in a healthcare setting	0.266	0.443
AGP	=1 if subjects were performing AGP	0.133	0.340
vaccination available	=1 if vaccination was available during the period and country in which a study was performed	0.121	0.327
random trial	=1 if a study is of random trial design	0.164	0.371
individual level	=1 if a study was performed on an individual level	0.672	0.470
control masked	=1 if the control group was using a lower grade of mask	0.137	0.344
Data characteristics			
panel data	=1 if the data is of panel structure	0.641	0.481
sample size	logarithm of a sample size of a study	7.904	2.060
year data	the year in which a study was performed (average for more years)	2020.262	0.440
Country and individual characteristics			
min temperature	average minimum temperature for a study's time period and area	3.555	10.559
max temperature	average maximum temperature for a study's time period and area	27.012	5.772
latitude	logarithm of latitude of study's area	3.627	0.429
age	logarithm of the average age of study's subjects	3.684	0.211
Publication characteristics			
peer review	=1 if published in peer-reviewed journal	0.992	0.088
impact	logarithm of the impact factor of a journal	2.348	1.588
year publication	logarithm of the year in which a study was published	2021.039	0.619
citations	logarithm of the number of citations on Google Scholar	4.262	1.395

Note: The table displays the mean and standard deviation of variables eligible for use in meta-regressions, AGP = Aerosol Generating Procedure

Table 5.1 displays the mean, standard deviation and a description of described variables that could be causing the different values of the collected effects. The year data and year publication variables were not shown in the logarithmic scale because of low variability in these variables. The temperature variables are also not shown in the logarithmic scale. The reason is the negative values in the variable minimum temperature.

5.2 Estimation

Now that we have characterised potential drivers behind the heterogeneity of the effects, we follow with the estimation. Since many of the collected variables were correlated, we needed to exclude some of them from the analysis not to have multicollinearity present in the models. To decide which of the correlated variables to remove, we proceed as follows. Firstly, we calculated the Variance Inflation Factor (VIF) scores for all of the variables. The variables with VIF score below 10 were selected for the analysis. Next, the correlation coefficients were calculated among all of the variables. Additional variables were removed based on which variables were already selected for the analysis and their corresponding correlation coefficients. After the removal, VIF scores were calculated again and the procedure was repeated two times. Additionally, dummy variables impact factor and random trial were highly correlated. We decided to remove the random trial variable and prioritize including the impact factor in the model. Variables minimum temperature and latitude were correlated as well. Since the maximum temperature variable was not correlated with the latitude, we decided to include the latitude in the model and remove the minimum temperature variable. The maximum temperature in contrast to the minimum temperature variable did not contain any negative values, hence it was adjusted to the logarithmic scale. Lastly, the variable publication year and data year did not have almost any variation, so we excluded them. The VIF scores were large too. As a result, 8 variables were removed and 18 were selected for the analysis. The selected variables together with their VIF scores can be seen in Table 5.2. To estimate the effect of the selected variables on the risk of Covid-19 transmission we could utilize the following equation.

$$\hat{risk}_{ij} = \beta_0 + \sum_{l=1}^{18} \beta_l X_{l,ij} + \gamma SE(\hat{risk}_{ij}) + \mu_{ij} \tag{5.1}$$

Table 5.2: Variables and their VIF scores

Variable	VIF		
Standard error	2.55	Control masked	3.27
Risk ratio	5.29	Panel data	4.34
Effect from data	3.20	Sample size	3.74
Logit	6.53	Vaccination available	2.17
Cox	2.76	Max temperature	2.35
Personal controls	1.91	Latitude	3.01
Policy controls	4.38	Age	1.33
Healthcare	4.76	Impact	1.91
AGP	2.38	Citations	2.35

Where $ris\hat{k}_{ij}$ represents the i-th estimate of risk from the j-th study. β_0 stands for the effect beyond bias conditional on X. The estimate of β_0 cannot be interpreted on its own. $X_{l,ij}$ is the matrix of control variables listed in Table 5.2 with their corresponding estimates β_l. γ is the estimate of publication bias and μ_{ij} represents the error term. Unfortunately, it would be problematic to estimate Equation 5.1 using OLS. The selected variables account for different settings and methodological approaches of primary studies. It is not unlikely that the inclusion of all variables into a single model would cause over-specification. On the other hand, selecting only some of the variables would not be wise, because of the model uncertainty. Additionally, the results of such an estimation would be likely biased and imprecise. If we were to estimate models with all of the possible combinations of variables, the number of these models would be 2^{18}. As a solution, we selected the approach commonly used in meta-analyses, Bayesian Model Averageing (BMA)(Havranek *et al.* 2018; Havranek & Sokolova 2020; Havranek *et al.* 2021; Gechert *et al.* 2022).

The idea behind the BMA is selecting the most appropriate subset of regressors. To do so, each estimated model is assigned a score called Posterior Model Probability (PMP) which is a representation of the model's performance. Next, the score called Posterior Inclusion probability (PIP) is assigned to a regressor. PIP is calculated as a sum of all PMP for models where the given regressor was included (Eicher *et al.* 2011; Steel 2020). Still, the number of all possible specifications is large. Thus, it would not be feasible to estimate all of the possible models. Hence, we use a Markov chain Monte Carlo (Madigan *et al.* 1995) algorithm to select for estimation only those models from the model space where PMP would be high. For BMA we need to specify the weight of the priors for each coefficient. This is referred to as g-prior (Havranek *et al.* 2018). We set the g-prior to the common practice in meta-analyses a unit information prior (Havranek *et al.* 2018). We use the unit information prior, meaning the weights are set to give the prior the same importance as one individual observation (Eicher *et al.* 2011; Havranek *et al.* 2018). In addition, we need to choose the prior for the model probability. As a baseline, we selected two approaches. The uniform model prior, which gives each model the same prior probability and the dilution prior. The latter is more suitable when dealing with potential collinearity. For small sample sizes - as is the case for this thesis, the models are prone to suffer from collinearity. The dilution prior tackles the issue by giving less weight to the models suffering from a lot of collinearity (George *et al.* 2010). The robustness checks with different g-priors and model priors are included in the appendices.

In addition to BMA we implement the Frequentist Model Averageing (FMA). Following the practice of Gechert *et al.* (2022) we employ Mallow's criteria as weights (Hansen 2007). Analogously to the BMA, we need to subset the space of all models to perform the estimation in a reasonable time. Since the Markov chain Monte Carlo algorithm is not applicable for FMA, the orthogonalization of the covariate space is used (Amini & Parmeter 2012). The results are presented in Table 5.3, Figure 5.1 and Figure 5.2 show the graphical results of BMA models.

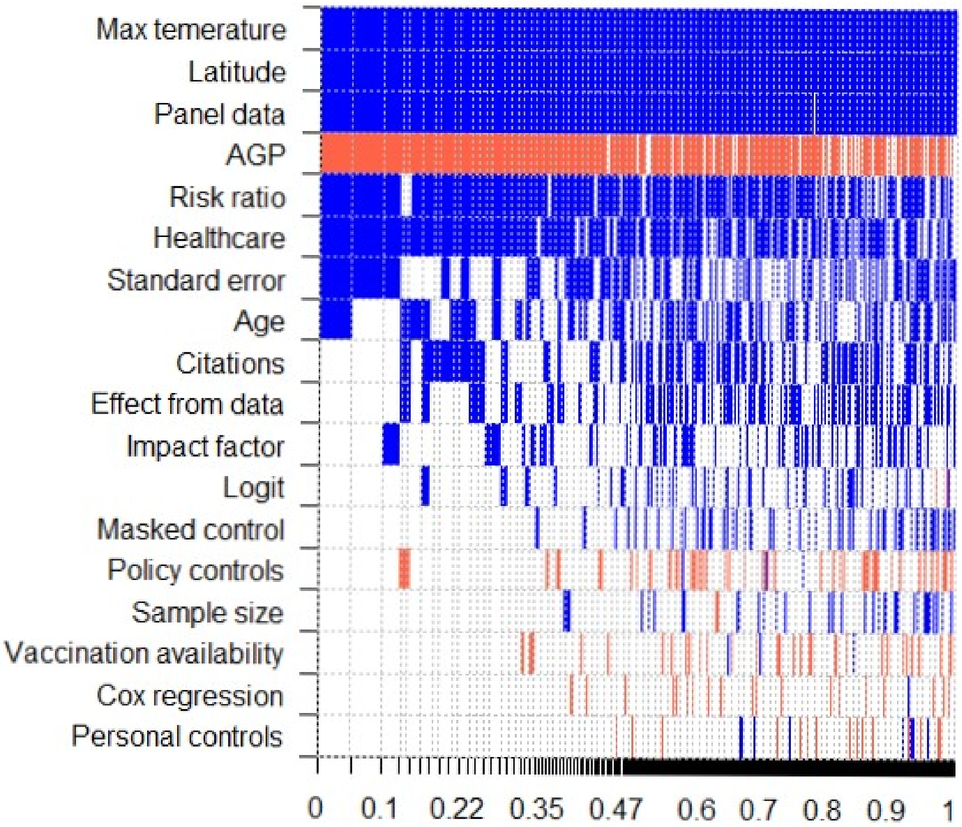
The highest posterior probability inclusion can be seen for variables representing the maximum average temperature and geographical latitude. The coefficient for max temperature is positive, which can be interpreted as follows. With increasing maximum temperature the protection provided by masks is lower. This is probably caused by the lower transmission of Covid-19 during summer periods (Shi *et al.* 2020). For latitude, the interpretation is the following: with increasing latitude, the masks are less effective. This would be caused by lower temperatures for regions with higher latitudes. For variable panel data, we got a positive estimate as well. The reason is that the panel data variable was correlated with the random trial variable which was not included in the model. The reasoning would be that these estimates are higher because the panel structure of the data would likely decrease the probability of estimating the effect at a non-representative point in the time. As expected, the healthcare variable has also a positive effect. Meaning that the masks are less protective in the healthcare environment because healthcare professionals are in contact with infected individuals.

