ESSAYS ON META-REGRESSION ANALYSIS:

TWO EMPIRICAL STUDIES ON TAX AND ECONOMIC GROWTH AND A SIMULATION

A thesis submitted in partial fulfilment of the requirements for the Degree

of Doctor of Philosophy in Economics

in the University of Canterbury

by Nazila Alinaghi

Supervised by: Prof. William Robert Reed

Co-Supervised by: Dr. Philip Gunby

University of Canterbury

2017

ABSTRACT

This thesis consists of three essays linked together by the methodology of meta-regression analysis. The first two essays address a long standing question of interest to both economists and policymakers which is whether taxes exert an important influence on economic growth and, if they do, how large this effect might be. To answer this question, I study two different setups. The first involves OECD countries and the second studies U.S. states. The last essay studies the performance of the FAT-PET-PEESE (FPP) procedure, a commonly employed approach for addressing publication selection bias in meta-regression analysis studies in economics and business.

In my first meta-regression analysis, I combine results from 42 studies containing 713 comparable estimates, all which endeavour to estimate the effect of taxes on economic growth in OECD countries (Chapter 2). I then switch from an institutionally and culturally diverse setup to a setup in which there is a set of common institutional features, U.S. states. I integrate 966 estimates derived from 29 studies investigating the effect of taxes on economic growth in the U.S. states (Chapter 3). The objective of these two studies is to answer the following questions: (Q1) What is the overall, mean effect of taxes on economic growth?; (Q2) Are some taxes (e.g., personal income taxes) more distortionary than others (e.g., value added taxes)?; (Q3) Is there any empirical evidence to support the conventional wisdom that "distortionary taxes" used to fund "unproductive expenditures" are especially harmful for economic growth?; and (Q4) What are the factors causing researchers to encounter different or even contradictory results? My results for OECD countries suggest that there is publication bias towards negative estimates. Controlling for publication bias, I find that the overall effect

of taxes on economic growth is statistically insignificant and negligibly small. An increase in unproductive expenditure funded by distortionary taxes has a significant negative effect on growth. I find weak evidence to support the idea that some taxes are more distortionary than others. Lastly, several factors are present that can explain discrepancies among the reported estimates, such as estimation methods, types of standard errors, whether the original study was published in a peer-reviewed journal, the publication date, and so on. I find the following outcomes in the study of taxes in U.S. states: estimates in the literature are characterized by statistically significant negative publication bias. Once I control for publication bias, the overall effect is not particularly meaningful since it lumps together different kinds of tax policies. With respect to particular types of taxes, I could not find enough evidence to support that taxes on labour are more growth retarding than other types of taxes. Evidence regarding other types of taxes is mixed. Finally, as with results for OECD countries, there are several factors that appear to explain discrepancies among the reported estimates for U.S. states.

In the Chapter 4, I conduct a Monte Carlo analysis to evaluate the performance of the FPP procedure in detecting and correcting for publication bias. The main three objectives of applying FPP procedure are: (i) Funnel Asymmetry Testing (FAT) to test whether the sample of estimates is influenced by publication selection bias; (ii) Precision Effect Testing (PET) to test whether there is a genuine non-zero true effect of estimates once the publication bias is accommodated and corrected; and (iii) Precision Effect Estimate with Standard Error (PEESE) to obtain an improved estimate of the overall mean effect. I simulate two common types of publication bias. These are publication bias against insignificant results and publication bias against wrong-signed results (according to associated theory). I run these simulations in three different environments, Fixed Effects, Random Effects, and Panel Random Effects. My findings indicate that the FPP procedure performs well in the basic but unrealistic environment of “Fixed Effects”, when there is one true effect and sampling error is the only reason why studies produce different estimates. However, once I study its performance in more realistic data environments, where there is heterogeneity in the population effects between and within studies, the FPP procedure becomes unreliable for the first two objectives, and is less efficient than other estimators when estimating overall mean effects. Further, hypothesis tests about the overall, mean effect cannot be trusted. These results call into question the efficacy of using the FPP procedure to test and correct for publication selection bias in meta-regression analysis studies.

ACKNOWLEDGEMENTS

I would firstly like to thank my primary supervisor, Professor Bob Reed for providing me with continuous support and excellent guidance throughout my thesis journey. I would also like to thank my associate supervisor, Dr. Philip Gunby for his guidance and encouragement during my study.

I would like to thank the University of Canterbury for funding my research with the UC Doctoral Scholarship. The entire Economics and Finance Department at the University of Canterbury deserve praise for providing me with several teaching assistant and also lecturing roles over the past three years. The experience I have gained from teaching has been invaluable. I am also extremely appreciative of the administrative support provided by Ms. Meredith Henderson.

Finally, I would like to thank my husband, Reza E Sedgh and my parents, Mohammad Ebrahim Alinaghi and Azam Edalat and my sisters, Elham, Maryam, and Masume for their unwavering support and love. They all have provided me with all possible opportunities and I would like to dedicate this thesis to them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

Chapter 1. Introduction

The rapid growth in research publications has often not made scientific research more illuminating, but on the contrary, this exponential growth has made our understanding of the world and the people in it more confusing. The major problem arises when conflicting and contradictory results are reported in similar studies of the same research question. Conventional narrative reviews may deal with aggregating different studies which have matching findings to some extent. However, those reviews use a "vote counting" method to summarize the results available in the literature and may advocate certain results according to the reviewer's points of view. Moreover, they are not normally easily reproduced by other scholars.

To overcome these and other shortcomings, the technique of meta-analysis was introduced by Glass in 1976. Meta-analysis is the statistical analysis of estimates from multiple studies that are concerned with measuring a similar "effect". By "effect" I mean the relationship between the main predictor and the dependent variable in primary studies. Two main goals of meta-analysis are (i) to reach a single conclusion about the size and significance of that effect, and (ii) to clarify the causes of variations of estimates of the effect across studies.

Thus, this method can offer a consistent reproducible guide to the reader to make sense of the rapidly expanding research literature on a given topic. Furthermore, by looking at study-variant characteristics, meta-regression analysis studies are able to provide additional insights to existing knowledge. This method has long been applied in various fields including medicine, psychology and education. In recent years, however, it has become an increasingly popular research tool in economics and business. Figure 1.1 represents a time series bar chart that lists all Web of Science journal articles in economics and business that have the word "meta-analysis" in their titles. The trend is clearly upward.

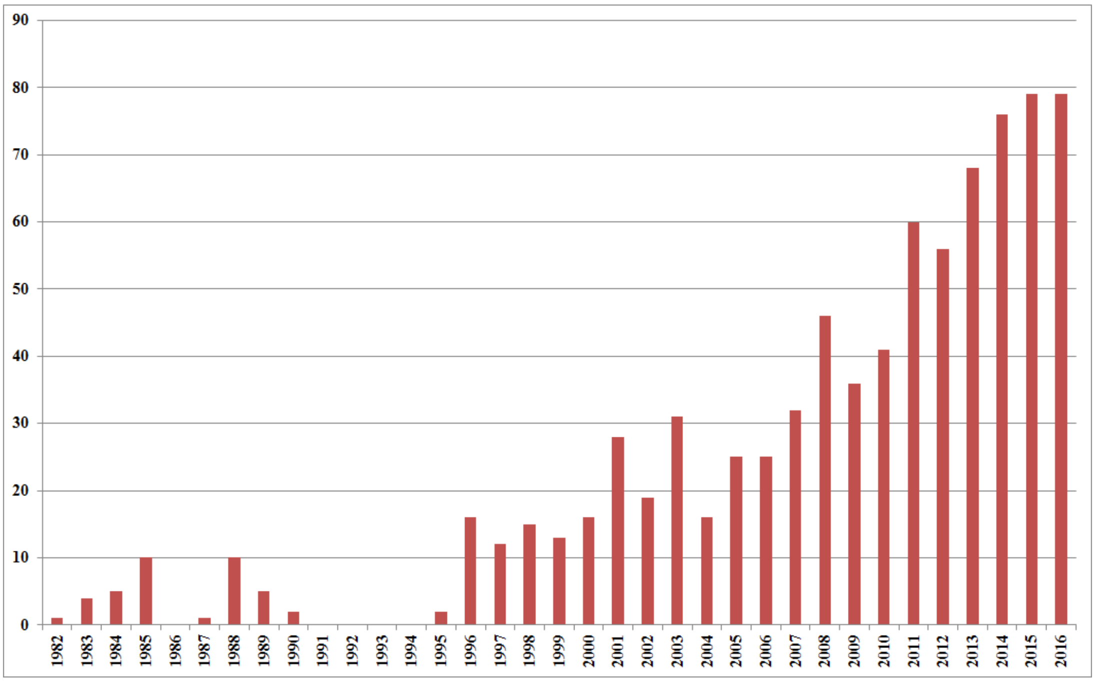


Figure 1.1: Number of Articles in Economics and Business Listed in Web of Science with "Meta-Analysis" in the Title

Note: Web of Science categories are: Economics, Business Finance, Business, Management, Criminology Penology, Urban Studies, and Social Sciences Interdisciplinary (813 articles).

There are several objectives in using the meta-analysis method. Some of which are as follows:

- Summarizing and integrating results from many original studies.
- Analysing differences in the results among studies.
- Overcoming small sample sizes of individual studies to detect the precise effects of an intervention, especially once a larger sample size is required.
- Increasing precision in estimating effects.
- Diagnosing and correcting misspecification bias in the original studies.
- Determining if new studies are needed to further investigate an issue.
- Generating new hypotheses for future studies.

The six basic steps of conducting high quality research to meet the above outlined objectives are as follows:

- Defining a research question.
- Conducting a literature search as comprehensive as possible to find all the relevant studies on a given topic, and including not only published papers but all the studies conducted by government advisors, research firms, policy think thanks and interest groups.
- Data coding to extract information from the divergent results.
- Calculating effect sizes which measure the relationship or association between the variables of interest.
- Explaining the differences in effect sizes.
- Identifying new areas of research, since conducting a meta-analysis highlights which parts of the topics need further research.

Like all empirical studies, meta-analysis starts with defining a research question. However, the main feature of meta-analysis is that it requires an exhaustive search of not only peer-reviewed journal papers but also unpublished papers on a given research question to identify the most relevant studies.[1] As such, this literature search includes an electronic search along with a complementary manual search. A set of inclusion/exclusion criteria is explicitly defined to narrow down the search for relevant studies. Being transparent in this way allows other researchers to replicate a meta-study.

The population of studies is identified by conducting both a traditional backward citation search strategy as well as a forward citation search strategy on the core studies collected in the previous stage. Backward citation refers to all the resources available at the

[1] Most meta-regression analyses conducted in economics only consider peer-reviewed or published literature due to the general belief that those studies have a higher quality compared to the other types of outlets, since they go through a refereeing process once.

end of core studies which have influenced the authors' thinking whereas forward citation indicates all the papers citing these core studies after their publication.

Another useful strategy to retrieve more relevant studies is to contact the leading authors who have conducted studies in the same research area. The main purpose of doing that is to identify any other unpublished papers written by these authors since the meta-analyst is often unaware of their existence. It is also useful in identifying any other papers written by relatively new scholars such as PhD students who are working on the same research question. All relevant studies collected in this manner are then read carefully for acceptability according to the pre-set criteria created by the researcher.

Once the literature search is completed, the next step is to identify important characteristics of studies and then code them. This step can be the most difficult and time consuming part of ameta-regression analysis. Two main issues in terms of coding are what to code and how to code. As stated by Ringquist (2013), three levels of information including search, study, and effect need to be considered for coding studies:

1. **Search-level information**: This includes all primary specifications assigned to each paper to identify for double checking at a later stage or for any replication purposes. Papers are given a unique ordinary number according to the authors' names (alphabetical order) to identify each relevant study. Other coding categories include the publication year, the name of the first three authors and the databases used for the relevant study.

2. **Study-level information**: This involves reading each paper in its entirety to exclude irrelevant studies which initially appear to be potentially relevant. This means that this level of information is coded just for the relevant studies. The coding at this stage

includes a full bibliographic citation including authors' name, publication date, publication status or some other specifications like the academic discipline of the outlet. These categories might be especially helpful in terms of estimating the presence of publication bias. It can also be used to study the evolution of a topic over time.

A complication when coding at the study-level is the existence of different formats of a single study. The standard practice in this situation is to choose the published version of the study. However, if all the available formats are unpublished then the preference is the latest version of the study.

3. **Effect-level information**: This part of coding is the most time consuming part as it involves exercising different levels of personal judgement. The most essential part of the coding is all the information which allows the meta-analyst to calculate effect sizes, such as regression coefficients, standard errors, t-statistics/p-values, which are extracted from the tables in the text or even the appendices. If the writer just reports the coefficients, but no other statistical details such as t-statistics or p-values, then the paper should be dropped from the final sample.

Almost every single study in economics will generate multiple effect sizes. Keeping a record of them all is crucial in conducting a comprehensive meta-analysis. For quality assurance purposes, the search and data coding procedure for a Meta-Regression Analysis (MRA) study in economics should follow the MAER-NET protocol (Stanley et al., 2013).

The next step after the coding of studies is to analyse the data. Doing so requires a standardized measure of association between two variables of interest (the explanatory variable and the response variable) such as effect size. The effect size is a measure of a relationship which can be validly compared across studies or even within a single study.

Several different effect types are available in meta-analysis research and include estimated elasticities, regression coefficients, odds ratio, partial correlation coefficients, and Fisher's *Z*-transformed partial correlation coefficients (see Hunter and Schmidt, 2004).

Due to inconsistencies of measurement units for regression variables available in the literature, many meta-analysts prefer to convert all the estimates into a common and comparable measure called the Partial Correlation Coefficient (PCC). The main advantage of using the PCC is that it is a unitless measure and can easily be compared across different estimates within a study and also between studies. The regression coefficients on the other hand are ideal measures for effect size as they refer to economic effect than statistical effect. Unfortunately, they are not independent of units of measurements. Therefore, they are not comparable unless the meta-analyst makes sure all the studies use the same scale (Stanley and Doucouliagos, 2013).

The final step in conducting a meta-regression analysis is to identify which characteristics can explain the large variation among the reported estimates. This also tells us how future studies can be better designed than existing studies.

All of the above mentioned steps lay out the standards of good MRA practice. But when it comes to application, "judgement calls" are inevitable and the steps are not as straightforward as suggested earlier. Further, some commonly used procedures in analysing the coded data have not received much scrutiny, which calls into question how valid they are in practice.

In addition to undertaking meta-regression analysis studies, my thesis is also concerned with identifying possible shortcomings in current practice. The first part of my thesis involves studying the effects of taxes on economic growth. Apart from naturally being a topic which is interesting in its own right, coding tax effects throws up interesting and unusual issues. For example, how to deal with dynamic effects, different units of measurement, and the presence of multiple, possibly conflicting simultaneous effects as a result of the composition of fiscal policy. The second part of my thesis is concerned with the issue of publication bias. In particular, I am interested in studying the performance of the FAT-PET-PEESE (FPP) procedure, a commonly employed approach for addressing publication selection bias in the economics and business meta-regression analysis literature.

In the first two essays, the effects of taxes on economic growth in OECD countries and U.S. states are examined, respectively. The last essay studies the performance of the FAT-PET-PEESE (FPP) procedure. Thus, while the three essays included in this dissertation examine separate topics, they are linked together by the methodology of meta-regression analysis.

Over the decades, there have been hundreds of studies estimating the effects of taxes on economic growth. Despite the fact that many studies use similar data and study many of the same states or countries, and time periods, estimates vary widely. The result is a lack of a consensus on whether taxes have an important influence on economic growth and, if they do, how large the effect might be. There are many possible reasons for this state of affairs. Tax policy is necessarily a two-sided activity. This is because taxes are used to fund expenditures and/or reduce deficits. As a result, tax effects are always net effects, and will differ depending on how the tax revenues are spent. Relatedly, different types of taxes may have different consequences for economic growth, as may different types of expenditures. For example, a distortionary tax on personal income used to fund transfer payments may be expected to have different growth effects than a non-distortionary tax on goods and services used to build productive infrastructure.[2] Further, the empirical models used to estimate tax effects may measure short-, medium-, or long-run effects depending on particular ways that regression equations are specified. For these and other reasons, even studies that use similar data can produce dissimilar estimates of tax effects.

In an attempt to provide a clear picture of the research papers investigating the effects of taxes on economic growth given the inclusion criteria, I conduct two meta-regression analyses. The first, contained in Chapter 2, is “Taxes and Economic Growth in OECD Countries: A Meta-Regression Analysis”. In this study, I use meta-regression analysis methods to evaluate the results of 42 studies and 641 individual estimates of the effect of taxes on economic growth in OECD countries. The second meta-regression analysis, presented in Chapter 3, is “Taxes and Economic Growth in U.S. States: A Meta-Regression Analysis”. As in the previous study, I employ meta-regression analysis techniques to synthesize 868 individual estimates derived from 29 studies investigating the effect of taxes on economic growth in U.S. states. The first two studies address a number of difficult coding issues including: implications of government budget constraints for interpretation of tax effects; units of measurement for economic growth rates and tax rates; implications of equation specifications that measure short-run, medium-run, and long-run effects; length of time period (annual data versus multi-year periods); as well as several others.

One of the main concerns in the meta-regression analysis literature is publication selection bias. Publication selection bias occurs when papers that are “published”, either in journals or that advance to the stage of publicly available working papers or reports, do not accurately represent the underlying empirical research. Publication selection bias occurs

[2] Distortionary taxes are those distorting the private sector’s incentive to invest such as taxes on personal income and property. An example of non-distortionary tax would be taxes on consumption. Accordingly, government expenditures are devided into two categories productive versus unproductive expenditures. If an expenditure are included in private production function, then it is assumed as productive expenditures such as educational expenditures. Welfare expenditure is an example for unproductive category (Kneller et al., 1999).

because there is a tendency amongst researchers, reviewers, and editors to submit or accept the manuscripts for publications based upon the direction or strength of the results. As the data for meta-regression analysis consists of estimated effects from the literature, if that distribution is distorted, so will the conclusions that derive from them. Thus, a crucial component of a meta-regression analysis is to detect and correct for publication bias.

In my third essay explained in Chapter 4, “How Well does the FAT-PET-PEESE Procedure Work?”, I study the performance of the FAT-PET-PEESE (FPP) procedure in presence of publication selection bias. This is a commonly employed approach for addressing publication bias in the economics and business meta-regression analysis literature. The FPP procedure is generally used for three purposes: (i) to test whether a sample of estimates suffers from publication bias, (ii) to test whether the estimates indicate that the effect of interest is statistically different from zero, and (iii) to obtain an estimate of the overall mean effect.

Chapter 2. Taxes and Economic Growth in OECD Countries: A Meta-Regression Analysis

2.1. Introduction

The effect of taxes on economic activity is one of the highly contested research areas in macroeconomics. Many studies have examined the effects of taxes on economic performance such as Agell, Lindh, and Ohlsson (1997); Mendoza, Milesi-Ferretti, and Asea (1997); Fölster and Henrekson (1999); Kneller, Bleaney, and Gemmell (1999); Daveri and Tabellini (2000); Bassanini and Hemmings (2001); Bleaney, Gemmell, and Kneller (2001); Fölster and Henrekson (2001); Afonso and Furceri (2010); Alesina and Ardagna (2010); and Arnold et al. (2011). But against expectation, there is no consensus among economists on whether taxes have any influential effect on economic growth, and if they do, how large the effect might be. While theory may not provide enough guidance on the ultimate effect of taxes on growth, so that the issue becomes an empirical one, the empirical results have a number of complications that make it challenging to draw general conclusions.

There are many possible reasons for the existence of a lack of consensus. Let's first see why there is no clear a priori theoretical prediction about the effects of taxes on economic growth. In the neoclassical growth model introduced by Solow (1956), fiscal variables such as taxes and spending may have transitional effects on output levels but they have no impact on the rate of economic growth in the long run. The steady-state growth rate is driven by exogenous factors such as the rate of technical progress and population growth. However, the endogenous growth model introduced by Barro (1990) and King and Rebelo (1990) challenged the traditional neoclassical growth model and predicted that long-run growth will be affected by productive expenditures and distortionary taxation. As taxes have no permanent effects on per capita GDP growth in the neoclassical model, most researchers assume that the endogenous model can better explain growth. Further, the reported growth effects of taxes depend not only on the type of taxes/expenditures considered (Barro, 1990;

Barro and Sala-i- Martin, 1992; Futagami et al., 1993; and Deverajan et al., 1996) but the net effect of taxes on growth also depends on how public spending and deficits are financed (Kneller et al., 1999; Bleanet et al., 2001; Gemmell et al., 2009). For example, a distortionary tax such as a progressive personal income tax used to fund unproductive expenditure such as transfer payments may have a different growth effect than the situation in which the same distortionary taxes are used to fund productive expenditure on public infrastructure.

Like the theoretical literature, empirical studies provide ambiguous results on the growth effects of tax policy due mainly to the lack of a uniform frame of reference. The difficulty in finding robust evidence of the effect of taxes on growth may be explained by several methodological choices, such as what countries to include, how to measure taxes and economic performance, the problem of omitted variables, particularly the exclusion of different types of expenditures, differences in the inclusion of control variables, the selection of estimation methods, and the duration of estimated tax effects. For these and other reasons, it is hardly surprising that these conflicting results exist.

Since the literature lacks any visible patterns, conventional narrative reviews can be used to compare estimates across studies and therefore highlight the reasons for the heterogeneity observed. However, these reviews suffer from the following shortcomings: (i) they reflect the reviewers' points of view and can certainly vary from one reviewer to another; (ii) bias might be an inherent part of these kinds of reviews; (iii) no clear inclusion/exclusion criteria are typically reported and therefore they cannot be replicated by other scholars; (iv) there is no objective standard for how to weight alternative estimates, and (v) as a result, they cannot be relied upon to provide clear and concrete guidance to policy makers and other researchers concerning the relationship in the research question.

To overcome the above-mentioned shortcomings and in order to be able to provide a clear picture of the existing literature investigating the effects of taxes on economic growth in OECD countries, I apply a meta-regression analysis (MRA). An MRA is a quantitative method for reviewing research of the existing literature in order to aggregate the empirical findings on a given research question. One of the main advantages of an MRA is that it allows one to disentangle various factors causing the conflicting results among researchers (Stanley, 2001). M-As have been used in the sciences and other disciplines, and MRAs have been used predominantly in economics.

To do so, I collect the estimates from this literature and carefully track the factors that can cause heterogeneity across studies and then by the use of this technique, I am able to compare and synthesize the estimates across the different studies.

This study aims to answer the following questions by applying a meta-regression analysis: (Q1) What is the overall, mean effect of taxes on economic growth? (Q2) Are some taxes (e,g., personal income taxes) more distortionary than others (e.g., value added taxes)? (Q3) Is there any empirical evidence to support the conventional wisdom that "distortionary taxes" used to fund "unproductive expenditures" are especially harmful for economic growth? (Q4) What are the factors causing researchers to encounter different or even contradictory results? As part of this research, I check for publication bias, by which I mean some estimates are disproportionately under-reported either due to statistical insignificance or for reporting the "wrong-direction" according to the associated theory (Stanley and Doucouliagos, 2012; Havranek and Irsova, 2012). I calculate an "overall tax effect" after accommodating and correcting for publication bias. It is worth mentioning that any measure of the "overall tax effect" on growth is not informative enough mainly because it encompasses estimated effects as a result of various kinds of fiscal policies. Accordingly, I compare estimated tax effects from two types of policies: (i) tax effects that are theoretically predicted to have a negative impact on economic growth versus (ii) tax effects that are theoretically predicted to have a positive impact on economic growth. The differences between these two sets of estimated tax effects will provide a measure of the impact of tax policy on economic growth.

To answer the aforementioned four research questions, this study collects 713 comparable estimates of tax effects on economic growth in OECD countries derived from 42 primary studies. According to a final sample of 641 estimates, I find strong evidence that the empirical literature suffers from negative publication bias. In other words, there is a tendency to over-report negative estimates. Once I control for this bias, I then calculate that the "overall effect" of taxes on economic growth is small and statistically insignificant. However, as mentioned earlier, this "overall tax effect" is not very informative because it includes estimated effects from different kinds of fiscal policies.

After accommodating and correcting for publication bias, once I turn to analysing different types of tax policies, I find evidence that the composition of fiscal policy matters. For example, increases in productive expenditures and/or government surpluses funded by non-distortionary taxes have a statistically significant, positive effect on economic growth. However, increases in unproductive expenditures funded by distortionary taxes and/or deficits have a statistically significant, negative effect on economic growth. These differences in the policy compositions may explain the heterogeneity reported among the literature. Further, I find weak evidence that progressive taxes on personal income are more growth-retarding than other types of taxes. Evidence regarding other types of taxes is mixed.

The remainder of this paper is organized as follows. Section 2.2 explains how I collected the sample of estimates. Section 2.3 discusses some of the reasons why studies of tax effects can produce different estimates. Section 2.4 represents my empirical results, addressing the above-mentioned research questions. Section 2.5 summarizes the main findings of this research.

2.2. Selection of Studies and Construction of Dataset

This meta-regression analysis collects estimated tax effects derived from all the studies estimating the following specification:

$$g = \alpha_0 + \alpha_1 tr + error, \tag{2.1}$$

where g is a measure of economic growth, tr is a measure of the tax rate, and the data are taken from OECD countries. I conducted a comprehensive research strategy including both electronic and manual search procedures. It is worth noting that studies estimating interaction and/or non-linear transformation of tax effects, such as the "growth hills" of Bania, Grey and Stone (2007) and also studies estimating interactive terms, such as Deskins and Hill (2010) are not included in this MRA mainly because if there is an interactive term in the model, the total effect is an outcome of both the term and its interaction. In these cases, the meta-analyst rarely has the data necessary to calculate the marginal effects and their respective standard errors.

The electronic search used three categories of keywords: (i) "TAX" keywords (ii) "ECONOMIC GROWTH" keywords, and (iii) "OECD" keywords in the following combination: "TAX" and "ECONOMIC GROWTH" and "OECD". A variety of keywords were substituted into each of the three categories. All the potential alternatives are reported in Appendix 2.1. I searched several keyword combinations in various electronic search engines such as EconLit, Google Scholar, JSTOR, Web of Science, Scopus, RePEc, EBSCO, and ProQuest. The primary search yielded a total of 303 papers.

The abstracts and conclusions of these studies were then read carefully to eliminate any studies that did not meet the inclusion/exclusion criteria. To be included in this meta-regression analysis each study needs to: (i) report an estimate of a growth equation with a tax variable; (ii) focus on a full set or a subset of OECD countries (e.g., EU15, G7, EU members); and (iii) provide standard errors (or the statistics through which standard errors can be computed) associated with each regression coefficient. Backwards and forwards citation search strategies were then applied to identify any additional relevant original studies. This produced a list of 51 studies, some of which were multiple versions of the same study, and included peer-reviewed journals, conference proceedings, reports released by government agencies, think tanks and research firms, theses and dissertations, and working papers and other unpublished or grey literature. [3]

The list including all the studies collected until that period was emailed to 64 scholars who had written at least one research paper on the topic of taxes and economic growth in OECD countries. The researchers were asked to assist me in identifying any additional research papers of their own or Masters/PhD students who are working with them.[4] The responses I received from the researchers resulted in a revised list of 54 studies.[5]

Each study in the revised list was then read thoroughly to see whether they were eligible according to the inclusion criteria defined at the earlier stage. The dependent variable had to be a measure of GDP growth. Alternatively, the dependent variable could be the level of income, as long as the lagged dependent variable was included in the specification. The growth equation had to include at least one tax variable that was measured in units of percent

[3] When reported estimates differ in multiple versions of the study, the peer-reviewed journal articles is considered as a benchmark. However, if there are additional estimates in previous versions of the study, I kept track of the outlet of the study, coded, and then pooled the estimates across versions.
[4] The letter along with the bibliography of the core studies emailed to the prominent authors in this research is available in Appendix 2.2 and Appendix 2.3.
[5] I am grateful for helpful suggestion received from all the scholars.

of income. Studies in which the "tax variable" consisted of all revenues, such as the ratio of total revenues to GDP, were not included. This is because they lump together tax and non-tax revenues. The countries included in a given regression equation had to consist of a full set of OECD countries, though they could be restricted to a subset of OECD countries – all, G7, EU-15 or a larger set of EU member nations. Further, all studies that included only a single country were dropped from this meta-study. To be included, estimates had to include multiple countries. The reason being that it was felt that aggregating the growth experiences across multiple countries provide the greatest opportunity to generate externally valid results. They also offer more degrees of freedom which improves the efficiency of the economic estimates. All estimated tax effects had to report standard errors or associated t-statistics/p-values. Finally, only studies written in English were included. I closed my search on 13 January 2016. The final sample of 42 studies is listed in Appendix 2.4.[6]

Once the final set of estimates was determined, I then went through each equation/estimate and coded a set of regression and study characteristics (more details provided in the next section). The coding was done independently by at least two coders with a careful reconciliation of any discrepancies or inconsistencies.[7] All search and coding procedures followed the guidelines for the reporting of MRA studies (Stanley et al., 2013).

2.3. Factors that Cause Tax Estimates to Differ Across Studies

<u>The government budget constraint.</u> To estimate the precise effects of taxes on economic growth it is important to address a number of issues. The first and foremost is how to deal with the government budget constraint:

$$0 = Taxes + OtherRevenues - Expenditures - Surplus \tag{2.2}$$

[6] Appendix 2.5 clarifies the steps which is undertaken to reach to the 42 final studies.
[7] The two plus coders includes myself, a PhD student recruited as a research assistant, and Prof. Reed to provide us the right direction once there is discrepencies in the reconciliation process.

The following specification is obtained by dividing both sides by $Income$:

$$0 = tr + \left(\frac{Other\ Revenues}{Income}\right) - \left(\frac{Expenditures}{Income}\right) - \left(\frac{Surplus}{Income}\right), \tag{2.3}$$

where the tax rate is considered as the ratio of taxes over income, $tr = \left(\frac{Taxes}{Income}\right)$.

The regression coefficient can be misinterpreted easily if one ignores the role of the government budget constraint. The main argument is that the regression coefficient on α_1 in Equation (2.1) should be interpreted as the growth effect of tax financed by the omitted categories and it may differ depending on which category(ies) has been omitted from the regression. If $\left(\frac{\text{Expenditures}}{\text{Income}}\right)$ is omitted, then α_1 measures the net effect of an increase in expenditures funded by taxes. Alternatively, if $\left(\frac{Surplus}{Income}\right)$ is omitted and expenditures are held constant, then α_1 measures the net effect of an increase in taxes used to cut the deficit (or increase the surplus).

The interpretation becomes even more complicated once taxes and expenditures are decomposed into their parts: distortionary versus non-distortionary taxes; productive versus unproductive expenditures. Table 2.1 summerizes the predicted tax-growth effects once one categorizes taxes and public expenditures into its components.

Table 2.1: Barro's predicted tax-growth effects

Financed by:		Public Spending		Deficit
		Productive	Unproductive	
Taxes	Distortionary	+/-	-	+/-
	Non-distortionary	+	0	+

Source: Adapted from Gemmell et al., 2009.

This can be seen in the following specification:

$$0 = \; tr(Non - distortionary) + tr(Distortionary) + \left(\frac{Other\ Revenues}{Income}\right) - \left(\frac{Productive\ Expenditures}{Income}\right) - \left(\frac{Unproductive\ Expenditures}{Income}\right) - \left(\frac{Surplus}{Income}\right) \tag{2.4}$$

If $\left(\frac{Productive\ Expenditures}{Income}\right)$ is omitted, the coefficient on the non-distortionary tax rate variable measures the net effect of an increase in productive expenditures funded by an increase in non-distortionary taxes. As discussed below, it is generally accepted that growth theory predicts a positive value for α_1 in this case. In contrast, if $\left(\frac{Unproductive\ Expenditures}{Income}\right)$ is omitted, the coefficient on the distortionary tax rate variable measures the net effect of an increase in unproductive expenditures funded by an increase in distortionary taxes. In this case, a negative value for α_1 would be expected. As a result, the two "tax rate" variables might legitimately produce opposite signs by virtue of the kind of tax variable that is being investigated, and depending on which other variables in the government budget constraint are omitted.

To address this issue, I go through each estimated tax effect and identify both the operative tax types and the use of the tax revenues implied by the government budget constraint. Tax types and expenditures are then categorized as distortionary/non-distortionary, productive/unproductive, or other according to the taxonomy provided in Table 2.2, taken from Kneller, Bleaney, and Gemmell (1999).[8]

[8] I use the Kneller, Bleaney, and Gemmell (1999) taxonomy because it is broadly representative of the fiscal policy literature. It may be best thought of as representing relative categories. Distortionary taxes are those distorting investment decisions (Barro, 1990).

Table 2.2: Matching of Functional and Theoretical Classifications

Functional classification	***Theoretical classification***
Taxation on income and profit Social security contributions Taxation on payroll and manpower Taxation on property	Distortionary taxation
Taxation on domestic goods and services	Non-distortionary taxations
Taxation on international trade Non-tax revenues Other tax revenues	Other revenues
General public services expenditure Defense expenditure Educational expenditure Health expenditure Housing expenditure Transport and communication expenditure	Productive expenditures
Social security and welfare expenditure Expenditure on recreation Expenditure on economic services	Unproductive expenditures
Other expenditures (unclassified)	Other expenditures

Note: The categorizations in the table are taken from Kneller, Bleaney, and Gemmell (1999).

Table 2.3 summarizes the predicted effect of distortionary/non-distortionary taxes on economic growth given the omitted fiscal category. This is taken from Gemmell, Kneller, and Sanz (2009), however, it is adjusted to accommodate the totality of cases encountered in my sample. Accordingly, every estimated tax effect in my sample is assigned a predicted effect with respect to its impact on growth (negative, positive, or ambiguous/zero).

Table 2.3: Predicted Tax Effects

Type of Tax	***Omitted Fiscal Category***	***Predicted Effect***
Distortionary	Productive Expenditures	Ambiguous
Distortionary	Unproductive expenditures	Negative
Distortionary	All the expenditures(Pro&Unpro)	Ambiguous
Distortionary	Other Expenditures	Ambiguous
Distortionary	Deficit/Surplus	Ambiguous
Distortionary	Other Revenue	Ambiguous
Distortionary	Distortionary Taxes	Ambiguous
Distortionary	Non-distortionary Taxes	Negative
Distortionary	Intergovernmental Revenue	Ambiguous
Distortionary	Net Utility Expenditures	Ambiguous
Non-distortionary	Productive Expenditures	Positive
Non-distortionary	Unproductive Expenditures	Ambiguous
Non-distortionary	Productive & Unproductive Expenditures	Ambiguous
Non-distortionary	Other Expenditures	Ambiguous
Non-distortionary	Deficit/Surplus	Positive
Non-distortionary	Other Revenue	Ambiguous
Non-distortionary	Distortionary Taxes	Positive
Non-distortionary	Non-distortionary Taxes	Ambiguous
Non-distortionary	Intergovernmental Revenue	Ambiguous
Non-distortionary	Net Utility Expenditures	Ambiguous

Note: The categorizations in the table are taken from Gemmell, Kneller, and Sanz (2009), where I combine the original categories of "zero" and "ambiguous" to "ambiguous".

There is another possible classification, in this case according to tax types. Taxes are classified as Labour taxes, Capital taxes, Consumption taxes, Mixed taxes, Other taxes, and Overall taxes. The classification system for assigning each tax to a tax type is presented in Table 2.4.

Table 2.4: Types of Taxes

Tax Type	***Examples***
Labour	Personal income tax Payroll tax Social security contributions
Capital	Corporate income tax Capital tax (tax on dividends)
Consumption	Consumption tax Taxes on goods and services Sales tax Value added tax (VAT) International trade tax
Other tax	Property tax Taxes not listed above
Mixed tax	Taxes that are a combination of the above types
Overall tax	Total taxes (e.g., Total Tax Revenues/GDP)

Units of measurement. The second issue that deserves careful attention is the units of measurement for both economic growth (g) and tax rate (tr) variables. Each of these variables can be measured in percentage points (e.g., 10%) or in decimals (0.1). This will clearly effect the size of the tax coefficient, α_1. For example, if a one-percentage point increase in the tax rate lowers growth by 0.1%, and if both g and tr are measured in percentage points, or both are measured in decimals, then the corresponding value of α_1 will be -0.1. However, if g is measured in percentage points, and tr is measured in decimals, then the corresponding value of α_1 should be multiplied by 100 and therefore the corresponding effect will be -10. And if g is measured in decimals, and tr is measured in percentage points, then the value of α_1 should be divided by 100 and therefore the corresponding effect will be -0.001. Accordingly, I adjust all estimated effects so that $\alpha_1 = X$ means that a one-percentage point increase in the tax rate is associated with an X percentage point increase in economic growth. If the original study lacks summary/descriptive statistics or the proper interpretation of the estimated results, it would be then difficult to determine the measurement units. In these cases, I contacted the author(s) to cross check the units. Those estimates were dropped from my analysis in the rare cases (one study) where I was unable to locate the author(s), or they did not respond to my emails.

Countries. The third issue has to do with the specific countries included in a given study. While the countries considered as an OECD member are fairly homogeneous, this grouping also involves developing countries such as Turkey. OECD membership is granted on the basis of both (i) economic performance and (ii) democratic and institutional development. Heterogeneities across OECD countries may yield systematically different results. Some of the studies available in the literature limit their sample to a sub-set of OECD countries including G-7, EU-15, and EU, with the idea that those subsets consist of more homogeneous countries. Appendix 2.6 lists the 34 OECD countries, ordered by their year of admission to the OECD.[9] This meta-regression analysis controls for these different groupings to identify whether the estimated tax effects vary systematically across the different sets of countries included in the original studies.

Duration of time periods. A fourth issue concerns the time frames of the data employed in the original studies. If the time periods of Equation (2.1) differ across studies, that could

[9] Latvia, the 35th member, was admitted to the OECD on July 1st, 2016.

cause estimates of α_1 to differ, even when the underlying effect is the same. For example, suppose there were two growth studies, one used 5-year time periods, the other used annual data. Suppose the former measured the cumulative rate of growth over each five-year period, while the latter reported annual growth rates. All things constant, one might expect α_1 to be larger in the former case. Accordingly, I adjust all growth measures to be (average) annual rates of growth.

Duration of estimated tax effects. Since most growth models agree that tax-growth effects occur in the short-run, the distinction between short-, medium-, and long-run effects of tax may explain discrepancies observed in the literature. Thus, a fifth issue has to do with the duration of the estimated tax effect as implied by the specification of the regression equation. Let the estimated relationship between growth, g, and the tax rate variable, *tr*, be given by the finite distributed lag model,

$$g_t = \alpha_0 + \alpha_1 tr_t + \alpha_2 tr_{t-1} + \varepsilon_t. \tag{2.5}$$

If this is the model estimated by the original study, then α_1 and α_2 represent the "short-run/immediate" effects of a one-percentage point increase in taxes in years *t* and *t-1* on economic growth in year *t*.

By adding and subtracting $\alpha_2 tr_t$ to the right hand side, one can rewrite the above as:

$$g_t = \alpha_0 + \tau\, tr_t - \alpha_2 \Delta tr_t + \varepsilon_t, \tag{2.6}$$

where $\tau = (\alpha_1 + \alpha_2)$. If this is the model estimated in the original study, then the coefficient on the current tax rate, τ, represents the "cumulative/intermediate" effect of a one-percentage point increase in taxes in year *t* and *t-1* on economic growth in year *t*.

An alternative specification to Equation (2.5) is the auto-regressive, distributed lag model,

$$g_t = \alpha_0 + \alpha_1 tr_t + \alpha_2 tr_{t-1} + \gamma g_{t-1} + \varepsilon_t. \tag{2.7}$$

Subtracting g_{t-1} from both sides gives:

$$\Delta g_t = \alpha_0 + \alpha_1 tr_t + \alpha_2 tr_{t-1} + (\gamma - 1) g_{t-1} + \varepsilon_t, \tag{2.8}$$

which can be rewritten in error correction form as:

$$\Delta g_t = \alpha_0 + \delta(g_{t-1} - \theta tr_t) - \alpha_2 \Delta tr_t + \varepsilon_t\ , \tag{2.9}$$

where $\delta = (\gamma - 1)$ and $\theta = \frac{(\alpha_1 + \alpha_2)}{(1-\gamma)}$. This specification is common in recent mean group and pooled mean group studies of economic growth. In Equation (2.9), the coefficient on tr_t in the cointegrating equation, θ, represents the total, long-run effect of a permanent, one-percentage point increase in the tax rate on steady-state economic growth.[10]

Specifications (2.5), (2.6), and (2.9) lead to three different measures of the effect of taxes on economic growth. My meta-regression analysis controls for this by noting the specification of the growth equation in the original study and categorizing the duration of the estimated tax effect as short-run, medium-run, or long-run.

<u>Different measures for economic growth and tax rates</u>. A final issue to be addressed is how the primary studies define the tax rate and economic growth variables. While some studies use nominal GDP as a measure of economic growth, others use real GDP. I keep track of both measures, however, because as long as a given study applied the nominal GDP (in log form) and also included time dummies then there is no distinction between nominal and real GDP. Per capita GDP and total GDP are the other forms of measuring economic growth in

[10] I have noticed that Equation (2.9) is sometimes estimated using an equivalent, alternative specification: $\Delta g_t = \alpha_0 + \delta(g_{t-1} - \theta tr_{t-1}) + \alpha_1 \Delta tr_t + \varepsilon_t$, where δ and θ are defined as above.

the literature. One of the main challenges faced by empirical studies investigating the effects of tax is how to identify an accurate measure of tax rates (Mendoza et al., 1997). Since economic decisions depend on the marginal effective tax rate, this measure is more appropriate for examining the tax-growth effects. However, marginal effective tax rates are not observable and there is no obvious estimate of them. Therefore, several proxies have been proposed in the literature. The most commonly used proxy is "tax burden" defined as tax revenues over a given measure of income. But this specification creates a potential collinearity with government expenditures (Easterly and Rebelo, 1993). The other available alternatives are average effective tax rates and statutory tax rates - typically the top marginal rate. These are more sophisticated measures. The last two measures are believed to perform better as opposed to the former in capturing the complexity of the tax system. And some studies attempt to distinguish marginal from average tax rates. I use dummy variables to indicate the specific measures underlying a given estimate.

Control variables. In addition to the issues explained earlier, I code many other study characteristics. These include estimation methods, types of standard errors, whether the original study was published in a peer-reviewed journal, the publication date, the sample period length, the midyear of the sample period, and whether specific variables such as country fixed effects, human capital, trade openness, inflation, and others are included in the estimating equations. A full list of the variables used in this study is discussed in the next section.

2.4. Empirical Analysis

Preliminary analysis. My search strategy identified 42 comparable empirical studies that offer regression based estimates of tax-growth effects. By coding various characteristics discussed earlier I was able to produce a dataset including 713 estimated tax effects. Table 2.5 reports descriptive statistics for both these estimates and the associated *t*-statistics.[11] For the full dataset, the median estimated tax effect is -0.073, implying that a ten percentage point increase in the tax rate is associated with a 0.73 percentage point decrease in annual economic growth. This should be compared to an average, annual growth rate for OECD countries of approximately 2.5 percent over the period 1970-2000, a period which roughly corresponds to the "average" sample period of the studies included in this meta-regression analysis.[12,13] The median *t*-statistic is -1.27.

Table 2.5 indicates that the estimated tax effects reported in primary studies range from a minimum of -3.52 to a maximum of 12.72. The max value seems unreasonable given the annual average growth rate of 2.5 percent. It suggests that a one percentage point increase in the tax rate is associated with over a 12 percentage point increase in annual growth rate, ceteris paribus. I check other unreasonable estimates to avoid any potential coding errors. Some of the primary studies report estimates that could be considered, in the context of the average growth rate, outliers. These outlier estimates can lead to inflated error rates and substantial distortion of the coefficients and their associated statistical significance, so I delete the top and bottom 5 percent of estimates and as a result obtain a sample including 641 tax effects. Accordingly, the subsequent analysis works with a truncated sample of estimates (641 estimates) rather than initial full set (713 estimates).

The descriptive statistics for the truncated sample are also reported in Table 2.5. The range of estimated tax effects for this sample is from a minimum of -0.524 to a maximum of 0.166 which seems reasonable. The median t-statistic still indicates insignificance, while the

[11] Excel spreadsheet that allows the user to replicate all the results of Table 2.5 through 2.11 can be downloaded from Dataverse: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/KNQEYB.
[12] This is calculated by taking the average beginning and average ending dates for the sample ranges of the respective studies.
[13] Growth rate is the average, annual growth rate over the period 1970-2000 for the 22 countries that belonged to the OECD in 1970.

sample of t-statistics ranges from a minimum of -14.50 to a maximum of 7.78, with a mean absolute value of 2.09.

Table 2.5: Descriptive Statistics for Estimated Effects and t-statistics

	Estimated Tax Effects		***t-statistics***	
	Full	***Truncated***	***Full***	***Truncated***
Mean	-0.097	-0.109	2.16*	2.09*
Median	-0.073	-0.073	-1.27	-1.32
Minimum	-3.520	-0.524	-14.50	-14.50
Maximum	12.720	0.166	8.03	7.78
Std. Dev.	0.649	0.147	2.49	2.35
1%	-1.320	-0.480	-7.91	-8.29
5%	-0.530	-0.420	-6.17	-6.18
10%	-0.411	-0.342	-4.72	-4.72
90%	0.078	0.041	1.07	0.67
95%	0.167	0.082	1.67	1.25
99%	0.820	0.143	4.59	3.09
Obs	713	641	713	641

Figure 2.1 plots the estimated tax of the truncated sample. If tax effects were homogeneous across studies and sampling error is the only reason making the estimated effects differ, one would then expect a bell-shaped (standard) histogram. However, as can be clearly seen in Figure 2.1 this is not the case, implying that the distribution is skewed towards negative values. This histogram can also be used to identify if there is any publication selection in the literature. Lack of symmetry in this plot suggests that there might be publication selection bias towards negative estimates. The results from this simple visual test should also be confirmed using a more formal test (i.e. the Funnel Asymmetry Test).

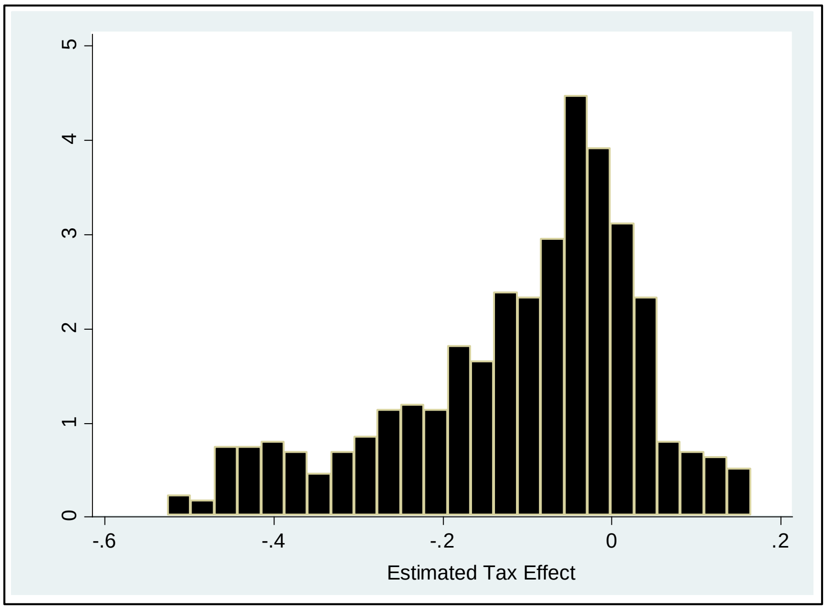


Figure 2.1: Histogram of Estimated Tax Effects (Truncated)

Figure 2.2 depicts a forest plot of the respective studies using a "Fixed Effects" weighting scheme. Note that the concept of both "Fixed Effects" and "Random Effects" in the meta-regression analysis context is quite different from the definitions used in the panel data literature (Reed, 2015). In the current context it simply means that the estimated tax effects are weighted by the inverse of their standard errors. For each study, a weighted average along with a 95 percent confidence interval is computed.

Looking at Figure 2.2, there are a couple of points which deserve particular attention. First, most of the studies estimate small effects with tight confidence intervals, although, study 39 (Abd Hakim et al., 2013) is a notable exception with respect to the confidence intervals. Second, there is a large amount of heterogeneity across studies, given the large I^2 computed and represented at the bottom of this figure (Higgins and Thompson, 2002). As discussed earlier, there are several reasons that may explain why this is the case. These include different measures of tax rate and economic growth, how primary studies deal with government budget constraints (GBC), different time periods as well as different samples of countries, differences in estimation methods applied, whether the effect is short-, medium, or long-run and so on. The large value of I^2 suggests that the heterogeneity across studies is far beyond just sampling error.[14]

Third, the last column calculates the percentage weight assigned to each study in calculating the overall weighted average. Study 26 (Hanson, 2010) is weighted substantially larger than all the other studies combined (81.39% versus 18.61%). The disproportionately large weight assigned to study number 26 is not a real concern as long as this study is truly more reliable. However, it might be a good reason to switch to the "Random Effects" weighting scheme.

[14] I^2measures heterogeneity, $I^2 = \sigma_h^2/(\sigma_h^2 + \sigma_\varepsilon^2)$. $I^2$is analalous to R^2in regression analysis.

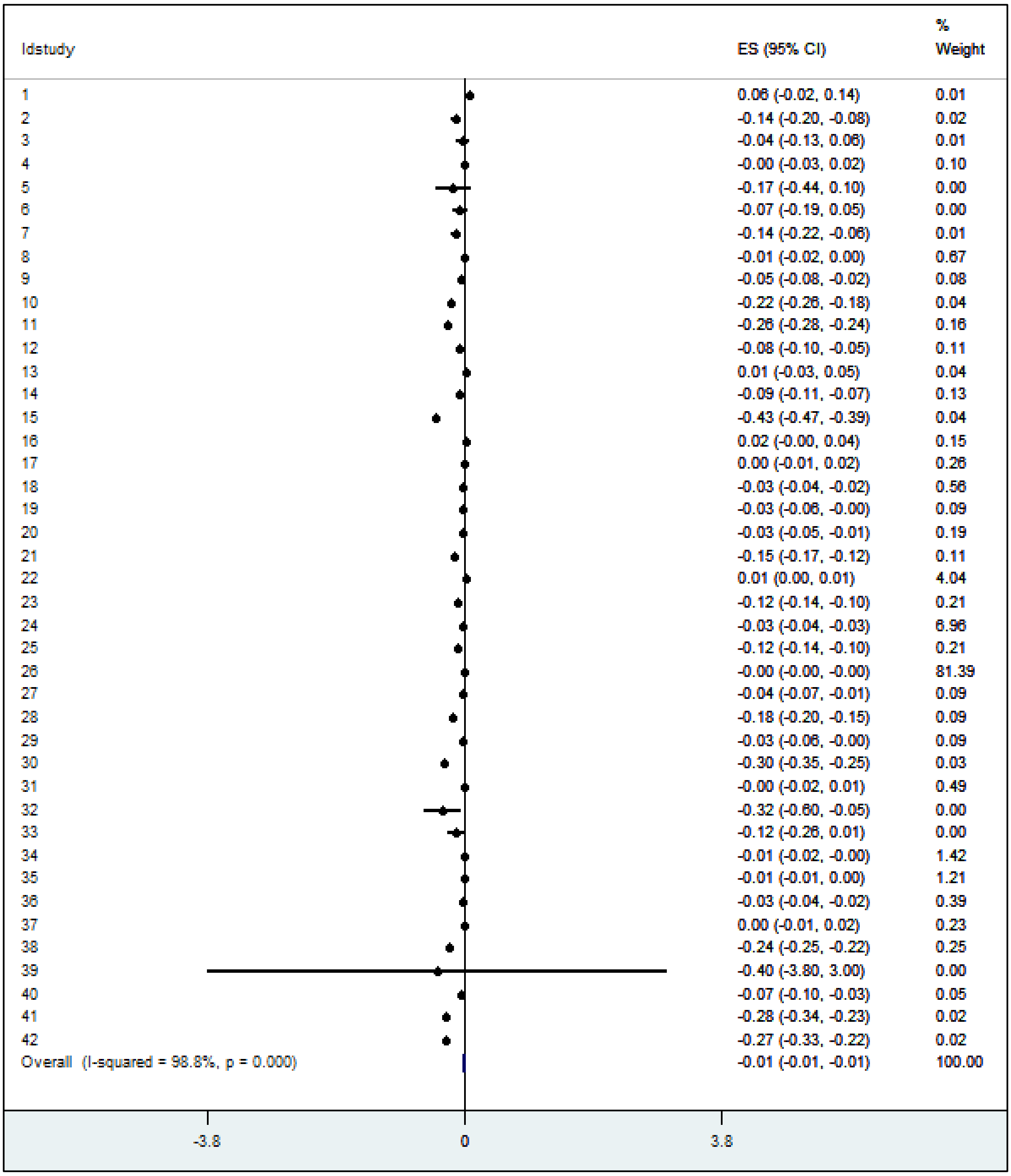


Figure 2.2: Forrest Plot of Studies (Fixed Effects)

The general assumption under the "Fixed Effects" framework is that there is an identical true effect size across all studies included in an MRA, and the only reason estimates differ is because of sampling error. Thus, it is not a concern if the estimates in the larger studies receive substantially more weight, because their "signal" is less distorted by "noise," since the estimates are more precise. In this framework, the optimal weight to assign each estimate is the inverse of its standard error.

In contrast, the general assumption under the "Random Effects" framework is that there is not just one true effect but a distribution of effects. This means that we cannot simply ignore a small study by assigning a smaller weight because these studies provide valuable information about the distribution of effects. Note that the weight implemented in the "Random Effects" model consists of two parts: (i) within-study variances (same as FE), and (ii) between-study variances (Borenstein et al., 2010).

Accordingly, the subsequent empirical work in this chapter emphasizes the "Random Effects" estimates where tax effects are weighted by their standard error (within-study heterogeneity) plus another term that captures the between-study heterogeneity. This will have the effect of equalizing the weights given to individual studies because cross-study heterogeneity is so great. Figure 2.3 displays the forest plot using "Random Effects". The study weights are much more balanced.

