ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics

The Tax Multiplier in the Netherlands: New Estimates from a Narrative Study

Master's Thesis
September 2017

Sander Geenen
344058

supervised by
prof.dr. B. Jacobs

Author note

Correspondence concerning this thesis and requests for the companion paper "Narrative Dataset of Legislated Tax Measures in the Netherlands: 1955-2015" should be addressed to the author at s.geenen@live.nl.

Abstract

This paper studies the effect of tax changes on output for the Netherlands between 1960 and 2015. I use the narrative record of parliamentary reports, the Budget Memorandum (BM) and Tax Plan to identify tax changes that are exogenous to output. Using simple OLS, I show that the effect of tax on output is significant and contractionary, yet mostly in the same quarter as the introduction of a tax change. The output effect has faded after two years. I estimate implied impact multipliers of around -0.9% of GDP and, depending on the model specification, a total multiplier of between -1% and -2% of GDP following a tax shock of 1% of GDP.

Keywords: tax, fiscal, multiplier, narrative

Acknowledgements and Preface

I am greatly indebted to my supervisor Bas Jacobs for his encouragement in and advice on writing this thesis. Professor Jacobs taught me a great deal and I have taken much pleasure in attending his lectures and seminars. His enthusiasm and way of teaching encouraged me to continuously improve my knowledge and ability (always explain the mechanism!) and to pursue a career as policy economist.

I would also like to thank my girlfriend Joyce. She is the living proof that brilliant economists can be warm and loving people. I treasure our many discussions, on economic policy and much more. These and her feedback on my work have undoubtedly enriched this thesis.

Finally, but not last, I would like to thank my parents and sister. They have been an invaluable rock during my studies and have always encouraged me to have confidence in my abilities. I am very grateful to my parents for the opportunities they have given me to grow as a person and to work to realise my potential.

The scope of this thesis required me to dive into modern history of Dutch tax policy. Though reading through numerous policy documents has proven to be an arduous undertaking at times, it was never dull. Dutch tax policy, I found out, has evolved to be much more than a way to finance government spending. Tax laws capture the zeitgeist. They depict, for example, a changing approach of the government in addressing issues related to the labour market, the education system and family affairs. Reading their motivation leads me to policy discussions of the time, from those on the emancipation of working women in the 1960s to the need to green the tax system in the 1990s. Exploring historic policy documents helps me to put current policy discussions in context.

These first words of this thesis are my last as a student in Policy Economics at the Erasmus University Rotterdam. They not only mark the end of my student days, but also the start of working life. The start of putting the things I learnt into practice. A moment, I admit, both exciting and thrilling.

Sander Geenen
September 2017

Table of Contents

List of Figures

List of Tables

I. Introduction

The global financial crisis and Great Recession led policymakers to explore the full extent of policy options to mitigate adverse effects on the economy. Tax policy reemerged[1] as an instrument to raise output and stabilise the economy in the short run. US Congress passed the *American Recovery and Reinvestment Act*, a stimulus package of spending increases and tax cuts totalling 6% of GDP (CBO, 2010).[2] In the EU national governments stimulated their economies with about 1.2% of GDP on average under the *2008 European Union stimulus plan* (Commission of the European Communities, 2008). The Netherlands increased tax rebates for households and firms, cancelled a proposed increase of VAT and suppressed employment contributions, amounting to 0.4% of the total economy. In March 2009 another emergency package consisting of tax rebates, social benefits and public expenditure followed, with a size of 1% of GDP (Ministry of Finance, 2010). This renewed interest in tax and budget policy is what Ramey (2016) considers "one of the few positive effects of the recent financial crisis" (p. 1).

On the other hand, the European Union continued on a path of austerity since 2011 to reduce budget deficits and public debt. The Dutch government introduced austerity measures worth 46 billion euro, of which 16 billion euro was related to the revenue side (Suyker, 2016). While fiscal austerity was hastily imposed in many countries in the EU, fiscal consolidation was far from quickly achieved (Semmler, 2013). The debate on the need and effectiveness of austerity at the time is characterised by considerable disharmony among researchers and policymakers (De Grauwe & Ji, 2013).

Literature in recent years attempted to shed light on the macroeconomic effects of tax policy. This tax multiplier is defined as the ratio of the change in output to a change in government taxes (Chinn, 2013). Studies produce tax multiplier estimates that support the full spectrum of views regarding the effect of tax on output. Results of the tax multiplier for the US, for example, range from close to 0 to nearly -5 (Mountford & Uhlig, 2009).

Measurement difficulty is cited to be a main reason why estimations of the effects of a tax change on output diverge (Mertens & Ravn, 2014; Caldara & Kamps, 2017). This has to do with the relationship between commonly used tax measures and output. Most empirical work takes a change in realised tax revenues as a measure of taxation (Cloyne, 2013). Yet, discretionary tax changes are not the only elements influencing realised tax revenues. Automatic stabilisers, global influences and tax policy responding to

[1] Beetsma (2008) describes how "until the early eighties fiscal policy was widely regarded as a useful tool for economic stabilisation... [but lost support to supply-side policy] when fiscal policies did not prevent the widespread increases in unemployment... Discretionary fiscal policy is undergoing a revival" (p. 2).

[2] I use output and GDP interchangeably throughout.

macroeconomic conditions also explain tax revenue movements. A feedback effect further exacerbates this effect: a changing level of tax revenues affects output, but a changing output contemporaneously influences tax revenues. Endogeneity of tax revenue leads to biased estimates of tax effects on output when an aggregate tax measure is used in the estimation of the macroeconomic effects of tax policy.

To deal with endogeneity, researchers attempt to identify exogenous shocks from observable tax changes. To do this, they isolate tax policy uncorrelated with GDP or with other factors affecting GDP. The literature gives account of multiple identification approaches. Their strategies and contribution to the debate on the tax multiplier are discussed in the literature review of this thesis.

A novel approach to estimating the tax multiplier is to construct an exogenous tax series from a narrative record. This narrative approach identifies exogenous tax changes by their stated motivation in policy documents. Only tax measures plausibly uncorrelated with output are used to derive estimates of the effects of tax changes on output. A key feature of the narrative approach is that it uses *projected* tax revenue changes instead of *realised* tax revenues as a measure of taxation. Romer and Romer (2010) (henceforth RR) have been the first to study the effects of tax on output in the US in this way, following the same approach as in their work on monetary policy (Romer & Romer, 2004).

This paper uses the Romer-Romer framework to estimate the tax multiplier in the Netherlands. The period of study ranges from 1960 to 2015. I develop a narrative record of legislated tax measures from policy documents and classify the changes according to their stated motivation. I estimate the tax multiplier by means of a simple regression of exogenous tax changes on output. I show that the initial effect of a 1% increase in the average tax rate in the Netherlands on average leads to a 0.9% decline in GDP growth in the same quarter. The effect of the tax shock is most pronounced in the first quarters and fades after approximately two years. The estimation results are lower than the results of RR, who observe a maximum output decline of 3% following a 1% increase in the average tax rate. I attribute these differences to the openness of the Dutch economy and the presence of strong automatic stabilisers.

This study contributes to the literature in multiple ways. This paper is the first, to the best of my knowledge, to present empirical tax multiplier estimations for the Netherlands.[3] Second, this study allows for international comparison of tax multipliers estimated in the same way. Beside RR who study the US, there are comparable studies on the tax multiplier, notably for the UK (Cloyne, 2013) and Germany (Hayo & Uhl, 2014). These

[3]The first and only other attempt is, to the best of my knowledge, Coudret (2013), who uses a Structural Vector Autoregression to estimate average tax and spending multipliers between 1996 and 2013, building on the approach of Blanchard and Perotti (2002) using a narrative method of identifying tax shocks from policy documents. He finds spending multipliers between 0.8 and 1.2, but does not reach conclusive results on the tax multiplier.

works find multipliers that are generally higher than the results of other estimation methods (multipliers are estimated to be 2.5 for the UK and 2.4 for Germany). Third, I construct a narrative record and dataset documenting all substantial postwar tax measures of the Netherlands, which may potentially form a foundation for further work on the macroeconomic effects of tax policy in the Netherlands.

The only other attempt to record tax changes for the Netherlands in this way is, to the best of my knowledge, Devries, Guajardo, Leigh, and Pescatori (2011) of the IMF. Their work presents a dataset of fiscal consolidation for 17 OECD economies for the period 1979 to 2009. The authors construct a dataset of exogenous changes in taxes and government spending. This yields eight years of tax changes and does not deliver sufficient data to study the effect of tax on output for the Netherlands. The dataset constructed for the present study is more detailed and better suited to this research.

A number of reasons make output effects of tax shocks in the Netherlands interesting to study. Dutch tax policy, as this study shows, plays a meaningful role in achieving a range of policy objectives. This results in many (289) legitimate exogenous observations. In peak years 1995 and 2003 around 25 measures were introduced. As will be shown in Section III, the majority of tax changes aims to address structural problems in the economy and is therefore suitable for the estimation of the output effects of tax changes. In addition, Dutch tax policy is well-documented. The announcement of proposed tax measures is part of an annual event and is accompanied by the publication of the extensive Budget Memorandum. The Budget Memorandum and its supplementary Tax Plan provide detailed forecasts of revenue changes caused by a change in tax rates, rebates, tax structures or fiscal facilities. Nearly all tax changes take effect on 1 January of the year following the tax announcement, resulting in an implementation lag of about 3 months. These months are used for discussion in parliament. The fruits of these discussions materialise in policy documents and parliamentary reports, making it feasible to identify exogenous tax measures. Revenue collection in the Netherlands is highly centralised,[4] diminishing the possibility of overlooking major policy changes.

