Discussion Paper No. 11-029

What Does Ex-post Evidence Tell us About the Output Effects of Future Tax Reforms?

Richard Kneller and Florian Misch

ZEW
Zentrum für Europäische Wirtschaftsforschung GmbH
Centre for European Economic Research

Discussion Paper No. 11-029

What Does Ex-post Evidence Tell us About the Output Effects of Future Tax Reforms?

Richard Kneller and Florian Misch

Non-Technical Summary

As a result of the global economic recession there is renewed interest in the effects of taxation on economic performance: many countries attempted to use tax cuts as a means to stimulate aggregate demand and now require future tax increases to reduce large public sector deficits partly in response to the crisis. Correctly forecasting the short-run output effects and the long-run growth impact of such tax changes is central from a policy perspective. In practice however, this task is complex: tax changes affect output through a number of transmission channels, and they impact both on demand and supply conditions.

There is a large empirical literature estimating the effects of tax changes ex-post that is a potentially useful source of information for policy makers. These studies can be broadly grouped according to whether they present structural estimates, which mostly refer to the medium- and long-run relationships and employ panel methods, or whether they are based on time series for particular countries, generally refer to shorter run effects, make few structural assumptions and are primarily based on vector auto regressions. Both strands of the literature and the short- and long-run effects of tax changes are usually considered separately, even though from a policy perspective and to understand potential trade-offs, this division is not justified. However, irrespective of this distinction a general conclusion would be that, while most studies agree that changes in at least some types of taxation are associated with changes in growth rates or output levels in the short and long run, studies neither agree on the magnitude nor on the direction of the effects.

The objective of this paper is to summarize and to evaluate existing ex-post evidence on the short- and long-run output effects (i.e. short-run tax multipliers and long-run growth effects) effects of changes in taxation from the point of view of a policy maker in an OECD country who wishes to predict the output effects of a tax reform ex-ante. For this reason we propose a set of criteria against which we evaluate existing evidence that will ensure distinctness from a more traditional literature review. In particular, a given parameter estimate is useful for predicting the effects of tax reforms according to whether 1) it is relevant, 2) it can be replicated and 3) it is robust and reliable. While it is relatively straight forward to determine whether a particular study is relevant and whether the results are robust, determining whether the tax reforms can be replicated turns out to be much more complex and requires the papers to recognize the government budget constraint in the empirical specification (that every fiscal change needs to be somehow offset by some compensating fiscal item) among other issues. By applying these criteria we narrow down the number of papers that we consider as useful for policy. We include in the paper detailed summary tables that allow a more systematic and enlightened comparison of conflicting estimates between papers.

Overall, our review suggests that based on the literature surveyed, at least the direction of the long-run and to a lesser extent of short-run output responses to tax changes can be predicted with some degree of certainty. While the magnitudes of the estimated long-run output effects differ, these differences can often be attributed to differences in the measurement of the tax burden or the assumed offsetting change. In contrast to this, differences in terms of the magnitude of time series studies are often more difficult to reconcile. In addition, the long-run literature generally takes greater care to identify the exact fiscal change underlying the estimates contrary to the time series literature. Our review also suggests that there may be trade-offs between short-run output stabilization and long-run growth depending on the tax reform considered, but more research focusing on this question is required.

Das Wichtigste in Kürze

Im Zuge der globalen Rezession lässt sich wieder ein deutlich stärkeres Interesse an den volkswirtschaftlichen Effekten von Steuerreformen beobachten: eine große Zahl von Ländern hat versucht, mit Hilfe von Steuersenkungen die gesamtwirtschaftliche Nachfrage zu stimulieren, so dass in naher Zukunft Steuererhöhungen zur Senkung der öffentlichen Verschuldung unausweichlich erscheinen. Die Vorhersage der kurzfristigen Outputeffekte und der Auswirkungen auf langfristiges Wachstum von Steuerreformen erscheint aus wirtschaftspolitischer Perspektive daher elementar. In der Praxis ist dies allerdings nicht einfach: Steueränderungen beeinflussen potentiell die gesamtwirtschaftliche Nachfrage und können außerdem unterschiedliche angebotsseitige Effekte haben.

Es existieren viele Studien, die die Effekte von Steueränderungen empirisch ex-post schätzen, und die eine potenziell wichtige Informationsquelle darstellen. Diese lassen sich grob in zwei Kategorien einteilen. In der einen werden die mittel- und langfristigen Outputeffekte auf Basis von Paneldatensätzen geschätzt. In der anderen Kategorie werden vor allem die kurzfristigen Outputeffekte auf Basis von Zeitreihen einzelner Länder geschätzt. Beide Literaturbereiche werden typischerweise getrennt betrachtet, obwohl aus einer wirtschaftspolitischen Perspektive dies nicht gerechtfertigt erscheint, da sich so etwaige Zielkonflikte zwischen kurz- und langfristigen Zielen nicht identifizieren lassen. Während die meisten Studien unabhängig von dieser Unterteilung übereinstimmend zeigen, dass Outputeffekte als Folge von Steuerreformen auftreten, herrscht jedoch kein Konsens in punkto Größe und Richtung dieser Effekte.

Das Ziel der vorliegenden Studie besteht darin, die bestehenden Schätzungen zu den kurzfristigen Steuermultiplikatoren und den langfristigen Wachstumseffekten zusammenzufassen und zu beurteilen, ob sie für die Vorhersage von kurz- und langfristigen Outputeffekten zukünftiger Steueränderungen in OECD Ländern geeignet sind. In diesem Zusammenhang spielen vor allem drei zentrale Faktoren eine Rolle. Erstens müssen die Ergebnisse relevant sein, das heißt, dass die Schätzungen auf Basis von neueren OECD-Daten vorgenommen wurden. Zweitens muss bei den Studien Klarheit herrschen, welche Art von Steuerreform den Schätzungen zugrunde liegt, so dass diese theoretisch replizierbar sind. Drittens müssen die Ergebnisse robust sein. Die Bestimmung, ob eine Schätzung replizierbar ist, ist am komplexesten und erfordert u.a., dass die Studien klar definieren, welche Steuerreformen genau den Schätzungen zugrunde liegen, wie sie finanziert werden und welche anderen fiskalpolitischen Variablen im Zuge der Steuerreform geändert werden. Durch Anwendung dieser Kriterien reduziert sich die Anzahl der Studien, deren Schätzungen in der Praxis brauchbar sind. In detaillierte Tabellen werden die Studien zusammengefasst, die einen systematischen Vergleich von sich widersprechenden Schätzungen erlauben.

Die Bewertung der Schätzungen zeigt, dass auf Basis der zusammengefassten Literatur zumindest die Richtung der Outputeffekte von Steueränderungen mit relativ großer Wahrscheinlichkeit vorhergesagt werden kann. Unterschiede bei der Höhe der langfristigen Wachstumseffekte können oft durch Unterschiede bei den Steuervariablen oder durch unterschiedliche Annahmen bei der Finanzierung von Steuerreformen erklärt werden. Demgegenüber sind Unterschiede bei den geschätzten kurzfristigen Steuermultiplikatoren schwerer erklärbar, und die zugrunde liegenden Steuerreformen sind hier oft nicht klar identifizierbar. Die Ergebnisse dieser Studie weisen zudem auf mögliche Zielkonflikte zwischen dem Ziel der kurzfristigen Outputstabilisierung und der Förderung langfristigen Wachstums hin, die möglicherweise bei bestimmten Steuerreformen auftreten können.

What does ex-post evidence tell us about the output effects of future tax reforms?[1]

Richard Kneller[2]
and
Florian Misch[3]

Abstract

This paper reviews the existing evidence on the effects of tax reforms on output levels and growth over the short and long run from different strands of the literature. It develops and applies criteria to evaluate the usefulness of ex-post estimates to predict the effects of tax reforms ex-ante. These include whether the estimated policy change can be replicated in practice and whether the estimates are reliable. Based on these criteria we present detailed tables summarizing and comparing ex-post estimates of the effects of tax reforms. Overall, our review suggests that at least the direction of the long-run growth effects can be predicted with a reasonable degree of certainty. However, our review also suggests that depending on the tax change, trade-offs between short-run stabilization and long-run growth may arise and that more research on this question is needed.

JEL code: E62, H20, O20
Keywords: Tax Reforms, Tax Policy, Aggregate Growth, Tax Multipliers, Fiscal Policy

[1] We are very grateful for initial funding from the New Zealand Treasury. We thank Norman Gemmell for detailed comments on an earlier draft of the paper.
[2] School of Economics, University of Nottingham, Nottingham NG7 2RD, UK
[3] Corresponding author. Centre for European Economic Research, 68161 Mannheim, Germany; E-mail address: misch@zew.de / Telephone: +49 (0)621 123-5394 / Fax: +49 (0)621 1235-223

1 Introduction

Given the large and predominantly debt-financed public spending increases in response to the recent global economic crisis, in many OECD countries major adjustments in taxation have either already been implemented or are likely in the near future. Understanding and correctly forecasting the effects of changes to the level and structure of taxation on output levels and output growth (which we term 'output effects') ex-ante is, as a consequence, essential for governments. The output effects are not only relevant for the design and choice of tax reforms, but also to evaluate potential trade-offs with equity-related or other objectives that policy makers might seek to pursue. In practice however, this task is complex: tax changes affect output through a number of transmission channels, and they impact both on demand and supply conditions. There is a large empirical literature estimating the effects of tax changes ex-post that is a potentially useful source of information for policy makers. The objective of this paper is to summarize and to evaluate existing ex-post evidence on the short- and long-run output effects (i.e. short-run tax multipliers and long-run growth effects) effects of changes in taxation from the point of view of a policy maker in an OECD country who wishes to predict the output effects of a tax reform ex-ante.[4]

Empirical studies that analyze the growth effects of tax reforms can be broadly grouped according to whether they present structural estimates, which mostly refer to the medium- and long-run relationships and employ panel methods, or whether they are based on time series for particular countries, generally refer to shorter run effects, make few structural assumptions and are primarily based on vector auto regressions. Both strands of the literature and the short- and long-run effects of tax changes are usually considered separately, even though from a policy perspective and to understand potential trade-offs, this division is not justified. However, irrespective of this distinction a general conclusion would be that, while most studies agree that changes in at least some types of taxation are associated with changes in growth rates or output levels in the short and long run, estimates differ widely in terms of the magnitude and even in terms of the sign of the impact (whether increased taxation lowers or raises output and growth), although to a lesser extent.

From an academic perspective this raises the question of what might explain this parameter heterogeneity.[5] From a policy perspective, while the robustness of the parameter estimate is of course important, an alternative set of criteria might be used to judge the usefulness of a given study however.

[4] We recognize that there are general concerns about using ex-post estimates to forecast the effects of tax changes (see for example Auerbach, 2009). However, we abstract from these issues in this paper.