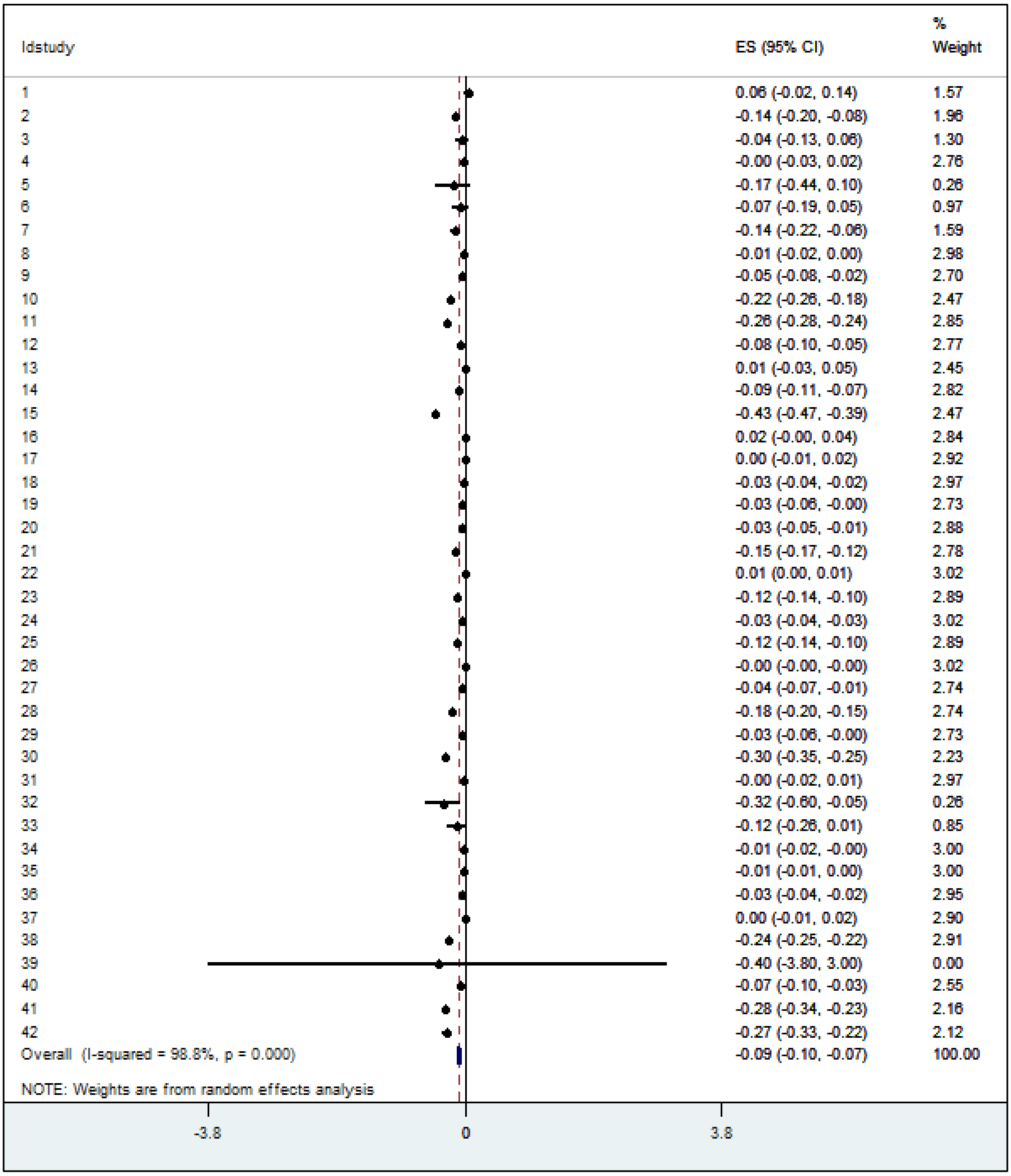


Figure 2.3: Forrest Plot of Studies (Random Effects)

The distribution of the reported estimates is illustrated in Figure 2.4 and Figure 2.5 in a form of a funnel plot. The funnel plot is a scatter diagram of effect sizes (here regression coefficients of the tax rate variable) versus some measure of their precision, typically the inverse of the standard error $(1/SE_i)$. It can be used as a simple visual tool to identify if there is any publication selection bias available in the literature (Stanley and Doucouliagos, 2010). It also provides further insight into the distribution of estimated tax effects.

Figure 2.4 displays individual estimates. In Figure 2.5 each study is represented by a single point relating its mean estimate to its mean standard error.[15]

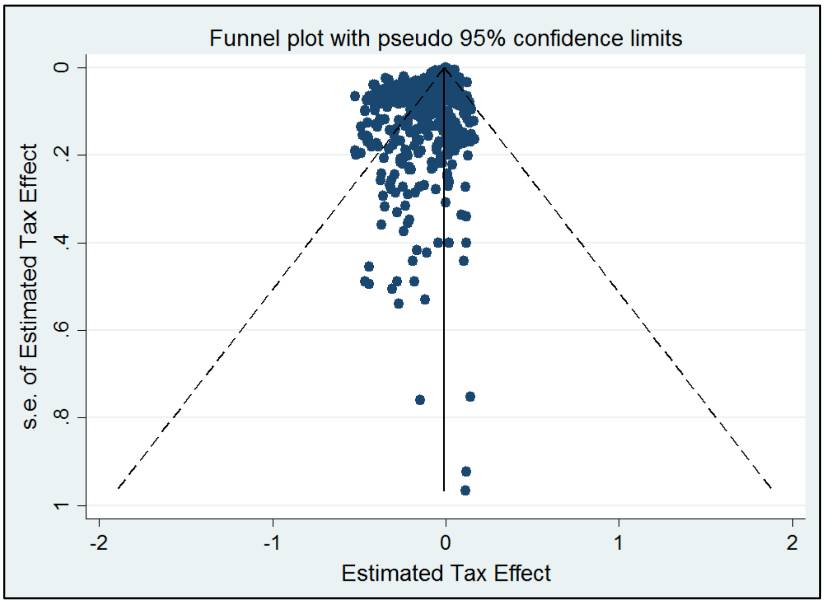


Figure 2.4: Funnel Plot (All Estimates)

[15] Both funnel plots omit observations where the standard error is greater than 1. This allows one to better observe the pattern of points at the top of the funnel.

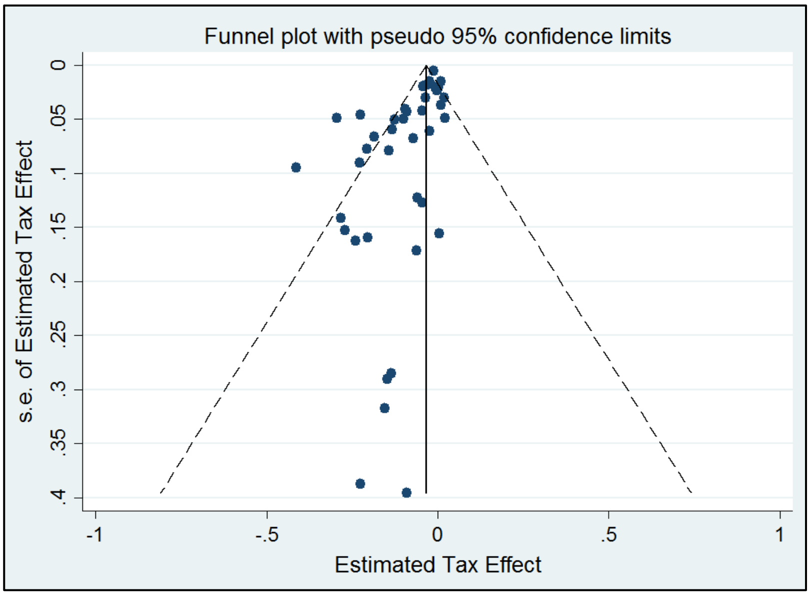


Figure 2.5: Funnel Plot (Mean of Study Estimates)

The solid line in both plots indicates the mean of estimated tax effects, and the dash lines that fan out from the top of the funnel shows the 95% confidence area where most of the estimates would fall if the dispersion in estimates was driven solely by sampling error. Publication bias is indicated whenever a disproportionate number of estimates lie on one side of the inverted, V-shaped confidence area. Both funnel plots suggest there is publication bias in favour of negative estimates. Further, the wide dispersion at the top of the funnel is consistent with substantial heterogeneity previously shown by the I^2 value reported with the forest plot.

FAT/PET tests. Table 2.6 reports the results of two tests: the Funnel Asymmetry Test (FAT) which is a conventional way to detect whether the literature suffers from publication selection bias (Egger et al., 1997; Stanley, 2008), and the Precision Effect Test (PET), which tests for the significance of the overall effect (Stanley and Doucoulliagos,2012; Shemilt et al., 2011). Both tests are obtained from estimating the following specification using weighted least squares (WLS),

$$\hat{\alpha}_{1,ij} = \beta_0 + \beta_1 SE_{ij}, \tag{2.10}$$

where $\hat{\alpha}_{1,ij}$ is the estimated tax effect from regression j in study i. The null hypotheses for the FAT and PET are $H_0\colon \beta_1 = 0$ and $H_0\colon \beta_0 = 0$, respectively.

My analysis uses four different weights to estimate Equation (2.10). The "Fixed Effects" and "Random Effects" estimators use weights $\left(\frac{1}{SE_{ij}}\right)$ and $\left(\frac{1}{\sqrt{\left(SE_{ij}\right)^2+\tau^2}}\right)$, respectively, where τ^2 is the estimated variance of population tax effect across studies. This set of weights ignores the fact that some studies report more estimates compared to others. As a result, a study including 10 estimates would be weighted 10 times more than a study including one single estimate, ceteris paribus. To address this issue, I multiply both sets of weights by the inverse of the number of estimates reported per study, $\left(\frac{1}{N_i}\right)$. Doing so, I assign each given study approximately the same weight as others even though the number of reported estimates differs from one study to another. Thus, "Weight 1" refers to the standard weighting scheme in which the number of reported estimates matter and studies with higher number of estimates receive the higher weight. However, by using "Weight 2" I assign each study the same importance.

Table 2.6: Funnel Asymmetry and Precision Effect Test (FAT/PET)

	Fixed Effects (Weight1) (1)	***Fixed Effects (Weight2) (2)***	***Random Effects (Weight1) (3)***	***Random Effects (Weight2) (4)***	***Random Effects (Weight1)[16] (5)***	***Random Effects (Weight2)[16] (6)***
(1) FAT	-1.660*** *(-5.47)*	-1.562*** *(-6.00)*	-1.245*** *(-3.31)*	-1.462*** *(-4.60)*	---	---
(2) PET	-0.001 *(-0.58)*	0.000 *(0.16)*	-0.001 *(-0.04)*	0.018 *(1.18)*	-0.065*** *(-4.27)*	-0.053*** *(-4.34)*
Observations	641	641	641	641	641	641

Note: Values in Row (1) and Row (2) come from estimating β_1 and β_0, respectively, in Equation (2.10) in the text. In both cases, the top value is the coefficient estimate, and the bottom value in parentheses is the associated *t*-statistic. The four WLS estimators (*Fixed Effects-Weight1, Fixed Effects-Weight2, Random Effects-Weight1,* and *Random Effects-Weight2*) are described in the "FAT/PET tests" subsection of Section 2.4 in the text. All four of the estimation procedures calculate cluster robust standard errors. *, **, and *** indicate statistical significance at the 10-, 5-, and 1-percent level, respectively.

Heteroskedasticity is always an issue for meta-regression analysis, because the original estimates, which are the dependent variable, come from very different datasets with different sample sizes and different estimation techniques. Thus, some variation of weighted least squares (WLS) should always be employed. Furthermore, authors in this literature typically report multiple estimates and therefore estimates within the study cannot be assumed independent. To account for these data complexities, the first four columns of Table 2.6 report the results of estimating Equation (2.10) using WLS with respect to the four different weighting schemes described above, and calculating cluster robust standard errors, with clustering by study. The FAT is reported in the first row. For all four estimators, the null hypothesis of no publication bias is rejected at the 1 percent level of significance. The negative coefficients imply that there is a selection bias in favour of negative estimated tax

[16] Column 5 and 6 report random effects estimates once it is not corrected for publication selection bias.

effects, perhaps due to researchers choosing to disproportionately report negative estimates, or reviewers in peer-reviewed journals discriminating against positive results. These results are consistent with earlier visual inspection of the estimated effects histogram and also the visual evidence of publication bias from the funnel plots represented in Figure 2.4 and 2.5.

The first four columns of the second row of Table 2.6 report the PET. All four estimators show that the overall tax effect, controlling for publication bias, is statistically insignificant and relatively small in economic terms. According to the "Random Effects (Weight1)", a 10-percentage point increase in the tax rate is associated with a 0.01 percentage point decrease in annual GDP growth, ceteris paribus.

The last two columns report random effects estimates of Equation (2.10) when the publication bias term (SE_{ij}) is not included, so that the overall estimate is not corrected for publication bias. The corresponding estimates of the overall tax effects are now substantially larger in absolute value (compared to previous results), and statistically significant at the 1 percent level. According to the "Random Effects (Weight1) in Column (5), a 10-percentage point increase in the tax rate is associated with a 0.65 percentage point decrease in annual GDP growth. These results indicate that the statistically and economically significant results reported in the literature are influenced by negative publication bias. Once one controls for that, the estimated tax-growth effect is substantially smaller and statistically insignificant. As a result, I want to be sure that my subsequent analysis corrects for this.

This section has addressed one of the main objectives of this research, to obtain an "overall estimate" of the effect of taxes on economic growth in OECD countries. I find that once I correct and accommodate for publication bias then the overall effect on taxes is statistically insignificant and negligibly small in economic terms. However, my previous discussion on factors that cause tax estimates to differ across studies (cf. Section III) makes clear that any estimate of overall tax effects is not particularly meaningful. The same fiscal policy intervention can be estimated as a positive or negative tax effect depending on the omitted fiscal categories from the primary study's regression equation. Accordingly, the next section undertakes a meta-regression that allows tax effects to vary systematically according to study and data characteristics.

Meta-regression. Section 2.3 identified factors that may cause heterogeneity in the reported estimates. In this section I compare tax effects associated with fiscal policies that are predicted to have negative growth effects with those predicted to have positive effects. I also investigate whether some types of taxes are more growth-retarding than others. To do that, it will be necessary to control for the factors that may influence estimates of tax effects.

Table 2.7 reports the variables used in the subsequent meta-regression analysis. The first set of variables were previously discussed and match each tax effect to a prediction. A little more than a fourth of the estimated tax effects allow a definite sign prediction, with 22.8 percent predicted to be negative, 5.9 percent predicted to be positive, and the rest ambiguous. As these three variables comprise the full set of possibilities, at least one variable must be omitted in the empirical analysis. Here and elsewhere in the table, I indicate the omitted variable with an asterisk.

The second set of variables assigns each tax effect to one of six types of taxes (Labour, Capital, Consumption, Other, Mixed, and Overall). The most common tax variable is constructed by taking the ratio of total tax revenues over GDP. Approximately 34.5 percent of tax effects are of this type. However, many studies disaggregate tax effects into separate types. For example, 18.6 percent of estimated tax effects involve Labour taxes (e.g., personal income taxes, payroll taxes, social security contributions). Another 12.5 percent are associated with Capital taxes (e.g., corporate income taxes, taxes on capital gains and dividends) and 13.3 percent are related to Consumption taxes (e.g., ad valorem taxes on goods and services, VAT). The remainder of tax effects mostly involve a mix of different types of taxes.

Other variables are grouped according to the following categories: Country Group, Economic Growth Measure, Tax Variable Measure, Duration of Tax Effect, etc. Most of the observed tax effects are estimated using data from the larger set of OECD countries (78.8%), as opposed to smaller groupings such as the G-7 countries (11.7%) or EU countries (6.4% and 3.1%). In most cases economic growth is measured in per capita terms (74.1%). Most taxes are measured as average rather than marginal rates (91.0% versus 9.0%); are specified in level rather than differenced form (82.8% versus 17.2%); and are effective rather than statutory tax rates (90.6% versus 9.4%). Most estimated tax effects measure the immediate effect of a tax change (70.2%) versus a medium- or long-run effect (5.3% and 24.5%).

Two thirds of the estimated tax effects in my meta-regression come from peer-reviewed journal articles and the mean year of publication was 2007. Almost all of the original studies used panel data to estimate tax effects (99.1%). The average sample length in the original studies was 31.4 years, and the average mid-point was 1985. About two-thirds of the tax effects were estimated using Ordinary Least Squares (OLS) or a related procedure that assumed errors to be independently and identically distributed across observations (such as mean group or pooled mean group procedures). Of the remainder, 15.4 percent used Generalized Laest Squares (GLS), and 16.8 percent attempted to correct for endogeneity using a procedure such as Two-Stage Least Squares (2SLS) or Generalized Method of Moments (GMM).

I categorized standard errors into three groupings because the standard error plays such a significant role in meta-regression analysis: *SE-OLS* (58.7%); *SE-HET* (24.5%), where standard errors were estimated using a heteroskedastic-robust estimator; and *SE-Other* (16.8%), whenever allowance was made for off-diagonal terms in the error variance-covariance matrix to be nonzero. Lastly, dummy variables were used to indicate the presence of important control variables, the most common of which were country fixed effects (83.3%), and measures of investment (58.5%), initial income (55.9%), human capital such as educational achievement (44.0%), employment growth (37.8%), and population growth (24.3%).

Table 2.7: Summary Statistics of Study Characteristics

Variable	***Description***	***Mean***	***Min***	***Max***
	PREDICTED TAX EFFECTS			
Prediction-Negative	=1, if the theoretical prediction of the coefficient is negative	0.228	0	1
*Prediction-Ambiguous**	=1, if the theoretical prediction of the coefficient is ambiguous	0.713	0	1
Prediction-Positive	=1, if the theoretical prediction of the coefficient is positive	0.059	0	1
	TAX TYPE			
Labour-Tax	=1, if labour tax	0.186	0	1
Capital-Tax	=1, if capital tax	0.125	0	1
*Consumption-Tax**	=1, if consumption tax	0.133	0	1
Other-Tax	=1, if other type of tax	0.005	0	1
Mixed-Tax	=1, if multiple tax types (but not overall tax)	0.207	0	1
Overall-Tax	=1, if overall tax	0.345	0	1
	COUNTRY GROUP			
G-7	=1, if G7 countries	0.117	0	1
EU-15	=1, if EU-15 countries	0.064	0	1
EU	=1, if EU countries but not EU-15	0.031	0	1
*OECD**	=1, if OECD countries but not G7, EU-15, or EU	0.788	0	1
	ECONOMIC GROWTH MEASURE			
GDP	=1, if dependent variable is GDP growth	0.259	0	1
*PC-GDP**	=1, if dependent variable is per capita GDP growth	0.741	0	1
	TAX VARIABLE MEASURE			
Marginal	=1, if marginal tax rate (as opposed to average tax rate)	0.090	0	1
Differenced	=1, if change in tax rate (as opposed to level of tax rate)	0.172	0	1
ETR	=1, if effective tax rate (as opposed to statutory tax rate)	0.906	0	1
	DURATION OF TAX EFFECT			
*Short-run**	=1, if tax variable measures immediate/short-run effect	0.702	0	1
Medium-run	=1, if tax variable measures cumulative/medium-run effect	0.053	0	1
Long-run	=1, if tax variable measures long-run, steady-state effect	0.245	0	1

Variable	*Description*	*Mean*	*Min*	*Max*
	STUDY TYPE			
Peer-reviewed	=1, if study published in peer-reviewed journal	0.661	0.48	0.75
Publication Year	Year in which the last version of study was "published."	2007	1993	2015
	DATA TYPE			
Cross-section	=1, if data are cross-sectional.	0.009	0	1
*Panel**	=1, if data are panel	0.991	0	1
Length	Length of sample time period	31.4	5	40
Mid-Year	Midpoint of the sample time period	1985	1970.5	2004.5
	ESTIMATION TYPE			
*OLS**	=1, if OLS estimator is used.	0.677	0	1
GLS	=1, if Generalized Least Squares estimator is used.	0.154	0	1
TSLS/GMM	=1, if estimator corrects for endogeneity, e.g. 2SLS, 3SLS, or GMM.	0.168	0	1
	STANDARD ERROR TYPE			
*SE-OLS**	=1, if OLS standard error is considered.	0.587	0	1
SE-HET	=1, if heteroskedasticity standard error is considered.	0.245	0	1
SE-Other	=1, if both heteroskedasticity and autocorrelation standard error are considered.	0.168	0	1
	INCLUDED VARIABLES			
Initial income	=1, if initial level of income included	0.559	0	1
Lagged DV	=1, if lagged dependent variable included	0.167	0	1
CountryFE	=1, if the country fixed effects are included	0.833	0	1
Investment	=1, if investment included	0.585	0	1
Trade Openness	=1, if trade openness included	0.170	0	1
Human Capital	=1, if human capital included	0.440	0	1
Population Growth	=1, if population growth included	0.243	0	1
Employment Growth	=1, if employment growth included	0.378	0	1
Unemployment	=1, if unemployment rate included	0.090	0	1
Inflation	=1, if inflation rate included	0.131	0	1

Note: The grouped variables include all possible categories, where the categories omitted in the subsequent analysis are indicated by an asterisk, where applicable.

In my investigation of tax effects, I adopt the following empirical procedure. First I separate out the two sets of tax variables: *Prediction-Negative* and *Prediction-Positive;* and *Labour-Tax, Capital-Tax, Other-Tax, Mixed-Tax,* and *Overall-Tax*. I do this because the two sets of tax variables are significantly correlated. For example, Labour and Capital taxes are significantly associated with tax policies that are predicted to have negative effects. I then combine the two sets of tax variables to check for robustness.

For each set of regressions, I also include two sets of control variables. The top panel of each of the following tables reports the regression results when all control variables are included in the equation. The bottom panel reports the results when a stepwise procedure is used to select control variables, even while the tax variables are fixed to remain in each equation.[17] Since the tax variables are locked into each regression, the use of the stepwise procedure does not invalidate testing for their significance. All regressions also include the publication bias variable, *SE*, and thus control for publication bias.

The results of this analysis are given in Table 2.8 through Table 2.10. Table 2.8 reports the results when the prediction variables (*Prediction-Negative* and *Prediction-Positive*) are included in the meta-regression, while holding out the tax type variables. Across all four estimation procedures, and for both sets of control variables, I estimate a negative and statistically significant coefficient for the variable *Prediction-Negative*, and a positive and statistically significant coefficient for *Prediction-Positive.* These results are consistent with the predictions of growth theory.

[17] I use a backwards stepwise regression procedure that selects variables so as to minimize the Schwarz Information Criterion. I employed the user-written, Stata program *vselect* to implement the stepwise procedure.

Table 2.8: Meta-Regression Analysis (Omitting Tax Type Variables)

Variable	***Fixed Effects (Weight1) (1)***	***Fixed Effects (Weight2) (2)***	***Random Effects (Weight1) (3)***	***Random Effects (Weight2) (4)***
All Control Variables Included				
SE	-1.150*** *(-4.38)*	-1.172*** *(-5.25)*	-0.581*** *(-3.55)*	-0.508** *(-2.37)*
Prediction-Negative	-0.046*** *(-2.70)*	-0.037** *(-2.42)*	-0.096** *(-2.57)*	-0.115*** *(-3.06)*
Prediction-Positive	0.039*** *(4.38)*	0.041*** *(5.83)*	0.073** *(2.68)*	0.066** *(2.30)*
Control Variables Selected Via Backwards Stepwise Regression				
SE	-1.090*** *(-4.21)*	-1.144*** *(-4.74)*	-0.543*** *(-4.10)*	-0.430*** *(-3.31)*
Prediction-Negative	-0.044*** *(-3.75)*	-0.042*** *(-4.31)*	-0.102** *(-2.58)*	-0.113*** *(-5.69)*
Prediction-Positive	0.039*** *(4.41)*	0.042*** *(5.99)*	0.071*** *(2.80)*	0.081*** *(4.95)*

Note: The top panel reports the results of estimating Equation (2.10) with the addition of the two tax variables, *Prediction-Negative* and *Prediction-Positive.* The bottom panel adds control variables selected through a backwards stepwise regression procedure that selects variables so as to minimize the Schwarz Information Criterion (see Footnote #12). The top value in each cell is the coefficient estimate, and the bottom value in parentheses is the associated *t*-statistic. The four WLS estimators (*Fixed Effects-Weight1, Fixed Effects-Weight2, Random Effects-Weight1,* and *Random Effects-Weight2*) are described in the "FAT/PET tests" subsection of Section 2.4 in the text. All four estimation procedures calculate cluster robust standard errors. *, **, and *** indicate statistical significance at the 10-, 5-, and 1-percent level, respectively.

Table 2.9: Meta-Regression Analysis (Omitting Prediction Variables)

Variable	***Fixed Effects (Weight1) (1)***	***Fixed Effects (Weight2) (2)***	***Random Effects (Weight1) (3)***	***Random Effects (Weight2) (4)***
All Control Variables Included				
SE	-1.108*** *(-4.18)*	-1.144*** *(-5.14)*	-0.725*** *(-3.96)*	-0.612** *(-2.64)*
Labour-Tax	-0.037*** *(-3.38)*	-0.027*** *(-3.13)*	-0.064*** *(-2.73)*	-0.047** *(-2.03)*
Capital-Tax	-0.021** *(-2.44)*	-0.017** *(-2.23)*	-0.009 *(-0.49)*	-0.005 *(-0.19)*
Other-Tax	0.345** *(2.60)*	0.356*** *(2.82)*	0.151 *(1.36)*	0.109 *(0.81)*
Mixed-Tax	-0.049*** *(-6.77)*	-0.045*** *(-8.47)*	-0.099*** *(-3.49)*	-0.070* *(-1.92)*
Overall-Tax	-0.034 *(-1.63)*	-0.039** *(-2.36)*	-0.005 *(-1.05)*	-0.003 *(-0.88)*
Control Variables Selected Via Backwards Stepwise Regression				
SE	-1.147*** *(-4.19)*	-1.219*** *(-4.80)*	-0.651*** *(-4.95)*	-0.528*** *(-3.45)*
Labour-Tax	-0.040*** *(-5.28)*	-0.028*** *(-3.45)*	-0.057** *(-2.32)*	-0.038* *(-1.74)*
Capital-Tax	-0.023*** *(-2.86)*	-0.018** *(-2.58)*	-0.005 *(-0.24)*	-0.001 *(-0.07)*
Other-Tax	0.414** *(2.43)*	0.434** *(2.64)*	0.135 *(1.23)*	0.126 *(0.87)*
Mixed-Tax	-0.051*** *(-6.91)*	-0.046*** *(-8.86)*	-0.085*** *(-2.81)*	-0.052*** *(-3.09)*
Overall-Tax	-0.046*** *(-3.88)*	-0.051*** *(-4.07)*	-0.002 *(-0.53)*	0.000 *(0.18)*

Note: The top panel reports the results of estimating Equation (2.10) with the addition of the five tax variables, *Labour*, *Capital*, *Other*, *Mixed*, and *Overall* taxes. The bottom panel adds control variables selected through a backwards stepwise regression procedure that selects variables so as to minimize the Schwarz Information Criterion (see Footnote #12). The top value in each cell is the coefficient estimate, and the bottom value in parentheses is the associated *t*-statistic. The four WLS estimators (*Fixed Effects-Weight1, Fixed Effects-Weight2, Random Effects-Weight1,* and *Random Effects-Weight2*) are described in the "FAT/PET tests" subsection of Section 2.4 in the text. All four estimation procedures calculate cluster robust standard errors. *, **, and *** indicate statistical significance at the 10-, 5-, and 1-percent level, respectively.

Table 2.10: Meta-Regression Analysis (All Tax Variables Included)

Variable	***Fixed Effects (Weight1) (1)***	***Fixed Effects (Weight2) (2)***	***Random Effects (Weight1) (3)***	***Random Effects (Weight2) (4)***
All Control Variables Included				
SE	-0.963*** *(-4.02)*	-1.024*** *(-4.70)*	-0.647*** *(-4.22)*	-0.525** *(-2.48)*
Prediction-Negative	-0.045** *(-2.44)*	-0.038** *(-2.31)*	-0.085** *(-2.25)*	-0.108*** *(-2.91)*
Prediction-Positive	-0.001 *(-0.11)*	0.005 *(0.43)*	0.062** *(2.07)*	0.060** *(2.03)*
Labour-Tax	-0.031** *(-2.42)*	-0.020 *(-1.48)*	-0.023 *(-0.82)*	-0.011 *(-0.43)*
Capital-Tax	-0.015 *(-0.97)*	-0.009 *(-0.62)*	0.026 *(1.16)*	0.027 *(1.07)*
Other-Tax	0.285** *(2.21)*	0.313** *(2.48)*	0.154 *(1.39)*	0.111 *(0.87)*
Mixed-Tax	-0.045*** *(-3.13)*	-0.038*** *(-2.82)*	-0.062* *(-2.20)*	-0.035 *(-0.99)*
Overall-Tax	-0.031 *(-1.21)*	-0.031 *(-1.52)*	-0.000 *(-0.02)*	0.001 *(0.18)*
Control Variables Selected Via Backwards Stepwise Regression				
SE	-0.925*** *(-3.89)*	-0.997*** *(-4.28)*	-0.623*** *(-5.03)*	-0.402*** *(-2.94)*
Prediction-Negative	-0.039*** *(-6.56)*	-0.040*** *(-3.29)*	-0.089** *(-2.60)*	-0.112*** *(-5.61)*
Prediction-Positive	-0.012 *(-1.29)*	0.007 *(0.53)*	0.063** *(2.08)*	0.070*** *(3.55)*
Labour-Tax	-0.041*** *(-4.57)*	-0.021 *(-1.47)*	-0.023 *(-0.78)*	0.010 *(0.48)*
Capital-Tax	-0.022** *(-2.48)*	-0.008 *(-0.59)*	0.021 *(0.89)*	0.046*** *(3.02)*
Other-Tax	0.316* *(2.00)*	0.368** *(2.38)*	0.145 *(1.33)*	0.075 *(0.60)*
Mixed-Tax	-0.055*** *(-6.26)*	-0.037*** *(-2.97)*	-0.050* *(-1.79)*	-0.017 *(-0.95)*
Overall-Tax	-0.048*** *(-4.03)*	-0.026** *(-2.07)*	0.001 *(0.33)*	0.005*** *(4.42)*

The results are only slightly less supportive of growth theory when the tax type variables are added to the specification. Table 2.10 reports the corresponding estimates. The coefficient for *Prediction-Negative* remains negative and statistically significant across all four estimation procedures. *Prediction-Positive* is positive and statistically significant in the two random effects regressions (Columns 3 and 4), but insignificant in the two fixed effects regressions (Columns 1 and 2). As noted above, I consider the random effects estimator to be more reliable, so that the results from Table 2.10 are generally consistent with those from Table 2.8.

Not only do these findings constitute general statistical support in favour of the predictions of growth theory, but the respective coefficients indicate that tax policy can have a substantial economic impact. For example, the difference between the coefficients for *Prediction-Negative* and *Prediction-Positive* range from a minimum of 0.027 (Table 2.10, Bottom panel, Column 1) to a maximum of 0.194 (Table 2.8, Bottom panel, Column 4), with a midpoint value of approximately 0.11.

Let me now consider the following thought experiment. Suppose fiscal policy underwent the following policy switch: distortionary taxes and unproductive expenditures were reduced by 10 percentage points while, simultaneously, non-distortionary taxes and productive expenditures were increased by the same amount. Using a point estimate of 0.11, my meta-regression results indicate that this would increase annual growth of GDP by 1.1 percentage points. As noted above, the average annual growth rate for OECD countries over the sample range of the studies included in this meta-regression analysis was approximately 2.5 percent. Thus a 1.1 percentage point increase in annual growth would constitute a substantial increase. Admittedly, this thought experiment is an extreme case, both in the absolute size of the tax changes and in the swing in fiscal policy from one extreme of the growth pole to the other. Nevertheless, it does indicate that there is a role for tax-based fiscal policy to increase economic growth amongst OECD countries.

The last tax issue addressed in this study investigates whether some types of taxes are more growth-retarding than others. As noted in Table 2.2, Labour and Capital taxes are commonly classified as distortionary, while Consumption taxes are classified as non-distortionary.

Table 2.9 estimates a meta-regression with the tax type variables but with prediction variables omitted, while Table 2.10 includes both. As the omitted category is Consumption taxes, I expect the coefficient on Labour and Capital taxes to be negative, whereas there is no sign expectation for the other tax type coefficients.

With respect to Labour taxes, the results from Table 2.9 across all four estimation procedures and with both sets of control variables show negative and statistically significant coefficients. However, when prediction variables are added to the regression (cf. Table 2.10), the coefficient on *Labour-Tax* becomes insignificant in the preferred random effects regressions. In terms of economic significance, the estimates range from -0.064 (Table 2.9, Top panel, Column 3) to 0.010 (Table 2.10, Bottom panel, Column 4). The more negative estimates indicate that raising revenues from Labour taxes rather than Consumption taxes can have important growth consequences. However, given that some of the preferred *Random Effects* estimates are statistically insignificant, my overall assessment is that these estimates constitute weak evidence that Labour taxes are more growth-retarding than Consumption taxes.

The evidence that Capital taxes are more distortionary than Consumption taxes is even weaker. While the coefficients on the *Capital-Tax* variable are negative in all Table 2.9 regressions, they are insignificant in the preferred *Random Effects* estimations. When the prediction variables are added, the respective coefficients are generally insignificant (cf. Table 2.10). One of the regressions even produces a significant positive coefficient (bottom panel, *Random Effects-Weight2*). As a result, I conclude that the evidence that Capital taxes are more distortionary than Consumption taxes is mixed.

Bayesian model averaging of control variables. In order to address one of the main objectives of this study I now turn to an analysis of the control variables. The problem is that other than the two sets of tax variables, there are 28 control variables and it is not clear which ones should be included. In other words, multicollinearity may be an issue with the inclusion of so many variables. For example, when all 28 variables are included with both sets of tax variables and the meta-regression is estimated using the "Random Effects (Weight2)" estimator, as in Column (4) of the top panel of Table 2.10, only 5 of the 28 control variables are statistically significant at the 5 percent level. In contrast, when a general-to-specific (G-to-S) approach is used -- in this case, backwards selection -- only 9 of the 28 control variables are significant (cf. bottom panel of Table 2.10). Further, one of the variables that is significant in the top panel is not significant in the bottom panel's specification. Thus, variable selection matters when trying to determine the effect of various control variables on estimated tax effects.

To tackle the problem of specification uncertainty, I use a technique called Bayesian Model Averaging, or BMA (Zeugner, 2011). BMA is not specifically designed for meta-regression studies. But because model uncertainty is an issue in these studies, it is an appropropriate method to apply. BMA runs a vast number of regressions with different subsets of the explanatory variables, and then constructs a weighted average over the set of estimated coefficients.

Table 2.11 reports the results of an analysis where I lock in the tax variables *Prediction-Negative* and *Prediction-Positive* and then apply BMA to the 28 control variables. All specifications adjust for publication bias. The results differ with respect to the estimation procedure used. However, they are more consistent across analyses than would be the case, say, if I reported the results from specifications that included all variables and those that employed stepwise regression. I report results for both the "Fixed Effects (Weight 1)" and "Random Effects (Weight 2)" estimators. These two estimators use very different weighting schemes. Previous tables indicated that the estimates from these two estimators sometimes vary substantially. As a result, they provide an indication of robustness across estimation procedures.

I report three summary measures. For each variable I compute a Posterior Inclusion Probability (*PIP*), which is the sum of posterior model probabilities of the regressions in which the variable is included. It can capture how well the model is designed and may be compared to the adjusted R^2, or to information criteria. With 28 control variables, there are 2^{28} potential regressions with various variable specifications. Variables that appear in specifications with high likelihood values will have larger *PIP* values. By construction, every variable appears in 50 percent of all possible specifications. However, the *PIP* can be very close to 100 percent if the specifications that include a variable have much greater likelihood values than those in which it is omitted.

The Posterior Mean (*Post. Mean*) uses the above-mentioned probability values to weight the estimated coefficients from each specification. Specifications in which a variable is not included assign an "estimated value" of zero to construct the Posterior Mean. Lastly, BMA also calculates the probability that a given coefficient has a positive sign (*Cond. Pos. Sign*). This is constructed in the same manner as the Posterior Mean, except that it uses a dummy variable indicating positive values rather than the estimated coefficient in constructing a weighted average.

Table 2.11 uses yellow to highlight all the control variables that: (i) have a *PIP* greater than 50%; (ii) have a Conditional Positive Sign of either 1.00 or 0.00 – indicating that the respective coefficient is consistently estimated to be either positive or negative in the most likely specifications; and (iii) have the same Conditional Positive Sign value for both the *Fixed Effects(Weight1)* and *Random Effects(Weight2)* estimators.

Studies that estimate tax effects for G-7 and EU-15 countries produce consistently less negative/more positive estimates than studies that include a large sample of countries from the OECD. To place the size of the Posterior Mean values in context, it helpful to recall that the median estimated tax effect from Table 2.5 is -0.073. By this standard, the effect of belonging to a G-7 country is relatively large (0.184 in the FE model and 0.181 in the RE model). The effect associated with being a EU-15 member, while still positive, is substantially smaller.

I find that studies that measure economic growth using total GDP (*GDP*) rather than per capita GDP, and that employ a marginal (as opposed to average) measure of tax rates (*Marginal*), generally produce tax effects that are less negative/more positive. Compared to the short-run effects of taxes, studies that estimate medium-run tax effects (*Medium-run*) produce estimates of tax effects that are less negative/more positive; while studies that estimate long-run, steady-state tax effects (*Long-run*) produce estimates that are more negative/less positive. There is evidence to indicate that more recent studies (*Publication Year*) produce less negative/more positive estimates as do cross-sectional studies (*Cross-section*) compared to panel studies. However, there is also evidence that studies using more recent data (*Mid-Year*) find more negative/less positive tax effects.

With respect to estimation procedures, studies that use GLS rather than OLS (*GLS*) generally produce more negative/less positive estimates of tax effects. Interestingly, correcting for endogeneity (*TSLS/GMM*) does not appear to have much impact. Meta-regressions using the *Fixed Effects(Weight1)* estimator find that studies that employ TSLS/GMM generally estimate more negative/less positive effects. Meta-regressions using the *Random EffectsWeight2)* estimator find the opposite. However, in both cases the Posterior Mean values are negligibly small (-0.001 and 0.009), suggesting either that tax policy is not endogenous or that the instruments that have been employed in previous studies are not effective in correcting endogeneity. There is evidence that it makes a difference as to how standard errors are calculated, with studies that incorporate serial correlation, cross-sectional correlation and the like in calculating standard errors (*SE-Other*) associated with less negative/more positive effects.

Lastly, I find that studies that include initial income, employment growth, and unemployment rates in the growth equations are likely to produce less negative/more positive estimates; with studies that include country fixed effects, population growth, and inflation producing more negative/less positive tax effects. While the above findings are robust across variable specifications and the two estimation procedures, I again emphasize that the sizes of the associated effects are small.

Table 2.11: Bayesian Model Averaging Analysis (Control Variables)

Variable	***Fixed Effects(Weight1)***			***Random Effects(Weight2)***		
	PIP	***Post. Mean***	***Cond. Pos. Sign***	***PIP***	***Post. Mean***	***Cond. Pos. Sign***
G-7	**1.00**	**0.184**	**1.00**	**1.00**	**0.181**	**1.00**
EU-15	**0.97**	**0.032**	**1.00**	**0.99**	**0.066**	**1.00**
EU	0.81	0.064	1.00	0.59	0.000	0.56
GDP	**0.99**	**0.025**	**1.00**	**1.00**	**0.065**	**1.00**
Marginal	**0.80**	**0.006**	**1.00**	**0.76**	**0.023**	**1.00**
Differenced	0.84	-0.018	0.01	1.00	-0.096	0.00
ETR	1.00	0.027	1.00	1.00	-0.091	0.00
Medium-run	**1.00**	**0.081**	**1.00**	**0.98**	**0.052**	**1.00**
Long-run	**0.99**	**-0.015**	**0.00**	**1.00**	**-0.079**	**0.00**
Peer-reviewed	1.00	0.056	1.00	0.63	-0.004	0.00
Publication Year	**0.98**	**0.004**	**1.00**	**1.00**	**0.009**	**1.00**
Cross-section	**0.76**	**0.009**	**1.00**	**0.73**	**0.015**	**1.00**
Length	0.94	-0.002	0.00	0.61	0.000	0.08
Mid-Year	**0.93**	**-0.003**	**0.00**	**1.00**	**-0.006**	**0.00**

Variable	*Fixed Effects(Weight1)*			*Random Effects(Weight2)*		
	PIP	*Post. Mean*	*Cond. Pos. Sign*	*PIP*	*Post. Mean*	*Cond. Pos. Sign*
GLS	**1.00**	**-0.043**	**0.00**	**0.84**	**-0.021**	**0.00**
TSLS/GMM	**0.73**	**-0.001**	**0.00**	**0.78**	**0.009**	**1.00**
SE-HET	0.70	-0.001	0.15	0.73	0.009	1.00
SE-Other	**1.00**	**0.051**	**1.00**	**0.69**	**0.013**	**1.00**
Initial income	**0.92**	**0.013**	**1.00**	**0.99**	**0.048**	**1.00**
Lagged DV	0.89	-0.027	0.00	0.71	0.016	1.00
Country FE	**1.00**	**-0.047**	**0.00**	**1.00**	**-0.062**	**0.00**
Investment	0.82	0.004	1.00	0.78	-0.011	0.00
Trade Openness	0.73	0.003	1.00	0.67	-0.006	0.00
Human Capital	0.84	0.007	1.00	0.87	-0.013	0.00
Population Growth	**1.00**	**-0.050**	**0.00**	**1.00**	**-0.074**	**0.00**
Employment Growth	**0.98**	**0.028**	**1.00**	**0.87**	**0.019**	**1.00**
Unemployment	**1.00**	**0.066**	**1.00**	**1.00**	**0.046**	**1.00**
Inflation	**0.75**	**-0.006**	**0.00**	**0.76**	**-0.009**	**0.00**

Note: The column headings *PIP*, *Post. Mean*, and *Cond. Pos. Sign* stand for Posterior Inclusion Probability, Posterior Mean, and the likelihood-weighted probability that the respective coefficient takes a positive sign. These are described in the "Bayesian model averaging of control variables" subsection of Section 2.4 in the text. The Bayesian Model Averaging (BMA) analysis was done using the R package BMS, described in Zeugner (2011). The WLS estimators *Fixed Effects-Weight1* and *Random Effects-Weight2* are described in the "FAT/PET tests" subsection of Section IV. All specifications included the tax variables *Prediction_Negative* and *Prediction_Positive,* which were forced into all model specifications, and adjusted for publication bias. The table yellow-highlights all the control variables that (i) have a *PIP* greater than 50%; (ii) have a *Conditional Positive Sign* of either 1.00 or 0.00 – indicating that the respective coefficient is consistently estimated to be either positive or negative in the most likely specifications; and (iii) have the same *Conditional Positive Sign* value for both the *Fixed Effects(Weight1)* and *Random Effects(Weight2)* estimators.

Figure 2.6 provides a visual representation of the BMA analysis for the tax (*Prediction-Negative* and *Prediction-Positive*) and control variables using the *Fixed Effects(Weight1)* estimator.[18] The figure reports estimates from the top 1000 models, with most likely models ordered from left to right. These 1000 models, out of 10^{28} possible models, account for a cumulative probability of approximately 30 percent. Red (blue) squares indicate that the respective coefficient is negative (positive) in the given model. A white square indicates that the variable is omitted from that model. A solid band of the same colour across the figure indicates that the respective variable is consistently estimated to have the same sign across all 1000 models. In addition to confirming the results from Table 2.11 the figure also indicates the variable specifications of the top models. These closely match the *PIP* values in Table 2.11. The corresponding figure for the *Random Effects(Weight2)* estimator is quite similar and is reproduced in Appendix 2.7.

[18] Note that in the associated specifications, the variable *Precision* corresponds to the constant term, while the constant term corresponds to the publication bias correction factor which is (1/*SE)*.

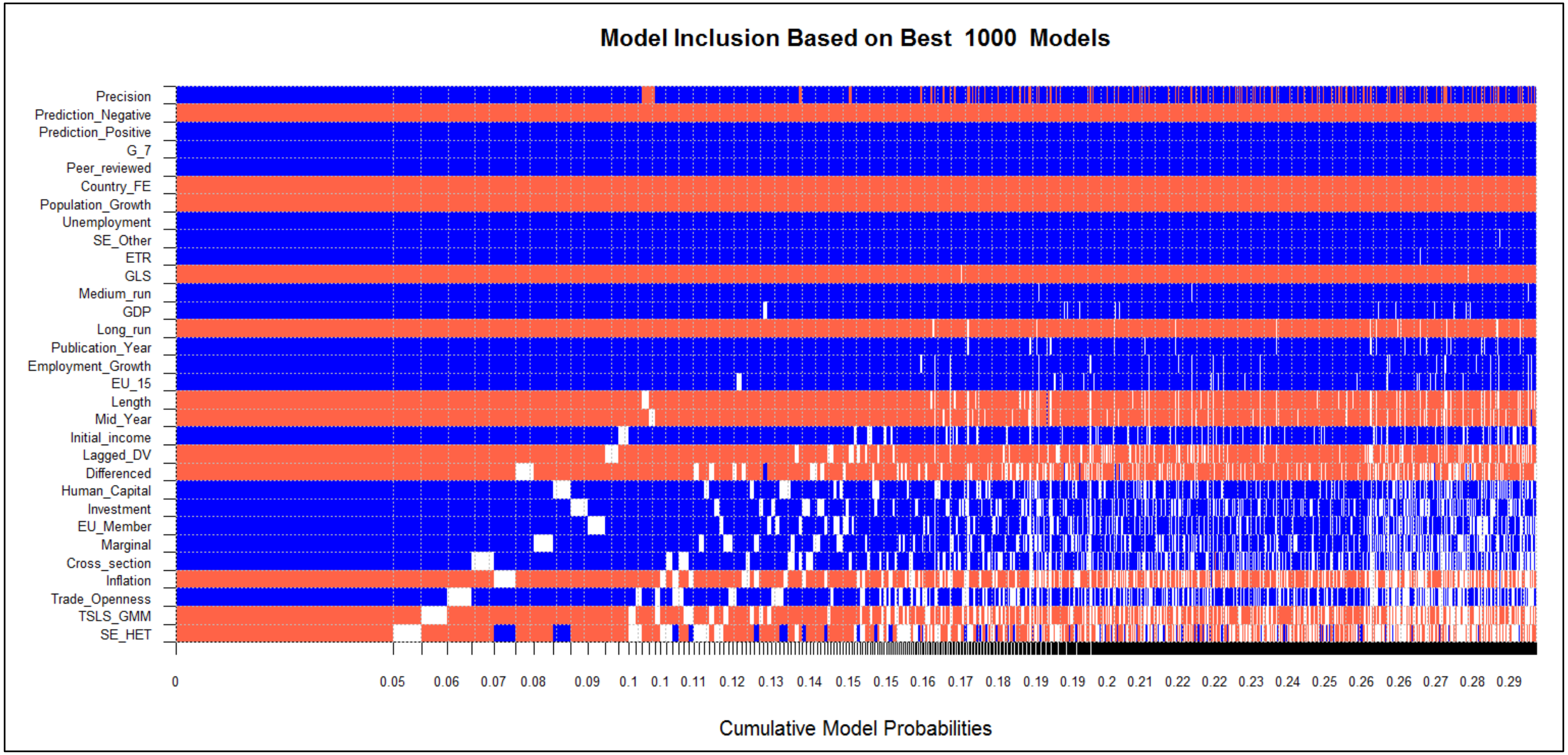


Figure 2.6: Visual Representation of BMA Analysis (Fixed Effects-Weight1)

Note: Each column represents a single model. Variables are listed in descending order of posterior inclusion probability (PIP) and have all been weighted according to the *Fixed Effects – Weight 1* case. Blue (dark) indicates that the variable is included in that model and estimated to be positive. Red (light) indicates the variable is included and estimated to be negative. No colour (white) indicates the variable is not included in that model. Further detail about this plot is given in Zeugner (2011).

2.5. Conclusion

The effect of taxation on economic growth has been an enduring question. Despite the large body of research devoted to taxes and economic growth in OECD countries, the general picture that emerged from the empirical evidence is inconclusive. One reason for the seemingly contradictory findings is that estimates of tax effects are often estimating different things. Because of the government budget constraint, the same tax effect can be estimated to be positive or negative, depending on the other fiscal categories omitted from the specification. For this and other reasons, it is valuable to collect the estimates from this literature and carefully track the differences across studies so that the estimates can be combined to provide an overall assessment of the growth effects of taxes.

This study combines results of 713 estimates from 42 studies, all of which attempt to estimate the effect of taxes on economic growth in OECD countries. I drop outlier estimates from both top and bottom of the sample range, and apply meta-regression analysis to analyse a final sample of 641 estimates. First, there is statistical evidence to support that estimates in the literature suffer from negative publication bias. Second, by accommodating and correcting for publication bias, the overall effect of taxes on economic growth is negligibly small and statistically insignificant. However, this overall effect is not particularly meaningful because it lumps together different tax policies.

Third, to provide a clear picture of the scope of tax policy to effect economic growth, I categorize tax policies by their predicted effects on economic growth according to the findings in public finance. Once I control for publication bias, increases in unproductive expenditures funded by distortionary taxes and/or deficits have a statistically significant, negative effect on economic growth. On the contrary, increases in non-distortionary taxes to fund productive expenditures and/or government surpluses have a statistically significant, positive effect on economic growth. The difference between these “best” and “worst” tax policies can be economically important. For example, using a midpoint estimate from my meta-regression analysis, I calculate that if distortionary taxes and unproductive expenditures were reduced by 10 percentage points while, simultaneously, non-distortionary taxes and productive expenditures were increased by the same amount, the net effect would be an increase of 1.1 percentage points in annual GDP growth. While this represents an extreme case, both in the absolute size of the tax changes, and in the swing in fiscal policy from one extreme of the growth pole to the other, it does indicate that there is scope for tax-based fiscal policy to increase economic growth.

Fourth, with respect to particular types of taxes, I find weak evidence that taxes on labour are more growth retarding than other types of taxes. Evidence regarding other types of taxes is mixed. Finally, I find evidence that data and study characteristics account for much systematic variation in tax estimates across studies, though the effects from any one characteristic is generally small. The one exception is that studies that focus their analysis on G-7 countries find less negative/more positive tax effects than those that use a wider sample of OECD countries.

One of the great advantages of meta-regression analysis compared to the original studies and also narrative reviews is that it can avoid some of the problems associated with publication bias and selective reporting of results. Further, it can control for differences across studies that might otherwise mask significant effects. It can also add new information relevant to the literature (Stanley and Doucouliagos, 2012). This is particularly of interest when estimating the effects of tax policy. The results of this study indicate that once these factors are taken into account, the combined weight of the evidence from the literature indicates that tax policy can have an economically important impact on economic growth.

2.6. References

Afonso, A., and Alegre, J. G. (2008). Economic growth and budgetary components: A panel assessment for the EU. Working Paper Series 848, European Central Bank, Frankfurt.

Afonso, A., and Alegre, J. G. (2011). Economic growth and budgetary components: A panel assessment for the EU. *Empirical Economics*, *41*(3), 703-723.

Afonso, A., and Furceri, D. (2010). Government size, composition, volatility and economic growth. *European Journal of Political Economy*, *26*(4), 517-532.

Afonso, A., and Jalles, J. T. (2014). Fiscal composition and long-term growth. *Applied Economics*, *46*(3), 349-358.

Afonso, A., and Jalles, J. T. (2013). Fiscal composition and long-term growth.Working Paper Series 1518, European Central Bank, Frankfurt.

Agell, J., Lindh, T., and Ohlsson, H. (1997). Growth and the public sector: A critical review essay. *European Journal of Political Economy*, *13*(1), 33-52.

Agell, J., Lindh, T., and Ohlsson, H. (1999). Growth and the public sector: A reply. *European Journal of Political Economy*, *15*(2), 359-366.

Agell, J., Ohlsson, H., and Thoursie, P. S. (2006). Growth effects of government expenditure and taxation in rich countries: A comment. *European Economic Review*, *50*(1), 211-218.

Alesina, A., and Ardagna, S. (2010). Large changes in fiscal policy: Taxes versus spending. *Tax Policy and the Economy*, *24*, 35-68.

Angelopoulos, K., Economides, G., and Kammas, P. (2007). Tax-spending policies and economic growth: Theoretical predictions and evidence from the OECD. *European Journal of Political Economy*, *23*(4), 885-902.

Arin, K. P. (2004). Fiscal Policy, private investment and economic growth: Evidence from G-7 Countries. Unpublished manuscript.

Arnold, J. M. (2008). Do tax structures affect aggregate economic growth? Empirical evidence from a panel of OECD countries. Working Paper Series 643, Organization for Economic Co-operation and Development. Economic Department of the OECD.

Arnold, J. M., Brys, B., Heady, C., Johansson, Å., Schwellnus, C., and Vartia, L. (2011). Tax policy for economic recovery and growth. *The Economic Journal*, *121*(550), F59-F80.

Bania, N., Gray, J.A., and Stone, J.A. (2007). Growth, taxes, and government expenditures: Growth hills for U.S. states. *National Tax Journal*, Vol. LX(2), 193-204.

Barro R. (1990) Government spending in a simple model of endogenous growth, *Journal of Political Economy*, 98, s103-117.

Barro, R. and Sala-i-Martin, X. (1992) Public finance in models of economic growth, *Review of Economic Studies*, 59, 645-61.

Baskaran, T., and Feld, L. P. (2013). Fiscal decentralization and economic growth in OECD countries: Is there a relationship? *Public Finance Review*, *41*(4), 421-445.

Bassanini, A., Scarpetta, S., and Hemmings, P. (2001). Economic growth: the role of policies and institutions. Panel data evidence from OECD countries.

Benos, N. (2009). Fiscal policy and economic growth: Empirical evidence from EU countries. Working Paper Series 107. Center for Planning and Economic Research.

Bergh, A., and Karlsson, M. (2010). Government size and growth: Accounting for economic freedom and globalization. *Public Choice*, *142*(1-2), 195-213.

Bergh, A., and Öhrn, N. (2011). Growth effects of fiscal policies: A critical appraisal of Colombier's (2009) study. Working Paper Series 865, Research Institute of Industrial Economics (IFN).

Bleaney, M., Gemmell, N., and Kneller, R. (2001). Testing the endogenous growth model: Public expenditure, taxation, and growth over the long run. *Canadian Journal of Economics/Revue Canadienne D'économique*, *34*(1), 36-57.