Knowledge on the tax multiplier improves the precision of tax as a policy instrument. Policymakers could take interest in how tax changes affect GDP with three different policy motivations in mind. First, when influencing output is not the policy objective, policymakers may want to neutralise spill-over effects on output by taking off-setting measures. Second, policymakers may respond to a downturn in the economy by cutting taxes, aiming to influence output in the short-run. Third, when policymakers aim to reduce a budget deficit, its success will hinge on the multiplier size. In all cases, the size of the effect of tax on output is crucial in determining the size of the policy action and will yield

[4]The OECD Tax Database shows that only 4% of tax revenue in the Netherlands is raised locally.

insights in its possible output effects.

The thesis has the following structure. Section II begins by explaining the econometric problem of endogeneity in more detail and goes on to give an overview of the literature on estimation methods to date. Section III explains the identification strategy of the narrative method. Section IV describes the construction of the new dataset of exogenous tax changes. Section V describes its properties. Section VI presents the estimations of the effect of tax changes on output. These results are extended with robustness checks in Section VII. In Section VIII I develop several policy implications and I perform a counterfactual analysis. Section IX concludes that tax increases in the Netherlands have a strong and negative effect on output, yet mostly statistically significant in the same quarter as the introduction of the tax change.

II. Literature review

The literature presents a wide variety of multiplier estimations. Consulting recent surveys of the effects on tax policy on output, such as Gechert and Will (2012) and Ramey (2016), shows that there is no single value for the tax multiplier for all times. Results of the tax multiplier for the US range from close to 0 to nearly -5 (Mountford & Uhlig, 2009). Put differently, the literature empirically supports insensitivity of output to a tax change as well as an output decline of a multiple of 5.

There are roughly three reasons that likely explain the variation in multiplier estimations. First, it is difficult to identify suitable exogenous tax shocks to estimate the effect of tax on output (Cloyne, 2013). Second, characteristics of the economy play a role in determining the size of the multiplier (Ilzetzki et al., 2013). Third, multipliers may vary across stages of the business cycle (Auerbach & Gorodnichenko, 2012a).

This literature review consists of three parts. I first focus on issues related to the identification and estimation of the tax multiplier. This leads up to the second part, where I give an overview of different identification approach in the literature. The third part discusses the economic reasons for the range in estimations of the multiplier.

A. The problem of identification

In measuring the tax multiplier, the aim is to identify the causal effect running from tax policy changes to output. A relevant question a researcher can therefore ask is to what extent output is affected by a change in taxes. This relation can be written in the following way:

$$\Delta T \rightarrow \Delta Y,$$

where T is a measure of taxes and Y a measure of output.

Possibly the most simple and intuitive way the researcher can estimate this relationship is by estimating a linear regression model, most likely using Ordinary Least Squares (OLS). In this model, the independent variable could be a measure of tax changes and the dependent variable a measure of changing output. The estimated coefficient of the tax measure would give the multiplier by which output is affected by a change in taxes.

But for estimating an unbiased and consistent coefficient for the tax variable more is required. This has to do with the relationship between commonly used tax measures and output. First, many more elements than just tax changes are likely to affect output, also

contemporaneously. Second, a change in government revenues is often used as a measure of tax changes. Government revenues, however, do not only explain a change in output, but also respond to it. When output rises so will tax revenue, simply because many taxes are some function of output.

For our simple linear regression model, the relationship between tax measures and output means that the model is likely to suffer from endogeneity. A model is endogenous when the explanatory variable correlates with the error term. This identification problem arises when tax and output are related, but not (only) in the way that we expect them to be: with a causal relation that runs from tax to output.

B. Types of endogeneity in multiplier estimations

The problem of identification in tax multiplier estimations could arise as a result of three types of endogeneity. I address reverse causation, omitted variable bias and simultaneity in turn.

Reverse causation

In estimating the effects of tax on output, the aim is to measure the causal effect running from tax policy changes to output. Yet, this might not be the only relationship between tax and output. Governments that observe a change in output growth may decide to respond by changing taxes. For example, they may cut taxes when they see a recession coming. Conversely, the government may raise taxes to prevent the economy from overheating in pursuing procyclical tax policy. A change in taxes could therefore be a cause as well as a result of a changing output. Fiscal variables do therefore not only affect output, but also respond to it. So, the relation between tax and output may run both ways:

$$\Delta T \rightarrow \Delta Y$$
$$\Delta Y \rightarrow \Delta T,$$

where T is again some measure of taxes and Y a measure of output.

How can reverse causation be observed in practice? Barro (2013) points to welfare-related transfers such as unemployment benefits, which tend to be higher in bad economic times (dampening a downward effect on GDP) and lower in good economic times (dampening an upward effect on GDP). In addition, automatic stabilisers tend to be particularly

effective in trough and peak cycles, precisely at times when governments consider countercyclical policy. Multipliers may therefore be biased downwards as increased welfare-related transfers generate a moderating effect on output fluctuations, including output changes caused by tax measures.

Omitted variable bias

Omitted variable bias arises when a relevant explanatory variable is omitted from the model or when omitted factors are unobservable. Tax changes may not be taken in isolation but could be correlated with other factors affecting output. For example, policymakers may raise taxes to finance an increase in government spending. The tax change will affect output, but so will more government spending. In this case, accounting for only the tax change and not the spending increase will lead to underestimations of the multiplier, as output effects of the tax hike are partly offset by higher government spending.

A way to eliminate bias in the estimation is to control for all other factors affecting output. This in turn introduces new difficulties. First, even the best statistical offices do not record all possible factors affecting output, at least not at the required time intervals. Historical figures on government spending in the Netherlands are only available on the yearly level and not the commonly used quarterly level. Second, operationalisation of such elements is challenging. Many factors may lead to lower output. For example, a decrease in exports may coincide with a fall in domestic consumption. Disentangling the effect of tax on output from other factors affecting output is inherently difficult. Using OLS, failing to account for other factors affecting output will mean that the error term is correlated with the regressors. This violates one of the assumptions of the Gauss-Markov theorem to produce a best linear unbiased estimator (*Introductory Econometrics: A modern approach*, n.d.). Whether omitted variables will lead to over- or underestimation of the effect of tax on output depends on the magnitude of the correlation of the omitted variable with the dependent variable (tax) and independent variable (GDP).

Simultaneity

Identification problems related to simultaneity arise when tax and output are simultaneously determined. This is typically encountered when using tax revenues as a tax measure in the multiplier estimation. To illustrate this, think of tax revenue as a multiple of the tax rate (tax imposed by the government) and the tax base (total of taxable assets). Both a change in the tax rate and the tax base leads to a change in tax revenue.

$$\underbrace{\text{tax revenue}}_{\text{policy outcome}} = \underbrace{\text{tax rate}}_{\text{policy instrument}} \cdot \underbrace{\text{tax base}}_{\text{influenced by output}}$$

The tax rate is under direct control of policymakers, but the tax base is (mostly) not. The government's tax revenue, however, depends on both the tax rate and the tax base (Riera-Crichton, Vegh, & Vuletin, 2016). This means that tax revenue may change as a result of changing output through the tax base, even though no change in the tax rate is legislated. Using tax revenue as a measure of tax changes is therefore likely to also reflect non-discretionary elements, such as the movement of the business cycle.

To counter this problem, cyclically-adjusted tax revenues have been widely used as a measure of tax changes (see for example "Can Severe Fiscal Contractions Be Expansionary? Tales of Two Small European Countries" (n.d.); Blanchard and Perotti (2002); Ilzetzki et al. (2013). Here, revenues are adjusted for cyclical fluctuations of the business cycle so that revenues are taken to be at a normal level (trend growth) of economic activity.

While using cyclically-adjusted revenues will remove business cycle effects, revenue-based measures of tax policy may suffer from substantial measurement error. Riera-Crichton et al. (2016) argue that using a cyclically-adjusted tax measure erroneously attributes any change in tax revenue to discretionary behaviour of policymakers. The tax base may still change as a result of factors not under policy control. Examples of this are changing output elasticities of tax revenues over time, a changing income distribution and taxpayers's willingness and possibility to evade taxes. Though taking cyclically-adjusted revenues removes simultaneity, it likely introduces omitted variable bias when these other factors influencing the tax base are not accounted for.

C. How the literature deals with endogeneity

To estimate the effects of government spending on output, Ramey and Zubairy (2014) would ideally ask the IMF to conduct a control trial across countries, randomly assigning changes in government spending. Then simple statistical regression techniques would be suitable to estimate the effects of spending on output. Following this logic, the best way to estimate the tax multiplier would be to produce a shock to average tax rates varying across otherwise identical countries. For obvious reasons, such a natural experiment does not belong to one of our options. Researchers therefore attempt to exploit available natural experiments or to using using historical data. The literature presents multiple approaches to addressing endogeneity issues, with each their benefits

and drawbacks. I present the most notable strategies[5] and show how they differ.

The greatest share of multiplier research studies the spending multiplier. Work on the tax multiplier is scarce. Data unavailability is the main problem (Riera-Crichton et al., 2016), with especially a lack of information on tax rates. In addition, as Perotti (2008) points out, "[the effect of tax shocks] are more difficult to identify [than shocks to government spending].. and, when taxation is distortionary, their theoretical effects depend crucially on the time profile of the tax response" (p. 3).

Research on the tax and spending multiplier is mostly similar in set-up. In the overview of recent literature I therefore draw on both spending and tax multiplier studies to show the identification approaches. I confine the summary of findings at the end of this section to tax multiplier estimation results.

I present several estimation results in this section. How to interpret the signs? When I mention spending multipliers, a *positive* sign means that for any spending increase, output is positively effected, and vice versa. When I mention tax multipliers, a *negative* sign means that for any tax increase, output is negatively effected, and vice versa.