Table 5.3: The results of BMA and FMA

Response variable:	*BMA uniform*			*BMA dilution*			*FMA*		
estimate of risk	post. mean	post. SD	PIP	post. mean	post. SD	PIP	coef.	SE	p-value
standard error	0.043	0.050	0.501	0.043	0.051	0.492	0.039	0.042	0.355
intercept	-2.848	NA	1.000	-2.797	NA	1.000	-3.466	0.444	0.000
Methodology and effect type									
RR	0.118	0.064	0.854	0.119	0.065	0.850	0.115	0.066	0.082
effect from data	0.071	0.106	0.377	0.069	0.106	0.365	0.150	0.093	0.106
logit	0.013	0.034	0.176	0.012	0.033	0.164	0.074	0.061	0.228
cox	-0.004	0.020	0.086	-0.003	0.019	0.080	-0.004	0.053	0.942
Study set-up									
personal controls	-0.000	0.008	0.060	-0.000	0.008	0.058	-0.014	0.031	0.648
policy controls	-0.009	0.030	0.147	-0.008	0.029	0.136	-0.032	0.051	0.531
healthcare	0.120	0.072	0.807	0.110	0.076	0.747	0.127	0.060	0.034
AGP	-0.164	0.080	0.890	-0.152	0.087	0.832	-0.170	0.056	0.002
vaccination available	-0.006	0.025	0.107	-0.005	0.024	0.099	-0.026	0.052	0.620
control masked	0.012	0.038	0.156	0.012	0.037	0.149	0.026	0.060	0.660
Data characteristics									
panel data	0.196	0.046	0.998	0.196	0.047	0.997	0.209	0.053	0.000
sample size	0.001	0.006	0.110	0.001	0.006	0.109	0.006	0.011	0.564
Country and individual characteristics									
max temperature	0.379	0.077	1.000	0.372	0.079	1.000	0.431	0.074	0.000
latitude	0.237	0.048	1.000	0.236	0.049	1.000	0.265	0.050	0.000
age	0.064	0.080	0.459	0.058	0.079	0.419	0.126	0.067	0.060
Publication characteristics									
impact	0.004	0.009	0.250	0.004	0.009	0.229	0.008	0.011	0.446
citations	0.012	0.016	0.441	0.011	0.016	0.410	0.024	0.014	0.073

Note: The table displays the results of Bayesian Model Averaging with uniform and dilution model priors and Frequentist Model Averaging results, PIP = Posterior Inclusion probability, AGP = Aerosol Generating Procedure

Figure 5.1: BMA with a uniform model prior and unit information g-prior



Note: The figure shows the Bayesian Model Averaging with the uniform model prior and unit information g-prior. The response variable is the risk of Covid-19 infection. The horizontal axis represents the cumulative posterior model probability. The regressors are ordered in descending order based on their posterior inclusion probabilities. The included regressors with positive signs are displayed in blue (dark in grayscale) colour and with negative signs in red (light in grayscale) colour. The Regressors not included in the model are left without any colour.

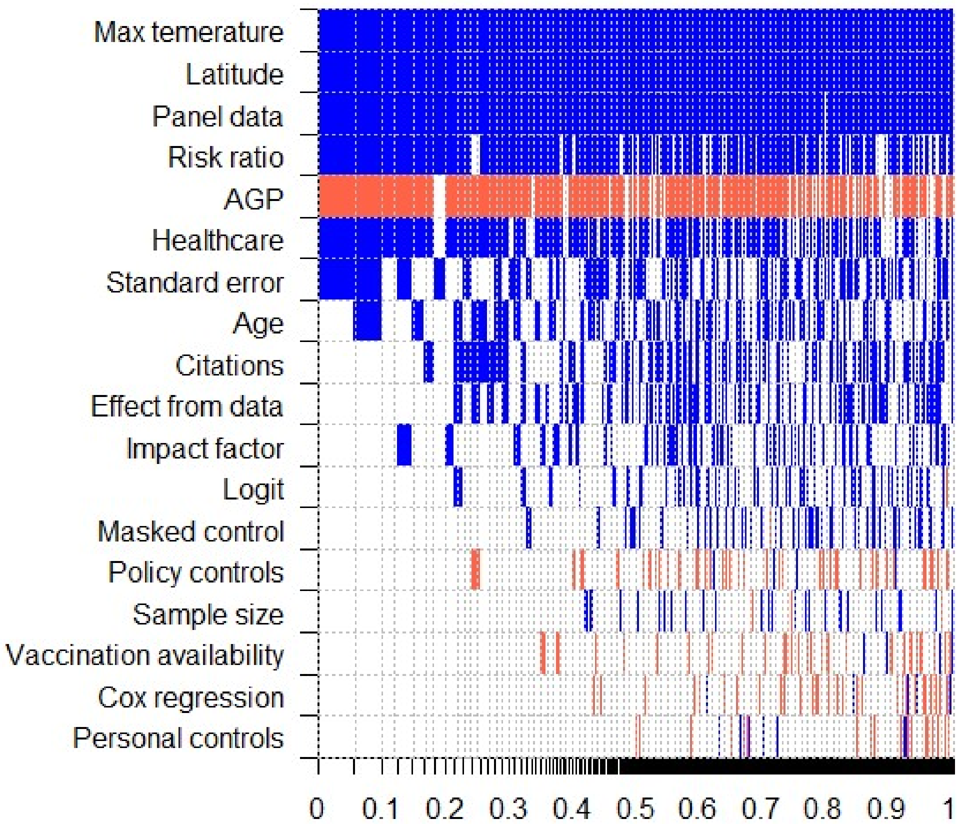
Additionally, the AGP variable has a negative effect. According to the present author, it can be interpreted as follows: using a face mask during procedures that generate aerosols is essential for decreasing the risk of infection. The risk estimated in the form of RR seems to be systematically higher (lower protection of masks). Lastly, the posterior inclusion probability for the standard error is just above the 0.5 bound for BMA with the uniform prior. For BMA with the dilution prior standard error did not cross the 0.5 bound.

Moreover, to provide an explanation of posterior inclusion probability and its meaning, we would like to present the following scale by Kass & Raftery (1995):

- PIP $\in [0.5, 0.75)$ - week evidence
- PIP $\in [0.75, 0.9)$ - positive effect
- PIP $\in [0.9, 0.99)$ - strong effect
- PIP $\in [0.99, 1)$ - decisive effect

Thus, the posterior inclusion probability can be considered analogous to the statistical significance of a variable. As apparent from the Table 5.3 the results

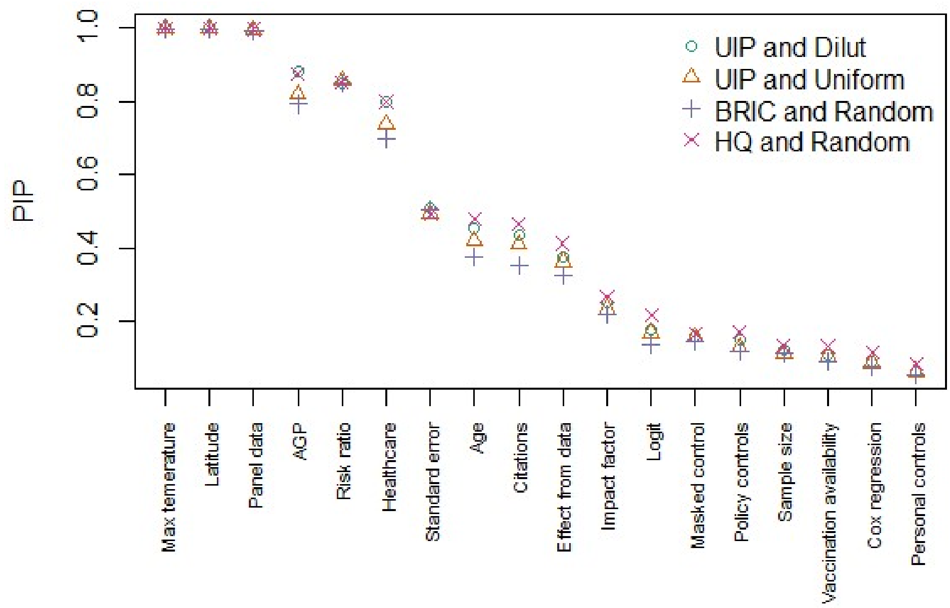
Figure 5.2: BMA with a dilution model prior and unit information g-prior



Note: The figure shows the Bayesian Model Averaging with the dilution model prior and unit information g-prior. The response variable is the risk of Covid-19 infection. The horizontal axis represents the cumulative posterior model probability. The regressors are ordered in descending order based on their posterior inclusion probabilities. The included regressors with positive signs are displayed in blue (dark in grayscale) colour and with negative signs in red (light in grayscale) colour. The Regressors not included in the model are left without any colour.

of all averaging methods are comparable. In addition to the already presented models, we performed BMA with different g-priors and model priors. The reader can access them in Appendix C. Furthermore, Figure 5.3 graphically compares the posterior inclusion probabilities of all variables among the four performed BMA models. We can conclude that the posterior inclusion probabilities are similar for all of the models and no apparent differences are present.