[5] There are already several papers that review the existing evidence of the macroeconomic effects of fiscal policy which can be classified according to which aspect of fiscal policy they focus on, whether they consider short-run or long-run effects and which macroeconomic effects (i.e. which macroeconomic variable) they consider. Hall (2009) reviews the evidence of the short-run impact of public spending on output. Romp and de Haan (2007) and Straub (2008) focus on the long-run output effects of public investment. Hebous (2009) reviews the short-run impact of various fiscal policy shocks on the macroeconomy more generally. Gemmell (2004) summarizes the existing literature on fiscal policy and long-run growth at that time from an academic perspective.

We order our criteria for selecting whether a given parameter estimate is of value for predicting the effects of tax reforms according to whether it can be *replicated* by government and is econometrically *robust*. While it is relatively straight forward to agree on a set of criteria against which researchers would be willing to describe an estimated parameter value as robust (it should be free from endogeneity bias for example), determining whether the tax reforms can be replicated turns out to be much more complex. Often papers deal with issues of econometric robustness but do not consider whether their implicit policy choices are replicable by policy makers.

In principle replicability of any empirical estimate of tax reforms therefore requires identifying the tax policy change that took place, at least in broad terms. Generally speaking, tax changes can be distinguished by the type of tax involved and by whether the change affects the level (revenue changing effects), the tax mix (share of each tax in total revenue) and the structure of a particular tax (its progressivity). For time series studies, indentifying the exact tax change underlying the estimates is mostly not feasible as they consider changes to some measure of total revenue. However, it may be possible to justify this based on the possibility that alternative tax changes have identical output effects over the short run.

In addition, key within the framework we adopt to judge replicability is that no tax change can be understood in isolation.[6] Changes in the level of any given tax *must* be financed by changes to some other tax, changes to government expenditure, or a change in the budget deficit, and may be associated with a change to the progressivity of taxation, or perhaps more likely some combination of all of those. Changes to taxation therefore imply (compensating) changes to other aspects of fiscal policy, which themselves may have an effect on output levels and growth, thereby reinforcing or offsetting the output effects of the original tax change. The outcomes from a particular policy change may, in this sense, be unique, and this is why parameter estimates on tax variables vary across different studies (in addition to endogeneity bias or changes in the sample of countries). Recognizing that another aspect of fiscal policy must change when a particular aspect of taxation changes does not imply that a single parameter estimate derived from any particular study is more or less 'correct' than another; each will capture the effect on growth of a change in that fiscal variable offset by some other changes. It does however imply that some studies produce parameter estimates that are more useful for policy forecasting the output effects from a given tax change. In principle, this implies that a given parameter estimate is only useful for forecasting growth outcomes if it is possible to *identify* and therefore *replicate* what the combination of other compensating fiscal changes were, although we make some exceptions to this general rule.

Related to this point is our interest to be able to compare the output level and growth effects from *different* compensating changes. The ability to identify what compensating changes are implied by a particular regression equation allows us to compare between the alternative compensating policy choices that arise

[6] Unless it is assumed that the tax variables of interest is the only fiscal policy change that can affect growth. The existing empirical evidence would lead one to argue against such a view.

out of a given tax change. In a perfect world there would exist a menu of other possible changes to fiscal policy for each possible tax change, such that the effects from these points in the policy space could be accurately compared and then applied. Unfortunately, the empirical literature is not yet at such a point that would allow us to construct a menu of choices, but we are able to at least identify some points on that menu.

In contrast to analyzing the extent to which tax reforms can be replicated, analyzing the robustness and the reliability of the econometric estimates is relatively straight forward. We apply standard criteria and examine whether panel studies control for endogeneity, and whether unobserved heterogeneity is likely to represent a problem. For time series studies, we examine some aspects of the econometric specifications including the identification strategy that affect the results. For both types of studies, we ignore specifications with coefficients that are not significant (which are however clearly outnumbered by those specifications with significant results).

By applying these criteria we narrow down the number of papers that we consider as useful for policy. We also present detailed summary tables which allow a more systematic and enlightened comparison of conflicting estimates between papers. Overall, our review suggests that at least the direction of the long-run growth response to tax changes can be predicted with a reasonable degree of certainty. While the magnitudes of the estimated long-run output effects differ, these differences can often be attributed to differences in the measurement of the tax burden or the assumed offsetting change. Most time series studies also agree in terms of the direction of the output effects, but differences in terms of the magnitude are often more difficult to reconcile. In addition, it is generally harder to identify the exact fiscal change underlying the estimates contrary compared to panel studies making them less useful for policy purposes. Our review also suggests that there may be trade-offs between short-run output stabilization and long-run growth depending on the tax reform considered, but more research focusing on this question is required.

This paper is organized as follows. Section 2 presents the criteria that we use to select studies that we include in this review which are then evaluated based on the criteria of replicability and reliability. We consider both studies that estimate the long-run growth effects of taxation as well as studies that estimate short-run tax multipliers. Section 3 discusses the replicability of the results and in particular measurement issues of tax changes as well as whether studies have considered how changes in taxation are offset. Section 4 evaluates the econometric methods used and the robustness of the result. Based on our analysis about which specifications in the selected studies are useful to predict the output effects, we then present the exact empirical estimates of the output effects over the long and the short run from the literature in Section 5. Section 6 concludes and points out directions for further policy-relevant research.

2 Some Preliminaries

One explanation for the lack of robustness of the estimated effects of tax reforms across studies would be standard econometric issues such as differences in the definitions of fiscal policy between countries (measurement bias) and differences in the effects across countries (heterogeneity) in the case of panel studies for example. Such issues are clearly important. In order to minimize the effects of heterogeneity and measurement bias at the outset we therefore mostly concentrate on studies based on data from developed countries only and for instance exclude paper like Lee and Gordon (2005) who use data from developed and developing countries.

We also only discuss studies from around the previous decade of research. This criterion is based on the assumption that a) more recent studies are of better quality because they are more likely to employ advanced econometric techniques and that b) more recent evidence is also likely based on more recent data and therefore more relevant for contemporaneous policy making. We also exclude studies that only measure the effects of government size on output, or studies that estimate only the indirect impact of various fiscal policy variables on output (i.e. studies that estimate the effects of fiscal policy on private investment or human capital accumulation for example).

The time horizon for the forecast of the effects of tax reforms matters because there is much theoretical evidence that suggests that the effects of changes in taxation differ over the short and the long run. Broadly we include studies that use period averages of the data (most typically 5 years) or dynamic panel data methods with annual data and that calculate long-run coefficients and label these as “long-run studies” or “panel studies”. Those that examine shorter time horizons we group as “short-run studies” or “time series studies”. A second reason for distinguishing between these strands of the literature is based on the observation that most of the time series studies pay little attention to potential long-run effects of taxation in the sense that they never distinguish productive from non-productive spending and on occasion (for instance Mertens and Ravn, 2009) assume that tax shocks are not permanent. For these short-run studies the growth effects of taxation are typically studied up to 3 years. Table 1 provides an overview of the papers that we review.

3 Replicability

3.1 Measures of the Tax Burden and the Tax Changes in Panel Studies

This sub-section reviews the measures of the tax burden in panel studies. The way the tax burden is measured matters for two reasons. First, it is debatable what type of measure captures best the growth effects of taxation. One simplifying assumption often made as outlined above is that the growth effects of taxation are fully captured by a single variable, the measure of the ‘implicit’ average tax rate (the revenue

derived from a given tax as a ratio to GDP). We might anticipate that growth is instead affected by the level of taxation (i.e. the revenue derived), the mix of taxes (i.e. the contribution of different taxes to the overall tax burden) *and* aspects of the structure of each tax (their progressivity). Bretschger and Hettich (2002), Devereux et al. (2008) and Myles (2007) all argue that other measures of taxation, such as the effective average tax rate and statutory tax rates, better reflect the 'real' tax burden. Second, the choice of the tax variable further determines what specific change in taxation is being considered. Using imprecise tax measures may imply that the tax change cannot be replicated and is therefore not useful for policy in practice.

It is not clear what the relevant measures of progressivity that one should consider are. The review by Myles (2007, p 89) succinctly summarizes both the importance and the empirical difficulties of this:

> *"What should matter for the economic outcome is the distortion caused by the tax (how much it changes decisions). An aggregate measure of the tax rate can never capture the varying degrees of distortion that individuals or firms with different incomes will face. […] it still remains the case that all of the regressions are limited by the fact that they are unable to work with the rate of tax that affects individual decisions. For decisions at the margin we would think of the marginal rate of tax as being important. But there are discrete choices (such as choice of location) for which the average rate matters. What the regressions end up using is an aggregate average rate, or constructed marginal rate, that probably does not affect the rate that any particular economic decision maker is facing."*

As already noted, many previous studies have mostly relied on implicit average tax rates (IATRs) measured using tax revenue data either as a share of GDP (e.g. Kneller et al., 1999; Bleaney et al., 2001, Arin, 2004; Aiginger and Falk, 2005; Romero-Ávila and Strauch, 2008; Benos, 2009). Others have alternatively calculated 'effective' rates, tax revenue data as ratios of the relevant tax base such as personal incomes or corporate profits (Padovano and Galli, 2002; Angelopoulos et al., 2007; Romero-Ávila and Strauch, 2008). A smaller number have reported tax-growth effects using statutory tax rates (e.g. Wildmalm, 2001; Angelopoulos et al., 2007). We might think of these as capturing the level and structure effects of taxation to different degrees. There is more sophisticated data available on effective tax rates (for instance provided by Devereux et al., 2009). However, their use in panel studies will only be possible once longer time series are available. Arnold (2008) reports the effects of changes in the shares of particular taxes in total revenue.

Given the discussion from the literature we might consider IATR measures are capturing more of the level effects than effects from the structure of taxation, while statutory tax rates as capturing relatively more information about the structure than the level. We view effective tax rates as most likely laying between these two. However, to some extent all capture the growth effects that occur from changes in the level of taxation. This justifies that we choose to ignore from the tables that summarize the estimates (Tables 2 to 5) those studies that include only a small number of other fiscal variables irrespective of whether the tax

variable are IATR measures, effective tax rates or statutory tax rates because in these cases, the offsetting fiscal change is not clear.

Assuming that both the level and the progressivity of taxation have non-zero output effects (the discussion is typically framed in relation to distortionary taxes), and assuming there is a tendency for both to be altered at the same time, then there will be an omitted variable bias of an unknown direction on the included tax variables (depending on whether they have a tendency to both become more distortionary at the same time, or whether they move in opposite directions). If the structure of taxation and the level of taxes are usually moved in the same direction, both become more distortionary, then by omitting the relevant measures of the structure of taxes, the level of taxes may pick up part or all of the structure's growth effect. It is in effect this form of omitted variable bias that the proponents of using non-IATR measures are assuming. Alternatively it is equally plausible that the level and structure of taxes are moved in opposite directions at the same time (one becomes more distortionary and the other less so). Here the coefficient on the level of taxes will be biased towards zero (its negative effects are being offset by the structure of taxation becoming less distortionary).[7] In practice it seems likely that the direction of the bias will vary across countries and time periods. More generally this suggests a need to give some thought to both the level and the structure of taxation.