Structural macromodels

In one of the earliest approaches, behavioural equations were estimated in large-scale structural macro econometric models (Chinn, 2013).[6] The models consist of equations of consumption, investment and price adjustment. The origin of this identification type is associated with the econometric approach of the Cowles Commission and the pioneering work of Jan Tinbergen. The Cowles Foundation for Research in Economics institute was the first to introduce a probabilistic framework[7] to link Keynesian economic theory with mathematics and econometrics (Christ, 1994). At about the same time, in 1936, Tinbergen developed the first national model for the economy of the Netherlands (Hansen, 1969). The macroeconomic responses to budget and tax policy are not estimated as such, but follow implicitly once models are calibrated to the data. Multipliers are obtained by dividing a projected change in output by the projected change in government revenues. The Netherlands Bureau for Economic Policy Analysis (CPB) uses the Saffier II model, a large structural model based on 30 years of historical data for the Netherlands

[5] See the excellent survey of Ramey (2016) for a recent and extensive overview of the effects of shocks on several macroeconomic variables.

[6] Some authors do not prefer using the term 'tax multiplier' where they mean the effect of tax on output. This is because tax multiplier implies a constant tax measure—usually government revenues expressed in some currency unit—and a direct estimation of the multiplier, whereas the effect of tax on output may implicitly follow by estimating a set of equations. Following the New Palgrave Dictionary of Economics, I use tax multiplier to mean the ratio of change in output to a change in government taxes, where the measure of government taxes depends on the context and may be realised tax revenues, projected revenues or some other measure. Simply put, 'the effect of tax on output' and 'tax multiplier' are used to denote the same concept: by what multiple does output change following a tax change, either as some currency unit or as a percentage of the average tax rate.

[7] A probabilistic framework is used to estimate, on the basis of historical data, the probability of an event occurring again.

(CPB, 2010). The model computes a tax multiplier of -0.4 for the wage and income tax in the first year, implying that 1 euro increase in the wage or income tax reduces output by 0.4 euro in the first year. The maximum effect is a reduction in GDP of -1.6% after 8 years. For the VAT, the model computes that a rise in the general VAT-rate of 1% leads to an output decline of -0.2% in the first year, with a maximum contractionary effect of -1.4% in year 8.

Structural macroeconometric models have been widely used for forecasting and policy analysis in Western economies (Chinn, 2013), yet they have also received substantial criticism. The models generally need many exogenous variables uncorrelated with policy parameters to effectively model the effect of a policy shock. The main objection is perhaps best articulated by the Lucas critique to econometric policy evaluation. Robert Lucas (1976) criticised the approach of estimating statistical relationships from past data to forecast the effects of adopting a new policy, as correlations between aggregate variables tend to change whenever policy is changed. The models describe historical relations and responses to policy based on aggregate data but new policy may change the nature of these responses. Decision rules of agents may therefore not be invariant to policy changes. Superior to models based on historical data, Lucas argued, are models based on theoretical behaviour of agents invariant to policy changes.

Dynamic Stochastic General Equilibrium models

Partly in response to this criticism, Dynamic Stochastic General Equilibrium (DSGE) models were developed in recent decades. Instead of building behavioural equations based on statistical relationships of aggregate quantities, the models feature microfoundations and intertemporal considerations. This is a theoretical approach to the behaviour of agents such as firms and households. These agents maximise their objective (profit or utility) under certain constraints. The macroeconomics in the model therefore has the individual behaviour of agents as its foundation (Barro, 2007). This makes the models consistent with the principles of both macroeconomics (the outcomes usually fit aggregate economic behaviour) and microeconomics (behaviour based on microfoundations). DSGE models are widely used in macroeconomic forecasting by governments, their statistics offices and central banks. An increasing amount of studies employ the models to study the effect of spending and tax shocks. Recent examples on spending multipliers include Cogan, Cwik, Taylor, and Wieland (2010) and Zubairy (2014) for the US and Bhattarai and Trzeciakiewicz (2017) for the UK. Leeper, Traum, and Walker (n.d.) use a monetary DSGE model that incorporate different monetary-fiscal policy regimes and show, for government spending multipliers in the US, that differences in specification of the model can produce a "morass-like range of multipliers" (p. 2452).

DSGE models incorporate rational consumers. When the government increases spending or cuts taxes, agents anticipate that the government needs to finance the current expansion in the future by doing the opposite. Under this Ricardian equivalence proposition, individuals may—as a response to expansionary fiscal policy—decide to decrease consumption and start saving today (Barro, 1974). Multipliers in neoclassical models with Ricardian consumers are generally low as consumers inversely mirror tax and budget policy, potentially reducing real effects of the initial expenditure increase or tax cut (Baxter & King, 1993). New Keynesian models additionally feature nominal rigidities and imperfect competition. In these models spending multipliers are notably higher (up to 1), depending on the type of distortions and 'stringency' of the rational expectation assumption (Woodford, 2011).

The Delfi model of De Nederlandsche Bank, the Dutch central bank (DNB), uses optimising agents and clearing markets with imperfections and frictions affecting the short-run dynamics of product markets, the labour market and financial markets (DNB, 2011). Following a wage and income tax cut of 1% of GDP, output initially rises by 0.1%, and has a maximum expansionary effect of about 0.8% after 4 years.

The widely used NiGEM model of the National Institute of Economic and Social Research computes a multiplier of -0.2 for a temporary increase of the income tax of 1% of GDP for the Netherlands in the first year. A permanent increase of direct taxes with the same size leads to an output decline of 0.15%, while an indirect tax increase lowers GDP by -0.05% (Barrell, Holland, & Hurst, 2012). The contractionary effect on output in the NiGEM model fades after the first year and disappears after 10 years.

Using DSGE models to compute the responses of output to a tax policy change is a good way to avoid endogeneity encountered in linear regression models. After all, the macroeconomic outcomes of the DSGE model follow from the individual behaviour of agents and do not rely on historical data, as is the case with structural macromodels.

This is not to say that DSGE models are bias free. Korinek (2015) names several sources of potential bias. Frictions in the model, for example, imply that welfare theorems in the system need to hold. The frictions are therefore often expressed in "well-behaved analytical forms" (p. 9), making the model easier to solve, but introducing a normative bias as welfare theorems may hold more frequently than actually observed in the economy. In addition, microfoundations have been developed in a way to fit macroeconomic outcomes. Consumers are assumed to be (mostly) homogeneous and the elasticitiy of labour supply is often higher than observed to fit observed employment responses in recessions. In addition, New Keynesian literature exhibits unrealistically strong habit persistence to suit inflation rate behaviour.

Although "in the comparatively brief space of 30 years, macroeconomists went

from writing prototype models of rational expectations... similar to jumping from the Wright brothers to an Airbus 380 in one generation", (Fernández-Villaverde, 2010, p. 2), precisely the many possible specifications and choice of frictions in the model pose limitations. Much of the variation in the implicit multipliers can be ascribed to differences in frictions and transmission mechanisms of policy shocks (Lukkezen, 2013).

Vector Autoregressive models

Sims (1980) also proposes an alternative to structural macromodels and pioneers with the estimation of unrestricted reduced-form equations, now known as Vector Autoregressive models (VARs), treating all variables as endogenous. A VAR models several series in terms of their own past. This makes it possible to only use a small system of equations, with each variable modelled as a function of lags of all variables.

The result is an extension from the univariate autoregressive model (one variable depending linearly on its previous values and a stochastic term) to multivariate time series (also explaining the interactions and co-movements among a group of time series). VARs are nowadays commonly used to describe the dynamic behaviour of economic and financial time series.

The outcomes of VAR estimations can be shown by an impulse response function. These show the "response of current and future values of each of the variables to a one-unit increase in the current value of one of the VAR errors" (Stock & Watson, n.d., p. 6). In estimations of the tax multiplier, the impulse response function will show the effect of a tax shock of 1% of GDP on output. The shock reverts to zero immediately, but the effects on the output series is traced out over time, usually over quarters or years.

Sims original VAR is recursive. This means that the error terms are constructed in each regression to be uncorrelated with the error term in the preceding equation. The researcher will need to place the variables in the order that that they affect each other. For example, in the first equation of the corresponding recursive VAR, output is the dependent variable and the regressors are lagged values of both output and the tax measure. In the second equation, the tax measure is the dependent variable and the regressors are lags of both output and the tax measure *plus* the current value of output. Changing the order of variables subsequently changes the VAR equations, coefficients and residuals.

This makes VARs in principle a useful technique to estimate the output response to a policy change using timeseries, but not a suitable identification method. The VAR produces unbiased estimates of the effect of tax on output only when causality runs from tax to output exclusively. This means that the endogeneity problem will be solved when a fully exogenous measure of tax is used. When an aggregate measure of tax is used, endogeneity as described under B is likely to persist. This, on the other hand, does imply

that a VAR model is useful in combination with a measure of taxes that is fully exogenous to output. When the researcher is able to construct an exogenous tax series in other ways, a VAR may still be used to derive consistent and unbiased estimates of the effects of tax policy on output.

Structural Vector Autoregressive models

What if such exogenous measure of tax changes is unavailable? In this case, the standard VAR could be extended by imposing assumptions on the model. This is done by Structural Vector Autoregression (SVAR) models, pioneered by Blanchard and Watson (1986). The structural component of the SVAR refers to how variables are related in practice. Assumptions are based on these actual relationships. SVARs decompose all variables into expected business cycle movements and unexpected policy innovation parts, and impose restrictions on the unexpected movements "where identifying assumptions are easier to find" (Gottschalk, 2001, p. 24).

SVARs make explicit identifying assumptions to isolate the effects of policy to recover policy shocks. Blanchard and Quah (1989) use long-run assumptions, assuming that one variable may not be affected by another in the long run. They attribute variation in U.S. real GDP and unemployment to aggregate demand and aggregate supply shocks. The structural policy shocks are identified by imposing the restriction that aggregate demand has no long-run effect on the level of real GDP.

Short-run assumptions are more common in recent literature. They may state that macroeconomic shocks only affect government spending in the next period. It follows that budget changes in the current period can be explained by policy changes. A regression of these policy shocks on output provides an estimation of the multiplier. The study of Blanchard and Perotti (2002) is a standard-setting example with this type of restriction and presents an implied tax multiplier of -0.78 to -1.33.