Figure 5.3: Comparison of Posterior Inclusion Probability for performed BMA models



Note: The figure shows the Posterior Inclusion Probability on the vertical axis for all regressors on the horizontal axis. UIP and Dilut = unit information g-prior and dilution model prior, UIP and Uniform = unit information g-prior and uniform model prior, BRIC and Random = benchmark g-prior and random theta model prior, HQ and Random = $log(n)^3$ g-prior and random theta model prior.

Chapter 6

The best practice estimate

After examination of publication bias and heterogeneity of the estimates, we would like to derive the best practice estimate. The best practice estimate can be seen as the bottom line of the meta-analysis.

Firstly, we consider it necessary to mention that deriving the best practice estimate is a subjective process. The process reflects the opinions and knowledge of the present author acquired by studying the papers included in this thesis and corresponding literature. For the subjective best practice estimate we set the values as follows. The standard error was set to zero because we would like to know the value of the effect after correcting for publication bias. The variable effect from data represents the effect calculated from data only. We would prefer the effect to be estimated by a regression method. Thus we set the variable equal to zero. The personal control and policy control dummy variables were put equal to one because we would like to know the effect of face masks after controlling for other policies and characteristics of subjects. Since the large sample size provides precise estimates, the sample size was set to the maximum value. The panel data structure would be also prioritised. The impact factor and the number of citations were set to their maxima to derive a best practice estimate of high trustworthiness. Also, the estimate should reflect the current state, thus we set the vaccination availability to one. Lastly, it would be preferred to have a control group not masked at all. The rest of the variables were set to their means.

In addition to the subjective best practice estimate, we derive the estimates of the prominent studies on the effect of face masks on Covid-19 transmission corrected for publication bias and misspecifications. Firstly, we chose the study by Karaivanov *et al.* (2021). The study was published in the Journal of Health

Economics with over 130 citations in Google Scholar. The econometric models estimated in the study are described in great detail together with all included control variables. Compared to the majority of studies regarding the effect of face masks on Covid-19 transmission, the study stands out for its transparency. The derived best practice estimate equals -0.136.

Secondly, the best practice estimate was derived for the study by Bundgaard *et al.* (2021). The reason for choosing the study is its random trial design with proper randomisation of the control and treatment groups. The study was published in the Annals of Internal Medicine with over 350 citations in Google Scholar. The derived best practice estimate equals -0.129. The third study, for which we derived the best practice estimate is the one by Nguyen *et al.* (2020). In contrast to the first two chosen studies, this one is performed in a healthcare setting. In addition, it was published in The Lancet Public Health journal and has over 2,000 citations in Google Scholar. Its best practice estimate equals -0.157. All of the best practice estimates derived for the studies are negative including the upper bounds of their 95% confidence intervals (Table 6.1).

Table 6.1: Best practice estimates

Study	Best practice estimate	95%CI
Subjective	-0.141	(-0.308, 0.025)
Karaivanov *et al.* (2021)	-0.136	(-0.161, -0.111)
Bundgaard *et al.* (2021)	-0.129	(-0.229, -0.030)
Nguyen *et al.* (2020)	-0.157	(-0.288, -0.025)

Chapter 7

Conclusion

One might think that the effect of face masks on Covid-19 transmission is strictly a medical topic. However, we would like to emphasise its economic consequences. The Covid-19 pandemic and related social-distancing measures caused a sharp decline in the GDP of major economies (Jena *et al.* 2021). Bagepally *et al.* (2021) suggest that the costs associated with the surgical mask-wearing amount to almost one billion USD. Resulting in avoiding more than 1,100 per million cases of Covid-19. However, these costs depend on the value of the true unbiased effect. In addition, the results of this thesis could be important for policymakers.

This thesis performs a meta-analysis on the effect of face masks on Covid transmission. We collected 258 estimates of the effect from 44 studies together with corresponding variables on the methodology and effect type, study set-up, data, country and individual, and publication characteristics. Together more than 9,300 data points were collected. Firstly, we examine the publication bias by employing many modern tests. The performed methods can be divided into three categories. The linear methods for publication bias detection include a graphical method: funnel plot (Egger *et al.* 1997) and numerical FAT-PET with different weights. We performed non-linear tests such as endogenous kink model (Bom & Rachinger 2019), stem-based method (Furukawa 2019), selection model (Andrews & Kasy 2019) and more. The last category includes the methods allowing for endogeneity: FAT-PET with instrumental variable, p-uniform* method and Caliper test. As a result, only some of the tests yielded significant estimates of publication bias. Nevertheless, these significant estimates imply only mild evidence of publication bias. Such a result is in line with Chu *et al.* (2020). Apart from the detection of publication bias, these methods estimate the effect beyond bias. The estimate was statistically significant for almost all of the methods ranging from -0.187 to -0.440 which can be interpreted as face masks being effective in reducing the risk of Covid-19 infection by 18.7% to 44%. In contrast, Chu *et al.* (2020) found a more protective effect of face masks. Nevertheless, Jefferson *et al.* (2023) found the protective effect of masks to be small to none.

In the second part of the thesis, we focused on model averaging to examine the heterogeneity. We performed Frequentist and Bayesian model averaging with different priors. The purpose of implementing the averaging method is to identify the important variables influencing the effect of face masks on Covid-19 transmission. 18 out of 26 eligible variables were used for the averaging. We found the following variables to have a positive effect on the risk of transmission associated with mask-wearing (decreasing the effectiveness of masks): the temperature, geographical latitude, panel data structure, risk ratio estimate type, healthcare set-up, standard error and age. Performing aerosol-generating procedures have a negative effect on risk (increasing the effectiveness of masks). Unfortunately, we cannot compare these results to the results of other authors, since they used a different design for their meta-analyses and did not evaluate the heterogeneity and its drivers in greater detail. Nevertheless, the results are in line with what we expected to find. In addition, performed robustness checks were yielding very similar outcomes. As a bottom line of this thesis, we derived the best practice estimate representing the effect of prominent studies after correcting for publication bias and misspecifications. The derived best practice estimates ranged from -0.129 to -0.157. Meaning the masks reduce the risk of transmission by 12.9% to 15.7% for the set-ups of these studies.

Lastly, we would like to present some drawbacks. In spite of including 44 primary studies in the meta-analysis, we were able to collect only above 250 estimates. In addition, we were not able to collect specific controls included in the models of primary studies. This issue is caused by the low transparency of medical studies. We at least collected dummy variables for policy and personal characteristics controls. The two dummy variables were not statistically significant. Lastly, an extension of this thesis could be performing a cost-effectiveness analysis as presented by Bagepally *et al.* (2021) with the adjusted estimate of the effect of face masks on Covid-19 transmission.

Bibliography

Abaluck, J., L. H. Kwong, A. Styczynski, A. Haque, M. A. Kabir, E. Bates-Jefferys, E. Crawford, J. Benjamin-Chung, S. Raihan, S. Rahman, S. Benhachmi, N. Z. Bintee, P. J. Winch, M. Hossain, H. M. Reza, A. A. Jaber, S. G. Momen, A. Rahman, F. L. Banti, T. S. Huq, S. P. Luby, & A. M. Mobarak (2022): "Impact of community masking on covid-19: A cluster-randomized trial in bangladesh." *Science* **375(6577)**: p. eabi9069.

van Aert, R. C. & M. Van Assen (2021): "Correcting for publication bias in a meta-analysis with the p-uniform* method." .

Akinbami, L. J., N. Vuong, L. R. Petersen, S. Sami, A. Patel, S. L. Lukacs, L. Mackey, L. A. Grohskopf, A. Shehu, & J. Atas (2020): "Sars-cov-2 seroprevalence among healthcare, first response, and public safety personnel, detroit metropolitan area, michigan, usa, may–june 2020." *Emerging infectious diseases* **26(12)**: p. 2863.

Allenby, G. M. & P. E. Rossi (2006): "Hierarchical bayes models." *The handbook of marketing research: Uses, misuses, and future advances* pp. 418–440.