Borenstein, M., Hedges, L. V., Higgins, J., and Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods*, *1*(2), 97-111.

Colombier, C. (2009). Growth effects of fiscal policies: An application of robust modified M-estimator. *Applied Economics*, *41*(7), 899-912.

Dackehag, M., and Hansson, Å. (2012). Taxation of income and economic growth: An empirical analysis of 25 rich OECD countries. Working Paper Series 2012:6, Lund University.

Daveri, F., and Tabellini, G. (2000). Unemployment, growth and taxation in industrial countries. *Economic Policy*, *15*(30), 47-104.

De la Fuente, A. (1997). Fiscal policy and growth in the OECD. Discussion Paper Series 1755, Center for Economic Policy Research (CEPR).

Devarajan S., Swaroop, V., and Zou, H. (1996) The composition of public expenditure and economic growth, *Journal of Monetary Economics*, 37, 313-344.

Fölster, S., and Henrekson, M. (1999). Growth and the public sector: A critique of the critics. *European Journal of Political Economy, 15*(2), 337-358.

Fölster, S., and Henrekson, M. (2001). Growth effects of government expenditure and taxation in rich countries. *European Economic Review*, *45*(8), 1501-1520.

Futagami, K.Y, Morita, Y., and Shibata, A. (1993) Dynamic analysis of an endogenous growth model with public capital, *Scandinavian Journal of Economics*, 95, 606-625.

Furceri, D., and Karras, G. (2009). Tax changes and economic growth: Empirical evidence for a panel of OECD countries. Unpublished manuscript.

Gemmell, N., Kneller, R., and Sanz, I. (2008). Fiscal policy impacts on growth in the OECD: Are they long-run? Mimeo, University of Nottingham.

Gemmell, N., Kneller, R., and Sanz, I, (2009). The composition of government expenditure and economic growth: some new evidence for OECD countries. In European Commission, *The quality of public finances and economic growth: Proceedings to the annual workshop on public finances (Brussels, 28 November 2008*), European Economy, Occasional Papers No. 45. Directorate-General for Economic and Financial Affairs, Publications, Brussels, Belgium, pages 17-46.

Gemmell, N., Kneller, R., and Sanz, I. (2011). The timing and persistence of fiscal policy impacts on growth: Evidence from OECD countries. *The Economic Journal*, *121*(550), F33-F58.

Gemmell, N., Kneller, R., and Sanz, I. (2014). The growth effects of tax rates in the OECD. *Canadian Journal of Economics/Revue Canadienne d'Economique*, *47*(4), 1217-1255.

Gemmell, N., Kneller, R., and Sanz, I. (2015). Does the Composition of Government Expenditure Matter for Long-Run GDP Levels? *Oxford Bulletin of Economics and Statistics*. DOI: 10.1111/obes.12121

Hakim, T. A., Bujang, I., and Ahmad, I. (2013) Tax structure and economic indicators in the modern era: Developing vs. high-income OECD countries. Conference paper.

Hansson, Å. (2010). Is the wealth tax harmful to economic growth? *World Tax Journal*, *2*(1), 19-34.

Havranek, T., and Irsova, Z. (2012). Survey article: Publication bias in the literature on foreign direct investment spillovers. *The Journal of Development Studies*, *48*(10), 1375-1396.

Heitger, B. (1993). Convergence, the “tax-state” and economic dynamics. *Weltwirtschaftliches Archiv*, *129*(2), 254-274.

Higgins, J.P. and Thompson, S.G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21*(11), 1539-58.

Karras, G. (1999). Taxes and growth: Testing the neoclassical and endogenous growth models. *Contemporary Economic Policy*, *17*(2), 177-188.

Karras, G., and Furceri, D. (2009). Taxes and growth in Europe. *South-Eastern Europe Journal of Economics*, *2*, 181-204.

Kneller, R., Bleaney, M. F., and Gemmell, N. (1999). Fiscal policy and growth: Evidence from OECD countries. *Journal of Public Economics*, *74*(2), 171-190.

Mendoza, E. G., Milesi-Ferretti, G. M., and Asea, P. (1997). On the ineffectiveness of tax policy in altering long-run growth: Harberger's superneutrality conjecture. *Journal of Public Economics*, *66*(1), 99-126.

Miller, S. M., and Russek, F. S. (1997). Fiscal structures and economic growth: International evidence. *Economic Inquiry*, *35*(3), 603-613.

Muinelo-Gallo, L., and Roca-Sagalés, O. (2013). Joint determinants of fiscal policy, income inequality and economic growth. *Economic Modelling*, *30*, 814-824.

Padovano, F., and Galli, E. (2001). Tax rates and economic growth in the OECD countries. *Economic Inquiry*, *39*(1), 44-57.

Paparas, D., Richter, C., and Paparas, A. (2015). Fiscal Policy and Economic Growth, Empirical Evidence in European Union. *Turkish Economic Review*, *2*(4), 239-268.

Reed, W.R. (2015). A Monte Carlo analysis of alternative meta-analysis estimators in the presence of publication bias. *Economics: The Open-Access, Open-Assessment E-Journal, 9*(2015-30), 1-40.

Romero-Avila, D., and Strauch, R. (2008). Public finances and long-term growth in Europe: Evidence from a panel data analysis. *European Journal of Political Economy*, *24*(1), 172-191.

Shemilt, I., Mugford, M., Vale, L., Marsh, K. and Donaldson, C., (eds.) (2011). *Evidence-based decisions and economics: Health care, social welfare, education, and criminal justice*, 2nd edition. New York: John Wiley & Sons.

Stanley, T.D., and Doucouliagos, H. (2012). *Meta-regression analysis in economics and business*. London: Routledge.

Stanley, T.D., Doucouliagos, H., Giles, M., Heckemeyer, J.H., Johnston, R.J., Laroche, P., Nelson, J.P., Paldam, M., Poot, J., Pugh, G. and Rosenberger, R.S. (2013). Meta-analysis of economics research reporting guidelines. *Journal of Economic Surveys*, *27*(2), 390-394.

Volkerink, B., Sturm, J. E., and De Haan, J. (2002). Tax ratios in macroeconomics: Do taxes really matter? *Empirica*, *29*(3), 209-224.

Widmalm, F. (2001). Tax structure and growth: Are some taxes better than others? *Public Choice*, *107*(3-4), 199-219.

Xing, J. (2011). Does tax structure affect economic growth? Empirical evidence from OECD countries. Working Paper Series 11/20. Center for Business Taxation, Said Business School, University of Oxford.

Xing, J. (2012). Tax structure and growth: How robust is the empirical evidence? *Economics Letters*, *117*(1), 379-382.

Zeugner, S. (2011). Bayesian model averaging with BMS. Available at http://bms.zeugner.eu/ [Accessed 15 Oct. 2015].

2.7. Appendices

Appendix 2.1: List of Terms Used in Electronic Search by Category

TAX	ECONOMIC GROWTH	OECD
Tax(es) /Tax rate(s)/Taxation	Economic growth	OECD countries
Tax policy(policies)	Growth	EU countries
Tax ratios	Economic indicators	G-7 countries
Tax changes	Long-term growth	High income OECD countries
Tax rate change	Long-run growth	Industrial countries
Fiscal policy(policies)		Rich countries
Tax structures/Fiscal structures		Europe
Fiscal decentralization		Cross-national study
Public finances		

Appendix 2.2: Letter to the Authors (OECD)

Dear Sir/Madam,

I am a Professor of economics at the University of Canterbury in New Zealand. We have a research team here undertaking a "meta-analysis" of the relationship between taxes and economic growth in the OECD countries.

A thorough meta-analysis involves collecting as many papers as possible on a subject, including unpublished research. The latter is known as "grey literature", and includes conference proceedings, reports from research firms or think thanks, theses and dissertations, etc. The unpublished literature is particularly important for addressing publication bias.

In this context, I am asking for your help.

Attached to this email is a listing of research on the topic of taxes and economic growth in the OECD countries. To be included, the research had to (i) include data from OECD countries (ii) have a dependent variable that was the growth of per capita personal income (PCPI) or GDP, and (iii) include one or more measures of taxes.

I am contacting you because you have researched in this area in the past.

Would you please look over this list and see if there are any notable omissions? I have broken the list down to the following categories: (i) journal articles, (ii) conference proceedings, (iii) studies from think tanks and research firms, (iv) theses/dissertations, and (v) working papers and unpublished research.

The last two categories are especially difficult to get information on. I would be greatly appreciative if you could identify any research we may have omitted.

Finally, if you are aware of any researchers who are currently researching in this area, it would be great if you could reply back with their names, and I will follow up with them directly.

I am sure you would agree that the subject of taxes and economic growth in OECD countries is very important. There is now a substantial enough literature that a careful meta-analysis can help to organize an empirical consensus of the existing literature.

Thank you so much for any help you can provide.

Sincerely,

Appendix 2.3: Bibliography (Attachment to the Above Letter)

Bibliography

(i) Journal Articles

Afonso, A., & Alegre, J. G. (2011). Economic growth and budgetary components: A panel assessment for the EU. *Empirical Economics*, *41*(3), 703-723.

Afonso, A., & Furceri, D. (2010). Government size, composition, volatility and economic growth. *European Journal of Political Economy*, *26*(4), 517-532.

Afonso, A., & Jalles, J. T. (2014). Fiscal composition and long-term growth.*Applied Economics*, *46*(3), 349-358.

Agell, J., Lindh, T., & Ohlsson, H. (1997). Growth and the public sector: A critical review essay. *European Journal of Political Economy*, *13*(1), 33-52.

Agell, J., Lindh, T., & Ohlsson, H. (1999). Growth and the public sector: A reply. *European Journal of political economy*, *15*(2), 359-366.

Agell, J., Ohlsson, H., & Thoursie, P. S. (2006). Growth effects of government expenditure and taxation in rich countries: A comment. *European Economic Review*, *50*(1), 211-218.

Alesina, A., & Ardagna, S. (2010). Large changes in fiscal policy: Taxes versus spending. *Tax Policy and the Economy*, *24*, 35-68.

Angelopoulos, K., Economides, G., & Kammas, P. (2007). Tax-spending policies and economic growth: Theoretical predictions and evidence from the OECD. *European Journal of Political Economy*, *23*(4), 885-902.

Arnold, J. M., Brys, B., Heady, C., Johansson, Å., Schwellnus, C., & Vartia, L. (2011). Tax policy for economic recovery and growth. *The Economic Journal*, *121*(550), F59-F80.

Baskaran, T., & Feld, L. P. (2013). Fiscal decentralization and economic growth in OECD countries: Is there a relationship?. *Public Finance Review*,*41*(4), 421-445.

Bergh, A., & Karlsson, M. (2010). Government size and growth: Accounting for economic freedom and globalization. *Public Choice*, *142*(1-2), 195-213.

Blankenau, W. F., Simpson, N. B., & Tomljanovich, M. (2007). Public education expenditures, taxation, and growth: Linking data to theory. *The American Economic Review*, 393-397.

Bleaney, M., Gemmell, N., & Kneller, R. (2001). Testing the endogenous growth model: Public expenditure, taxation, and growth over the long run.*Canadian Journal of Economics/Revue canadienne d'économique*, *34*(1), 36-57.

Bujang, I., Hakim, T. A., & Ahmad, I. (2013). Tax structure and economic indicators in developing and high-income OECD countries: Panel cointegration analysis. *Procedia Economics and Finance*, *7*, 164-173.

Colombier, C. (2009). Growth effects of fiscal policies: An application of robust modified M-estimator. *Applied Economics*, *41*(7), 899-912.

Daveri, F., & Tabellini, G. (2000). Unemployment, growth and taxation in industrial countries. *Economic policy*, *15*(30), 47-104.

Fölster, S., & Henrekson, M. (2001). Growth effects of government expenditure and taxation in rich countries. *European Economic Review*,*45*(8), 1501-1520.

Fölster, S., & Henrekson, M. (1999). Growth and the public sector: A critique of the critics. *European Journal of Political Economy*, 15(2), 337-358.

Gemmell, N., Kneller, R., & Sanz, I. (2015). Does the Composition of Government Expenditure Matter for Long-Run GDP Levels?. *Oxford Bulletin of Economics and Statistics*. DOI: 10.1111/obes.12121

Gemmell, N., Kneller, R., & Sanz, I. (2014). The growth effects of tax rates in the OECD. *Canadian Journal of Economics/Revue canadienne d'économique*, *47*(4), 1217-1255.

Gemmell, N., Kneller, R., & Sanz, I. (2011). The timing and persistence of fiscal policy impacts on growth: Evidence from OECD countries. *The Economic Journal*, *121*(550), F33-F58.

Hansson, Å. (2010). Is the wealth tax harmful to economic growth?. *World Tax Journal*, *2*(1), 19-34.

Heitger, B. (1993). Convergence, the “tax-state” and economic dynamics.*Weltwirtschaftliches Archiv*, *129*(2), 254-274.

Karras, G., & Furceri, D. (2009). Taxes and growth in Europe. *South-Eastern Europe Journal of Economics*, *2*, 181-204.

Karras, G. (1999). Taxes and growth: Testing the neoclassical and endogenous growth models. *Contemporary Economic Policy*, *17*(2), 177-188.

Katz, C. J., Mahler, V. A., & Franz, M. G. (1983). The impact of taxes on growth and distribution in developed capitalist countries: A cross-national study. *American Political Science Review*, *77*(04), 871-886.

Kneller, R., Bleaney, M. F., & Gemmell, N. (1999). Fiscal policy and growth: Evidence from OECD countries. *Journal of Public Economics*, *74*(2), 171-190.

Mendoza, E. G., Milesi-Ferretti, G. M., & Asea, P. (1997). On the ineffectiveness of tax policy in altering long-run growth: Harberger's superneutrality conjecture. *Journal of Public Economics*, *66*(1), 99-126.

Miller, S. M., & Russek, F. S. (1997). Fiscal structures and economic growth: International evidence. *Economic Inquiry*, *35*(3), 603-613.

Muinelo-Gallo, L., & Roca-Sagalés, O. (2013). Joint determinants of fiscal policy, income inequality and economic growth. *Economic Modelling*, *30*, 814-824.

Padovano, F., & Galli, E. (2001). Tax rates and economic growth in the OECD countries. *Economic Inquiry*, *39*(1), 44-57.

Romero-Avila, D., & Strauch, R. (2008). Public finances and long-term growth in Europe: Evidence from a panel data analysis. *European Journal of Political Economy*, *24*(1), 172-191.

Volkerink, B., Sturm, J. E., & De Haan, J. (2002). Tax ratios in macroeconomics: Do taxes really matter?. *Empirica*, *29*(3), 209-224.

Widmalm, F. (2001). Tax structure and growth: Are some taxes better than others?. *Public Choice*, *107*(3-4), 199-219.

(ii) Conference Proceedings

Gemmell, N., Kneller, R., & Sanz, I. (2009). The growth effects of corporate and personal tax rates in the OECD. *Paper presented to the Victoria University of Wellington Tax Policy Conference*, Wellington.

(iii) Studies from Think Thanks and Research Firms

Afonso, A., & Alegre, J. G. (2008). Economic growth and budgetary components: A panel assessment for the EU. *European Central Bank Working Paper, 848.*

Arnold, J. (2008). Do tax structures affect aggregate aconomic growth?: Empirical evidence from a panel of OECD countries. OECD Economics Department Working Papers, *643, OECD Publishing.*DOI:http://dx.doi.org/10.1787/236001777843

Bassanini, A., Scarpetta, S., & Hemmings, P. (2001). Economic growth: The role of policies and institutions: Panel data. Evidence from OECD countries. OECD Economics Department Working Papers, *283, OECD Publishing*. DOI: http://dx.doi.org/10.1787/18151973

Benos, N. (2009). Fiscal policy and economic growth: Empirical evidence from EU countries. *Center for Planning and Economic Research, Working Paper, 107*.

De la Fuente, A. (1997). Fiscal policy and growth in the OECD. CEPR Discussion Paper, NO. 1755.

Bergh, A., & Öhrn, N. (2011). Growth effects of fiscal policies: A critical appraisal of Colombier's (2009) study. *IFN Working Paper*, *865*.

Garcia-Escribano, M., & Mehrez, G. (2004). The impact of government size and the composition of revenue and expenditure on growth. *IMF Country Report, 04/237*.

Daveri, F., & Tabellini, G. (1997). Unemployment, growth and taxation in industrial countries. *World Bank Research Project.*

(iv) Theses/ Dissertations

Arin, K. P. (2003). An empirical investigation of tax policy in G-7 countries(Doctoral dissertation, Louisiana State University).

(v) Working papers and other unpublished research

Arin, K. P. (2004). Fiscal Policy, private investment and economic growth: Evidence from G-7 countries. *Massey University*.

Dackehag, M., & Hansson, Å. (2012). Taxation of income and economic growth: An empirical analysis of 25 rich OECD countries. *Lund University*, *Working Paper*, 2012:6.

Furceri, D., & Karras, G. (2009). Tax changes and economic growth: Empirical evidence for a panel of OECD countries. *Manuscript, University of Illinois*.

Gemmell, N., Kneller, R., & Sanz, I. (2008). Fiscal policy impacts on growth in the OECD: Are they long-run?. *University of Nottingham.*

Gemmell, N., Kneller, R., & Sanz,I (2012). Does the composition of government expenditure matter for economic growth. *manuscript, University of Nottingham*.

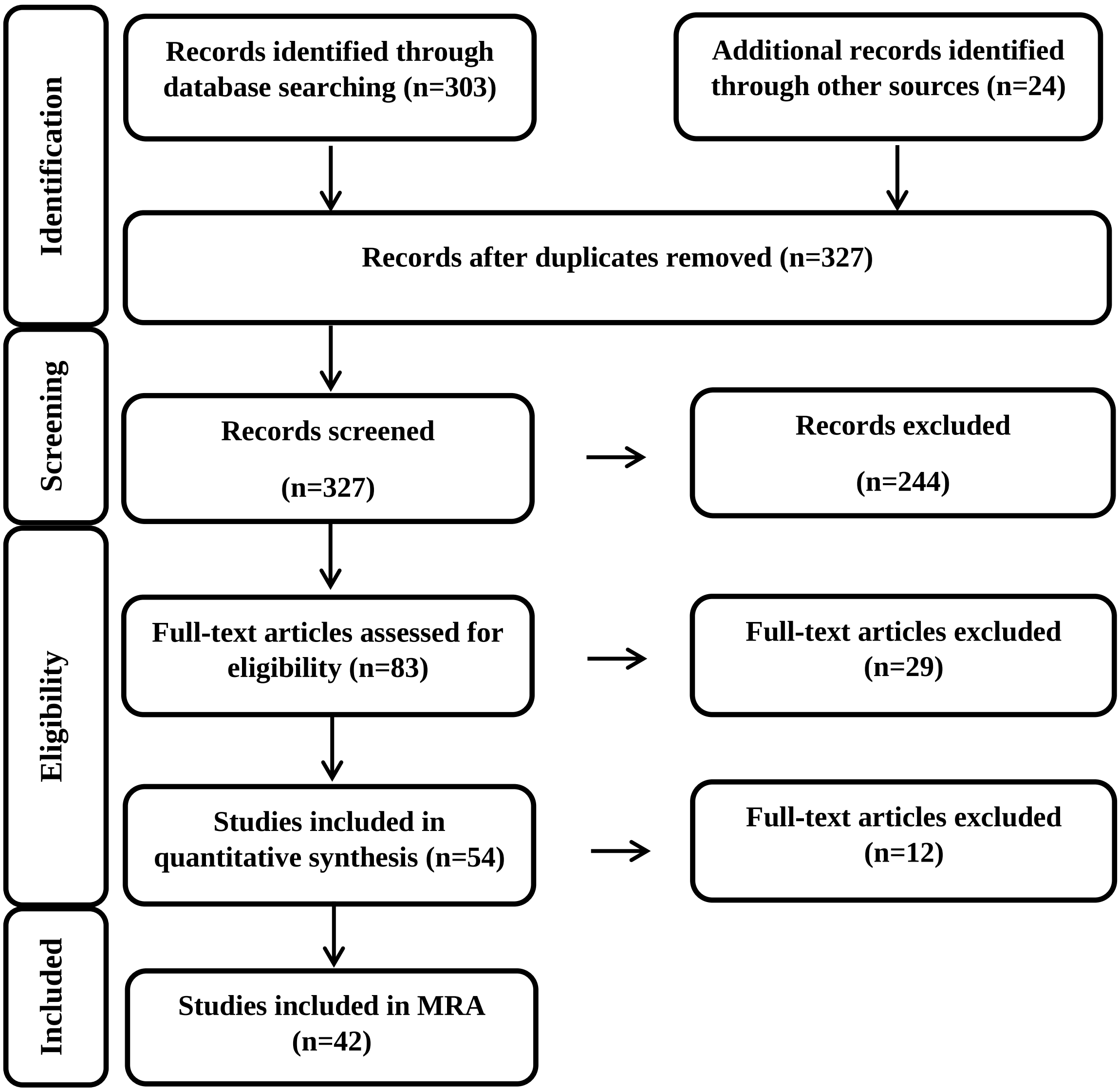
Gemmell, N., Kneller, R., & Sanz, I. (2011). The growth effects of corporate and personal tax rates in the OECD.*Bond University* , *Working Paper, 49*.

Xing, J. (2011). Does tax structure affect economic growth? Empirical evidence from OECD countries.

Appendix 2.4: Final Sample of Studies

ID	Study	Publication Status	Number of estimates
1	Afonso and Alegre (2008, 2011)	Working Paper + Journal	12
2	Afonso and Furceri (2010)	Journal	6
3	Afonso and Jalles (2013, 2014)	Working Paper + Journal	21
4	Agell et al. (1997)	Journal	3
5	Agell et al. (1999)	Journal	4
6	Agell et al. (2006)	Journal	4
7	Alesina and Ardagna (2010)	Journal	26
8	Angelopoulos et al. (2007)	Journal	36
9	Arin (2004)	Working Paper	80
10	Arnold et al. (2011)	Journal	5
11	Arnold (2008)	Working Paper	18
12	Baskaran and Feld (2013)	Journal	12
13	Bergh and Karlsson (2010)	Journal	3
14	Bergh and Ohrn (2011)	Working Paper	10
15	Bleaney et al. (2001)	Journal	19
16	Colombier (2009)	Journal	13
17	Daveri et al. (1997, 2000)	Working Paper + Journal	6
18	De La Fuente (1997)	Discussion Paper	15
19	Folster and Henkerson (2001)	Journal	7
20	Folster and Henkerson (1999)	Journal	7
21	Furceri and Karras (2009)	Working Paper	43
22	Gemmell et al. (2015)	Journal	10
23	Gemmell et al. (2008)	Working Paper	18
24	Gemmell et al. (2014)	Journal	53
25	Gemmell et al. (2011)	Journal	19
26	Hansson (2010)	Journal	23
27	Heitger (1993)	Journal	2
28	Karras and Furceri (2009)	Journal	32
29	Karras (1999)	Journal	28
30	Kneller et al. (1999)	Journal	35
31	Mendoza et al. (1997)	Journal	11
32	Miller and Russek (1997)	Journal	12
33	Muinelo-Gallo and Roca-Sagales (2013)	Journal	6
34	Padovano and Galli (2001)	Journal	2
35	Romero-Avila and Strauch (2008)	Journal	15
36	Volkerink et al. (2002)	Journal	26
37	Widmalm (2001)	Journal	6
38	Xing (2011)	Working Paper	34
39	Abd Hakim et al. (2013)	Conference Paper	2
40	Arin et al. (2015)	Working Paper	6
41	Paparas et al. (2015)	Journal	16
42	Xing (2012)	Journal	7

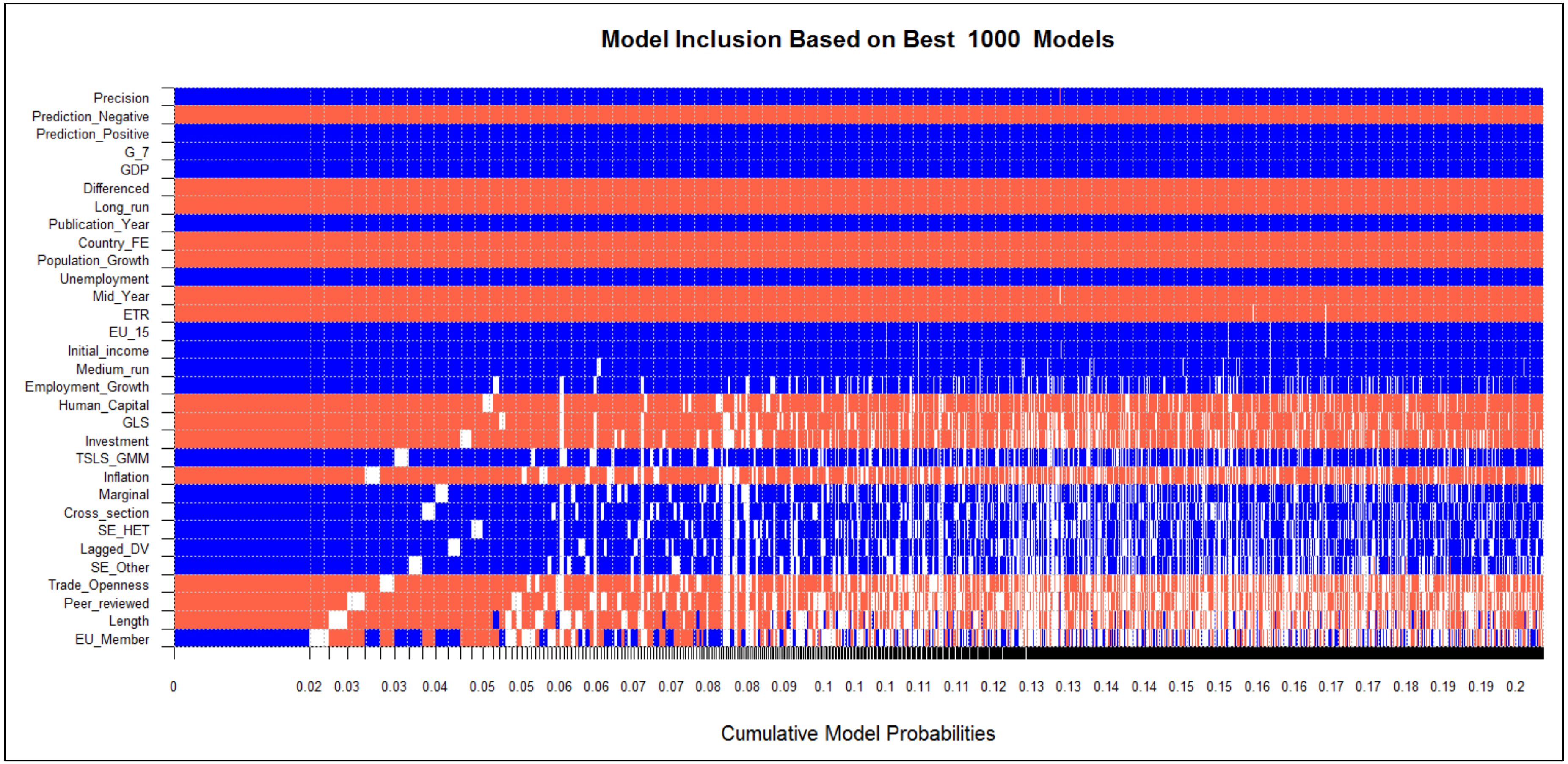
Appendix 2.5: Literature Search Flow Diagram



Appendix 2.6: List of Countries with Groupings

Year	*OECD*	*EU*	*EU-15*	*G-7*
1961	Austria	Austria	Austria	
1961	Belgium	Belgium	Belgium	
1961	Canada			Canada
1961	Denmark	Denmark	Denmark	
1961	France	France	France	France
1961	Germany	Germany	Germany	Germany
1961	Greece	Greece	Greece	
1961	Iceland			
1961	Ireland	Ireland	Ireland	
1961	Luxembourg	Luxembourg	Luxembourg	
1961	Netherlands	Netherlands	Netherlands	
1961	Norway			
1961	Portugal	Portugal	Portugal	
1961	Spain	Spain	Spain	
1961	Sweden	Sweden	Sweden	
1961	Switzerland			
1961	Turkey			
1961	United Kingdom	UK	UK	UK
1961	United States			USA
1962	Italy	Italy	Italy	Italy
1964	Japan			Japan
1969	Finland	Finland	Finland	
1971	Australia			
1973	New Zealand			
1994	Mexico			
1995	Czech Republic	Czech Republic		
1996	Hungary	Hungry		
1996	Korea			
1996	Poland	Poland		
2000	Slovak Republic	Slovak Republic		
2010	Chile			
2010	Estonia	Estonia		
2010	Israel			
2010	Slovenia	Slovenia		

Appendix 2.7: Visual Representation of BMA Analysis (Random Effects-Weight2)



Note: Each column represents a single model. Variables are listed in descending order of posterior inclusion probability (PIP) and have all been weighted according to the *Fixed Effects – Weight 1* case. Blue (dark) indicates that the variable is included in that model and estimated to be positive. Red (light) indicates the variable is included and estimated to be negative. No colour (white) indicates the variable is not included in that model. Further details about this plot is given in Zeugner (2011).

Appendix 2.8: Stata Codes

.do file for TABLE 2.4

```
clear
log using "\\file\UsersN$\nal53\Home\Desktop\(Nazila's program)\OECD\SampleMeans.smcl", replace
set more off
set type double
graph drop _all
import excel "\\file\UsersN$\nal53\Home\Desktop\(Nazila's program)\OECD\TAX.xlsx", sheet("Stata") firstrow
case (lower)
//TABLE 4-Descriptive Statistica, Column 1(Estimated Tax Effects)-Full
summ coefficient, detail
//TABLE 4-Descriptive Statistica, Column 3(t-statistics)-Full
summ tstatistics, detail
//Mean absolute value of t-statitics
gen abststat=abs(tstat)
summ abststat
//TABLE 4-Descriptive Statistica, Column 2(Estimated Tax Effects)-Truncated
summ coefficient, detail
scalar low = r(p5)
scalar high = r(p95)
gen endog = (tsls == 1 | gmm == 1)
keep if coefficient > low & coefficient < high
summ coefficient, detail
//TABLE 4-Descriptive Statistica, Column 4(t-statistics)-Truncated
summ tstatistics, detail
//Mean absolute value of t-statitics
replace abststat=abs(tstat)
summ abststat
log close
```

.do file for TABLE 2.5

```
clear
log using "\\file\UsersN$\nal53\Home\Desktop\(Nazila's program)\OECD\Part1 Results.smcl", replace
set more off
set type double
graph drop _all
import excel "\\file\UsersN$\nal53\Home\Desktop\(Nazila's program)\OECD\TAX.xlsx", sheet("Stata") firstrow
case (lower)
summ coefficient, detail
scalar low = r(p5)
scalar high = r(p95)
summ tstat, detail
summ coefficient tstat if coefficient > low & coefficient < high, detail

keep if coefficient > low & coefficient < high

// Generating transformed variables for FE and RE
gen feprecision = 1/se
gen fetstat = coefficient/se

*-------------------------------------------------------------------------------*
*   TABLE 5: FAT/PET                                                            *
*-------------------------------------------------------------------------------*

// Fixed Effects with SE
//NOTE: If the coefficient on the constant term is significant, that
//is evidence of publication bias
// Fixed Effects
//This regression gives equal weight to each observation
// TABLE 5 - Column 1
regress fetstat feprecision, vce(cluster idstudy)
//This regression gives equal weight to each study
// TABLE 5 - Column 2
regress fetstat feprecision [pweight = weight], vce(cluster idstudy)
// Fixed Effects without SE
//NOTE: If the coefficient on the constant term is significant, that
//is evidence of publication bias
// Fixed Effects
//This regression gives equal weight to each observation
// TABLE 5 - (not reported)
regress fetstat feprecision, noc vce(cluster idstudy)
//This regression gives equal weight to each study
// TABLE 5 - (not reported)
regress fetstat feprecision [pweight = weight], noc vce(cluster idstudy)


metan coefficient se, random
scalar tau2 = r(tau2)
gen revar = se^2 + tau2
gen rese = sqrt(revar)
gen reprecision = 1/rese
gen retstat = coefficient/rese
```

```
// Random Effects with SE
//This regression gives equal weight to each observation
// TABLE 5 - Column 3
regress retstat reprecision,  vce(cluster idstudy)
//This regression gives equal weight to each study
// TABLE 5 - Column 4
regress retstat reprecision [pweight = weight], vce(cluster idstudy)
// TABLE 5 - Column 5
regress retstat reprecision, noc vce(cluster idstudy)
//This regression gives equal weight to each study
// TABLE 5 - Column 6
regress retstat reprecision [pweight = weight], noc vce(cluster idstudy)

log close
```

.do file for TABLE 2.6

```
clear
log using "\\file\UsersN$\nal53\Home\Desktop\(Nazila's program)\OECD\Part1 Results.smcl", replace
set more off
set type double
graph drop _all
import excel "\\file\UsersN$\nal53\Home\Desktop\(Nazila's program)\OECD\TAX.xlsx", sheet("Stata") firstrow
case (lower)
summ coefficient, detail
scalar low = r(p5)
scalar high = r(p95)

keep if coefficient > low & coefficient < high
gen endog = (tsls == 1 | gmm == 1)
summ predneg predother predpos ///
labourtax capitaltax consumptiontax othertaxes mixedtaxes overalltax ///
g7 eu15 eumem oecd ///
gdp pcgdp ///
marginal differenced etr ///
lrcase1 lrcase2 lrcase3 ///
peerreviewed originalpubyear ///
cs panel ///
length originalmidyear ///
ols gls endog ///
seols sehet sehac ///
income laggeddv countryfe ///
investment tradeopenness human popgrowth employgrowth unemploymentrate inflation
```

.do file for TABLE 2.7

```
clear
log using "\\file\UsersN$\nal53\Home\Desktop\(Nazila's program)\OECD\Part1 Results.smcl", replace
set more off
set type double
graph drop _all
import excel "\\file\UsersN$\nal53\Home\Desktop\(Nazila's program)\OECD\TAX.xlsx", sheet("Stata") firstrow
case (lower)
quietly summ coefficient, detail
scalar low = r(p5)
scalar high = r(p95)
keep if coefficient > low & coefficient < high

gen endog = (tsls == 1 | gmm == 1)

// Generating transformed variables for FE and RE
gen feprecision = 1/se
gen fetstat = coefficient/se
gen peerreviewedd = peerreviewed/se
gen pubyearr = pubyear/se
gen css = cs/se
gen lengthh = length/se
gen midyearr = midyear/se
gen gdpp = gdp/se
gen labourtaxx = labourtax/se
gen capitaltaxx = capitaltax/se
gen overalltaxx = overalltax/se
gen othertaxess = othertaxes/se
gen mixedtaxess = mixedtaxes/se
gen marginall = marginal/se
gen differencedd = differenced/se
gen etrr = etr/se
gen prednegg = predneg/se
gen predposs = predpos/se
gen investmentt = investment/se
gen tradeopennesss = tradeopenness/se
gen humann = human/se
gen popgrowthh = popgrowth/se
gen employgrowthh = employgrowth/se
gen unemploymentt = unemploymentrate/se
gen inflationn = inflation/se
gen incomee = income/se
gen laggeddvv = laggeddv/se
gen countryfee = countryfe/se
gen sehacc = sehac/se
gen sehett = sehet/se
gen lrcase22 = lrcase2/se
gen lrcase33 = lrcase3/se
gen glss = gls/se
gen endogg = endog/se
gen eu155 = eu15/se
gen g77 = g7/se
```

```
gen eumemm = eumem/se



*********************************************
*********** FIXED EFFECTS ********************
*********************************************

// These specifications include the SeR term
// NOTE: The constant term is the SER term

//THIS SECTION GIVES EQUAL WEIGHT TO EACH OBSERVATION
// Basic regression with no control variables
reg fetstat feprecision prednegg predposs,  vce(cluster idstudy)
display e(r2_a)

//TABLE 7, Collumn 1, Top Panel
// Regression with all control variables
reg fetstat feprecision prednegg predposs ///
peerreviewedd pubyearr css lengthh midyearr gdpp marginall differencedd etrr  ///
investmentt tradeopennesss humann popgrowthh employgrowthh unemploymentt ///
inflationn incomee laggeddvv countryfee sehacc sehett lrcase22 lrcase33 ///
glss endogg eu155 g77 eumemm,  vce(cluster idstudy)
display e(r2_a)

// Here we use backwards stepwise regression selecting on the best regression using the BIC criterion
// We lock in the variables: feprecision labourtaxx capitaltaxx consumptiontaxx othertaxess prednegg predposs
vselect fetstat peerreviewedd pubyearr css lengthh midyearr gdpp marginall differencedd etrr  ///
investmentt tradeopennesss humann popgrowthh employgrowthh unemploymentt inflationn incomee laggeddvv ///
countryfee sehacc sehett lrcase22 lrcase33 glss endogg eu155 g77 eumemm, backward bic ///
fix(feprecision prednegg predposs)

//TABLE 7, Collumn 1, Bottom Panel
// We take the results from the preceding stepwise regression and reestimate the best model
// but this time with cluster robust standard errors
reg fetstat feprecision   ///
prednegg predposs peerreviewedd gdpp popgrowthh employgrowthh unemploymentt laggeddvv ///
countryfee sehacc lrcase22 glss eu155 g77 ,  vce(cluster idstudy)
display e(r2_a)

//THIS SECTION GIVES EQUAL WEIGHT TO EACH STUDY
// Basic regression with no control variables
reg fetstat feprecision prednegg predposs  [pweight = weight], vce(cluster idstudy)
display e(r2_a)

//TABLE 7, Collumn 2, Top Panel
// Regression with all control variables
reg fetstat feprecision prednegg predposs ///
peerreviewedd pubyearr css lengthh midyearr gdpp marginall differencedd etrr  ///
investmentt tradeopennesss humann popgrowthh employgrowthh unemploymentt ///
inflationn incomee laggeddvv countryfee sehacc sehett lrcase22 lrcase33 ///
glss endogg eu155 g77 eumemm [pweight = weight],  vce(cluster idstudy)
display e(r2_a)

// Here we use backwards stepwise regression selecting on the best regression using the BIC criterion
// We lock in the variables: feprecision labourtaxx capitaltaxx consumptiontaxx othertaxess prednegg predposs
```

```
vselect fetstat peerreviewedd pubyearr css lengthh midyearr gdpp marginall differencedd etrr  ///
investmentt tradeopennesss humann popgrowthh employgrowthh unemploymentt inflationn incomee laggeddvv ///
countryfee sehacc sehett lrcase22 lrcase33 glss endogg eu155 g77 eumemm [pweight = weight], backward bic ///
fix(feprecision prednegg predposs)

//TABLE 7, Collumn 2, Bottom Panel
// We take the results from the preceding stepwise regression and reestimate the best model
// but this time with cluster robust standard errors
reg fetstat feprecision   ///
prednegg predposs peerreviewedd pubyearr lengthh midyearr differencedd popgrowthh employgrowthh unemploymentt ///
incomee countryfee sehacc lrcase33 glss eu155 g77 [pweight = weight], vce(cluster idstudy)
display e(r2_a)

**********************************************
************ RANDOM EFFECTS ******************
**********************************************

metareg coefficient predneg predpos, wsse(se)
scalar tau2 = e(tau2)
display tau2
gen revar = se^2 + tau2
gen rese = sqrt(revar)

gen reprecision = 1/rese
gen retstat = coefficient/rese
replace peerreviewedd = peerreviewed/rese
replace pubyearr = pubyear/rese
replace css = cs/rese
replace lengthh = length/rese
replace midyearr = midyear/rese
replace gdpp = gdp/rese
replace labourtaxx = labourtax/rese
replace capitaltaxx = capitaltax/rese
replace overalltaxx = overalltax/se
replace othertaxess = othertaxes/rese
replace mixedtaxess = mixedtaxes/rese
replace marginall = marginal/rese
replace differencedd = differenced/rese
replace etrr = etr/rese
replace prednegg = predneg/rese
replace predposs = predpos/rese
replace investmentt = investment/rese
replace tradeopennesss = tradeopenness/rese
replace humann = human/rese
replace popgrowthh = popgrowth/rese
replace employgrowthh = employgrowth/rese
replace unemploymentt = unemploymentrate/rese
replace inflationn = inflation/rese
replace incomee = income/rese
replace laggeddvv = laggeddv/rese
replace countryfee = countryfe/rese
replace sehacc = sehac/rese
replace sehett = sehet/rese
replace lrcase22 = lrcase2/rese
```

```
replace lrcase33 = lrcase3/rese
replace glss = gls/rese
replace endogg = endog/rese
replace eu155 = eu15/rese
replace g77 = g7/rese
replace eumemm = eumem/rese


//THIS SECTION GIVES EQUAL WEIGHT TO EACH OBSERVATION
// Basic regression with no control variables
reg retstat reprecision prednegg predposs,  vce(cluster idstudy)
display e(r2_a)

//TABLE 7, Collumn 3, Top Panel
// Regression with all control variables
reg retstat reprecision prednegg predposs ///
peerreviewedd pubyearr css lengthh midyearr gdpp marginall differencedd etrr  ///
investmentt tradeopennesss humann popgrowthh employgrowthh unemploymentt ///
inflationn incomee laggeddvv countryfee sehacc sehett lrcase22 lrcase33 ///
glss endogg eu155 g77 eumemm,  vce(cluster idstudy)
display e(r2_a)

// Here we use backwards stepwise regression selecting on the best regression using the BIC criterion
// We lock in the variables: feprecision labourtaxx capitaltaxx consumptiontaxx othertaxess prednegg predposs
vselect retstat peerreviewedd pubyearr css lengthh midyearr gdpp marginall differencedd etrr  ///
investmentt tradeopennesss humann popgrowthh employgrowthh unemploymentt inflationn incomee laggeddvv ///
countryfee sehacc sehett lrcase22 lrcase33 glss endogg eu155 g77 eumemm, backward bic ///
fix(reprecision prednegg predposs)

//TABLE 7, Collumn 3, Bottom Panel
// We take the results from the preceding stepwise regression and reestimate the best model
// but this time with cluster robust standard errors
reg retstat reprecision   ///
prednegg predposs pubyearr midyearr gdpp marginall differencedd popgrowthh unemploymentt ///
countryfee lrcase33 g77,  vce(cluster idstudy)
display e(r2_a)

//THIS SECTION GIVES EQUAL WEIGHT TO EACH STUDY
// Basic regression with no control variables
reg retstat reprecision prednegg predposs [pweight = weight], vce(cluster id)
display e(r2_a)

//TABLE 7, Collumn 4, Top Panel
// Regression with all control variables
reg retstat reprecision prednegg predposs ///
peerreviewedd pubyearr css lengthh midyearr gdpp marginall differencedd etrr  ///
investmentt tradeopennesss humann popgrowthh employgrowthh unemploymentt ///
inflationn incomee laggeddvv countryfee sehacc sehett lrcase22 lrcase33 ///
glss endogg eu155 g77 eumemm [pweight = weight],  vce(cluster idstudy)
display e(r2_a)

// Here we use backwards stepwise regression selecting on the best regression using the BIC criterion
// We lock in the variables: feprecision labourtaxx capitaltaxx consumptiontaxx othertaxess prednegg predposs
vselect retstat peerreviewedd pubyearr css lengthh midyearr gdpp marginall differencedd etrr  ///
investmentt tradeopennesss humann popgrowthh employgrowthh unemploymentt inflationn incomee laggeddvv ///
```

```
countryfee sehacc sehett lrcase22 lrcase33 glss endogg eu155 g77 eumemm [pweight = weight], backward bic
///
fix(reprecision prednegg predposs)

//TABLE 7, Collumn 4, Bottom Panel
// We take the results from the preceding stepwise regression and reestimate the best model
// but this time with cluster robust standard errors
reg retstat reprecision ///
prednegg predposs pubyearr gdpp marginall differencedd popgrowthh ///
unemploymentt incomee countryfee lrcase33 endogg [pweight = weight], vce(cluster id)
display e(r2_a)

log close
```

.do file for TABLE 2.8

```
clear
log using "\\file\UsersN$\nal53\Home\Desktop\(Nazila's program)\OECD\Type of Tax Results.smcl", replace
set more off
set type double
graph drop _all
import excel "\\file\UsersN$\nal53\Home\Desktop\(Nazila's program)\OECD\TAX.xlsx", sheet("Stata") firstrow
case (lower)
quietly summ coefficient, detail
scalar low = r(p5)
scalar high = r(p95)
keep if coefficient > low & coefficient < high

gen endog = (tsls == 1 | gmm == 1)

// Generating transformed variables for FE and RE
gen feprecision = 1/se
gen fetstat = coefficient/se
gen peerreviewedd = peerreviewed/se
gen pubyearr = pubyear/se
gen css = cs/se
gen lengthh = length/se
gen midyearr = midyear/se
gen gdpp = gdp/se
gen labourtaxx = labourtax/se
gen capitaltaxx = capitaltax/se
gen overalltaxx = overalltax/se
//gen consumptiontaxx = consumptiontax/se
gen othertaxess = othertaxes/se
gen mixedtaxess = mixedtaxes/se
gen marginall = marginal/se
gen differencedd = differenced/se
gen etrr = etr/se
gen prednegg = predneg/se
gen predposs = predpos/se
gen investmentt = investment/se
gen tradeopennesss = tradeopenness/se
gen humann = human/se
gen popgrowthh = popgrowth/se
gen employgrowthh = employgrowth/se
gen unemploymentt = unemploymentrate/se
gen inflationn = inflation/se
gen incomee = income/se
gen laggeddvv = laggeddv/se
gen countryfee = countryfe/se
gen sehacc = sehac/se
gen sehett = sehet/se
gen lrcase22 = lrcase2/se
gen lrcase33 = lrcase3/se
gen glss = gls/se
gen endogg = endog/se
gen eu155 = eu15/se
```

```
gen g77 = g7/se
gen eumemm = eumem/se


*********************************************
*********** FIXED EFFECTS ********************
*********************************************

// These specifications include the SeR term
// NOTE: The constant term is the SER term

//THIS SECTION GIVES EQUAL WEIGHT TO EACH OBSERVATION
// Basic regression with no control variables
reg fetstat feprecision labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx  ///
,  vce(cluster idstudy)
display e(r2_a)

//TABLE 8, Collumn 1, Top Panel
// Regression with all control variables
reg fetstat feprecision labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx  ///
peerreviewedd pubyearr css lengthh midyearr gdpp marginall differencedd etrr  ///
investmentt tradeopennesss humann popgrowthh employgrowthh unemploymentt inflationn incomee laggeddvv ///
countryfee sehacc sehett lrcase22 lrcase33 glss endogg eu155 g77 eumemm,  vce(cluster idstudy)
display e(r2_a)

// Here we use backwards stepwise regression selecting on the best regression using the BIC criterion
// We lock in the variables: feprecision labourtaxx capitaltaxx consumptiontaxx othertaxess prednegg predposs
vselect fetstat peerreviewedd pubyearr css lengthh midyearr gspp pcgspp pii marginall differencedd etrr  ///
hlkk incomee laggeddvv statefee sehacc sehett lrcase22 lrcase33 ///
glss endogg nonee akk dcc akdcc akhidcotherss, backward bic ///
fix(feprecision labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess overalltaxx )

//TABLE 8, Collumn 1, Bottom Panel
// We take the results from the preceding stepwise regression and reestimate the best model
// but this time with cluster robust standard errors
reg fetstat feprecision prednegg predposs labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx  ///
 peerreviewedd popgrowthh employgrowthh unemploymentt laggeddvv ///
countryfee sehacc lrcase22 eu155 g77 ,  vce(cluster idstudy)
display e(r2_a)

//THIS SECTION GIVES EQUAL WEIGHT TO EACH STUDY
// Basic regression with no control variables
reg fetstat feprecision prednegg predposs labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx  ///
 [pweight = weight], vce(cluster idstudy)
display e(r2_a)

//TABLE 8, Collumn 2, Top Panel
// Regression with all control variables
reg fetstat feprecision prednegg predposs labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx  ///
peerreviewedd pubyearr css lengthh midyearr gdpp marginall differencedd etrr  ///
investmentt tradeopennesss humann popgrowthh employgrowthh unemploymentt inflationn incomee laggeddvv ///
countryfee sehacc sehett lrcase22 lrcase33 glss endogg eu155 g77 eumemm  [pweight = weight],  vce(cluster idstudy)
display e(r2_a)
```

```
// Here we use backwards stepwise regression selecting on the best regression using the BIC criterion
// We lock in the variables: feprecision labourtaxx capitaltaxx consumptiontaxx othertaxess prednegg predposs
vselect fetstat peerreviewedd pubyearr css lengthh midyearr gdpp marginall differencedd etrr  ///
investmentt tradeopennesss humann popgrowthh employgrowthh unemploymentt inflationn incomee laggeddvv ///
countryfee sehacc sehett lrcase22 lrcase33 glss endogg eu155 g77 eumemm [pweight = weight], backward bic ///
fix(feprecision prednegg predposs labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx)

//TABLE 8, Collumn 2, Bottom Panel
// We take the results from the preceding stepwise regression and reestimate the best model
// but this time with cluster robust standard errors
reg fetstat feprecision prednegg predposs labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx  ///
 peerreviewedd pubyearr lengthh midyearr humann popgrowthh employgrowthh unemploymentt ///
incomee laggeddvv countryfee lrcase22 lrcase33 eu155 g77 [pweight = weight], vce(cluster idstudy)
display e(r2_a)

**********************************************
************ RANDOM EFFECTS ******************
**********************************************

metareg coefficient labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess overalltaxx  ///
, wsse(se)
scalar tau2 = e(tau2)
display tau2
gen revar = se^2 + tau2
gen rese = sqrt(revar)

gen reprecision = 1/rese
gen retstat = coefficient/rese
replace peerreviewedd = peerreviewed/rese
replace pubyearr = pubyear/rese
replace css = cs/rese
replace lengthh = length/rese
replace midyearr = midyear/rese
replace gdpp = gdp/rese
replace labourtaxx = labourtax/rese
replace capitaltaxx = capitaltax/rese
replace overalltaxx = overalltax/se
replace othertaxess = othertaxes/rese
replace mixedtaxess = mixedtaxes/rese
replace marginall = marginal/rese
replace differencedd = differenced/rese
replace etrr = etr/rese
replace prednegg = predneg/rese
replace predposs = predpos/rese
replace investmentt = investment/rese
replace tradeopennesss = tradeopenness/rese
replace humann = human/rese
replace popgrowthh = popgrowth/rese
replace employgrowthh = employgrowth/rese
replace unemploymentt = unemploymentrate/rese
replace inflationn = inflation/rese
replace incomee = income/rese
replace laggeddvv = laggeddv/rese
replace countryfee = countryfe/rese
replace sehacc = sehac/rese
```

```
replace sehett = sehet/rese
replace lrcase22 = lrcase2/rese
replace lrcase33 = lrcase3/rese
replace glss = gls/rese
replace endogg = endog/rese
replace eu155 = eu15/rese
replace g77 = g7/rese
replace eumemm = eumem/rese

//THIS SECTION GIVES EQUAL WEIGHT TO EACH OBSERVATION
// Basic regression with no control variables
reg retstat reprecision prednegg predposs labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx  ///
,  vce(cluster idstudy)
display e(r2_a)

//TABLE 8, Collumn 3, Top Panel
// Regression with all control variables
reg retstat reprecision prednegg predposs labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx  ///
peerreviewedd pubyearr css lengthh midyearr gdpp marginall differencedd etrr  ///
investmentt tradeopennesss humann popgrowthh employgrowthh unemploymentt inflationn incomee laggeddvv ///
countryfee sehacc sehett lrcase22 lrcase33 glss endogg eu155 g77 eumemm,  vce(cluster idstudy)
display e(r2_a)

// Here we use backwards stepwise regression selecting on the best regression using the BIC criterion
// We lock in the variables: feprecision labourtaxx capitaltaxx consumptiontaxx othertaxess prednegg predposs
vselect retstat peerreviewedd pubyearr css lengthh midyearr gdpp marginall differencedd etrr  ///
investmentt tradeopennesss humann popgrowthh employgrowthh unemploymentt inflationn incomee laggeddvv ///
countryfee sehacc sehett lrcase22 lrcase33 glss endogg eu155 g77 eumemm, backward bic ///
fix(reprecision prednegg predposs labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx)

//TABLE 8, Collumn 3, Bottom Panel
// We take the results from the preceding stepwise regression and reestimate the best model
// but this time with cluster robust standard errors
reg retstat reprecision prednegg predposs labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx  ///
 pubyearr midyearr gdpp marginall differencedd popgrowthh unemploymentt ///
inflationn incomee countryfee sehacc lrcase22 lrcase33 glss eu155 g77,  vce(cluster idstudy)
display e(r2_a)

//THIS SECTION GIVES EQUAL WEIGHT TO EACH STUDY
// Basic regression with no control variables
reg retstat reprecision prednegg predposs labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx  ///
 [pweight = weight], vce(cluster id)
display e(r2_a)

//TABLE 8, Collumn 4, Top Panel
//Regression with all control variables
reg retstat reprecision prednegg predposs labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx  ///
peerreviewedd pubyearr css lengthh midyearr gdpp marginall differencedd etrr  ///
investmentt tradeopennesss humann popgrowthh employgrowthh unemploymentt inflationn incomee laggeddvv ///
countryfee sehacc sehett lrcase22 lrcase33 glss endogg eu155 g77 eumemm [pweight = weight], vce(cluster idstudy)
display e(r2_a)

// Here we use backwards stepwise regression selecting on the best regression using the BIC criterion
```

```
// We lock in the variables: feprecision labourtaxx capitaltaxx consumptiontaxx othertaxess prednegg predposs
vselect retstat peerreviewedd pubyearr css lengthh midyearr gdpp marginall differencedd etrr  ///
investmentt tradeopennesss humann popgrowthh employgrowthh unemploymentt inflationn incomee laggeddvv ///
countryfee sehacc sehett lrcase22 lrcase33 glss endogg eu155 g77 eumemm [pweight = weight], backward bic ///
fix(reprecision prednegg predposs labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx)

//TABLE 8, Collumn 4, Bottom Panel
// We take the results from the preceding stepwise regression and reestimate the best model
// but this time with cluster robust standard errors
reg retstat reprecision prednegg predposs labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx  ///
pubyearr gdpp differencedd etrr humann popgrowthh ///
unemploymentt incomee countryfee lrcase33 eu155 [pweight = weight], vce(cluster id)
display e(r2_a)

log close
```