Sign restrictions are another type of identifying assumption. In this set-up, researchers restrict the signs of certain responses to identify a business cycle shock and a policy shock. Mountford and Uhlig (2009) were to first to impose restrictions on fiscal variables by "restricting responses for a year following the shock in order to rule out shocks where government spending rises on impact but then subsequently falls after one or two quarters" (p. 961). Using U.S. data over the period 1955-2000, the authors arrive at a relatively large multiplier of -5 for deficit-financed tax cuts. Caldara and Kamps (2008) identify a business cycle shock by imposing that the impulse responses of output and taxes are positive for at least the four quarters following the shock, and a tax shock is identified by imposing that the impulse responses of taxes are positive for at least the four quarters following the shock. Using data for the US over the period 1955-2006, the study produces

an output increase of 2% following a tax cut of 1% of GDP after five quarters.

There have been numerous other innovations to SVAR models. Examples are allowing the model to switch between recessions and 'normal' economic times (Auerbach & Gorodnichenko, 2012b), or including government budget constraints and rational expectations of households (Chung & Leeper, 2007). These works only find multipliers for the effects of spending changes.

By imposing assumptions, SVARs do not require an extensive dataset. But the estimates are only as good as the assumptions that are imposed on the model. The identifying assumptions to a large extent determine how well policy shocks are isolated from other shocks to the tax measure or GDP. Endogeneity is therefore only successfully eliminated when assumptions are in accordance with actual observations suggested by the economy.

A key identifying assumption of SVARs is that structural innovations are orthogonal (Gottschalk, 2001). Orthogonality implies that the innovations of the error term of non-policy movements are uncorrelated with innovations of the error term of the policy innovation. In the specific case of tax multiplier estimations, this can be problematic. Orthogonal tax shocks cannot be correlated with other shocks to output. As the analysis in subsection *B* has shown, this is only partly true for legislated tax policy. Tax changes may be discretionary but also correlated with output, such as countercyclical policy. There is therefore a possibility that even under imposed assumptions that adhere closest to actual macroeconomic behaviour, policy and non-policy innovations are not orthogonal. This can possibly bias estimation results.

There are additional drawbacks to using SVARs for tax multiplier estimations. The model usually produces an average of the output response to policy changes. This does not account for the possibility that the economy undergoes structural changes over time and could make it difficult to find assumptions that hold true for all periods. Also, not all relations between policy and output are modelled. IMF (2010) points to the failure to account for asset price and commodity movements. SVAR estimations are furthermore linear, which fails to capture that multiplier size may depend on the state of the economy. Among others, Auerbach and Gorodnichenko (2012a) have attempted to apply non-linear SVARs to see if multipliers vary across business cycles (they do).

Military instrumental variable approach

Another approach to dealing with possible endogeneity of variables is to find an alternative spending or tax measure plausibly uncorrelated with output. This instrumental variable (IV) method uses the exogenous instrument instead of the endogenously determined variable in the estimation of the effects of policy shocks on output. Hall (1980)

and Barro (1981) argue that military spending is a suitable instrument for government spending, as military build-ups are driven by military rather than macroeconomic events.

With this in mind, Ramey and Shapiro (1998) build a narrative record of military spending shocks described in the American weekly magazine Business Week. The authors use a simple VAR to estimate the effect of these spending shocks on output and arrive at a spending multiplier of 1.2. The narrative method of using military variables has been extended since, see for example Burnside, Eichenbaum, and Fisher (2004); Ramey (2011); Barro and Redlick (2011). Availability of data poses the main challenge to using military build-ups as an instrument. Wars occur rather infrequently with large intervals and variation in military spending is limited. Hall (2009) finds that only WWII and the Korean War provide sufficient variation in military spending to use this a measure to estimate the multiplier. He argues that "historical experience on the magnitude of the multipliers only makes the case that the multiplier is above 0.5" (p. 185). Estimation results are not biased up or down, but rather too uncertain to produce a reliable value for the multiplier.

Ramey (2011) points to the importance of timing and anticipation to shocks in determining the size of the multiplier. In the specific case of shocks to military spending, she shows that consumers may anticipate increases in defence spending and that the timing of shocks matters: contrary to earlier results, she presents a fall in most components of consumption as a result of an increase in military spending.

Exploiting regional variation

Nakamura and Steinsson (2014) criticise the 'aggregate' set-up of Ramey-Shapiro and argue that "military spending is notoriously political... likely to be endogenous to regional economic conditions" (p. 755). They challenge the crucial assumption that military spending is uncorrelated with other macroeconomic variables: factors such as patriotism, a rise in taxes to finance the war effort and other macroeconomic effects plausibly affect output contemporaneously.

The authors extend the aggregate IV by exploiting local variation in the federal military funds allocation to states in the US. They use two characteristics of military spending. First, geopolitical events drive national military spending. Second, when national military spending rises by 1% of GDP, some states receive a disproportionate amount of spending (about 3%) compared to others (receiving about 0.5%). This heterogeneity is used to identify the effects of government spending on output.

Exploiting regional variation in this way comes close to a natural experiment, with a treatment group (states receiving military spending) and a control group (states receiving none or little). Endogeneity is effectively eliminated under the assumption that

the US does not embark on military build-ups because some states receive more spending than others. Spending multipliers resulting from these sub-national studies are notably higher than those found for aggregates. This is due to the fact that effects of financing are differenced out, as taxes that finance federal spending are levied at the national level (Ramey, 2016). Nakamura and Steinsson find "an open economy relative multiplier of approximately 1.5" (p. 753).

Research with a similar empirical approach of exploiting regional variation includes Acconcia, Corsetti, and Simonelli (2014), who use sudden contractions in public spending as an instrument for local government spending resulting from mafia infiltration in city councils. They find a spending multiplier of 1.7. Chodorow-Reich, Feiveson, Liscow, and Woolston (2012) study the (endogenous) transfers to states under the American Recovery and Reinvestment Act by using pre-recession Medicaid spending levels. The research looks at the creation of jobs and does not present a multiplier. Similarly, Clemens and Miran (2012) identify exogenous variation by fiscal institutions on state level, which differ in their stringency to state budget rules, and find a multiplier that is closer to findings on the aggregate level of 0.5. Shoag (2013) uses differences in returns of state pension funds, windfall shocks, that serve as a predictor of future government spending, and comes to relatively large multipliers of 2.2. Ongoing work of Serrato and Wingender (2016) exploits variation in the changing methodology of the U.S. Census to allocate federal spending for local populations. Chodorow-Reich (2017) in a recent survey shows that the average of empirical estimations for geographical cross-section fiscal spending multipliers is about 1.8. Drawing on theoretical evidence and the results from the regional studies, the analysis presents a national closed economy zero lower bound deficit-financed multiplier of 1.7 or above.

These approaches are widely used in estimating a spending multiplier, but are much less suitable for estimating the tax multiplier. The US is especially suited for exploiting regional variation given the system of states under a federal government, whereby the federal government assigns varying levels of spending to an arguably homogeneous set of states. Estimating the tax multiplier this way is more difficult because the federal government cannot assign different sub-national tax rates (a prerogative of the states).

An attempt to study the effect of a change in federal taxes on regional state-level income is provided by Hayo and Uhl (2015), using the exogenous tax series of Romer and Romer (2009). They find considerable variation in how regional income is affected. A 1% rise in tax on personal income leads to a decline in income of -0.2% in Utah to -3.7% in Hawaii.

Narrative approach

The narrative approach identifies macroeconomic shocks through non-statistical procedures in the same way as Ramey and Shapiro (1998) do for military spending. The military variable approach and the narrative method are similar in that they identify shocks from documents such as newspapers, reports and parliamentary proceedings. I differentiate, however, between the approach of Ramey-Shapiro and the narrative method by the need to use an instrumental variable. Ramey-Shapiro identify exogenous shocks by choosing an arguably exogenous measure (military spending), while identification of exogenous shocks in the narrative method takes place by evaluating motivations of policymakers. In the latter, endogeneity is eliminated when exogenous shocks are successfully isolated from endogenous changes. By only using the identified shocks exogenous to the variable of interest (e.g. output or inflation), it is possible to derive consistent and unbiased estimates of the effects of macroeconomic shocks.

The work of Friedman and Schwartz (1963) is regarded as the first to identify shocks in this way. The study recovers monetary disturbances from policy documents to estimate effects on several macroeconomic variables, showing that monetary shocks have large and real effects. Romer and Romer (1989) revisit their work and extend the dataset by adding shocks observed after WWII. In Romer and Romer (2004), the same authors present a new measure of monetary shocks of a period between 1969-1996. They identify exogenous monetary policy shocks by Federal Reserve's intentions for the federal funds rate around Federal Open Market Committee meetings. Monetary policy, they find, has a "large, relatively rapid, and statistically significant effects on both output and inflation" (p. 1055).

Hamilton (1985) and Hoover and Perez (1994) identify oil shocks, by isolating political events that caused global disruptions in the oil market. Poterba (1986) studies the impact of changing tax expectations on the taxable and tax-exempt yield spread. He exploits four major events in recent U.S. history that changed personal and corporate taxes.

Romer and Romer (2009) perform a narrative analysis of postwar tax changes in the US from 1945 to 2007. They find that legislated tax changes fall into one of the four categories: responding to a current or planned change in government spending, offsetting other influences on economic activity, reducing an inherited budget deficit, and attempting to increase long-run growth. The size as a projection of the change government revenue, timing and motivation of the tax changes are documented.

Exogenous tax changes are identified from all legislated tax changes by their stated motivation. Legitimate exogenous tax measures are 1) uncorrelated with past and future shocks to output (not taken as a result of changing output or projections of changing output), and 2) uncorrelated with past and present contemporaneous shocks to output

(not related to other factors affecting output). This results in a timeseries of exogenous tax changes. If the exogenous shocks are identified correctly, simple OLS or a VAR estimation is sufficient to yield an unbiased estimate of the effect of tax on output.