If the tax level and tax structure have independent effects, changing one does not imply changes to the other, then they can be captured using different variables and the empirical solution would be to include measures of both in the regression equation. More likely the measures of taxation that are used by researchers capture aspects of both the level and the structure, making them interdependent. Once again issues relating to the budget constraint arise; if the level of taxation changes consideration has to be given to what other fiscal revenues, expenditures and the deficit have been included and excluded from the regression.

Therefore, there also exists some motivation for including in the same regression average and marginal measures of the same tax. As discussed in Gemmell et al. (2009) top personal statutory rates capture the marginal rates relevant to higher income earners, they are likely to be close to the personal rates most relevant to many human capital accumulation and entrepreneurial decisions. Likewise, effective tax rates (ETRs), have been argued to reflect the tax rates relevant to corporate investment decisions under a variety of assumptions and, as Devereux et al. (2008) emphasize, the *statutory* corporate tax rate is the relevant rate applicable to corporate profit-shifting decisions. If done however, the research should recognize that the effects of changes in one measure of taxation, say the marginal tax rate, are conditional on the other included aspects, say the implicit average.

[7] Dahlby and Ferede (2008) regress total tax revenues against the top rate of personal and corporate income tax and sales tax. Conditional on income levels they find the coefficients on the top rate of corporate and personal income taxes to be negatively associated with the total tax level.

A small number of studies have performed analyses of this type. Angelopoulos et al. (2007) use data from 23 OECD countries from 1970 to 2000 (5-year averages) and have measures of implicit average, effective average and statutory tax rates. Unfortunately however they never include all measures within the regression equation, choosing instead to group them and include all effective average, implicit and statutory tax rates together in separate regressions. The budget deficit is excluded and thus forms the compensating fiscal item. For effective average income taxation they find a significant negative effect on growth that is robust in all bar two specifications, while for capital and consumption taxation they find no significant effects. Interestingly when using the top individual income tax rate, the top corporate income tax rate and the statutory manufacturing tax rate they find no significant effects at all. This would suggest that effective average and top marginal income tax rates have different growth outcomes.

In contrast to the above Arnold (2008) and Gemmell et al. (2009) include multiple measures of the same tax in the same regression, but in slightly different ways. Arnold (2008) conditions the effects of taxation on the overall tax burden. He therefore considers the effect of a shift away from one type of tax to another, holding constant total tax revenues. Gemmell et al. (2009) alternatively control for the level of distortionary taxes and include the budget surplus/deficit and productive expenditures. Conditional on the level of taxes Arnold (2008) finds that an increase in the progressivity of income taxes decreases growth, even when conditioned on the share of personal income taxes. Gemmell et al. (2009) also uncover differences in the growth effects of average and marginal income taxes. The average distortionary tax variable, with non-productive expenditure and non-distortionary taxes as the compensating fiscal items, has a negative correlation with growth, although this result is not robust to the inclusion of initial income. Conditional on this, the top statutory rates of income tax and corporate tax are also both negative.

Changes to the mix of taxes are straight forward to deal with. Some studies include separate variables for different types of taxation. Arnold (2008) for example includes the share of revenue from each tax in total tax revenue which facilitates estimating changes in the mix of taxation.

There are other studies that only include very broad revenue measures. Bania et al. (2007) use total state revenue which includes state and local taxes, fees and intergovernmental revenue as their only tax-related variable. Reed (2008) likewise uses a tax measure comprising all state and local tax revenue which may include transfers from the central government. We choose to exclude both studies from the final tables because their ‘tax’ variables might include a significant share of intergovernmental transfers.

3.2 Measurement of Tax Shocks in Time Series Studies

This sub-section reviews the measures of the tax burden in time series studies. Tax variables vary and include net tax measures (i.e. revenue minus transfers), total tax measures (i.e. total revenue) and disaggregate tax measures (i.e. revenue from a particular type of tax). Romer and Romer (2010) and

others for example, use tax liabilities to GDP which is probably close to the implicit average tax rate. Blanchard and Perotti (2002) in contrast use a different measure, total tax revenues minus transfers. Caldara and Kamps (2008) also use a different measure: government receipts minus transfers and interest payments. Barro and Redlick (2009) is the only time series study that uses a measure of marginal tax rates. They use newly compiled time series on average marginal income tax rates for the U.S. spanning several decades up to 2006.

Most of the measures used are quite broad, and the estimated output effects are difficult to interpret for policy. For instance, it is not possible to unambiguously identify the underlying policy change of an increase of net tax revenue measures (i.e. tax revenue minus spending items). Nevertheless, given that most studies use such measures, we include them in the final tables. It is also unclear to what extent the composition of tax shocks (i.e. the contribution of distortionary versus non-distortionary taxes to an increase in overall taxation for instance) matters over the short run. Among the few papers that consider shocks to different types of taxes are de Castro and de Cos (2008), Arin and Koray (2005) and Arin and Koray (2006). According to their estimates, the difference between shocks to revenue from different tax types appears to be small in some instances. Nevertheless, these issues limit the usefulness of these studies for forecasting the effects of tax reforms ex-ante.

Another issue that is important for the dynamic response of output is the assumption about when the tax shock occurs. In principle, tax shocks can be anticipated or unanticipated so that the output effects at the date of the implementation or at the date of the announcement could be chosen. In addition, even after implementation, it is conceivable that the output response differs between anticipated and unanticipated tax shocks. However, few studies control for expectations. Using the dataset compiled by Romer and Romer (2010), Mertens and Ravn (2009) for instance estimate the effects of anticipated and unanticipated tax shocks that are exogenous and not correlated to other developments in the economy. For each tax shock, they define the announcement and the implementation dates. If these dates are sufficiently close together, they classify the tax shock as ‘unanticipated’. They find that the dynamics of anticipated shocks resemble those of unanticipated tax shocks after implementation. However, they also show that there are significant pre-implementation effects when tax shocks are announced but not yet implemented, and these differ in sign and magnitude to the post-implementation output effects. In particular, they find that during the pre-implementation period an anticipated tax cut gives rise to significant declines in output contrary to the post-implementation period. Blanchard and Perotti (2002) also find qualitatively similar results between anticipated and unanticipated tax shocks once tax changes are implemented. However, they find no output effects in the pre-implementation period. Part of the reason could be attributed to their identification strategy and their method of estimating pre-implementation effects, which differs from the one used in Mertens and Ravn (2009).[8]

[8] Whether the tax shock underlying particular estimates is permanent or not may also be important for replication as suggested by Auerbach (2009). However, we do not discuss this issue further.

3.3 The Government Budget Constraint

A key element when defining the empirical strategy to understand the output effects from tax changes is the recognition of the interdependence of fiscal variables across the government budget constraint: changes in one form of taxation must either be financed through changes in other taxes, changes to expenditures, or the budget deficit. There is an empirical counterpart to this: the estimated coefficient on a given fiscal variable, say income taxation, does not solely measure its direct effect on growth, but also includes information on the growth effects from the fiscal item(s) used to finance or offset it. Being able to identify the offsetting change is therefore in principle an important requirement for replication.

Panel data studies

For panel studies, Kneller et al. (1999) show that the correct interpretation of the coefficient on each fiscal category is the effect of a unit change in the relevant variable *offset by a unit change in the omitted fiscal* category.[9] To understand the implications of this point, suppose that growth, g_{it}, in country *i* at time *t* is a function of conditioning (non-fiscal) variables, Y_{it} , and a vector of fiscal variables, X_{jt}.

(1) $$g_{it} = \alpha + \sum_{i=1}^{k} \beta_i Y_{it} + \sum_{j=1}^{m} \gamma_j X_{jt} + u_{it}$$

Since the budget constraint requires that expenditures are financed by the sum of taxes and the deficit, one element of *X* must be omitted in order to avoid perfect collinearity. The omitted variable is effectively the assumed compensating element within the government's budget constraint. Thus, rewriting (1) as:

(2) $$g_{it} = \alpha + \sum_{i=1}^{k} \beta_i Y_{it} + \sum_{j=1}^{m-1} \gamma_j X_{jt} + \gamma_m X_{mt} + u_{it}$$

where X_{mt} is the omitted fiscal category, the condition $\sum_{j=1}^{m} X_{jt} = 0$ allows (2) to be rearranged to give:

(3) $$g_{it} = \alpha + \sum_{i=1}^{k} \beta_i Y_{it} + \sum_{j=1}^{m-1} (\gamma_j - \gamma_m) X_{jt} + u_{it}$$

It follows that adjusting the omitted category changes the estimated coefficients of the included categories and (possibly but not necessarily) the resulting parameter estimate. Interpretation of the growth effects from a given fiscal change therefore requires consideration of what elements of the government budget are being excluded from the regression. In studies that directly apply the methodology to take into account the government budget constraint (GBC) outlined in Kneller et al. (1999) this is usually made clear. In those that do not formally use the GBC methodology it has to be inferred from what other forms of expenditure, taxation and the budget surplus/deficit have been included in the regression. Ideally then,

[9] If distortionary taxes are the only element of the government budget that affects growth, as assumed in some early models of fiscal policy and growth, then this problem disappears and the parameter estimates directly capture those effects.

studies would properly specify the government budget constraint and then rotate a single omitted fiscal item such that all combinations of fiscal changes were identified. The sheer number of fiscal items that might be included in the regression means this is typically not done.[10] Instead either broader aggregates, labelled as distortionary taxes for example, or elements from different parts of the GBC are omitted.

The GBC generates greater problems for interpreting coefficient estimates when more than one type of fiscal variable is being excluded from the regression. In such a case we might think of the growth effect from a change in a given fiscal parameter as equal to its own direct effect on growth multiplied by the size of the change, minus the sum of the changes in each of the compensating fiscal components, multiplied by the effect that each of those has on growth. This implies that interpretation becomes difficult because we do not directly observe those changes made to the other fiscal categories or the direct growth effects associated with those fiscal variables. While the particular parameter estimate provides information of the growth effect compensated by a given set of changes elsewhere in the GBC, we therefore cannot be sure what the complete policy change was. Therefore, it may not be feasible to replicate the exact policy change so that the estimates cannot be used to forecast the output effects that result from tax reforms.

To take a specific example: suppose that a measure of income taxation were included within the regression and the omitted fiscal items, the compensating aspects of fiscal policy, were capital taxation, productive government expenditures and the budget deficit. We could only be sure we correctly forecast the growth effect from a given change in income taxation using that parameter if we could exactly identify the changes that also occurred to capital taxation, productive expenditures and the deficit and knew the direct effects each of those had on growth. Even if we knew the average set of compensating changes typically associated with a change in income taxation, we do not know the direct growth effect of each of the omitted fiscal elements. From this we derive the rule of thumb: the fewer the fiscal variables included in the regression, the longer is the list its estimated effect on growth is conditional on, the more difficult it is to identify and to replicate the tax reform underlying the estimates in practice, and therefore the less useful it is for policy advice. Below, we review how and to what extent panel and time series have incorporated the GBC.

