

Publication Bias in Measuring Anthropogenic Climate Change

May 1, 2015

Abstract

We present a meta-regression analysis of the relation between the concentration of carbon dioxide in the atmosphere and changes in global temperature. The relation is captured by “climate sensitivity”, which measures the response to a doubling of carbon dioxide concentrations compared to pre-industrial levels. Estimates of climate sensitivity play a crucial role in evaluating the impacts of climate change and constitute one of the most important inputs into the computation of the social cost of carbon, which reflects the socially optimal value of a carbon tax. Climate sensitivity has been estimated by many researchers, but their results vary significantly. We collect 48 estimates from 16 studies and analyze the literature quantitatively. We find evidence for publication selection bias: researchers tend to report preferentially large estimates of climate sensitivity. Corrected for publication bias, the bulk of the literature is consistent with climate sensitivity lying between 1.4 and 2.3°C.

Keywords: Climate sensitivity, climate change, CO_2 , publication bias, meta-analysis

JEL Codes: Q53, Q54, C42

1 Introduction

Hundreds of researchers have tried to estimate the influence of human beings on climate change. We focus on estimates of equilibrium climate sensitivity, often simply termed climate sensitivity (CS), which is basically a measure of the climatic response to a doubling of CO_2 concentrations compared to pre-industrial levels (Solomon *et al.* 2007). These estimates play a crucial role in evaluating the impacts of anthropogenic climate change and constitute one of the most important inputs into the computation of the social cost of carbon, which reflects the socially optimal value of a carbon tax. Researchers report diverse results across studies, though the estimate of climate sensitivity most frequently oscillates around 3°C.

Our main objective is to find out, based on a collected data sample of published estimates, whether the reported estimates of climate sensitivity suffer from publication bias. No such analysis has previously been published. The 48 CS estimates collected from 16 studies range from 0.7 to 10.4, with a mean of 3.27. We summarize and quantify the reported estimates using

meta-regression analysis (MRA). To avoid possible problems in the MRA we focus on a narrow definition of climate sensitivity. The analysis is based on the assumption that the reported estimates are not correlated with their standard errors. Graphical tests reveal that such a relationship is present, indicating publication selection bias at first glance. As we cannot be sure about the true distribution of the CS estimates, we assume the standard normal distribution to be the best approximation. We provide a broader analysis by employing ordinary least squares (OLS), weighted least squares (WLS), fixed-effects (FE), and mixed-effects multilevel regressions of the CS estimates on their standard errors. We check for asymmetry of distribution of the CS estimates, which could give a false impression of publication bias, and also analyze subsets of median and mean CS estimates separately. Aside from that, we estimate the underlying effect of climate sensitivity corrected for publication bias.

The main contribution of this study is that it provides a quantitative survey of climate sensitivity estimates. Governments worldwide spend huge amounts of money on natural research, technological research and development, and so on to reduce greenhouse gas emissions and avoid man-made global warming. The results of this analysis could influence current policy decisions, which concentrate in first place on cutting CO_2 emissions.

The rest of this paper is organized as follows: Section 2 introduces the issue of climate sensitivity and publication bias; Section 3 summarizes data set collection; Section 4 describes the method of calculation of the standard error; Section 5 lists the regressions used in this analysis; Section 6 and Section 7 refer to methods for the detection of publication bias; Section 8 presents the results; and Section 9 concludes.

2 Climate Sensitivity

“The climate sensitivity is the equilibrium response of global surface temperature to a doubling of equivalent CO_2 concentration” (Houghton *et al.* 2001). This is the common definition of equilibrium climate sensitivity, but other sources provide different definitions. One defines equilibrium climate sensitivity as: *“the response in global-mean near-surface temperature to a doubling of atmospheric CO_2 concentrations from preindustrial levels”* (Klocke *et al.* 2011). Such discrepancies in the definition of climate sensitivity could damage the meta-analysis. When sampling estimates, we therefore focused on what exactly is predicted. Not all studies provide a definition of climate sensitivity, but many give the definition as the change from pre-industrial levels. The character of the data used in the studies indicates that the two definitions given above are saying the same thing and that the estimates collected are therefore all comparable with each other.

The issue of climate change, however, is more complicated, since CO_2 concentrations are not the only factor influencing the temperature change. According to Edwards *et al.* (2007) the size of forested area, ice melting, cloudiness, the frequency of extreme events, change in land cover, and other aspects can affect the global temperature. Some of these aspects can both warm and cool the atmosphere. Take clouds, for instance: low, white clouds reflect solar radiation back into space, thereby cooling the atmosphere, while high, dark clouds have exactly the opposite

effect. Besides that, the prediction remains uncertain due to imperfect knowledge of the ocean uptake of CO_2 , the terrestrial carbon cycle, and above all the sensitivity of the climate system to change.

Edwards *et al.* (2007) also quotes techniques for estimating climate sensitivity. These include inferring it directly from observations, comparing model simulations with observations, and weighting climate sensitivity predictions from several different models. The climate sensitivity predictions collected were estimated using either the comparison or weighting technique only.

Estimates of climate change and climate sensitivity occur only rarely in the scientific literature. For instance, the fifth assessment report of the Intergovernmental Panel on Climate Change (IPCC) predicts only that climate sensitivity probably ranges from 1.5 to 4.5 with high confidence and is extremely unlikely to be lower than 1, again with high confidence (Stocker 2013). For comparison, the IPCC's third assessment report estimates that climate sensitivity "likely" ranges between 2 and 4.5 and is "very unlikely" to be less than 1.5 (Pachauri & Reisinger 2007). Andronova & Schlesinger (2001) disagree with the third IPCC report and argue that climate sensitivity lies with 54% likelihood outside the IPCC range. They find that the 90% confidence interval for CS is 1 to 9.3.

Masters (2013) notes a robust relationship between the modeled rate of heat uptake in global oceans and the modeled climate sensitivity. This signals that researchers could have ways of influencing their results. We apply a common tool, meta-regression analysis, to detect publication bias in the literature about climate sensitivity. Michaels (2008) analyses 116 issues of two journals that forecast climate change: *Science* and *Nature*. Through vote-counting he found bias towards "worse" results.

Publication bias arises from the various motivations of different people. Both the scientist and the journal editor may only want to publish attractive or reliable results. Their motivations to bias the results or publish only selected results are similar: first, the selection of significant estimates (type II bias in the terminology of Stanley (2005)), and second, the selection of estimates with intuitive magnitude (type I bias). Publishing only selected results is called the "file drawer problem" (Rosenthal 1979). Although the selection of significant estimates is more benign (Stanley 2005), it still causes publication bias and precludes an accurate overview of the problem (De Long & Lang 1992).

3 Data

The collected studies estimating climate sensitivity are restricted to studies written in English.¹ Furthermore, to allow the use of modern meta-analysis methods this analysis only includes estimates with a reported standard error or a statistic from which the standard error can be computed. We collect all the estimates from the papers and also codify 13 variables reflecting the context in which researchers obtain their estimates of climate sensitivities, including information about the character of the estimate. The literature provides multiple types of climate sensitivity estimates. For the analysis, we use only one type of estimate from each study in this preference

¹Searches were performed in Web of Knowledge and Google Scholar.

order: mean, median, mode, best estimate. We add the last study to the data set on March 3, 2014 and terminate the search. The oldest study was published in 2001 and the most recent in 2013.

Table 1: **List of primary studies used**

Andronova & Schlesinger (2001)	Lindzen & Choi (2011)
Forest <i>et al.</i> (2006)	Murphy <i>et al.</i> (2004)
Frame <i>et al.</i> (2005)	Piani <i>et al.</i> (2005)
Gregory <i>et al.</i> (2002)	Scafetta (2013a)
Hargreaves & Annan (2009)	Scafetta (2013b)
Hegerl <i>et al.</i> (2006)	Schmittner <i>et al.</i> (2011)
Huber (2011)	Webb <i>et al.</i> (2006)
Knutti <i>et al.</i> (2006)	Wigley <i>et al.</i> (2005)

Notes: The search for primary studies was terminated on March 3, 2014.

Although some meta-analysts argue for using estimates from all available studies in order to avoid publication bias, we decide to not use estimates from unpublished papers, as the magnitude of any bias caused by failure to include unpublished papers has never been well quantified (Thornton & Lee 2000). Moreover, collecting estimates only from studies published in peer-reviewed journals serves as simple guaranty of quality and avoids multiple inclusion of the same results.