.do file for TABLE 2.9

```
clear
log using "\\file\UsersN$\nal53\Home\Desktop\(Nazila's program)\OECD\All Tax Var Results.smcl", replace
set more off
set type double
graph drop _all
import excel "\\file\UsersN$\nal53\Home\Desktop\(Nazila's program)\OECD\TAX.xlsx", sheet("Stata") firstrow
case (lower)
quietly summ coefficient, detail
scalar low = r(p5)
scalar high = r(p95)
keep if coefficient > low & coefficient < high

gen endog = (tsls == 1 | gmm == 1)

// Generating transformed variables for FE and RE
gen feprecision = 1/se
gen fetstat = coefficient/se
gen peerreviewedd = peerreviewed/se
gen pubyearr = pubyear/se
gen css = cs/se
gen lengthh = length/se
gen midyearr = midyear/se
gen gdpp = gdp/se
gen labourtaxx = labourtax/se
gen capitaltaxx = capitaltax/se
gen overalltaxx = overalltax/se
//gen consumptiontaxx = consumptiontax/se
gen othertaxess = othertaxes/se
gen mixedtaxess = mixedtaxes/se
gen marginall = marginal/se
gen differencedd = differenced/se
gen etrr = etr/se
gen prednegg = predneg/se
gen predposs = predpos/se
gen investmentt = investment/se
gen tradeopennesss = tradeopenness/se
gen humann = human/se
gen popgrowthh = popgrowth/se
gen employgrowthh = employgrowth/se
gen unemploymentt = unemploymentrate/se
gen inflationn = inflation/se
gen incomee = income/se
gen laggeddvv = laggeddv/se
gen countryfee = countryfe/se
gen sehacc = sehac/se
gen sehett = sehet/se
gen lrcase22 = lrcase2/se
gen lrcase33 = lrcase3/se
gen glss = gls/se
gen endogg = endog/se
gen eu155 = eu15/se
gen g77 = g7/se
```

```
gen eumemm = eumem/se


*********************************************
************ FIXED EFFECTS ********************
*********************************************
// These specifications include the SeR term
// NOTE: The constant term is the SER term

//THIS SECTION GIVES EQUAL WEIGHT TO EACH OBSERVATION
// Basic regression with no control variables
reg fetstat feprecision labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx  ///
,  vce(cluster idstudy)
display e(r2_a)

//TABLE 9, Collumn 1, Top Panel
// Regression with all control variables
reg fetstat feprecision labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx  ///
peerreviewedd pubyearr css lengthh midyearr gdpp marginall differencedd etrr  ///
investmentt tradeopennesss humann popgrowthh employgrowthh unemploymentt inflationn incomee laggeddvv ///
countryfee sehacc sehett lrcase22 lrcase33 glss endogg eu155 g77 eumemm,  vce(cluster idstudy)
display e(r2_a)

// Here we use backwards stepwise regression selecting on the best regression using the BIC criterion
// We lock in the variables: feprecision labourtaxx capitaltaxx consumptiontaxx othertaxess prednegg predposs
vselect fetstat peerreviewedd pubyearr css lengthh midyearr gdpp marginall differencedd etrr  ///
investmentt tradeopennesss humann popgrowthh employgrowthh unemploymentt inflationn incomee laggeddvv ///
countryfee sehacc sehett lrcase22 lrcase33 glss endogg eu155 g77 eumemm, backward bic ///
fix(feprecision labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx )

//TABLE 9, Collumn 1, Bottom Panel
// We take the results from the preceding stepwise regression and reestimate the best model
// but this time with cluster robust standard errors
reg fetstat feprecision labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx  ///
peerreviewedd etrr popgrowthh employgrowthh unemploymentt laggeddvv ///
countryfee sehacc lrcase22 lrcase33 eu155 g77 ,  vce(cluster idstudy)
display e(r2_a)

//THIS SECTION GIVES EQUAL WEIGHT TO EACH STUDY
// Basic regression with no control variables
reg fetstat feprecision labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx  ///
 [pweight = weight], vce(cluster idstudy)
display e(r2_a)

//TABLE 9, Collumn 2, Top Panel
// Regression with all control variables
reg fetstat feprecision labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx  ///
peerreviewedd pubyearr css lengthh midyearr gdpp marginall differencedd etrr  ///
investmentt tradeopennesss humann popgrowthh employgrowthh unemploymentt inflationn incomee laggeddvv ///
countryfee sehacc sehett lrcase22 lrcase33 glss endogg eu155 g77 eumemm [pweight = weight],  vce(cluster idstudy)
display e(r2_a)

// Here we use backwards stepwise regression selecting on the best regression using the BIC criterion
```

```
// We lock in the variables: feprecision labourtaxx capitaltaxx consumptiontaxx othertaxess prednegg predposs
vselect fetstat peerreviewedd pubyearr css lengthh midyearr gdpp marginall differencedd etrr  ///
investmentt tradeopennesss humann popgrowthh employgrowthh unemploymentt inflationn incomee laggeddvv ///
countryfee sehacc sehett lrcase22 lrcase33 glss endogg eu155 g77 eumemm [pweight = weight], backward bic ///
fix(feprecision labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx)

//TABLE 9, Collumn 2, Bottom Panel
//We take the results from the preceding stepwise regression and reestimate the best model
// but this time with cluster robust standard errors
reg fetstat feprecision labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx  ///
peerreviewedd css midyearr gdpp humann employgrowthh unemploymentt ///
incomee sehacc lrcase33 glss eu155 g77 [pweight = weight], vce(cluster idstudy)
display e(r2_a)

**********************************************
************ RANDOM EFFECTS ******************
**********************************************
metareg coefficient labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess overalltaxx  ///
predneg predpos, wsse(se)
scalar tau2 = e(tau2)
display tau2
gen revar = se^2 + tau2
gen rese = sqrt(revar)

gen reprecision = 1/rese
gen retstat = coefficient/rese
replace peerreviewedd = peerreviewed/rese
replace pubyearr = pubyear/rese
replace css = cs/rese
replace lengthh = length/rese
replace midyearr = midyear/rese
replace gdpp = gdp/rese
replace labourtaxx = labourtax/rese
replace capitaltaxx = capitaltax/rese
replace overalltaxx = overalltax/se
replace othertaxess = othertaxes/rese
replace mixedtaxess = mixedtaxes/rese
replace marginall = marginal/rese
replace differencedd = differenced/rese
replace etrr = etr/rese
replace prednegg = predneg/rese
replace predposs = predpos/rese
replace investmentt = investment/rese
replace tradeopennesss = tradeopenness/rese
replace humann = human/rese
replace popgrowthh = popgrowth/rese
replace employgrowthh = employgrowth/rese
replace unemploymentt = unemploymentrate/rese
replace inflationn = inflation/rese
replace incomee = income/rese
replace laggeddvv = laggeddv/rese
replace countryfee = countryfe/rese
replace sehacc = sehac/rese
replace sehett = sehet/rese
replace lrcase22 = lrcase2/rese
```

```
replace lrcase33 = lrcase3/rese
replace glss = gls/rese
replace endogg = endog/rese
replace eu155 = eu15/rese
replace g77 = g7/rese
replace eumemm = eumem/rese

//THIS SECTION GIVES EQUAL WEIGHT TO EACH OBSERVATION
// Basic regression with no control variables
reg retstat reprecision labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx  ///
,  vce(cluster idstudy)
display e(r2_a)

//TABLE 9, Collumn 3, Top Panel
// Regression with all control variables
reg retstat reprecision labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx  ///
peerreviewedd pubyearr css lengthh midyearr gdpp marginall differencedd etrr  ///
investmentt tradeopennesss humann popgrowthh employgrowthh unemploymentt inflationn incomee laggeddvv ///
countryfee sehacc sehett lrcase22 lrcase33 glss endogg eu155 g77 eumemm,  vce(cluster idstudy)
display e(r2_a)

// Here we use backwards stepwise regression selecting on the best regression using the BIC criterion
// We lock in the variables: feprecision labourtaxx capitaltaxx consumptiontaxx othertaxess prednegg predposs
vselect retstat peerreviewedd pubyearr css lengthh midyearr gdpp marginall differencedd etrr  ///
investmentt tradeopennesss humann popgrowthh employgrowthh unemploymentt inflationn incomee laggeddvv ///
countryfee sehacc sehett lrcase22 lrcase33 glss endogg eu155 g77 eumemm, backward bic ///
fix(reprecision labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx)

//TABLE 9, Collumn 3, Bottom Panel
// We take the results from the preceding stepwise regression and reestimate the best model
// but this time with cluster robust standard errors
reg retstat reprecision labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx  ///
peerreviewedd pubyearr midyearr gdpp popgrowthh unemploymentt ///
incomee countryfee sehacc sehett lrcase33 glss eu155 g77,  vce(cluster idstudy)
display e(r2_a)

//THIS SECTION GIVES EQUAL WEIGHT TO EACH STUDY
// Basic regression with no control variables
reg retstat reprecision labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx  ///
 [pweight = weight], vce(cluster id)
display e(r2_a)

//TABLE 9, Collumn 4, Top Panel
// Regression with all control variables
reg retstat reprecision labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx  ///
peerreviewedd pubyearr css lengthh midyearr gdpp marginall differencedd etrr  ///
investmentt tradeopennesss humann popgrowthh employgrowthh unemploymentt inflationn incomee laggeddvv ///
countryfee sehacc sehett lrcase22 lrcase33 glss endogg eu155 g77 eumemm [pweight = weight], vce(cluster idstudy)
display e(r2_a)

// Here we use backwards stepwise regression selecting on the best regression using the BIC criterion
// We lock in the variables: feprecision labourtaxx capitaltaxx consumptiontaxx othertaxess prednegg predposs
vselect retstat peerreviewedd pubyearr css lengthh midyearr gdpp marginall differencedd etrr  ///
```

```
investmentt tradeopennesss humann popgrowthh employgrowthh unemploymentt inflationn incomee
laggeddvv ///
countryfee sehacc sehett lrcase22 lrcase33 glss endogg eu155 g77 eumemm [pweight = weight], backward bic
///
fix(reprecision labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx)

//TABLE 9, Collumn 4, Bottom Panel
// We take the results from the preceding stepwise regression and reestimate the best model
// but this time with cluster robust standard errors
reg retstat reprecision labourtaxx capitaltaxx othertaxess mixedtaxess overalltaxx  ///
pubyearr midyearr differencedd etrr humann popgrowthh ///
unemploymentt incomee countryfee lrcase33 eu155 g77 [pweight = weight], vce(cluster id)
display e(r2_a)



log close
```

R Commands for TABLE 2.10

Download R from the following link:
https://cran.r-project.org/src/base/R-3/
The one I am applying is R-3.2.1.tar.gz

After opening up the R, type the following commands:
install.packages() ⟶ New Zealand ⟶ ok ⟶ BMS ⟶ ok

Library(BMS)

The data file should have the dependent variable as the first column.

Open the data file (Excel spreadsheet) ⟶ copy data

```
TAX1=read.table("clipboard-512", sep="\t", header=TRUE)

TAX11 = bms(TAX1, burn=10000000, iter=10000000, g="hyper", mprior="random",
fixed.reg=c("Precision","Prediction_Negative","Prediction_Positive"),  nmodel=1000, mcmc="bd",
user.int=FALSE)

plot(TAX11)

summary(TAX11)

coef(TAX11, order.by.pip = T, exact=T, include.constant=T)

image(TAX11, cex.axis=0.7, order.by.pip = T, yprop2pip=F)
```

Chapter 3. Taxes and Economic Growth in U.S. States: A Meta-Regression Analysis

3.1. Introduction

A fundamental goal of economic policy makers is to encourage economic growth. Tax policy is considered to be one of the principal policy instruments for government to achieve this goal. Most economists would agree that taxes and spending are essential for economic growth. But they are uncertain as to what extent the negative economic effects of increasing taxes starts to outweigh the positive effects of increasing spending funded by these increased taxes. The previous chapter examined the effect of taxes on economic growth in OECD countries. This chapter pursues a similar line of study with respect to the effect of taxes on economic growth in U.S. states.

The effect of tax policies on economic growth among U.S. states has been the focus of numerous academic studies over the last decades. Prominent examples include Helms (1985), Miller and Russek (1997), Reed (2008), Mullen and Williams (1994), Tomljanovich (2004), and Yamarik (2000). While some empirical studies find that state and local taxes have a measurable and consistently adverse impact on state economic growth, other studies reach the opposite conclusion. Many more are mixed, ambivalent, or show any adverse impacts are small (Mazerove, 2013). Despite the fact that many of these studies use similar data and examine many of the same states and time periods, estimates vary widely. The result is a lack of consensus among economists about whether in the US taxes have any impact on economic growth and, if they do, how large the size of the effect might be.

There are several potential reasons that can explain heterogeneity across the reported estimates in the tax-growth literature. First, there is no settled theoretical prediction about the main determinants of economic growth. For example, the neoclassical growth model, introduced by Solow (1956), predicts that fiscal policies such as taxation and expenditures may have transitional effects on output level but not the long-run growth rate. In this class of models, the long-run growth rate is determined by exogenous factors such as technical progress and population growth. The endogenous growth models introduced by Barro (1990) and King and Rebelo (1990) have challenged the traditional neoclassical growth model and provide a mechanism through which taxes and public expenditures can determine both the level of output and the steady-state growth rate. According to endogenous growth theory, Helms (1985) emphasizes that in order to evaluate the true effect of tax or expenditure on growth both sides of government budget constraints (GBC) including the sources and uses of funds must be taken into account.

Partly in step with these theoretical issues, empirical studies have evolved over time. There are two distinct strands of literature among the studies investigating the role of taxation as a determinant of economic growth in U.S. states: (i) studies in which the complete specification of the government budget constraint (GBC) is taken into account, and (ii) studies ignoring the role of the government budget constraint. Further, it is important to recognize that the net effects of alternative policies may differ depending on the types of taxes/expenditures considered, what a government produces, and how the government output is financed. For example, if the revenue generated by distortionary taxes such as a personal income tax is used to fund productive expenditures such as infrastructure and/or education, then the expected net effect might differ from a situation in which same distortionary tax is used to fund unproductive expenditures such as welfare and/or recreation. Another important issue is to consider the duration of any tax effects. In other words, are the effects of taxes on economic growth being measured over short, medium, or long periods of time. For these and other reasons, even studies that use similar data can produce dissimilar estimates of tax effects.

To overcome the above-mentioned shortcomings, and also in an attempt to offer a clear picture of the extensive and dispersed research investigating the effect of state and local taxes on state economic growth, I conduct a meta-regression analysis (MRA) on existing studies. The main objectives of this study are to answer the following questions: (i) what is the overall, mean effect of taxes on economic growth in U.S. states?; (ii) is there any empirical evidence to support the public finance argument that "productive expenditures" financed by "distortionary taxes" are less growth retarding than same spending financed by "unproductive expenditures"?; (iii) are some taxes more growth retarding than others?; and (iv) what are the possible causes of heterogeneity in the results observed in the literature?

To achieve the stated objectives, this study collects estimates of tax effects on economic growth in U.S. states from 29 empirical studies. According to a final sample of 868 estimates, I find strong evidence that the empirical literature on estimated tax effects is influenced by negative publication bias by which I mean that negative estimates are over-reported in publicly available studies. Once I control for this bias, I calculate that the "overall effect" of taxes on economic growth is small and statistically insignificant. However, as mentioned earlier, this "overall tax effect" is not very informative. Once I turn to analysing different types of tax policies, and after controlling for publication bias, the evidence regarding the composition of fiscal policy is mixed.

The remainder of the chapter is organized as follows. Section 3.2 describes how I collected the sample of estimates. Section 3.3 discusses some of the reasons why studies of tax effects can produce different estimates. Section 3.4 presents my empirical results, addressing my objectives. Section 3.5 summarizes the main findings of this research.

3.2. Selection of Studies and Construction of Dataset

This meta-regression analysis collects estimated tax effects derived from all studies that regress a variation of the following specification:

$$g = \alpha_0 + \alpha_1 tr + error, \tag{3.1}$$

where g is a measure of economic growth, tr is a measure of the tax rate, and the data are taken from 50 U.S. states plus the District of Columbia. To identify these studies, I conducted a comprehensive search including both electronic and manual search procedures. It is worth noting that studies estimating non-linear transformation of tax effects (a squared term), such as the "growth hills" of Bania, Grey and Stone (2007) and also studies estimating interactive terms, such as Deskins and Hill (2010), are not included in this MRA. This is because if there is an interaction term in the model, the total effect is an outcome of both the term and its interaction. Unfortunately, the meta-analyst usually does not have access to the data necessary to calculate the associated marginal effect and its statistical significance.

The electronic search used three categories of keywords: (i) "STATE and LOCAL TAXES" keywords, (ii) "STATE ECONOMIC GROWTH" keywords, and (iii) "U.S. STATES" keywords in the following combination: "STATE and LOCAL TAXES" and "STATE ECONOMIC GROWTH" and "U.S. STATES". A variety of keywords were substituted for each of these three categories. All the possible alternatives are reported in Appendix 3.1. I then implemented the search on a variety of comprehensive electronic search engines such as: EconLit, Google Scholar, JSTOR, Web of Science, Scopus, RePEc, EBSCO, and ProQuest by searching various keyword combinations. A total of 459 studies were identified in this manner.

As mentioned earlier, numerous studies have examined the relationship between taxes and economic growth. However, a meaningful meta-regression analysis requires comparable original studies and therefore the inclusion/exclusion selection criteria are designed to fulfil this requirement. Thus, to be included in this data set each study must meet the following criteria. First, a growth equation must have a tax variable. Second, the regional focus is U.S. states, so each study must include at least 44 contiguous U.S. states. The reason being that even though the U.S. includes 50 states plus the District of Columbia, most studies exclude the non-contiguous states of Alaska and Hawaii, the District of Columbia, and sometimes other states as well, for various reasons. Finally, each included study must report sufficient statistical information to allow the calculation of the effect sizes such as regression coefficients, standard errors, and t-statistics or p-values.

The abstracts and conclusions of the 459 studies were then read carefully to eliminate any studies that did not meet these criteria. Many studies, including government reports and narrative reviews, were excluded from my dataset as a result. Backwards and forwards citation search strategies were then implemented to locate additional studies. This process eventually resulted in a list of 43 studies, some of which were multiple versions of the same study, and included journal articles, conference proceedings, reports released by government agencies, think thanks and research firms, thesis and dissertations, and finally working papers and other unpublished or grey literature.[19]

This list was then emailed to 56 researchers who had previously written at least one research paper on the topic of taxes and economic growth using U.S. data. The researchers were asked to assist in identifying any additional research papers of their own or of any master/PhD students who were working with them.[20] A revised list of 53 studies was then compiled based on the responses I received from the researchers.[21, 22]

Each study in the revised list was then read carefully and thoroughly to see whether it was eligible according to the defined inclusion criteria. The dependent variable had to be a measure of state income growth (usually GSP or Personal Income growth, but not

---

[19] When reported estimates differ in multiple versions of the study, the peer-reviewed journal articles is considered as a benchmark. However, if there are additional estimates in previous versions of the study, I kept track of the outlet of the study, coded, and then pooled the estimates across versions.
[20] The letter along with the bibliography of the core studies emailed to the prominent authors in this research area is available in Appendix 3.2 and Appendix 3.3.
[21] I am grateful for helpful suggestion received from all the scholars.
[22] The two plus coders includes myself, a PhD student recruited as a research assistant, and Prof. Reed to provide us the right direction once there is discrepencies in the reconciliation process.

employment growth). Alternatively, the dependent variable could be in level form, as long as the lagged dependent variable was included in the specification. The growth equation had to include at least one tax variable that was measured in units of percent of income. Studies in which the "tax variable" consists of all revenues, such as the ratio of total revenues to GSP, were not included because they lump together tax and non-tax revenues.

The states included in a given regression equation had to consist of at least 44 contiguous U.S. states. While some studies included all the 50 states plus District of Columbia, others excluded outlier states such as Alaska and Hawaii plus the District of Columbia. There are several reasons for this. Outlier states such as Alaska and Hawaii have limited labour mobility compared to the contiguous 48 states. Alaska is also an outlier because a large share of its tax revenues comes from severance taxes on oil. And the District of Columbia is not a state. Given the above-mentioned reasons one would expect the remaining studies to include 48 states. However, some studies dropped additional states, such as Wyoming (because it also receives a large portion of severance tax revenues from oil), Nebraska (because it has a unicameral legislature), or other states because of the absence of a sales and/or income tax. Setting the threshold at 44 contiguous states allowed me to include a larger number of studies. All estimated tax effects had to report a standard error or associated t-statistic. Finally, only studies written in English were included. I closed my research on October 2015. The final sample of 29 studies is listed in Appendix 3.4.[23]

Once the final set of estimates was determined, I then went through each equation/estimate and coded a set of regressions and study characteristics (more details are provided in the next section). The coding was done independently by at least two coders with

[23] Appendix 3.5 clarifies the steps which is undertaken to reach to the 42 final studies.

a careful reconciliation of any discrepancies or inconsistencies.[24] All search and coding procedures followed the MAER-NET protocols closely (Stanley et al., 2013).

3.3. Factors that Cause Tax Estimates to Differ Across Studies

The government budget constraint. As explained in the previous chapter, to estimate the precise effects of tax on economic growth it is important to deal adequately with a number of issues. The first issue has to do with the government budget constraint and the importance of implicit financing.

The regression coefficient on the tax rate can sometimes be misinterpreted once the importance of implicit financing and/or the role of the government budget constraint are ignored. The main argument is that the regression coefficient on tax variable should be interpreted as the growth effect of a tax financed by the omitted fiscal categories. For example, if public expenditure is omitted from the specification, the coefficient on tax variable measures the net effect of an increase in expenditure funded by taxes.

The precise interpretation is further complicated by the finer gradations of taxes (distortionary versus non-distortionary) and expenditures (productive versus unproductive expenditures). Thus, the net effect of tax on growth depends on the simultaneous change in taxes and/or expenditures. For example, if capital spending such as infrastructure is omitted from the regression equation, the coefficient on the consumption tax rate variable measures the net effect of an increase in productive expenditures financed by an increase in non-distortionary taxes. In which case, according to theory, a positive value for the tax coefficient would be expected (cf. Section 2.3 in the previous chapter for further details). As a result, similar "tax rate" variables might legitimately produce negative, positive or zero/ambiguous

[24] The two plus coders includes myself, a PhD student recruited as a research assistant, and Prof. Reed to provide us the right direction once there is discrepencies in the reconciliation process.

effects by virtue of the kind of tax variable that was being investigated as well as which other variables in the government budget constraint were omitted.

To address this issue, I go through each estimated tax effect and identify both the operative tax types and the use of the tax revenues implied by the government budget constraint. Tax types and expenditures are then categorized as distortionary/non-distortionary, productive/unproductive, or other according to the taxonomy represented in Table 3.1, taken from Kneller, Bleaney, and Gemmell (1999).[25]

[25] I use the Kneller, Bleaney, and Gemmell (1999) taxonomy because it is broadly representative of the fiscal policy literature. Distortionary taxes are those distorting the private sector's incentive to invest such as taxes on income and property. An example of non-distortionary tax would be taxes on consumption.

Table 3.1: Matching of Functional and Theoretical Classifications

Functional classification	***Theoretical classification***
Taxation on income and profit Social security contributions Taxation on payroll and manpower Taxation on property	Distortionary taxation
Taxation on domestic goods and services	Non-distortionary taxations
Non-tax revenues Other tax revenues	Other revenues
General public services expenditure Public safety expenditure Educational expenditure Health expenditure Housing expenditure Transport and communication expenditure	Productive expenditures
Social security and welfare expenditure Expenditure on recreation Expenditure on economic services	Unproductive expenditures
Other expenditures (unclassified)	Other expenditures

Note: The categorizations in the table are taken from Kneller, Bleaney, and Gemmell (1999) with some accommodation for the fact that the categories refer to revenues and expenditures for U.S. states.

Table 3.2 summarizes the predicted effect of distortionary/non-distortionary taxes on economic growth given the omitted fiscal category. This is taken from Gemmell, Kneller and Sanz (2009) but has been adapted to the various cases available in my sample. Doing so, every tax effect is assigned a prediction with respect to its impact on growth (negative, positive, or ambiguous/zero), where I merge the original categories of “zero” and “ambiguous” to “ambiguous”.

Table 3.2: Predicted Tax Effects

Type of Tax	***Omitted Fiscal Category***	***Predicted Effect***
Distortionary	Productive Expenditures	Ambiguous
Distortionary	Unproductive expenditures	Negative
Distortionary	Productive & Unproductive Expenditures	Ambiguous
Distortionary	Other Expenditures	Ambiguous
Distortionary	Deficit/Surplus	Ambiguous
Distortionary	Other Revenue	Ambiguous
Distortionary	Distortionary Taxes	Ambiguous
Distortionary	Non-distortionary Taxes	Negative
Distortionary	Intergovernmental Revenue	Ambiguous
Distortionary	Net Utility Expenditures	Ambiguous
Non-distortionary	Productive Expenditures	Positive
Non-distortionary	Unproductive Expenditures	Ambiguous
Non-distortionary	Productive & Unproductive Expenditures	Ambiguous
Non-distortionary	Other Expenditures	Ambiguous
Non-distortionary	Deficit/Surplus	Positive
Non-distortionary	Other Revenue	Ambiguous
Non-distortionary	Distortionary Taxes	Positive
Non-distortionary	Non-distortionary Taxes	Ambiguous
Non-distortionary	Intergovernmental Revenue	Ambiguous
Non-distortionary	Net Utility Expenditures	Ambiguous

Note: The categorizations in the table are taken from Gemmell, Kneller, and Sanz (2009), where I combine the original categories of "zero" and "ambiguous" to "ambiguous".

I also classify each estimated tax effect according to its type. Taxes are classified as Labour taxes, Capital taxes, Consumption taxes, Property taxes, Mixed taxes, Other taxes, and Overall taxes. The classification system for assigning each tax to a tax type is given in Table 3.3.[26]

[26] Mixed taxes are a combination of various types of taxes but not all.

Table 3.3: Types of Taxes

Tax Type	***Examples***
Labour	Personal income tax Payroll tax Social security contributions
Capital	Corporate income tax Capital tax (tax on dividends)
Consumption	Consumption tax Taxes on goods and services Sales tax Value added tax (VAT)
Property	Property tax
Other tax	Taxes not listed above
Mixed tax	Taxes that are a combination of the above types
Overall tax	Total taxes (e.g., Total Tax Revenues/GSP)

Units of measurement. The second issue refers to the units of measurement for economic growth and the measure of the tax rate variables in Equation (3.1). Each of these variables can be measured in percentage points (e.g., 2%) or in decimals (e.g., 0.02). This will obviously affect the size of the coefficients, α_1. For example, if a one-percentage point increase in the tax rate lowers growth by 0.1%, and if both g and tr are measured in percentage points, or both are measured in decimals, then the corresponding value of α_1 will be -0.1. However, if g is measured in percentage points, and tr is measured in decimals, then the corresponding value of α_1 will be -10. And if g is measured in decimals, and tr is measured in percentage points, then the value of α_1 will be -0.001. Accordingly, I adjust all estimated effects so that $\alpha_1 = X$ means that a one-percentage point increase in the tax rate is associated with an X percentage point increase in economic growth.[27]

[27] Sometimes it was difficult to determine the units of measurement of the respective variables from the study so as to properly interpret the coefficient. When this would happen, we could contact the original author(s). When

States. The third issue relates to the specific states excluded from a given study. While there are studies in which all 50 U.S. states plus the District of Columbia are included, the convention established in the literature is to exclude Alaska and Hawaii as their economies are thought to behave differently from the continental states. Their isolation noticeably reduces the mobility of their labour force. In addition, the Alaskan economy has been strongly affected by construction of a pipeline in 1977, and therefore a substantial portion of tax revenue is received in the form of severance taxes as a result of oil exporting capabilities. The District of Columbia is also excluded from most studies since it is not a state. Many studies only exclude one of the above-mentioned states, although there are studies that exclude a subset of these states. I further categorize the excluded states by the groupings None (once all 50 states plus DC are included), AK, DC, AKDC, AKHIDC, AKHIDCOTHERS, with the idea that by leaving the outlier states aside the remaining sample consists of more homogeneous economies. My meta-regression analysis controls for these different groupings in order to identify whether the estimated tax effects vary systematically across the different sets of states excluded from the original studies.

Duration of time periods. A fourth issue concerns the time frames of the data employed in the different studies. If the time periods of Equation (3.1) differ across studies, that could cause estimates of α_1 to differ, even when the underlying effect is the same. For example, suppose there were two growth studies, but one used 5-year time periods while the other used annual data. Suppose the former measured the cumulative rate of growth over each five-year period while the latter reported annual growth rates. Ceteris paribus, one might expect α_1 to be larger in the former case. Accordingly, I adjust all growth measures to be (average) annual rates of growth.

---

there was substantial uncertainty about the interpretation of the coefficients, the estimate was dropped from my analysis.

Duration of estimated tax effects. Since there is a consensus amongst growth models that tax-growth effects occur in the short-run, distinction between short-, medium-, and long-run may partly explain discrepancies observed in the reported estimates. Therefore, the fifth issue is related to the duration of the estimated tax effect as implied by the specification of the regression equation.

In the finite distributed lag (FDL) model, the coefficient on the current tax rate as well as the lagged tax rate represent the “short-run/immediate” effects of a one-percentage point increase in taxes in the current year and the previous year, respectively, on the current economic growth. Modifications of the regression specification may alter the effects from “short-run/immediate” to “cumulative/intermediate” or “long-run/permanent”.

My meta-regression analysis controls for this by noting the specification of the growth equation in the original study and categorizing the duration of the estimated tax effect as short-run, medium-run, or long-run.

Different measures for economic growth and tax rates. A final issue has to do with how the economic growth and tax rate variables are defined in the primary studies. While some studies measure economic growth in terms of nominal GSP/Personal Income, others apply real GSP/Personal Income. Other measures available in the literature are per capita GSP/Personal Income and total GSP/Personal Income. It is worth nothing that I don’t distinguish between real and nominal growth as long as the economic growth measure used in the primary studies is a log transformation of nominal GSP/Personal Income and time dummies are included in the specification.

When it comes to measuring “the tax rate”, the main question is how to accurately measure tax rates. Unfortunately, economic theory provides no clear answer to this question. As a result various measures can be seen in the literature. Most studies use effective tax rates, defined as the ratio of tax revenues over a given measure of income. However, Engen and Skinner (1992) and Easterly and Rebelo (1993) show that average tax rates are strongly correlated with public spending. Others use statutory tax rates typically the top marginal rate. And some studies attempt to distinguish marginal from average tax rates. Marginal tax rates are defined as the additional taxes paid when personal income rises by a small amount. For example, for a personal income tax, the marginal tax rate describes a person's tax bracket and shows how much taxes are paid on the last dollar earned from working and investing. Since economic decisions depend on the marginal tax rate, this measure is more appropriate for investigating the effect of taxes on growth. However, marginal tax rates are not easily observable. I use dummy variables to indicate the specific measures underlying a given estimate.

Control variables. In addition to all the challenging issues discussed earlier, I code many other study characteristics, including estimation methods; types of standard errors; whether the original study was published in a peer-reviewed journal; publication year; sample period length; midpoint of the sample; and inclusion of various specific variables such as state fixed effects, human capital, capital investment, employment/population growth, and others. A full list of the variables used in this study is discussed in the following section.

3.4. Empirical Analysis

Preliminary analysis. As mentioned earlier, my literature search produced a dataset of 29 studies containing a total of 966 estimated tax effects. Table 3.4 reports the descriptive statistics for both the estimated tax effects and their associated t-statistics. The median estimated tax effect is -0.055 for the full dataset, implying that a ten percentage point increase in the tax rate is associated with a 0.55 percentage point decrease in annual economic growth. This compares to an average, annual PCPI growth for the 48 contiguous U.S. states of approximately 6.07 per cent over the period 1975-2003.[28]This period roughly corresponds to the "average" sample period of the studies included in this analysis. [29,30] The median t-statistic is -0.67.

Table 3.4 immediately identifies a problem in that the minimum and maximum estimated effects are -7.21 and 9.58.[31] These values indicate a tax effect size that is considerably outside the bounds of reasonable. While researchers differ in their estimates of the effects of taxes, nobody suggests that a one percentage point increase in the tax rate would lower annual economic growth by over 7 percentage points, or increase it by over 9 percentage points. Accordingly, the subsequent analysis works with a truncated sample of estimates.

I delete the top and bottom 5 percent of estimates and obtain a sample of 868 estimated tax effects. The descriptive statistics for this truncated sample are also reported in Table 3.4. The range of estimated tax effects for this sample is from a minimum -1.47 to a maximum of

[28] Note that average, annual GSP growth was not available over the above mentioned period.
[29] This is calculated by taking the average beginning and average ending dates for the sample ranges of the respective studies.
[30] Growth rate is the average, annual PCPI growth rate over the period 1975-2003 for the 48 states. Alaska, Hawaii and the District of Columbia are excluded. The reason why I report annual PCPI growth rather than annual GSP growth is due to the gap in data availability (data for GSP is not available untill 1977).
[31] Excel spreadsheet that allows the user to replicate all the results of Table 3.4 through 3.10 can be downloaded from Dataverse: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/IW8UEY

0.639. The median t-statistic still indicates insignificance, while the samples of t-statistics range from a minimum of -17.29 to a maximum of 8.77, with a mean absolute value of 1.87.

Table 3.4: Descriptive Statistics for Estimated Effects and t-statistics

	Estimated Tax Effects		***t-statistics***	
	Full	***Truncated***	***Full***	***Truncated***
Mean	-0.140	-0.121	2.12*	1.87*
Median	-0.055	-0.055	-0.67	-0.71
Minimum	-7.210	-1.47	-18.54	-17.29
Maximum	9.581	0.639	8.77	8.77
Std. Dev.	1.023	0.349	3.00	2.55
1%	-3.47	-1.33	-13.23	-10.76
5%	-1.50	-0.83	-5.8	-4.87
10%	-0.797	-0.535	-3.63	-3.45
90%	0.358	0.262	2.13	1.78
95%	0.66	0.37	1.67	2.60
99%	2.841	0.587	4.95	3.59
Obs	966	868	966	868

Note: The mean absolute value of t-statistics is indicated by an asterisk.

Figure 3.1 plots the estimated tax of the truncated sample. If tax effects were homogeneous across studies and differed solely due to sampling error, we would expect a bell-shaped histogram. This is clearly not the case in Figure 3.1. The distribution is skewed to the left, and suggests that there may be sample selection favouring negative estimates, perhaps due to publication bias. I test for this possible publication bias below.

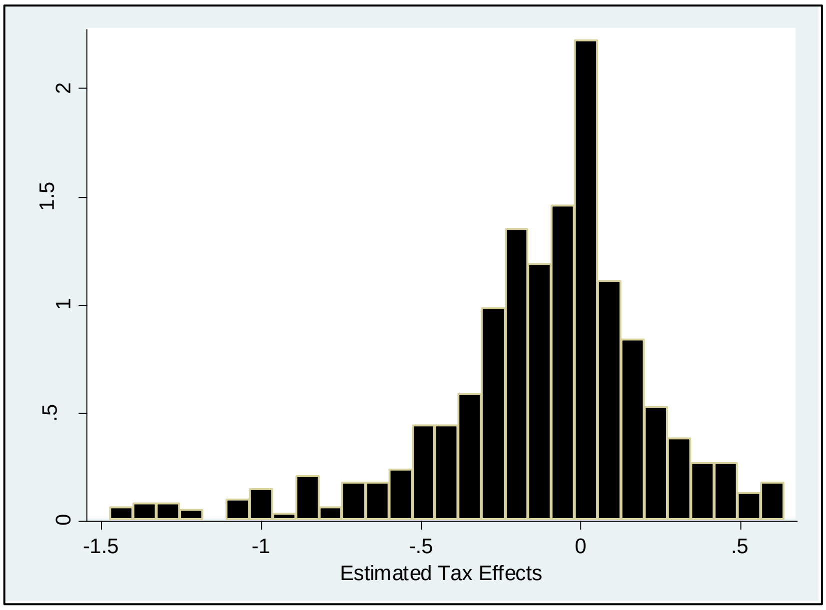


Figure 3.1: Histogram of Estimated Tax Effects (Truncated)

Fixed Effects versus Random Effects. In order to achieve the first goal of this study of estimating the "overall tax effect", an appropriate weight must be assigned to each study/estimate. The reason for this is that the studies/estimates are not equally precise. Therefore, it is reasonable to give more weight to the more precise studies/estimates. There are two models used in the MRA literature, "fixed effects" and "random effects". It is worth mentioning that "fixed effects" and "random effects" mean something entirely different in meta-analysis than they do in panel data econometrics (Reed, 2015).

Fixed Effects. The underlying assumption for the fixed effects model is that there is one true effect size and the main sources of effect size variation is due to sampling error. Accordingly, the weight assigned to each estimate is the inverse of its associated standard error.

Random Effects. For the random effects model, on the other hand, we assume that there is a distribution of "true effects". In this case, which is more realistic than the previous one, the differences in the estimated effects across studies are assumed to be due to a combination of (i) sampling error, and (ii) genuine differences in the underlying effects. The corresponding weight in this case is the inverse of these two components.

Figure 3.2 presents a forest plot of the respective studies using a “fixed effects” weighting scheme. In this figure, the estimated tax effects are weighted by the inverse of their estiamted standard errors. A weighted average is constructed for each study, along with a 95 percent confidence interval. Several features of the forest plot are noteworthy. First, most of the studies estimate small effects with tight confidence intervals, though study 5 (Chernick, 2010) is a notable counterexample.

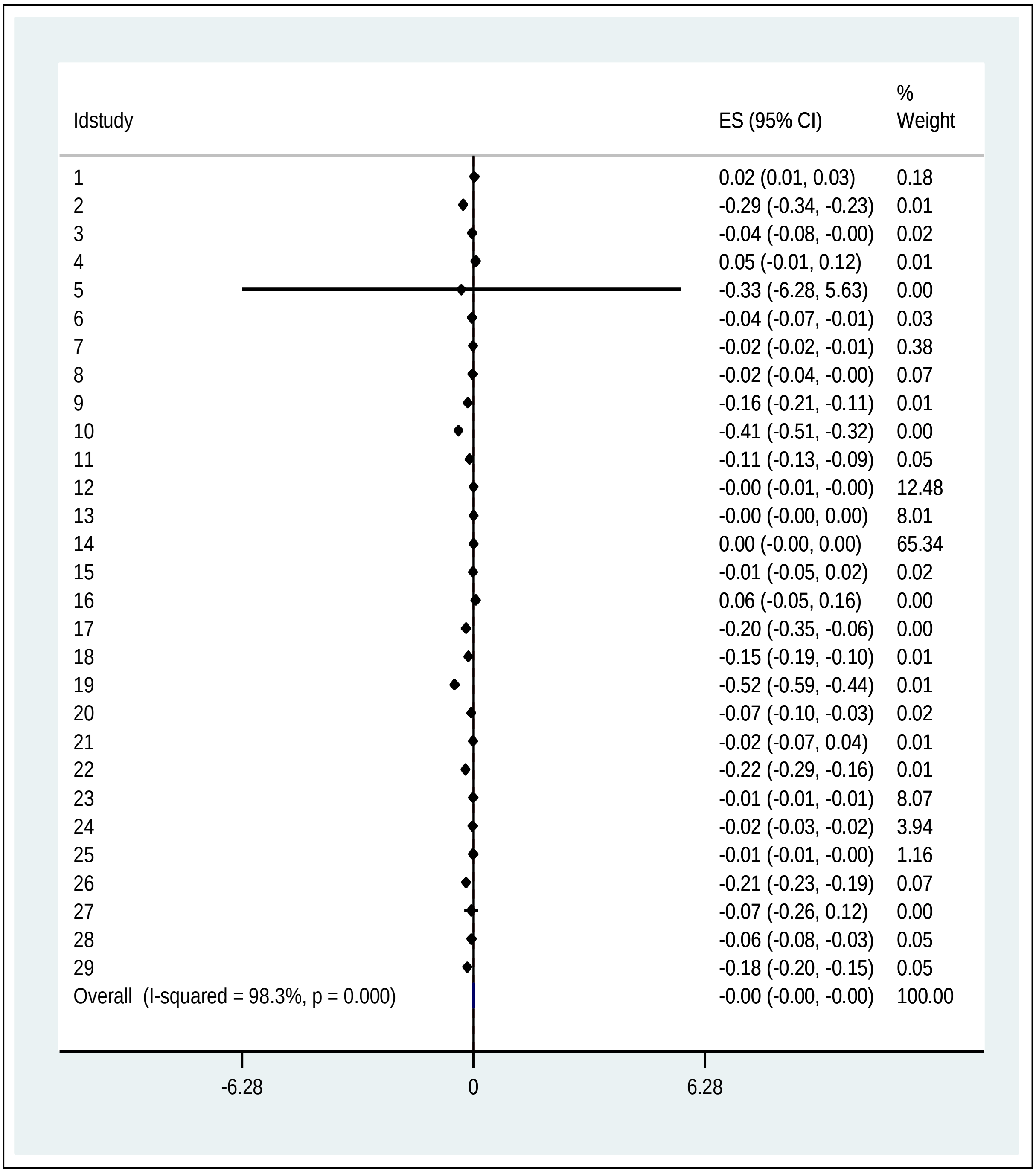


Figure 3.2: Forrest Plot of Studies (Fixed Effects)

Second, there is a substantial amount of cross-study heterogeneity, indicated by an exceptionally large I^2 value of 98.3% (Higgins and Thompson, 2002). As discussed above, different studies have different ways of incorporating the government budget constraint, of measuring tax effects of different durations, study different samples of states, and so on. The high I^2 value indicates that the differences across studies overwhelm the variation that would be expected solely from sampling error. Finally, the last column shows the percentage weight each study receives in calculating the overall weighted average. Study 14 (Reed, 2008) is weighted substantially more than all the other studies combined (65.34% versus 34.66%). The disproportionately large weight given to one study is not necessarily a concern if that one study is truly, substantially more reliable than the others. However, it may be prudent to use a more dispersed weighting scheme.

Accordingly, the subsequent empirical work emphasizes the "random effects" estimates, where tax effects are weighted by their standard error plus a term that captures the cross-study heterogeneity. Because cross-study heterogeneity is so great, this will have the effect of equalizing the weights given to individual studies. Figure 3.3 displays the forest plot using random effects. The study weights are much more balanced. The Reed study now receives a weight of 7.11 percent, substantially less than the 65.34 percent using the fixed effects.

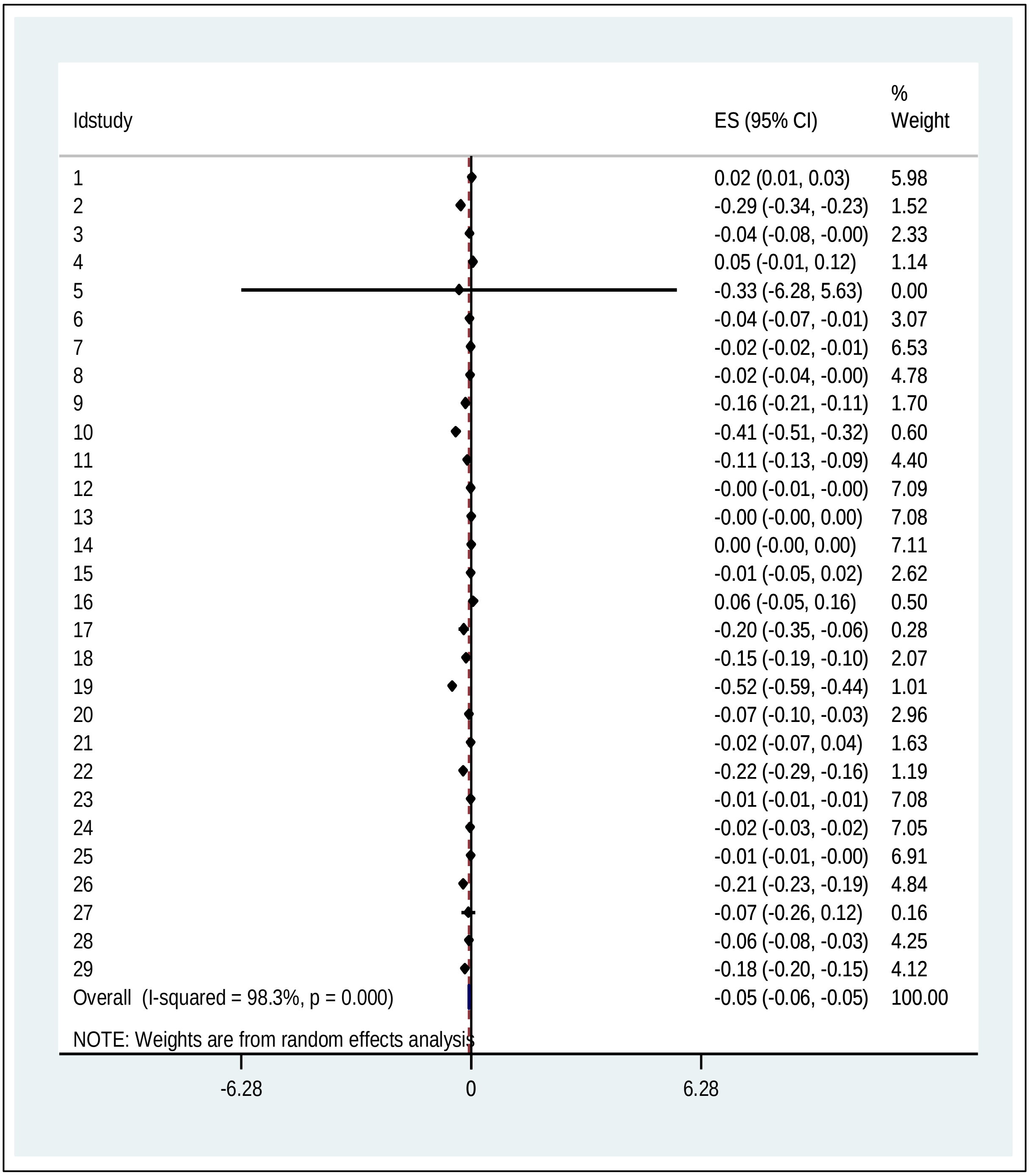


Figure 3.3: Forrest Plot of Studies (Random Effects)

Figure 3.4 and Figure 3.5 reports two funnel plots, with estimates plotted against their standard errors. The top figure displays individual estimates. In the bottom figure, each study is represented by a single point relating its mean estimate to its mean standard error. [32] The solid line in both plots shows the overall mean of estimated tax effects. The dotted lines that fan out from the top of the funnel demarcate the 95% confidence area where most of the estimates would fall if the dispersion in estimates was driven solely by sampling error. Publication bias is indicated whenever a disproportionate number of estimates lie on one side of the inverted, V-shaped confidence area. Both funnel plots indicate publication bias, with a preference for negative estimates over positive ones. Further, the wide dispersion at the top of the funnel is consistent with substantial heterogeneity, as previously indicated by the I^2 value.

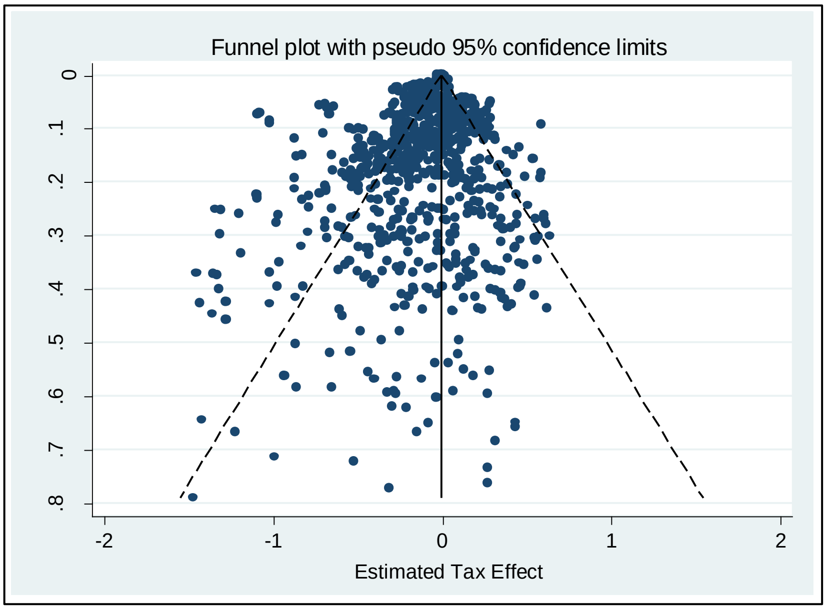


Figure 3.4: Funnel Plot (Individual Estimates)

[32] Both funnel plots omit observations where the standard error is greater than 1. This allows the reader to better observe the pattern of points at the top of the funnel.

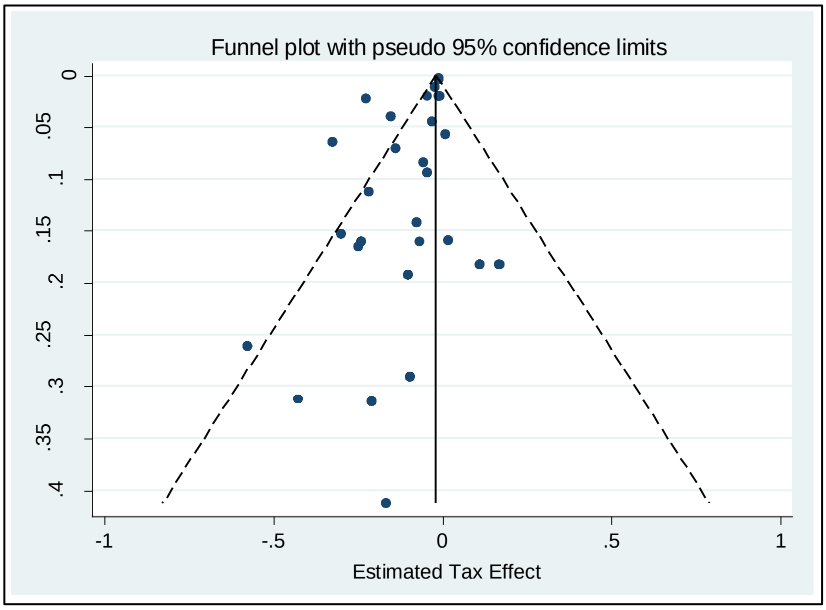


Figure 3.5: Funnel Plot (Mean of Study Estimates)

FAT/PET tests. Table 3.5 reports the results of two tests: the Funnel Asymmetry Test (FAT) to test for publication bias, and the Precision Effect Test (PET), which tests for the significance of the overall effect.[33] Both tests are obtained from estimating the following specification using Weighted Least Squares (WLS),

$$\hat{\alpha}_{1,ij} = \beta_0 + \beta_1 SE_{ij}, \tag{3.2}$$

where $\hat{\alpha}_{1,ij}$ is the estimated tax effect from regression *j* in study *i*. The null hypotheses for the FAT and PET are $H_0\colon \beta_1 = 0$ and $H_0\colon \beta_0 = 0$, respectively.

This analysis uses four different weights to estimate Equation (3.2). The *Fixed Effects (Weight 1)* and *Random Effects (Weight 1)* estimators use weights $\left(\frac{1}{SE_{ij}}\right)$ and $\left(\frac{1}{\sqrt{(SE_{ij})^2+\tau^2}}\right)$, respectively, where τ^2 is the estimated variance of the population tax effect across studies. This set of weights makes no allowance for the fact that some studies report more estimates than others. As a result, a study with 50 estimates is weighted 50 times more than a study that

[33] Detailed discussions of these tests are provided in Stanley and Doucouliagos (2012) and Shemilt et al. (2011).

reports a single estimate, ceteris paribus. To address this, we multiply both sets of weights by $\left(\frac{1}{N_i}\right)$, where N_i is the number of estimated tax effects reported in study i. The corresponding *Fixed Effects (Weight 2)* and *Random Effects (Weight 2)* estimators attempt to give equal weight to each study regardless of the number of tax effects each study reports.

Table 3.5: Funnel Asymmetry and Precision Effect Tests (FAT/PET)

	Fixed Effects (Weight1) (1)	***Fixed Effects (Weight2) (2)***	***Random Effects (Weight1) (3)***	***Random Effects (Weight2) (4)***	***Random Effects (Weight1) (5)***	***Random Effects (Weight2) (6)***
(1) FAT	-0.915** *(-2.72)*	-1.425*** *(-3.05)*	-0.688*** *(-2.97)*	-1.081*** *(-3.39)*	---	---
(2) PET	-0.0002 *(-0.12)*	-0.004 *(-0.89)*	-0.009 *(-0.65)*	-0.007 *(-0.40)*	-0.038** *(-2.61)*	-0.041** *(-2.52)*
Observations	868	868	868	868	868	868

Note: Values in Row (1) and Row (2) come from estimating β_1 and β_0, respectively, in Equation (3.2) in the text. In both cases, the top value is the coefficient estimate, and the bottom value in parentheses is the associated *t*-statistic. The four WLS estimators (*Fixed Effects-Weight1, Fixed Effects-Weight2, Random Effects-Weight1*, and *Random Effects-Weight2*) are described in the "FAT/PET tests" subsection of Section 3.4 in the text. All four of the estimation procedures calculate cluster robust standard errors. *, **, and *** indicate statistical significance at the 10-, 5-, and 1-percent level, respectively.

The first four columns of Table 3.5 report the results of estimating Equation (3.2) with WLS, using the four different weighting schemes described above. The FAT is reported in the first row. Apart from the first column in which the null hypothesis of no publication bias is rejected at the 5 percent level of significance, all the three remaining estimators reject the null at the 1 percent level of significance. The negative coefficients indicate that sample selection favours negative estimated tax effects, perhaps due to researchers choosing to disproportionately report negative estimates, or journals discriminating against positive results. These results are consistent with earlier observations about the histogram of estimated effects and the visual evidence of publication bias from the funnel plots represented in Figure 3.4 and Figure 3.5.

The first four columns of the second row of Table 3.5 report the PET estimates. All four estimators support the conclusion that the overall tax effect, controlling for publication bias, is statistically insignificant and relatively small in economic terms. For example, the Random Effects (Weight 1) estimate indicates that a ten percentage point increase in the tax rate is associated with a 0.09 percentage point decrease in annual GSP/Personal Income growth. The last two columns report random effects estimates of Equation (3.2) when the publication bias term $\left(\mathrm{SE}_{ij}\right)$ is not included, so that the overall estimate is not corrected for publication bias. The corresponding estimates of the overall tax effects are now substantially larger in absolute value and statistically significant at the 5 percent level. According to the Random Effects (Weight 1) estimate in Column (5), a ten percentage point increase in the tax rate is associated with a 0.38 percentage point decrease in annual GSP growth. These results indicate that the statistically and economically significant results reported in the literature are a consequence of publication bias that favours negative estimates of tax effects, while suppressing the publication of positive tax effects. As a result, I want to be sure that my subsequent analysis corrects for this.