For panel studies, economic theory allows making an exception to the inclusion of a large number of fiscal variables which is useful when including many fiscal variables is not feasible or desirable. Following the Barro (1990) model we might classify some expenditure types as unproductive and some taxes as non-distortionary. In such cases the growth effects from those particular fiscal items are zero (or in practice at least very small), so if omitted together, it does not matter what the extent of the change across those two groups were, as neither has an effect on growth. This is the approach taken in Kneller et al. (1999), Bleaney et al. (2001), Benos (2009), and Gemmell et al. (2009). They use economic theory to suggest the

[10] As most of the categories are near zero, this often leads to problems of multicollinearity, so that most variables tend to be poorly identified.

categories that are predicted to have zero effects on growth and therefore aid with the interpretation of the estimated coefficients. As an additional exception, we also assume that papers which exclude non-tax revenue and other taxes alongside other fiscal variables adequately consider the government budget constraint and therefore include their estimates in the final tables. The reasons include that the magnitude of these items is often relatively small compared to revenue from broad-based taxes in many OECD countries, non-tax revenue is to some extent exogenously given so that it may not be used to offset tax changes, and the growth effects that changes of this type of revenue entail may also be small.

Kneller et al. (1999) find that the coefficient on distortionary taxation (taxation of income and profit, social security contributions, taxation on payroll and manpower, taxation on property), financed by either changes to unproductive expenditures, non-distortionary taxation (taxation of domestic goods and services) or both, yields a robust negative effect on growth. These effects are conditional on the inclusion of other revenues, which include taxes on international trade, other tax revenues and non-tax revenues. The same approach is taken in Benos (2009) who assumes that non-distortionary taxes and other revenues are the compensating fiscal items that are used to offset changes, amongst other things, in a measure of distortionary taxation (calculated as the sum of tax revenues on income and wealth, capital and social security contributions). There is some sensitivity to the exact parameter estimate according to the methodology used; he attempts to control for endogeneity problems, but in general finds a significant negative growth effect from distortionary taxes.

One study that does not formally introduce the government budget constraint, but that includes a sufficient number of fiscal categories for us to imply its use is Romero-Ávila and Strauch (2008). Romero-Ávila and Strauch (2008) include in their regressions measures of consumption spending, total transfers, public investment and rotate the tax variables. They exclude the budget deficit/surplus. They therefore consider the effect of a change in different taxes financed by changes to the budget deficit.[11]

Arin (2004) further disaggregates taxation into income taxes, corporate taxes, indirect taxes and social security taxes, changing across the regressions exactly which tax variable is being included (which is the reason why we include the paper in the tables). The paper does not detail either whether the budget constraint has been fully specified. However, the specifications also include government outlays (government purchases, subsidies, transfers and wages) and capital outlays (government fixed capital value formation) as control variables. Broadly we might therefore interpret the coefficients on the tax variables as capturing the effect of the included tax categories financed by a change in the budget surplus/deficit and the omitted tax. While there are two items omitted with potentially non-zero growth effects, we still include the paper because it disaggregates taxation so that the omitted tax element is at least well identified.

[11] A failure to recognize the implications of excluding the budget surplus/deficit probably accounts for the sense of disappointment the authors have with their tax results.

Conditional on the overall tax burden Arnold (2008) includes measures of income taxes (in total and separated into personal and corporate income taxes) and consumption and property taxes. The results therefore show that a shift in the tax burden away from consumption and property taxes and towards income taxes is associated with lower growth, with somewhat stronger effects from corporate income taxes. A shift towards property and consumption taxes and away from income taxation has the opposite effect; it is associated with higher growth. Indeed fairly uniquely within the literature he investigates the effect of dropping different tax variables one at a time, thereby allowing comparisons of different compensating changes from the same tax policy change.

Bania et al. (2007) employ a dataset on subnational fiscal policy and explicitly consider the government budget constraint by specifying that the only offsetting (omitted) fiscal variable is productive expenditure. Reed (2008) reports one specification where the omitted fiscal variables include productive public spending and the deficit. They argue that at the level of the U.S. states, deficits are usually negligibly small compared to overall revenues and expenditures. Given that this may have held prior to the recent economic crisis, we do not exclude Reed (2008) on the basis that they do not consider the GBC. However, as argued above, we exclude both studies because the tax measures they include are fairly broad. Whether Angelopoulos et al. (2007) incorporate the GBC is debatable: they include various tax variables and the ratio of productive spending over total government expenditure. This set-up is not ideal because an increase in taxation may be offset by a decrease in the deficit, the level of total public spending or other revenue. However, given that an increase of total spending with unchanged allocation between productive and unproductive spending seems unlikely, and given that we ignore other revenue for the reason explained above, the offsetting change is most likely the deficit. We therefore include their estimates in the final tables.

Of those papers that are based on subnational data, Alm and Rogers (forthcoming) employ the richest fiscal data in the sense that they consider most tax variables including total taxes, corporate taxes, income taxes, sales taxes and property taxes in the same specifications. However, given that the focus of their paper is on whether fiscal policy affects growth at all and on the direction of the growth effects rather on their exact magnitude, the offsetting change is not clear despite the large number of tax variables. The reason is that whether the tax variables are measured in terms of state GDP or total tax revenue is not made explicit for each set of results. This is problematic as the measurement of the tax variables implicitly determines the offsetting change, for instance, whether it is public spending or another type of tax. We therefore exclude this study from the summary tables.[12]

[12] In addition, the authors admit that the coefficients are not robust which may be attributed to the fact that the offsetting change differs across specifications.

In contrast, Denaux (2005) only includes one tax variable and the share of public infrastructure investment in total spending in their regressions which is arguably a narrower fiscal aggregate compared to productive expenditure. We therefore exclude Denaux (2005) from our final tables. The results by Denaux et al. (2005) are likewise difficult to interpret because they only include proxies of infrastructure public spending (a measure of road density) and education expenditure per student (and not as a ratio to total spending or GDP) as control variables. The number of fiscal variables included in the specifications of Wildmalm (2001), Castro (2006), Colombier (2009), Doménech and García (2001) and Padovano and Galli (2002) is not sufficient to identify the offsetting change. We therefore exclude their estimates in the summary tables.

Time Series Studies

There has been only little discussion about the role of the government budget constraint in the time series literature, and, as a result, there are often several elements of fiscal policy that implicitly offset changes in taxation. From our discussion above, we can therefore not be sure what the exact policy change was and that the exact fiscal policy change can be replicated. This applies even to some of the more prominent papers including Blanchard and Perotti (2002), Dungey and Fry (2008), Mertens and Ravn (2009) and Perotti (2004).

While there are no time series studies that clearly identify the public spending category that offsets tax shocks, there are nevertheless a few papers that recognize the GBC and that suggest that the offsetting change, at least in broad terms, matters. Mountford and Uhlig (2008) is one example: they recognize that the government budget constraint gives rise to a large set of different tax shocks even if one ignores the possibility that it is possible to consider shocks to different tax measures and tax types which would increase their number further. Mountford and Uhlig (2008) first estimate the impulse response function to two 'basic' fiscal shocks including a shock to net taxes and a shock to general spending. They then consider different combinations of these 'basic' fiscal shocks which allows to specify what the offsetting change is. On the one hand, they consider the output effects of a deficit-financed tax cut where public spending remains unchanged for four quarters following the shock. On the other hand, they consider an increase in taxation that is matched by an increase in public spending. Mountford and Uhlig (2008) find that while both tax shocks imply a fall of output, a deficit-financed tax increase is less harmful for output according to their estimates.

Caldara and Kamps (2008) take a similar approach. Their results suggest that differences between the specifications where tax changes are explicitly offset by the deficit and those where the offsetting change is not considered are small. However and in contrast to Mountford and Uhlig (2008), they find that the effects of tax shocks offset by the deficit and offset by spending may have opposing signs.

Romer and Romer (2010) also consider tax shocks where the offsetting change is the deficit. Their approach is different in the sense that their dataset allows differentiating between deficit-financed tax changes and tax changes that were implemented for other reasons. However, the discussions in Romer and Romer (2010) imply that there may be some 'contamination' in the sense that deficit reductions in the past were often accompanied by public spending cuts which they do not control for. Nevertheless, the results by Romer and Romer (2010) suggest that the effects of tax changes offset by the deficit and those offset by a combination of public spending and the deficit differ in terms of the sign.

In other papers, the offsetting change is less explicit, but it can still be imperfectly identified because they also estimate the response of government expenditure to changes in taxation. When the change of expenditure is close to zero as in Caldara and Kamps (2008) and Claus et al. (2006), the government budget constraint implies that changes in taxation are offset by the deficit. In these cases, the reason why this is the case may be related to the fact that these shocks are temporary (and hence by definition do not persist so that public spending patterns are not adjusted).

Given that only few papers unambiguously identify the offsetting change, we are unable to exclude studies based on this criterion because this would leave us with too few estimates. However, whether or under what conditions the offsetting change matters for the output effects of tax changes over the short run is not fully clear. For instance, under certain conditions, the short-run output effects of different public spending categories may be similar. It could be argued that the supply-side effects differ to a greater extent than the demand-side effects. While the latter materialize instantly, the former do not, for instance due to time-to-build requirements of infrastructure for example. This would imply that the short-run output effects are driven by the demand-side effects and may therefore be similar. For instance, spending on salaries of teachers and salaries of other civil servants could be expected to have the same demand-side effects, but the former can be expected to contribute more likely to human capital accumulation and long-run growth. Nevertheless, not being able to identify the offsetting change means that the estimates are generally less useful for policy.

4 Reliability

Panel Studies

As a last step to evaluate the usefulness of the panel studies, we assess whether their estimates are reliable. Generally speaking however, we regard studies that recognize the GBC as reliable. We therefore do not limit the number of studies further based on lacking reliability, but we only consider significant estimates.

Panel studies frequently suffer from unobserved heterogeneity and measurement error. While recognizing that the GBC eliminates one source of unobserved heterogeneity, it may be questionable if tax and

expenditure measures are comparable across countries. These problems frequently apply to studies based on cross-country data so that one solution is to use subnational data where the extent of unobserved heterogeneity is arguably smaller. However, Dahlby and Ferede (2008) is the only paper that we are able to include in the summary tables.

Identification is another issue that is important to assess the reliability of the estimates. A branch of the literature has begun to examine what happens to the domestic growth rates as a result of changes in foreign country tax rates. While interesting, as foreign taxes lie outside of the policy influence of most countries, we do not discuss this point in detail. Their inclusion or exclusion, outside of any omitted variable bias, should not affect the interpretation of any other the other domestic fiscal variables. Interestingly the evidence from Gemmell et al. (2009) suggests a strong omitted variable bias. Including foreign corporation tax measures causes the coefficient on the top rate of domestic corporate income taxation to switch from positive to negative. This is perhaps caused by the general downward trend that has occurred across all OECD countries over this time period.