By using the newly constructed dataset of exogenous tax measures, the first application of the narrative approach to estimate the macroeconomic effect of tax changes is Romer and Romer (2010). The authors run a simple regression and a VAR of all identified exogenous tax changes on output. Following a 1% increase in the implied average tax rate in the US, they find, output decreases by nearly 3% after 10 quarters. The results are robust across several estimation specifications.

Cloyne (2013) follows the same identification strategy for the UK. He constructs a dataset of 2,500 "non-negligible" tax changes from 1945 to 2009 by analysing the Financial Statement and Budget Report and speeches of the Chancellor of the Exchequer to parliament. Estimating the multiplier by OLS and a VAR model, Cloyne finds effects of tax on output of -2.5 after 10 quarters, remarkably close to the RR estimates. In the same spirit, Hayo and Uhl (2014) find a maximum effect of -2.4 for (West) Germany for a period of 1974 to 2009.

Output effects to tax found by the narrative approach are generally large. Favero and Giavazzi (2009) even claim that the responses estimated by RR are implausibly large. They argue that the RR specification cannot be interpreted as a moving average representation of the output process. They use a proper truncated moving average representation and, using the RR dataset, produce results of typically -0.5% of GDP following a 1% increase in the average tax rate.

Perotti (2012) further investigates this claim and argues on theoretical grounds that "the discretionary component of taxation should be allowed to have different effects than the automatic response of tax revenues to macroeconomic variables" (p. 214), and that the results by Favero and Giavazzi (2009) are biased to zero. Perotti finds multipliers about half the size of RR results: a decline in output of 1.5% after 3 years following a 1% increase in the average tax rate. Mertens and Ravn (2014) reconcile these differences and argue that earlier findings of lower multipliers can be explained by an output elasticity of tax revenues assumption that is contradicted by empirical evidence, the failure to account for a changing tax base due a tax change or measurement error in narrative series of tax shocks.

D. The economics: why results vary widely

The wide range of identification approaches are likely to explain some of the variety in multiplier estimates (Reichling & Whalen, 2015). After all, the tax multiplier is not a structural parameter, but rather "a function of structural parameters and policy reaction

parameters" (Chinn, 2013, p. 1). What economic reasoning underlies this variation? The literature presents several empirical and theoretical explanations, in detail summarised by Batini, Eyraud, Forni, and Weber (2014).

Structural country characteristics affect the economy's response to fiscal shocks in normal times. Examples of such structural effects are 1) import propensities, with less trade openness leading to higher multipliers (Barrell et al., 2012), 2) labour market rigidities, with less flexibility leading to higher multipliers (Gorodnichenko, Mendoza, & Tesar, 2012) and 3) automatic stabilisers, with large stabilisers leading to lower multipliers (Dolls, Fuest, & Peichl, 2012). Also economies with flexible exchange rate regimes and countries with high government debt levels tend to have smaller multipliers (Ilzetzki et al., 2013), as well as economies with governments that have a poor revenue administration (Batini et al., 2014).

Business cycle factors could also influence the size of output effects to fiscal shocks. Tax and spending multipliers tend to be notably larger in recessions than in expansions. First, multiplier size is influenced by income and liquidity constraints of households (Mittnik & Semmler, 2012). In bad economic times consumers are less likely to save and they may find it harder to borrow. Credit constrained households, Roeger and in 't Veld (2009) argue, raises the "marginal propensity to consume out of transitory tax reductions and increases in transfers" (p. 2). Through consumption, the effects of a spending increase or tax cut on output are therefore likely to be higher when credit constraints are binding.

Second, multipliers might depend on the degree of monetary accommodation to fiscal shocks. When the interest rate is near or at zero, the transmission mechanism of monetary policy is impaired (Woodford, 2011; Ramey & Zubairy, 2014). This might lead to higher tax and spending multipliers. In normal times, the effects of an increase in government spending or a tax cut on the economy are off-set by a higher interest rate. But deflationary pressures in times of recession will push interest rates down to zero, to the point where people are indifferent between holding government bonds and cash. After all, the opportunity cost of holding cash (forgone interest payments) is zero. A monetary expansion is then fully absorbed by the excess demand for money and when the zero lower bound is binding, central banks will be reluctant to raise interest rates (Lukkezen, Jacobs, & Kool, 2016). Fiscal expansion, on the other hand, is likely to be effective in raising output as monetary accommodation is absent and—when the stimulus is large enough—in raising interest rates subsequently.

Research in this field accelerated in light of the financial crisis and subsequent Great Recession. By using a New Keynesian model, Farhi and Werning (2016) show that government spending is "especially potent during a liquidity trap" (p. 2419). A fiscal expansion promotes inflation, and with fixed nominal interest rates, real interest

rates are reduced, in turn increasing current spending. While it is plausible that tax cuts in a liquidity trap will have similar effects, most studies focus on spending effects. Auerbach and Gorodnichenko (2012a) find multipliers ranging from 0 in expansion and 1.7 in recession for the US .

Finally, Blanchard and Leigh (2013) weigh into this debate and show that stronger fiscal consolidation was associated with lower growth than expected, especially early in the crisis. The authors argue that the multipliers must have been higher than the ones used in forecasts. Forecasters learnt and used larger multiplier later in the crisis. Subsequently, smaller estimating errors provide ground for the claim that output responses to fiscal policy were larger in bad economic times.

E. Overview of results found in the literature

I summarise the most notable estimations of the tax multiplier found in the literature in Table 1. The shown multipliers are the largest estimated output response to a tax shock.

Table 1—: Empirical literature on tax multipliers

Study	Country	Method	Maximum multiplier	Notes
Romer and Romer (2010)	US	Narrative	-2.94	
Favero and Giavazzi (2009)	US	Narrative	-0.5	RR dataset
Perotti (2012)	US	Narrative	-1.5	RR dataset
Cloyne (2013)	UK	Narrative	-2.5	Cumulative multiplier
Hayo & Uhl (2013)	DE	Narrative	-2.4	Cumulative multiplier
Saffier II	NL	Structural Macromodel	-1.4 to -1.6	Effect after 8 years
Delfi	NL	DSGE-model	-0.8	Effect after 8 years
NiGEM	NL	DSGE-model	-0.05 to -0.2	Average maximum effect in first year
Mertens and Ravn (2014)	US	Narrative combined with SVAR	-3	RR data subset
Blanchard (2002)	US	SVAR	-0.78 to -1.33	
Blanchard and Perotti (2012)	US	Narrative combined with SVAR	-0.8	
Mountford and Uhlig (2009)	US	SVAR	-5	Deficit financed tax-cuts
Caldara and Kamps (2017)	US	Proxy SVAR	-2	RR data subset, combining narrative with SVAR

The table tells us three things. First, about any view on the size of the multiplier is supported by the literature. The smallest multiplier is the one produced by NiGEM (maximum of -0.2), the largest result is presented by Mountford and Uhlig (2009) who find -5. Second, the narrative method produces mostly large multipliers, with Favero and Giavazzi (2009) being the exception. Third, the Saffier II structural macromodel and the Delfi DSGE model may produce small multipliers on impact, but they rise over the

years, reaching a peak after 8 years. This window is standard in the DSGE literature, but much longer than the estimated mulipliers in for example the narrative method (3 years). Fourth, the signs for all tax multiplier estimations are negative, implying that all studies find that a tax increase leads to a reduction in output or output growth.

Evaluating the different identification strategies, I argue that the narrative approach is most suitable for the identification of exogenous tax shocks. First, using projected changes in tax revenue from the policy documents instead of cyclically adjusted tax revenues helps to determine the size of the tax shock. It also avoids simultaneity as described under *B*. Second, there is no need to impose assumptions on the models as in SVARs. Multiplier results depend critically on assumptions and frictions. Third, one could regard the narrative method as the 'purest'[8] way to identify exogenous shocks. After all, no statistical methods are necessary and other researchers can easily verify exogeneity of tax shocks. Fourth, following closely the identification approach of RR, I can compare results to see if multipliers for the Netherlands are as large as the estimations of RR, Cloyne (2013) and Hayo and Uhl (2014) suggest.

[8]Beetsma (2008) argues that the contribution in an earlier 2007 working paper of Romer and Romer (2010) "likely yields the most reliable results up to now..." (p. 20).

III. The narrative method of identifying exogenous tax changes

I employ the narrative method to identify exogenous tax changes from all tax changes mentioned in policy documents. This section explains the identification strategy in more depth. Following the RR framework I begin by showing a simple equation of how tax changes affect output.

A. Set-up

Suppose taxes affect output in the following way:

$$\Delta Y_t = \alpha_0 + \beta \Delta T_t + \varepsilon_t, \tag{1}$$

where ΔY_t is a change in real GDP growth, α_0 is a constant, ΔT_t is an aggregate measure of tax changes and ε_t is an error term. A change in taxes might lead to a change in GDP growth with some multiplier coefficient β. The time subscript t indicates that taxes affect output in the same period. Tax changes, though, might have a lagged effect on output or last multiple periods. Like RR, I ignore this for now for simplicity.

In estimating the effects of tax on output, we are interested in the value of the β. The coefficient could be positive (in case an increase in tax causes an increase in output), negative (in case an increase in tax causes a decrease in output) or zero (tax changes do not affect output).

It would be possible to estimate the linear model (1) using OLS to find the effect of a change in the aggregate tax measure on output growth. But is it likely that (1) will also lead to the best linear unbiased estimator? The Gauss-Markov theorem is helpful in answering that question (*Introductory Econometrics: A modern approach*, n.d.). To obtain an unbiased estimate, the assumptions of the theorem need to hold. A central assumption is that the expected value of the ε_t, given the explanatory variables for all time periods, is zero. So:

$$\mathrm{E}[\varepsilon_t | T_1, T_2, ..., T_k] = 0\,. \tag{2}$$

The assumption (2) states that regardless of the realisation of the value of tax measure T_t, the expected error term ε_t would still on the average be zero. The distribution of the errors is therefore not dependent on the covariates. Put differently, T_t should not be correlated with the errors. This is given by (3).

$$\text{Cov}[\text{T}_t, \varepsilon_t] = 0\,, \quad \text{for all } t \text{ in the data.} \tag{3}$$

For the estimation to be unbiased and consistent the error term ε_t cannot correlate with the tax measure T_t, for all time periods in the data. If the error term correlates with the tax measure, the model may suffer from endogeneity. The estimated α_0 and β in equation (1) will be biased, and therefore not reflect the true relationship between ΔT_t and Y_t.