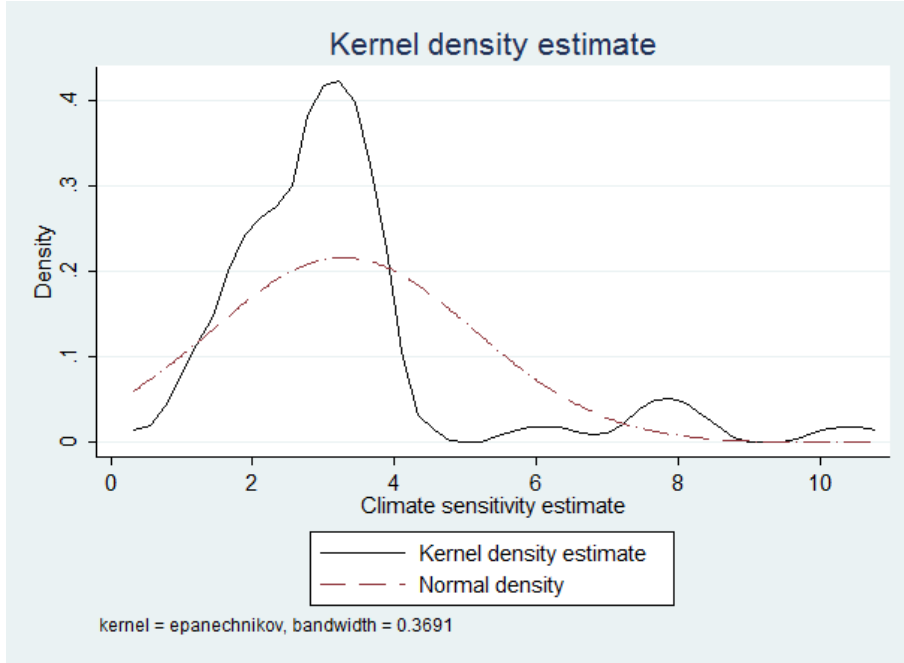
A total of 15 published papers and one dissertation from Switzerland provide 49 estimates of climate sensitivity. However, we decide to exclude one estimate of infinity, which would bias the meta-analysis. It comes from a study with multiple estimates computed using different models, and even the study itself fails to explain how it is possible to have a value of infinity. All 16 papers included are listed in Table 1. The estimates of climate sensitivity range from 0.7 to 10.4, with an average of 3.27. Full summary statistics of the estimates and a list of dummy variables are reported in Table 11 and Table 12 in Appendix A.

Figure 1 depicts the kernel density of the estimated climate sensitivity with the use of the Epanechnikov kernel (Samiuddin & El-Sayyad 1990). It indicates that the distribution is skewed. The right tail of the distribution is much longer than the left one. A common assumption made in meta-analysis is that in the absence of bias the estimates are normally distributed around the hypothetical true effect (Stanley 2001). Figure 1 depicts the normal density as a long-dash dot line for comparison.

4 Calculating the Standard Errors

Standard techniques for correcting publication bias suppose that the ratio of the effect estimates to their standard errors has a t -distribution, which implies approximately normally distributed estimates of climate sensitivity (Stanley 2005). Therefore, the estimates should not be cor-

Figure 1: The kernel density of climate sensitivity estimates



related with their standard errors. With asymmetric distributions this assumption does not necessarily hold, but there is no reason why climate sensitivity estimates should not be distributed symmetrically. None of the collected studies reports the standard error, so we use the confidence interval to calculate it. Assuming standard normal distribution of climate sensitivity estimates we construct the standard errors according to Wooldridge (2012) in the following way:

$$SE_{up} = \frac{up - est}{z}, \quad (1)$$

where SE_{up} denotes the standard error, up represents the upper bound of the confidence interval, est is the estimate of climate sensitivity, and z is the corresponding value for the magnitude of the confidence intervals. In most of the studies, 90% confidence intervals are reported, implying that z equals 1.645. One study reports only 66% confidence intervals and for that we use a z of 0.955. Because the estimates are in most cases asymmetrically distributed we compute two values of the standard error depending on which bound we use for the calculation.

$$SE_{low} = \frac{est - low}{z}, \quad (2)$$

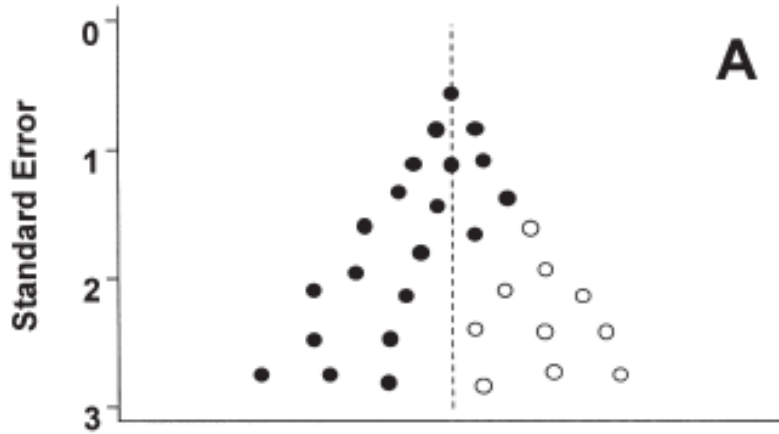
where SE_{low} denotes the standard error and low represents the lower bound of the confidence interval. The assumption of normality is the best choice even if it is inaccurate. Law *et al.* (1994, p. 427) mention that the assumption of normality “is commonly made in social science research under circumstances in which it cannot be known with certainty to be correct.” According Stanley (2001), in the absence of bias normality is a common assumption in meta-analysis. Moreover, Cohen (1983, p. 252) notes that “the failure of normality assumption, unless extreme, bears only marginally on the validity of the conclusions drawn.” Therefore, we consider

standard normal distribution of estimates of climate sensitivity to be a good approximation. For assurance, we check whether the possible relationship between the estimates and their standard errors is caused only by huge asymmetry in the distribution using the ratio of the standard errors constructed from “below” and from “above”. The calculated standard errors are summarized in Table 14 and Table 16 in Appendix A.

5 The Concept of Publication Bias

Graphical tests offer a simple way of detecting publication bias. Probably the most common is the funnel plot (Sterne *et al.* 2001; Stanley 2005). The name comes from the shape of the diagram. It depicts the estimated climate sensitivity on the horizontal axis, while the vertical axis measures the precision of the estimates, i.e., the inverse of the standard error. Without any bias the diagram should look like an inverted funnel. The estimates should be symmetrical around the values with the highest precision, since all estimates have the same chance of being reported (Havranek 2015). Imprecise estimates should also be present, although these will be infrequent and more dispersed, as shown in Figure 2.

Figure 2: Hypothetical funnel plot in the absence of publication bias



Notes: Source: Sterne *et al.* (2000). The inverted y-axis of the standard error is the same as the y-axis of precision. The x-axis shows the effect sizes (climate sensitivities). This is an example of a funnel plot in the absence of publication bias. If the open circles were missing, bias would be present.

However, the formal test for publication bias, often called the funnel asymmetry test or FAT, is described as the relation between the reported estimates and their standard errors (Murphy *et al.* 2004):

$$cs_i = c_0 + \beta_0 \cdot Se(cs_i) + \beta_1 \cdot mea + u_i, \quad u_i | Se(cs_i) \sim N(0, \delta^2), \quad (3)$$

where cs_i denotes the estimate of climate sensitivity, c_0 is the average climate sensitivity, $Se(cs_i)$ is the standard error of cs_i , β_0 measures the magnitude of publication bias, mea is a dummy

variable that equals 1 if the estimate of CS is the mean, thus β_1 corrects for the differences between mean and median estimates, and u_i is a disturbance term. In the absence of publication bias the estimates are randomly distributed around the true mean climate sensitivity, c_0 . If, however, some estimates fall into the “file drawer” because they are insignificant or just too low in magnitude, the reported estimates will be correlated with their standard errors and β_0 will be positive. As the estimates with a low standard error lie close to the mean climate sensitivity, the bigger the standard error, the more dispersed the estimates get, with some becoming very small and others large. Therefore, if researchers omit estimates that are low in magnitude but keep large imprecise ones, correlation arises between cs_i and $Se(cs_i)$. A significant estimate of β_0 provides formal evidence for publication bias and funnel asymmetry.

However, estimates taken from one study are likely to be dependent. We therefore employ study-level clusters to avoid within-study heterogeneity. For the same reason we also apply the fixed-effects model. Besides that, excessive asymmetry in the distribution of the climate sensitivity estimates could give rise to correlation between the estimates and their standard errors. For that reason we add variables detecting the magnitude of the asymmetry in the distribution of the estimates: first $Se_{lowup}(cs_{ij})$, defined as the ratio of the standard errors computed using the lower and upper bounds ($\frac{Se(cs_{ij})}{Se_{up}(cs_{ij})}$), and second $inter(cs_{ij})$, which detects whether this ratio is correlated with the standard error of the climate sensitivity ($Se(cs_{ij}) \cdot \frac{Se(cs_{ij})}{Se_{up}(cs_{ij})}$). If their coefficient, $\beta_2 = 0$, is statistically significant, the correlation between the climate sensitivity estimates and their standard errors does not necessarily signal publication selection bias. On the other hand, if we cannot reject the null hypothesis $H_0 : \beta_2 = 0$ we have evidence that it is not asymmetrical distribution that causes the relationship between the climate sensitivity estimates and their standard errors.

Specification (3) obviously suffers from heteroscedasticity, since the explanatory variable, $Se(cs_i)$, is a sample estimate of the standard error deviation of the response variable, cs_i . For that reason, meta-analysts prefer to use weighted least squares (Stanley (2005); Stanley *et al.* (2008)):

$$\frac{cs_{ij}}{Se(cs_{ij})} = t_{ij} = c_0 \cdot 1/Se(cs_{ij}) + \beta_1 \cdot \frac{mea}{Se(cs_{ij})} + \beta_0 + \xi_{ij}, \quad \xi_{ij}|Se(cs_{ij}) \sim N(0, \delta^2), \quad (4)$$

where i and j denote estimate and study subscripts and t_{ij} correspond to the t -statistics of the climate sensitivity estimates from the primary studies; the other characteristics remain the same as in equation (3), but the interpretation is now different, since we employ the precision term, $1/Se(cs_i)$, instead of $Se(cs_i)$. The intercept and slope coefficients are reversed from equation (3), and precision becomes the key variable in this meta-analysis. As Stanley *et al.* (2008) and many other meta-analysts argue, significance of coefficient c_0 may correspond to significance of the authentic effect of climate sensitivity beyond publication bias; testing $H_0 : \beta_0 = 0$ is effective in detecting publication bias.