Endogeneity potentially creates another problem. The fact that faster growth induces changes in *total* government expenditure or taxation is well known. Distortionary taxes such as capital and personal income taxes would be expected to be income-elastic and therefore pro-cyclical (both absolutely and relative to non-distortionary taxes). As a result they would be expected to rise, as a share of GDP, when income grows more rapidly and fall (or rise more slowly) when income grows slowly. Endogeneity arguments may therefore lower (downward bias) the estimated relationship between taxes and growth. There may be additional effects through the budget constraint: expenditures and budget surpluses are also known to vary over the business cycle.

The standard econometric response to this problem is to find some instruments for the endogenous variable. Unfortunately good instruments are difficult to find and so researchers more commonly used lagged values of the fiscal variables. Using such an approach Gemmell et al. (2009, 2011) find that the corrected coefficients on taxation do not differ strongly from those estimated without instrumenting. As well as instrumenting using lagged values, Arnold (2008) also tries removing the correlation with the business cycle by regressing his measures on the output gap. Unfortunately the regressions are somewhat modified compared to those that do not use an instrumental variable approach so that it is not possible to comment exactly on the size of any bias. However the total tax variable, which should be robust to these other changes in specification, has a very similarly sized effect. The idea that endogeneity has only a small effect on the parameter estimates contrasts with the evidence presented in Benos (2009). Here, again using lagged values, the coefficient on distortionary taxation is between 7 and 19 times larger depending on the exact specification chosen. Alternatively, studies have used 5-year averages to minimize problems related to unobserved endogeneity.

Time Series Studies

In the summary tables, we include most time series studies, but omit all specifications which result in insignificant estimates (i.e. when the error band includes zero). However, the studies seem to differ in terms of the reliability of their estimates with more recent studies appearing to provide more robust estimates.

One central issue that affects the reliability of the time series estimates is the identification strategy of tax shocks. There are different ones used in the literature. The sign-restriction approach identifies fiscal shocks via sign restrictions on the impulse responses. The identification approach by Blanchard and Perotti (2002) relies on institutional information about the tax and transfer systems and about the timing of tax collections in order to identify the automatic response of taxes and government spending to economic activity. The recursive approach implies a causal ordering of the model variables (Caldara and Kamps, 2008). The empirical results are often not robust to the choice of the identification strategy which is demonstrated by Caldara and Kamps (2008). They estimate impulse response functions to tax shocks using quarterly data of the U.S. from 1955 to 2006 and find that the choice of the identifying assumption matters.

Romer and Romer (2010) adopt a narrative approach, but use a single equation framework. They argue that tax changes occur for a number of reasons, some of which are correlated with other developments in the economy, which gives rise to an omitted variable bias. They use narrative sources such as presidential speeches or reports prepared by the government or the Congress to identify exogenous tax changes. These tax changes were implemented for exogenous reasons, which include the reduction of an inherited budget deficit or the promotion of long-run growth, and not in response to changes in output or to offset changes in spending. Using narrative information to identify tax changes has a number of advantages. Romer and Romer (2010) argue that this may be the only way to successfully deal with the omitted variable bias as many tax changes are correlated with other macro variables. Mertens and Ravn (2009) and Favero and Giavazzi (2010) further argue that a structural VAR framework is unattractive to study the impact of anticipated tax shocks.

Most time series studies may suffer from unobserved heterogeneity because the offsetting change is often unclear and the tax change itself is often only vaguely specified and may entail changes in the tax mix and the tax structure. However, it is equally possible that alternative offsetting changes and tax mix and tax structure changes have similar effects over the short-run implying that this is not a source of unobserved heterogeneity.

5 Estimated Output Responses to Tax Reforms

Panel Studies

In Tables 2 to 5 we summarize the long-run and significant coefficient estimates with respect to the growth effects of personal income taxes, corporate taxes, distortionary taxes in general and consumption as well as property taxes, respectively.[13] These tables only summarize those estimates that we regard as useful for policy based on the criteria discussed in the previous sections. In most cases, there are a several offsetting changes considered which are specified in the tables and labeled as 'omitted variables'. On a number of occasions where the same compensating changes are being considered, different measures of the same tax are employed. Most papers include similar control variables in each specification which are summarized in the tables as well and which can be categorized as follows: initial GDP, measures of human capital, physical capital, or labor force growth.

Before discussing these estimates, we note an issue about the comparability of the coefficients. Most of the studies do not take logs of the fiscal variables (this is not possible if the budget deficit/surplus is being used). As such the estimated coefficients imply that the marginal effect on growth (or the log difference in GDP) is constant at all values of the fiscal variable (as opposed to the elasticity being constant in a double log model). This makes it difficult to directly compare the estimated coefficients where different measures of the same tax variable are used, for example comparing implicit average and top statutory rates, as they will depend on the unit of measurement. Unfortunately most of the studies do not report summary statistics and so it is also difficult to calculate elasticity values from the data (we would normally do this at one specific value of the right-hand side variables, most commonly the mean). Where they use the same type of measure, the magnitude of the coefficient estimates are broadly comparable. In some cases, comparisons are however further complicated because the studies use different dependent variables including per capita growth, absolute growth and per capita GDP levels.

The estimates with respect to personal income taxation are reported in Table 2. For personal taxation, several measures of the tax burden are considered including effective rates, implicit average rates, top statutory rates and the shares of personal income tax revenue in total revenue. The estimated coefficients appear to be consistently negative. One interesting comparison is that between Arin (2004) and Dahlby and Ferede (2008). Both consider the effect of an increase in personal income taxation financed by a reduction in the budget deficit and other tax revenues. Arin (2004) finds the estimated coefficient of -0.2 using an implicit average measure and Dahlby and Ferede (2008) -0.05 to -0.065 using a top statutory tax rate. The coefficients partly reflect differences in the units of measurement. As the papers both report summary statistics we use these to estimate elasticities (calculated at the mean values). The elasticities on income taxation in the two studies are very similar at around -0.67 from Dahlby and Ferede (2008) and

[13] One paper that we mainly exclude because the coefficients are not significant includes Aiginger and Falk (2005) who include total taxation and social security contributions in their specification.

-0.8 from Arin (2004).[14] The estimates provided by Gemmell et al. (2009) allow us to compare how the growth effects of the implicit average measures differ when top statutory taxes are included. When excluding both non-distortionary taxes and non-productive expenditures the coefficient on distortionary taxation is significant and negative and those on the top statutory rate of personal income taxation are in the range of -0.02 to -0.04 and therefore smaller. The remaining estimates on personal income taxation are taken from Arnold (2008). These capture the effect of a change in the tax mix, financed by reductions in other types of taxes holding total revenues constant, and they use the log of GDP per capita as dependent variable. In addition, given that the unit of measurement of the tax burden is different, direct comparisons of the results are difficult; his coefficient estimates are in the range of -0.2 to -1.4.

For corporate taxation, the coefficients which are summarized in Table 3 are also generally negative. The exception to this occurs in the study by Angelopoulos et al. (2007) where using a measure of the top marginal tax rate and essentially holding public spending constant as argued above (i.e. this cannot be the compensating fiscal term) results in a coefficient of 0.05. This contrasts with the estimates of Dahlby and Ferede (2008) which differ in using the top statutory rate and in including total expenditures and the top rate of personal income taxation. They find that the effect of corporate taxation is negative and between -0.11 and -0.16. In turn these estimates are similar to those in Gemmell et al. (2009), also using the top statutory rate and including productive expenditures and the deficit (estimates are in the range of -0.04 to -0.13).

As a final exercise, in order to compare these estimates to the estimated elasticity for income taxation, we calculate the elasticity between corporate taxation and growth implied by the Dahlby and Ferede (2008) study. The elasticity is greater than that for personal income taxation, between -0.91 and -1.26 per cent depending on the exact parameter chosen (for the mean value of corporate income taxes). This general conclusion, that the growth effects of corporate taxes are stronger than those for income taxes, is supported by the evidence presented in Gemmell et al. (2009) and Arnold (2008).

Table 4 summarizes the estimates with respect to distortionary taxation in general. Inevitably, the range of tax measures used is narrower because several tax types are summed together and include the implicit average tax rate and tax shares. Close consideration of Table 4 suggests that for distortionary taxation the most common compensating changes chosen are non-productive expenditures, non-distortionary taxation or both. Here the effect of an increase in taxation on growth is around -0.4 in most cases when the dependent variable is some measure of growth. The stability of the coefficient across several specifications is perhaps explained by the choice of compensating fiscal items, both of which are predicted to have zero growth effects. There are some exceptions to this. In Gemmell et al. (2009) the coefficient on distortionary taxation is smaller at around -0.21. These regressions include measures of the top statutory

[14] Arin (2004) does not present standard deviations so that we cannot take this comparison further.

rate of the income tax, suggesting that the implicit average tax measures used elsewhere probably capture in part the effects of changes to this variable. The estimates from Benos (2009) are outliers compared to these other studies. There the coefficients of distortionary taxation are between -0.08 and -1.52. The estimated coefficient of -0.08 comes from an estimate by OLS (-0.10 using fixed effects), and the higher figures refer to GMM (instrumental variable) estimates. Under one of the GMM estimates listed the coefficient estimate on distortionary taxation is -0.54. We therefore consider the plausible range for the coefficient estimate to be up to -0.54. The estimated coefficient on distortionary taxes in Arnold (2008) is -0.98, about twice the effect from above, but in this case the variables are expressed as a ratio to total revenue rather than GDP, and GDP per capita in levels is the dependent variable. Here the coefficient reflects a change in the tax mix towards income taxation, by decreasing all non-distortionary taxes and holding total revenue constant. The paper does not report summary statistics which would allow calculating elasticities to ease comparisons with the results above. Finally, Romero-Ávilla and Strauch (2008) consider a change in direct taxation compensated by a change in the budget surplus/deficit (and other tax revenues). Here the effect of an increase in taxation on growth would seem poorly identified (it differs between -0.04 and 0.05). The regressions from which these come are some attempts to control for endogeneity problems. The most standard approach, a GMM estimator, yields a coefficient that is not statistically significantly different from zero at -0.01 which nevertheless suggests that the GMM methodology generates results that can be very different from other panel approaches. In addition, Romero-Ávilla and Strauch (2008) include relatively few control variables which might be another reason about why their results diverge from the remaining papers.

The paper by Arnold (2008) is of further interest because it rotates the implicit financing variable such that we can consider the effects of changes in property or consumption taxation, conditional on reductions in personal or corporate income taxation. The coefficients which are summarized in Table 5 suggest that increasing the share of these types of taxes benefits growth, and that the output effects from changes in property taxation are larger than those for consumption taxes.