Section II argued that because of the nature of the relationship between tax and output, endogeneity in the model is likely. Equation (1) therefore needs modification to reflect the possible ways in which the tax measure and the error term are correlated.

Tax changes are not the only factors affecting output. GDP is affected by any element influencing consumption, investment, government spending, imports or exports. Examples include monetary policy, expectations about the economy, non-tax budget policy and global influences. I call all these factors output determinants. Clearly, equation (1) does not account for the output determinants as separate explanatory variables. The equation captures them in the error term ε_t. This means that the error term contains information of all other elements (than tax) affecting GDP. So:

$$\varepsilon_t = \sum_{i=1}^{K} \varepsilon_t^i\,, \tag{4}$$

where ε_t is the same error as before and the superscript i denotes disparate output determinants. The summation sign indicates that ε_t is a composition of all K disparate elements—other than tax—that influence output. Therefore, ε_t in equation (1) could contain information on output growth. If that is the case, the covariance between the tax measure and the error term is not zero, and assumption (2) is violated.

Now turn to the tax measure in equation (1). The measure of tax change ΔT_t now contains any kind of tax change. It may consist of tax changes correlated with output (endogenous changes) and tax changes uncorrelated with output (exogenous changes). Exogenous changes are not affected by output or by other factors affecting output, while endogenous tax changes are. In other words, exogenous changes are not correlated with the error term, while endogenous tax changes are. To reflect this, I rewrite ΔT_t to get:

$$\Delta T_t = \sum_{i=1}^{K} b_t^i \varepsilon_t^i + \sum_{j=1}^{L} w_t^j \,, \tag{5}$$

where ΔT_t is the same as before, b_t^i is an endogenous tax change, the error term ε_t is the same as before and w_t^j is an exogenous tax change. Equation (5) shows that the total tax measure T_t in (1) is composed of the sum of endogenous tax measures, $\sum_{j=1}^{K} b_t^i \varepsilon_t^i$, and the sum of exogenous tax measures, $\sum_{j=1}^{L} w_t^j$.

Equation (5) is further crucial in showing that some tax changes may be a response to a changing output growth. Policymakers may lower taxes to return output growth to normal in times of recession, or they may raise taxes to finance a spending increase. These tax changes, the b_t^i's, correlate with output through the error term ε_t. On the other hand, some taxes do not correlate with a changing output. Governments may lower taxes to alleviate distortions on the labour supply, or they may change rate structures to redistribute wealth. These tax changes, the w_t^j's, do not correlate with the error term ε_t.

B. Implications

Expression (5) now shows that the full set of legislated tax changes consists of two kinds of tax measures: 1) tax changes caused by or related to other factors affecting output (endogenous tax changes), and 2) tax changes *not* caused by or related to other factors affecting output (exogenous tax changes). I combine (1) and (5), to get:

$$\Delta Y_t = \alpha_0 + \beta \left[\sum_{i=1}^{K} b_t^i \varepsilon_t^i + \sum_{j=1}^{L} w_t^j \right] + \varepsilon_t \,. \tag{6}$$

Equation (6) reflects the fact that the full set of legislated tax measures T_t consists of both exogenous and endogenous tax changes. This way of writing follows Wold's decomposition theorem, which states that any covariance-stationary time series can be written as the sum of a stochastic (correlated with the error term) and deterministic (uncorrelated with the error term) system.

Equation (6) shows what needs to be done in order to rid the linear model (1) of possible endogeneity. For an unbiased and consistent estimation of the effect of tax on output, only tax measures uncorrelated with the error term ε_t, so the w_t^j's, should be used.

Identifying exogenous tax measures is the point of the narrative method. I use the stated motivation in policy documents to isolate exogenous from endogenous tax measures and only use exogenous tax changes for the estimation of the effect of tax on output.

The basic model is therefore written in this way:

$$\Delta Y_t = \alpha_0 + \beta \sum_{j=1}^{L} w_t^j + v_t \,, \tag{7}$$

where ΔT_t, α_0 and w_t^j are the same as before and $v_t = \sum_{i=1}^{K} b_t^i(1 + \beta b_t^i)\varepsilon_t^i$. This last term follows from rewriting (6), putting all shocks to output except the exogenous tax changes in the error term. By leaving tax shocks correlated with the error term out of the estimation, equation (7) allows for a consistent and unbiased estimation of the β of all exogenous tax changes $\sum_{j=1}^{L} w_t^j$ on a change in output growth ΔY_t.

Cloyne (2013) points out that the narrative approach is useful precisely because it effectively identifies exogenous tax measures from the full set of legislated measures. By doing so, it assures that the Gauss-Markov identifying assumptions (2) and (3) hold.

IV. Constructing a dataset of exogenous tax series for the Netherlands

The framework in the previous section shows that identifying exogenous tax changes is necessary to estimate the unbiased effect of tax on output. For this, I need to find the motivation of each tax change. I will also record the size and the timing of each tax change to construct a dataset of exogenous tax series for the Netherlands.

A. Sources

I construct a dataset of exogenous tax changes ranging from 1955 to 2015. For the estimation of the effect of tax on output, I use a slightly smaller window, from 1960 to 2015. The year 1960 is the first for which quarterly GDP data for the Netherlands is available. Digital archives of the most important policy document, the Budget Memorandum, begin at 1955.[9] The last tax changes I consider are the ones that take effect in January 2015. This allows me to use the most recent economic growth figures.

There is also an economic reason for taking the years from 1960 and up. Circumstances in the Netherlands in the first decade after World War II were exceptional. Between 1950 and 1960 the average annual real GDP growth was 4.2%. This was a direct effect of suppressed wages, spending increases through the Marshall plan and a sharp rise in employment (De Kam, 2015). Moreover, government institutions were rebuilt and expanded in a rapid fashion, affecting tax collection and economic growth in several ways. Even if my identification of exogenous taxes in this period would be correct, there are many other factors that may effect economic growth contemporaneously, possibly biasing my estimation results. The 55 consecutive postwar years should in principle yield enough observations for the estimation of the effect of tax on output (RR use 60 consecutive postwar years).

With centralised taxation comes the advantage of comprehensive sources that cover most of the tax changes taking place in a year. A key policy document is the Budget Memorandum (henceforth also BM), a yearly enclosure of the cabinet on projected government revenues and spending. The document is traditionally published on the third Tuesday in September and is part of an event called Prince's Day ('Prinsjesdag'). What had formerly been the celebration of the birthday of Prince Willem V in the Dutch Republic now marks the first day of the new parliamentary year. The monarch delivers the King's Speech ('Troonrede') to the joint session of the Senate and House of Representatives, setting out policy of the government for the coming year. Afterwards, the minister of finance—traditionally in tail-coat and in earlier times also wearing a top hat—presents the BM to

[9]The Budget Memorandum is the official English name for the 'Nota over de toestand van 's Rijks financiën', informally known as the *Miljoenennota*, or "Million's Bill".

the House of Representatives.

The Budget Memorandum presents an overview of the state of the economy, government finance and encompasses plans related to the budget for the next calendar year. A supplement to the Budget Memorandum is the Tax Plan. This document outlines tax proposals in more detail, provides a motivation for each, mentions the desired introduction date and their projected effects to government revenues.

The Budget Memorandum and the Tax Plan will be the main sources for the construction of the exogenous tax series. However, both documents present tax *proposals*: draft tax laws may be amended in Parliament and their introduction may differ from the original submissions. To account for this possibility I consult the Explanatory Note ('Memorie van Toelichting') that accompanies legislated tax changes. This document outlines policy objectives and projected budgetary effects in more detail and mentions possible associated tax changes or spending measures. I also follow up on proposed tax changes in the next Budget Memorandum. When draft laws are amended substantially, I find the changes in the Notes of Change ('Nota van Wijziging'). Most tax proposals, however, are introduced without changes to their size, timing or motivation.[10]

In case these documents are inconclusive on the size, timing or motivation of a tax change I consult parliamentary proceedings in the extensive digital archive of Dutch parliamentary history.[11]

I turn to statistical databases for GDP figures and the construction of several control variables. I retrieve quarterly seasonally adjusted GDP from the OECD database. The figures are in current prices.

I control for several factors plausibly affecting GDP as robustness checks. First, I do so for government expenditure, for which I retrieve data from Statistics Netherlands. Although I do not include tax measures related to changes in spending (see next section), government expenditure is still the most likely omitted variable. Given the feature of the legislative system of the Netherlands in which most policy changes take effect on 1 January, controlling for government spending as a robustness check is a logic choice. Second, I obtain the consumer price index from Statistics Netherlands to control for inflation. The argument behind this is that governments might raise taxes when inflation is high. Controlling for inflation accounts for the possibility that I may have erroneously classified such endogenous counter-inflation measures as exogenous. Third, I obtain the relative price of crude oil from the RR dataset to account for supply shocks. Doing so, I control for global influences that could affect GDP through imports and exports. They might bias

[10]The Cabinet holds a majority in Parliament for nearly all years, which makes it more likely that initial tax proposals gather sufficient support without much change to proposals.

[11]The electronic database on www.statengeneraaldigitaal.nl runs back to an impressive 1814, which marks the first year of modern parliamentary history of the Netherlands.

my results when they coincide with the introduction of a tax change. Fourth, I control for the policy interest rate. Output and interest are determined simultaneously in the goods and money market: the interest rate affects output through investment and output affects the interest rate through money demand. Controlling for the interest rate checks for the possibility that a change in output growth is explained by a changing interest rate, rather than by discretionary tax policy. For the period prior to 1999, I use the rate on ordinary end-day advances of De Nederlandsche Bank, the Dutch central bank. For the years after 1999, I obtain the marginal lending facility from the ECB.