The dependence of estimates from one study probably originates in the use of various data sets or specific explanatory variables for estimation across studies. This produces between-

study heterogeneity. To cope with this problem, meta-analysts often apply the mixed-effects multilevel model (Doucouliagos & Stanley 2009), which allows for unobserved between-study heterogeneity. We use the model specified by Havranek & Irsova (2011) as follows:

$$t_{ij} = c_0 \cdot 1/Se(cs_{ij}) + \beta_1 \cdot \frac{mea}{Se(cs_{ij})} + \beta_0 + \zeta_j + \epsilon_{ij}, \quad \zeta_j | Se(cs_{ij}) \sim N(0, \psi), \epsilon_{ij} | Se(cs_{ij}), \zeta_j \sim N(0, \theta) \quad (5)$$

The new model divides the overall error term ξ_{ij} into study-level random effects ζ_j and estimate-level disturbances ϵ_{ij} . Because the model assumes both components of the error term to be independent, we can calculate the overall error variance as follows: $\text{Var } \xi_{ij} = \psi + \theta$, where ψ explains the between-study variance (that is, between-study heterogeneity) and θ describes the within-study variance. If ψ is zero, the simple ordinary least squares (OLS) model would be equally suitable as the mixed-effects multilevel estimator. We employ the likelihood-ratio test (LR) to consider which estimator to use.

As Havranek *et al.* (2012) mention, the mixed-effects multilevel model resembles the random-effects model commonly used in panel data econometrics. However, the mixed-effects multilevel model is closer to the hierarchical ordering of the data. The name mixed-effects originates in the presence of a fixed effect c_0 and a random part ζ_j . The multilevel framework is thus more appropriate in meta-analysis, as it considers the imbalances in the data set. It uses the maximum likelihood estimator (MLE) instead of generalized least squares (GLS) and allows for multiple random effects (author-, study-, or country- level), so the mixed-effects multilevel model is more flexible. According to Havranek *et al.* (2012), because the data in meta-analyses are often unbalanced, it is difficult to test competently for the exogeneity assumptions that lie behind the mixed-effects multilevel model. But as Nelson & Kennedy (2009) note, unless the violation of the exogeneity assumptions is huge the damage caused by the mixed-effects does not outweigh their benefits. We will control for such violation through comparison with OLS, clustered at the study level. Huge differences between the estimates from OLS and mixed effects would imply violation of the exogeneity assumptions.

So far, only the significance of the true effect of climate sensitivity has been tested. To examine the magnitude of the authentic effect beyond publication selection bias we follow Havranek *et al.* (2012) and apply Heckman meta-regression. It is based on the existence of a nonlinear relationship between the estimates and their standard errors (Stanley & Doucouliagos 2007). The specification modifies equation (5) taking into account heteroscedasticity and between-study heterogeneity; it assumes a quadratic relationship between the standard errors and publication bias:

$$t_{ij} = c_0 \cdot 1/Se(cs_{ij}) + \beta_1 \cdot \frac{mea}{Se(cs_{ij})} + \beta_2 \cdot Se(cs_{ij}) + \beta_0 + \zeta_j + \epsilon_{ij}, \quad (6)$$

where c_0 measures the magnitude of the average climate sensitivity corrected for publication bias.

6 Graphical Tests of Publication Bias

Figure 3 depicts the funnel plot for the estimate of climate sensitivity using the standard error constructed with the lower tail of the confidence interval, and Figure 4 shows the same using the construction with the upper tail. The funnels are heavily asymmetrical: the left-hand side of the funnels is almost completely missing, hence we have good reason to believe that publication selection bias is strong in this literature.

In Figure 3 the dotted lines pick out climate sensitivity with magnitudes 1, 2, and 3, while the dashed line represents precision 2 (that is, standard error 0.5). With increasing precision the estimates converge to climate sensitivity 1.² Most of the estimates lie between 2 and 4, with quite low precision between 0.6 and 2. The most precise estimates differ in magnitude: one of them predicts climate sensitivity at around 3.5, while four predict it at between 1 and 2. Although the magnitude of the climate sensitivity estimates varies, Figure 3 clearly displays the relationship between the estimates and their precision: the higher the precision, the lower the estimate of climate sensitivity. In the absence of publication bias these figures should look like an inverted funnel. However, Figure 3 depicts only the right-hand side of the inverted funnel and the left-hand side is completely missing, indicating publication selectivity bias.

Figure 4 represents a check on whether the situation is similar with the use of the standard error constructed from the upper bound of confidence interval for the CS estimates. Figure 4 again signals publication bias, since the left-hand side of the inverted funnel is missing. The relationship between the estimates and their standard errors, however, is not so straightforward in Figure 4 as in previous figures. Estimates with very low precision (lower than 1, which means standard error higher than 1) converge with increasing precision to a CS value of 4. However, the high-precision estimates range around a CS of 2 and increase with decreasing precision. In addition, there are six estimates excluded from Figure 4, since the collected sample includes one missing and five infinity values of the upper limit, so the standard error cannot be calculated. The limit of infinity alone assumes that the estimate is more and more imprecise going to infinity. It provides evidence that the upper tail of the confidence interval for the estimates is not well constrained.³ For that reason, and because upward publication bias is expected, the subsequent analysis employs only the lower tail. The meta-analysis literature argues that the funnel plot explains both sources of publication bias – selection of significant estimates and selection of the expected magnitude or sign, although the explanation is often poor.

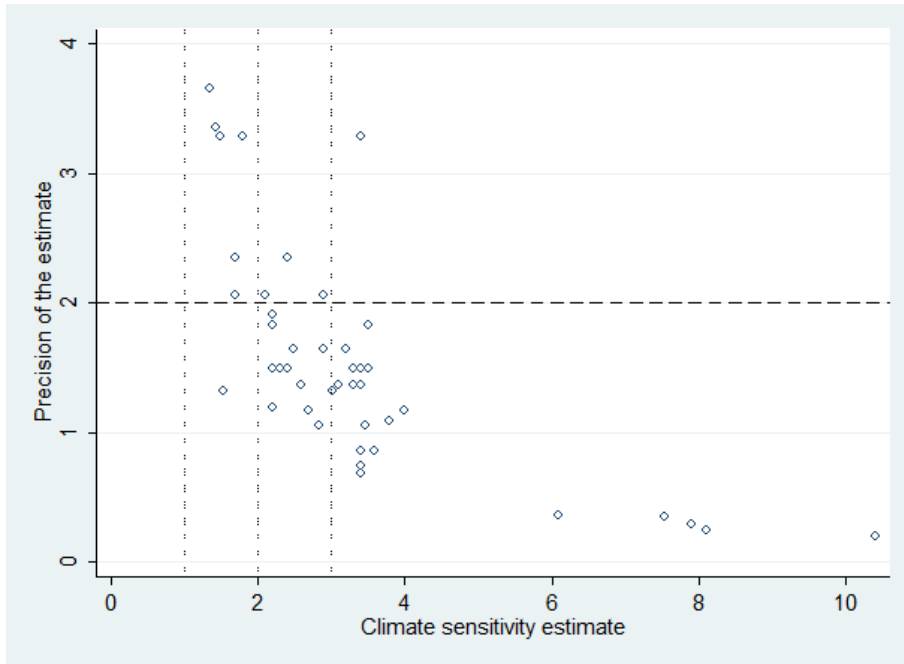
Figure 1 depicts the density of the estimated climate sensitivity using the Epanechnikov kernel. Again, in the absence of publication bias the distribution should be symmetrical, which is not the case with Figure 1. The left-hand side of the graph is completely missing and the shape of the solid line representing the kernel density of the CS estimates does not correspond to the normal density, shown as the long-dash dot line.

All the figures indicate publication selectivity bias. However, asymmetry in funnel plots can

²The single most precise and lowest estimate of climate sensitivity is excluded to depict the funnel in more detail.

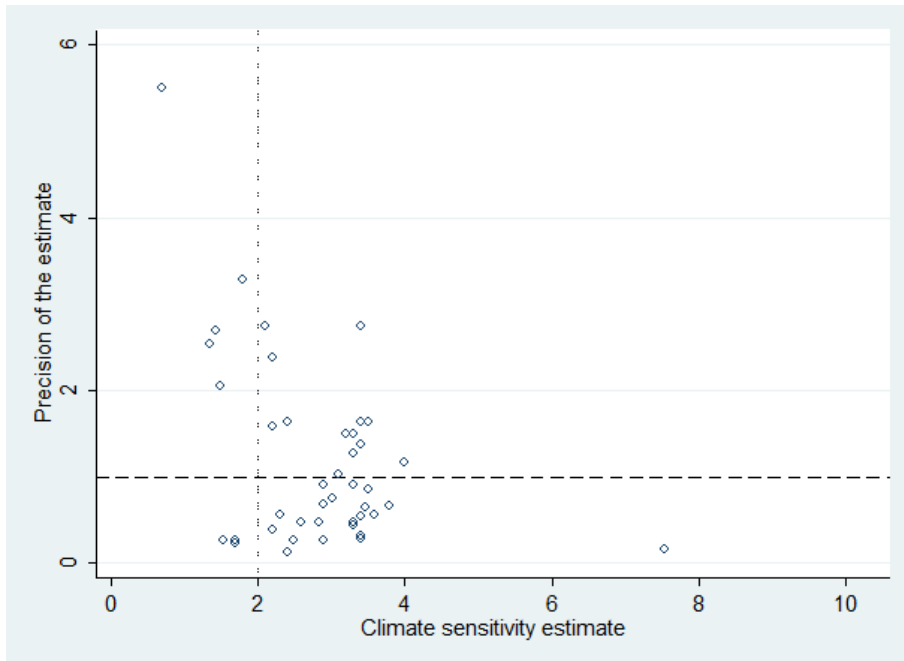
³Sterne *et al.* (2000) describe the contrary convergence as a signal of overestimation of treatment effects with lower methodological quality.