This section addresses the first goal of this research to obtain an "overall estimate" of the effect of taxes on economic growth in US States. We find that a publication bias adjusted estimate of the overall effect on tax is statistically insignificant and negligibly small in economic terms. However, my previous discussion on factors that cause tax estimates to differ across studies (cf. Section 3.3) makes clear that any estimate of overall tax effects is not particularly meaningful. The same fiscal policy action can be estimated as a positive or negative tax effect depending on the elements of the government budget constraint that are omitted from the original study's regression equation. Accordingly, the next section undertakes a meta-regression that allows tax effects to vary systematically according to study and data characteristics.

Meta-regression. In this section, we compare tax effects associated with fiscal policies that are predicted to have negative growth effects, with those predicted to have positive effects. We also investigate whether some types of taxes are more growth-retarding than others. To do that, it will be necessary to control for the myriad factors that affect estimates of tax effects.

Table 3.6 reports the variables used in the subsequent meta-regression analysis. The first sets of variables were previously discussed and match each tax effect to a prediction. A little more than an eighth of the estimated tax effects allow a definite sign prediction, with 9.9 percent predicted to be negative, 3.3 percent predicted to be positive, and the rest ambiguous. As these three variables comprise the full set of possibilities, at least one variable must be omitted in the empirical analysis. Here and elsewhere we indicate the omitted variable with an asterisk.

The second set of variables assign each tax effect to one of seven types of taxes (Labour, Capital, Consumption, Property, Other, Mixed, and Overall). The most common tax variable is constructed by taking the ratio of total tax revenues over GSP. Approximately 32.9 percent of tax effects are of this type. However, many studies disaggregate tax effects into separate types. For example, 31.2 percent of estimated tax effects involve Labour taxes (e.g., personal income taxes, payroll taxes, social security contributions). Another 11.1 percent are associated with Capital taxes (e.g., corporate income taxes, occupational and business licencing tax, taxes on capital gains and dividends), 9.1 percent are related to Consumption taxes (e.g., sales tax, ad valorem taxes on goods and services, VAT), and 9.2 percent are associated with Property taxes (e.g., taxes on property, taxes on immovable property, and land taxes). The remainder of tax effects mostly involve either a mix of different tax types or other tax types.

Subsequent variables are grouped into numerous categories: State Group, Economic Growth Measure, Tax Variable Measure, Duration of Tax Effect, etc. Most of the observed tax effects are estimated using data from a set of 48 contiguous states (50.1%) where Alaska, Hawaii and the District of Columbia are excluded from the sample. This is in contrast to studies that included all the 50 states plus the District of Columbia (7%), or the ones that excluded Alaska, Hawaii, and District of Columbia as well as other states (11.6%). The remaining portion constitutes those excluding either Alaska (18.7%) or the District of Columbia (10.3%) or both (2.3%). In most cases, economic growth is measured in per capita personal income terms (63%), as opposed to GSP (20%), per capita GSP (10.8%), and personal income (6.1%). Most taxes are measured as average rates, rather than marginal (74.8% versus 25.2%); are specified in level rather than differenced form (80% versus 20%); and are effective rather than statutory tax rates (70.2% versus 29.8%). Most estimated tax effects measure the immediate effect of a tax change (86%), versus a medium- or long-run effect (11.5% and 2.5%, respectively).

A little more than a quarter of the estimated tax effects in my meta-regression come from peer-reviewed journal articles with the mean year of publication being 2010.[34] Most of the original studies used panel data to estimate tax effects (91%) as opposed to cross-section data (9%). The average sample length in the original studies was 28.6 years and the average mid-point year was 1989. About half of the tax effects were estimated using GLS. Of the remainder, 35.1 percent used OLS or a related procedure, and 9.8 percent attempted to correct

[34] If the estimated tax effects come from a working paper then the publication year indicates the year in which the latest version of the paper is publicly available.

for endogeneity using a procedure such as TSLS or GMM. Because the standard error plays such a significant role in meta-regression analysis, we categorize standard errors into three groupings: *SE-OLS* (47.9%); *SE-HET* (18.8%), where standard errors were estimated using a heteroskedastic-robust estimator; and *SE-Other* (33.3%), whenever allowance was made for off-diagonal terms in the error variance-covariance matrix to be nonzero. Lastly, dummy variables were used to indicate the presence of important control variables, the most common of which were state fixed effects (77.4%) and measures of initial income (45.9%). A fifth of the estimated tax effects control for at least two of three measures of investment (e.g., capital investment, private capital, spending), employment growth (e.g., unemployment rate, population growth, wage), and human capital (e.g., educational achievement or education as a proxy for human capital).

Table 3.6: Summary Statistics of Study Characteristics

Variable	*Description*	*Mean*	*Min*	*Max*
	PREDICTED TAX EFFECTS			
Prediction-Negative	=1, if the theoretical prediction of the coefficient is negative	0.099	0	1
*Prediction-Ambiguous**	=1, if the theoretical prediction of the coefficient is ambiguous	0.868	0	1
Prediction-Positive	=1, if the theoretical prediction of the coefficient is positive	0.033	0	1
	TAX TYPE			
Labour-Tax	=1, if labour tax	0.312	0	1
Capital-Tax	=1, if capital tax	0.111	0	1
*Consumption-Tax**	=1, if consumption tax	0.091	0	1
Property-Tax	=1, if property tax	0.092	0	1
Other-Tax	=1, if other type of tax	0.046	0	1
Mixed Tax	=1, if multiple tax types (but not overall tax)	0.018	0	1
Overall-Tax	=1, if overall tax	0.329	0	1
	STATE GROUP			
None	=1, if all states and DoC are included	0.070	0	1
AK	=1, if Alaska is excluded	0.187	0	1
DC	=1, if District of Columbia is excluded	0.103	0	1
AK, DC	=1, if both Alaska and District of Columbia are excluded	0.023	0	1
*AK, HI, DC**	=1, if states including (AK, HI, and DC) are excluded	0.501	0	1
AKHIDCOthers	=1, if states including (AK, HI, DC, and Others) are excluded	0.116	0	1
	ECONOMIC GROWTH MEASURE			
GSP	=1, if dependent variable is GSP growth	0.200	0	1
PC-GSP	=1, if dependent variable is per capita GSP growth	0.108	0	1
PI	=1, if dependent variable is PI growth	0.061	0	1
*PC-PI**	=1, if dependent variable is per capita PI growth	0.630	0	1

Variable	***Description***	***Mean***	***Min***	***Max***
	TAX VARIABLE MEASURE			
Marginal	=1, if marginal tax rate (as opposed to average tax rate)	0.252	0	1
Differenced	=1, if change in tax rate (as opposed to level of tax rate)	0.200	0	1
ETR	=1, if effective tax rate (as opposed to statutory tax rate)	0.702	0	1
	DURATION OF TAX EFFECT			
*Short-run**	=1, if tax variable measures immediate/short-run effect	0.860	0	1
Medium-run	=1, if tax variable measures cumulative/medium-run effect	0.115	0	1
Long-run	=1, if tax variable measures long-run, steady-state effect	0.025	0	0
	STUDY TYPE			
Peer-reviewed	=1, if study published in peer-reviewed journal	0.274	0	1
Publication Year	Year in which the last version of study was "published."	2010	1985	2015
	DATA TYPE			
Cross-section	=1, if data are cross-sectional.	0.092	0	1
*Panel**	=1, if data are panel.	0.908	0	1
Length	Length of sample time period.	28.6	5	65
Mid-Year	Midpoint of the sample time period.	1989	1921	2003.5
	ESTIMATION TYPE			
*OLS**	=1, if OLS estimator is used.	0.351	0	1
GLS	=1, if Generalized Least Squares estimator is used.	0.551	0	1
TSLS/GMM	=1, if estimator corrects for endogeneity, e.g. 2SLS, 3SLS, or GMM.	0.098	0	1
	STANDARD ERROR TYPE			
*SE-OLS**	=1, if OLS standard error is considered.	0.479	0	1
SE-HET	=1, if heteroskedasticity standard error is considered.	0.188	0	1
SE-Other	=1, if both heteroskedasticity and autocorrelation standard error are considered.	0.333	0	1

Variable	***Description***	***Mean***	***Min***	***Max***
	INCLUDED VARIABLES			
Initial income	=1, if initial level of income included.	0.459	0	1
Lagged DV	=1, if lagged dependent variable included.	0.234	0	1
StateFE	=1, if the state fixed effects are included.	0.774	0	1
HLK	=1, if at least two out of three control variables (H, L, K) are included.	0.211	0	1

Note: The grouped variables include all possible categories, where the categories omitted in the subsequent analysis are indicated by an asterisk, where applicable.

In addressing the twin questions regarding the predictions of growth theory and differences in tax types on growth, we adopt the following empirical procedure. First we separate out the two sets of tax variables: *Prediction-Negative* and *Prediction-Positive* for addressing the first of the two research questions; and *Labour-Tax*, *Capital-Tax*, *Property-Tax*, *Other-Tax*, *Mixed-Tax* and *Overall-Tax* to address the second question. We do this because the two sets of tax variables are significantly correlated. For example, Labour and Capital taxes are significantly associated with tax policies that are predicted to have negative effects. We then combine the two sets of tax variables to check for robustness.

For each set of regressions, I also include two sets of control variables. The top panel of each table reports the regression results when all control variables are included in the equation. The bottom panel reports the regression results when a backwards stepwise procedure is used to select control variables, even while the tax variables are fixed to remain in each equation.[35] The use of the stepwise procedure does not invalidate their significance testing, since the tax variables are locked into each regression. All regressions also include the publication bias variable, SE, and thus control for publication bias.

The results of this analysis are given in Table 3.7 through Table 3.9. Table 3.7 reports the results when the prediction variables (Prediction-Negative and Prediction-Positive) are included in the meta-regression, while holding out the tax type variables. Prediction-Negative is negative but statistically insignificant in the two fixed effects regressions for both sets of control variables (Column 1 and 2, top and bottom panel). However, despite the expectation from growth theory, it is positive and significant and positive and insignificant in the following two random effects regressions (Column 3 and 4). The results are not supportive at

[35] I use a backwards stepwise regression procedure that selects variables so as to minimize the Schwarz Information Criterion. I employed the user-written, Stata program *vselect* to implement the stepwise procedure.

all when it comes to the Prediction-Positive case. Almost in all cases (except in column 2, bottom panel), we estimate a negative and significant coefficient for this variable.

Table 3.7: Meta-Regression Analysis (Omitting Tax Type Variables)

Variable	***Fixed Effects (Weight1) (1)***	***Fixed Effects (Weight2) (2)***	***Random Effects (Weight1) (3)***	***Random Effects (Weight2) (4)***
All Control Variables Included				
SE	-0.984*** *(-3.26)*	-0.950** *(-2.42)*	-0.677** *(-2.08)*	-0.649** *(-2.56)*
Prediction-Negative	-0.015 *(-1.06)*	-0.016 *(-1.03)*	0.135*** *(3.72)*	0.067 *(1.59)*
Prediction-Positive	-0.285*** *(-6.50)*	-0.356*** *(-2.83)*	-0.328*** *(-5.15)*	-0.464** *(-2.47)*
Control Variables Selected Via Backwards Stepwise Regression				
SE	-0.820*** *(-3.84)*	-0.948*** *(-3.45)*	-0.791*** *(-2.87)*	-0.680*** *(-3.22)*
Prediction-Negative	-0.001 *(-0.27)*	-0.017 *(-1.17)*	0.078** *(2.56)*	0.061 *(1.46)*
Prediction-Positive	-0.283*** *(-6.57)*	-0.180 *(-1.32)*	-0.339*** *(-4.69)*	-0.470*** *(-4.43)*

Note: The top panel reports the results of estimating Equation (3.2) with the addition of the two tax variables, *Prediction-Negative* and *Prediction-Positive.* The bottom panel adds control variables selected through a backwards stepwise regression procedure that selects variables so as to minimize the Schwarz Information Criterion (see Footnote #25). The top value in each cell is the coefficient estimate, and the bottom value in parentheses is the associated *t*-statistic. The four WLS estimators (*Fixed Effects-Weight1*, *Fixed Effects-Weight2*, *Random Effects-Weight1*, and *Random Effects-Weight2*) are described in the "FAT/PET tests" subsection of Section 3.4 in the text. All four estimation procedures calculate cluster robust standard errors. *, **, and *** indicate statistical significance at the 10-, 5-, and 1-percent level, respectively.

Table 3.8: Meta-Regression Analysis (Omitting Prediction Variables)

Variable	***Fixed Effects (Weight1) (1)***	***Fixed Effects (Weight2) (2)***	***Random Effects (Weight1) (3)***	***Random Effects (Weight2) (4)***
All Control Variables Included				
SE	-0.961*** *(-2.98)*	-0.968** *(-2.35)*	-1.010*** *(-3.45)*	-0.696** *(-2.65)*
Labour-Tax	0.002 *(0.49)*	-0.003 *(-0.64)*	0.242*** *(2.91)*	0.091 *(0.99)*
Capital-Tax	0.01 *(1.12)*	0.007 *(0.89)*	0.308*** *(2.91)*	0.118 *(1.31)*
Property-Tax	-0.007* *(-1.82)*	-0.006* *(-1.80)*	0.042 *(0.51)*	0.014 *(0.22)*
Other-Tax	-0.003 *(-1.15)*	-0.003 *(-1.52)*	0.074 *(1.20)*	0.011 *(0.20)*
Mixed-Tax	-0.298*** *(-5.06)*	-0.303*** *(-6.77)*	-0.459*** *(-4.64)*	-0.592*** *(-5.02)*
Overall-Tax	0.015 *(1.11)*	0.012 *(1.07)*	0.054 *(0.60)*	0.033 *(0.37)*
Control Variables Selected Via Backwards Stepwise Regression				
SE	-0.987*** *(-3.88)*	-0.870*** *(-3.45)*	-1.054*** *(-3.45)*	-0.637*** *(-3.33)*
Labour-Tax	0.003 *(0.58)*	-0.002 *(-0.46)*	0.218** *(2.55)*	0.071 *(1.43)*
Capital-Tax	0.010 *(1.24)*	0.006 *(0.83)*	0.290** *(2.62)*	0.088* *(1.84)*
Property-Tax	-0.006* *(-1.86)*	-0.005* *(-1.88)*	0.032 *(0.46)*	-0.006 *(-0.12)*
Other-Tax	-0.002 *(-1.08)*	-0.003 *(-1.22)*	0.065 *(1.10)*	-0.030 *(-0.60)*
Mixed-Tax	-0.315*** *(-7.61)*	0.301*** *(-13.58)*	-0.453*** *(-6.05)*	-0.641*** *(-6.83)*
Overall-Tax	0.010 *(1.34)*	0.001 *(0.50)*	0.029 *(0.38)*	-0.035 *(-0.80)*

Note: The top panel reports the results of estimating Equation (3.2) with the addition of the five tax variables, *Labour, Capital, Other, Mixed,* and *Overall* taxes. The bottom panel adds control variables selected through a backwards stepwise regression procedure that selects variables so as to minimize the Schwarz Information Criterion (see Footnote #25). The top value in each cell is the coefficient estimate, and the bottom value in parentheses is the associated *t*-statistic. The four WLS estimators (*Fixed Effects-Weight1, Fixed Effects-Weight2, Random Effects-Weight1,* and *Random Effects-Weight2*) are described in the "FAT/PET tests" subsection of Section 3.4 in the text. All four estimation procedures calculate cluster robust standard errors. *, **, and *** indicate statistical significance at the 10-, 5-, and 1-percent level, respectively.

Table 3.9 : Meta-Regression Analysis (All Tax Variables Included)

Variable	***Fixed Effects (Weight1) (1)***	***Fixed Effects (Weight2) (2)***	***Random Effects (Weight1) (3)***	***Random Effects (Weight2) (4)***
All Control Variables Included				
SE	-0.949*** *(-3.02)*	-0.968** *(-2.36)*	-0.954*** *(-3.23)*	-0.658** *(-2.55)*
Prediction-Negative	-0.013 *(-0.99)*	-0.012 *(-0.86)*	0.122*** *(3.53)*	0.092* *(1.79)*
Prediction-Positive	-0.82*** *(-6.27)*	-0.354*** *(-2.82)*	-0.272*** *(-3.44)*	-0.455** *(-2.49)*
Labour-Tax	0.006 *(0.91)*	0.0005 *(0.09)*	0.153** *(2.61)*	0.046 *(0.65)*
Capital-Tax	0.014 *(1.20)*	0.011 *(0.99)*	0.211** *(2.62)*	0.067 *(1.00)*
Property-Tax	-0.002 *(-0.66)*	-0.002 *(-0.76)*	-0.058 *(-0.73)*	-0.048 *(-0.83)*
Other-Tax	0.001 *(0.30)*	0.0001 *(0.04)*	-0.012 *(-0.37)*	-0.058 *(-1.23)*
Mixed-Tax	-0.293*** *(-5.00)*	-0.293*** *(-6.59)*	-0.540*** *(-6.45)*	-0.638*** *(-6.26)*
Overall-Tax	0.017 *(1.11)*	0.014 *(1.07)*	-0.012 *(-0.15)*	-0.003 *(-0.04)*

Note: This panel reports the results of estimating Equation (3.2) with the addition of the seven tax variables, *Prediction-Negative, Prediction-Positive, Labour, Capital, Other, Mixed,* and *Overall* taxes. The top value in each cell is the coefficient estimate, and the bottom value in parentheses is the associated *t*-statistic. The four WLS estimators (*Fixed Effects-Weight1, Fixed Effects-Weight2, Random Effects-Weight1,* and *Random Effects-Weight2*) are described in the "FAT/PET tests" subsection of Section 3.4 in the text. All four estimation procedures calculate cluster robust standard errors. *, **, and *** indicate statistical significance at the 10-, 5-, and 1-percent level, respectively.

Table 3.9, continued: Meta-Regression Analysis (All Tax Variables Included)

Variable	*Fixed Effects (Weight1) (1)*	*Fixed Effects (Weight2) (2)*	*Random Effects (Weight1) (3)*	*Random Effects (Weight2) (4)*
Control Variables Selected Via Backwards Stepwise Regression				
SE	-0.848*** *(-3.73)*	-0.888*** *(-2.89)*	-0.993*** *(-3.25)*	-0.685*** *(-3.23)*
Prediction-Negative	-0.006 *(-0.61)*	-0.019 *(-1.22)*	0.073** *(2.41)*	0.050 *(1.10)*
Prediction-Positive	-0.280*** *(-6.59)*	-0.192 *(-1.36)*	-0.295*** *(-3.42)*	-0.470*** *(-4.10)*
Labour-Tax	0.005 *(0.57)*	0.002 *(0.24)*	0.136** *(2.37)*	0.040 *(1.12)*
Capital-Tax	0.012 *(1.11)*	0.012 *(0.99)*	0.202** *(2.44)*	0.065 *(1.62)*
Property-Tax	-0.003 *(-0.97)*	-0.0002 *(-0.06)*	-0.073 *(-1.16)*	-0.024 *(-0.70)*
Other-Tax	0.0003 *(0.07)*	0.002 *(0.41)*	-0.027 *(-1.09)*	-0.030 *(-0.83)*
Mixed-Tax	-0.306*** *(-8.06)*	-0.315*** *(-11.93)*	-0.515*** *(-9.58)*	-0.554*** *(-6.67)*
Overall-Tax	0.013 *(1.11)*	0.007 *(1.00)*	-0.024 *(-0.42)*	-0.026 *(-0.83)*

Note: This panel adds control variables selected through a backwards stepwise regression procedure that selects variables so as to minimize the Schwarz Information Criterion (see Footnote #25). The top value in each cell is the coefficient estimate, and the bottom value in parentheses is the associated *t*-statistic. The four WLS estimators (*Fixed Effects-Weight1*, *Fixed Effects-Weight2*, *Random Effects-Weight1*, and *Random Effects-Weight2*) are described in the "FAT/PET tests" subsection of Section 3.4 in the text. All four estimation procedures calculate cluster robust standard errors. *, **, and *** indicate statistical significance at the 10-, 5-, and 1-percent level, respectively.

The results are even more inconsistent with the theoretical prodictions when the tax type variables are added to the specification. Table 3.9 reports the corresponding estimates. The coefficient for *Prediction-Negative* remains negative and statistically insignificant in the two fixed effects regressions (Column 1 and 2), but positive and statistically significant in the two random effects regressions (the top panel of column 3 and 4). *Prediction-Positive* is negative and statistically significant across all four estimation procedures, and for both sets of control variables (the exception is the bottom panel, second Column).

The last tax issue addressed in this study investigates whether some types of taxes are more growth retarding than others. As noted in Table 3.1, Labour and Capital taxes are commonly classified as distortionary, while Consumption taxes are classified as non-distortionary. Table 3.8 estimates a meta-regression with the tax type variables but with prediction variables omitted, while Table 3.9 includes both. As the omitted category is Consumption taxes, we expect the coefficient on Labour and Capital taxes to be negative, whereas there is no sign expectation for the other tax type coefficients.

With respect to Labour taxes, the results from Table 3.8, with both sets of control variables, are negative but statistically insignificant in the *Fixed Effects (Weight2)* case, positive and statistically significant when using the *Random Effects (Weight1)* estimation, and positive and statistically insignificant in the two remaining estimation procedures. When prediction variables are added to the regression (cf. Table 3.9), the coefficient on Labour-Tax becomes all positive and statistically insignificant in three out of four cases. We can conclude that there is insufficient evidence that Labour taxes are more growth-retarding than Consumption taxes is mixed.

There is also no evidence that Capital taxes are more distortionary than Consumption taxes. The coefficients on the *Capital-Tax* variable are positive in all Table 3.8 regressions. They are insignificant across the fixed effects estimations and with both sets of control variables, but in contrast they are generally significant across the random effects estimations. When the prediction variables are added, the respective coefficients are generally positive and insignificant (cf. Table 3.9). One of the regressions even produces a significant positive coefficient (top and bottom panel, *Random Effects-Weight1).* As a result, we conclude that there is no evidence that Capital taxes are more distortionary than Consumption taxes.

<u>Bayesian model averaging of control variables</u>. Having addressed the major goals of this study, we turn to an analysis of the control variables. There are 26 control variables not counting the two sets of tax variables. Multicollinearity is always a potential problem with so many variables. For example, when all the 26 variables are included with both sets of tax variables and the meta-regression is estimated using the *Random Effects-Weight2* estimator, as in Column (4) of the top panel of Table 3.9, 5 of the 26 control variables are statistically significant at the 5 percent level. When the backwards stepwise routine is employed, as in the bottom panel of Table 3.9, 9 of the 26 control variables are significant. Two of the variables that are significant in the top panel are no longer significant in the bottom panel's specification. Thus, variable selection makes a difference. This was not so much of a problem when we estimated tax effects, because the variables were locked into the respective specifications without regard to statistical significance. However, it is a problem when trying to decide which control variables to include in a parsimonious specification.

We use Bayesian Model averaging (BMA) to address this issue (Zeugner, 2011). BMA involves running many regressions with various subsets of these 26 control variables and then constructing the weighted average over these regressions. Table 3.10 reports the results of an analysis where we lock in the tax variables *Prediction-Negative* and *Prediction-Positive* and then apply BMA to the 26 control variables. All specifications adjust for publication bias. The results differ somewhat depending on the estimation procedure used. However, they are more consistent across analyses than would be the case, say, if we reported the results from the specifications that included all variables and those that employed stepwise regression. We report results for both the *Fixed effects-Weight1* and *Random Effects-Weight2* estimators. These two estimators use very different weighting schemes. Previous tables indicated that the estimates from these two estimators sometimes vary substantially. As a result, they provide an indication of robustness across estimation procedures.

We report three summary measures. The Posterior Inclusion Probability *(PIP)* is a weighted probability that uses the likelihood values of specifications to construct a "probability" that a given specification is "true". There are 10^{26} possible variable specifications with 26 control variables. Variables that appear in specifications with high likelihood values will have larger *PIP* values. By construction, every variable appears in 50 percent of all possible specifications. However, the *PIP* can be very close to 100 percent if the specifications that include a variable have much greater likelihood values than those in which it is omitted.

The Posterior Mean *(Post. Mean)* uses the above-mentioned probability values to weight the estimated coefficients from each specification. Specifications in which a variable is not included assign an "estimated value" of zero to construct the Posterior Mean. Lastly, BMA also calculates the probability that a given coefficient has a positive sign *(Cond. Pos. Sign)*. This is constructed in the same manner as the Posterior Mean, except that it uses a dummy variable indicating positive value in constructing a weighted average rather than the estimated coefficients.

Yellow rows in Table 3.10 highlight all the control variables that: (i) have a *PIP* greater than 50%; (ii) have a Conditional Positive Sign of either 1.00 or 0.00 – indicating that the respective coefficient is consistently estimated to be either positive or negative in the most likely specifications; and (iii) have the same Conditional Positive Sign value for both the *Fixed Effects (Weight1)* and *Random Effects (Weight2)* estimators.

Studies that estimate tax effects for all the states plus the District of Columbia while excluding Alaska from the sample produce consistently less negative/more positive estimates than studies excluding Alaska, Hawaii, and the District of Columbia as is conventional in the literature. To place the size of the Posterior Mean values in context, it is helpful to recall that the median estimated tax effect from Table 3.4 is -0.055. By this standard, the effect of excluding AK from the sample is relatively large (0.180 and 0.175, respectively). The effect associated with excluding both Alaska and the District of Columbia is negative and smaller. We find that studies that employ Differenced (as opposed to level), generally produce more negative/less positive tax effects. Compared to the short-run effects of taxes, studies that estimate long-run effects produce estimates that are more negative/less positive.

Table 3.10: Bayesian Model Averaging Analysis (Control Variables)

Variable	***Fixed Effects(Weight1)***			***Random Effects(Weight2)***		
	PIP	***Post. Mean***	***Cond. Pos. Sign***	***PIP***	***Post. Mean***	***Cond. Pos. Sign***
Peer-reviewed	0.79	-0.0023	0.14	0.07	-0.0011	0.10
Publication Year	1.00	-0.0010	0.00	0.07	0.0001	0.00
Cross-section	0.97	-0.0363	0.00	0.14	0.0064	1.00
Length	1.00	0.0009	1.00	0.06	0.0000	0.31
Mid-Year	0.92	0.0004	1.00	0.06	0.0000	0.92
GSP	0.77	0.0081	0.97	0.06	0.0007	0.99
PC-GSP	0.96	0.0498	1.00	0.05	-0.0003	0.22
PI	0.95	-0.0075	0.00	0.15	-0.0096	0.00
Marginal	0.88	0.0190	1.00	0.06	0.0008	0.98
Differenced	**1.00**	**-0.0326**	**0.00**	**1.00**	**-0.1054**	**0.00**
ETR	0.88	0.0123	1.00	1.00	-0.0933	0.00
HLK	0.99	-0.0166	0.00	0.06	-0.0005	0.25
Initial Income	0.89	-0.0096	0.01	0.08	-0.0017	0.01
Lagged DV	1.00	0.0571	1.00	0.47	-0.0276	0.00
State FE	0.80	0.0027	0.98	0.53	-0.0281	0.00

Variable	***Fixed Effects(Weight1)***			***Random Effects(Weight2)***		
	PIP	***Post. Mean***	***Cond. Pos. Sign***	***PIP***	***Post. Mean***	***Cond. Pos. Sign***
SE-HET	1.00	-0.0401	0.00	0.12	0.0043	0.95
SE-Other	0.99	-0.0127	0.00	0.05	-0.0002	0.22
Medium-run	1.00	-0.0323	0.00	0.08	0.0019	0.84
Long-run	**0.83**	**-0.0551**	**0.00**	**0.86**	**-0.1535**	**0.00**
GLS	0.86	0.0065	0.98	0.68	0.0362	1.00
TSLS/GMM	0.74	0.0003	0.71	0.07	0.0012	0.85
None	0.74	-0.004	0.18	0.06	-0.0009	0.19
AK	**1.00**	**0.1798**	**1.00**	**1.00**	**0.1746**	**1.00**
DC	1.00	-0.0982	0.00	0.06	-0.0005	0.39
AKDC	**1.00**	**-0.1222**	**0.00**	**1.00**	**-0.2995**	**0.00**
AKDCHIOTHERS	0.79	0.0114	1.00	0.06	-0.0009	0.04

Figure 3.6 and Figure 3.7 provide visual representations of the BMA analysis for the tax *(Prediction-Negative* and *Prediction-Positive)* and control variables using the *Fixed Effects (Weight1) and Random Effects (Weight2)* estimators, respectively.[36] The figures report estimates from the top 1000 models, out of 10^{26} possible models, account for a cumulative probability of approximately 37 and 23 percent. Red (blue) squares indicate that the respective coefficient is negative (positive) in the given model. A white square indicates that the variable is omitted from that model. Solid bands of the same colour across the figures indicate that the respective variable is consistently estimated to have the same sign across all 1000 models. In addition to confirming the results from Table 3.10, the figures also indicate the variable specifications of the top models. These closely match the *PIP* values in Table 3.10.

[36] Note that in the associated specifications, the variable *Precision* corresponds to the constant term, while the constant term corresponds to the publication biascorrection factor which is (1/*SE)*.

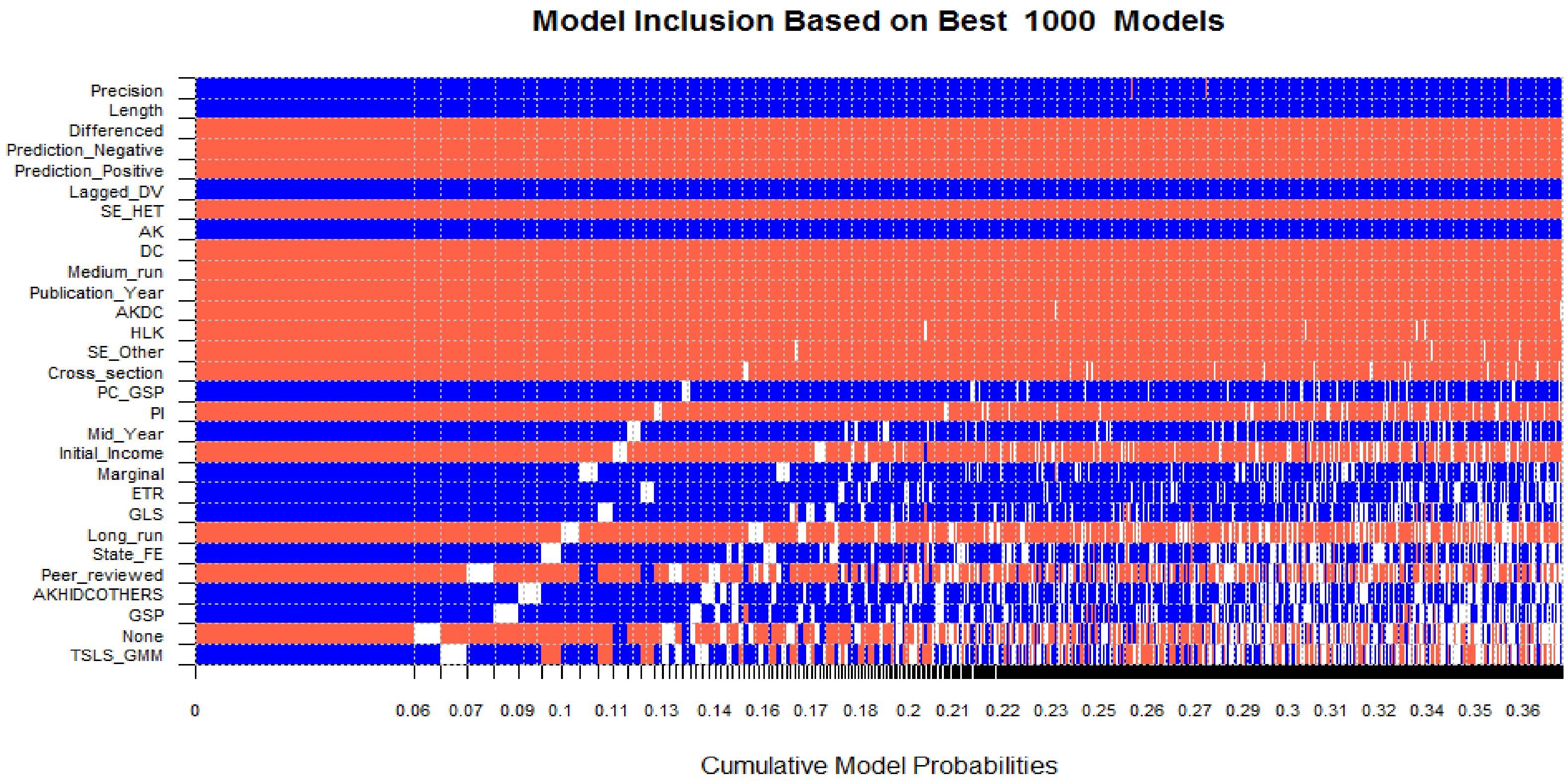


Figure 3.6: Visual Representation of BMA Analysis (Fixed Effects-Weight1)

<u>Note</u>: Each column represents a single model. Variables are listed in descending order of Posterior Inclusion Probability (PIP) and have all been weighted according to the *Fixed Effects-Weight1* case. Blue (dark) indicates that the variable is included in that model and estimated to be positive. Red (light) indicates that the variable is included and estimated to be negative. No colour (white) indicates that the variable is not included in that mode. Further details about this plot are given in Zeugner (2011).

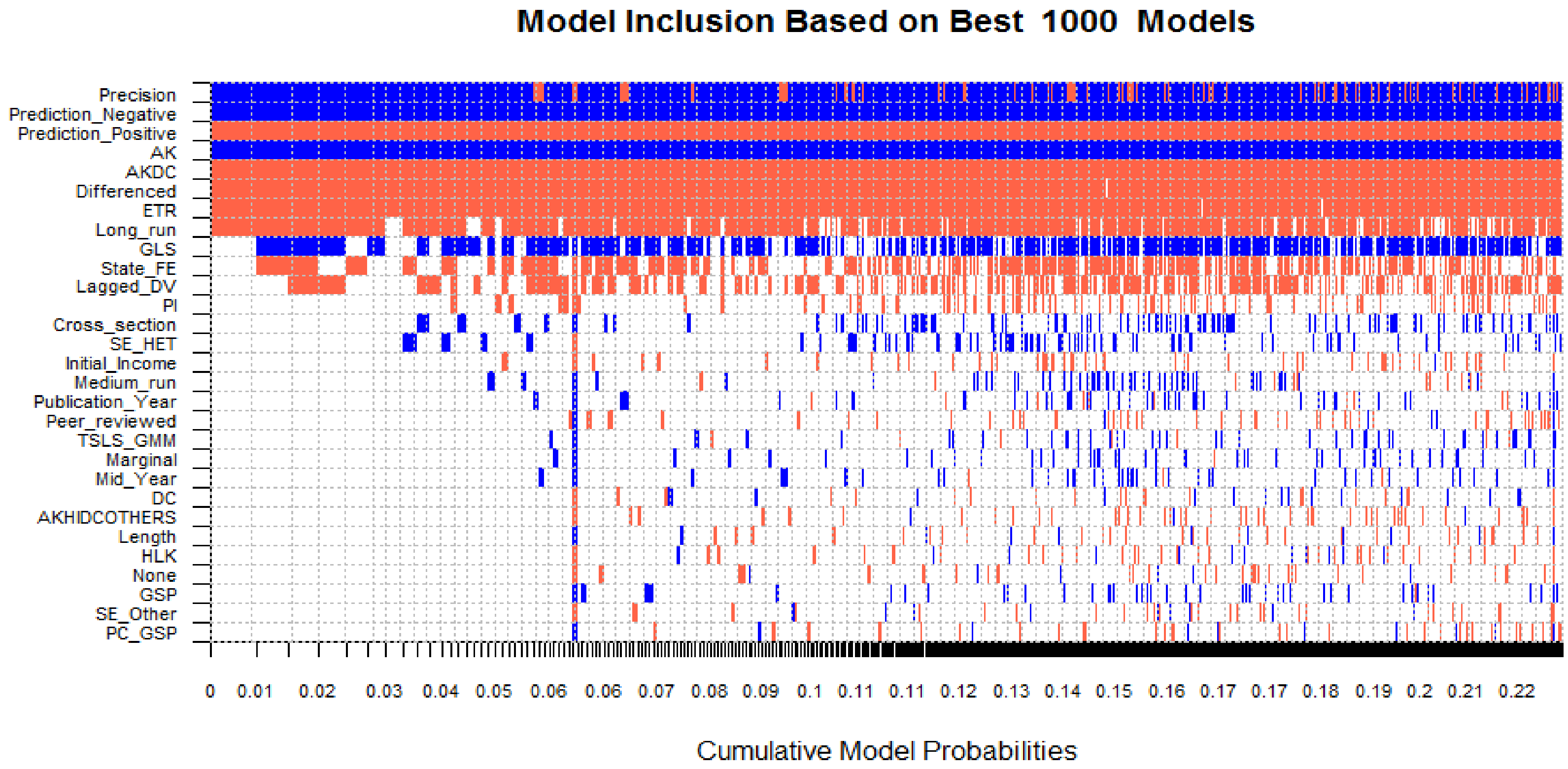


Figure 3.7: Visual Representation of BMA Analysis (Random Effects-Weight2)

__Note__: Each column represents a single model. Variables are listed in descending order of Posterior Inclusion Probability (PIP) and have all been weighted according to the Random Effects-Weight2 case. Blue (dark) indicates that the variable is included in that model and estimated to be positive. Red (light) indicates that the variable is included and estimated to be negative. No colour (white) indicates that the variable is not included in that mode. Further details about this plot are given in Zeugner (2011).

3.5. Conclusion

The impact of state and local taxes on state economic growth in the US has been a long-lasting question. The importance of this issue is reflected by the large number of empirical studies investigating this relationship. However, the literature has generated a large number of frequently conflicting estimates. The main reason being that the estimates of tax effects of different studies are often estimating different things. It is important to realize that the same tax effect can be estimated to be positive or negative, depending on the omitted fiscal categories from the specification. As a result, ignoring the role of government budget constraints may change the results dramatically. For this and other reasons, it is valuable to collect the estimates from this literature and carefully track the differences within and between studies so that the estimates can be integrated in order to provide an overall assessment of the growth effects of taxes.

This study combines 966 estimates derived from 29 studies, all of which examine the effect of taxes on economic growth in U.S. states. Extreme outliers are dropped from both ends of the sample range, and all the analyses discussed for a final sample of 868 estimates. The results show evidence of negative publication selection bias in the literature. Controlling for publication bias, the overall effect of state and local taxes on state economic growth is small in magnitude and statistically insignificant. However, this overall effect is not particularly meaningful since it lumps together different kinds of tax policies. With respect to particular types of taxes, I could not find enough evidence to support that taxes on labour are more growth retarding than other types of taxes. Evidence regarding other types of taxes is mixed.

Overall, these results are very surprising and stand in stark contrast to the results I obtained in the previous chapter when analysing studies of tax effects in OECD countries.

Ironically, there is a temptation to keep manipulating the data until I get results from US states that are consistent with those from OECD countries. However, I am mindful that this procedure violates the research plan with which I began my study. To further manipulate the analysis to obtain consistent results would contribute to the very publication bias I am trying to overcome in my analysis. As a result, I believe it is important to report the results for US states using the exact same procedures that I used for the OECD countries, without alteration. I leave it as a future research subject to further explore these inconsistencies.

While further investigation is required to precisely explore the reasons for this state of affairs, there are three hypotheses that may explain the observed contrast in the reported results in two consecutive chapters. First, a simple comparison of summary statistics on the predicted tax effects respectively reported in Table 2.7 and Table 3.6 shows that research undertaken in US context generally ignores the government budget constraints. Second, Econometric methods applied in OECD studies are generally more sophisticated with some attempts to distinguish between the short- versus long-run effects. Finally, the econometrics techniques applied in OECD studies have tackled the problem of endogeneity.

3.6. References

Arkes, Jeremy. (2013). The effect of state income tax rates on economic growth. *Naval Postgraduate School.*

Bania, N., Gray, J.A., and Stone, J.A. (2007). Growth, taxes, and government expenditures: Growth hills for U.S. states. *National Tax Journal*, *60*(2), 193-204.

Bartik, T. J. (1992). The effects of state and local taxes on economic development: A review of recent research. *Economic Development Quarterly*, *6*(1), 102-111.

Brady, K. K. (2007). State government size and economic growth: A panel data analysis of the United States over the period 1986-2003. *AJ Palumbo School of business administration, Duquesne University*.

Briem, C. (1998). State fiscal policy and regional growth. Retrieved from http://www.briem.com/papers/3yp11.PDF

Bruce, D., Deskins, J. A., Hill, B. C., and Rork, J. C. (2009). (Small) business activity and state economic growth: does size matter? *Regional Studies*, *43*(2), 229-245.

Carroll, R., and Wasylenko, M. (1994). Do state business climates still matter?Evidence of a structural change. *National Tax Journal*, *47*(1), 19-37.

Chernick, H. (1997). Tax progressivity and state economic performance. *Economic Development Quarterly, 11*(3), 249-267.

Chernick, H., and Sturm, P. (2004, January). Redistribution at the State and Local Level: Consequences for Economic Growth. In *Proceedings. Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association* (Vol. 97, pp. 190-201). National Tax Association.

Chernick, H. (2010). Redistribution at the state and local level: Consequences for economic growth. *Public Finance Review, 38*(4), 409-449.

Deskins, J., and Hill, B. (2010). State taxes and economic growth revisited: have distortions changed? *The Annals of Regional Science*, *44*(2), 331-348.

Easterly, W., and Rebelo, S. (1993). Fiscal policy and economic growth. *Journal of Monetary Economics*, *32*(3), 417-458.

Engen, E. M., and Skinner, J. (1992). Fiscal policy and economic growth. Working paper No. w4223. *National Bureau of Economic Research*.

Feng, J., and Young, D. J. (2003). Taxes, growth and the convergence of incomes among US states: 1959-99. *Department of Agricultural Economics and Economics (Montana State University).*

Gemmell, N., Kneller, R., and Sanz, I, (2009). The composition of government expenditure and economic growth: some new evidence for OECD countries. In European Commission, *The quality of public finances and economic growth: Proceedings to the annual workshop on public finances (Brussels, 28 November 2008*), European Economy, Occasional Papers No. 45. Directorate-General for Economic and Financial Affairs, Publications, Brussels, Belgium, pages 17-46.

Goff, B., Lebedinsky, A., and Lile, S. (2012). A matched pairs analysis of state growth differences. *Contemporary Economic Policy, 30*(2), 293-305.

Havranek, T., and Irsova, Z. (2012). Survey article: Publication bias in the literature on foreign direct investment spillovers. *The Journal of Development Studies*, *48*(10), 1375-1396.

Helms, L. J. (1985). The effect of state and local taxes on economic growth: A time series-cross section approach. *The Review of Economics and Statistics*, 574-582.

Higgins, J.P. and Thompson, S.G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21*(11), 1539-58.

Hunter, J. E., and Schmidt, F. L. (Eds.). (2004). Methods of meta-analysis: Correcting error and bias in research findings. London: Sage.

Kneller, R., Bleaney, M. F., and Gemmell, N. (1999). Fiscal policy and growth: evidence from OECD countries. *Journal of Public Economics*, *74*(2), 171-190.

Mazerov, M. (2013). Academic research lacks consensus on the impact of state tax cuts on economic growth. *Center on Budget and Policy Priorities*.

McBride, W. (2012). What is the evidence on taxes and growth. *Tax Foundation*, 207.

McCracken, C. (2006). Whether state fiscal policy affects state economic growth. *Economics Department (Stanford, CA: Stanford University).*

Meier, K. J., and Artiz, S. (2012). Taxes, incentives, and economic growth: Assessing the impact of pro-business taxes on US state economies. In *APSA 2012 Annual Meeting Paper*.

Miller, S. M., and Russek, F. S. (1997). Fiscal structures and economic growth at the state and local level. *Public Finance Review, 25*(2), 213-237.

Miller, S. M., and Clarke, C. A. (2014). Can state and local revenue and expenditure enhance economic growth? A cross-state panel study of fiscal activity. (February 3, 2014).

Mullen, J. K., and Williams, M. (1994). Marginal tax rates and state economic growth. *Regional Science and Urban Economics, 24*(6), 687-705.

Nijkamp, P., and Poot, J. (2004). Meta-analysis of the effect of fiscal policies on long-run growth. *European Journal of Political Economy*, *20*(1), 91-124.

Ojede, A., and Yamarik, S. (2012). Tax policy and state economic growth: The long-run and short-run of it. *Economics Letters, 116*(2), 161-165.

Ojede, A., Atems, B., and Yamarik, S. Disentangling direct and spillover effects of fiscal policy on economic growth using disaggregate U.S. state level data. Retrieved from http://www.researchgate.net/

Phillips, J. M., and Goss, E. P. (1995). The effect of state and local taxes on economic development: A meta-analysis. *Southern Economic Journal*, 320-333.

Pjesky, R. J. (2002). A reevaluation and extention of previous research on state and local taxes and economic development. Doctoral dissertation, University of Oklahoma.

Reed, W. R. (2008). The robust relationship between taxes and US state income growth. *National Tax Journal*, *61*(1), 57-80.

Reed, W.R. (2015). A Monte Carlo analysis of alternative meta-analysis estimators in the presence of publication bias. *Economics: The Open-Access, Open-Assessment E-Journal*, *9*(2015-30): 1-40.

Rhee, T. H. (2008). Macroeconomic effects of progressive taxation. Retrieved from http://www.aeaweb.org/aea/2013conference/program/retrieve.php?pdfid=394

Shemilt, I., Mugford, M., Vale, L., Marsh, K. and Donaldson, C., (Eds.) (2011). *Evidence-based decisions and economics: Health care, social welfare, education, and criminal justice*. San Francisco, CA : John Wiley & Sons.

Solow, R. M. (1956). A contribution to the theory of economic growth. *The Quarterly Journal of Economics*, *70*(1), 65-94.

Srithongrung, A., and Kriz, K. A. (2014). The impact of subnational fiscal policies on economic growth: a dynamic analysis approach. *Journal of Policy Analysis and Management*, *33*(4), 912-928.

Stanley, T.D., and Doucouliagos, H. (2012). Meta-regression analysis in economics and business. London: Routledge.

Stanley, T.D., Doucouliagos, H., Giles, M., Heckemeyer, J.H., Johnston, R.J., Laroche, P., Nelson, J.P., Paldam, M., Poot, J., Pugh, G. and Rosenberger, R.S. (2013). Meta-analysis of economics research reporting guidelines. *Journal of Economic Surveys*, *27*(2), 390-394.

Sutton, A. J., Duval, S. J., Tweedie, R. L., Abrams, K. R., and Jones, D. R. (2000). Empirical assessment of effect of publication bias on meta-analyses. *Bmj*, 320(7249), 1574-1577.

Swan, T. W. (1956). Economic growth and capital accumulation. *Economic Record*, *32*(2), 334-361.

Tomljanovich, M. (2004). The role of state fiscal policy in state economic growth. *Contemporary Economic Policy*, *22*(3), 318-330.

Vedder, R. (1990). Tiebout, taxes, and economic growth. *Cato Journal*, *10*, 91.

Wasylenko, M. (1997). Taxation and economic development: the state of the economic literature. *New England Economic Review*, 37-52.

Wibowo, K. (2003). An empirical analysis of taxation and state economic growth. Doctoral dissertation, University of Oklahoma.

Yamarik, S. (2000). Can tax policy help explain state-level macroeconomic growth? *Economics Letters*, *68*(2), 211-215.

Yu, W., Wallace, M. S., and Nardinelli, C. (1991). State growth rates: taxes, spending, and catching up. *Public Finance Review*, *19*(1), 80-93.

Zeugner, S. (2011). Bayesian model averaging with BMS. Available at http://bms.zeugner.eu/ [Accessed 15 Oct. 2015].

3.7. Appendices

Appendix 3.1: List of Terms Used in Electronic Search by Category

TAX	***ECONOMIC GROWTH***	***U.S. STATES***
Tax(es) /Tax rate(s)/Taxation	Economic growth	Contiguous states
State and local taxes/taxation	Regional economic growth	U.S. States
Tax policy(policies)	State economic growth	American States
Tax ratios	State economic performance	State and local level
Tax changes	Regional economic activity	
Tax rate change	Economic indicators	
State fiscal policy(policies)	Long-term growth	
Tax structures/Fiscal structures	Long-run growth	
Fiscal decentralization	Growth	
Public finances		

Appendix 3.2: Letter to the Authors (U.S. States)

Dear Sir/Madam,

I am a Professor of economics at the University of Canterbury in New Zealand. We have a research team here undertaking a "meta-analysis" of the relationship between taxes and economic growth in US states.

A thorough meta-analysis involves collecting as many papers as possible on a subject, including unpublished research. The latter is known as "grey literature", and includes conference proceedings, reports from research firms or think thanks, theses and dissertations, etc. The unpublished literature is particularly important for addressing publication bias.

In this context, I am asking for your help.

Attached to this email is a listing of research on the topic of taxes and economic growth in US states. To be included, the research had to (i) include data from at least 45 US states, (ii) have a dependent variable that was the growth of state per capita personal income (PCPI) or GSP, and (iii) include one or more measures of state-level taxes.

I am contacting you because you have researched in this area in the past.

Would you please look over this list and see if there any notable omissions? I have broken the list down to the following categories: (i) journal articles, (ii) conference proceedings, (iii) studies from think tanks and research firms, (iv) theses/dissertations, and (v) working papers and unpublished research.

The last two categories are especially difficult to get information on. I would be greatly appreciative if you could identify any research we may have omitted.

Finally, if you are aware of any researchers who are currently researching in this area, it would be great if you could reply back with their names, and I will follow up with them directly.

I am sure you would agree that the subject of taxes and economic growth in US states is very important. There is now a substantial enough literature that a careful meta-analysis can help to organize an empirical consensus of the existing literature.

Thank you so much for any help you can provide.

Sincerely,

Appendix 3.3: Bibliography (Attachment to the Above Letter)

BIBLIOGRAPHY

(vi) Journal Articles

Adkisson, R. V., & Mohammed, M. (2014). Tax structure and state economic growth during the Great Recession. *The Social Science Journal, 51*(1), 79-89.

Alm, J., & Rogers, J. (2011). Do state fiscal policies affect state economic growth? *Public Finance Review, 39*(4), 483-526.

Bania, N., Gray, J. A., & Stone, J. A. (2007). Growth, taxes, and government expenditures: Growth hills for US states. *National Tax Journal*, *60(2)*,193-204.

Bania, N., & Stone, J. A. (2008). Ranking state fiscal structures using theory and evidence. *Journal of Policy Analysis and Management*, 27(4), 751-770.

Becsi, Z. (1996). Do state and local taxes affect relative state growth? *Economic Review, 81*(2), 18-36.

Canto, V. A., & Webb, R. I. (1987). The effect of state fiscal policy on state relative economic performance. *Southern Economic Journal, 54*(1), 186-202.

Chernick, H. (1997). Tax progressivity and state economic performance. *Economic Development Quarterly*, *11*(3), 249-267.

Chernick, H. (2010). Redistribution at the state and local level: Consequences for economic growth. *Public Finance Review*, *38*(4), 409-449.

Dahl, D. S., & Gane, S. H. (1980). The impact of state and local taxes on economic growth. *Regional Science Perspectives, 10*(2), 3-9.

Deskins, J., & Hill, B. (2010). State taxes and economic growth revisited: Have distortions changed? *The Annals of Regional Science, 44*(2), 331-348.

Dye, T. R. (1980). Taxing, spending, and economic growth in the American states. *The Journal of Politics*, *42*(04), 1085-1107.

Goff, B., Lebedinsky, A., & Lile, S. (2012). A matched pairs analysis of state growth differences. *Contemporary Economic Policy*, *30*(2), 293-305.

Holcombe, R. G., & Lacombe, D. J. (2004). The effect of state income taxation on per capita income growth. *Public Finance Review*, *32*(3), 292-312.

Howard, B. (2011). Does the mix of a state's tax portfolio matter? An empirical analysis of United States tax portfolios and the relationship to levels and growth in real per capita gross state product. *Journal of Applied Business Research (JABR)*, *19*(3), 97-110.

Miller, S. M., & Russek, F. S. (1997). Fiscal structures and economic growth at the state and local level. *Public Finance Review*, *25*(2), 213-237.

Mullen, J. K., & Williams, M. (1994). Marginal tax rates and state economic growth. *Regional Science and Urban Economics*, *24*(6), 687-705.

Ojede, A., & Yamarik, S. (2012). Tax policy and state economic growth: The long-run and short-run of it. *Economics Letters*, *116*(2), 161-165.

Prillaman, S. A., & Meier, K. J. (2014). Taxes, incentives, and economic growth: Assessing the impact of pro-business taxes on US state economies.*The Journal of Politics*, *76(2)*, 364-379.

Reed, W. R. (2008). The robust relationship between taxes and US state income growth. *National Tax Journal*, *61(1)*, 57-80.

Reed, W. R. (2009). The determinants of US state economic growth: A less extreme bounds analysis. *Economic Inquiry, 47*(4), 685-700.

Romans, T., & Subrahmanyam, G. (1979). State and local taxes, transfers and regional economic growth. *Southern Economic Journal*, *46*(2), 435-444.

Tomljanovich, M. (2004). The role of state fiscal policy in state economic growth. *Contemporary Economic Policy, 22*(3), 318-330.

Yamarik, S. (2000). Can tax policy help explain state-level macroeconomic growth? *Economics Letters, 68*(2), 211-215.

Yu, W., Wallace, M. S., & Nardinelli, C. (1991). State growth rates: Taxes, spending, and catching up. *Public Finance Review, 19*(1), 80-93.

(vii) Conference Proceedings

Chernick, H., & Sturm, P. (2004). Redistribution at the state and local level: Consequences for economic growth. *Paper presented at the Proceedings. Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association*.

Reed, W. R. (2006). The determinants of U.S. state economic growth: A search for control variables. *Paper presented at the Annual Meeting of the National Tax Association*, Boston, MA.