Altman, D. G. (1990): *Practical statistics for medical research.* CRC press.

Amini, S. M. & C. F. Parmeter (2012): "Comparison of model averaging techniques: Assessing growth determinants." *Journal of Applied Econometrics* **27(5)**: pp. 870–876.

Andrejko, K. L., J. M. Pry, J. F. Myers, N. Fukui, J. L. DeGuzman, J. Openshaw, J. P. Watt, J. A. Lewnard, S. Jain, C. COVID *et al.* (2022): "Effectiveness of face mask or respirator use in indoor public settings for prevention of sars-cov-2 infection—california, february–december 2021." *Morbidity and Mortality Weekly Report* **71(6)**: p. 212.

Andrews, I. & M. Kasy (2019): "Identification of and correction for publication bias." *American Economic Review* **109(8)**: pp. 2766–94.

Bagepally, B. S., M. Haridoss, M. Natarajan, K. Jeyashree, & M. Ponnaiah (2021): "Cost-effectiveness of surgical mask, N-95 respirator, hand-hygiene and surgical mask with hand hygiene in the prevention of COVID-19: Cost effectiveness analysis from Indian context." *Clinical Epidemiology and Global Health* **10**: p. 100702.

Bom, P. R. & H. Rachinger (2019): "A kinked meta-regression model for publication bias correction." *Research synthesis methods* **10(4)**: pp. 497–514.

Brodeur, A., N. Cook, & A. Heyes (2018): "Methods matter: P-hacking and causal inference in economics and finance." *IZA DP* **(11796)**.

Brodeur, A., D. Gray, A. Islam, & S. Bhuiyan (2021): "A literature review of the economics of covid-19." *Journal of Economic Surveys* **35(4)**: pp. 1007–1044.

Brodeur, A., M. Lé, M. Sangnier, & Y. Zylberberg (2016): "Star wars: The empirics strike back." *American Economic Journal: Applied Economics* **8(1)**: pp. 1–32.

van den Broek-Altenburg, E. M., A. J. Atherly, S. A. Diehl, K. M. Gleason, V. C. Hart, C. D. MacLean, D. A. Barkhuff, M. A. Levine, & J. K. Carney (2021): "Jobs, housing, and mask wearing: cross-sectional study of risk factors for covid-19." *JMIR Public Health and Surveillance* **7(1)**: p. e24320.

Brooks-Pollock, E., J. M. Read, A. R. McLean, M. J. Keeling, & L. Danon (2021): "Mapping social distancing measures to the reproduction number for covid-19." *Philosophical Transactions of the Royal Society B* **376(1829)**: p. 20200276.

Budzyn, S. E., M. J. Panaggio, S. E. Parks, M. Papazian, J. Magid, M. Eng, & L. C. Barrios (2021): "Pediatric covid-19 cases in counties with and without school mask requirements—united states, july 1–september 4, 2021." *Morbidity and Mortality Weekly Report* **70(39)**: p. 1377.

Bundgaard, H., J. S. Bundgaard, D. E. T. Raaschou-Pedersen, C. von Buchwald, T. Todsen, J. B. Norsk, M. M. Pries-Heje, C. R. Vissing,

P. B. Nielsen, U. C. Winsløw *et al.* (2021): "Effectiveness of adding a mask recommendation to other public health measures to prevent sars-cov-2 infection in danish mask wearers: a randomized controlled trial." *Annals of internal medicine* **174(3)**: pp. 335–343.

Burke, R. M., S. Balter, E. Barnes, V. Barry, K. Bartlett, K. D. Beer, I. Benowitz, H. M. Biggs, H. Bruce, J. Bryant-Genevier *et al.* (2020): "Enhanced contact investigations for nine early travel-related cases of sars-cov-2 in the united states." *PloS one* **15(9)**: p. e0238342.

Cameron, A. C. & P. K. Trivedi (2005): *Microeconometrics: methods and applications.* Cambridge university press.

Catching, A., S. Capponi, M. T. Yeh, S. Bianco, & R. Andino (2021): "Examining the interplay between face mask usage, asymptomatic transmission, and social distancing on the spread of covid-19." *Scientific reports* **11(1)**: pp. 1–11.

Chaabna, K., S. Doraiswamy, R. Mamtani, & S. Cheema (2021): "Facemask use in community settings to prevent respiratory infection transmission: A rapid review and meta-analysis." *International Journal of Infectious Diseases* **104**: pp. 198–206.

Chatterjee, P., T. Anand, K. J. Singh, R. Rasaily, R. Singh, S. Das, H. Singh, I. Praharaj, R. R. Gangakhedkar, B. Bhargava *et al.* (2020): "Healthcare workers & sars-cov-2 infection in india: A case-control investigation in the time of covid-19." *The Indian journal of medical research* **151(5)**: p. 459.

Chaudhry, M. R. A. (2022): "Chapter 5 - coronavirus infection outbreak: comparison with other viral infection outbreak." In A. I. Qureshi, O. Saeed, & U. Syed (editors), "Coronavirus Disease," pp. 47–57. Academic Press.

Chen, Y., X. Tong, J. Wang, W. Huang, S. Yin, R. Huang, H. Yang, Y. Chen, A. Huang, Y. Liu *et al.* (2020): "High sars-cov-2 antibody prevalence among healthcare workers exposed to covid-19 patients." *Journal of Infection* **81(3)**: pp. 420–426.

Cheng, H.-Y., S.-W. Jian, D.-P. Liu, T.-C. Ng, W.-T. Huang, , & H.-H. Lin (2020): "High transmissibility of covid-19 near symptom onset." https://doi.org/10.1101/2020.03.18.20034561.

Chernozhukov, V., H. Kasahara, & P. Schrimpf (2021): "Causal impact of masks, policies, behavior on early covid-19 pandemic in the us." *Journal of econometrics* **220(1)**: pp. 23–62.

Chu, D. K., E. A. Akl, S. Duda, K. Solo, S. Yaacoub, H. J. Schünemann, A. El-Harakeh, A. Bognanni, T. Lotfi, M. Loeb *et al.* (2020): "Physical distancing, face masks, and eye protection to prevent person-to-person transmission of sars-cov-2 and covid-19: a systematic review and meta-analysis." *The lancet* **395(10242)**: pp. 1973–1987.

Ciotti, M., M. Ciccozzi, A. Terrinoni, W.-C. Jiang, C.-B. Wang, & S. Bernardini (2020): "The covid-19 pandemic." *Critical reviews in clinical laboratory sciences* **57(6)**: pp. 365–388.

Covid19 vaccine tracker (2022): "11 vaccines granted emergency use listing (eul) by who." https://covid19.trackvaccines.org/agency/who/, Last accessed on 2022-12-30.

Cox, D. R. (1972): "Regression models and life-tables." *Journal of the Royal Statistical Society: Series B (Methodological)* **34(2)**: pp. 187–202.

Davido, B., S. Gautier, I. Riom, S. Landowski, C. Lawrence, A. Thiebaut, S. Bessis, V. Perronne, H. Mascitti, L. Noussair *et al.* (2021): "The first wave of covid-19 in hospital staff members of a tertiary care hospital in the greater paris area: A surveillance and risk factors study." *International Journal of Infectious Diseases* **105**: pp. 172–179.

Dörr, T., S. Haller, M. F. Müller, A. Friedl, D. Vuichard, C. R. Kahlert, & P. Kohler (2022): "Risk of sars-cov-2 acquisition in health care workers according to cumulative patient exposure and preferred mask type." *JAMA Network Open* **5(8)**: pp. e2226816–e2226816.

Doucouliagos, C. & T. D. Stanley (2013): "Are all economic facts greatly exaggerated? theory competition and selectivity." *Journal of Economic Surveys* **27(2)**: pp. 316–339.

Doung-Ngern, P., R. Suphanchaimat, A. Panjangampatthana, C. Janekrongtham, D. Ruampoom, N. Daochaeng, N. Eungkanit, N. Pisitpayat, N. Srisong, O. Yasopa *et al.* (2020): "Case-control study of use of personal protective measures and risk for sars-cov 2 infection, thailand." *Emerging infectious diseases* **26(11)**: p. 2607.

Dôvera (2022): "Covid-19 v číslach." `https://www.dovera.sk/poistenec/potrebujem-poradit/o-dovere/statistiky/covid-19-v-cislach`, Last accessed on 2022-12-30.

Egger, M., G. D. Smith, M. Schneider, & C. Minder (1997): "Bias in meta-analysis detected by a simple, graphical test." *Bmj* **315(7109)**: pp. 629–634.

Eicher, T. S., C. Papageorgiou, & A. E. Raftery (2011): "Default priors and predictive performance in bayesian model averaging, with application to growth determinants." *Journal of Applied Econometrics* **26(1)**: pp. 30–55.

Eikenberry, S. E., M. Mancuso, E. Iboi, T. Phan, K. Eikenberry, Y. Kuang, E. Kostelich, & A. B. Gumel (2020): "To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the covid-19 pandemic." *Infectious disease modelling* **5**: pp. 293–308.