B. Recording all legislated tax changes

I use the policy documents to record all legislated tax changes for the time period considered. For each tax change, I record its size in million guilders (before 2002) and euro (after 2002), the exact motivation of the tax change as found in the policy sources and the source location. For every tax change, I try to give enough citations from the policy documents for other researchers to check the motivation. I furthermore take (parts of) the introductory remarks of the minister of finance in each Budget Memorandum (in Dutch). Together with the motivation of the tax change, this helps to identify exogenous changes from the full dataset in the next stage. I assemble all legislated tax changes in a companion paper titled "Narrative Dataset of Legislated Tax Measures in the Netherlands: 1955-2015".

C. Classification

Section III has shown how identifying exogenous shocks is a possible way of removing endogeneity from the model. But which tax changes are exogenous and which are not? In constructing an identification scheme, I follow the requirements set out by RR, Cloyne (2013) and Hayo and Uhl (2014). I further complement the classification scheme with criteria suited to the Dutch context. I come to the criteria summarised in Table 2 and explain each criterion in more detail below.

Table 2—: Classification scheme

Motivation of typical exogenous changes	Motivation of typical endogenous tax changes
Raising output in the long-run (e.g. labour market participation)	Offsetting developments (i.e. inflation) that cause output to be away from structural growth
Paternalistic motives (e.g. tobacco excise)	Tax changes to offset macroeconomic effects of a change in spending
Redistributive motives	Returning the economy to structural growth path
Fixing an inherited budget deficit	Short-term considerations
Fixing an inherited collective tax burden	Raising employment in the short run
Rulings/directives of external bodies	
Simply absence of any motivation to return growth to normal	
Increasing the efficiency of the tax system	
Increasing the efficacy of the tax system	

Exogenous tax changes

Non-countercyclical tax changes mostly classify as exogenous. First, these are tax changes with the objective to raise output in the medium and long run. I call these tax measures structural tax changes. One could argue that in a way all tax measures have some objective to enhance long term growth (or are at least designed with that intention in mind), but I confine my classification to tax measures which structurally lower tax rates over a long time period or make substantial changes to the tax structure itself. There are ample examples of these policies in the record, such as lowering of labour and income taxes in the 1960s to increase consumption in the long run, withdrawal of tax benefits such as the VUT[12] for early retirees in light of the ageing population in the 1990s and the modernisation of several taxes, such as in 2001 (wage and income tax).

A second exogenous type are all tax changes with a value-driven or philosophical motivation. These tax measures are in essence of a structural nature. The difference with the previous category is that their objective is to benefit society in a different manner than by raising output in the long run. The motivation of these tax measures vary. For example, 'greening' of the tax system reflects policymakers's desire to address sustainability issues.

[12]The 'Vervroegde uittreding en pre-pensioen' (English Early withdrawal and pre-pension) was a long-time fiscal incentive for early retirement. This measure was phased out halfway the 1990s, simply because it became apparent that the pay-as-you-go system became unsustainable in light of the ageing population.

Environmental taxes were introduced in the 1990s and their implementation accelerated at times when political parties with an extensive sustainability agenda held a majority in parliament. Another example is alleviating distortions on labour supply. Favourable taxation towards married couples with working husband and non-working wives or, later on, introducing policies to equalise fiscal treatment between married and unmarried couples reflect changing views on marriage and households.[13] I also consider many redistributive policies as value-driven, but not all. Following the implementation of a tax package, policymakers evaluate whether changes have affected the progressiveness of the tax system. When changing the progressive nature of the tax system was not the objective of the package, policymakers may propose some measures to 'smooth rough edges'.[14] I classify redestributive policies as exogenous when they stand on their own *or* when they accompany another exogenous tax change. I classify redistributive tax measures as endogenous when their objective is to correct any effects on the progressiveness of the tax system following an endogenous tax change.

Third, governments that are confronted with a large budget deficit from earlier cabinets may have the desire to reduce it. Like RR, I classify measures to reduce such inherited deficits as exogenous. A prime example in the record are the extensive efforts by Ministers of Finance Ruding and Kok to decrease the large deficit that accumulated over the postwar decades.[15]

Fourth, changes to the tax system to increase efficacy (the extent to which the measure achieves its policy objective) or efficiency (the extent to which potential spillover effects or related costs are acceptable and proportional to the policy objective) usually classify as exogenous. Especially in recent decades it is common practice to evaluate the implications of tax policy. A (parliamentary) report could lead to amendments to the tax proposal or even to its withdrawal.[16]

A fifth type of tax measures that I classify as exogenous are 'external' changes. These are measures imposed on policymakers by external bodies, such as a court ruling a tax unconstitutional. I encounter them throughout, with a more prominent role for the national Court of Justice in the first half of the considered time period, and for the Court of Justice of the European Union in the second. The implementation of directives of the

[13] A tax change reducing these differences in 1960 would be a good example.

[14] Progressiveness of the tax system or the notion that 'the broadest shoulders carry most weight' is mentioned countless times in the narrative records, and it is discussed in what way new policies might affect purchasing power for income groups. To correct for possibly regressive measures, policymakers could decide to make additional tax changes.

[15] From the narrative record in 1986: "The government has been fighting a sustained deficit for several years. 'Core of the problem is that expenditure is too high, while the pillar on which spending rests [gas mining revenues] fall away'. The new cabinet inherits the budget deficit and imposes changes out of fiscal prudence. The high deficit and the fact that the cabinet deemed average tax rates to be to high, finds it cause in a process of multiple years..."

[16] Also simplification of the system like the revision of the wage and income tax in 2001 follows this line of thinking.

European Union also fall in this category.

Endogenous changes

I classify tax changes affected by output in the broadest sense as endogenous. A main category is tax policy with a countercyclical motivation. The most obvious piece of evidence that policymakers try to influence output growth is that taxes are changed to make output growth return to normal in the short run. I take 'normal' to be trend growth. The provided context is crucial. Notions in the policy documents that the economy is in a prolonged period of low economic growth or the very Dutch expression (and its literal translation) "we are facing economic bad weather" (e.g. BM 2010) are strong signs that the economy is in a downturn. It is important to note that tax changes introduced at a time when growth is away from normal is on itself not enough to classify a tax measure as endogenous; the motivation of the tax measure is crucial.

Related to this are tax measures to offset the effects of developments causing output to be away from normal. Here the tax change targets a specific element that influences output. Examples are measures to reduce surging inflation in 1993,[17] which was of such levels that it led to higher wages, reduced competitiveness and stalling economic growth (De Kam, 2015).

I classify tax changes as endogenous in case they are implemented to offset a spending shock. Policymakers may want to increase spending, for example, and finance the spending increase by raising taxes. Obviously, both a change in spending and tax policy affect output. The estimation would suffer from omitted variable bias when including the tax measure but not accounting for the change in spending.

Tax changes related to the unemployment rate are classified as endogenous, but under strict conditions. Often tax changes to increase employment follow times of low economic growth. Low economic growth is in that case a driver of rising unemployment, and any tax measure aimed at reducing unemployment is in fact responding to a changing output. That would be a strong case to classify the measure as endogenous. However, the relation between output and employment is not always clear. Policy documents account for motivated tax changes in times of high unemployment but normal growth. When I find no clear relation to output or another factor affecting output I classify the change as exogenous. Additionally, it is important not to confuse employment policies with measures taken to increase labour force participation.[18] Those measures are a longer

[17]Taken from the discussion in the narrative record for 1993: "As stated in the introduction of the Budget Memorandum, compared to other countries in Europe, the economy is performing relatively well. Though it seems policymakers expect some headwind. Nevertheless, the net tax changes are an increase in government revenues. This tax measure is mentioned in relation to the high inflation and therefore considered endogenous."

[18]Fiscal measures to increase labour force participation were especially popular in the 1970s to increase participation among women, and in the early 1990s to raise the participation rate among citizens with a migrant

term consideration and therefore exogenous.

I classify tax measures as endogenous when I cannot effectively disentangle policy events from a tax change. Examples are tax measures that accompany the completion of the European single market. Here the single market completion is likely to affect output in numerous ways (general optimism, ease of trans-border movement, and so forth). When the tax change is part of a larger policy event that may also affect output, I classify the measure as endogenous.

Dropping tax changes

There are several tax measures that I do not consider in the dataset. First, like RR, I leave out 'tax escalators', non-legislated measures such as automatic increases of excise duties on cigarettes. These changes often have small budgetary effects and even their news value is minimal. Second, I do not document tax changes with a revenue effect of less than 0.1% of GDP. These small changes are often poorly documented, are likely to have a minor effect on output and their improvement of the estimation does not weigh up against the considerable identification effort. I do, however, consider these tax changes with a minor budgetary effect when they come in the form of packages, as a tax package usually has a single motivation, which helps to classify them precisely.

Temporary tax changes

I document temporary measures and classify them in the usual way. Temporary are usually minor (change in tax revenue relative to GDP) and most temporary taxes are short term, often effective for a year. The literature mentions that macroeconomic effects of temporary taxes may be different from structural tax changes, as households may anticipate a return to a previous tax level, especially when the date that the measure is withdrawn is announced before or along with its implementation. What I find, is that temporary tax changes are mostly either taken to offset an increase in spending or to provide consumption incentives consumption in an effort to increase output.[19] In practice, most temporary shocks will therefore be classified as endogenous and excluded from the estimation.

Taking motivations at face value

Like RR, I take stated motivations of policymakers at face value. By default, I will not argue against a clearly stated motivation of policy makers. The included discussion in the narrative record is confined to a precise classification by the stated motivation.

background

[19] What comes to mind are the temporary reductions in certain transfer taxes to stimulate the housing market in the Netherlands, in 2014.