(viii) Studies from Think Thanks and Research Firms

Genetski, R. J., & Chin, Y. D. (1978). The impact of state and local taxes on economic growth. *Harris Bank, Harris Economic Research Office Service.*

Poulson, B. W., & Kaplan, J. G. (2008). State income taxes and economic growth. *Cato Journal*, *28*, 53.

Scully, G. W. (1991). How state and local taxes affect economic growth. *National Center for Policy Anallysis.*

Vedder, R. (1990). Tiebout, taxes, and economic growth. *Cato Journal, 10*, 91.

(ix) Theses/ Dissertations

Brady, K. K. (2007). State government size and economic growth: A panel data analysis of the United States over the period 1986-2003 (Bachelor thesis, Duquesne University)

Clarke, C. A. (2013). Can state and local revenue and expenditure enhance economic growth? A cross-state panel study of fiscal activity (Master thesis, University of Nevada, Las Vegas).

Feng, J. (2002). State and local taxes and economic performance: An empirical study of how much taxes matter to economies (Master thesis, Montana State University-Bozeman, College of Agriculture).

Pjesky, R. J. (2002). A reevaluation and extention of previous research on state and local taxes and economic development (Doctoral dissertation, University of Oklahoma).

Wibowo, K. (2003). An empirical analysis of taxation and state economic growth (Doctoral dissertation, University of Oklahoma).

(x) Working papers and other unpublished research

Arkes, Jeremy. (2013). The effect of state income tax rates on economic growth. Naval Postgraduate School.

Bania, N., & Gray, J. A. (2006). Taxes, government expenditures, and state economic growth: The role of nonlinearities. *Department of Economics (University of Oregon), v1.5.*

Briem, C. (1998). State fiscal policy and regional growth. Retrieved from http://www.briem.com/papers/3yp11.PDF

Feng, J., & Young, D. J. (2003). Taxes, growth and the convergence of incomes among US states: 1959-99. *Department of Agricultural Economics and Economics (Montana State University).*

McCarty, R., & Bruce, D. (2010). Corporate income tax systems and state economic activity. *(University of Tennessee).*

McCracken, C. (2006). Whether state fiscal policy affects state economic growth. *Economics Department (Stanford, CA: Stanford University).*

Ojede, A., Atems, B., & Yamarik, S. Disentangling direct and spillover effects of fiscal policy on economic growth using disaggregate U.S. state level data. Retrieved from http://www.researchgate.net/

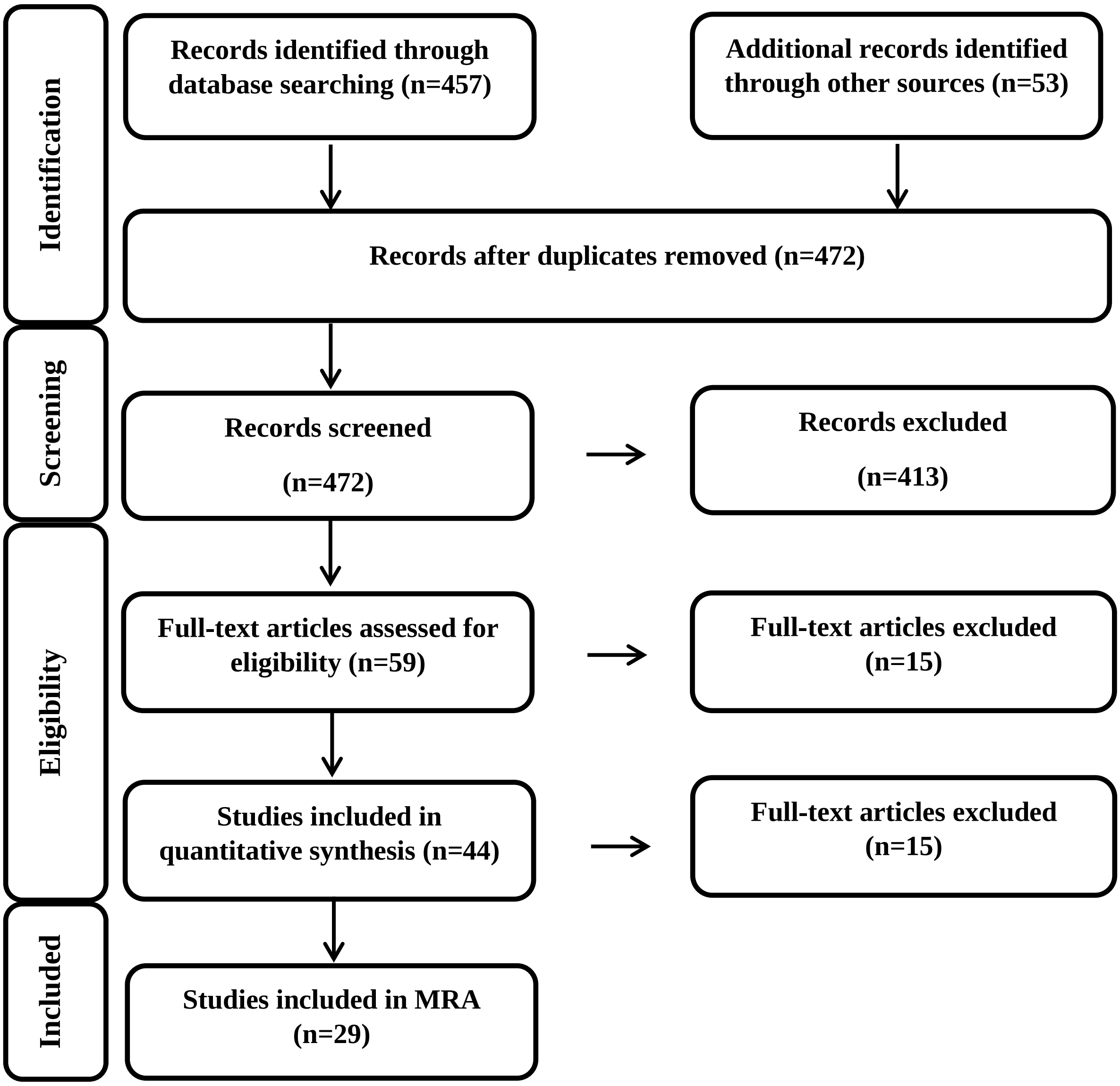
Reed, W. R., & Rogers, C. L. (2000). Measurement error and endogeneity in studies of state tax policy and economic growth. *Department of Economics (University of Oklahoma)*

Rhee, T. H. (2008). Macroeconomic effects of progressive taxation. Retrieved from http://www.aeaweb.org/aea/2013conference/program/retrieve.php?pdfid=394

Appendix 3.4: Final Sample of Studies

ID	Study	Publication Status	Number of estimates
1	Arkes (2014)	Working Paper	184
2	Prillaman and Meier (2014)	Journal	68
3	Bradly (2007)	Thesis	1
4	Chernick (1997)	Journal	4
5	Chernick (2010)	Journal	3
6	Clarke and Miller (2013)	Working Paper	180
7	Feng and Young (2003)	Working Paper	28
8	Goff, Lebedinsky and Lile (2012)	Journal	12
9	Mc Cracken (2006)	Thesis	14
10	Miller and Russek (1997)	Journal	50
11	Mullen and Williams (1994)	Journal	22
12	Ojede, atems and Yamarik (2014)	Working Paper	18
13	Pjesky (2002)	Thesis	27
14	Reed (2008)	Journal	55
15	Rhee (2012)	Working Paper	40
16	Srithongrung (2013)	Working Paper	6
17	Tomljanovich (2004)	Journal	33
18	Vedder (1990, 1996)	Journal + Working Paper	7
19	Wibow (2003)	Thesis	24
20	Yamarik (2000)	Journal	5
21	Yu, Wallace and Nardinelli (1991)	Journal	20
22	ECIN-Jan-2014-1240 (2014)	Working Paper	22
23	Helms (1985)	Journal	12
24	Bauer et l. (2012)	Journal	6
25	Rasmussen and Zuehlke (1990)	Journal	4
26	Caporale and Leirer (2010)	Journal	5
27	Goetz et al. (2010)	Journal	2
28	Gale et al. (2015)	Journal	94
29	Biswas et al. (2015)	Working Paper	20

Appendix 3.5: Litearure Search Flow Diagram



Appendix 3.6: List of States of the United States

ID	***State***	***State Abbreviations***
1	Alaska	AK
2	Alabama	AL
3	Arkansas	AR
4	Arizona	AZ
5	California	CA
6	Colorado	CO
7	Connecticut	CT
8	Delaware	DE
9	Florida	FL
10	Georgia	GA
11	Hawaii	HI
12	Iowa	IA
13	Idaho	ID
14	Illinois	IL
15	Indiana	IN
16	Kansas	KS
17	Kentucky	KY
18	Louisiana	LA
19	Massachusetts	MA
20	Maryland	MD
21	Maine	ME
22	Michigan	MI
23	Minnesota	MN
24	Missouri	MO
25	Mississippi	MS
26	Montana	MT
27	North Carolina	NC
28	North Dakota	ND
29	Nebraska	NE
30	New Hampshire	NH
31	New Jersey	NJ
32	New Mexico	NM

ID	***State***	***State Abbreviations***
33	Nevada	NV
34	New York	NY
35	Ohio	OH
36	Oklahoma	OK
37	Oregon	OR
38	Pennsylvania	PA
39	Rhode Island	RI
40	South Carolina	SC
41	South Dakota	SD
42	Tennessee	TN
43	Texas	TX
44	Utah	UT
45	Virginia	VA
46	Vermont	VT
47	Washington	WA
48	Wisconsin	WI
49	West Virginia	WV
50	Wyoming	WY
26	Montana	MT

Appendix 3.7: Stata Codes

.do file for TABLE 3.4

```
log using "\\file\UsersN$\nal53\Home\Desktop\(Nazila's program)\US\SampleMeans.smcl", replace
set more off
set type double
graph drop _all
import excel "\\file\UsersN$\nal53\Home\Desktop\(Nazila's program)\US\TAX.xlsx", sheet("Stata") firstrow
case (lower)
//TABLE 4-Descriptive Statistica, Column 1(Estimated Tax Effects)-Full
summ coefficient, detail
//TABLE 4-Descriptive Statistica, Column 3(t-statistics)-Full
summ tstatistics, detail
//Mean absolute value of t-statitics
gen abststat=abs(tstat)
summ abststat
//TABLE 4-Descriptive Statistica, Column 2(Estimated Tax Effects)-Truncated
summ coefficient, detail
scalar low = r(p5)
scalar high = r(p95)
gen endog = (tsls == 1 | gmm == 1)
keep if coefficient > low & coefficient < high
summ coefficient, detail
//TABLE 4-Descriptive Statistica, Column 4(t-statistics)-Truncated
summ tstatistics, detail
//Mean absolute value of t-statitics
replace abststat=abs(tstat)
summ abststat
log close
```

.do file for TABLE 3.5

```
clear
log using "\\file\UsersN$\nal53\Home\Desktop\(Nazila's program)\US\Part1 Results.smcl", replace
set more off
set type double
graph drop _all
import excel "\\file\UsersN$\nal53\Home\Desktop\(Nazila's program)\US\TAX.xlsx", sheet("Stata") firstrow
case (lower)
summ coefficient, detail
scalar low = r(p5)
scalar high = r(p95)
summ tstat, detail
summ coefficient tstat if coefficient > low & coefficient < high, detail

keep if coefficient > low & coefficient < high

// Generating transformed variables for FE and RE
gen feprecision = 1/se
gen fetstat = coefficient/se

*------------------------------------------------------------------------------*
*   TABLE 5: FAT/PET                                                          *
*------------------------------------------------------------------------------*

// Fixed Effects with SE
//NOTE: If the coefficient on the constant term is significant, that
//is evidence of publication bias
// Fixed Effects
//This regression gives equal weight to each observation
// TABLE 5 - Column 1
regress fetstat feprecision, vce(cluster idstudy)
//This regression gives equal weight to each study
// TABLE 5 - Column 2
regress fetstat feprecision [pweight = weight], vce(cluster idstudy)
// Fixed Effects without SE
//NOTE: If the coefficient on the constant term is significant, that
//is evidence of publication bias
// Fixed Effects
//This regression gives equal weight to each observation
// TABLE 5 - (not reported)
regress fetstat feprecision, noc vce(cluster idstudy)
//This regression gives equal weight to each study
// TABLE 5 - (not reported)
regress fetstat feprecision [pweight = weight], noc vce(cluster idstudy)


metan coefficient se, random
scalar tau2 = r(tau2)
gen revar = se^2 + tau2
gen rese = sqrt(revar)
gen reprecision = 1/rese
gen retstat = coefficient/rese

// Random Effects with SE
//This regression gives equal weight to each observation
```

```
// TABLE 5 - Column 3
regress retstat reprecision,  vce(cluster idstudy)
//This regression gives equal weight to each study
// TABLE 5 - Column 4
regress retstat reprecision [pweight = weight], vce(cluster idstudy)
// TABLE 5 - Column 5
regress retstat reprecision, noc vce(cluster idstudy)
//This regression gives equal weight to each study
// TABLE 5 - Column 6
regress retstat reprecision [pweight = weight], noc vce(cluster idstudy)

log close
```

.do file for TABLE 3.6

```
clear
log using "\\file\UsersN$\nal53\Home\Desktop\(Nazila's program)\US\Part1 Results.smcl", replace
set more off
set type double
graph drop _all
import excel "\\file\UsersN$\nal53\Home\Desktop\(Nazila's program)\US\TAX.xlsx", sheet("Stata") firstrow
case (lower)
summ coefficient, detail
scalar low = r(p5)
scalar high = r(p95)

keep if coefficient > low & coefficient < high
gen endog = (tsls == 1 | gmm == 1)
summ predneg predother predpos ///
labourtax capitaltax consumptiontax othertaxes mixedtaxes propertytax overalltax ///
none ak dc akdc akhidc akhidcothers ///
gsp pcgsp pi pcpi ///
marginal differenced etr ///
lrcase1 lrcase2 lrcase3 ///
peerreviewed pubyear ///
cs panel ///
length midyear ///
ols gls endog ///
seols sehet sehac ///
income laggeddv statefe hlk
```

.do file for TABLE 3.7

```
clear
log using "\\file\UsersN$\nal53\Home\Desktop\(Nazila's program)\US\Part1 Results.smcl", replace
set more off
set type double
graph drop _all
import excel "\\file\UsersN$\nal53\Home\Desktop\(Nazila's program)\US\TAX.xlsx", sheet("Stata") firstrow
case (lower)
quietly summ coefficient, detail
scalar low = r(p5)
scalar high = r(p95)
keep if coefficient > low & coefficient < high

gen endog = (tsls == 1 | gmm == 1)

// Generating transformed variables for FE and RE
gen feprecision = 1/se
gen fetstat = coefficient/se
gen peerreviewedd = peerreviewed/se
gen pubyearr = pubyear/se
gen css = cs/se
gen lengthh = length/se
gen midyearr = midyear/se
gen gspp = gsp/se
gen pcgspp = pcgsp/se
gen pii = pi/se
gen labourtaxx = labourtax/se
gen capitaltaxx = capitaltax/se
gen propertytaxx = propertytax/se
gen overalltaxx = overalltax/se
gen othertaxess = othertaxes/se
gen mixedtaxess = mixedtaxes/se
gen marginall = marginal/se
gen differencedd = differenced/se
gen etrr = etr/se
gen prednegg = predneg/se
gen predposs = predpos/se
gen hlkk = hlk/se
gen incomee = income/se
gen laggeddvv = laggeddv/se
gen statefee = statefe/se
gen sehacc = sehac/se
gen sehett = sehet/se
gen lrcase22 = lrcase2/se
gen lrcase33 = lrcase3/se
gen glss = gls/se
gen endogg = endog/se
gen nonee = none/se
gen akk = ak/se
gen dcc = dc/se
gen akdcc = akdc/se
gen akhidcotherss = akhidcothers/se
```

```
*********************************************
************ FIXED EFFECTS ********************
*********************************************

// These specifications include the SeR term
// NOTE: The constant term is the SER term

//THIS SECTION GIVES EQUAL WEIGHT TO EACH OBSERVATION
// Basic regression with no control variables
reg fetstat feprecision prednegg predposs,  vce(cluster idstudy)
display e(r2_a)

//TABLE 7, Collumn 1, Top Panel
// Regression with all control variables
reg fetstat feprecision prednegg predposs ///
peerreviewedd pubyearr css lengthh midyearr gspp pcgspp pii marginall differencedd etrr  ///
hlkk incomee laggeddvv statefee sehacc sehett lrcase22 lrcase33 ///
glss endogg nonee akk dcc akdcc akhidcotherss,  vce(cluster idstudy)
display e(r2_a)

// Here we use backwards stepwise regression selecting on the best regression using the BIC criterion
// We lock in the variables: feprecision labourtaxx capitaltaxx consumptiontaxx othertaxess prednegg predposs
vselect fetstat peerreviewedd pubyearr css lengthh midyearr gspp pcgspp pii marginall differencedd etrr  ///
hlkk incomee laggeddvv statefee sehacc sehett lrcase22 lrcase33 ///
glss endogg nonee akk dcc akdcc akhidcotherss, backward bic ///
fix(feprecision prednegg predposs)

//TABLE 7, Collumn 1, Bottom Panel
// We take the results from the preceding stepwise regression and reestimate the best model
// but this time with cluster robust standard errors
reg fetstat feprecision   ///
prednegg predposs pubyearr lengthh midyearr differencedd ///
laggeddvv sehett lrcase22 akk dcc akdcc,  vce(cluster idstudy)
display e(r2_a)


//THIS SECTION GIVES EQUAL WEIGHT TO EACH STUDY
// Basic regression with no control variables
reg fetstat feprecision prednegg predposs  [pweight = weight], vce(cluster idstudy)
display e(r2_a)

//TABLE 7, Collumn 2, Top Panel
// Regression with all control variables
reg fetstat feprecision prednegg predposs ///
peerreviewedd pubyearr css lengthh midyearr gspp pcgspp pii marginall differencedd etrr  ///
hlkk incomee laggeddvv statefee sehacc sehett lrcase22 lrcase33 ///
glss endogg nonee akk dcc akdcc akhidcotherss [pweight = weight],  vce(cluster idstudy)
display e(r2_a)

// Here we use backwards stepwise regression selecting on the best regression using the BIC criterion
// We lock in the variables: feprecision labourtaxx capitaltaxx consumptiontaxx othertaxess prednegg predposs
vselect fetstat peerreviewedd pubyearr css lengthh midyearr gspp pcgspp pii marginall differencedd etrr  ///
hlkk incomee laggeddvv statefee sehacc sehett lrcase22 lrcase33 ///
glss endogg nonee akk dcc akdcc akhidcotherss  [pweight = weight], backward bic ///
fix(feprecision prednegg predposs)
```

```
//TABLE 7, Collumn 2, Bottom Panel
// We take the results from the preceding stepwise regression and reestimate the best model
// but this time with cluster robust standard errors
reg fetstat feprecision   ///
prednegg predposs pubyearr lengthh midyearr differencedd hlkk ///
laggeddvv sehacc sehett glss dcc akdcc [pweight = weight], vce(cluster idstudy)
display e(r2_a)


*********************************************
************ RANDOM EFFECTS ******************
*********************************************

metareg coefficient predneg predpos, wsse(se)
scalar tau2 = e(tau2)
display tau2
gen revar = se^2 + tau2
gen rese = sqrt(revar)

gen reprecision = 1/rese
gen retstat = coefficient/rese
replace peerreviewedd = peerreviewed/rese
replace pubyearr = pubyear/rese
replace css = cs/rese
replace lengthh = length/rese
replace midyearr = midyear/rese
replace gspp = gsp/rese
replace pcgspp = pcgsp/rese
replace pii = pi/rese
replace labourtaxx = labourtax/rese
replace capitaltaxx = capitaltax/rese
replace propertytaxx = propertytax/rese
replace overalltaxx = overalltax/rese
replace othertaxess = othertaxes/rese
replace mixedtaxess = mixedtaxes/rese
replace marginall = marginal/rese
replace differencedd = differenced/rese
replace etrr = etr/rese
replace prednegg = predneg/rese
replace predposs = predpos/rese
replace hlkk = hlk/rese
replace incomee = income/rese
replace laggeddvv = laggeddv/rese
replace statefee = statefe/rese
replace sehacc = sehac/rese
replace sehett = sehet/rese
replace lrcase22 = lrcase2/rese
replace lrcase33 = lrcase3/rese
replace glss = gls/rese
replace endogg = endog/rese
replace nonee = none/rese
replace akk = ak/rese
replace dcc = dc/rese
replace akdcc = akdc/rese
replace akhidcotherss = akhidcothers/rese
```

```
//THIS SECTION GIVES EQUAL WEIGHT TO EACH OBSERVATION
// Basic regression with no control variables
reg retstat reprecision prednegg predposs,  vce(cluster idstudy)
display e(r2_a)

//TABLE 7, Collumn 3, Top Panel
// Regression with all control variables
reg retstat reprecision prednegg predposs ///
peerreviewedd pubyearr css lengthh midyearr gspp pcgspp pii marginall differencedd etrr  ///
hlkk incomee laggeddvv statefee sehacc sehett lrcase22 lrcase33 ///
glss endogg nonee akk dcc akdcc akhidcotherss,  vce(cluster idstudy)
display e(r2_a)

// Here we use backwards stepwise regression selecting on the best regression using the BIC criterion
// We lock in the variables: feprecision labourtaxx capitaltaxx consumptiontaxx othertaxess prednegg predposs
vselect retstat peerreviewedd pubyearr css lengthh midyearr gspp pcgspp pii marginall differencedd etrr  ///
hlkk incomee laggeddvv statefee sehacc sehett lrcase22 lrcase33 ///
glss endogg nonee akk dcc akdcc akhidcotherss, backward bic ///
fix(reprecision prednegg predposs)

//TABLE 7, Collumn 3, Bottom Panel
// We take the results from the preceding stepwise regression and reestimate the best model
// but this time with cluster robust standard errors
reg retstat reprecision  ///
prednegg predposs differencedd etrr lrcase33 ///
akk akdcc,  vce(cluster idstudy)
display e(r2_a)

//THIS SECTION GIVES EQUAL WEIGHT TO EACH STUDY
// Basic regression with no control variables
reg retstat reprecision prednegg predposs [pweight = weight], vce(cluster id)
display e(r2_a)

//TABLE 7, Collumn 4, Top Panel
// Regression with all control variables
reg retstat reprecision prednegg predposs ///
peerreviewedd pubyearr css lengthh midyearr gspp pcgspp pii marginall differencedd etrr  ///
hlkk incomee laggeddvv statefee sehacc sehett lrcase22 lrcase33 ///
glss endogg nonee akk dcc akdcc akhidcotherss [pweight = weight],  vce(cluster idstudy)
display e(r2_a)

// Here we use backwards stepwise regression selecting on the best regression using the BIC criterion
// We lock in the variables: feprecision labourtaxx capitaltaxx consumptiontaxx othertaxess prednegg predposs
vselect retstat peerreviewedd pubyearr css lengthh midyearr gspp pcgspp pii marginall differencedd etrr  ///
hlkk incomee laggeddvv statefee sehacc sehett lrcase22 lrcase33 ///
glss endogg nonee akk dcc akdcc akhidcotherss [pweight = weight], backward bic ///
fix(reprecision prednegg predposs)

//TABLE 7, Collumn 4, Bottom Panel
// We take the results from the preceding stepwise regression and reestimate the best model
// but this time with cluster robust standard errors
reg retstat reprecision ///
prednegg predposs pubyearr css midyearr pii ///
differencedd statefee glss akk akdcc [pweight = weight], vce(cluster id)
display e(r2_a)

log close
```

.do file for TABLE 3.8

```
clear
log using "\\file\UsersN$\nal53\Home\Desktop\(Nazila's program)\US\Type of Tax Results.smcl", replace
set more off
set type double
graph drop _all
import excel "\\file\UsersN$\nal53\Home\Desktop\(Nazila's program)\US\TAX.xlsx", sheet("Stata") firstrow
case (lower)
quietly summ coefficient, detail
scalar low = r(p5)
scalar high = r(p95)
keep if coefficient > low & coefficient < high

gen endog = (tsls == 1 | gmm == 1)

// Generating transformed variables for FE and RE
gen feprecision = 1/se
gen fetstat = coefficient/se
gen peerreviewedd = peerreviewed/se
gen pubyearr = pubyear/se
gen css = cs/se
gen lengthh = length/se
gen midyearr = midyear/se
gen gspp = gsp/se
gen pcgspp = pcgsp/se
gen pii = pi/se
gen labourtaxx = labourtax/se
gen capitaltaxx = capitaltax/se
gen propertytaxx = propertytax/se
gen overalltaxx = overalltax/se
gen othertaxess = othertaxes/se
gen mixedtaxess= mixedtaxes/se
gen marginall = marginal/se
gen differencedd = differenced/se
gen etrr = etr/se
gen prednegg = predneg/se
gen predposs = predpos/se
gen hlkk = hlk/se
gen incomee = income/se
gen laggeddvv = laggeddv/se
gen statefee = statefe/se
gen sehacc = sehac/se
gen sehett = sehet/se
gen lrcase22 = lrcase2/se
gen lrcase33 = lrcase3/se
gen glss = gls/se
gen endogg = endog/se
gen nonee = none/se
gen akk = ak/se
gen dcc = dc/se
gen akdcc = akdc/se
gen akhidcotherss = akhidcothers/se
```

```
*********************************************
************ FIXED EFFECTS ********************
*********************************************

// These specifications include the SeR term
// NOTE: The constant term is the SER term

//THIS SECTION GIVES EQUAL WEIGHT TO EACH OBSERVATION
// Basic regression with no control variables
reg fetstat feprecision labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess overalltaxx  ///
,  vce(cluster idstudy)
display e(r2_a)

//TABLE 8, Collumn 1, Top Panel
// Regression with all control variables
reg fetstat feprecision labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess overalltaxx  ///
peerreviewedd pubyearr css lengthh midyearr gspp pcgspp pii marginall differencedd etrr  ///
hlkk incomee laggeddvv statefee sehacc sehett lrcase22 lrcase33 ///
glss endogg nonee akk dcc akdcc akhidcotherss,  vce(cluster idstudy)
display e(r2_a)

// Here we use backwards stepwise regression selecting on the best regression using the BIC criterion
// We lock in the variables: feprecision labourtaxx capitaltaxx consumptiontaxx othertaxess prednegg predposs
vselect fetstat peerreviewedd pubyearr css lengthh midyearr gspp pcgspp pii marginall differencedd etrr  ///
hlkk incomee laggeddvv statefee sehacc sehett lrcase22 lrcase33 ///
glss endogg nonee akk dcc akdcc akhidcotherss, backward bic ///
fix(feprecision labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess overalltaxx )

//TABLE 8, Collumn 1, Bottom Panel
// We take the results from the preceding stepwise regression and reestimate the best model
// but this time with cluster robust standard errors
reg fetstat feprecision labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess overalltaxx  ///
pubyearr lengthh midyearr differencedd hlkk incomee laggeddvv ///
sehacc sehett akk dcc akdcc,  vce(cluster idstudy)
display e(r2_a)


//THIS SECTION GIVES EQUAL WEIGHT TO EACH STUDY
// Basic regression with no control variables
reg fetstat feprecision labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess overalltaxx  ///
 [pweight = weight], vce(cluster idstudy)
display e(r2_a)

//TABLE 8, Collumn 2, Top Panel
// Regression with all control variables
reg fetstat feprecision labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess overalltaxx  ///
peerreviewedd pubyearr css lengthh midyearr gspp pcgspp pii marginall differencedd etrr  ///
hlkk incomee laggeddvv statefee sehacc sehett lrcase22 lrcase33 ///
glss endogg nonee akk dcc akdcc akhidcotherss [pweight = weight],  vce(cluster idstudy)
display e(r2_a)

// Here we use backwards stepwise regression selecting on the best regression using the BIC criterion
// We lock in the variables: feprecision labourtaxx capitaltaxx consumptiontaxx othertaxess prednegg predposs
vselect fetstat peerreviewedd pubyearr css lengthh midyearr gspp pcgspp pii marginall differencedd etrr  ///
hlkk incomee laggeddvv statefee sehacc sehett lrcase22 lrcase33 ///
glss endogg nonee akk dcc akdcc akhidcotherss [pweight = weight], backward bic ///
```

```
fix(feprecision labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess overalltaxx)

//TABLE 8, Collumn 2, Bottom Panel
// We take the results from the preceding stepwise regression and reestimate the best model
// but this time with cluster robust standard errors
reg fetstat feprecision labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess overalltaxx  ///
pubyearr lengthh midyearr differencedd laggeddvv sehett dcc akdcc [pweight = weight], vce(cluster idstudy)
display e(r2_a)

*********************************************
************ RANDOM EFFECTS ******************
*********************************************

metareg coefficient labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess overalltaxx  ///
, wsse(se)
scalar tau2 = e(tau2)
display tau2
gen revar = se^2 + tau2
gen rese = sqrt(revar)

gen reprecision = 1/rese
gen retstat = coefficient/rese
replace peerreviewedd = peerreviewed/rese
replace pubyearr = pubyear/rese
replace css = cs/rese
replace lengthh = length/rese
replace midyearr = midyear/rese
replace gspp = gsp/rese
replace pcgspp = pcgsp/rese
replace pii = pi/rese
replace labourtaxx = labourtax/rese
replace capitaltaxx = capitaltax/rese
replace propertytaxx = propertytax/rese
replace overalltaxx = overalltax/rese
replace othertaxess = othertaxes/rese
replace mixedtaxess = mixedtaxes/rese
replace marginall = marginal/rese
replace differencedd = differenced/rese
replace etrr = etr/rese
replace prednegg = predneg/rese
replace predposs = predpos/rese
replace hlkk = hlk/rese
replace incomee = income/rese
replace laggeddvv = laggeddv/rese
replace statefee = statefe/rese
replace sehacc = sehac/rese
replace sehett = sehet/rese
replace lrcase22 = lrcase2/rese
replace lrcase33 = lrcase3/rese
replace glss = gls/rese
replace endogg = endog/rese
replace nonee = none/rese
replace akk = ak/rese
replace dcc = dc/rese
replace akdcc = akdc/rese
replace akhidcotherss = akhidcothers/rese
```

```
//THIS SECTION GIVES EQUAL WEIGHT TO EACH OBSERVATION
// Basic regression with no control variables
reg retstat reprecision labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess overalltaxx  ///
,  vce(cluster idstudy)
display e(r2_a)

//TABLE 8, Collumn 3, Top Panel
// Regression with all control variables
reg retstat reprecision labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess overalltaxx  ///
peerreviewedd pubyearr css lengthh midyearr gspp pcgspp pii marginall differencedd etrr  ///
hlkk incomee laggeddvv statefee sehacc sehett lrcase22 lrcase33 ///
glss endogg nonee akk dcc akdcc akhidcotherss,  vce(cluster idstudy)
display e(r2_a)

// Here we use backwards stepwise regression selecting on the best regression using the BIC criterion
// We lock in the variables: feprecision labourtaxx capitaltaxx consumptiontaxx othertaxess prednegg predposs
vselect retstat peerreviewedd pubyearr css lengthh midyearr gspp pcgspp pii marginall differencedd etrr  ///
hlkk incomee laggeddvv statefee sehacc sehett lrcase22 lrcase33 ///
glss endogg nonee akk dcc akdcc akhidcotherss, backward bic ///
fix(reprecision labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess overalltaxx)

//TABLE 8, Collumn 3, Bottom Panel
// We take the results from the preceding stepwise regression and reestimate the best model
// but this time with cluster robust standard errors
reg retstat reprecision labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess overalltaxx  ///
marginall differencedd akk akdcc,  vce(cluster idstudy)
display e(r2_a)

//THIS SECTION GIVES EQUAL WEIGHT TO EACH STUDY
// Basic regression with no control variables
reg retstat reprecision labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess overalltaxx  ///
 [pweight = weight], vce(cluster id)
display e(r2_a)

//TABLE 8, Collumn 4, Top Panel
//Regression with all control variables
reg retstat reprecision labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess overalltaxx  ///
peerreviewedd pubyearr css lengthh midyearr gspp pcgspp pii marginall differencedd etrr  ///
hlkk incomee laggeddvv statefee sehacc sehett lrcase22 lrcase33 ///
glss endogg nonee akk dcc akdcc akhidcotherss [pweight = weight], vce(cluster idstudy)
display e(r2_a)

// Here we use backwards stepwise regression selecting on the best regression using the BIC criterion
// We lock in the variables: feprecision labourtaxx capitaltaxx consumptiontaxx othertaxess prednegg predposs
vselect retstat peerreviewedd pubyearr css lengthh midyearr gspp pcgspp pii marginall differencedd etrr  ///
hlkk incomee laggeddvv statefee sehacc sehett lrcase22 lrcase33 ///
glss endogg nonee akk dcc akdcc akhidcotherss [pweight = weight], backward bic ///
fix(reprecision labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess overalltaxx)

//TABLE 8, Collumn 4, Bottom Panel
// We take the results from the preceding stepwise regression and reestimate the best model
// but this time with cluster robust standard errors
reg retstat reprecision labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess overalltaxx  ///
peerreviewedd pubyearr css midyearr pcgspp pii marginall differencedd sehacc sehett glss akhidcotherss
[pweight = weight], vce(cluster id)
display e(r2_a)
log close
```

.do file for TABLE 3.9

```
clear
log using "\\file\UsersN$\nal53\Home\Desktop\(Nazila's program)\US\All Tax Var Results.smcl", replace
set more off
set type double
graph drop _all
import excel "\\file\UsersN$\nal53\Home\Desktop\(Nazila's program)\US\TAX.xlsx", sheet("Stata") firstrow
case (lower)
quietly summ coefficient, detail
scalar low = r(p5)
scalar high = r(p95)
keep if coefficient > low & coefficient < high

gen endog = (tsls == 1 | gmm == 1)

// Generating transformed variables for FE and RE
gen feprecision = 1/se
gen fetstat = coefficient/se
gen peerreviewedd = peerreviewed/se
gen pubyearr = pubyear/se
gen css = cs/se
gen lengthh = length/se
gen midyearr = midyear/se
gen gspp = gsp/se
gen pcgspp = pcgsp/se
gen pii = pi/se
gen labourtaxx = labourtax/se
gen capitaltaxx = capitaltax/se
gen propertytaxx = propertytax/se
gen overalltaxx = overalltax/se
gen othertaxess = othertaxes/se
gen mixedtaxess= mixedtaxes/se
gen marginall = marginal/se
gen differencedd = differenced/se
gen etrr = etr/se
gen prednegg = predneg/se
gen predposs = predpos/se
gen hlkk = hlk/se
gen incomee = income/se
gen laggeddvv = laggeddv/se
gen statefee = statefe/se
gen sehacc = sehac/se
gen sehett = sehet/se
gen lrcase22 = lrcase2/se
gen lrcase33 = lrcase3/se
gen glss = gls/se
gen endogg = endog/se
gen nonee = none/se
gen akk = ak/se
gen dcc = dc/se
gen akdcc = akdc/se
gen akhidcotherss = akhidcothers/se
```

```
*********************************************
************ FIXED EFFECTS ********************
*********************************************
// These specifications include the SeR term
// NOTE: The constant term is the SER term

//THIS SECTION GIVES EQUAL WEIGHT TO EACH OBSERVATION
// Basic regression with no control variables
reg fetstat feprecision prednegg predposs labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess
overalltaxx  ///
,  vce(cluster idstudy)
display e(r2_a)

//TABLE 9, Collumn 1, Top Panel
// Regression with all control variables
reg fetstat feprecision prednegg predposs labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess
overalltaxx  ///
peerreviewedd pubyearr css lengthh midyearr gspp pcgspp pii marginall differencedd etrr  ///
hlkk incomee laggeddvv statefee sehacc sehett lrcase22 lrcase33 ///
glss endogg nonee akk dcc akdcc akhidcotherss,  vce(cluster idstudy)
display e(r2_a)

// Here we use backwards stepwise regression selecting on the best regression using the BIC criterion
// We lock in the variables: feprecision labourtaxx capitaltaxx consumptiontaxx othertaxess prednegg predposs
vselect fetstat peerreviewedd pubyearr css lengthh midyearr gspp pcgspp pii marginall differencedd etrr  ///
hlkk incomee laggeddvv statefee sehacc sehett lrcase22 lrcase33 ///
glss endogg nonee akk dcc akdcc akhidcotherss, backward bic ///
fix(feprecision prednegg predposs labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess overalltaxx)

//TABLE 9, Collumn 1, Bottom Panel
// We take the results from the preceding stepwise regression and reestimate the best model
// but this time with cluster robust standard errors
reg fetstat feprecision prednegg predposs labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess
overalltaxx  ///
peerreviewedd pubyearr lengthh differencedd ///
laggeddvv sehett lrcase22 akk dcc akdcc ,  vce(cluster idstudy)
display e(r2_a)

//THIS SECTION GIVES EQUAL WEIGHT TO EACH STUDY
// Basic regression with no control variables
reg fetstat feprecision prednegg predposs labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess
overalltaxx  ///
 [pweight = weight], vce(cluster idstudy)
display e(r2_a)

//TABLE 9, Collumn 2, Top Panel
// Regression with all control variables
reg fetstat feprecision prednegg predposs labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess
overalltaxx  ///
peerreviewedd pubyearr css lengthh midyearr gspp pcgspp pii marginall differencedd etrr  ///
hlkk incomee laggeddvv statefee sehacc sehett lrcase22 lrcase33 ///
glss endogg nonee akk dcc akdcc akhidcotherss [pweight = weight],  vce(cluster idstudy)
display e(r2_a)

// Here we use backwards stepwise regression selecting on the best regression using the BIC criterion
// We lock in the variables: feprecision labourtaxx capitaltaxx consumptiontaxx othertaxess prednegg predposs
```

```
vselect fetstat peerreviewedd pubyearr css lengthh midyearr gspp pcgspp pii marginall differencedd etrr  ///
hlkk incomee laggeddvv statefee sehacc sehett lrcase22 lrcase33 ///
glss endogg nonee akk dcc akdcc akhidcotherss [pweight = weight], backward bic ///
fix(feprecision prednegg predposs labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess overalltaxx)

//TABLE 9, Collumn 2, Bottom Panel
//We take the results from the preceding stepwise regression and reestimate the best model
// but this time with cluster robust standard errors
reg fetstat feprecision prednegg predposs labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess
overalltaxx  ///
pubyearr lengthh midyearr marginall differencedd hlkk laggeddvv sehacc sehett glss ///
dcc  [pweight = weight], vce(cluster idstudy)
display e(r2_a)

**********************************************
************ RANDOM EFFECTS ******************
**********************************************
metareg coefficient labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess overalltaxx  ///
predneg predpos, wsse(se)
scalar tau2 = e(tau2)
display tau2
gen revar = se^2 + tau2
gen rese = sqrt(revar)

gen reprecision = 1/rese
gen retstat = coefficient/rese
replace peerreviewedd = peerreviewed/rese
replace pubyearr = pubyear/rese
replace css = cs/rese
replace lengthh = length/rese
replace midyearr = midyear/rese
replace gspp = gsp/rese
replace pcgspp = pcgsp/rese
replace pii = pi/rese
replace labourtaxx = labourtax/rese
replace capitaltaxx = capitaltax/rese
replace propertytaxx = propertytax/rese
replace overalltaxx = overalltax/rese
replace othertaxess = othertaxes/rese
replace mixedtaxess = mixedtaxes/rese
replace marginall = marginal/rese
replace differencedd = differenced/rese
replace etrr = etr/rese
replace prednegg = predneg/rese
replace predposs = predpos/rese
replace hlkk = hlk/rese
replace incomee = income/rese
replace laggeddvv = laggeddv/rese
replace statefee = statefe/rese
replace sehacc = sehac/rese
replace sehett = sehet/rese
replace lrcase22 = lrcase2/rese
replace lrcase33 = lrcase3/rese
replace glss = gls/rese
replace endogg = endog/rese
replace nonee = none/rese
replace akk = ak/rese
```

```
replace dcc = dc/rese
replace akdcc = akdc/rese
replace akhidcotherss = akhidcothers/rese

//THIS SECTION GIVES EQUAL WEIGHT TO EACH OBSERVATION
// Basic regression with no control variables
reg retstat reprecision prednegg predposs labourtaxx capitaltaxx propertytaxx othertaxess mixedtaxess
overalltaxx  ///
,  vce(cluster idstudy)
display e(r2_a)

//TABLE 9, Collumn 3, Top Panel
// Regression with all control variables
reg retstat reprecision prednegg predposs labourtaxx capitaltaxx propertytaxx mixedtaxess othertaxess
overalltaxx  ///
peerreviewedd pubyearr css lengthh midyearr gspp pcgspp pii marginall differencedd etrr  ///
hlkk incomee laggeddvv statefee sehacc sehett lrcase22 lrcase33 ///
glss endogg nonee akk dcc akdcc akhidcotherss,  vce(cluster idstudy)
display e(r2_a)

// Here we use backwards stepwise regression selecting on the best regression using the BIC criterion
// We lock in the variables: feprecision labourtaxx capitaltaxx consumptiontaxx othertaxess prednegg predposs
vselect retstat peerreviewedd pubyearr css lengthh midyearr gspp pcgspp pii marginall differencedd etrr  ///
hlkk incomee laggeddvv statefee sehacc sehett lrcase22 lrcase33 ///
glss endogg nonee akk dcc akdcc akhidcotherss, backward bic ///
fix(reprecision prednegg predposs labourtaxx capitaltaxx propertytaxx mixedtaxess othertaxess overalltaxx)

//TABLE 9, Collumn 3, Bottom Panel
// We take the results from the preceding stepwise regression and reestimate the best model
// but this time with cluster robust standard errors
reg retstat reprecision prednegg predposs labourtaxx capitaltaxx propertytaxx mixedtaxess othertaxess
overalltaxx  ///
marginall differencedd lrcase33 akk akdcc , vce(cluster idstudy)
display e(r2_a)

//THIS SECTION GIVES EQUAL WEIGHT TO EACH STUDY
// Basic regression with no control variables
reg retstat reprecision prednegg predposs labourtaxx capitaltaxx propertytaxx mixedtaxess othertaxess
overalltaxx  ///
 [pweight = weight], vce(cluster id)
display e(r2_a)

//TABLE 9, Collumn 4, Top Panel
// Regression with all control variables
reg retstat reprecision prednegg predposs labourtaxx capitaltaxx propertytaxx mixedtaxess othertaxess
overalltaxx  ///
peerreviewedd pubyearr css lengthh midyearr gspp pcgspp pii marginall differencedd etrr  ///
hlkk incomee laggeddvv statefee sehacc sehett lrcase22 lrcase33 ///
glss endogg nonee akk dcc akdcc akhidcotherss [pweight = weight], vce(cluster idstudy)
display e(r2_a)

// Here we use backwards stepwise regression selecting on the best regression using the BIC criterion
// We lock in the variables: feprecision labourtaxx capitaltaxx consumptiontaxx othertaxess prednegg predposs
vselect retstat peerreviewedd pubyearr css lengthh midyearr gspp pcgspp pii marginall differencedd etrr  ///
hlkk incomee laggeddvv statefee sehacc sehett lrcase22 lrcase33 ///
glss endogg nonee akk dcc akdcc akhidcotherss [pweight = weight], backward bic ///
fix(reprecision prednegg predposs labourtaxx capitaltaxx propertytaxx mixedtaxess othertaxess overalltaxx)
```

```
//TABLE 9, Collumn 4, Bottom Panel
// We take the results from the preceding stepwise regression and reestimate the best model
// but this time with cluster robust standard errors
reg retstat reprecision prednegg predposs labourtaxx capitaltaxx propertytaxx mixedtaxess othertaxess
overalltaxx  ///
pubyearr css midyearr pii differencedd statefee akk akdcc akhidcotherss [pweight = weight], vce(cluster id)
display e(r2_a)

log close
```

R Commands for TABLE 3.10

Download R from the following link:
https://cran.r-project.org/src/base/R-3/
The one I am applying is R-3.2.1.tar.gz

After opening up the R, type the following commands:
install.packages() → New Zealand → ok → BMS → ok

Library(BMS)

The data file should have the dependent variable as the first column.

Open the data file (Excel spreadsheet) → copy data

TAX1=read.table("clipboard-512", sep="\t", header=TRUE)

TAX11 = bms(TAX1, burn=10000000, iter=10000000, g="hyper", mprior="random", fixed.reg=c("Precision","Prediction_Negative","Prediction_Positive"), nmodel=1000, mcmc="bd", user.int=FALSE)

plot(TAX11)

summary(TAX11)

coef(TAX11, order.by.pip = T, exact=T, include.constant=T)

image(TAX11, cex.axis=0.7, order.by.pip = T, yprop2pip=F)

Chapter 4. Meta-Analysis and Publication Bias: How Well Does the FAT-PET-PEESE Procedure Work?

4.1. Introduction

Meta-regression analysis offers a statistical analysis through which conflicting theoretical and/or empirical findings on a given topic can be summarized and compared. Two main objectives of meta-regression analysis are (i) to reach a single conclusion about the magnitude and significance of the results, and (ii) to compare findings yielded from various studies and explain potential reasons for the heterogeneity observed among estimates. Meta-regression analysis has become an increasingly popular method in economics and business. Figure 4.1 depicts a time series bar chart that lists all Web of Science journal articles in economics and business that have the word "meta-analysis" in the title. The trend is clearly upward reflecting the fact that the number of studies applying this tool is increasing over time.

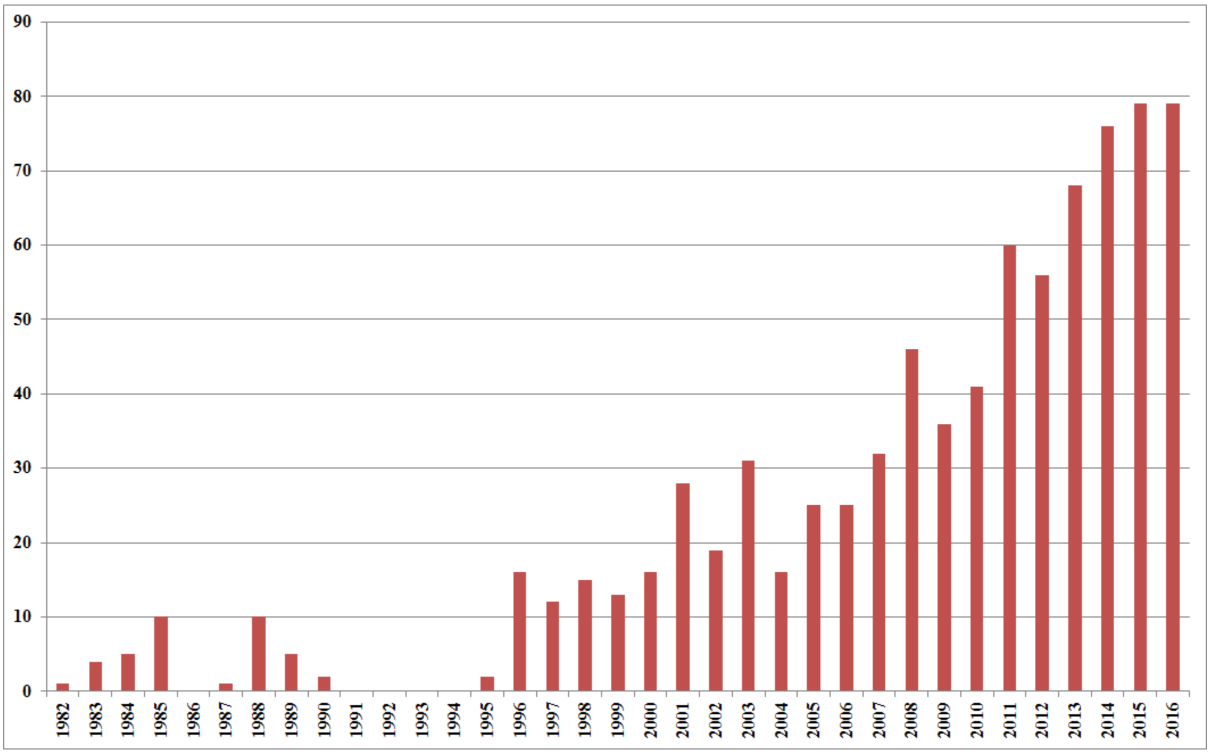


Figure 4.1: Number of Articles in Economics and Business Listed in Web of Science with "Meta-Analysis" in the Title

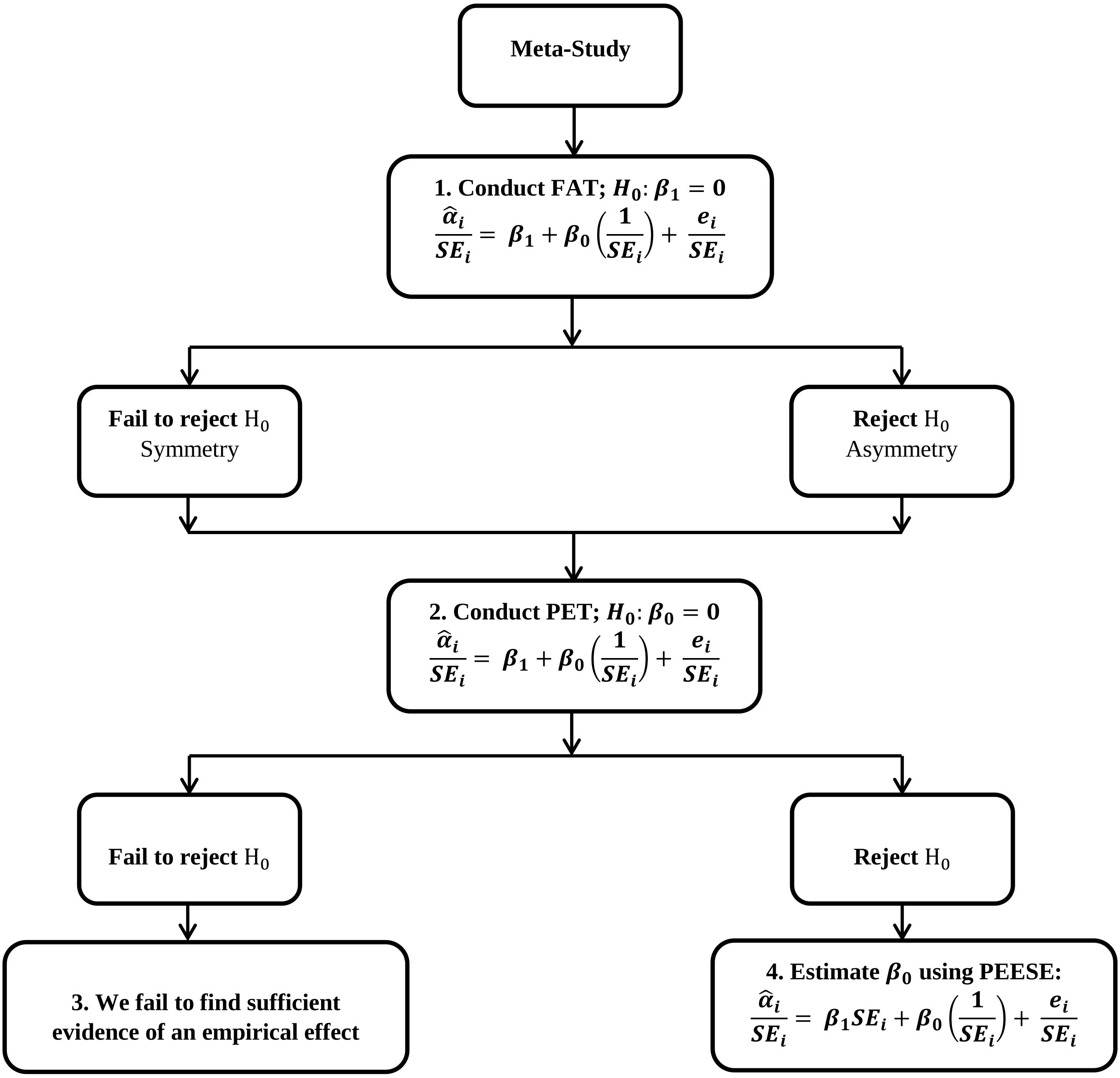
Note: Web of Science categories are: Economics, Business Finance, Business, Management, Criminology Penology, Urban Studies, and Social Sciences Interdisciplinary (813 articles).

It is well known that publication selection bias or "selectivity bias" can distort the distribution of estimated effects in the literature. Publication bias might happen because there is a tendency amongst researchers, reviewers, and editors to avoid reporting and publishing statistically insignificant estimates or estimates which are inconsistent with well-established theories. As a result, the true effect of the focal predictor on the response variable might be over- or under-estimated. An example of the second type of bias called directional publication bias was provided by Stanley (2005). He documented how the price elasticity of water demand is exaggerated fourfold in the literature as a direct result of publication selection bias. It is generally accepted that positive estimates of price elasticities of demand are inconsistent with theory.

The data used for meta-analysis consist of estimated effects from studies on a particular phenomenon. If the distribution of those effects is distorted, so will be any conclusion derived from them. It is therefore crucial to identify whether the literature on a given topic suffers from publication selection bias and if there is, how it should be corrected.

A common procedure for doing this in the economics and business literature is through the FAT-PET-PEESE procedure (Stanley and Doucouliagos, 2012; 2014a). Figure 4.2 shows the associated four steps procedure. The first is the Funnel Asymmetry Test (FAT) to test whether the sample of estimates is influenced by publication selection bias. It uses Weighted Least Squares (WLS) to regress the estimated effects $\left(\widehat{\alpha}_j\right)$ on a constant term (β_0) and the standard errors of the estimated effects $\left(\mathrm{SE_j}\right)$; where weights $\omega_i = \left(\frac{1}{\mathrm{SE_i}}\right)$ are applied to correct for heteroskedasticity in the estimates (which is inevitable in meta-regression analysis). If the estimated coefficient on the standard error variable, $\widehat{\beta}_1$, is statistically significant, then the estimates suffer from publication bias.

Figure 4.2: FAT-PET-PEESE Procedure



Source: Adapted from Stanley and Doucouliagos (2012, page 79).

Regardless of the results in the previous step, the next step is to conduct a Precision Effect Test (PET) to test whether there is a genuine, non-zero, true effect of estimates once publication bias is accommodated and corrected. It uses the same equation as the FAT, but tests whether $\beta_0 = 0$. If the SE_i variable were not included in the equation and if OLS was used rather than WLS, then the estimate of β_0 would simply be the arithmetic mean of the estimated effects in the literature. Thus, $\hat{\beta}_0$ is an estimate of the overall effect, and the PET tests $\hat{\beta}_0$ for statistical significance, correcting for publication bias.