European Medicines Agency (2022): "Covid-19 vaccines: authorised." `https://www.ema.europa.eu/en/human-regulatory/overview/public-health-threats/coronavirus-disease-covid-19/treatments-vaccines/vaccines-covid-19/covid-19-vaccines-authorised`, Last accessed on 2022-12-30.

Fawcett, S. E., M. S. Madhusudhan, E. N. Gaddam, M. J. Almario, S. R. Masih, D. D. Klute-Evans, J. C. Johnson, C. D. Stroud, J. A. Dolan-Caren, M. A. Ben-Aderet, & et al. (2023): "Transmission risk of severe acute respiratory coronavirus virus 2 (sars-cov-2) to healthcare personnel following unanticipated exposure to aerosol-generating procedures: Experience from epidemiologic investigations at an academic medical center." *Infection Control amp; Hospital Epidemiology* **44(2)**: p. 325–327.

Fletcher, J. J., E. C. Feucht, P. Y. Hahn, T. N. McGoff, D. J. Dehart, M. E. El Mortada, & R. G. Grifka (2022): "Healthcare-acquired coronavirus disease 2019 (covid-19) is less symptomatic than community-acquired disease among healthcare workers." *Infection Control & Hospital Epidemiology* **43(4)**: pp. 490–496.

Fox, J. & S. Weisberg (2002): "Cox proportional-hazards regression for survival data." *An R and S-PLUS companion to applied regression* **2002**.

Furukawa, C. (2019): "Publication bias under aggregation frictions: Theory, evidence, and a new correction method." *Evidence, and a New Correction Method (March 29, 2019)* .

Gechert, S., T. Havranek, Z. Irsova, & D. Kolcunova (2022): "Measuring capital-labor substitution: The importance of method choices and publication bias." *Review of Economic Dynamics* **45**: pp. 55–82.

George, E. I. *et al.* (2010): "Dilution priors: Compensating for model space redundancy." *Borrowing Strength: Theory Powering Applications–A Festschrift for Lawrence D. Brown* **6**: pp. 158–165.

Gerber, A., N. Malhotra *et al.* (2008): "Do statistical reporting standards affect what is published? publication bias in two leading political science journals." *Quarterly Journal of Political Science* **3(3)**: pp. 313–326.

Gonçalves, M. R., R. C. P. Dos Reis, R. P. Tólio, L. C. Pellanda, M. I. Schmidt, N. Katz, S. S. Mengue, P. C. Hallal, B. L. Horta, M. F. Silveira *et al.* (2021): "Social distancing, mask use, and transmission of severe acute respiratory syndrome coronavirus 2, brazil, april–june 2020." *Emerging infectious diseases* **27(8)**: p. 2135.

Grambsch, P. M. (1995): "Goodness-of-fit and diagnostics for proportional hazards regression models." *Recent Advances in Clinical Trial Design and Analysis* pp. 95–112.

Grambsch, P. M. & T. M. Therneau (1994): "Proportional hazards tests and diagnostics based on weighted residuals." *Biometrika* **81(3)**: pp. 515–526.

Greene, W. H. (2018): *Econometric Analysis, 8th edition.* Pearson.

Guo, X., J. Wang, D. Hu, L. Wu, L. Gu, Y. Wang, J. Zhao, L. Zeng, J. Zhang, & Y. Wu (2020a): "Survey of covid-19 disease among orthopaedic surgeons in wuhan, people's republic of china." *The Journal of bone and joint surgery. American volume* .

Guo, Y.-R., Q.-D. Cao, Z.-S. Hong, Y.-Y. Tan, S.-D. Chen, H.-J. Jin, K.-S. Tan, D.-Y. Wang, & Y. Yan (2020b): "The origin, transmission and clinical therapies on coronavirus disease 2019 (covid-19) outbreak–an update on the status." *Military medical research* **7(1)**: pp. 1–10.

Guy Jr, G. P., F. C. Lee, G. Sunshine, R. McCord, M. Howard-Williams, L. Kompaniyets, C. Dunphy, M. Gakh, R. Weber, E. Sauber-Schatz *et al.* (2021): "Association of state-issued mask mandates and allowing on-premises restaurant dining with county-level covid-19 case and death growth rates—united states, march 1–december 31, 2020." *Morbidity and Mortality Weekly Report* **70(10)**: p. 350.

Haller, S., S. Güsewell, T. Egger, G. Scanferla, R. Thoma, O. B. Leal-Neto, D. Flury, A. Brucher, E. Lemmenmeier, J. C. Möller *et al.* (2022): "Impact of respirator versus surgical masks on sars-cov-2 acquisition in healthcare workers: a prospective multicentre cohort." *Antimicrobial Resistance & Infection Control* **11(1)**: pp. 1–11.

Hansen, B. E. (2007): "Least squares model averaging." *Econometrica* **75(4)**: pp. 1175–1189.

Havranek, T., Z. Irsova, L. Laslopova, & O. Zeynalova (2021): "Skilled and unskilled labor are less substitutable than commonly thought." .

Havranek, T., Z. Irsova, & O. Zeynalova (2018): "Tuition fees and university enrolment: a meta-regression analysis." *Oxford Bulletin of Economics and Statistics* **80(6)**: pp. 1145–1184.

Havranek, T. & A. Sokolova (2020): "Do consumers really follow a rule of thumb? three thousand estimates from 144 studies say "probably not"." *Review of Economic Dynamics* **35**: pp. 97–122.

Havránek, T., T. D. Stanley, H. Doucouliagos, P. Bom, J. Geyer-Klingeberg, I. Iwasaki, W. R. Reed, K. Rost, & R. C. van Aert (2020): "Reporting guidelines for meta-analysis in economics." *Journal of Economic Surveys* **34(3)**: pp. 469–475.

Heinzerling, A., M. J. Stuckey, T. Scheuer, K. Xu, K. M. Perkins, H. Resseger, S. Magill, J. R. Verani, S. Jain, M. Acosta *et al.* (2020): "Transmission of covid-19 to health care personnel during exposures to a hospitalized patient—solano county, california, february 2020." *Morbidity and Mortality Weekly Report* **69(15)**: p. 472.

Higgins, J. P. T., J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, & V. A. Welch (2019): *Cochrane Handbook for Systematic Reviews of Interventions*. John Wiley & Sons. Google-Books-ID: cTqyDwAAQBAJ.

Holland, P. W. (1989): "A note on the covariance of the mantel-haenszel log-odds ratio estimator and the sample marginal rates." *ETS Research Report Series* **1989(1)**: pp. i–11.

Ioannidis, J. P., T. D. Stanley, & H. Doucouliagos (2017): "The power of bias in economics research."

Jefferson, T., L. Dooley, E. Ferroni, L. A. Al-Ansary, M. L. van Driel, G. A. Bawazeer, M. A. Jones, T. C. Hoffmann, J. Clark, E. M. Beller *et al.* (2023): "Physical interventions to interrupt or reduce the spread of respiratory viruses." `https://doi.org/10.1002/14651858.CD006207.pub6`.

Jehn, M., J. Mac McCullough, A. P. Dale, M. Gue, B. Eller, T. Cullen, & S. E. Scott (2021): "Association between k–12 school mask policies and school-associated covid-19 outbreaks—maricopa and pima counties, arizona, july–august 2021." *Morbidity and Mortality Weekly Report* **70(39)**: p. 1372.

Jena, P. R., R. Majhi, R. Kalli, S. Managi, & B. Majhi (2021): "Impact of covid-19 on gdp of major economies: Application of the artificial neural network forecaster." *Economic Analysis and Policy* **69**: pp. 324–339.

Joo, H., G. F. Miller, G. Sunshine, M. Gakh, J. Pike, F. P. Havers, L. Kim, R. Weber, S. Dugmeoglu, C. Watson *et al.* (2021): "Decline in covid-19 hospitalization growth rates associated with statewide mask mandates—10 states, march–october 2020." *Morbidity and Mortality Weekly Report* **70(6)**: p. 212.

Kahlert, C. R., R. Persi, S. Güsewell, T. Egger, O. B. Leal-Neto, J. Sumer, D. Flury, A. Brucher, E. Lemmenmeier, J. C. Möller *et al.*

(2021): "Non-occupational and occupational factors associated with specific sars-cov-2 antibodies among hospital workers–a multicentre cross-sectional study." *Clinical Microbiology and Infection* **27(9)**: pp. 1336–1344.

Karaivanov, A., S. E. Lu, H. Shigeoka, C. Chen, & S. Pamplona (2021): "Face masks, public policies and slowing the spread of covid-19: Evidence from canada." *Journal of Health Economics* **78**: p. 102475.

Kass, R. E. & A. E. Raftery (1995): "Bayes factors." *Journal of the american statistical association* **90(430)**: pp. 773–795.

Kelly, J., T. Sadeghieh, & K. Adeli (2014): "Peer review in scientific publications: benefits, critiques, & a survival guide." *Ejifcc* **25(3)**: p. 227.