Time Series Studies

In Tables 6 to 9 we summarize the estimated short-term output response to positive tax shocks (i.e. tax increases) in terms of tax multipliers for time studies that are significant.[15] We make the estimates comparable as follows. We calculate tax multipliers as the ratio of the estimated output change to the change in tax revenue if not reported in the papers. For papers that report the output response to a tax cut, we change the sign of the estimate. In the tables, we report the peak response (which we approximate from the graphs that show the impulse response functions if not reported in the paper) and in the 'time horizon' column the number of quarters when the peak is reached. While we recognize that differences in peak estimates may conceal similarities in the overall evolution of the output response, the

[15] See de Castro and de Cos (2008) and Claus et al. (2006) for additional tables that summarize the estimated output responses to tax shocks.

advantage of approximations of peak estimates is that they can be extracted from every paper that includes an impulse response. There are only two papers that only present point estimates including Perotti (2011) and Barro and Redlick (2009). However, their estimated tax multipliers are mostly within the range of estimated peak responses.

We subdivide the studies by the type of tax variable they consider. Table 6 summarizes the estimated effects of an increase of revenue from net taxation, Table 7 summarizes the estimated effects of an increase of revenue from total taxation, Table 8 summarizes the estimated effects of an increase of revenue from some specific tax type. While Barro and Redlick (2009) partly estimate the effects of changes of the average marginal tax rate on GDP growth, they calculate from these estimates the tax multiplier with respect to total tax revenue.

In general, the estimated output response of positive shocks to total and net revenue is negative which mostly holds irrespective of whether the offsetting change is considered or not. There are exceptions to this, which can however be reconciled with the remaining estimates to some extent. Auerbach and Gorodnichenko (2010) find positive output effects for tax shocks during recessions. However, they suggest that this result is not robust to the choice of the elasticity of tax revenue with respect to output that is required for the Blanchard-Perotti identification. Caldara and Kamps (2008) find positive effects of tax shocks that finance spending increases for the recursive and Blanchard-Perotti identification approaches, which squares with the fact that find no output effects for deficit-financed tax cuts for the same identification strategies. These results contrast however with those based on the sign restriction approach. Claus et al. (2006) finds partially positive effects, but this result is not robust to the trend specification. In addition, the overall evolution of the impulse response still shares some similarities with the other specifications in the same paper where the output effects are negative.

Perroti (2005) likewise finds positive effects for several countries which cannot be explained easily. Giordano et al. (2007) find positive and statistically significant effects for Italy. Likewise, Marcellino (2002) who, like Giordano et al. (2007), uses Italian data also finds positive effects which are difficult to reconcile with the remaining studies. He suggests that this may be a consequence of the improvement in the government deficit as the spending response to the tax shock is close to zero.

A final exception is one specification in Romer and Romer (2010) which takes the deficit as the offsetting change and which also results in positive output effects. However, their results contrast with the ones by Caldara and Kamps (2008) and Mountford and Uhlig (2008) find partly negative output effects resulting from a deficit-reducing increase in net taxation.

In contrast, the magnitude of the estimates diverges, and it is difficult to make general statements. For instance, whether the shocks are anticipated or not does not seem to affect the magnitudes of the peak

tax multiplier in a specific way. We only discuss the differences in terms of the magnitudes of the effects for the narrative identification of the tax shocks because this identification strategy is increasingly seen as the most appropriate one. Romer and Romer (2010) find tax multipliers of up to -3. Favero and Giavazzi (2010) argue that tax multipliers should ideally be estimated based on the narrative identification introduced by Romer and Romer (2010) but within a multivariate dynamic model contrary to the approach taken by Romer and Romer (2010). As a result, they find much smaller multipliers of about -1. In contrast, Perotti (2011) argues that the output effects of discretionary tax changes and endogenous response of taxes are likely to differ. He then shows that once this holds true, there is a downward bias of the specification by Favero and Giavazzi (2010). He therefore proposes an instrumental variable estimation as a remedy which results in tax multipliers of around 1.5 which are in between the ones estimated by Favero and Giavazzi (2010) and Perotti (2011).

Table 8 shows that the estimated output effects resulting from changes of different *tax types* likewise diverge and sometimes contradict the predictions of economic theory. While for some specifications, large parts of the impulse response functions are not significant in Arin and Koray (2005) and Arin and Koray (2006) (apart from the peak estimate), de Castro and Hernandéz de Cos (2008) also find that the peak output response to an increase in social security contributions is positive.

6 Conclusions

This paper has evaluated the usefulness of empirical studies to predict the effects of tax changes on output in OECD countries. We have selected only those studies that are based on recent OECD data and have distinguished two branches of the literature: panel studies mostly estimate the long-run impact, whereas time series studies estimate the short-run response of output to changes in taxation. We have discussed two issues in detail. First, we have evaluated if the tax changes whose effects are estimated are replicable based on whether and to what extent they take into account the government budget constraint and based on the tax measures they use. Second, we have discussed the reliability of the results.

For the panel studies, we showed that difficulties to replicate the tax changes and to interpret the results implied that some studies are not suitable for policy advice. In contrast, lacking reliability appeared not to be a major problem for the remaining studies. While we have most confidence in Gemmell et al. (2009) and Arnold (2008), this does not mean that we discard the remaining studies that consider the government budget constraint. Generally, the growth effects of tax changes depend on the tax type and the offsetting change considered.

For the time series studies, it may be difficult to replicate the estimates in practice. However, we have not applied this criterion rigorously to time series studies because this would imply excluding most results, and because it is not clear to what extent this criterion matters over the short run: in general, there seems to

be increasingly the consensus that the output effects are negative possibly suggesting that various different tax reforms have similar affects. Nevertheless, predicting the short-run effects using these estimates for policy purposes is therefore associated with a greater uncertainty. Reliability depends on the identification strategy, and there seems to be a consensus emerging that identification based on narrative evidence of the type introduced by Romer and Romer (2010) is the most appropriate strategy. With respect to the measure of the tax variables, the use of marginal tax rates in time series studies as in the case of Barro and Redlick (2009) seems to be a promising route for future research.

Considering the time series and panel literature together implies that depending on the tax change, short- and long-run effects may have opposing signs. Long-run estimates are largely in line with theory and mostly refer to changes in distortionary taxation offset by unproductive spending or non-distortionary taxation. These tax changes generally entail negative growth effects similarly to negative tax multipliers usually found by the time series studies. In contrast, the long-run growth effects of increases of non-distortionary taxation offset by decreases in distortionary taxation may be positive. While the time series studies do not specifically estimate the effects of this particular tax shock, the fact that output multipliers are negative irrespective of the offsetting change and whether or not the role of the government budget constraint is considered suggests that the short-run output response of tax changes with positive long-run effects may well be negative. However, future research will ideally address the possibility of trade-offs between the short and long run in greater detail.

7 Bibliography

Aiginger, K. and Falk, M (2005) Explaining differences in growth among OECD countries. *Empirica*, 32, 19–43.

Alm, J. and Rogers, J. (forthcoming) Do State Fiscal Policies Affect State Economic Growth? *Public Finance Review*.

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, 885–902.

Arin, P. (2004) Fiscal Policy, Private Investment and Economic Growth: Evidence from G-7 Countries. Unpublished manuscript.

Arin, P. and Koray, F. (2005.) Fiscal Policy and Economic Activity: U.S. Evidence. Unpublished manuscript.

Arin, P. and Koray, F. (2006) Are some taxes different than others? An empirical investigation of the effects of tax policy in Canada. *Empirical Economics*, 31(1), 183–193.

Arin, P., Berlemann, M. and Koray, F. (2008) The Taxation-Growth-Nexus Revisited. Unpublished manuscript.

Arnold, J. (2008) Do tax structures affect aggregate economic growth: empirical evidence from a panel of OECD countries. *OECD Economics Department Working Paper*, No. 643.

Auerbach, A. (2009) Activist Fiscal Policy to Stabilize Economic Activity. *NBER Working Papers*, 15407.

Auerbach, A. and Gorodnichenko, Y. (2010) Measuring the Output Responses to Fiscal Policy. *NBER Working Papers*, 16311.

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, 193–204.

Barro R. (1990) Government spending in a simple model of endogenous growth. *Journal of Political Economy*, 98, 103–117.

Barro, R. and Redlick, J. (2009) Macroeconomic Effects from Government Purchases and Taxes. *NBER Working Papers*, 15369.

Benos, N. (2009) Fiscal Policy and Economic Growth: Empirical Evidence from EU Countries. Unpublished manuscript.

Blanchard, O. and Perotti, R. (2002) An empirical characterization of the dynamic effects of changes in government spending and taxes on output. Quarterly Journal of Economics 117: 1329–68.

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*, 34, 36–57.

Bretschger, L. and Hettich, F. (2002) Globalisation, capital mobility and tax competition: theory and evidence for OECD countries. *European Journal of Political Economy*, 18(4), 695–716.

Caldara, D. and Kamps, C. (2008) What are the effects of fiscal policy shocks? A VAR-based comparative analysis. European Central Bank Working Paper 877.

Canzoneri, M., Cumby, R. and Diba, B. (2002) Should the European Central Bank and the Federal Reserve be concerned about fiscal policy? *Proceedings, Federal Reserve Bank of Kansas City*, 333–389.

Castro, C. (2006) Política fiscal e crescimento económico. *Polytechnical Studies Review*, 5/6. 87–118.

Claus, I., Gill, A., Lee, B. and McLellan, N. (2006) An empirical investigation of fiscal policy in New Zealand. *New Zealand Treasury Working Paper*, 06/08.

Colombier, C. (2009). Growth Effects of Fiscal Policies: An Application of Robust Modified M-Estimator. Applied Economics, 41(7), 899–912.

Dahlby, B. and Ferede, E. (2008) Tax Cuts, Economic Growth, and the Marginal Cost of Public Funds for Canadian Provincial Governments. Unpublished manuscript.

de Castro, F. and Hernández de Cos, P. (2008) The economic effects of fiscal policy: The case of Spain. *Journal of Macroeconomics*, 30(3), 1005–1028.

Denaux, Z.S. (2005) A Cross-County Evaluation of Fiscal Policy. *Journal of Business and Public Affairs*, 32(1), 14-20.

Denaux, Z.S., Cunningham, D.J. and Allen, R.C. (2005) Spend More of Tax Less? Which way to State Economic Growth? *Journal of Business and Public Affairs*, 32(1), 21-25.

Devereux, M.P., Elschner, C. and Spengel, C. (2009) Effective tax rates in the EU, 1998-2007, Springer.

Devereux, M.P., Lockwood, B. and Redoano, M. (2008) Do countries compete over tax rates? *Journal of Public Economics*, 92, 1210–1235.

Doménech, R. and García, J. R. (2001) Estructura Fiscal y Crecimiento Económico en la OECD. Investigaciones Económicas, 15(3), 441–472.

Dungey, M and Fry, R (2007) The identification of fiscal and monetary policy in a structural VAR. Australian National University, CAMA Working Paper, 29.

Favero, C. and Giavazzi, F. (2010) Reconciling VAR-based and Narrative Measures of the Tax Multiplier. Unpublished manuscript.