Figure 3: Funnel plot of the estimated CS



Notes: This figure excludes the single most precise estimate from the data set to zoom in on the relationship.

Figure 4: Funnel plot of the estimated climate sensitivities with the use of $S_{e_{up}}$



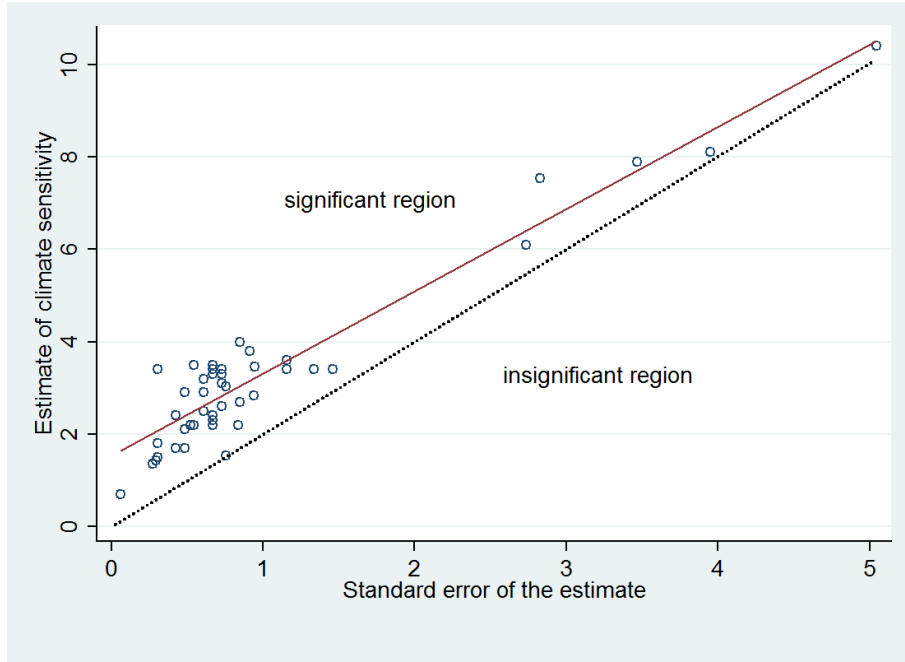
Notes: In the absence of publication bias, the funnel plot should be symmetrical around the most precise estimates of climate sensitivities. This funnel is asymmetrical. This suggests publication bias – nonnegative or very low positive estimates are reported even though, according to the law of chance, there should be at least a few of them.

be caused by factors other than publication bias, such as data irregularities or heterogeneity in the data set (Sterne *et al.* 2000). Further analysis is therefore needed.

7 Econometric Tests of Publication Bias

Let us proceed to the formal test of publication bias, described by regression (3). It is often called the funnel asymmetry test, or FAT, since it follows directly from the funnel plot. Though regression (3) only depends on the standard errors, according to many sources in the meta-analysis literature (for instance, Duval & Tweedie 2000) it still captures both sources of publication bias: first, the selection of significant estimates (type II bias in the terminology of Stanley 2005), and second, the selection of estimates with intuitive magnitude (type I bias). The suitability of funnel plots for detecting both sources of publication bias needs to be discussed, as the literature seldom explains it in detail (Havranek *et al.* 2012).

Figure 5: Visualization of the funnel asymmetry test



Notes: The dotted line denotes the combinations of the estimates of climate sensitivity and their standard errors for which the t -statistic equals two. The solid line denotes a linear fit of the points – that is, regression (3); its positive slope suggests publication bias.

Figure 5 visualizes the regression relationship (3) between the estimates and their standard errors. (Compared with the funnel plot, the axes are switched and the values on the new horizontal axis are inverted.) In the absence of publication bias, regression (3) would yield no significant slope coefficient β_0 , as the estimates should be randomly distributed around the true mean climate sensitivity c_0 . Moreover, Figure 5 would depict an isosceles triangle with the most precise estimates at the tip. In Figure 5 the tip would estimate the predicted true climate sensitivity, 1.69. First, let us assume that only enough high estimates, with no dependence on

their statistical significance, were reported. In such case, the triangle would lose its lower part. Regression (3) would yield a positive slope coefficient, which is evidence of publication bias. Second, let us suppose that researchers omit to report estimates insignificant at the 5% level, irrespective of the magnitude of the estimates. In such case, the imaginary triangle would lose its middle part. The boundary of significance at the 5% level as depicted by the dotted line in Figure 5 isolates the significant part from the insignificant one, as it represents the t -statistic 2 (since the estimates from primary studies are all positive, the figure does not picture the t -statistic -2). In the case of type II bias researchers do not report estimates with $|t| < 2$. Regression (3) would again predict a positive slope coefficient, indicating publication bias.

The steep positive slope of the regression line in Figure 5 signals the presence of strong upward publication bias. The source of it is identified as lying in both types of bias. The missing values in the right-hand lower corner of the imaginary isosceles triangle provide evidence for type I publication bias, since only 5 out of the 48 estimates are lower than 1.69. The low number of estimates with high standard errors (higher than 2) indicates the presence of type II bias. According to the law of chance, in Figure 5 there should be more estimates lower than 1.69, the tip of the hypothetical triangle, including insignificant ones. Negative estimates rarely occur in the field of climate sensitivity. Still, it is possible to estimate negative climate sensitivity and there should be at least few negative insignificant estimates and definitely more low positive estimates of climate sensitivity.

To sum up, the solid line in Figure 5 shows a linear fit based on regression (3), and its positive slope indicates strong publication bias. Assuming the standard error to be close to zero, the regression calculates the average estimate of climate sensitivity. In other words, if the precision were infinite, the hypothetical estimate of climate sensitivity would be shown in Figure 5 as the intercept of the solid line with the vertical axis (climate sensitivity of 1.69 in Figure 5). Table 2 shows the results of regression (3) compared to regression (3) with robust standard errors clustered at the study level. Both specifications detect publication bias, since the coefficient of the standard error is significant even at the 1% level.

However, because of heteroscedasticity and between-study heterogeneity, regression (3) is not commonly estimated itself. Moreover, excessive asymmetry in the distribution of climate sensitivity estimates could be the cause of the correlation between the estimates and their standard errors. Table 3 summarizes the results based on specifications which control for the magnitude of the asymmetric distribution. As *inter* and the share of SE are not significant at the 14% level, we have good reason to believe that the relationship between the climate sensitivity estimates and their standard errors stems from publication bias. The WLS regression (4) corrects for heteroscedasticity, the fixed-effects (FE) regression for within-study heterogeneity, and the mixed-effects multilevel (ME) regression (5) for between- and within-study heterogeneity. We can see in Table 4 that the results of the OLS and ME regressions are consistent. This serves as a robustness check of the mixed-effects multilevel regression, since testing the exogeneity assumptions behind this model is difficult because of the high degree of unbalancedness of the

Table 2: **Test of publication bias using OLS regression**

Response variable:	OLS	Clustered OLS
Estimate of CS		
Se (evidence of publication bias)	1.817*** (0.088)	1.817*** (0.065)
Constant (average true effect of CS)	1.692*** (0.138)	1.692*** (0.177)
mea (correction for mean estimates)	-0.365** (0.171)	-0.365* (0.203)
Observations	48	48
R^2	0.906	0.906

Notes: Standard errors are shown in parentheses and for second OLS clustered at the study level. ***, **, and * denote statistical significance at the 1%, 5%, and 10% level.

Table 3: **Test of asymmetric distribution of CS estimates**

Response variable:	Model-Share of SE	Model-Interaction term
Estimate of CS		
Se	2.131*** (0.153)	1.713*** (0.351)
inter		0.848 (0.587)
Se_{low}/Se_{up}	0.538 (0.344)	
mea	-0.389* (0.182)	-0.395** (0.178)
Constant	1.18*** (0.285)	1.451*** (0.203)
Observations	42	42
R^2	0.741	0.741

Notes: Standard errors, clustered at the study level for OLS, are shown in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% level. The interaction term ($inter = Se(cs_{ij}) \cdot \frac{Se(cs_{ij})}{Se_{up}(cs_{ij})}$) and the share of SE detect the magnitude of the asymmetry in the distribution of the estimates.

data. As the differences between the mixed-effects multilevel and clustered OLS regressions are negligible, the exogeneity assumption behind the mixed-effects model is not seriously violated. Likelihood-ratio tests reject the null hypothesis of the absence of between-study heterogeneity, which suggests that the OLS is misspecified and the mixed-effects model is more reliable.