Khalil, M., M. M. Alam, M. K. Arefin, M. R. Chowdhury, M. R. Huq, J. A. Chowdhury, A. M. Khan *et al.* (2020): "Role of personal protective measures in prevention of covid-19 spread among physicians in bangladesh: a multicenter cross-sectional comparative study." *SN comprehensive clinical medicine* **2(10)**: pp. 1733–1739.

Li, A., J. Slezak, A. M. Maldonado, J. Concepcion, C. V. Maier, & G. Rieg (2021a): "Sars-cov-2 positivity and mask utilization among health care workers." *JAMA Network Open* **4(6)**: pp. e2114325–e2114325.

Li, Y., M. Liang, L. Gao, M. Ayaz Ahmed, J. P. Uy, C. Cheng, Q. Zhou, & C. Sun (2021b): "Face masks to prevent transmission of covid-19: A systematic review and meta-analysis." *American Journal of Infection Control* **49(7)**: pp. 900–906.

Liang, M., L. Gao, C. Cheng, Q. Zhou, J. P. Uy, K. Heiner, & C. Sun (2020): "Efficacy of face mask in preventing respiratory virus transmission: A systematic review and meta-analysis." *Travel Medicine and Infectious Disease* **36**: p. 101751.

Lio, C. F., H. H. Cheong, C. I. Lei, I. L. Lo, L. Yao, C. Lam, & I. H. Leong (2021): "Effectiveness of personal protective health behaviour against covid-19." *BMC Public Health* **21(1)**: p. 827.

Liu, Y., A. A. Gayle, A. Wilder-Smith, & J. Rocklöv (2020a): "The reproductive number of covid-19 is higher compared to sars coronavirus." *Journal of travel medicine* **27(2)**.

Liu, Y., Z. Ning, Y. Chen, M. Guo, Y. Liu, N. K. Gali, L. Sun, Y. Duan, J. Cai, D. Westerdahl *et al.* (2020b): "Aerodynamic analysis of sars-cov-2 in two wuhan hospitals." *Nature* **582(7813)**: pp. 557–560.

Loeb, M., A. Bartholomew, M. Hashmi, W. Tarhuni, M. Hassany, I. Youngster, R. Somayaji, O. Larios, J. Kim, B. Missaghi *et al.* (2022): "Medical masks versus n95 respirators for preventing covid-19 among health care workers: a randomized trial." *Annals of Internal Medicine* **175(12)**: pp. 1629–1638.

Lotfi, M., M. R. Hamblin, & N. Rezaei (2020): "Covid-19: Transmission, prevention, and potential therapeutic opportunities." *Clinica Chimica Acta* **508**: pp. 254–266.

Lu, R., X. Zhao, J. Li, P. Niu, B. Yang, H. Wu, W. Wang, H. Song, B. Huang, N. Zhu *et al.* (2020): "Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding." *The lancet* **395(10224)**: pp. 565–574.

Madigan, D., J. York, & D. Allard (1995): "Bayesian graphical models for discrete data." *International Statistical Review/Revue Internationale de Statistique* pp. 215–232.

Maltezou, H. C., X. Dedoukou, M. Tseroni, P. Tsonou, V. Raftopoulos, K. Papadima, E. Mouratidou, S. Poufta, G. Panagiotakopoulos, D. Hatzigeorgiou *et al.* (2020): "Sars-cov-2 infection in healthcare personnel with high-risk occupational exposure: evaluation of 7-day exclusion from work policy." *Clinical Infectious Diseases* **71(12)**: pp. 3182–3187.

Martischang, R., A. Iten, I. Arm, M. Abbas, B. Meyer, S. Yerly, I. Eckerle, J. Pralong, J. Sauser, J.-C. Suard *et al.* (2022): "Severe acute respiratory coronavirus virus 2 (sars-cov-2) seroconversion and occupational exposure of employees at a swiss university hospital: A large longitudinal cohort study." *Infection Control & Hospital Epidemiology* **43(3)**: pp. 326–333.

Mittal, R., C. Meneveau, & W. Wu (2020): "A mathematical framework for estimating risk of airborne transmission of covid-19 with application to face mask use and social distancing." *Physics of Fluids* **32(10)**: p. 101903.

Mitze, T., R. Kosfeld, J. Rode, & K. Wälde (2020): "Face masks considerably reduce covid-19 cases in germany." *Proceedings of the National Academy of Sciences* **117(51)**: pp. 32293–32301.

National Cancer Institute (2022a): "National cancer institute dictionary of cancer terms." `https://www.cancer.gov/publications/dictionaries/cancer-terms/def/hazard-ratio`, Last accessed on 2022-12-06.

National Cancer Institute (2022b): "National cancer institute dictionary of cancer terms." `https://www.cancer.gov/publications/dictionaries/cancer-terms/def/risk-ratio`, Last accessed on 2022-12-29.

National Cancer Institute (2022c): "National cancer institute dictionary of cancer terms." `https://www.cancer.gov/publications/dictionaries/cancer-terms/def/odds-ratio`, Last accessed on 2022-12-06.

National Cancer Institute (2022d): "National cancer institute dictionary of cancer terms." `https://www.cancer.gov/publications/dictionaries/cancer-terms/def/randomized-clinical-trial`, Last accessed on 2022-12-29.

Nelson, S. B., C. M. Dugdale, A. Bilinski, D. Cosar, N. R. Pollock, & A. Ciaranello (2021): "Prevalence and risk factors for in-school transmission of sars-cov-2 in massachusetts k-12 public schools, 2020-2021." `https://doi.org/10.1101/2021.09.22.21263900`.

Nguyen, L. H., D. A. Drew, M. S. Graham, A. D. Joshi, C.-G. Guo, W. Ma, R. S. Mehta, E. T. Warner, D. R. Sikavi, C.-H. Lo, S. Kwon, M. Song, L. A. Mucci, M. J. Stampfer, W. C. Willett, A. H. Eliassen, J. E. Hart, J. E. Chavarro, J. W. Rich-Edwards, R. Davies, J. Capdevila, K. A. Lee, M. N. Lochlainn, T. Varsavsky, C. H. Sudre, M. J. Cardoso, J. Wolf, T. D. Spector, S. Ourselin, C. J. Steves, A. T. Chan, C. M. Albert, G. Andreotti, B. Bala, B. A. Balasubramanian, L. E. Beane-Freeman, J. S. Brownstein, F. J. Bruinsma, J. Coresh, R. Costa, A. N. Cowan, A. Deka, S. L. Deming-Halverson, M. Elena Martinez, M. E. Ernst, J. C. Figueiredo, P. Fortuna, P. W. Franks, L. B. Freeman, C. D. Gardner, I. M. Ghobrial, C. A. Haiman, J. E. Hall, J. H. Kang, B. Kirpach, K. C. Koenen, L. D. Kubzansky, J. V. Lacey, Jr, L. Le Marchand, X. Lin, P. Lutsey, C. R. Marinac, M. E. Martinez, R. L. Milne, A. M. Murray, D. Nash,

J. R. Palmer, A. V. Patel, E. Pierce, M. M. Robertson, L. Rosenberg, D. P. Sandler, S. H. Schurman, K. Sewalk, S. V. Sharma, C. J. Sidey-Gibbons, L. Slevin, J. W. Smoller, C. J. Steves, M. I. Tiirikainen, S. T. Weiss, L. R. Wilkens, & F. Zhang (2020): "Risk of covid-19 among front-line health-care workers and the general community: a prospective cohort study." *The Lancet Public Health* **5(9)**: pp. e475–e483.

Notari, A. (2021): "Temperature dependence of covid-19 transmission." *Science of The Total Environment* **763**: p. 144390.

Page, M. J., J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, L. Shamseer, J. M. Tetzlaff, E. A. Akl, S. E. Brennan *et al.* (2021): "The prisma 2020 statement: an updated guideline for reporting systematic reviews." *International journal of surgery* **88**: p. 105906.

Payne, D. C., S. E. Smith-Jeffcoat, G. Nowak, U. Chukwuma, J. R. Geibe, R. J. Hawkins, J. A. Johnson, N. J. Thornburg, J. Schiffer, Z. Weiner *et al.* (2020): "Sars-cov-2 infections and serologic responses from a sample of us navy service members—uss theodore roosevelt, april 2020." *Morbidity and Mortality Weekly Report* **69(23)**: p. 714.

Piapan, L., P. De Michieli, F. Ronchese, F. Rui, M. Mauro, M. Peresson, L. Segat, P. D'Agaro, C. Negro, M. Bovenzi *et al.* (2020): "Covid-19 outbreak in healthcare workers in hospitals in trieste, north-east italy." *Journal of Hospital Infection* **106(3)**: pp. 626–628.

Pienthong, T., T. Khawcharoenporn, P. Apisarnthanarak, D. J. Weber, & A. Apisarnthanarak (2022): "Factors associated with coronavirus disease 2019 (covid-19) among thai healthcare personnel with high-risk exposures: The important roles of double masking and physical distancing while eating." *Infection Control & Hospital Epidemiology* **43(12)**: pp. 1978–1980.