If the PET fails to reject the null hypothesis of no effect, then $\hat{\beta}_0$ is taken as the estimate of overall effect with the understanding that it is not statistically significantly different from zero. In other words, there is not enough evidence to support the existence of any empirical effect. If the PET, however, rejects the null, then one concludes that there is a genuine non-zero true effect. In that case, one estimates a new specification known as the PEESE, or Precision Effect Estimate with Standard Error. The corresponding estimate of β_0 then becomes the "best" estimate of the overall effect.

Given the wide application of the FAT-PET-PEESE (FPP) procedure in the economics and business literature (e.g. Costa-Font; Gemmill, and Rubert (2011); Doucouliagos, Stanley, and Viscusi (2014); Doucouliagos and Paldam (2013); Efendic, Pugh, and Adnett (2011); Haelermans and Borghans (2012); Havránek (2010); Iwasaki and Tokunaga (2014); Laroche (2016); Lazzaroni and Van Bergeijik (2014); Linde Leonard, Stanley, and Doucouliagos (2014); and Nelson (2013)), it is surprising that there have not been any comparative evaulations of its performance. That is the purpose of this chapter.

The three objectives of this study are to evaluate how well the FAT-PET-PEESE procedure (i) correctly detects the existence of publication selection bias, (ii) correctly tests the existence of a genuine non-zero true effect, and (iii) compares with three common meta-analysis estimators that do not correct for publication bias.

I use Monte Carlo experiments to demonstrate that the FPP procedure does not perform well in the kind of statistical environments one would expect to encounter in economics and business. Section 4.2 describes my Monte Carlo experiments including associated terminology. Section 4.3 describes my experimental design and the simulated datasets used in my analysis, and also presents the results. Section 4.4 presents the main conclusions from this research.

4.2. Description of the Monte Carlo Experiments

4.2.1. Publication Bias

It is widely recognized that publication selection bias or "selectivity bias" distorts the distribution of estimated effects that appear in the literature. This arises because there is a tendency amongst researchers, reviewers, and editors to submit or accept manuscripts for publication based upon the direction or the strength of their results. Thus, it is unlikely that papers with statistically insignificant results or results which are not in line with an established theory could get published in a peer-reviewed journal. These studies usually end up sitting, unpublished, in file drawers of researchers. That is why this problem is called the "file-drawer problem" (Rosenthal, 1979).

Two popular types of publication bias modelled in these experiments are: (i) publication bias against insignificant results and (ii) publication bias against wrong-sign results. An example of the latter is a price elasticity where there is a strong presumption that the estimate should be negative, so that positive estimates will find it difficult to get published.

In my analysis, I model the first type of publication bias assuming that there is a tendency in favour of the strength of the results. Therefore, if the absolute values of the reported t-statistics associated with the estimates are greater than or equal to 2, then they will get published. Studies with insignificant estimates can still get published, but with a relatively small probability. For the second type of publication bias, I assume that theory predicts the "correct" sign should be positive. The publication process then works against negative estimates. While negative estimates can still get published, they can do so only with a relatively small probability.

4.2.2. Estimators

I use Monte Carlo experiments to compare the performance of seven different estimators. For the sake of comparison, I use estimators studied by Reed (2015).[37] However, the main focus of this study is on a new estimator, the estimator described above as part of the FAT-PET-PEESE procedure. I compare these estimators using three performance measures: Bias, Mean-Squared Error (MSE) as an efficiency test, and Type I error rates associated with testing whether the estimate of α equals its true value. The remainder of this section describes the respective estimators.

The "Unadjusted" estimator. The unadjusted estimator of the mean true effect of x on y is given by OLS estimates of β_0 in the following equation:

$$\hat{\alpha}_{i1} = \beta_0 + e_i, i = 1,2, \dots, M \quad (4.1)$$

where $\hat{\alpha}_{i1}$ is the ith estimated effect of y on x, and M is the number of estimates in the "Post-Publication" bias sample. The unadjusted estimator simply calculates the arithmetic mean of estimated effects across studies. It is used as a benchmark to compare the various meta-analysis (MA) estimators against.

The "Fixed Effects" (FE) estimator. Under this model I assume that there is one true underlying effect size and the only reason for the studies to obtain different estimated effect sizes is due solely to sampling error. This is why Borenstein et al. (2010) call this model the "common-effect model," which conveys the message more precisely. The fixed effect

[37] The conceptual design for my Monte Carlo experiments is based on Reed (2015), the replicated results are identical.

estimator weights all the observations by the inverse of the estimated standard error of $\hat{\alpha}_i$, SE_i. The FE estimator of the mean true effect is the weighted least squares estimate of β_0, except that the residuals are standardized to produce a sample variance of 1.

$$\frac{\hat{\alpha}_{i1}}{SE_i} = \beta_0\left(\frac{1}{SE_i}\right) + \frac{e_i}{SE_i}, i = 1,2,\ldots,M \quad (4.2)$$

<u>The "Weighted Least Squares" (WLS) estimator</u>. The WLS estimator is identical to the FE estimator except that the residual remains unstandardized. It is worthwhile to note that both FE and WLS estimators produce identical estimates of β_0, but the associated standard errors are different.

<u>The "Random Effects" (RE) estimator</u>. While the fixed effects models assume that there is one underlying true effect for all studies, this assumption seems quite implausible for most meta-analyses conducted in economics and business. Thus, under this model I assume that there is a distribution of true underlying effects and the goal is to estimate the mean of this distribution of true effects. The RE estimator is motivated by the assumption that differences in estimated effects across studies are due to (i) sampling error and also (ii) genuine differences in the underlying effects. The second component is represented by τ, which is the "between studies" variance. If the two variances are independent of each other, then,

$$SE(\hat{\alpha}_i) = \sqrt{(SE_i)^2 + \tau^2} = \omega_i \quad (4.3)$$

The RE estimator of the mean true effect is given by weighted least squares estimation of β_0, with weights equal to ω_i:

$$\frac{\hat{\alpha}_{i1}}{\omega_i} = \beta_0\left(\frac{1}{\omega_i}\right) + \frac{e_i}{\omega_i}, i = 1,2,\ldots,M \quad (4.4)$$

The "Precision Effect Testing" (PET) estimator. The PET estimator is designed to report the genuine underlying empirical effect after accommodating and correcting for publication bias. The PET adds the ith study's estimated standard error of the estimated effect, (SE_i), as an explanatory variable to control for publication bias. It then estimates the value of the mean effect as follows:

$$\hat{\alpha}_{i1} = \beta_0 + \beta_1 SE_i + e_i, i = 1,2, \dots, M \quad (4.5)$$

WLS estimation of β_0 provides an estimate of the mean true effect of x on y, weighting by $\left(\frac{1}{SE_i}\right)$, where SE_i is the same term used to correct for publication bias:

$$\frac{\hat{\alpha}_{i1}}{SE_i} = \beta_0 \left(\frac{1}{SE_i}\right) + \beta_1 + \frac{e_i}{SE_i}, i = 1,2, \dots, M \quad (4.6)$$

The "Precision-Effect Estimate with Standard Error" (PEESE) estimator. The PEESE estimator is designed to provide a better estimate of the actual empirical effect corrected for publication bias. What makes this estimator different from the PET is that it replaces SE_i with $(SE_i)^2$ in equation (4.5):

$$\hat{\alpha}_{i1} = \beta_0 + \beta_1 (SE_i)^2 + e_i, i = 1,2, \dots, M \quad (4.7)$$

This yields the following weighted least squares specification:

$$\frac{\hat{\alpha}_{i1}}{SE_i} = \beta_0 \left(\frac{1}{SE_i}\right) + \beta_1 SE_i + \frac{e_i}{SE_i}, i = 1,2, \dots, M \quad (4.8)$$

The last estimator, the FPP estimator, which is the main focus of this chapter, combines the "FAT" with elements of both the "PET" and "PEESE."

The "FAT-PET-PEESE" (FPP) estimator. Three main elements available in the "FAT-PET-PEESE" approach are summarized as follows: (i) identify the existence of publication bias (FAT); (ii) identify the presence of a genuine non-zero "true" effect (PET), corrected for publication bias; and (iii) estimate the magnitude of this "true" effect after accommodating and correcting for publication bias (both PET and PEESE).

The "FAT-PET-PEESE" (FPP) estimator – Step One. The first step involved in implementing the FPP estimator carries out the Funnel Asymmetry Test (FAT). This test is designed to test for publication bias. As can be seen in Equation (4.5), it uses Weighted Least Squares (WLS) to regress the estimated effects ($\hat{\alpha}_i$) on a constant term and the standard errors of the estimated effects (SE_i); where weights $\omega_i = \left(\frac{1}{SE_i}\right)$ are applied to correct for heteroskedasticity in estimates. Whereas the PET focuses on the coefficient β_0 in Equation (4.5), the FAT tests whether the coefficient on the *SE* variable, β_1, is significantly different from zero. If $\hat{\beta}_1$ is significant, then the null hypothesis, $H_0\colon \beta_1 = 0$, which indicates there is no publication bias, is rejected. Note that the bias can be positive or negative. If the conclusion of the FAT is a failure to reject the null, then there is not enough evidence to support the existence of publication bias. As part of my analysis of the performance of the FAT-PET-PEESE estimator, I will also record the performance of this FAT.

The "FAT-PET-PEESE" (FPP) estimator – Steps Two and Three. Regardless of the results on the previous step, one proceeds to the Precision Effect Test (PET), which is designed to identify whether there is genuine non-zero empirical effect after accounting for publication bias. According to Stanley and Doucouliagos (2012), the reason why the test is called precision effect testing is because β_0 is the coefficient on precision (the inverse of the standard error). It uses the same equation as the FAT, but tests whether $H_0\colon \beta_0 = 0$.

If the SE_jvariable was not included in the equation and if OLS was used rather than WLS, then the estimate of β_0 would simply be the arithmetic mean of the estimated effects in the literature. Thus, $\hat{\beta}_0$ is an estimate of the overall effect, and the PET tests $\hat{\beta}_0$ for statistical significance, corrected for publication bias.

If one fails to reject the null hypothesis of no effect, $H_0: \beta_0 = 0$. then $\hat{\beta}_0$ is taken as the estimate of overall effect with the understanding that it is statistically indistinguishable from zero. However, if the null is rejected, then one proceeds to Step Four and a new specification is estimated.

The "FAT-PET-PEESE" (FPP) estimator – Step Four. If the previous step determines that the true effect/mean value of the distribution of true effects is statistically different from zero, Stanley and Doucouliagos (2007, 2011) recommend that one estimate Equation (4.8) rather than Equation (4.6). In this case, the associated estimate of β_0 represents the best estimate of overall true effect.

In this study the four-step procedure explained above produces two test results (the FAT and PET), and a single estimate of overall effect (which I identify as the "FPP" estimate). To summarize, each simulation starts with conducting the FAT. Following that, regardless of the results for the FAT, the PET is conducted. If the PET produces a failure to reject conclusion, the coefficient on the precision term (the inverse of the standard error), is taken as the estimate of overall effect. If the PET produces a reject conclusion, the PEESE specification is estimated (cf. Equation 4.8), and the coefficient on the precision term from this specification is taken as the estimate of overall effect. This procedure is represented in the Figure 4.2 flowchart.

4.3. The Experiments

I create three different simulation environments to conduct my Monte Carlo experiments. In the first two simulation environments ("Fixed Effects" and "Random Effects") only one estimate per study is produced. In the last one ("Panel Random Effects"), multiple estimates per study are produced. The latter case is far more realistic, as most studies in the economics and business literature produce more than one estimate of the effect that is being studied.

In the Fixed Effects environment, there is only one underlying true effect and the only reason for the studies to obtain different estimated effect sizes is because of sampling error. In contrast, the true effect of x on y differs across studies in the Random Effects environment. In the last data environment, the true effects are heterogeneous both within and between studies. Given that studies differ in various characteristics such as sample sizes, estimation methods, sets of control variables, geographical units, and time periods, the more realistic data environment is when there is a distribution of the true effects.

In the Fixed Effects environment, the experiments begin by simulating a common true effect. The common true effect, α, is used to generate individual (y, x) observations, from which a single estimate is derived. In the Random Effects environment, the experiments begin by simulating a distribution of true effects that is normally distributed with mean value α. Random draws from this distribution generate study-specific "true effects",α_i. The α_i's are used to generate individual (y, x) observations, from which a single estimate is derived. The Panel Random Effects environment also builds in heterogeneity across regressions within a study. Each of these environments is described in greater detail below.

The estimates derived within each of these environments are then put through a publication bias "filter", with the number of estimates in the post-publication bias sample, M, being determined endogenously. The resulting sample constitutes the sample of estimates available to the hypothetical meta-analyst.

The respective estimators are applied to this post-publication bias sample to produce estimates of α, the true effect in the Fixed Effects environment, and the mean of the distribution of true effects in the Random Effects and Panel Random Effects environments. This process simulates a single meta-analysis study.

The process is repeated to produce 10,000 simulated meta-analysis studies. The estimates for each of the estimators are then aggregated over these simulated studies and compared on the dimensions of Bias, MSE, and Type I error rates.

For each of the tree environments, I run experiments for nine different values of α including: 0 (i.e., no overall effect), 0.5, 1, 1.5, 2, 2.5, 3, 3.5, and 4. When the distribution of true effects is centred on zero, there will be more statistically insignificant estimates, and more wrong-signed estimates, than when the distribution shifts to the right. As a result, the percent of studies excluded by publication bias will be greatest at $\alpha = 0$. As α increases and the distribution shifts to the right, fewer studies are impacted by publication bias. Eventually, for a sufficiently large value of α , all studies are "published", and the post-publication bias sample is identical to the pre-publication bias sample. As will be demonstrated below, the consequence of increasing α will differ depending on the nature of the publication bias (statistical insignificance versus wrong-signed (or wrong-direction) estimates).

<u>Performance Tests</u>. Table 4.2 through 4.9 compare seven different estimators across three different performance dimensions: (i) Average Estimate of Mean True Effect, (ii) Mean Squared Error (MSE), and (iii) Type I error rates.

<u>Mean Squared Error (MSE)</u>. MSE is one of the three performance dimensions investigated in this study. The MSE measures the average squared difference between the estimator $\hat{\alpha}$ and the parameter α, which represents either the true effect (Fixed Effects environment), or the mean of the distribution of true effects (Random Effects and Panel Random Effects environments).

$$MSE = E(\hat{\alpha} - \alpha)^2 = Var(\hat{\alpha}) + [E(\hat{\alpha}) - \alpha]^2 = Var(\hat{\alpha}) + Bias(\hat{\alpha})^2, \quad (4.9)$$

where

$$Bias\ (\hat{\alpha}) = E(\hat{\alpha}) - \alpha \tag{4.10}$$

This is also called the risk function of an estimator, with $(\hat{\alpha} - \alpha)^2$ comprising a quadratic loss function. Thus, MSE contains two components. The first component measures the variability of the estimator (precision), while the second measures its bias (accuracy). One of the properties of a good estimator is that it should have a relatively small MSE.

Type I Error Rates. Another measure of an estimator's performance is the Type I error rate associated with testing whether the estimate of α equals its true value. In the context of my experiments, the associated null hypothesis is:

$$H_0: \beta_0 = \alpha$$

I test this null at the 95% confidence/5% significance level. As a result, the associated rejection rates should likewise be equal to 5 percent. Type I error rates substantially larger or smaller than 5% indicate that the results from hypothesis testing with the given estimator are not reliable. For example, a Type I error rate equal to 0.89 means that, in my experiments, the true null hypothesis is incorrectly rejected 89% of the time. This compares with an expected rejection rate of 5% given the 5% significance level employed in the tests.

4.3.1. The Fixed Effects Data Environment

Experimental Design. For the Fixed Effects (FE) data environment, the true effect is the same for all studies. The data generating process (DGP) for the experiments in this data environment is given by

$$y_{it} = 1 + \alpha x_{it} + \varepsilon_{it}, t = 1, 2, \dots, T \tag{4.11}$$

In my experiments, I set $T = 100$ observations. In order to generate different coefficient standard errors, I allow the DGP error term to have different variances across studies:

$$\varepsilon_{it} = \lambda_i \cdot NID(0,1), \text{where} \tag{4.12}$$

$$\lambda_i = 0.2 + UID(0{,}30) \tag{4.13}$$

λ_i controls the variance of the error term. The last specification, Equation (4.13) provides a realistic range of λ_i and ensures that the variance is always nonzero.

Table 4.1 is designed to give a picture of what a typical meta-analysis sample looks like, both before and after publication bias. The specific case that is represented is when the true effect equals 1 ($\alpha = 1$). The top panel represents average sample characteristics of an "empirical literature" consisting of 1000 estimated effects, before publication bias keeps some of them from seeing the light of day. This is the "Pre-Publication Bias" sample. The next two panels respectively report average sample characteristics after imposing the two types of publication bias: publication bias against insignificance and publication bias against wrong-signs. As noted earlier, we assume that theory predicts that the respective effect should be positive (as in a value-of-life study). These are each "Post-Publication Bias" samples, and comprise the samples that the hypothetical meta-analyst analyzes.

When $\alpha = 1$ and there is no publication bias, the (average) median value of estimated effect in the full sample is 1.00, as would be expected. Estimated effects range from an average minimum of -6.85 to an average maximum of 8.92. t-statistics range from an average minimum of -2.69 to an average maximum of 45.62. The median t-value in the full sample is, on average, statistically insignificant.

These estimated effects and corresponding statistics are unobserved to the meta-analyst, as the meta-analyst only sees the estimates that survive publication bias (the "Post-Publication Bias" samples. When $\alpha = 1$ and publication bias is in favour of statistical significance, the average meta-analysis sample reduces to 318 estimated effects. The associated median estimated effect is 1.18 (representing a bias of 18%), and the average median t-statistic has increased from 0.94 in the unbiased, full sample to 2.60, and is now statistically significant. Similar results can be seen when publication bias favours estimates that are positively signed, though the median t-statistic is, on average, not so large as to be significant.

Note that when $\alpha = 1$, both types of publication bias disproportionately omit negative estimated effects, inducing a positive bias in both estimated effects and t-statistics in the post-publication bias samples. Further, both post-publication bias samples look “reasonable”. The estimated ranges of t-statistics/precision are comparable to those typically reported in economics and business.

Table 4.1: Sample Characteristics for a Simulated Meta-Analysis Data Set (Fixed Effects ($\alpha = 1$))

Variable	***Median***	***Minimum***	***P5%***	***P95%***	***Maximum***
Pre-Publication Bias (100 percent of estimates):					
Estimated effect	1.00	-6.85	-1.99	3.99	8.92
t-statistic	0.94	-2.69	-0.96	6.08	45.62
Publication Bias Against Insignificance (31.8 percent of estimates):					
Estimated effect	1.18	-6.74	-0.84	5.19	8.86
t-statistic	2.60	-2.69	-0.45	14.95	45.44
Publication Bias Against Negative Effects (80.5 percent of estimates):					
Estimated effect	1.20	-4.61	0.11	4.26	8.92
t-statistic	1.27	-1.89	0.07	7.33	45.45

Fixed Effects: Performance Tests. Table 4.2 and Table 4.3 compare seven different estimators across the three different performance dimensions mentioned earlier. While Table 4.2 reports the results when publication bias is directed against statistical insignificance, Table 4.3 examines publication bias against wrong-signed estimates. Each of the estimators is studied for a set of mean true effect values (α) ranging from 0.0 to 4.0.

The top panel of Table 4.2 reports the average estimated effects for each of the respective estimators. The first two columns report the value of the true effect (α) and the average percent studies included in the simulated meta-analysis (MA) studies, where the average is taken over 10,000 simulated studies. The first thing to note is that there is a strong relationship between the size of the true effect and the number of studies that survive publication bias against statistical insignificance. When $\alpha = 0$, less than 15% of all studies appear in the meta-analyst's sample. As α increases and the mean of the distribution of estimated effects moves away from zero, more and more studies produce significant estimates. When $\alpha = 4$, approximately three-quarters of all studies survive publication bias and are included in the meta-analyst's sample.

The next column reports results for the Unadjusted estimator. When $\alpha = 0$ and publication bias discriminates against insignificant estimates, the average estimated value of α for the Unadjusted estimator (averaged across the 10,000 simulated MA studies) is 0.01, which is very close to its expected value of 0. The Unadjusted estimator is an unbiased estimator of the true effect when $\alpha = 0$ because sampling error is equally likely to produce significant estimates that are above and below the true effect. However, as α increases, publication bias disproportionately omits studies with estimates below the true effect since, ceteris paribus, studies with small estimates are more likely to have small t-values.

Table 4.2: Comparative Performance of Meta-Analysis Estimators (Fixed Effects/Publication Bias against Insignificance)

α	*Percent*	*Unadjusted*	*PET*	*PEESE*	*FPP*	*FE*	*WLS*	*RE*
Average Estimate of Mean True Effect								
0.0	14.3	0.01	-0.01	0.00	0.00	0.00	0.00	0.00
0.5	23.0	0.92	0.48	0.51	0.51	0.51	0.51	0.56
1.0	31.8	1.57	0.97	1.00	1.00	1.01	1.01	1.05
1.5	40.0	2.14	1.46	1.50	1.50	1.51	1.51	1.54
2.0	47.6	2.67	1.96	2.00	2.00	2.01	2.01	2.03
2.5	54.6	3.17	2.47	2.50	2.50	2.51	2.51	2.53
3.0	61.1	3.65	2.97	2.99	2.99	3.01	3.01	3.02
3.5	67.0	4.12	3.47	3.49	3.49	3.51	3.51	3.51
4.0	72.2	4.58	3.97	3.99	3.99	4.01	4.01	4.01
Mean Squared Error								
0.0	14.3	0.0520	0.0021	0.0016	0.0018	0.0015	0.0015	0.0071
0.5	23.0	0.1935	0.0005	0.0002	0.0002	0.0002	0.0002	0.0035
1.0	31.8	0.3378	0.0012	0.0001	0.0001	0.0002	0.0002	0.0024
1.5	40.0	0.4199	0.0014	0.0001	0.0001	0.0002	0.0002	0.0017
2.0	47.6	0.4504	0.0014	0.0001	0.0001	0.0002	0.0002	0.0012
2.5	54.6	0.4494	0.0012	0.0001	0.0001	0.0001	0.0001	0.0008
3.0	61.1	0.4245	0.0011	0.0001	0.0001	0.0001	0.0001	0.0005
3.5	67.0	0.3845	0.0009	0.0001	0.0001	0.0001	0.0001	0.0003
4.0	72.2	0.3353	0.0007	0.0001	0.0001	0.0001	0.0001	0.0002
Type I Error Rates (H_0: $\beta_0 = \alpha$)								
0.0	14.3	0.05	0.16	0.14	0.12	0.22	0.05	0.01
0.5	23.0	0.82	0.50	0.22	0.22	0.33	0.13	0.73
1.0	31.8	1.00	0.86	0.10	0.10	0.29	0.15	0.71
1.5	40.0	1.00	0.91	0.07	0.07	0.27	0.17	0.64
2.0	47.6	1.00	0.91	0.08	0.08	0.23	0.17	0.52
2.5	54.6	1.00	0.89	0.11	0.11	0.20	0.16	0.40
3.0	61.1	1.00	0.86	0.12	0.12	0.17	0.15	0.29
3.5	67.0	1.00	0.82	0.14	0.14	0.15	0.14	0.21
4.0	72.2	1.00	0.75	0.13	0.13	0.13	0.13	0.15

When $\alpha = 1.0$, the Unadjusted estimator overestimates the mean true effect by approximately 57%. As α increases, fewer and fewer studies are affected by publication bias. While the table does not show that, further increases in α would eventually cause the publication bias associated with the Unadjusted estimator to disappear.

Continuing with the top panel of Table 4.2, I turn my attention to the performances of the six MA estimators. I am particularly interested in the first three estimators, which are specifically designed to address publication bias. The last of these reports the overall effect derived from the four-step, FAT-PET-PEESE procedure ("FPP"). With respect to estimation bias, all three estimators do very well compare to the Unadjusted estimator. When $\alpha = 1$, the mean estimates of the true effect for the PET, PEESE, and FPP estimators are 0.97, 1.00 and 1.00, respectively. When $\alpha = 2$, the estimates are 1.96, 2.00 and 2.00. In fact, for every value of α, the PET reports a bias of 3% to 4%. However, the PEESE and FPP estimators eliminate estimation bias. This success seemingly validates the ability of the PET, PEESE and FPP estimators to correct for publication bias.

However, the next three columns demonstrate that the other MA estimators also perform well, even though they do not explicitly address publication bias. The explanation lies in how the study estimates are weighted. In one way or another, all six of these estimators weight by the inverse of the estimated coefficient's standard error.

Turning to the middle panel of Table 4.2, which focuses on MSE performance, I see that the Unadjusted estimator, unsurprisingly, performs by far the worst. Of the three MA estimators designed to address and accommodate for publication bias, the PEESE and FPP estimators are most efficient. However, the FE and WLS estimators perform extremely close to the efficient estimators.

The bottom panel of Table 4.2 is the first indication that the respective MA estimators suffer from performance inadequacies, and this includes the FPP estimator. Type I error rates associated with the hypothesis $H_0: \beta_0 = \alpha$ are often far in excess of their expected value of 5%.

In summary, all six MA estimators perform better than the Unadjusted estimator. Among those six, the FPP estimator performs among the best, but struggles when it comes to reliability in hypothesis testing.

Table 4.3 continues investigating the Fixed Effects case, where all studies share a common true effect, but it reports publication bias that is targeted towards wrong-signed estimates. As evidenced by the top panel, the Unadjusted estimator again produces effect estimates that can be substantially biased. In contrast to publication bias against statistical insignificance, the bias is greatest for small values of α. As α increases, studies estimate fewer negative effects, so more studies get "published". When α is very large (= 4), almost all studies get published (98%), and the Unadjusted estimator correspondingly has a relatively small estimation bias.

Turning now to the PET, PEESE and FPP estimators, the FPP estimator is superior. However, depending on whether $\alpha = 0$ or $\alpha > 0$, the relative performance of PET and PEESE are quite different. When $\alpha = 0$, the PET estimator performs very well; when $\alpha > 0$, the PEESE estimator dominates. As before, the other three MA estimators generally also do a good job of eliminating estimation bias; and have MSE performance similar to the PEESE and FPP estimators. None of the estimators except FPP is reliable for hypothesis testing.

Table 4.3: Comparative Performance of Meta-Analysis Estimators (Fixed Effects /Publication Bias against Wrong Sign)

α	*Percent*	*Unadjusted*	*PET*	*PEESE*	*FPP*	*FE*	*WLS*	*RE*
Average Estimate of Mean True Effect								
0.0	55.0	1.01	0.00	0.51	0.00	0.07	0.07	0.08
0.5	71.7	1.22	0.47	0.50	0.50	0.51	0.51	0.51
1.0	80.6	1.56	0.98	1.00	1.00	1.01	1.01	1.01
1.5	86.5	1.94	1.48	1.50	1.50	1.51	1.51	1.51
2.0	90.6	2.34	1.99	2.00	2.00	2.00	2.00	2.00
2.5	93.5	2.76	2.49	2.50	2.50	2.50	2.50	2.50
3.0	95.5	3.20	2.99	3.00	3.00	3.00	3.00	3.00
3.5	97.0	3.65	3.49	3.50	3.50	3.50	3.50	3.50
4.0	98.0	4.11	4.00	4.00	4.00	4.00	4.00	4.00
Mean Squared Error								
0.0	55.0	1.0164	0.0001	0.0028	0.0003	0.0054	0.0054	0.0068
0.5	71.7	0.5245	0.0010	0.0001	0.0001	0.0003	0.0003	0.0003
1.0	80.6	0.3174	0.0006	0.0001	0.0001	0.0002	0.0002	0.0002
1.5	86.5	0.1957	0.0004	0.0001	0.0001	0.0001	0.0001	0.0001
2.0	90.6	0.1179	0.0003	0.0001	0.0001	0.0001	0.0001	0.0001
2.5	93.5	0.0701	0.0002	0.0001	0.0001	0.0001	0.0001	0.0001
3.0	95.5	0.0412	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
3.5	97.0	0.0240	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
4.0	98.0	0.0140	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
Type I Error Rates ($H_0\colon \beta_0 = \alpha$)								
0.0	55.0	1.00	0.08	0.99	0.08	1.00	1.00	1.00
0.5	71.7	1.00	0.86	0.08	0.08	0.45	0.50	0.44
1.0	80.6	1.00	0.71	0.07	0.07	0.22	0.27	0.22
1.5	86.5	1.00	0.53	0.08	0.08	0.13	0.17	0.13
2.0	90.6	1.00	0.37	0.08	0.08	0.08	0.11	0.08
2.5	93.5	1.00	0.25	0.08	0.08	0.08	0.10	0.08
3.0	95.5	0.98	0.17	0.07	0.07	0.06	0.07	0.06
3.5	97.0	0.82	0.12	0.07	0.07	0.06	0.07	0.06
4.0	98.0	0.53	0.10	0.07	0.07	0.05	0.06	0.05

4.3.2. Random Effects Data Environment

Experimental Design. For the Random Effects data environments, I generate heterogeneity in true effects across studies by letting the true effect be normally and independently distributed with mean α and variance 1. In particular, the DGP producing individual observations for study i is given by:

$$y_{it} = 1 + \alpha_i x_{it} + \varepsilon_{it},\ t = 1, 2, \ldots, T, \text{ where} \tag{4.14}$$

$$\alpha_i = NID(\alpha, 1) \tag{4.15}$$

All the studies have $T = 100$ observations. In order to generate different coefficient standard errors, I allow the DGP error term to have different variances across studies as follows:

$$\varepsilon_{it} = \lambda_i NID(0,1), \text{ where} \tag{4.16}$$

$$\lambda_i = 0.5 + UID(0,30). \tag{4.17}$$

As before, λ_i controls the variance of the error term. The last specification sets the minimum and maximum value for λ_i.

The specific parameter values used in the experiments were selected to simultaneously satisfy four criteria:

1. Produce a realistic range of t-values for the estimated effects.
2. Produce realistic-looking funnel plots.
3. Cause the percent of studies eliminated by publication bias to range between 10 and 90 percent (so all the MA studies are impacted by publication bias to some degree)
4. Produce realistic values of "effect heterogeneity"

"Effect heterogeneity" refers to the differences in true effects across studies.

As discussed earlier, the experiments model two kinds of publication bias: selection against statistical insignificance, and selection against wrong-signed estimates. In both cases, statistically insignificant/wrong-signed estimates are allowed to be included in the post-publication bias sample, but with a relatively low probability. The experiments set this probability at 10 percent.

Table 4.4 gives average sample characteristics for a typical meta-analysis sample in the the Random Effects data environment when $\alpha = 1$. The associated parameter values in Equations (4.16) and (4.17) have been chosen to produce a range of estimated effects and t-statistics similar to those produced in the Fixed Effects data environment (cf. Table 4.1). The additional sample characteristics added to this table is a measure of effect heterogeneity, I^2.

I^2 takes values between 0 and 1 and measures the share of variation in the estimated effects that is not attributed to sampling error (Higgins and Thompson, 2002). As a point of comparison, I^2 values between 70-95% are commonly encountered in economics and business meta-analysis research.

Table 4.4: Sample Characteristics for a Simulated Meta-Analysis Data Set (Random Effects ($\alpha = 1$))

Variable	*Median*	*Minimum*	*P5%*	*P95%*	*Maximum*
Pre-Publication Selection Bias Sample (100 percent of estimates):					
Estimated effect	1.00	-7.47	-2.38	4.39	9.46
t-statistic	0.79	-13.19	-1.47	5.90	42.19
Percent significant	0.26	0.22	0.24	0.28	0.30
I^2	0.86	0.72	0.81	0.90	0.94
Publication Bias Against Insignificance (33.0 percent of estimates):					
Estimated effect	1.81	-7.43	-2.07	5.69	9.54
t-statistic	2.54	-13.21	-2.35	12.63	42.24
Percent significant	0.93	0.89	0.91	0.94	0.95
I^2	0.94	0.86	0.92	0.96	0.98
Publication Bias Against Negative Effects (74.6 percent of estimates):					
Estimated effect	1.55	-5.01	0.05	4.77	9.52
t-statistic	1.28	-5.14	0.04	7.33	42.05
Percent significant	0.49	0.44	0.46	0.51	0.53
I^2	0.81	0.65	0.73	0.88	0.91

<u>Random Effects: Performance Tests</u>. Table 4.5 is the Random Effects analogue to Table 4.2. It shows estimator performance in the presence of publication bias against statistical insignificance. As in the Fixed Effects case, when $\alpha = 0.0$, the Unadjusted estimator is an unbiased estimator of the true effect because sampling error is equally likely to produce significant estimates that are above and below the true effect. However, as α increases, publication bias disproportionately omits studies with estimates below the true effect since, ceteris paribus, studies with small estimates are more likely to have small t-values. When $\alpha = 1.0$, the Unadjusted estimator overestimates the mean true effect by approximately 82 percent. As α increases, fewer and fewer studies are affected by publication bias. While the table does not show that, further increases in α would eventually cause the publication bias associated with the Unadjusted estimator to disappear as is the case with the results of the Fixed Effects tests. Continuing with the top panel of Table 4.5, the PET, PEESE and FPP do very well compared to the Unadjusted estimator.

The FE estimator and its near twin, the WLS estimator perform almost as well as the PET, PEESE, and FPP estimators, even though they do not explicitly correct for publication bias. As before, the explanation lies in how the study estimates are weighted. In one way or another, all of these estimators weight by the inverse of the estimated coefficient's standard error.

The RE estimator consistently overestimates the mean true effect for nonzero value of α. Interestingly, it is tailored to match the data environment in which the simulations are conducted. Despite that fact, it is the most biased estimator of the six MA estimators. This seemingly paradoxical result has been noted by other researchers (Doucouliagos and Paldam, 2013, p.586; Stanley and Doucouliagos, 2012, p.83). On the dimension of estimation bias, when $0 < \alpha < 2$, the PET performs best of all MA estimators. For $\alpha > 2$, the PEESE and FPP estimators perform best.

The middle panel of Table 4.5 focuses on MSE performance, with smaller MSE values indicating greater efficiency. The Unadjusted estimator performs poorly compared to the MA estimators for all values of $\alpha > 0$. Among MA estimators when $\alpha > 0$, the PEESE, FPP and FE/WLS estimators generally perform best.

Finally, when it comes to hypothesis testing, the bottom panel of Table 4.5 suggests that caution is in order. The FE, WLS, and RE estimators all produce type I error rates that are unacceptably large. For example, when $\alpha = 0.0$, the FE and WLS estimators reject the hypothesis that $\alpha = 0.0$ in 89 percent and 47 percent of the tests, despite the fact that the hypothesis is true. This compares with an expected rejection rate of 5 percent given the 5 percent significance level employed in the tests. The PEESE and FPP are substantially better, though they also produce Type I error rates larger than 5 percent when $0.5 \leq \alpha \leq 1.5$. Given these unattractive choices, one might easily be tempted to conclude that the PET estimator is serviceable for hypothesis testing. However, the subsequent results will render this option less tempting.

Table 4.5: Comparative Performance of Meta-Analysis Estimators (Random Effects/Publication Bias against Insignificance)

α	*Percent*	*Unadjusted*	*PET*	*PEESE*	*FPP*	*FE*	*WLS*	*RE*
				Average Estimate of Mean True Effect				
0.0	27.1	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.5	28.7	1.01	0.52	0.59	0.55	0.61	0.61	0.89
1.0	33.0	1.82	1.02	1.12	1.12	1.15	1.15	1.58
1.5	39.1	2.44	1.48	1.60	1.60	1.63	1.63	2.09
2.0	45.9	2.94	1.95	2.06	2.06	2.09	2.09	2.53
2.5	52.8	3.40	2.43	2.52	2.52	2.56	2.56	2.96
3.0	59.2	3.84	2.93	3.00	3.00	3.04	3.04	3.40
3.5	65.1	4.28	3.42	3.49	3.49	3.53	3.53	3.84
4.0	70.4	4.71	3.93	3.99	3.99	4.02	4.02	4.29
				Mean Squared Error				
0.0	27.1	0.026	0.059	0.037	0.051	0.036	0.036	0.012
0.5	28.7	0.286	0.056	0.043	0.057	0.044	0.044	0.164
1.0	33.0	0.693	0.049	0.046	0.047	0.050	0.050	0.340
1.5	39.1	0.888	0.043	0.036	0.037	0.041	0.041	0.352
2.0	45.9	0.893	0.042	0.028	0.028	0.032	0.032	0.285
2.5	52.8	0.815	0.046	0.026	0.024	0.027	0.027	0.216
3.0	59.2	0.711	0.044	0.024	0.024	0.024	0.024	0.160
3.5	65.1	0.609	0.044	0.024	0.024	0.023	0.023	0.120
4.0	70.4	0.511	0.042	0.024	0.024	0.022	0.022	0.089
				Type I Error Rates ($H_0: \beta_0 = \alpha$)				
0.0	27.1	0.05	0.08	0.07	0.07	0.89	0.47	0.03
0.5	28.7	0.92	0.09	0.12	0.13	0.90	0.55	0.95
1.0	33.0	1.00	0.08	0.16	0.16	0.92	0.64	1.00
1.5	39.1	1.00	0.08	0.13	0.13	0.91	0.65	1.00
2.0	45.9	1.00	0.09	0.09	0.09	0.89	0.61	1.00
2.5	52.8	1.00	0.10	0.07	0.08	0.88	0.59	1.00
3.0	59.2	1.00	0.10	0.07	0.07	0.87	0.60	1.00
3.5	65.1	1.00	0.10	0.07	0.07	0.87	0.59	1.00
4.0	70.4	1.00	0.10	0.07	0.07	0.87	0.61	1.00

Table 4.6 repeats the preceding analysis for the case when publication bias discriminates against negative effect estimates. The Unadjusted estimator again produces substantially biased estimates of the mean true effect, now even when $\alpha = 0$. Unlike the previous case, the MA estimators also produce biased estimates when α is relatively small. For example, when $\alpha = 1$, the associated bias ranges from 21 percent to 49 percent. These biases get smaller as α increases and the proportion of included studies becomes larger.

Table 4.6 tells a story for MSE performance that is similar to Table 4.5. The FE/WLS estimator often performs as well, and sometimes slightly better, than the PET, PEESE and FPP estimators. Interestingly, when $\alpha \geq 3$, the RE estimator is most efficient, despite being the most biased. The explanation has to do with the fact that RE estimates have generally smaller variances than other MA estimators.

Finally, as in Table 4.5, Type I error rates for the FE, WLS, and RE estimators are unacceptably large. Unlike Table 4.5, the PET, PEESE and FPP estimators now also have unacceptably large Type I error rates for small values of α.

Summarizing the results for the Random Effects data environment, I find that the MA estimators that do not explicitly correct for publication bias often perform as well, if not better, than those that do. While the MA estimators always reduce estimation bias in our experiments, they do not always eliminate it. In other words, up to this point, there is little that distinguishes the FPP estimator from other MA estimators that do not correct for publication bias.

Table 4.6: Comparative Performance of Meta-Analysis Estimators (Random Effects /Publication Bias against Wrong Sign)

α	***Percent***	***Unadjusted***	***PET***	***PEESE***	***FPP***	***FE***	***WLS***	***RE***
				Average Estimate of Mean True Effect				
0.0	55.0	1.26	0.61	0.66	0.65	0.69	0.69	0.91
0.5	65.4	1.52	0.90	0.95	0.95	0.97	0.97	1.18
1.0	74.7	1.81	1.21	1.26	1.26	1.29	1.29	1.49
1.5	82.0	2.12	1.59	1.63	1.62	1.65	1.65	1.85
2.0	87.4	2.48	2.01	2.05	2.05	2.07	2.07	2.24
2.5	91.3	2.86	2.49	2.51	2.51	2.53	2.53	2.66
3.0	94.0	3.27	2.98	3.00	3.00	3.02	3.02	3.11
3.5	95.9	3.70	3.48	3.50	3.50	3.51	3.51	3.58
4.0	97.2	4.15	3.99	4.00	4.00	4.01	4.01	4.06
				Mean Squared Error				
0.0	55.0	1.602	0.405	0.461	0.456	0.498	0.498	0.828
0.5	65.4	1.053	0.184	0.218	0.218	0.241	0.241	0.467
1.0	74.7	0.654	0.073	0.087	0.088	0.099	0.099	0.245
1.5	82.0	0.392	0.038	0.037	0.036	0.041	0.041	0.122
2.0	87.4	0.229	0.032	0.024	0.025	0.025	0.025	0.060
2.5	91.3	0.133	0.033	0.023	0.022	0.022	0.022	0.030
3.0	94.0	0.078	0.035	0.023	0.023	0.022	0.022	0.016
3.5	95.9	0.045	0.035	0.024	0.024	0.022	0.022	0.009
4.0	97.2	0.026	0.035	0.024	0.023	0.022	0.022	0.006
				Type I Error Rates ($H_0: \beta_0 = \alpha$)				
0.0	55.0	1.00	0.89	0.96	0.90	1.00	1.00	1.00
0.5	65.4	1.00	0.74	0.91	0.92	1.00	1.00	1.00
1.0	74.7	1.00	0.29	0.53	0.53	0.98	0.96	1.00
1.5	82.0	1.00	0.10	0.18	0.17	0.91	0.79	1.00
2.0	87.4	1.00	0.08	0.09	0.09	0.87	0.68	1.00
2.5	91.3	1.00	0.08	0.07	0.07	0.87	0.65	0.92
3.0	94.0	1.00	0.08	0.07	0.07	0.87	0.65	0.63
3.5	95.9	0.94	0.08	0.07	0.07	0.87	0.66	0.36
4.0	97.2	0.70	0.08	0.07	0.07	0.87	0.66	0.20

4.3.3. Panel Random Effects Data Environment

Experimental Design. The last set of experiments examines the performance of the respective MA estimators when each study contains multiple regressions/effect estimates. The true effects are modelled as differing both across and within studies.

There is a debate in the literature as to whether MA studies should include all estimates from a study, or just one, or a selected few. To the extent a consensus exists, it is that MA estimators should include all the estimates, but correct for error correlation across estimates within studies (Stanley and Doucouliagos, 2012; Ringquist, 2013).

My Monte Carlo experiments fix the number of pre-publication bias studies at 100, each with 10 estimates per study, where each estimate is based upon 100 observations. True effects are modelled as differing both within and across studies, with the differences within studies, σ_1^2, being smaller than the differences across studies, σ_2^2, such that $var(\alpha_{ij}|\alpha_i) < var(\alpha_i)$.[38]

$y_{ijt} = 1 + \alpha_{ij}x_{ijt} + e_{ijt}, t = 1, 2, \ldots, 100$, where (4.18)

$\alpha_{ij} = \alpha_i + 0.5N(0,1), j = 1, 2, \ldots, 10$, and (4.19)

$\alpha_i = \alpha + 2N(0,1)\ , i = 1, 2, \ldots, 100$ (4.20)

The different weights on the standard normal variates in (4.19) and (4.20) are designed to capture the idea that effects are more likely to be similar within a study than across studies.

The error terms are modelled similarly, with error variances again differing both within and across studies, but with most of the variation occurring across studies.

$e_{ijt} = \lambda_{ij} \cdot NID(0,1)$, where (4.21)

[38] In my experiments, σ_1^2 and σ_2^2 are set equal to 0.25 and 4, respectively.

$$\lambda_{ij} = \lambda_i + UID(0,1) \text{ , and} \tag{4.22}$$

$$\lambda_i = 0.5 + 30 \cdot UID(0,1) \tag{4.23}$$

As in the Random Effects data environment, these DGP parameters are designed to simultaneously satisfy the four criteria listed above.

Publication bias is also treated differently in the panel random effects environment. The experiments assume that the bias works at the level of the study and not the individual estimate. In the case of bias against statistical insignificance, I assume that in order to be published, a study must have most of its estimates (at least 7 out of 10) be statistically significant. If the study meets that selection criterion, all the estimates from that study are "published". If the study does not meet that criterion, none of the estimates from that study are published. An identical "7 out of 10, or more" rule applies to publication bias against wrong-signed estimates.

Another difference has to do with the specification of the MA regressions. I modify Equation (4.1) to include multiple estimates per study:

$$\hat{\alpha}_{ij1} = \beta + e_{ij} \tag{4.24}$$

Dividing through by the appropriate standard error (either SE_{ij} or $\omega_{ij} = \sqrt{\left(SE_{ij}\right)^2 + \tau^2}$) produces the FE, WLS, and RE estimators as described above.

The PET estimator follows the recommendation of Stanley and Doucouliagos (2012, see (i) Equation 5.5, p.85, and (ii) Equation 5.9, p.101):

$$\hat{\alpha}_{ij1} = \beta + \textstyle\sum_i \gamma_i SE_{ij} D_i + e_{ij} \tag{4.25}$$

where D_i is a dummy variable that takes the value 1 for study i and 0 for other studies. Dividing through by SE_{ij} produces the following specification:

$$\frac{\hat{\alpha}_{ij1}}{SE_{ij}} = \beta\left(\frac{1}{SE_{ij}}\right) + \sum_i \gamma_i D_i + \frac{(e_{ij})}{SE_{ij}} \tag{4.26}$$

The panel version of PEESE estimator is given by:

$$\frac{\hat{\alpha}_{ij1}}{SE_{ij}} = \beta\left(\frac{1}{SE_{ij}}\right) + \sum_i \gamma_i SE_i D_i + \frac{(e_{ij})}{SE_{ij}} \tag{4.27}$$

For all estimators except the FE estimator, coefficient standard errors are calculated using a clustered robust procedure to allow for within-study correlation of error terms.

The above experimental design is intended to capture the fact that studies typically contain more than one estimate of a given “effect”, perhaps because separate regressions are estimated for different subsamples of the data, or because the regression equations differ in their specifications or econometric procedures used. Thus, a realistic study of meta-analysis performance should incorporate this feature.

Table 4.7 gives average sample characteristics for a typical meta-analysis sample in the Panel Random Effects data environment when $\alpha = 1$, both pre- and post-publication bias. The associated parameter values in the DGP above have been chosen to produce a range of estimated effects and t-statistics similar to those produced in the Fixed Effects and Random Effects data environments (cf. Tables 4.1 and 4.4). As in Table 4.4, I again report a measure of effect heterogeneity, I^2. As mentioned earlier, I^2 values between 70-95% are common in meta-analysis studies conducted in economics and business. Table 4.7 makes clear that the simulated meta-analysis samples that I use for analysing the performance of the FAT-PET-PEESE estimator “look like” the kinds of meta-analysis samples that researchers apply in practice.

Table 4.7: Sample Characteristics for a Simulated Meta-Analysis Data Set (Panel Data/Random Effects ($\alpha = 1$))

Variable	***Median***	***Minimum***	***P5%***	***P95%***	***Maximum***
Pre-Publication Bias (100 percent of estimates):					
Estimated effect	0.96	-8.95	-3.51	5.51	10.89
t-statistic	0.68	-17.76	-2.90	7.05	33.43
Percent significant	0.34	0.23	0.29	0.40	0.45
I^2	0.91	0.73	0.83	0.97	0.99
Publication Bias Against Insignificance (21.9 percent of estimates):					
Estimated effect	2.40	-5.34	-3.08	6.02	8.88
t-statistic	3.68	-17.57	-7.84	16.90	33.42
Percent signifiant	0.98	0.95	0.97	0.99	1.00
I^2	0.97	0.87	0.94	0.99	1.00
Publication Bias Against Negative Effects (80.5 percent of estimates):					
Estimated effect	2.23	-5.36	-0.84	6.21	10.85
t-statistic	1.72	-2.93	-0.50	10.15	33.42
Percent significant	0.68	0.57	0.62	0.74	0.80
I^2	0.83	0.51	0.69	0.94	0.98

Panel Random Effects: Performance Tests. Table 4.8 reports performance measures for the respective estimators when publication bias favours estimates that are statistically significant. As before, the Unadjusted estimator provides an unbiased estimate of the mean true effect when $\alpha = 0$. As α increases, publication bias at first worsens, and then eventually starts to improve as more studies are "published". The numerical bias can be quite substantial. For example, when $\alpha = 2.0$, the Unadjusted estimator estimates an average value of 3.36 for α.

With respect to bias, the PET, PEESE and FPP estimators perform best of all MA estimators, with the FPP performing marginally better. For example, when $\alpha = 2.0$, the PET and PEESE estimators produce a mean estimate of α equal to 2.24, compared to 2.37 and 3.13 for the MA estimators that do not correct for publication bias. The FPP estimator produces a least biased estimate of 2.21.

However, superiority on the dimension of bias does not necessarily translate into superiority in MSE performance. While the PET, PEESE, and also FPP estimators are least biased, they are also least efficient among the MA estimators, and sometimes even less efficient than the Unadjusted estimator (cf. $0 \leq \alpha \leq 1$). Among MA estimators, the FE/WLS estimators are generally most efficient, though the RE estimator is best for low values of α.

Finally, when it comes to hypothesis testing, the lesson from the bottom panel of Table 4.8 could perhaps be summarized as "don't". In almost every case, the Type I error rates are so much larger than 5 percent that any results derived from hypothesis testing about the mean true effect should be regarded as highly dubious.

Table 4.8: Comparative Performance of Meta-Analysis Estimators (Panel Random Effects/Publication Bias against Insignificance)

α	*Percent*	*Unadjusted*	*PET*	*PEESE*	*FPP*	*FE*	*WLS*	*RE*
Average Estimate of Mean True Effect								
0.0	19.2	0.01	0.01	0.01	0.01	0.01	0.01	0.01
0.5	19.9	1.09	0.61	0.61	0.59	0.66	0.66	1.01
1.0	22.0	2.05	1.20	1.21	1.18	1.29	1.29	1.90
1.5	25.2	2.78	1.73	1.74	1.71	1.85	1.85	2.59
2.0	29.5	3.36	2.24	2.24	2.21	2.37	2.37	3.13
2.5	34.7	3.84	2.74	2.75	2.73	2.86	2.86	3.60
3.0	40.4	4.26	3.21	3.21	3.20	3.31	3.31	4.00
3.5	46.4	4.65	3.66	3.67	3.66	3.76	3.76	4.39
4.0	52.8	5.03	4.11	4.12	4.11	4.20	4.20	4.77
Mean Squared Error								
0.0	19.2	0.506	1.765	1.553	1.629	0.874	0.874	0.443
0.5	19.9	0.796	1.767	1.554	1.628	0.879	0.879	0.655
1.0	22.0	1.435	1.700	1.506	1.577	0.880	0.880	1.111
1.5	25.2	1.866	1.673	1.465	1.548	0.851	0.851	1.387
2.0	29.5	2.000	1.531	1.341	1.415	0.782	0.782	1.428
2.5	34.7	1.916	1.461	1.277	1.338	0.722	0.722	1.312
3.0	40.4	1.671	1.415	1.231	1.281	0.652	0.652	1.094
3.5	46.4	1.397	1.335	1.159	1.198	0.577	0.577	0.874
4.0	52.8	1.126	1.287	1.107	1.138	0.527	0.527	0.670
Type I Error Rates ($H_0 : \beta_0 = \alpha$)								
0.0	19.2	0.05	0.29	0.28	0.25	0.97	0.17	0.05
0.5	19.9	0.15	0.29	0.29	0.29	0.97	0.17	0.14
1.0	22.0	0.43	0.29	0.29	0.30	0.97	0.19	0.37
1.5	25.2	0.71	0.30	0.30	0.31	0.97	0.22	0.62
2.0	29.5	0.89	0.29	0.29	0.29	0.97	0.23	0.80
2.5	34.7	0.95	0.29	0.29	0.29	0.97	0.23	0.88
3.0	40.4	0.98	0.29	0.29	0.29	0.96	0.21	0.90
3.5	46.4	0.98	0.27	0.26	0.27	0.96	0.18	0.89
4.0	52.8	0.98	0.28	0.27	0.27	0.96	0.17	0.84

Table 4.9 provides further support that superiority on the dimension of biasedness does not imply superiority on efficiency. The RE estimator is now either best or close to best on the dimension of MSE for all values of α. Meanwhile, the Unadjusted estimator is more efficient than every MA estimator except the RE estimator. Reliability in hypothesis testing for the estimators continues to be abysmal across the full range of α values.

Table 4.9: Comparative Performance of Meta-Analysis Estimators (Panel Random Effects /Publication Bias against Wrong Sign)

α	*Percent*	*Unadjusted*	*PET*	*PEESE*	*FPP*	*FE*	*WLS*	*RE*
Average Estimate of Mean True Effect								
0.0	38.4	2.01	1.74	1.74	1.69	1.77	1.77	1.88
0.5	47.7	2.19	1.92	1.92	1.89	1.94	1.94	2.07
1.0	56.8	2.40	2.14	2.15	2.12	2.17	2.17	2.29
1.5	65.6	2.63	2.40	2.40	2.38	2.41	2.41	2.53
2.0	73.6	2.89	2.66	2.66	2.64	2.68	2.68	2.80
2.5	80.6	3.19	3.00	3.00	2.98	3.01	3.01	3.11
3.0	86.2	3.51	3.35	3.35	3.34	3.36	3.36	3.45
3.5	90.6	3.87	3.73	3.73	3.72	3.73	3.73	3.82
4.0	93.9	4.26	4.14	4.14	4.13	4.14	4.14	4.22
Mean Squared Error								
0.0	38.4	4.090	3.897	3.664	3.591	3.414	3.414	3.592
0.5	47.7	2.900	2.884	2.672	2.648	2.388	2.388	2.513
1.0	56.8	2.002	2.176	1.999	1.997	1.672	1.672	1.709
1.5	65.6	1.312	1.703	1.522	1.540	1.143	1.143	1.106
2.0	73.6	0.830	1.362	1.194	1.221	0.796	0.796	0.689
2.5	80.6	0.507	1.231	1.061	1.084	0.609	0.609	0.418
3.0	86.2	0.299	1.173	1.004	1.027	0.502	0.502	0.245
3.5	90.6	0.172	1.143	0.980	1.000	0.445	0.445	0.144
4.0	93.9	0.105	1.142	0.973	0.989	0.423	0.423	0.092
Type I Error Rates ($H_0 \colon \beta_0 = \alpha$)								
0.0	38.4	1.00	0.78	0.90	0.78	1.00	0.99	1.00
0.5	47.7	1.00	0.67	0.77	0.75	1.00	0.92	1.00
1.0	56.8	1.00	0.54	0.61	0.61	1.00	0.77	1.00
1.5	65.6	1.00	0.44	0.47	0.48	0.99	0.56	1.00
2.0	73.6	1.00	0.36	0.38	0.38	0.97	0.38	0.98
2.5	80.6	0.97	0.32	0.33	0.33	0.96	0.28	0.87
3.0	86.2	0.79	0.30	0.30	0.30	0.96	0.21	0.60
3.5	90.6	0.49	0.28	0.27	0.27	0.95	0.17	0.34
4.0	93.9	0.26	0.28	0.27	0.27	0.96	0.16	0.18

4.3.4. Funnel Asymmetry and Precision Effect Tests

Almost all MRA studies start by testing whether or not there is a publication bias. This helps meta-analysts to determine whether accommodating and correcting for publication bias is required. Testing $\mathrm{H_0}: \beta_1 = 0$ in the following equation is designed to answer this question (FAT).

$$\frac{\hat{\alpha}_{i1}}{SE_i} = \beta_0 \left(\frac{1}{SE_i}\right) + \beta_1 + \frac{e_i}{SE_i} \tag{4.28}$$

Further, testing $\mathrm{H_0}: \beta_0 = 0$ is a test for the presence of a genuine non-zero effect (PET). The results from FAT and PET hypothesis testing are reported in Table 4.10.