Table 4: **Test of publication bias**

Response variable: t-statistic	ME	Clustered OLS	Clustered FE
Constant (publication bias)	2.192*** (0.328)	2.577*** (0.178)	2.043*** (0.08)
1/SE	1.689*** (0.188)	1.425*** (0.3)	2.15*** (0.085)
mea/SE	-1.105*** (0.186)	-0.832*** (0.274)	-1.573*** (0.077)
Observations	48	48	48
R^2		0.707	0.637
Likelihood-ratio test (χ^2)	8.82***		

Notes: Standard errors are shown in parentheses and clustered at the study level. ME denotes mixed-effects multilevel, OLS ordinary least squares, and FE fixed-effects regression. Null hypothesis for the likelihood-ratio test H_0 : no between-study heterogeneity (that is, mixed-effects multilevel has the same benefit as OLS). *** denotes statistical significance at the 1% level.

Table 5 compares the results of all the specifications – WLS, FE, and mixed-effects multilevel – using the standard error based on equation (1) constructed using the upper limit of the confidence interval of the estimated climate sensitivities. The result should verify the robustness of our analysis with the standard error based on equation (2) constructed using the lower tail of the confidence interval of the estimated CS. Both regressions yield a significant coefficient of publication bias at least at the 1% level. The magnitude of the bias based on the upper (lower) limit is 0.84 (2.58) in the case of WLS, 0.78 (2.19) in the case of mixed-effects multilevel, and 0.88 (2.04) in the fixed-effects regression. Therefore, we have good reason to believe that the choice of Se_{low} for implementation in the meta-analysis does not damage its results. We will explain later why significant publication bias of around $|2|$ signals publication selectivity.

The suitability of taking the estimate of climate sensitivity as mix of mean and median estimates should not be overlooked. The effect of the mode and the best estimate can be neglected, since there are only six such estimates in the sample. Some studies report both median and mean estimates, which creates a good subset for testing. These two new samples contain all the mean estimates that are in the original data set, but they also contain an extra 11 median estimates. Table 6 summarizes all the results, showing that there are quite huge differences in the magnitude of the bias depending on whether the mean or median estimate of climate sensitivity is used. The likelihood-ratio test does not reject the null hypothesis in the model with the median, which suggests that mixed-effects multilevel has no advantage over OLS

Table 5: **Test of publication bias using Se_{up}**

Response variable: t-statistic	Mixed-effects	Clustered OLS	Fixed-effects
Constant (publication bias)	0.78*** (0.245)	0.844*** (0.201)	0.882*** (0.206)
$1/SE_{up}$	2.227*** (0.192)	2.388*** (0.336)	2.327*** (0.275)
mea/SE_{up}	-0.59*** (0.083)	-1.216** (0.421)	-0.645*** (0.115)
Observations	42	42	42
Likelihood-ratio test (χ^2)	2.91**		
R^2		0.679	0.821

Notes: Standard errors are shown in parentheses and clustered at the study level. Null hypothesis for the likelihood-ratio test H_0 : no between-study heterogeneity (that is, mixed-effects multilevel has the same benefit as OLS). ***, **, and * denote statistical significance at the 1%, 5%, and 10% level.

regression (which means that between-study heterogeneity is present). However, the LR test rejects this hypothesis in the specification with mean estimates, suggesting that mixed-effects multilevel is more appropriate. The likely reason why only the model with median estimates suffers from between-study heterogeneity is that it contains estimates from 11 studies, while the model with mean values contains estimates from only half of them, i.e., from six studies. Concentrating the OLS regression by mean estimation and the multilevel regression by median estimation, the constants in both cases are significant at least at the 1% level, which indicates publication bias. The magnitudes differ: the bias in studies reporting mean estimates is more than twice that in studies reporting the median value.

Table 6: **Test of publication bias on subsets**

Response variable: t-statistic	Mixed-effects multilevel		Clustered OLS	
	mean	median	mean	median
Constant (publication bias)	4.154*** (1.068)	1.847*** (0.537)	4.154*** (0.828)	1.813** (0.616)
$1/SE$	0.564** (0.288)	1.596*** (0.273)	0.564*** (0.061)	1.663** (0.558)
Observations	25	28	25	28
Studies	6	11	6	11
Likelihood-ratio test (χ^2)	0	3.79**		

Notes: Standard errors are shown in parentheses and clustered at the study level. Null hypothesis for the likelihood-ratio test H_0 : no between-study heterogeneity (that is, mixed-effects multilevel has the same benefit as OLS). *** and ** denote statistical significance at the 1% and 5% level.

However, the whole sample provides more reliable results, since it contains 48 estimates of climate sensitivity, which means it is almost two times larger than each of the individual subsets. Although it does not include the 11 median estimates, it instead incorporates the mean values and six more estimates from other studies. Moreover, all the specifications correct for differences between the mean and other estimates using the *mea* variable. Its negative sign indicates that the mean estimates should be lower than the median estimates on average (or, except for OLS models, they have higher standard errors on average, which reduces their *t*-statistics and makes them less precise). This matches the data: the averages of the mean and median estimates differ slightly (3.42 and 3.21, respectively) and the averages of their standard errors differ more (1.16 and 0.79, respectively). Consequently, the median estimates have a higher *t*-statistic than the mean estimates on average. We therefore add an interaction term between the standard error and *mea* into equation (3) to distinguish whether the differences in the mean and median estimates imply variance in publication bias or variance in the magnitude of climate sensitivity (Stanley *et al.* 2008). The first row in Table 7 refers to specification (3) and indicates that the added interaction term is statistically insignificant. The interaction term, however, is highly correlated with the standard error of climate sensitivity and the two together are significant. After weighting by the standard errors of CS (which means the interaction term changes to *mea*) the “interaction” term becomes statistically significant at least at the 5% level in all models (WLS, fixed-effects, mixed-effects). Its sign is negative in all cases, which reduces the magnitude of the publication bias of the studies using mean estimates. However, the results are not consistent with the analysis of the subsets, where the selectivity in studies using mean estimates of climate sensitivity seems to be higher. To sum up, there is no solid evidence to suggest that the magnitude of publication bias varies between mean and median estimates.

Table 7: **Test of differences in publication bias between mean and median estimates**

Response variable:	OLS	WLS	FE	ME
	Estimate of CS		t-statistic	
SE* <i>mea</i>	-0.078 (0.195)			
<i>mea</i>		-0.962** (0.373)	-3.112*** (0.029)	-1.152*** (0.46)

Notes: ME denotes mixed-effects multilevel, OLS ordinary least squares, and FE fixed-effects regression. Standard errors are shown in parentheses and clustered at the study level. ***, **, and * denote statistical significance at the 1%, 5%, and 10% level.

The studies employ distinct characteristics and data sets, which may themselves lead to systematically different results and magnitudes for publication bias. Although it is not the main aim of this paper, we try to find out whether some aspects of primary studies can influence the differences in outcomes between individual estimates (that is, publication bias). With the

Table 8: **Multivariate meta-regression**

Response variable: t-statistic	Mixed-effects	Clustered OLS	Fixed-effects
Constant (publication bias)	1.929*** (0.344)	2.352*** (0.179)	1.764*** (0.119)
1/SE	1.74*** (0.187)	1.476*** (0.309)	2.255*** (0.099)
mea/SE	-1.152*** (0.184)	-0.873*** (0.276)	-1.676*** (0.089)
Ice	0.901** (0.452)	0.844*** (0.186)	0.941*** (0.136)
Observations	48	48	48
Likelihood-ratio test (χ^2)	7.61***		
R^2		0.788	0.659

Notes: Standard errors are shown in parentheses and clustered at the study level. Null hypothesis for the likelihood-ratio test H_0 : no between-study heterogeneity (that is, mixed-effects multilevel has the same benefit as OLS). ***, **, and * denote statistical significance at the 1%, 5%, and 10% level.

stepwise regression, only mea, ice, and cloud were chosen as possible drivers out of the 14 variables. However, the fixed-effects model omits cloud because of collinearity. Therefore, the final model contains only one extra variable compared to the previous specifications. The results in Table 8 again signal serious publication selectivity and put the true climate sensitivity in the range of 1.47–2.26. This serves as a robustness check of the analysis, since the results do not differ. Selectivity seems to be higher in the 10 studies using data about ice surface (it increases β_0 by approximately 0.9), but some studies do not report all the characteristics used to estimate climate sensitivity, hence we cannot take this result as a rule.

8 Discussion of the Results

We analyze 48 estimates of climate sensitivity from 16 studies. The estimates range from 0.7 to 10.4, with a mean of 3.27. The analysis yields interesting results. Although the estimates of climate sensitivity should not be correlated with their standard errors in the absence of publication bias, 14 models indicate the opposite. Publication bias is present in the climate sensitivity literature at least at the 5% significance level. Unfortunately, the analysis cannot precisely identify the reasons for such bias. Researchers and journal editors may be displaying selectivity in publishing only significant or preferred magnitude estimates.