Rebmann, T., T. M. Loux, L. D. Arnold, R. Charney, D. Horton, & A. Gomel (2021): "Sars-cov-2 transmission to masked and unmasked close contacts of university students with covid-19—st. louis, missouri, january–may 2021." *Morbidity and Mortality Weekly Report* **70(36)**: p. 1245.

Röhrig, B., J.-B. Du Prel, D. Wachtlin, & M. Blettner (2009): "Types of study in medical research: part 3 of a series on evaluation of scientific publications." *Deutsches Arzteblatt International* **106(15)**: p. 262.

Rothstein, H. R., A. J. Sutton, & M. Borenstein (2005): "Publication bias in meta-analysis." *Publication bias in meta-analysis: Prevention, assessment and adjustments* pp. 1–7.

Schmidt, C. O. & T. Kohlmann (2008): "When to use the odds ratio or the relative risk?" *International journal of public health* **53(3)**: p. 165.

Schoberer, D., S. Osmancevic, L. Reiter, N. Thonhofer, & M. Hoedl (2022): "Rapid review and meta-analysis of the effectiveness of personal protective equipment for healthcare workers during the covid-19 pandemic." *Public Health in Practice* **4**: p. 100280.

Sharif, N., K. J. Alzahrani, S. N. Ahmed, R. R. Opu, N. Ahmed, A. Talukder, R. Nunia, M. S. Chowdhury, I. J. Nodi, T. Saha *et al.* (2021): "Protective measures are associated with the reduction of transmission of covid-19 in bangladesh: A nationwide cross-sectional study." *PLoS One* **16(11)**: p. e0260287.

Shi, P., Y. Dong, H. Yan, C. Zhao, X. Li, W. Liu, M. He, S. Tang, & S. Xi (2020): "Impact of temperature on the dynamics of the covid-19 outbreak in china." *Science of the total environment* **728**: p. 138890.

Spruance, S. L., J. E. Reid, M. Grace, & M. Samore (2004): "Hazard ratio in clinical trials." *Antimicrobial agents and chemotherapy* **48(8)**: pp. 2787–2792.

Stanley, T. D. (2005): "Beyond publication bias." *Journal of economic surveys* **19(3)**: pp. 309–345.

Stanley, T. D., S. B. Jarrell, & H. Doucouliagos (2010): "Could it be better to discard 90% of the data? a statistical paradox." *The American Statistician* **64(1)**: pp. 70–77.

Steel, M. F. (2020): "Model averaging and its use in economics." *Journal of Economic Literature* **58(3)**: pp. 644–719.

Sugimura, M., O. Chimed-Ochir, Y. Yumiya, H. Ohge, N. Shime, T. Sakaguchi, J. Tanaka, T. Takafuta, M. Mimori, M. Kuwabara *et al.* (2021): "The association between wearing a mask and covid-19." *International Journal of Environmental Research and Public Health* **18(17)**: p. 9131.

Tabatabaeizadeh, S.-A. (2021): "Airborne transmission of covid-19 and the role of face mask to prevent it: a systematic review and meta-analysis." *European journal of medical research* **26(1)**: pp. 1–6.

Talic, S., S. Shah, H. Wild, D. Gasevic, A. Maharaj, Z. Ademi, X. Li, W. Xu, I. Mesa-Eguiagaray, J. Rostron, E. Theodoratou, X. Zhang, A. Motee, D. Liew, & D. Ilic (2021): "Effectiveness of public health measures in reducing the incidence of covid-19, sars-cov-2 transmission, and covid-19 mortality: systematic review and meta-analysis." *BMJ* **375**.

Tenny, S. & M. R. Hoffman (2017): *Odds Ratio*. StatPearls Publishing, Treasure Island (FL).

Toyokawa, T., T. Shimada, T. Hayamizu, T. Sekizuka, Y. Zukeyama, M. Yasuda, Y. Nakamura, S. Okano, J. Kudaka, T. Kakita *et al.* (2022): "Transmission of sars-cov-2 during a 2-h domestic flight to okinawa, japan, march 2020." *Influenza and Other Respiratory Viruses* **16(1)**: pp. 63–71.

Ueki, H., Y. Furusawa, K. Iwatsuki-Horimoto, M. Imai, H. Kabata, H. Nishimura, & Y. Kawaoka (2020): "Effectiveness of face masks in preventing airborne transmission of sars-cov-2." *mSphere* **5(5)**: pp. e00637–20.

Van Dyke, M. E., T. M. Rogers, E. Pevzner, C. L. Satterwhite, H. B. Shah, W. J. Beckman, F. Ahmed, D. C. Hunt, & J. Rule (2020): "Trends in county-level covid-19 incidence in counties with and without a mask mandate—kansas, june 1–august 23, 2020." *Morbidity and Mortality Weekly Report* **69(47)**: p. 1777.

Venugopal, U., N. Jilani, S. Rabah, M. A. Shariff, M. Jawed, A. M. Batres, M. Abubacker, S. Menon, A. Pillai, N. Shabarek *et al.* (2021): "Sars-cov-2 seroprevalence among health care workers in a new york city hospital: A cross-sectional analysis during the covid-19 pandemic." *International Journal of Infectious Diseases* **102**: pp. 63–69.

Wang, Q., X. Huang, Y. Bai, X. Wang, H. Wang, X. Hu, F. Wang, X. Wang, J. Chen, Q. Chen *et al.* (2020a): "Epidemiological characteristics of covid-19 in medical staff members of neurosurgery departments in hubei province: a multicentre descriptive study." https://doi.org/10.1101/2020.04.20.20064899.

Wang, X., X. Jiang, Q. Huang, H. Wang, D. Gurarie, M. Ndeffo-Mbah, F. Fan, P. Fu, M. A. Horn, A. Mondal, C. King, S. Xu, H. Zhao, & Y. Bai (2020b): "Risk factors of sars-cov-2 infection in healthcare workers: a retrospective study of a nosocomial outbreak." *Sleep Medicine: X* **2**: p. 100028.

Wang, X., Z. Pan, & Z. Cheng (2020c): "Association between 2019-ncov transmission and n95 respirator use." *Journal of Hospital Infection* **105(1)**: pp. 104–105.

Wang, Y., H. Tian, L. Zhang, M. Zhang, D. Guo, W. Wu, X. Zhang, G. L. Kan, L. Jia, D. Huo *et al.* (2020d): "Reduction of secondary transmission of sars-cov-2 in households by face mask use, disinfection and social distancing: a cohort study in beijing, china." *BMJ global health* **5(5)**: p. e002794.

Wilson, N., G. Marks, A. Eckhardt, A. Clarke, F. Young, F. Garden, W. Stewart, T. Cook, & E. Tovey (2021): "The effect of respiratory activity, non-invasive respiratory support and facemasks on aerosol generation and its relevance to covid-19." *Anaesthesia* **76(11)**: pp. 1465–1474.

Wohlin, C. (2014): "Guidelines for snowballing in systematic literature studies and a replication in software engineering." In "Proceedings of the 18th international conference on evaluation and assessment in software engineering," pp. 1–10.

World Health Organisation (2022a): "Coronavirus disease (covid-19)." https://www.who.int/health-topics/coronavirus#tab=tab_3, Last accessed on 2022-12-29.

World Health Organisation (2022b): "Coronavirus disease (covid-19) pandemic." https://www.who.int/emergencies/diseases/novel-coronavirus-2019, Last accessed on 2022-12-29.

World Health Organisation (2022c): "Who coronavirus (covid-19) dashboard overview." https://covid19.who.int/, Last accessed on 2022-12-29.

Xiao, M., L. Lin, J. S. Hodges, C. Xu, & H. Chu (2021): "Double-zero-event studies matter: A re-evaluation of physical distancing, face masks, and eye protection for preventing person-to-person transmission of covid-19 and its policy impact." *Journal of Clinical Epidemiology* **133**: pp. 158–160.

Zhang, J. & F. Y. Kai (1998): "What's the relative risk?: A method of correcting the odds ratio in cohort studies of common outcomes." *Jama* **280(19)**: pp. 1690–1691.

Zhang, J. & K. F. Yu (1998): "What's the Relative Risk? A Method of Correcting the Odds Ratio in Cohort Studies of Common Outcomes." *JAMA* **280(19)**: pp. 1690–1691.