Favero, C. and Giavazzi, F. (2007) Debt and the Effects of Fiscal Policy. *NBER Working Papers*, 12822.

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

Folster, S. and Henrekson, M. (2006) Growth effects of government expenditure and taxation in rich countries: a reply. *European Economic Review*, 50, 219–221.

Gemmell, N., Kneller, R. and I., Sanz (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 I., Sanz (2009) The growth effects of personal and corporate tax rates in the OECD. Unpublished manuscript.

Gemmell, N. (2004) Fiscal Policy in a Growth Framework. In: Addison, T. and Roe, A. (2004) *Fiscal Policy for Development. Poverty, Reconstruction and Growth,* Palgrave Macmillan, New York.

Giordano, R., Momigliano, S., Neri, S. and Perotti, R. (2007) The effects of fiscal policy in Italy: Evidence from a VAR model. *European Journal of Political Economy*, 23(3), 707–733.

Hall, R. (2009) By How Much Does GDP Rise if the Government Buys More Output? *NBER Working Papers*, 15496.

Hebous, S. (2009), The Effects of Discretionary Fiscal Policy on Macroeconomic Aggregates: A Reappraisal, unpublished manuscript. Available at: http://ssrn.com/abstract=1496681

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

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

Lee, Y. and Gordon, R.H. (2005) Tax structure and economic growth. *Journal of Public Economics*, 89, 1027–1043.

Marcellino, M. (2002) Some stylized facts on non-systematic fiscal policy in the Euro area. *Working Papers*, 225, IGIER, Bocconi University.

Mertens, K. and Ravn, M. (2009) Empirical Evidence on the Aggregate Effects of Anticipated and Unanticipated U.S. Tax Policy Shocks, *Centre for Economic Policy Research Discussion Paper*, 7370.

Miller S., and Russek F. (1997) Fiscal structures and economic growth: international evidence. *Economic Inquiry*, 35, 603–613.

Mountford, A. and Uhlig, H. (2008) What are the Effects of Fiscal Policy Shocks? *NBER Working Papers*, 14551.

Myles, D. (2000) Taxation and economic growth, *Fiscal Studies*, 21, 141–168.

Myles, D. (2007) Economic Growth and the Role of Taxation. Prepared for the OECD. Unpublished Report, University of Exeter and IFS.

Neri, S. (2001) Assessing the effects of monetary and fiscal policy. *Economic working papers*, 425, Bank of Italy, Economic Research Department.

Padovano, F. and Galli, E. (2002) Comparing the growth effects of marginal vs average tax rates and progressivity. European Journal of Political Economy, 18, 529–544.

Perotti (2011) The Effects of Tax Shocks on Output: Not So Large, But Not Small Either. *NBER Working Papers*, 16786.

Perotti, R. (2004) Estimating the effect of fiscal policy in OECD countries. Proceedings, Federal Reserve Bank of San Francisco.

Reed, R. (2008) The relationship between taxes and US state income growth. National Tax Journal, LXI, 57–80.

Romer, C.D. and Romer, D.H. (2010) The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks, American Economic Review, 100(3): 763–801.

Romero-Ávila, 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, 172–191.

Romp, W. and J. de Haan (2007) Public capital and economic growth: a critical survey, Perspektiven der Wirtschaftspolitik 8 (1), 6–52.

Straub, J. (2008): Infrastructure and growth in developing countries: recent advances and research challenges, Washington, DC: World Bank (World Bank Policy Research Working Paper 4460).

Wildmalm, F. (2001) Tax structure and growth: are some taxes better than others? Public Choice, 107, 199–219.

8 Appendix

Table 1: Overview of panel studies that are evaluated

Type	Paper	Data
Cross-country panel	Aiginger and Falk (2005)	OECD, 1970 – 1999
	Angelopoulos et al. (2007)	23 OECD, 1970 – 2000
	Arin (2004)	G7, 1965 – 2000
	Arin et al. (2008)	UK, US, Scandinavia, *years differ*
	Arnold (2008)	21 OECD, 1971 – 2004
	Romero-Ávila and Strauch (2007)	EU-15, 1971 – 2001
	Benos (2009	14 EU, 1990 – 2006
	Bleaney et al. (2001)	OECD, 1970 – 1995
	Castro (2006)	EU-15 (except Lux), 1970 – 2000
	Colombier (2009)	21 OECD, 1970 – 2001
	Doménech and García (2001)	OECD, 1960 – 1995
	Gemmell et al. (2009)	17 OECD, late 1970s - 2004
	Kneller et al. (1999)	22 OECD, 1970 – 1995
	Padovano and Galli (2002)	25 developed countries, 1970 – 1998
	Widmalm (2001)	23 OECD, 1965 – 1990
Subnational panel	Alm and Rogers (forthcoming)	48 U.S. states, 1947 – 1997
	Bania et al. (2007)	49 U.S. states, 1962 – 1997
	Dahlby and Ferede (2008)	Canadian provinces, 1977 – 2006
	Denaux (2005)	North Carolina counties, 1980 – 1995
	Denaux et al. (2005)	48 U.S. states, 1969 – 1988
	Reed, R. (2008)	48 U.S. states, 1970 – 1999
Time series	Arin and Koray (2005)	Several OECD, around 1960 – 2001
	Auerbach and Gorodnichenko (2010)	U.S., 1947 – 2009
	Barro and Redlick (2009)	U.S., 1950 – 2006
	Blanchard and Perotti (2002)	U.S., 1960 – 1997
	Caldara and Kamps (2008)	U.S., 1955 – 2006
	Claus et al. (2006)	New Zealand, 1982 – 2004
	de Castro and Hernández de Cos (2008)	Spain, 1980 – 2004
	Dungey and Fry (2008)	New Zealand, 1983 – 2006
	Favero and Giavazzi (2010)	U.S., 1950 – 2007, quarterly
	Giordano et al. (2007)	Italy, 1982 – 2004
	Marcellino (2002)	Italy, 1981 – 2001
	Mertens and Ravn (2009)	U.S., 1947 – 2006
	Mountford and Uhlig (2008)	U.S., 1955 – 2000
	Perotti (2004)	Several OECD, 1960 – 2001
	Perotti (2011)	U.S., 1945 – 2009, quarterly
	Romer and Romer (2010)	U.S., 1947 – 2007

Table 2: Overview of estimated growth effects of personal income taxation (Part 1)

Study		**Angelopoulos et al. (2007)**	**Arin (2004)**	**Dahlby and Ferede (2008)**	**Gemmell et al. (2009)**	**Gemmell et al. (2009)**
Tax measure		effective	IATR	top statutory	top statutory	top statutory
Endogenous Variable		growth (in %)	p.c. growth	p.c. growth (in %)	growth (in %)	growth (in %)
Coefficient	Low	-0.174	-0.199	-0.065	-0.033	-0.039
	High	-0.118		-0.054	-0.018	-0.034
Included (I) / omitted (O) 'level' variables						
total exp.				I		
tax revenue						
deficit		O	O	O	I	I
productive expenditure		I			I	I
non-productive exp.		I			O	O
other expenditure						
government cons.			I			
government invest.			I			
transfers						
Included (I) / omitted (O) 'tax mix' variables						
distortionary / direct						I
non-dist. / indirect			I		O	O
other revenue / tax		O	O	O		
personal		I	I	I	I	I
corporate		I	I	I	I	I
consumption		I				
property						
'tax structure' included (I) as separate variables						
progressivity of inc. tax						
Included (I) 'control' variables						
initial GDP		I		I		
human capital		I	I		I	I
other macro variables		I		I		
physical capital		I	I	I	I	I
labor force growth		I		I	I	I

Table 2 – continued – (Part 2)

Study		**Arnold (2008)**	**Arnold (2008)**	**Arnold (2008)**	**Arnold (2008)**	**Arnold (2008)**
Tax measure		% total tax	% total tax	% total tax	% total tax	% total tax
Endogenous Variable		p.c. GDP (in logs)	p.c. GDP (in logs)	p.c. GDP (in logs)	p.c. GDP (in logs)	p.c. GDP (in logs)
Coefficient	Low High	-1.13	-0.2	1.01	-0.96	-1.35
Included (I) / omitted (O) 'level' variables						
total exp.						
tax revenue		I	I	I	I	I
deficit						
productive expenditure						
non-productive exp.						
other expenditure						
government cons.						
government invest.						
transfers						
Included (I) / omitted (O) 'tax mix' variables						
distortionary / direct						
non-dist. / indirect		O				
other revenue / tax			O			
personal		I	I	I	I	I
corporate		I		O	I	I
consumption				I	O	I
property				I	I	O
'tax structure' included (I) as separate variables						
progressivity of inc. tax			I			
Included (I) / 'control' variables						
initial GDP						
human capital		I	I	I	I	I
other macro variables						
physical capital		I	I	I	I	I
labor force growth		I	I	I	I	I

Table 3: Overview of estimated growth effects of corporate taxation

Study		**Angelo-poulos et al. (2007)**	**Dahlby and Ferede (2008)**	**Gemmell et al. (2009)**	**Gemmell et al. (2009)**	**Gemmell et al. (2009)**	**Arnold (2008)**	**Arnold (2008)**	**Arnold (2008)**	**Arnold (2008)**
Tax measures		top statutory	top statutory	statutory	statutory	effective	% total tax	% total tax	% total tax	% total tax
Endogenous variable		growth (in %)	p.c. growth (in %)	growth (in %)	growth (in %)	growth (in %)	p.c. GDP (in logs)	p.c. GDP (in logs)	p.c. GDP (in logs)	p.c. GDP (in logs)
Coefficient	low	0.047	-0.158	-0.13	-0.13	-0.161	-2.01	-1.18	-2.04	-2.4
	high		-0.108	-0.035		-0.049				
Included (I) / omitted (O) 'level' variables										
total exp.			I							
tax revenue							I	I	I	I
deficit		O	O	I	I	I				
productive expenditure		I		I	I	I				
non-productive exp.				O	O	O				
other expenditure										
government cons.										
government invest.										
transfers										
Included (I) / omitted (O) 'tax mix' variables										
distortionary / direct					I					
non-dist. / indirect				O	O	O	O			
other revenue / tax		O	O							
personal		I	I	I	I	I	I	O	I	I
corporate		I	I	I	I	I	I	I	I	I
consumption								I	O	I
property								I	I	O
Included (I) 'control' variables										
initial GDP		I	I							
human capital		I		I	I	I	I	I	I	I
other macro variables		I	I							
physical capital		I	I	I	I	I	I	I	I	I
labor force growth		I	I	I	I	I	I	I	I	I

Table 4: Overview of estimated growth effects of distortionary taxation (Part 1)