In the studies in our sample, researchers report their estimates of climate sensitivity in the form of means, medians, modes or best estimates (here, “best” means as decided by the researchers without specifying whether the estimate is the mean, the median or anything else). Even their decision on what to report is fundamental. Mean or median estimates are reported most commonly, but only 11 studies state both of them. The mean estimates in the sample are

higher than the median estimates on average. At the same time, the magnitude of median estimates reported together with mean estimates is lower on average than that of median estimates reported alone. This suggests that researchers tend to report higher estimates because of their magnitude or in order to achieve higher significance, since the higher the t -statistic the higher the significance level, and the t -statistic is computed as the ratio of the estimate to its standard error.⁴

Both the mixed-effects and WLS models on the mean and median subsets indicate serious publication bias. According to the mean estimates the publication bias (β_0) is 4 at least at the 1% significance level, and according to the median estimates β_0 is 2 at least at the 5% significance level. The LR test detects between-study heterogeneity in the subset of median estimates. The definitions of the median as the middle value and the mean as the average offer a possible explanation, since the median may vary across different samples with the same mean. The mean and median subsets of the CS estimates are of similar magnitude: 25 coming from six studies and 28 coming from 11 studies, respectively. Samples of such magnitude should provide significant results. This analysis signals stronger selectivity in studies reporting mean estimates of climate sensitivity than in studies reporting median estimates. However, the analysis of the whole sample does not confirm such selectivity (see the results in Table 7). On the contrary, it indicates stronger publication bias according to the median estimates. Furthermore, the mean estimates still suffer from serious selectivity (the coefficient of publication bias, β_0 , ranges between 1.72 and 2.71 depending on the calculation method). Table 15 in Appendix A summarizes all the results.

As expected, after we create funnel plots for the whole data set, the meta-regression identifies upward publication selection bias significant at least at the 1% level for all the models applied. In all specifications the intensity of publication bias, β_0 , ranges between 1.8 and 3. Such magnitude of publication bias signals serious selection efforts. Doucouliagos & Stanley (2013) regard a FAT result of higher than 2 in absolute terms as “severe” selectivity: if the true climate sensitivity was zero and only statistically significant estimates of climate sensitivity were reported, the estimated coefficient of publication bias would be approximately 2 as the most common critical value of the t -statistic. The publication bias in this literature is hence strong enough to produce a significant average estimate of climate sensitivity that is much higher than the true value. Table 4 also shows that the estimate of the true effect after correcting for publication bias ranges between 1.4 and 2.2 at least at the 1% significance level in all the specifications. However, to estimate the true average climate sensitivity precisely, we employ the Heckman meta-regression specified in equation (6) in line with Stanley & Doucouliagos (2007) and Moreno *et al.* (2009). Table 9 summarizes the results. The likelihood-ratio test suggests again that at least at the 1% significance level, mixed-effects multilevel regression is more suitable. The models provide similar estimates of the true climate sensitivity: 1.3 (WLS), 1.6 (ME), and 2.1 (FE).

After correction for publication bias, the best estimate assumes that the mean climate

⁴In one model all the estimates have the same standard error.

sensitivity equals 1.6 with a 95% confidence interval (1.246, 1.989). This is one half of the simple uncorrected average, 3.27: the publication bias contains the estimate of the true CS approximately two times. Out of the 48 collected estimates, five are smaller than or equal to the average true effect; the lowest estimate is 0.7. This means that almost half of the estimates of climate sensitivity may be put into the “file drawer”.

Since the preferred mixed-effects model also adjusts for heteroscedasticity and other factors by applying WLS or study-level random effects, comparison with a simple average may not be straightforward (Havranek *et al.* 2012). Using the mixed-effects model on the unchanged original data set, the estimate of the uncorrected average is 3.274 with a 95% confidence interval (2.752, 3.797). That is as close as possible to the simple average (3.274). It suggests that the specification characteristics of the meta-regression do not imply differences between the simple averages and the corrected estimates. The results of this meta-analysis provide strong evidence of publication bias, and the estimated true effects do not significantly differ either. Table 10 compares the estimated true sensitivities across the model specifications. The estimates of the true CS range between 1.4 and 2.3 in the extreme cases, and the average is 1.74. That is very close to the preferred mixed-effects model estimate of 1.6, so there is good reason to believe that the result is robust.

Table 9: **Test of true climate sensitivity beyond publication bias**

Response variable: t-statistic	ME	Clustered OLS	Clustered FE
1/SE (true CS)	1.617*** (0.19)	1.276*** (0.316)	2.087*** (0.086)
mea/SE	-1.074*** (0.183)	-0.732** (0.286)	-1.55*** (0.079)
SE	-0.234* (0.132)	-0.316*** (0.086)	-0.226*** (0.017)
Constant (bias)	2.5*** (0.369)	3.054*** (0.232)	2.353*** (0.068)
Observations	48	48	48
Likelihood-ratio test (χ^2)	8***		
R^2		0.728	0.647

Notes: Standard errors are shown in parentheses and clustered at the study level. ME denotes mixed-effects multilevel, OLS ordinary least squares, and FE fixed-effects regression. Null hypothesis for the likelihood-ratio test H_0 : no between-study heterogeneity (that is, mixed-effects multilevel has the same benefit as OLS). ***, **, and * denote statistical significance at the 1%, 5%, and 10% level.

Table 10: **List of true effects of climate sensitivity**

Specification model:	True climate sensitivity
OLS	1.692*** (0.138)
clustered OLS	1.692*** 0.177
clustered WLS	1.425*** (0.3)
clustered WLS with dummy variables	1.476*** (0.309)
clustered WLS: correction of publication bias	1.276*** (0.316)
mixed-effects	1.689*** (0.188)
mixed-effects with dummy variables	1.74*** (0.187)
mixed-effects: correction of publication bias	1.617*** (0.19)
fixed-effects	2.15*** (0.085)
fixed-effects with dummy variables	2.255*** (0.099)
fixed-effects: correction of publication bias	2.087*** (0.086)

Notes: Standard errors are shown in parentheses and clustered at the study level. *** denotes statistical significance at the 1% level.

9 Conclusion

Anthropogenic climate change is a very topical issue. We consider climate sensitivity as an indicator and apply mixed-effects multilevel meta-regression to estimate potential publication selection bias and the underlying mean effect. The results confirm that publication bias is strong in this literature. After correction for the bias, the estimated true effect of climate sensitivity is approximately one half of the simple mean of all the estimates in the collected sample of literature. If the simple mean reflects climate scientists' impression of the magnitude of climate sensitivity, that impression exaggerates the true climate sensitivity two times.

We provide the first quantitative survey of journals estimating climate sensitivity (an indicator of anthropogenic climate change) and one of the first surveys of the literature concerning climate change. Michaels (2008) focuses on publication bias in the journals *Science* and *Nature* covering global warming. He collected a larger sample (116 articles), but his analysis does not take into account any exact measure of global warming and therefore does not use any econometric model. Using vote-counting meta-analysis he comes to the conclusion that the literature is biased.

We sample 48 climate sensitivity estimates. These are means, medians, modes or best estimates, but the majority are means or medians. An analysis conducted on subsets with 25 mean and 28 median CS estimates indicates that publication bias is twice as strong in the subset of means than in that of medians. Both subsets and the whole data set suffer from upward publication selectivity bias. The interpretation is not straightforward, however. The definition of median, which allows the estimate to be lower and higher than the mean estimate, causes the analysis of all the estimates together to signal stronger publication bias according to the median estimates.

Whether the magnitude of publication bias between mean and median estimates differs or not, the estimated climate sensitivity corrected for publication bias is approximately 1.6 when one accounts for a single estimate from each measurement collected (in this preference order: mean, median, mode, best estimate). Though meta-regression analysis is generally considered to be a statistically efficient tool, the corrected climate sensitivity estimate is a reference value. It averages across many methods, primary data sets, and factors influencing CS, and if there is another aspect influencing all the studies, this MRA will also be biased. The level of uncertainty in the prediction of climate sensitivity is high and is influenced by a huge number of factors. We tried to check for as many aspects as we could, but sometimes it was not possible to take them all into account. Still, publication selectivity is substantial in this literature, since its intensity in the full data set, β_0 , is around 2 and in the models corrected for heteroscedasticity and heterogeneity it approaches 4. This means that the literature may produce significant estimates of climate sensitivity that are twice as high as the true effect.

What consequences might this have? Predictions of climate change caused by humans influence policy decisions in most nations. Current environmental policy across many nations is focused on reducing emissions of greenhouse gases (GHG), especially CO_2 . For instance, the EU aims to cut GHG emissions by 20% by 2020 compared to 1990 (EU [2014]). Other countries

and territories, such as New Zealand, Australia, and Quebec in Canada, aim to reduce CO_2 emissions by implementing Emission Trading Systems (OECD 2013). The U.S. government has formed a special authority to investigate the social cost of carbon (SCC). It estimates the SCC as the economic damages associated with a small increase in emissions, that is, it puts a dollar figure on the benefit of a small reduction in emissions (EPA [2013]). The SCC measures the benefit of implementing a policy to reduce CO_2 emissions and can be understood as the amount of money spent on agriculture, human health and so on as a result of climate change (extreme weather) caused by a small increase in CO_2 emissions. This is exactly what climate sensitivity represents, since the SCC is calculated on the basis of the climatic response to an increase in emissions. It is possible that policy targets would be different if researchers reported lower climate sensitivities. A lower estimate of climate sensitivity would imply a lower estimate of the social cost of carbon. This, in turn, would influence the amount spent on reducing carbon dioxide in the atmosphere. This money could be spent on other areas of environmental protection.