Appendix A

Linear methods for publication bias: extention

Table A.1: Publication bias: linear methods extention

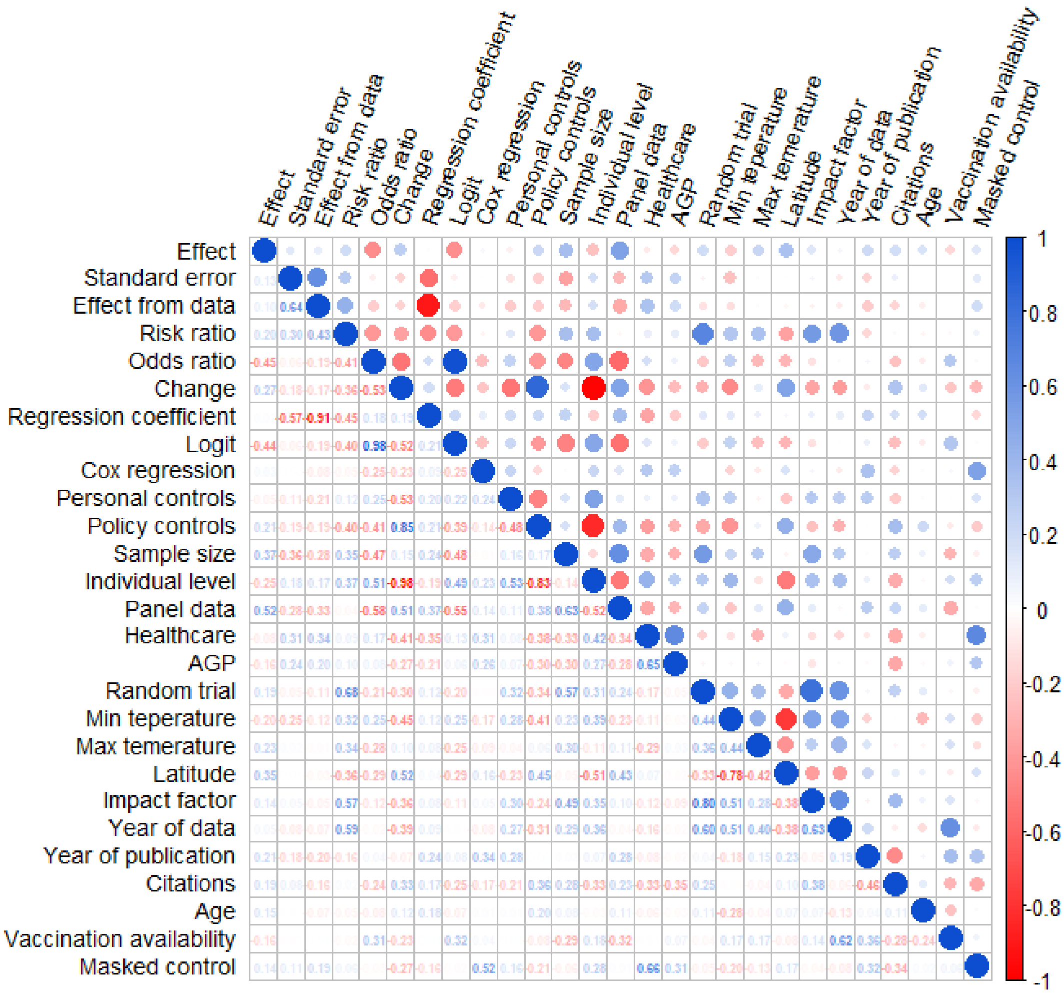
	OLS	**FE**	**BE**	**Study**	**Precision**
SE	-0.193*	-0.432***	-0.306	-0.255	-0.432
Publication bias	(0.162)	(0.001)	(0.280)	(0.175)	(2.211)
Constant	-0.253***	-0.187***	-0.243***	-0.170***	-0.187
Effect beyond bias	(0.024)	(0.001)	(0.032)	(0.015)	(0.158)
Studies	43	43	43	43	43
Observations	256	256	256	256	256

Note: The table displays linear methods for publication bias. Estimated on sub-sample that excluded estimates calculated from data. OLS = Ordinary Least Squares, FE = Fixed Effects, BE = Between Effects, Study = estimates were weighted by the inverse number of observations reported per study, Precision = estimates were weighted by the inverse of standard errors. Standard errors are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

Appendix B

Correlation coefficients table

Figure B.1: Correlation table for all eligible variables

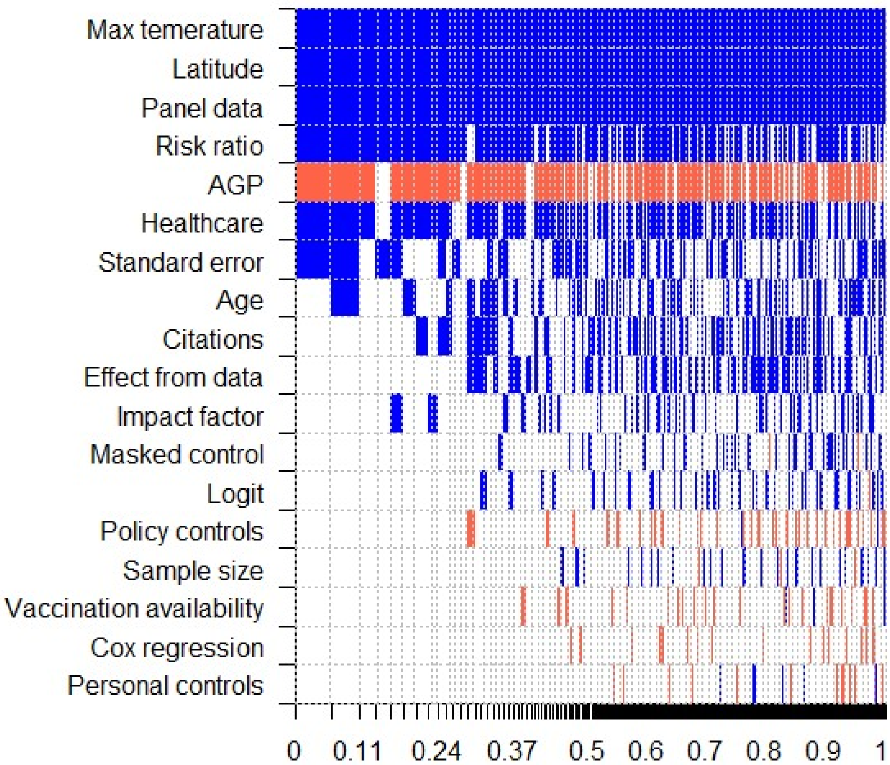


Note: The figure shows the correlation coefficients for all variables eligible for BMA, only 18 of these 27 variables were selected for the final model.

Appendix C

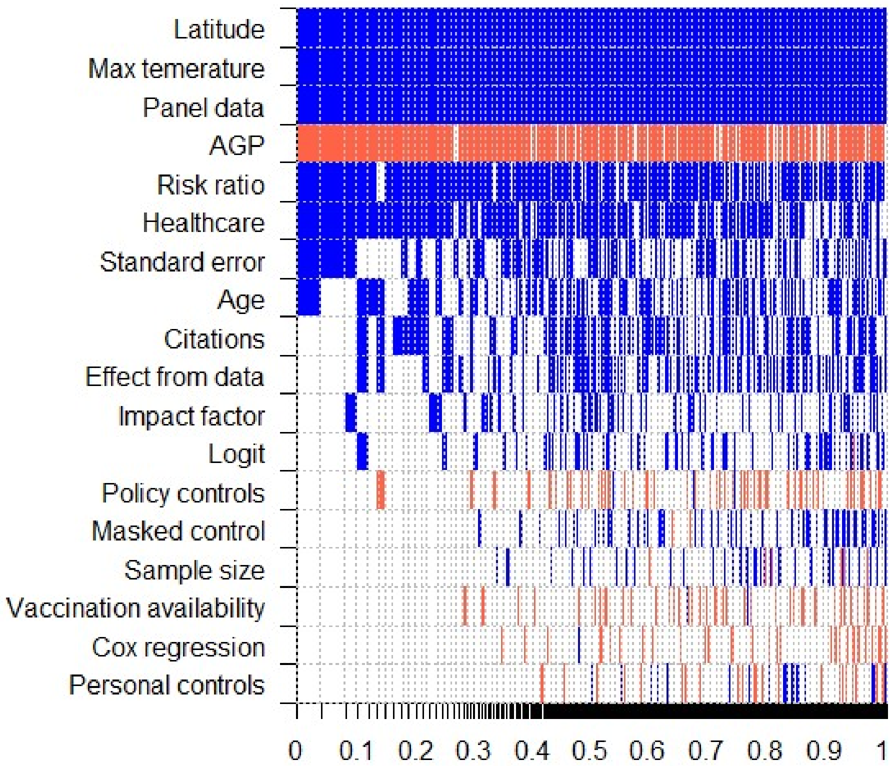
Robustness checks

Figure C.1: BMA with a benchmark g-prior and random theta model prior



Note: The figure shows the Bayesian Model Averaging with the benchmark g-prior and random theta model prior. The response variable is the risk of Covid-19 infection. The horizontal axis represents the cumulative posterior model probability. The regressors are ordered in descending order based on their posterior inclusion probabilities. The included regressors with positive signs are displayed in blue (dark in grayscale) colour and with negative signs in red (light in grayscale) colour. The Regressors not included in the model are left without any colour.

Figure C.2: BMA with a $log(n)^3$ g-prior and random theta model prior



Note: The figure shows the Bayesian Model Averaging with the $log(n)^3$ g-prior and random theta model prior. The response variable is the risk of Covid-19 infection. The horizontal axis represents the cumulative posterior model probability. The regressors are ordered in descending order based on their posterior inclusion probabilities. The included regressors with positive signs are displayed in blue (dark in grayscale) colour and with negative signs in red (light in grayscale) colour. The Regressors not included in the model are left without any colour.