As noted above, there are six classes of experiments based on the pairing of: (i) type of date environment (Fixed Effects, Random Effects, Panel Randome Effects), and (ii) type of publication bias. The table is divided vertically into three panels according to type of data environment, from least realistic (Fixed Effects) to most realistic (Panel Random Effects). It is divided horizontally into left and right halves based on type of publication bias. The far left column reports the true overall effect, α.

I start with the Fixed Effects data environment, where each study produces only one estimate and there is one population effect underlying all studies. Each cell in the table reports the results of testing 10,000, simulated (sp), post-publication bias, meta-analysis samples. Each meta-analysis sample starts with 1,000 estimates, but not all of these are observed by the meta-analysts due to publication bias.[39] For example, when $\alpha = 0$ and publication bias is directed against insignificance (cf. left side of the table), the average meta-analysis sample contains 143 studies/estimates (14.3 percent).

[39] Note that, for the FE and RE DGPs, there is one estimate per study. However, for the PRE DGP, there are 100 studies, each containing 10 estimates.

Each of these 10,000 meta-analysis samples is tested for publication bias (FAT). As discussed above, under publication bias against insignificance, when $\alpha = 0$, there is no bias in the estimate of the overall effect, so that the null hypothesis is true. The FAT performs very well in this case, producing a rejection rate of 6 percent--close to its 5 percent significance level. In contrast, the PET is oversized with a 16 percent rejection rate. Both the FAT and the PET show excellent power. Rejection rates for the null hypotheses of no publication bias and no effect are 100 percent whenever $\alpha > 0$.

Continuing with publication bias against insignificance (left side of the table), I move down a panel to the more realistic case of Random Effects. While the rejection rates of 0.08 for both the FAT and PET are close to their significance levels when $\alpha = 0$, the tests do not perform as well when $\alpha > 0$. For example, when $\alpha = 0.5$, the FAT rejects the (false) null of no publication bias only about 33 percent of the time. The PET fails to reject the (false) null of no effect approximately 35 percent (=1-0.65) of the time. While the performances of the FAT and PET generally improve as α increases, the tests are not as reliable as they were in the Fixed Effects data environment.

The bottom panel reports results for the most realistic data environment, Panel Random Effects, where studies contain more than one estimate and there is heterogeneity in true effects both within and across studies. Both the FAT and the PET perform substantially worse.[40] When $\alpha = 0$ and publication bias is directed towards statistical insignificance, the FAT rejects the (true) null of no publication bias over half of the time (55 percent). The PET rejects the true null 29 percent of the time, and that rejection rate increases slowly as α gets larger.

[40] Heteroskedasticity-robust standard errors were used when testing hypotheses in the FE and RE cases. Clustered robust standard errors were used in the PRE case.

Moving to the right side of the table and beginning again with the top panel, I see that the FAT again does well within the Fixed Effects data environment. When publication bias is directed against negatively signed estimates and $\alpha = 0$, so that approximately half of all estimates are wrong signed, the FAT rejects the null of no publication bias 100 percent of the time. As α increases, fewer and fewer estimates are eliminated via publication bias, so that publication bias diminishes. Correspondingly the rejection rate from the FAT also falls.

The PET also performs well. When $\alpha = 0$, 45 percent (=100-55) of the estimates are eliminated because of negative signs. This causes the remaining estimates to have a strong positive bias. Even so, the PET is not fooled, and generally leads to the correct conclusion of no effect: The rejection rate of 8 percent is close to its 5 percent significance level. Further, the PET accurately identifies the existence of a nonzero effect 100 percent of the time for all $\alpha > 0$.

As before, the performances of the FAT and PET decline as the data environments become more realistic. Compared to the 100 percent rejection rate for the FAT when $\alpha = 0$ in the Fixed Effects environment, the FAT falls to 62 percent for the same scenario in the Random Effects data environment. Likewise, the PET finds evidence of an effect 90 percent of the time under Random Effects when there is, in fact, no effect ($\alpha = 0$). Things decline further still in the most realistic environment of Panel Random Effects. The FAT is largely insensitive to changes in the degree of publication bias, and the ability of the PET to identify an effect when there really is one is worse. In summary, while the FPP procedure does very well in the basic, unrealistic case of a Fixed Effects data environment, its performance declines substantially when data environments become more realistic.

Table 4.10: Funnel Asymmetry Testing (FAT) and Precision Effect Testing (PET)

	Publication Bias against Insignificant			Publication Bias against Wrong Sign		
Fixed Effects Data Environments						
α	***Percent***	**FAT**	**PET**	***Percent***	**FAT**	**PET**
0.0	14.3	0.06	0.16	55.0	1.00	0.08
0.5	23.0	1.00	1.00	71.7	1.00	1.00
1.0	31.8	1.00	1.00	80.6	1.00	1.00
1.5	40.0	1.00	1.00	86.5	1.00	1.00
2.0	47.6	1.00	1.00	90.6	1.00	1.00
2.5	54.6	1.00	1.00	93.5	0.98	1.00
3.0	61.1	1.00	1.00	95.5	0.81	1.00
3.5	67.0	1.00	1.00	97.0	0.53	1.00
4.0	72.2	1.00	1.00	98.0	0.30	1.00
Random Effects Data Environments						
α	***Percent***	**FAT**	**PET**	***Percent***	**FAT**	**PET**
0.0	27.1	0.08	0.08	55.0	0.62	0.90
0.5	28.7	0.33	0.65	65.4	0.62	1.00
1.0	33.0	0.67	0.99	74.7	0.56	1.00
1.5	39.1	0.79	1.00	82.0	0.48	1.00
2.0	45.9	0.79	1.00	87.4	0.35	1.00
2.5	52.8	0.75	1.00	91.3	0.24	1.00
3.0	59.2	0.69	1.00	94.0	0.17	1.00
3.5	65.1	0.61	1.00	95.9	0.13	1.00
4.0	70.4	0.55	1.00	97.2	0.10	1.00
Panel Random Effects Data Environments						
α	***Percent***	**FAT**	**PET**	***Percent***	**FAT**	**PET**
0.0	19.2	0.55	0.29	38.4	0.45	0.78
0.5	19.9	0.58	0.34	47.7	0.47	0.84
1.0	22.0	0.66	0.46	56.8	0.44	0.87
1.5	25.2	0.72	0.60	65.6	0.43	0.91
2.0	29.5	0.66	0.73	73.6	0.46	0.93
2.5	34.7	0.59	0.83	80.6	0.50	0.95
3.0	40.4	0.66	0.90	86.2	0.46	0.97
3.5	46.4	0.69	0.94	90.6	0.45	0.98
4.0	52.8	0.65	0.97	93.9	0.41	0.99

4.4. Conclusion

This section reports on a Monte Carlo simulation used to evaluate the performance of the FAT-PET-PEESE (FPP) procedure, a commonly employed approach for detecting and correcting publication bias in economics and business meta-analyses. The FPP procedure addresses three main objectives: (i) testing whether the sample of estimates is influenced by publication selection bias; (ii) testing whether there is a genuine non-zero true effect of estimates once the publication bias is accommodated and corrected; and (iii) obtaining an estimate of the overall mean effect.

My analysis investigated two types of publication bias: (i) publication bias against insignificant results and (ii) publication bias against wrong-signs. I also considered three data environments: (i) the Fixed Effects data environment where each study only contains one estimated effect, and where there is one true effect underlying all studies, so that all differences in estimated effects are due to sampling error; (ii) the Random Effects data environment where each study still only has one estimated effect, but where there is a distribution of true effects across studies; and (iii) the Panel Random Effects data environment where studies contain multiple estimates and there is heterogeneity in true effects both across estimates and within studies. The Panel Random Effects data environment is the data environment that most realistically models what meta-analysts in business and economics are likely to encounter in their research.

My findings indicate that the FPP procedure performs very well in the basic environment of Fixed Effects. However, in more realistic data environments, where there is heterogeneity in true effects both across and within studies, the FPP procedure's performance is generally poor. It is unreliable for the first two objectives, and less efficient than some other estimators that are not particularly designed to correct for publication bias. Further, hypothesis tests about the overall mean effect cannot be trusted.

I attempt to corroborate these findings by recreating the simulation framework of Stanley and Doucouliagos (2017) and repeat my tests using their framework. This is an undergoing project, however, the preliminary results confirm the main findings with one exception: in the S&D data environments, the FPP procedure performs better in testing hypothesis about the overall mean. However, this is not surprising given that, the “Panel Random Effects” data environment has a very different error structure than S&D, making hypothesis testing more challenging.The main conclusions I draw from thie research in this chapter are as follows. First, meta-analyses should routinely report measures of heterogeneity such as I^2. This is not standard practice in the economics and business literatures and should be. Second, future research should more intensively explore the conditions under which FPP performs well. As noted elsewhere (Stanley, 2008; Moreno et al., 2009; Stanley and Doucouliagos, 2014a), publication bias is a serious problem and the FPP procedure has shown great promise in mitigating its deleterious consequences in some cases. Having a better understanding of where the FPP procedure can be successfully applied is an important topic for future research.

4.5. References

Borenstein, M., Hedges, L. V., Higgins, J., and Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods, 1*(2), 97-111.

Chliova, M., Brinckmann, J., and Rosenbusch, N. (2015). Is microcredit a blessing for the poor? A meta-analysis examining development outcomes and contextual considerations. *Journal of Business Venturing*, *30*(3), 467-487.

Costa-Font, J., Gemmill, M., and Rubert, G. (2011). Biases in the healthcare luxury good hypothesis? A meta-regression analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *174*(1), 95-107.

Doucouliagos, H., Haman, J., and Stanley, T.D. (2012). Pay for performance and corporate governance reform. *Industrial Relations*, *51*, 670- 703.

Doucouliagos, H., and Paldam, M. (2013). The robust result in meta-analysis of aid effectiveness: A response to Mekasha and Tarp. *The Journal of Development Studies*, *49*(4), 584-587.

Doucouliagos, H. and Stanley, T.D. (2009a). Publication selection bias in minimum-wage research? A meta-regression analysis. *British Journal of Industrial Relations*, *47*, 406-29.

Doucouliagos, H., and Paldam, M. (2009b). The aid effectiveness literature: the sad results of 40 years of research. *Journal of Economic Surveys*, *23*(3), 433-461.

Doucouliagos, H., Stanley, T.D., and Giles, M. (2012). Are estimates of the value of a statistical life exaggerated? *Journal of Health Economics*, *31*, 197-206.

Doucouliagos, H., Stanley, T. D., and Viscusi, W. K. (2014). Publication selection and the income elasticity of the value of a statistical life. *Journal of Health Economics*, *33*, 67-75.

Efendic, A., Pugh, G., and Adnett, N. (2011). Institutions and economic performance: A meta-regression analysis. *European Journal of Political Economy*, *27*(3), 586-599.

Eisend, M. (2015). Have we progressed marketing knowledge? A meta-analysis of effect sizes in marketing research. *Journal of Marketing*, *79*(3), 23-40.

Gunby, P., Jin, Y., and Reed, W.R. (2017). Did FDI really cause Chinese economic growth? A meta-analysis. *World Development, 90*, 242-255.

Haelermans, C., and Borghans, L. (2012). Wage effects of on-the-job training: A meta-analysis. *British Journal of Industrial Relations*, *50*(3), 502-528.

Havránek, T. (2010). Rose effect and the euro: is the magic gone? *Review of World Economics*, *146*, 241-261.

Higgins, J.P. and Thompson, S.G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539-58.

Hughes, A.M., Gregory, M.E., Joseph, D.L., Sonesh, S.C., Marlow, S.L., Lacerenza, C.N., Benishek, L.E., King, H.B., and Salas, E. (2016). Saving lives: A meta-analysis of team training in healthcare. *Journal of Applied Psychology*, *101*(9), 1266-1304.

Iwasaki, I., and Tokunaga, M. (2014). Macroeconomic impacts of FDI in transition economies: a meta-analysis. *World Development*, *61*, 53-69.

Khechine, H., Ndjambou, P., and Lakhal, S. (2016). A meta-analysis of the UTAUT model: Eleven years later. *Canadian Journal of Administrative Sciences*, *33*, 138-152.

Laroche, P. (2016). A meta-analysis of the union–job satisfaction relationship. *British Journal of Industrial Relations*, *54*(4), 709-741.

Lazzaroni, S., and van Bergeijk, P. A. (2014). Natural disasters' impact, factors of resilience and development: A meta-analysis of the macroeconomic literature. *Ecological Economics*, *107*, 333-346.

Linde Leonard, M., Stanley, T. D., and Doucouliagos, H. (2014). Does the UK minimum wage reduce employment? A meta-regression analysis. *British Journal of Industrial Relations*, *52*(3), 499-520.

Moayeri, F., Hsueh, Y.S., Clarke, P., and Dunt, D. (2016). Do model-based studies in chronic obstructive pulmonary disease measure correct values of utility? A meta-analysis. *Value in Health*, *19*(4), 363-373.

Nelson, J. P. (2013). Meta-analysis of alcohol price and income elasticities – with corrections for publication bias. *Health Economics Review*, *3*(1), 1-10.

Reed, W. R. (2015). A Monte Carlo analysis of alternative meta-analysis estimators in the presence of publication bias. *Economics: The Open-Access, Open-Assessment E-Journal*, 9 (2015-30): 1—40. http://dx.doi.org/10.5018/economics-ejournal.ja.2015-30

Reed, W. R., Florax, R. J. G. M., and Poot, J. (2015). A Monte Carlo analysis of alternative meta-analysis estimators in the presence of publication bias. Economics Discussion Papers, No 2015-9, Kiel Institute for the World Economy. http://www.economics-ejournal.org/economics/discussionpapers/2015-9

Ringquist, E. (2013). *Meta-analysis for public management and policy*. San Franciso, CA : John Willey & Sons.

Rodríguez, J. and Muro, J. (2015). On the size of sheepskin effects: A meta-analysis. *Economics: The Open-Access, Open-Assessment E-Journal*, *9*, (2015-37): 1—18. http:// dx.doi.org/10.5018/economics-ejournal.ja.2015-37

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological bulletin*, *86*(3), 638.

Sarooghi, H., Libaers, D., and Burkemper, A. (2015). Examining the relationship between creativity and environmental factors. *Journal of Business Venturing*, *30*(5), 714-731.

Stanley, T. D. (2008). Meta-Regression methods for detecting and estimating empirical effects in the presence of publication selection. *Oxford Bulletin of Economics and statistics*, *70*(1), 103-127.

Stanley, T.D., and Doucouliagos, H. (2012a). Meta-regression analysis in economics and business. London: Routledge.

Stanley, T. D., Jarrell, S. B., and Doucouliagos, H. (2012). Could it be better to discard 90% of the data? A statistical paradox. *The American Statistician*, *64*(1), 70-77.

Stanley, T.D., and Doucouliagos, H. (2014a). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods, 5*(1), 60-78.

Stanley, T.D., and Doucouliagos, H. (2014b). Better than random: Weighted least squares meta-regression analysis. Working paper, Deakin University, School of Accounting, Economics, and Finance, Economics Series, SWP 2013/2, updated February 2014.

Stanley, T. D., and Doucouliagos, H. (2015). Neither fixed nor random: weighted least squares meta-analysis. *Statistics in Medicine*, *34*(13), 2116-2127.

Stanley, T.D., and Doucouliagos, H. (2017). Neither fixed nor random: weighted least squares meta-regression. *Research Synthesis Methods, 8*(1), 19-42.

Triana, M. D. C., Jayasinghe, M., and Pieper, J. (2015). Perceived workplace racial discrimination and its correlates: A meta-analysis. *Journal of Organizational Behavior, 36*(4), 491-513.

4.6. Appendices

Appendix 4.1: Codes

.do file for TABLES 4.2 & 4.3 – Part A

```
set more off
program drop _all
program define FEpbias, rclass
version 13.1

syntax, alpha(real) studies(integer) select(real) obs(integer)

// Remove existing variables
drop _all

//We first create the matrix to store the results of each study
set matsize 5000
matrix A = J(`studies',1,0)
matrix B = J(`studies',1,0)
matrix C = J(`studies',1,0)

forvalues i = 1/`studies' {
	clear
	// STEP ONE: Create the data for each study and estimate an effect
	set obs `obs'
	generate x = rnormal()
	scalar lambda = 0.2 + 30*runiform()
	// Note that each "study" has the same number of observations (100)
	// but differ in the variance of their respective error terms.  This
	// causes the estimate of the effect to be estimated with varying degrees
	// of precision
	generate e = lambda*rnormal()
	generate y = 1 + `alpha'*x + e
	quietly regress y x
	scalar coef = _b[x]
	scalar secoef = _se[x]
	scalar tcoef = coef/secoef

	//if abs(tcoef) < 2 {
	// An alternative criterion is that results that show a negative effect
	// have a harder time getting published.  To study that case, substitute
	// the line below for the line above
	if coef < 0 {
	scalar dummy = cond(runiform()<`select',1,.)
	// The statement above creates a dummy variable that randomly selects which
	// studies will get "published" if they fail to meet the "publication criterion"
	// either (i) abs(tcoef) >= 2 or (ii) if coef >= 0.  Studies that are not "published"
	// receive missing values and thus are not included in the "meta-analysis."
	scalar coef = dummy*coef
	scalar secoef = dummy*secoef
	scalar tcoef = dummy*tcoef
	}


	matrix A[`i',1] = coef
```

```
		matrix B[`i',1] = secoef
		matrix C[`i',1] = tcoef
}

// The next set of commands moves the data out of matrices and reformats them as
// standard Stata data series.  We have now completed generating our individual
// studies and we now move into the "meta-analysis" stage.
matrix bob = A,B,C
svmat bob
rename bob1 effect
rename bob2 seeffect
rename bob3 teffect
generate pet = (1/seeffect)

// This estimate produces the OLS estimate of the effect
regress effect
return scalar effect_OLS = _b[_cons]
test _b[_cons] = `alpha'
return scalar pvalue_OLS = r(p)

// This estimate produces the PET version of the "publication bias"-
// corrected effect estimate
regress teffect pet, vce(robust)
return scalar effect_PET = _b[pet]
scalar effect_PET = _b[pet]
test _b[_cons] = 0
return scalar pvalue_FAT = r(p)
scalar pvalue_FAT = r(p)
test pet = `alpha'
return scalar pvalue_PET = r(p)
scalar pvalue_PET = r(p)
test pet = 0
return scalar pvalue_PETFPP = r(p)
scalar pvalue_PETFPP = r(p)

// This estimate produces the PEESE version of the "publication bias"-
// corrected effect estimate
regress teffect seeffect pet, noc vce(robust)
return scalar effect_PEESE = _b[pet]
scalar effect_PEESE = _b[pet]
test pet = `alpha'
return scalar pvalue_PEESE = r(p)
scalar pvalue_PEESE = r(p)

// This estimate produces the FE estimate of the effect
generate constant = 1
vwls effect constant, sd(seeffect) nocon
return scalar effect_FE = _b[constant]
test _b[constant] = `alpha'
return scalar pvalue_FE = r(p)

// This estimate produces the WLS estimate of the effect
regress teffect pet, noc
return scalar effect_WLS = _b[pet]
test pet = `alpha'
return scalar pvalue_WLS = r(p)
// This estimate produces the RE estimate of the effect
```

```
// NOTE: We use the Method of Moments (mm) option of metareg
// because the maximum likelihood procedure had too many instances
// of failure to optimize.  Method of Moments does not require
// iteration and thus avoids this problem.
metareg effect , wsse(seeffect) mm
matrix bill = e(b)
return scalar effect_RE = bill[1,1]
test _cons = `alpha'
return scalar pvalue_RE = r(p)

return scalar effect_FPP = effect_PET
return scalar pvalue_FPP = pvalue_PET

if pvalue_PETFPP < 0.05 {
return scalar effect_FPP = effect_PEESE
return scalar pvalue_FPP = pvalue_PEESE
}

// This last command keeps track of how many studies are in our "meta-analysis"
return scalar N = e(N)

end
```

.do file for TABLE 4.2 & 4.3 – Part B

```
// This program takes about 14 hours to run on my computer
etime, start
drop _all
clear
graph drop _all
set more off
set seed 13
matrix FAT = J(9,2,0)
matrix PETFPP = J(9,2,0)
matrix EFFECT = J(9,8,0)
matrix MSE = J(9,8,0)
matrix TEST = J(9,8,0)
local studies = 1000
// Select is used to set the probability of being included in the data set
// when the study is subject to publication bias, either because (i) abs(tstat) < 2, or
// coef < 0.
local select = 0.10
local obs = 100

local i = 1
foreach alpha in 0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 {
simulate effect_OLS = r(effect_OLS) effect_FE = r(effect_FE) effect_RE = r(effect_RE) ///
        effect_PET = r(effect_PET) effect_PEESE = r(effect_PEESE)  ///
        pvalue_OLS = r(pvalue_OLS) pvalue_FE = r(pvalue_FE) pvalue_RE = r(pvalue_RE) ///
        pvalue_PET = r(pvalue_PET) pvalue_PEESE = r(pvalue_PEESE) N = r(N) ///
        effect_WLS = r(effect_WLS) pvalue_WLS = r(pvalue_WLS) pvalue_FAT = r(pvalue_FAT) ///
        pvalue_PETFPP = r(pvalue_PETFPP) effect_FPP = r(effect_FPP) pvalue_FPP = r(pvalue_FPP) , ///
        reps(10000): FEpbias, alpha(`alpha') studies(`studies') select(`select') obs(`obs')

        summ N, meanonly
        matrix EFFECT[`i',1] = r(mean)
        summ effect_OLS, meanonly
        matrix EFFECT[`i',2] = r(mean)
        summ effect_PET, meanonly
        matrix EFFECT[`i',3] = r(mean)
        summ effect_PEESE, meanonly
        matrix EFFECT[`i',4] = r(mean)
        summ effect_FE, meanonly
        matrix EFFECT[`i',5] = r(mean)
        summ effect_WLS, meanonly
        matrix EFFECT[`i',6] = r(mean)
        summ effect_RE, meanonly
        matrix EFFECT[`i',7] = r(mean)
        summ effect_FPP, meanonly
        matrix EFFECT[`i',8] = r(mean)

        summ N, meanonly
        matrix MSE[`i',1] = r(mean)
        generate mse_OLS = (effect_OLS - `alpha')^2
        summ mse_OLS, meanonly
        matrix MSE[`i',2] = r(mean)
        generate mse_PET = (effect_PET - `alpha')^2
        summ mse_PET, meanonly
        matrix MSE[`i',3] = r(mean)
```

```
generate mse_PEESE = (effect_PEESE - `alpha')^2
summ mse_PEESE, meanonly
matrix MSE[`i',4] = r(mean)
generate mse_FE = (effect_FE - `alpha')^2
summ mse_FE, meanonly
matrix MSE[`i',5] = r(mean)
generate mse_WLS = (effect_WLS - `alpha')^2
summ mse_WLS, meanonly
matrix MSE[`i',6] = r(mean)
generate mse_RE = (effect_RE - `alpha')^2
summ mse_RE, meanonly
matrix MSE[`i',7] = r(mean)
generate mse_FPP = (effect_FPP - `alpha')^2
summ mse_FPP, meanonly
matrix MSE[`i',8] = r(mean)

summ N, meanonly
matrix FAT[`i',1] = r(mean)
generate RRFAT = 0
replace RRFAT = cond(pvalue_FAT<0.05,1,0)
summ RRFAT, meanonly
matrix FAT[`i',2] = r(mean)

summ N, meanonly
matrix PETFPP[`i',1] = r(mean)
generate RRPETFPP = 0
replace RRPETFPP = cond(pvalue_PETFPP<0.05,1,0)
summ RRPETFPP, meanonly
matrix PETFPP[`i',2] = r(mean)

summ N, meanonly
matrix TEST[`i',1] = r(mean)
generate RROLS = 0
replace RROLS = cond(pvalue_OLS<0.05,1,0)
summ RROLS, meanonly
matrix TEST[`i',2] = r(mean)
generate RRPET = 0
replace RRPET = cond(pvalue_PET<0.05,1,0)
summ RRPET, meanonly
matrix TEST[`i',3] = r(mean)
generate RRPEESE = 0
replace RRPEESE = cond(pvalue_PEESE<0.05,1,0)
summ RRPEESE, meanonly
matrix TEST[`i',4] = r(mean)
generate RRFE = 0
replace RRFE = cond(pvalue_FE<0.05,1,0)
summ RRFE, meanonly
matrix TEST[`i',5] = r(mean)
generate RRWLS = 0
replace RRWLS = cond(pvalue_WLS<0.05,1,0)
summ RRWLS, meanonly
matrix TEST[`i',6] = r(mean)
generate RRRE = 0
replace RRRE = cond(pvalue_RE<0.05,1,0)
summ RRRE, meanonly
matrix TEST[`i',7] = r(mean)
generate RRFPP = 0
```

```
	replace RRFPP = cond(pvalue_FPP<0.05,1,0)
	summ RRFPP, meanonly
	matrix TEST[`i',8] = r(mean)


	local `++i'
}

matrix colnames FAT = FAT
matrix rownames FAT = A0 A0P5 A1 A1P5 A2 A2P5 A3 A3P5 A4
matrix colnames PETFPP = PETFPP
matrix rownames PETFPP = A0 A0P5 A1 A1P5 A2 A2P5 A3 A3P5 A4
matrix colnames EFFECT = N OLS PET PEESE FE WLS RE  FPP
matrix rownames EFFECT = A0 A0P5 A1 A1P5 A2 A2P5 A3 A3P5 A4
matrix colnames MSE = N OLS PET PEESE FE WLS RE FPP
matrix rownames MSE = A0 A0P5 A1 A1P5 A2 A2P5 A3 A3P5 A4
matrix colnames TEST = N OLS PET PEESE FE WLS RE FPP
matrix rownames TEST = A0 A0P5 A1 A1P5 A2 A2P5 A3 A3P5 A4
matrix list FAT
matrix list PETFPP
matrix list EFFECT
matrix list MSE
matrix list TEST

etime
```

.do file for TABLES 4.5 & 4.6 – Part A

```
set more off
program drop _all
program define REpbias, rclass
version 13.1

syntax, alpha(real) studies(integer) select(real) obs(integer)

// Remove existing variables
drop _all

//We first create the matrix to store the results of each study
matrix A = J(`studies',1,0)
matrix B = J(`studies',1,0)
matrix C = J(`studies',1,0)

forvalues i = 1/`studies' {
	clear
	// STEP ONE: Create the data for each study and estimate an effect
	set obs `obs'
	generate x = rnormal()
	// Note that each "study" has the same number of observations (100)
	// but differ in the variance of their respective error terms.  This
	// causes the estimate of the effect to be estimated with varying degrees
	// of precision
	scalar lambdai = 0.5+30*runiform()
	generate e = lambdai*rnormal()
	scalar alphai = `alpha' + rnormal()
	generate y = 1 + alphai*x + e
	quietly regress y x
	scalar coef = _b[x]
	scalar secoef = _se[x]
	scalar tcoef = coef/secoef

	//if abs(tcoef) < 2 {
	// An alternative criterion is that results that show a negative effect
	// have a harder time getting published.  To study that case, substitute
	// the line below for the line above
	if coef < 0 {
	scalar dummy = cond(runiform()<`select',1,.)
	// The statement above creates a dummy variable that randomly selects which
	// studies will get "published" if they fail to meet the "publication criterion"
	// either (i) abs(tcoef) >= 2 or (ii) if coef >= 0.  Studies that are not "published"
	// receive missing values and thus are not included in the "meta-analysis."
	 scalar coef = dummy*coef
	scalar secoef = dummy*secoef
	scalar tcoef = dummy*tcoef
	}


	matrix A[`i',1] = coef
	matrix B[`i',1] = secoef
	matrix C[`i',1] = tcoef
}
```

```
// The next set of commands moves the data out of matrices and reformats them as
// standard Stata data series.  We have now completed generating our individual
// studies and we now move into the "meta-analysis" stage.
matrix bob = A,B,C
svmat bob
rename bob1 effect
rename bob2 seeffect
rename bob3 teffect
generate pet = (1/seeffect)

// This estimate produces the OLS estimate of the effect
regress effect
return scalar effect_OLS = _b[_cons]
test _b[_cons] = `alpha'
return scalar pvalue_OLS = r(p)

// This estimate produces the PET version of the "publication bias"-
// corrected effect estimate
regress teffect pet, vce(robust)
return scalar effect_PET = _b[pet]
scalar effect_PET = _b[pet]
test _b[_cons] = 0
return scalar pvalue_FAT = r(p)
scalar pvalue_FAT = r(p)
test pet = `alpha'
return scalar pvalue_PET = r(p)
scalar pvalue_PET = r(p)
test pet = 0
return scalar pvalue_PETFPP = r(p)
scalar pvalue_PETFPP = r(p)

// This estimate produces the PEESE version of the "publication bias"-
// corrected effect estimate
regress teffect seeffect pet, noc vce(robust)
return scalar effect_PEESE = _b[pet]
scalar effect_PEESE = _b[pet]
test pet = `alpha'
return scalar pvalue_PEESE = r(p)
scalar pvalue_PEESE = r(p)

// This estimate produces the FE estimate of the effect
generate constant = 1
vwls effect constant, sd(seeffect) nocon
return scalar effect_FE = _b[constant]
test _b[constant] = `alpha'
return scalar pvalue_FE = r(p)

// This estimate produces the WLS estimate of the effect
regress teffect pet, noc
return scalar effect_WLS = _b[pet]
test pet = `alpha'
return scalar pvalue_WLS = r(p)
// This estimate produces the RE estimate of the effect
// NOTE: We use the Method of Moments (mm) option of metareg
// because the maximum likelihood procedure had too many instances
// of failure to optimize.  Method of Moments does not require
// iteration and thus avoids this problem.
```

```
metareg effect , wsse(seeffect) mm
matrix bill = e(b)
return scalar effect_RE = bill[1,1]
test _cons = `alpha'
return scalar pvalue_RE = r(p)

return scalar effect_FPP = effect_PET
return scalar pvalue_FPP = pvalue_PET

if pvalue_PETFPP < 0.05 {
return scalar effect_FPP = effect_PEESE
return scalar pvalue_FPP = pvalue_PEESE
}

// This last command keeps track of how many studies are in our "meta-analysis"
return scalar N = e(N)

end
```

.do file for TABLE 4.5 & 4.6 – Part B

```
// This program takes about 14 days to run on my laptop
etime, start
drop _all
clear
graph drop _all
set more off
set seed 13
set matsize 5000
matrix FAT = J(9,2,0)
matrix PETFPP = J(9,2,0)
matrix EFFECT = J(9,8,0)
matrix MSE = J(9,8,0)
matrix TEST = J(9,8,0)
local studies = 1000
// Select is used to set the probability of being included in the data set
// when the study is subject to publication bias, either because (i) abs(tstat) < 2, or
// coef < 0.
local select = 0.10
local obs = 100

local i = 1
foreach alpha in 0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 {
simulate effect_OLS = r(effect_OLS) effect_FE = r(effect_FE) effect_RE = r(effect_RE) ///
        effect_PET = r(effect_PET) effect_PEESE = r(effect_PEESE)  ///
        pvalue_OLS = r(pvalue_OLS) pvalue_FE = r(pvalue_FE) pvalue_RE = r(pvalue_RE) ///
        pvalue_PET = r(pvalue_PET) pvalue_PEESE = r(pvalue_PEESE) N = r(N) ///
        effect_WLS = r(effect_WLS) pvalue_WLS = r(pvalue_WLS) pvalue_FAT = r(pvalue_FAT) ///
        pvalue_PETFPP = r(pvalue_PETFPP) effect_FPP = r(effect_FPP) pvalue_FPP = r(pvalue_FPP) , ///
        reps(10000): REpbias, alpha(`alpha') studies(`studies') select(`select') obs(`obs')

        summ N, meanonly
        matrix EFFECT[`i',1] = r(mean)
        summ effect_OLS, meanonly
        matrix EFFECT[`i',2] = r(mean)
        summ effect_PET, meanonly
        matrix EFFECT[`i',3] = r(mean)
        summ effect_PEESE, meanonly
        matrix EFFECT[`i',4] = r(mean)
        summ effect_FE, meanonly
        matrix EFFECT[`i',5] = r(mean)
        summ effect_WLS, meanonly
        matrix EFFECT[`i',6] = r(mean)
        summ effect_RE, meanonly
        matrix EFFECT[`i',7] = r(mean)
        summ effect_FPP, meanonly
        matrix EFFECT[`i',8] = r(mean)

        summ N, meanonly
        matrix MSE[`i',1] = r(mean)
        generate mse_OLS = (effect_OLS - `alpha')^2
        summ mse_OLS, meanonly
        matrix MSE[`i',2] = r(mean)
        generate mse_PET = (effect_PET - `alpha')^2
        summ mse_PET, meanonly
```

```
matrix MSE[`i',3] = r(mean)
generate mse_PEESE = (effect_PEESE - `alpha')^2
summ mse_PEESE, meanonly
matrix MSE[`i',4] = r(mean)
generate mse_FE = (effect_FE - `alpha')^2
summ mse_FE, meanonly
matrix MSE[`i',5] = r(mean)
generate mse_WLS = (effect_WLS - `alpha')^2
summ mse_WLS, meanonly
matrix MSE[`i',6] = r(mean)
generate mse_RE = (effect_RE - `alpha')^2
summ mse_RE, meanonly
matrix MSE[`i',7] = r(mean)
generate mse_FPP = (effect_FPP - `alpha')^2
summ mse_FPP, meanonly
matrix MSE[`i',8] = r(mean)

summ N, meanonly
matrix FAT[`i',1] = r(mean)
generate RRFAT = 0
replace RRFAT = cond(pvalue_FAT<0.05,1,0)
summ RRFAT, meanonly
matrix FAT[`i',2] = r(mean)

summ N, meanonly
matrix PETFPP[`i',1] = r(mean)
generate RRPETFPP = 0
replace RRPETFPP = cond(pvalue_PETFPP<0.05,1,0)
summ RRPETFPP, meanonly
matrix PETFPP[`i',2] = r(mean)

summ N, meanonly
matrix TEST[`i',1] = r(mean)
generate RROLS = 0
replace RROLS = cond(pvalue_OLS<0.05,1,0)
summ RROLS, meanonly
matrix TEST[`i',2] = r(mean)
generate RRPET = 0
replace RRPET = cond(pvalue_PET<0.05,1,0)
summ RRPET, meanonly
matrix TEST[`i',3] = r(mean)
generate RRPEESE = 0
replace RRPEESE = cond(pvalue_PEESE<0.05,1,0)
summ RRPEESE, meanonly
matrix TEST[`i',4] = r(mean)
generate RRFE = 0
replace RRFE = cond(pvalue_FE<0.05,1,0)
summ RRFE, meanonly
matrix TEST[`i',5] = r(mean)
generate RRWLS = 0
replace RRWLS = cond(pvalue_WLS<0.05,1,0)
summ RRWLS, meanonly
matrix TEST[`i',6] = r(mean)
generate RRRE = 0
replace RRRE = cond(pvalue_RE<0.05,1,0)
summ RRRE, meanonly
matrix TEST[`i',7] = r(mean)
```

```
	generate RRFPP = 0
	replace RRFPP = cond(pvalue_FPP<0.05,1,0)
	summ RRFPP, meanonly
	matrix TEST[`i',8] = r(mean)

	local `++i'
}

matrix colnames FAT = FAT
matrix rownames FAT = A0 A0P5 A1 A1P5 A2 A2P5 A3 A3P5 A4
matrix colnames PETFPP = PETFPP
matrix rownames PETFPP = A0 A0P5 A1 A1P5 A2 A2P5 A3 A3P5 A4
matrix colnames EFFECT = N OLS PET PEESE FE WLS RE  FPP
matrix rownames EFFECT = A0 A0P5 A1 A1P5 A2 A2P5 A3 A3P5 A4
matrix colnames MSE = N OLS PET PEESE FE WLS RE FPP
matrix rownames MSE = A0 A0P5 A1 A1P5 A2 A2P5 A3 A3P5 A4
matrix colnames TEST = N OLS PET PEESE FE WLS RE FPP
matrix rownames TEST = A0 A0P5 A1 A1P5 A2 A2P5 A3 A3P5 A4
matrix list FAT
matrix list PETFPP
matrix list EFFECT
matrix list MSE
matrix list TEST

etime
```

.do file for TABLES 4.8 & 4.9 – Part A

```
program drop _all
program define PANELpbias, rclass
version 13.1

syntax, studies(integer) estperstudy(integer) totalobs(integer) alpha(real) ///
theta(real) obs(integer)

// Remove existing variables
drop _all

//We first create the matrix to store the results of each study
set matsize 5000
matrix A = J(`totalobs',1,.)
matrix B = J(`totalobs',1,.)
matrix C = J(`totalobs',1,.)
matrix D = J(`totalobs',1,.)

forvalues i = 1/`studies' {
        scalar lambdai = 0.5+30*runiform()
        scalar alphai = `alpha'+2*rnormal()
        forvalues j = 1/`estperstudy' {
                // STEP ONE: Create the data for each study and estimate an effect
                clear
                set obs 100
                generate x = rnormal()
                // Note that each "study" has a difference error variance, causing the estimate
                // of the effect to be estimated with varying degrees of precision
                // This is for random effects
                scalar lambdaij = lambdai+`theta'*runiform()
                scalar alphaij = alphai+0.5*rnormal()
                generate e = lambdaij*rnormal()
                generate y = 1 + alphaij*x + e
                // This is for fixed effects
                // scalar lambda = 0.2+30*runiform()
                // generate e = lambda*rnormal()
                // generate y = 1 + `alpha'*x + e
                quietly regress y x
                scalar coef = _b[x]
                scalar secoef = _se[x]
                scalar tcoef = coef/secoef
                scalar ID = `i'
                scalar obsno = (`i'-1)*`estperstudy'+`j'

                // First run this program once to get the pre-publication study sample data
                // To get post-publication study sample data, uncomment one of the two sections
                // below.
                matrix A[obsno,1] = coef
                matrix B[obsno,1] = secoef


                matrix C[obsno,1] = tcoef
        matrix D[obsno,1] = ID
        }
}
```

```
// The next set of commands moves the data out of matrices and reformats them as
// standard Stata data series.  We have now completed generating our individual
// studies and we now move into the "meta-analysis" stage.
matrix bob = A,B,C,D
svmat bob
rename bob1 effect
rename bob2 seeffect
rename bob3 teffect
rename bob4 ID
generate pet = (1/seeffect)

// This set of commands imposese the publication bias, where the publication criterion
// is either that (i) the t-stat must be greater than or equal to 2, or (ii) the
// estimated effect is positive.  The commands below implement the assumption that a
// study must have at least 7 out of 10 estimates that satisfy the publication
// criterion in order for the study to be "published."
generate dummy = 1
//replace dummy = 0 if abs(teffect) < 2
replace dummy = 0 if effect < 0
by ID, sort: egen select = mean(dummy)
// Not sure why this happens, but if I put select<0.70, it kicks out the studies that
// have 7 estimates that satisfy the publication criterion.  So I set select<0.65.
// Studies are omitted from the "meta-analysis" sample by replacing the relevant variables
// with missing values.
replace effect = cond(select<0.65,.,effect)
replace seeffect = cond(select<0.65,.,seeffect)
replace teffect = cond(select<0.65,.,teffect)
replace pet = cond(select<0.65,.,pet)

// This creates dummy variables for each of the 100 studies
// The dummy variables take names dum1 to dum100
tab ID, gen(dum)

// This creates study-specific SE terms for use in the PEESE
// according to equation 5.7 on page 85 of S&D
forvalues i = 1/100 {
        generate SE`i' = seeffect*dum`i'
}

// This estimate produces the OLS estimate of the effect
regress effect, vce(cluster ID)
return scalar effect_OLS = _b[_cons]
test _b[_cons] = `alpha'
return scalar pvalue_OLS = r(p)

// This estimate produces the PET version of the "publication bias"-
// corrected effect estimate
// NOTE1:  According to equation 5.6 on page 85 of S&D, the bias-corrected
// effect is given by the coefficients on the respective precision terms, pet*.
// The specification below forces all the effects to be the same, while
// allowing for fixed effects to correct for bias-associated with estimate SEs.
// NOTE2: Also note that while all the dummy variables will not be included in the
// meta-analysis sample, this is not a problem because STATA will automatically
// kick out the ones that don't belong.
regress teffect dum1-dum100 pet, vce(cluster ID)
return scalar effect_PET = _b[pet]
```

```
scalar effect_PET = _b[pet]
test _b[_cons] = 0
return scalar pvalue_FAT = r(p)
scalar pvalue_FAT = r(p)
test pet = `alpha'
return scalar pvalue_PET = r(p)
scalar pvalue_PET = r(p)
test pet = 0
return scalar pvalue_PETFPP = r(p)
scalar pvalue_PETFPP = r(p)

// This estimate produces the PEESE version of the "publication bias"-
// corrected effect estimate.  It is based on equation 5.7 on page 85 of S&D.
// See notes from above.
regress teffect SE1-SE100 pet, noc vce(cluster ID)
return scalar effect_PEESE = _b[pet]
scalar effect_PEESE = _b[pet]
test pet = `alpha'
return scalar pvalue_PEESE = r(p)
scalar pvalue_PEESE = r(p)

// This estimate produces the FE estimate of the effect
// Note that the FE estimator cannot do cluster robust
generate constant = 1
vwls effect constant, sd(seeffect) nocon
return scalar effect_FE = _b[constant]
test _b[constant] = `alpha'
return scalar pvalue_FE = r(p)

// This estimate produces the WLS estimate of the effect
regress teffect pet, noc vce(cluster ID)
return scalar effect_WLS = _b[pet]
test pet = `alpha'
return scalar pvalue_WLS = r(p)

// This estimate produces the RE estimate of the effect
// NOTE: We use the Method of Moments (mm) option of metareg
// because the maximum likelihood procedure had too many instances
// of failure to optimize.  Method of Moments does not require
// iteration and thus avoids this problem.
quietly metareg effect, wsse(seeffect) mm
scalar tau2 = e(tau2)
gen revarR= seeffect^2 + tau2
gen reseR = sqrt(revarR)
gen reteffect = effect/reseR
gen repet = 1/reseR
regress reteffect repet, noc vce(cluster ID)
return scalar effect_RE = _b[repet]
test repet = `alpha'
return scalar pvalue_RE = r(p)

return scalar effect_FPP = effect_PET
return scalar pvalue_FPP = pvalue_PET
if pvalue_PETFPP < 0.05 {
return scalar effect_FPP = effect_PEESE
return scalar pvalue_FPP = pvalue_PEESE
}
```

```
// This last command keeps track of how many studies are in our "meta-analysis"
return scalar N = e(N)

end
```

.do file for TABLE 4.8 & 4.9 – Part B

```
// This program takes about 14 days to run on my laptop
etime, start
drop _all
clear
graph drop _all
set more off
set seed 13
set matsize 5000
matrix FAT = J(9,2,0)
matrix PETFPP = J(9,2,0)
matrix EFFECT = J(9,8,0)
matrix MSE = J(9,8,0)
matrix TEST = J(9,8,0)
local studies = 100
local estperstudy = 10
local totalobs = `studies'*`estperstudy'
local theta = 1
local obs = 100

local i = 1
foreach alpha in 0 0.5 1 1.5 2 2.5 3 3.5 4 {
	simulate effect_OLS = r(effect_OLS) effect_FE = r(effect_FE) effect_RE = r(effect_RE) ///
	effect_PET = r(effect_PET) effect_PEESE = r(effect_PEESE)  ///
	pvalue_OLS = r(pvalue_OLS) pvalue_FE = r(pvalue_FE) pvalue_RE = r(pvalue_RE) ///
	pvalue_PET = r(pvalue_PET) pvalue_PEESE = r(pvalue_PEESE) N = r(N) ///
	effect_WLS = r(effect_WLS) pvalue_WLS = r(pvalue_WLS) pvalue_FAT = r(pvalue_FAT) ///
	pvalue_PETFPP = r(pvalue_PETFPP) effect_FPP = r(effect_FPP) pvalue_FPP = r(pvalue_FPP) , ///
	reps(10000): PANELpbias, studies(`studies') estperstudy(`estperstudy') totalobs(`totalobs') ///
	alpha(`alpha') theta(`theta') obs(`obs')

	summ N, meanonly
	matrix EFFECT[`i',1] = r(mean)
	summ effect_OLS, meanonly
	matrix EFFECT[`i',2] = r(mean)
	summ effect_PET, meanonly
	matrix EFFECT[`i',3] = r(mean)
	summ effect_PEESE, meanonly
	matrix EFFECT[`i',4] = r(mean)
	summ effect_FE, meanonly
	matrix EFFECT[`i',5] = r(mean)
	summ effect_WLS, meanonly
	matrix EFFECT[`i',6] = r(mean)
	summ effect_RE, meanonly
	matrix EFFECT[`i',7] = r(mean)
	summ effect_FPP, meanonly
	matrix EFFECT[`i',8] = r(mean)

	summ N, meanonly
	matrix MSE[`i',1] = r(mean)
	generate mse_OLS = (effect_OLS - `alpha')^2
	summ mse_OLS, meanonly
	matrix MSE[`i',2] = r(mean)
	generate mse_PET = (effect_PET - `alpha')^2
	summ mse_PET, meanonly
```

```
matrix MSE[`i',3] = r(mean)
        generate mse_PEESE = (effect_PEESE - `alpha')^2
        summ mse_PEESE, meanonly
        matrix MSE[`i',4] = r(mean)
        generate mse_FE = (effect_FE - `alpha')^2
        summ mse_FE, meanonly
        matrix MSE[`i',5] = r(mean)
        generate mse_WLS = (effect_WLS - `alpha')^2
        summ mse_WLS, meanonly
        matrix MSE[`i',6] = r(mean)
        generate mse_RE = (effect_RE - `alpha')^2
        summ mse_RE, meanonly
        matrix MSE[`i',7] = r(mean)
        generate mse_FPP = (effect_FPP - `alpha')^2
        summ mse_FPP, meanonly
        matrix MSE[`i',8] = r(mean)

        summ N, meanonly
        matrix FAT[`i',1] = r(mean)
        generate RRFAT = 0
        replace RRFAT = cond(pvalue_FAT<0.05,1,0)
        summ RRFAT, meanonly
        matrix FAT[`i',2] = r(mean)

        summ N, meanonly
        matrix PETFPP[`i',1] = r(mean)
        generate RRPETFPP = 0
        replace RRPETFPP = cond(pvalue_PETFPP<0.05,1,0)
        summ RRPETFPP, meanonly
        matrix PETFPP[`i',2] = r(mean)

        summ N, meanonly
        matrix TEST[`i',1] = r(mean)
        generate RROLS = 0
        replace RROLS = cond(pvalue_OLS<0.05,1,0)
        summ RROLS, meanonly
        matrix TEST[`i',2] = r(mean)
        generate RRPET = 0
        replace RRPET = cond(pvalue_PET<0.05,1,0)
        summ RRPET, meanonly
        matrix TEST[`i',3] = r(mean)
        generate RRPEESE = 0
        replace RRPEESE = cond(pvalue_PEESE<0.05,1,0)
        summ RRPEESE, meanonly
        matrix TEST[`i',4] = r(mean)
        generate RRFE = 0
        replace RRFE = cond(pvalue_FE<0.05,1,0)
        summ RRFE, meanonly
        matrix TEST[`i',5] = r(mean)
        generate RRWLS = 0
        replace RRWLS = cond(pvalue_WLS<0.05,1,0)
        summ RRWLS, meanonly
        matrix TEST[`i',6] = r(mean)
        generate RRRE = 0
        replace RRRE = cond(pvalue_RE<0.05,1,0)
        summ RRRE, meanonly
        matrix TEST[`i',7] = r(mean)
```

```
	generate RRFPP = 0
	replace RRFPP = cond(pvalue_FPP<0.05,1,0)
	summ RRFPP, meanonly
	matrix TEST[`i',8] = r(mean)

	local `++i'
}

matrix colnames FAT = FAT
matrix rownames FAT = A0 A0P5 A1 A1P5 A2 A2P5 A3 A3P5 A4
matrix colnames PETFPP = PETFPP
matrix rownames PETFPP = A0 A0P5 A1 A1P5 A2 A2P5 A3 A3P5 A4
matrix colnames EFFECT = N OLS PET PEESE FE WLS RE FPP
matrix rownames EFFECT = A0 A0P5 A1 A1P5 A2 A2P5 A3 A3P5 A4
matrix colnames MSE = N OLS PET PEESE FE WLS RE FPP
matrix rownames MSE = A0 A0P5 A1 A1P5 A2 A2P5 A3 A3P5 A4
matrix colnames TEST = N OLS PET PEESE FE WLS RE FPP
matrix rownames TEST = A0 A0P5 A1 A1P5 A2 A2P5 A3 A3P5 A4
matrix list FAT
matrix list PETFPP
matrix list EFFECT
matrix list MSE
matrix list TEST

etime
```

Chapter 5. Conclusion

In this thesis I undertake three studies which are linked together by the methodology of meta-analysis. The first part of my thesis investigates the effect of taxes on economic growth, which continues to be a much-studied subject. Both policy makers and researchers have long been interested to know whether taxes exert an important influence on economic growth and, if they do, how large the effect might be. Despite the large number of studies devoted to this topic, there has hitherto not been a consensus among researchers on the size, nor even the direction, of the effect.

In an attempt to provide a clear picture of existing literature, I conduct two meta-regression analyses to compare and aggregate estimates across studies. To do so, I carefully track the factors that can cause tax effects to differ. My analyses address a number of important coding issues. These include but are not limited to the implications of the government budget constraint for the interpretation of tax effects, how to integrate different units of measurement for economic growth rates and tax rates, the empirical implications of equation specifications that measure short-, medium-, and long-run effects, and how to deal with different lengths of time periods (annual data versus multi-year periods).

Chapters 2 and 3 study the effects of taxes on economic growth by looking at two different literatures. Chapter 2 focuses on OECD countries where the economies are regarded as fairly homogeneous but institutionally and culturally diverse. Chapter 3 focuses on U.S. states. When it comes to American states, there are many common features such as language and legal systems. But within this set of common institutional features, each state sets an independent tax policy and therefore this provides 50 “laboratories” to evaluate the consequence of different tax policies.

In Chapter 2, I combine 713 estimates derived from 42 empirical studies, all which endeavour to estimate the effect of taxes on economic growth in OECD countries. After dropping extreme estimates from both ends of the sample range, I apply meta-analysis procedures to analyse a final sample of 641 estimates. My results suggest that there is a publication bias towards negative estimates in the literature. After controlling for publication bias, I find that the overall effect of taxes on economic growth is statistically insignificant and negligibly small. However, this measure of the overall effect of taxes combines estimates that measure different net effects, and thus is not particularly meaningful. When I tease out the various net effects of taxes, I find general statistical support in favour of the predictions of growth theory. Further, the estimates indicate that there is scope for tax-based fiscal policy to increase economic growth amongst OECD countries.

However, I obtain very different results when I analyse tax effects in the literature on U.S. states. In Chapter 3, I combine the results from 29 empirical studies containing 966 estimates, all of which investigate the effect of taxes on economic growth in U.S. states. As in Chapter 2, I drop extreme estimates from both ends of the sample range, producing a final meta-analysis sample of 868 estimates. As in Chapter 2, I find evidence that estimates are characterized by significant negative publication bias. However, unlike Chapter 2, I do not find that the estimates support the predictions of growth theory. Nor do I find evidence to support a role for tax-based fiscal policy to contribute to economic growth in U.S. states. The reasons for the different results between Chapter 2 and 3 are not clear. In both chapters I followed identical procedures. This is a topic I hope to pursue in further research.

The second part of my thesis is concerned with the issue of publication bias. In particular, I study the performance of the FAT-PET-PEESE (FPP) procedure, a commonly employed approach for addressing publication selection bias in meta-analysis studies in economics and business. I use Monte Carlo simulations to evaluate the performance of the FPP procedure, comparing it to other common meta-analysis estimators. The three primary objectives of the FPP procedure are: (i) Funnel Asymmetry Testing (FAT) to test whether the sample of estimates is influenced by publication selection bias; (ii) Precision Effect Testing (PET) to test whether there is a genuine non-zero true effect of estimates once publication bias is accommodated and corrected; and (iii) an estimate of the true effect. In my simulations, I model two types of publication bias. These are publication bias against insignificant results and publication bias against wrong-signed estimates. I do this in a variety of data environments.

My findings indicate that the FPP procedure performs well in the basic but unrealistic environment of "Fixed Effects," where studies contain only one estimate, and there is a single, true effect underlying all studies. However, when I study the performance of the FPP procedure in more realistic data environments, where there is a distribution of true effects and studies contain multiple estimates, I find that the performance of the FPP procedure deteriorates substantially. The FAT and PET procedures become unreliable, and the FPP estimate of the overall effect is not substantially better, and sometimes worse, than other meta-analysis estimators that do not correct for publication bias. I further find that hypothesis testing across all meta-analysis estimators is unreliable and cannot be trusted.

There are two main conclusions I draw from the second part of my study. The first is that meta-analyses should routinely report measures of heterogeneity such as I^2. This is not standard practice in the economics and business literature and should be. The second conclusion I draw from my study is that future research should more intensively explore the conditions under which FPP performs well. Publication bias is a well known problem and the FPP procedure has elsewhere shown promise in mitigating its deleterious consequences. Having a better understanding of where the FPP procedure can be successfully applied is an important topic for future research.