Study		**Romero-Ávila and Strauch (2008)**	**Benos (2009)**	**Gemmell et al. (2009)**	**Bleaney et al. (2001)**	**Bleaney et al. (2001)**	**Bleaney et al. (2001)**
Tax measures		IATR	IATR	IATR	IATR	IATR	IATR
Endogenous variable		p.c. growth (in %)	p.c. growth (in %)	growth (in %)	p.c. growth (in %)	p.c. growth (in %)	p.c. growth (in %)
Coefficient	**low**	-0.04	-1.516	-0.208	-0.467	-0.854	-0.427
	high	0.05	-0.077			-0.411	-0.393
Included (I) / omitted (O) 'level' variables							
total exp.							
tax revenue							
deficit		O	I	I	I	I	I
productive expenditure			I	I	I	I	I
non-productive exp.			I	O	O	O	I
other expenditure			O		I	I	I
government cons.		I					
government investment		I					
transfers		I					
Included (I) / omitted (O) 'tax mix' variables							
distortionary / direct		I	I	I	I	I	I
non-dist. / indirect		I	O	O	I	O	O
other revenue / tax					I	I	I
personal				I			
corporate				I			
consumption							
property							
Included (I) 'control' variables							
initial GDP							
human capital		I	I	I			
other macro variables			I		I	I	I
physical capital			I	I	I	I	I
labor force growth			I	I	I	I	I

Table 4 – continued – (Part 2)

Study		**Kneller et al. (1999)**	**Kneller et al. (1999)**	**Kneller et al. (1999)**	**Arnold (2008)**
Tax measures		IATR	IATR	IATR	% total tax
Endogenous variable		p.c. growth (in %)	p.c. growth (in %)	p.c. growth (in %)	p.c. GDP (in logs)
Coefficient	**low**	-0.467	-0.463	-0.446	-0.98
	high	-0.41	-0.41	-0.427	-0.28
Included (I) / omitted (O) 'level' variables					
total exp.					
tax revenue					I
deficit		I	I	I	
productive expenditure		I	I	I	
non-productive exp.		O	O	I	
other expenditure		I	I	I	
government cons.					
government investment					
transfers					
Included (I) / omitted (O) 'tax mix' variables					
distortionary / direct		I	I	I	I
non-dist. / indirect		I	O	O	O
other revenue / tax		I	I	I	
personal					
corporate					
consumption					
property					
Included (I) 'control' variables					
initial GDP		I	I	I	
human capital					I
other macro variables		I	I	I	
physical capital		I	I	I	I
labor force growth		I	I	I	I

Table 5: Overview of estimated growth effects of consumption and property taxes

Study	**Arnold (2008)**	**Arnold (2008)**	**Arnold (2008)**	**Arnold (2008)**	**Arnold (2008)**
Tax measures	% total tax	% total tax	% total tax	% total tax	% total tax
Log / level	p.c. GDP (in logs)	p.c. GDP (in logs)	p.c. GDP (in logs)	p.c. GDP (in logs)	p.c. GDP (in logs)
Tax categories					
Consumption	0.93	0.74	0.92	2.16	n/a
Property		1.45	1.55	2.71	0.71
Included (I) / omitted (O) 'level' variables					
total exp.					
tax revenue	I	I	I	I	I
deficit					
productive expenditure					
non-productive exp.					
other expenditure					
government cons.					
government investment					
transfers					
Included (I) / omitted (O) 'tax mix' variables					
distortionary / direct	O	O			
non-dist. / indirect					
other revenue / tax					
personal			O	I	I
corporate			I	O	I
consumption	I	I	I	I	O
property	I	I	I	I	I
Included (I) 'control' variables					
initial GDP					
human capital	I	I	I	I	I
other macro variables					
physical capital	I	I	I	I	I
labor force growth	I	I	I	I	I

Table 6: Overview of time series studies and estimates (net tax measures)

Paper	Data	Specification		Estimated Effects		
		Identification strategy	Offsetting change	Time horizon[16]	(Peak) tax multiplier[17]	Comments
Auerbach and Gorodnichenko (2010)	U.S., 1947 – 2009, quarterly	Blanchard and Perotti	?	12 quarters	-0.99	
Auerbach and Gorodnichenko (2010)	U.S., 1947 – 2009, quarterly	Blanchard and Perotti	?	3 quarters	0.3	Estimated tax multiplier refers to recessions
Auerbach and Gorodnichenko (2010)	U.S., 1947 – 2009, quarterly	Blanchard and Perotti	?	2 quarters	-0.50	Estimated tax multiplier refers to expansions
Blanchard and Perotti (2002)	U.S., 1960 – 1997, quarterly	Blanchard and Perotti	spending response small	5 quarters	-0.78	
Blanchard and Perotti (2002)	U.S., 1960 – 1997, quarterly	Blanchard and Perotti	spending response small	7 quarters	-1.33	
Blanchard and Perotti (2002)	U.S., 1960 – 1997, quarterly	Blanchard and Perotti	?	6 to 7 quarters	-0.8 to -1.5	Anticipated shocks
Caldara and Kamps (2008)	U.S., 1955 – 2006, quarterly	sign restriction	deficit	4 to 6 quarters	-1.2 to -1.9	
Caldara and Kamps (2008)	U.S., 1955 – 2006, quarterly	recursive	spending	14 quarters	1.6	
Caldara and Kamps (2008)	U.S., 1955 – 2006, quarterly	Blanchard and Perotti	spending	14 quarters	1.6	
Caldara and Kamps (2008)	U.S., 1955 – 2006, quarterly	sign restriction	spending	20 quarters	-1.6	
Claus et al. (2006)	New Zealand, 1982 – 2004, quarterly	Blanchard and Perotti	small spending response	1 – 3 quarters	-0.25 to 0.22	Different results arise due to different trend specifications

[16] Depending on the paper, the number of quarters to reach the peak response is an approximation from the impulse response function when the exact value is not reported in the paper.
[17] Depending on the paper, the tax multiplier may be an approximation from the impulse response function when the exact value is not reported in the paper.

Paper	Data	Specification		Estimated Effects		
		Identification strategy	Offsetting change	Time horizon[16]	(Peak) tax multiplier[17]	Comments
de Castro and Hernández de Cos (2008)	Spain, 1980 – 2004, quarterly	Blanchard and Perotti	?	approx. 28 quarters	-.125	
Giordano et al. (2007)	Italy, 1982 – 2004, quarterly	Blanchard and Perotti	?	5 to 8 quarters	0.08 to 0.16	Tax multiplier refers to GDP net of public consumption
Mountford and Uhlig (2008)	U.S., 1955 – 2000, quarterly	sign restriction	?	10 quarters	-0.5	
Mountford and Uhlig (2008)	U.S., 1955 – 2000, quarterly	sign restriction	?	14 quarters	-0.2	Anticipated shock
Mountford and Uhlig (2008)	U.S., 1955 – 2000, quarterly	sign restriction	deficit	15 quarters	-0.5	
Mountford and Uhlig (2008)	U.S., 1955 – 2000, quarterly	sign restriction	spending	12 quarters	-0.78	
Perotti (2004)	Canada, 1960 – 2001, quarterly	Blanchard and Perotti	?	6 quarters	-0.5	
Perotti (2004)	Germany, 1960 – 1989, quarterly	Blanchard and Perotti	?	6 quarters	-0.7	
Perotti (2004)	UK, 1960 – 2001, quarterly	Blanchard and Perotti	?	9 quarters	0.25	
Perotti (2004)	U.S., 1960 – 2001, quarterly	Blanchard and Perotti	?	7 quarters	-0.5	
Perotti (2004)	Australia, 1960 – 2001, quarterly	Blanchard and Perotti	?	4 quarters	0.5	Anticipated shock

Table 7: Overview of time series studies and estimates (total tax measures)

Paper	Data	Specification		Estimated Effects		
		Identification strategy	Offsetting change	Time horizon	(Peak) tax multiplier[18]	Comments
Arin and Koray (2006)	Canada, 1960 – 1999, quarterly	recursive	?	7 quarters	-0.16	
Barro and Redlick (2009)	U.S., 1950 – 2006, annual	IV	?	4 quarters	-1.14	Estimates based on effects of change in marginal tax rates on per capita growth
Favero and Giavazzi (2010)	U.S., 1950 – 2007, quarterly	narrative	?	9 quarters	-1	Multiplier refers to fiscal VAR specification
Favero and Giavazzi (2010)	U.S., 1950 – 2007, quarterly	Blanchard and Perotti	?	10 quarters	-1	
Marcellino (2002)	Italy, 1981 – 2001, semiannual	Blanchard and Perotti	small spending response	3 quarters	1	
Mertens and Ravn (2009)	U.S., 1947 – 2006, quarterly	narrative	no	10 quarters	-2	only baseline specification
Mertens and Ravn (2009)	U.S., 1947 – 2006, quarterly	narrative	no	8 quarters	-1.5	Anticipated shock
Perotti (2011)	U.S., 1945 – 2009, quarterly	narrative	?	12 quarters	-1.32 to -1.68	Tax multiplier point estimate which may not coincide with peak; IV specifications only
Perotti (2011)	U.S., 1945 – 2009, quarterly	narrative	?	12 quarters	-1.68	Anticipated shock
Romer and Romer (2010)	U.S., 1947 – 2007, quarterly	narrative	no	10 quarters	-2.93 to -3.08	
Romer and Romer (2010)	U.S., 1947 – 2007, quarterly	narrative	mostly deficit	10 quarters	2.48	

Table 8: Overview of time series studies and estimates (disaggregated tax measures)

Paper	Data	Specification		Estimated Effects			
		Identification strategy	Offsetting change	Tax variable	Time horizon	(Peak) tax multiplier	Comments
Arin and Koray (2005)	U.S., 1959 – 2001, quarterly	recursive	spending response not significant	indirect tax revenue	7 quarters	negative	
Arin and Koray (2005)	U.S., 1959 – 2001, quarterly	recursive	no	corporate tax revenue	3 quarters	negative	

[18] Depending on the paper, the tax multiplier may be an approximation from the impulse response function when the exact value is not reported in the paper.

Paper	Data	Specification		Estimated Effects			
		Identification strategy	Offsetting change	Tax variable	Time horizon	(Peak) tax multiplier	Comments
Arin and Koray (2006)	Canada, 1960 – 1999, quarterly	recursive	no	indirect tax revenue	6 quarters	-0.09	
Arin and Koray (2006)	Canada, 1960 – 1999, quarterly	recursive	no	corporate tax revenue	3 quarters	0.06	
Arin and Koray (2006)	Canada, 1960 – 1999, quarterly	recursive	no	personal income tax revenue	4 quarters	0.05	
Arin and Koray (2006)	Canada, 1960 – 1999, quarterly	recursive	no	social security tax revenue	6 quarters	0.01	
de Castro and Hernández de Cos (2008)	Spain, 1980 – 2004, quarterly	Blanchard and Perotti	no	direct taxes	25 quarters	-0.2	
de Castro and Hernández de Cos (2008)	Spain, 1980 – 2004, quarterly	Blanchard and Perotti	no	social security contributions	28 quarters	0.2	