References

- ANDRONOVA, N. G. & M. E. SCHLESINGER (2001): “Objective estimation of the probability density function for climate sensitivity.” *Journal of Geophysical Research: Atmospheres (1984–2012)* **106(D19)**: pp. 22605–22611.
- COHEN, J. (1983): “The cost of dichotomization.” *Applied Psychological Measurement* **7**: pp. 249–253.
- DE LONG, J. B. & K. LANG (1992): “Are all economic hypotheses false?” *Journal of Political Economy* **100**: pp. 1257–1257.
- DOUCOULIAGOS, C. & T. STANLEY (2013): “Are all economic facts greatly exaggerated? theory competition and selectivity.” *Journal of Economic Surveys* **27(2)**: pp. 316–339.
- DOUCOULIAGOS, H. & T. D. STANLEY (2009): “Publication selection bias in minimum-wage research? a meta-regression analysis.” *British Journal of Industrial Relations* **47(2)**: pp. 406–428.
- DUVAL, S. & R. TWEEDIE (2000): “Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis.” *Biometrics* **56(2)**: pp. 455–463.
- EDWARDS, T. L., M. CRUCIFIX, & S. P. HARRISON (2007): “Using the past to constrain the future: how the palaeorecord can improve estimates of global warming.” *Progress in Physical Geography* **31(5)**: pp. 481–500.
- EPA ([2013]): “The social cost of carbon.” <http://www.epa.gov/climatechange/EPAactivities/economics/scc.html>. [cit. 2014-05-07].
- EU ([2014]): “Eu greenhouse gas emissions and targets.” http://ec.europa.eu/clima/policies/g-gas/index_en.htm/. [cit. 2014-05-01].
- FOREST, C. E., P. H. STONE, & A. P. SOKOLOV (2006): “Estimated pdfs of climate system properties including natural and anthropogenic forcings.” *Geophysical Research Letters* **33(1)**: p. L01705.
- FRAME, D., B. BOOTH, J. KETTLEBOROUGH, D. STAINFORTH, J. GREGORY, M. COLLINS, & M. ALLEN (2005): “Constraining climate forecasts: The role of prior assumptions.” *Geophysical Research Letters* **32(9)**.

- GREGORY, J., R. STOUFFER, S. RAPER, P. STOTT, & N. RAYNER (2002): “An observationally based estimate of the climate sensitivity.” *Journal of Climate* **15(22)**.
- HARGREAVES, J. & J. ANNAN (2009): “Comment on” aerosol radiative forcing and climate sensitivity deduced from the last glacial maximum to holocene transition”, by p. chylek and u. lohmann, geophys. res. lett., 2008.” *Climate of the Past* **5(2)**.
- HAVRANEK, T. (2015): “Measuring intertemporal substitution: The importance of method choices and selective reporting.” *Journal of the European Economic Association* (**forthcoming**).
- HAVRANEK, T. & Z. IRSOVA (2011): “Estimating vertical spillovers from fdi: Why results vary and what the true effect is.” *Journal of International Economics* **85(2)**: pp. 234–244.
- HAVRANEK, T., Z. IRSOVA, & K. JANDA (2012): “Demand for gasoline is more price-inelastic than commonly thought.” *Energy Economics* **34(1)**: pp. 201–207.
- HEGERL, G. C., T. J. CROWLEY, W. T. HYDE, & D. J. FRAME (2006): “Climate sensitivity constrained by temperature reconstructions over the past seven centuries.” *Nature* **440(7087)**: pp. 1029–1032.
- HOUGHTON, J. T., Y. DING, D. J. GRIGGS, M. NOGUER, P. J. VAN DER LINDEN, X. DAI, K. MASKELL, & U. C. JOHNSON, CA (2001): “Projections of future climate change.” In “Climate change 2001: the scientific basis,” volume 881, pp. 67–71. Cambridge university press Cambridge.
- HUBER, M. (2011): *The Earth’s energy balance and its changes: Implications for past and future temperature change*. Ph.D. thesis, ETH.
- KLOCKE, D., R. PINCUS, & J. QUAAS (2011): “On constraining estimates of climate sensitivity with present-day observations through model weighting.” *Journal of Climate* **24(23)**.
- KNUTTI, R., G. A. MEEHL, M. R. ALLEN, & D. A. STAINFORTH (2006): “Constraining climate sensitivity from the seasonal cycle in surface temperature.” *Journal of Climate* **19(17)**.
- LAW, K. S., F. L. SCHMIDT, & J. E. HUNTER (1994): “Nonlinearity of range corrections in meta-analysis: Test of an improved procedure.” *Journal of Applied Psychology* **79(3)**: pp. 425–438.
- LINDZEN, R. S. & Y.-S. CHOI (2011): “On the observational determination of climate sensitivity and its implications.” *Asia-Pacific Journal of Atmospheric Sciences* **47(4)**: pp. 377–390.
- MASTERS, T. (2013): “Observational estimate of climate sensitivity from changes in the rate of ocean heat uptake and comparison to cmip5 models.” *Climate Dynamics* pp. 1–9.
- MICHAELS, P. J. (2008): “Evidence for” publication bias” concerning global warming in science and nature.” *Energy & environment* **19(2)**: pp. 287–301.
- MORENO, S. G., A. J. SUTTON, E. H. TURNER, K. R. ABRAMS, N. J. COOPER, T. M. PALMER, & A. ADES (2009): “Novel methods to deal with publication biases: secondary analysis of antidepressant trials in the fda trial registry database and related journal publications.” *BMJ: British Medical Journal* **339**.
- MURPHY, J. M., D. M. SEXTON, D. N. BARNETT, G. S. JONES, M. J. WEBB, M. COLLINS, & D. A. STAINFORTH (2004): “Quantification of modelling uncertainties in a large ensemble of climate change simulations.” *Nature* **430(7001)**: pp. 768–772.
- NELSON, J. P. & P. E. KENNEDY (2009): “The use (and abuse) of meta-analysis in environmental and natural resource economics: an assessment.” *Environmental and Resource Economics* **42(3)**: pp. 345–377.
- OECD (2013): “Climate and carbon aligning prices and policies.” **01**: pp. 11–23.

- PACHAURI, R. K. & A. REISINGER (2007): *Climate Change 2007 Synthesis Report: Summary for Policymakers*. IPCC Secretariat.
- PIANI, C., D. FRAME, D. STAINFORTH, & M. ALLEN (2005): “Constraints on climate change from a multi-thousand member ensemble of simulations.” *Geophysical Research Letters* **32(23)**.
- ROSENTHAL, R. (1979): “The file drawer problem and tolerance for null results.” *Psychological bulletin* **86(3)**: p. 638.
- SAMIUDDIN, M. & G. EL-SAYYAD (1990): “On nonparametric kernel density estimates.” *Biometrika* **77(4)**: pp. 865–874.
- SCAFETTA, N. (2013a): “Discussion on climate oscillations: Cmp5 general circulation models versus a semi-empirical harmonic model based on astronomical cycles.” *Earth-Science Reviews* **126**: pp. 321–357.
- SCAFETTA, N. (2013b): “Solar and planetary oscillation control on climate change: hind-cast, forecast and a comparison with the cmp5 gcms.” *Energy & Environment* **24(3)**: pp. 455–496.
- SCHMITTNER, A., N. M. URBAN, J. D. SHAKUN, N. M. MAHOWALD, P. U. CLARK, P. J. BARTLEIN, A. C. MIX, & A. ROSELL-MELE (2011): “Climate sensitivity estimated from temperature reconstructions of the last glacial maximum.” *Science* **334(6061)**: pp. 1385–1388.
- SOLOMON, S., D. QIN, M. MANNING, Z. CHEN, M. MARQUIS, K. AVERYT, M. TIGNOR, & H. MILLER (2007): “Ipcc, 2007: climate change 2007: the physical science basis.” *Contribution of Working Group I to the fourth assessment report of the Intergovernmental Panel on Climate Change* .
- STANLEY, T. & C. DOUCOULIAGOS (2007): “Identifying and correcting publication selection bias in the efficiency-wage literature: Heckman meta-regression.” *Economics Series* **11**: p. 2007.
- STANLEY, T. D. (2001): “Wheat from chaff: Meta-analysis as quantitative literature review.” *Journal of economic perspectives* pp. 131–150.
- STANLEY, T. D. (2005): “Beyond publication bias.” *Journal of Economic Surveys* **19(3)**: pp. 309–345.
- STANLEY, T. D., C. DOUCOULIAGOS, & S. B. JARRELL (2008): “Meta-regression analysis as the socio-economics of economics research.” *The Journal of Socio-Economics* **37(1)**: pp. 276–292.
- STERNE, J. A., M. EGGER, & G. D. SMITH (2001): “Investigating and dealing with publication and other biases in meta-analysis.” *Bmj* **323(7304)**: pp. 101–105.
- STERNE, J. A., D. GAVAGHAN, & M. EGGER (2000): “Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature.” *Journal of clinical epidemiology* **53(11)**: pp. 1119–1129.
- STOCKER, D. Q. (2013): “Climate change 2013: The physical science basis.” *Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Summary for Policymakers, IPCC* .
- THORNTON, A. & P. LEE (2000): “Publication bias in meta-analysis: its causes and consequences.” *Journal of clinical epidemiology* **53(2)**: pp. 207–216.
- WEBB, M., C. SENIOR, D. SEXTON, W. INGRAM, K. WILLIAMS, M. RINGER, B. MCAVANEY, R. COLMAN, B. SODEN, R. GUDGEL *et al.* (2006): “On the contribution of local feedback mechanisms to the range of climate sensitivity in two gcm ensembles.” *Climate Dynamics* **27(1)**: pp. 17–38.
- WIGLEY, T., C. AMMANN, B. SANTER, & S. C. RAPER (2005): “Effect of climate sensitivity on the response to volcanic forcing.” *Journal of Geophysical Research: Atmospheres (1984–2012)* **110(D9)**.

A Additional tables

Table 11: **Summary statistics of regression variables**

Variable	Observations	Mean	Std. dev.	Min	Max
Climate sensitivity	48	3.274	1.847	0.7	10.4
Mean estimate of CS	25	3.417	2.462	0.7	10.4
Median estimate of CS	28	3.139	1.026	1.38	6.1
Mode estimate of CS	2	2.9	0	2.9	2.9
Best estimate of CS	5	2.82	0.75	1.54	3.4
constructed t -statistic	48	4.316	1.86	2.043	11.475
low (confidence interval)	48	1.689	0.596	0.3	2.9
up (confidence interval)	42	5.796	3.171	1	17.8
Se_{low}	48	0.975	0.983	0.061	5.046
Se_{up}	42	6.695	32.139	0.182	218.28
Publication year	48	2007.813	3.535	2001	2013
No. of citations of study	48	124.646	186.317	2	918
No. of models in study	48	9.417	21.576	1	128

Notes: 11 studies report both mean and median estimates of climate sensitivity. In these cases, only the mean estimate is included in the data set for this meta-analysis (i.e., the climate sensitivity variable contains 17 median estimates), except for the specification where all the median estimates collected are used. Similarly, the climate sensitivity variable contains only the four best estimates. The t -statistics are computed in the following way: $cs_i/Se(cs_i)$, with the use of the standard error of the CS estimate constructed from the lower limit of the reported confidence intervals. The search for primary studies was terminated on March 3, 2014.

B Details about the Studies Collected

One of the studies included does not report what significance level is assumed by the confidence interval and only states that the mean value of the estimate of climate sensitivity “may very likely vary from 0.9 to 2” (Scafetta 2013b). We believe it should be the 90% confidence interval, as that is the most frequent one, so we use a z value of 1.645 in constructing the standard error terms. If we are wrong, the bias caused by this should not be significant, as it appears in just this one case.

Lindzen & Choi (2011) report two confidence intervals for one estimate of climate sensitivity. We would have to include this estimate twice if we were to calculate with both intervals. In order to keep the data as consistent as possible, we use only the 90% confidence interval of the estimate in constructing the standard error (the studies in the data set mostly report a 90% confidence interval).

Table 12: **List of dummy variables and their description**

Variable	Description	Number of studies that use it
ASA	Aerosol forcing	17
solar	Solar irradiation	32
cloud	Cloud feedback	15
ozone	Ozone forcing	29
volcanic	Volcanic forcing	10
ice	Information about the ice sheet	10
mo	The study uses models for generating data	19
mea	The estimate of CS is the mean	25
medi	The estimate of CS is the median	28

Notes: 11 studies report both mean and median estimates of climate sensitivity. In these cases, only the mean estimate is included in the data set for this meta-analysis (i.e., the data set contains 17 median estimates), except for the specification where all the median estimates collected are used. The search for primary studies was terminated on March 3, 2014.

Table 13: **Summary statistics of collected estimates of climate sensitivity**

Variable	Observations	Mean	Std. dev.	Min	Max
Median estimates of CS used in climate sensitivity variable	17	3.212	0.921	1.5	6.1
Median estimate of CS	28	3.139	1.026	1.38	6.1
Mean estimate of CS	25	3.417	2.462	0.7	10.4
Mode estimate of CS	2	2.9	0	2.9	2.9
Best estimate of CS	4	2.82	0.75	1.54	3.4
Climate sensitivity	48	3.274	1.847	0.7	10.4

Notes: The climate sensitivity variable contains the estimates included in the data set for the meta-analysis. 11 studies report both mean and median estimates of climate sensitivity. In these cases, only the mean estimate is included in the data set (i.e., the climate sensitivity variable contains 17 median estimates), except for the specification where all the median estimates collected are used. Similarly, the climate sensitivity variable contains only the four best estimates. The search for primary studies was terminated on March 3, 2014.

Table 14: **Summary statistics of standard errors**

Standard error of:	Observations	Mean	Std. dev.	Min	Max
Median estimates of CS used in the climate sensitivity variable	17	0.794	0.562	0.304	2.736
Median estimate of CS	28	0.861	0.607	0.298	2.827
Mean estimate of CS	25	1.165	1.264	0.061	5.046
Mode estimate of CS	2	0.547	0.086	0.486	0.608
Best estimate of CS	4	0.78	0.115	0.669	0.942
Climate sensitivity (all together)	48	0.975	0.983	0.061	5.046

Notes: The climate sensitivity variable contains the estimates included in the data set for the meta-analysis. 11 studies report both mean and median estimates of climate sensitivity. In these cases, only the mean estimate is included in the data set (i.e., the climate sensitivity variable contains 17 median estimates), except for the specification where all the median estimates collected are used. Similarly, the climate sensitivity variable contains only the four best estimates. The search for primary studies was terminated on March 3, 2014.

Table 15: **Differences in publication bias between mean and median estimates of climate sensitivity**

Response variable: t-statistic	ME	Clustered OLS	Clustered FE
Constant (publication bias)	3.716*** (0.331)	3.668*** (0.365)	4.829*** (0.103)
1/SE	0.614*** (0.064)	0.618*** (0.08)	0.596*** (0.047)
mea	-1.152** (0.46)	-0.962** (0.373)	-3.114*** (0.029)
Observations	48	48	48
R^2		0.602	0.421
Likelihood-ratio test (χ^2)	5.93***		

Notes: Standard errors are shown in parentheses and clustered at the study level. ME denotes mixed-effects multilevel, OLS ordinary least squares, and FE fixed-effects regression. Null hypothesis for the likelihood-ratio test H_0 : no between-study heterogeneity (that is, mixed-effects multilevel has the same benefit as OLS). *** and ** denote statistical significance at the 1% and 5% level.

Table 16: List of constructed standard errors

ID study	Estimate of CS	Lower limit	Upper limit	z	Se_{low}	Se_{up}
2	2.3	1.2	5.2	1.96	0.669	1.763
3	3.2	2.2	4.3	1.96	0.608	0.669
3	3.3	2.2	4.6	1.96	0.669	0.79
3	3.4	2.3	4.4	1.96	0.669	0.608
3	3.3	2.2	4.4	1.96	0.669	0.669
3	3.1	1.9	4.7	1.96	0.729	0.973
3	3.4	1.5	6.4	1.96	1.155	1.824
4	2.9	2.1	8.9	1.96	0.486	3.647
4	2.9	1.9	4.7	1.96	0.608	1.094
5	3.3	2.2	6.8	1.96	0.669	2.128
6	6.1	1.6	-	1.96	2.736	-
7	1.43	0.94	2.04	1.96	0.298	0.371
7	3.46	1.9	6.02	1.96	0.948	1.556
7	7.53	2.88	17.8	1.96	2.827	6.243
7	3.4	1	9.3	1.96	1.459	3.587
8	2.9	1.9	5.3	1.96	0.608	1.459
8	3.5	2.4	5.4	1.96	0.669	1.155
10	3.8	2.3	6.3	1.96	0.912	1.52
10	3.3	2.1	7.1	1.96	0.729	2.31
11	2.83	1.28	6.32	1.96	0.942	2.122
11	1.54	0.3	7.73	1.96	0.754	3.763
11	3.03	1.79	5.21	1.96	0.754	1.325
12	1.5	1	2.3	1.96	0.304	0.486
13	3.5	2.6	4.5	1.96	0.547	0.608
13	2.4	1.7	3.4	1.96	0.426	0.608
16	2.6	1.4	6.1	1.96	0.729	2.128
16	3.4	1.2	8.6	1.96	1.337	3.161
17	3.4	2.9	4	1.96	0.304	0.365
17	4	2.6	5.4	1.96	0.851	0.851
17	3.6	1.7	6.5	1.96	1.155	1.763
17	1.8	1.3	2.3	1.96	0.304	0.304
17	3.3	2.2	5.1	1.96	0.669	1.094
18	2.2	1.7	2.6	0.955	0.524	0.419
18	2.2	1.4	2.8	0.955	0.838	0.628
18	2.1	1.3	2.7	1.96	0.486	0.365
18	3.4	2.2	4.6	1.96	0.729	0.729
19	0.7	0.6	1	1.96	0.061	0.182
19	8.1	1.6	-	1.96	3.951	-
19	1.7	0.9	8	1.96	0.486	3.83
19	7.9	2.2	-	1.96	3.465	-
19	2.2	1.1	-	1.96	0.669	-
19	2.5	1.5	8.7	1.96	0.608	3.769
19	2.7	1.3	-	1.96	0.851	-
19	10.4	2.1	-	1.96	5.046	-
19	2.2	1.3	6.4	1.96	0.547	2.553
19	2.4	1.3	14.7	1.96	0.669	7.477
19	1.7	1	8.8	1.96	0.426	4.316
20	1.35	0.9	2	1.96	0.274	0.395