I do not question the motivation itself for several reasons. First, I am limited to the available information in the policy sources. Retrieving motivations from these documents is in a way second best: it would be challenging and in many instances impossible to ask policymakers directly why they changed certain taxes. Especially for the early years in my dataset, it proves to be demanding to find the exact motivation of tax changes. It is even more difficult to plausibly argue against a stated motivation and provide enough context to 'disprove' a motivation. The motivations in policy sources are therefore arguably the most reliable signals of actual policy intentions.

A second but weaker argument to take policymakers' intentions as stated is to account for news and anticipation effects. Even in case the stated motivation is different from the documented one, the behaviour of decision makers in the economy is likely to be based on the stated motivation of tax changes. Accounting for the tax change in accordance with the motivation of policymakers may be the most reliable reflection of their output effects.

D. Exhibit of the dataset

I apply the criteria to identify the exogenous tax policy measures from the full narrative record of legislated tax changes. It is not practical to include the narrative dataset in this thesis, but I provide an exhibit of the narrative record to familiarise the reader with its content. In Appendix A, I describe the dataset on a decade-by-decade basis and discuss the classification of the most notable tax changes in each.

Exhibit 1 is taken from the companion paper and shows tax changes taking effect in 1994. The extract is exemplary for the way that I record discretionary tax policy between 1960 and 2015. For each year I take the introduction (or parts thereof) from the Budget Memorandum, which usually consists of opening remarks by the minister of finance. Mostly, the introduction summarises the current state of the economy and the most notable tax and budget measures. It sketches the context and climate in which the government introduces new tax measures.

I record discretionary policy in a table, listing the original name of the tax policy change (in Dutch), its translation in English, the classification together with the type of measure (long-run considerations, spending driven or countercyclical), the size of the tax change in millions (guilders prior to 2002 and euro after), a justification of the classification, the motivation of policymakers as a citation and the source location. Whenever tax changes are introduced as a package and share the same motivation, I find evidence for the package and list them as one measure.

Exhibit 1: Sample year from the narrative analysis

1994

Het gaat Europa dit jaar economisch niet voor de wind. De groei is nog lager dan eerder werd verwacht en menig land - waaronder het buurland Duitsland - verkeert in een recessie. De groeiverwachtingen voor volgend jaar zijn matig en de werkloosheid dreigt fors op te lopen. Onder deze omstandigheden is er alle aanleiding om het in 1989 ingezette regeringsbeleid ook volgend jaar met kracht voort te zetten met als centrale elementen: versterking van de economische structuur, vergroting van de deelname aan betaalde arbeid en voortgaande gezondmaking van de overheidsfinanciën.

In de afgelopen jaren is de trend van dalende overheidsinvesteringen omgebogen. In de samenstelling van de overheidsuitgaven zijn verschuivingen aangebracht: relatief minder subsidies en meer structuurverbetering. Door nu en in de komende jaren meer te investeren in infrastructuur, milieu en kennis -de lijnen daarvoor worden in deze Miljoenennota uitgezet - leggen wij een steviger basis voor de toekomst van onze kinderen en voor houdbare gemeenschapsvoorzieningen. De werkloosheidsproblemen zijn nijpend. Het aantal banen dreigt in 1993 voor het eerst sinds vele jaren weer te dalen. Het is zaak de schade tot het uiterste te beperken. Alleen het overwinnen van de conjuncturele inzinking is daarvoor niet voldoende. Er zijn ook knelpunten van meer structurele aard.

De in deze Miljoenennota uiteengezette beleidsvoornemens zijn van tweeërlei aard: zij laten zien wat de regering zélf wil doen om de werkloosheidsgroei te beteugelen en zij zijn tevens een indringende uitnodiging aan werkgevers en werknemers om samen inhoud te geven aan een breed gedragen en consequente keuze voor arbeid boven inkomen. De gezondmaking van de overheidsfinanciën ondervindt onvermijdelijk enige vertraging door de lagere economische groei en de daaruit voortvloeiende tegenvallende belastingopbrengsten.

Desondanks kan het financieringstekort van 1992 op 1993 toch nog dalen en kan in 1994 een stijging worden voorkomen. Nederland wijkt daarmee in gunstige zin af van vrijwel alle andere landen in Europa. Daarom is gekozen voor een sober uitgavenbeleid en een pakket aanvullende ombuigingen dat in de komende jaren een sterk oplopend bedrag aan besparingen zal opleveren. In samenhang met de voorziene daling van de collectieve lasten levert het budgettair beleid aldus een positieve en ondersteunende bijdrage aan het zo noodzakelijke herstel van economie en werkgelegenheid in 1994. De Europese dimensie van het te voeren beleid verdient bijzondere aandacht. Het proces van economische en monetaire samenwerking maakt thans een moeizame fase door. Een gezamenlijke aanpak van de werkloosheidsproblematiek is van groot belang. De concurrentiepositie van Europa in de wereld moet worden versterkt. Met het Verdrag van Maastricht als gezamenlijk aanvaarde basis staan de lidstaten thans voor de opgave de cohesie in Europa te versterken en een krachtige bijdrage te leveren aan duurzaamheid, ontwikkeling en rechtvaardigheid op wereldschaal.

- Minister van Financiën,
  Wim Kok

Measure (Dutch)	Measure (English)	Class	Size (mln)	Reason	Evidence	Source
Verleggingsregeling OB Loonconfectie	Change in certain revenue taxes	Exo (LR)	13	Minor change, no relation the cycle.	No further evidence	Miljoennota 1994, Bijlage 4, p. 113
Verbruiksbelasting op milieugrondslag, tariefsverhoging voor grondwater/afval	Consumption tax on environmental basis	Exo (LR)	130	Environmental concerns.	*Dat geldt voor de CO_2-uitstoot in relatie tot het energiegebruik, de uitstoot van NO, uit verkeer en de lokaal hoge uitstoot van ammoniak uit intensieve veehouderijen. Om de CO_2-uitstoot conform de beleidsdoelstellingen te verminderen, is de inzet van het prijsinstrumentarium gewenst. Ook bij andere milieuthemars kan inzet van het prijsinstrumentarium bijdragen aan het bereiken van de doelstellingen. Met de verbruiksbelasting op milieugrondslag - die wordt uitgebreid met de grondslagen afvalstoffen en grondwater - is reeds een stap in die richting gezet.*	Miljoenennota 1994, p. 28
Beperking faciliëring kinderopvang	Reducing fiscal facilities for child care	Endo (SD)	35	To offset increase in spending for child care, therefore affected by another change in output. Endogenous.	*Vanaf 1994 wordt 85 miljoen extra uitgetrokken voor de uitgaven op het terrein van de kinderopvang. Aanvullend is in de jaren 1994 en 1995 een bedrag van 15 miljoen beschikbaar. Ter dekking is de fiscale aftrekbaarheid van kinderopvang beperkt, hetgeen vanuit de fiscaliteit 35 miljoen oplevert.*	Miljoenennota, p. 49
Fiscale faciliteit voor de zeescheepvaart en aanpassing	Fiscal facilities for shipping	Exo (LR)	-18	Specific policies for certain measures in shipping industry.	*De regelingen op basis waarvan steunverlening aan de zeescheepvaart wordt verleend, lopen eind 1995 af. Het betreft afnemende uitgaven voor de reeds afgelopen IPZ, de Interimregeling Stimulering Zeescheepvaart voor 1993 en de deels fiscaal begrootte Wet Stimulering Zeescheepvaart voor 1994 en 1995.*	Miljoenennota 1994, p. 142
Verscheidene veranderingen in accijnzen en MRB	Several changes in excise duties and tax on motorized	Endo (SD)	785	Excise taxes increased to offset spending on employment policies.	*Op het terrein van verkeer en vervoer is besloten tot een verhoging van de accijns op motorbrandstoffen. De budgettaire baten die daardoor beschikbaar komen, worden aangewend ter financiering van het werkgelegenheids- en structuurbeleid. Met de verhoging*	Miljoenennota 1994, p. 51

	vehicles				*van de accijns op benzine en diesel, alsmede de verhoging van de LPG- en dieseltoeslag in de motorrijtuigenbelasting, is een bedrag gemoeid van 905 miljoen in 1994 en 980 miljoen structureel. Een deel van de opbrengst zal in het Infrastructuurfonds vloeien respectievelijk worden aangewend voor de financiering van de bouw van extra cellen (zie paragraaf 3.3.4). Bovendien zal uit de opbrengst een gedeeltelijke compensatie ten behoeve van het wegvervoer voor de verhoging van de dieselaccijns worden gefinancierd.*	
Verhoging arbeidskostenforfait	Increase wage cost forfait	Exo (LR)	-520	Long-term considerations for economic growth. Increase in general tax exemption for the working to increase price of leisure and thereby contributing to labour participation.	*Voorts wordt ter verbetering van het functioneren van de arbeidsmarkt in 1994 het arbeidskostenforfait verhoogd met (maximaal) 500 gulden. Werken wordt zo financieel aantrekkelijker, en er ontstaat een betere aansluiting van vraag en aanbod in het lagere segment van de arbeidsmarkt. Ter verdere ondersteuning wordt enige nadere stroomlijning aangebracht in de regulering van de arbeidsmarkt (in de sfeer van de preventieve ontslagtoets, de Arbeidstijdenwet, en het Arbeidsomstandighedenbesluit).*	Miljoenennota 1994, p. 12
Verlaging tarief eerste schijf 0,65%-punt	Decrease rate of first rate in income tax	Exo (LR)	-1495	Political objective to smooth purchasing power across and for households.	*Met het oog op de inkomensontwikkeling in 1994 en de bevordering van de arbeidsparticipatie is besloten tot een samenstel van maatregelen in de loon- en inkomstenbelasting. Daarbij zal worden afgezien van de toepassing van de inflatiecorrectie op de schijven. Voorts zal het tarief van de eerste schijf worden verlaagd en zal een substantiële verhoging van het arbeidskostenforfait worden doorgevoerd.*	Miljoenennota 1994, p. 51
Vervallen inflatiecorrectie op belastingschijven	Cancelling of the inflation correction	Dropped	1205	Though legislated, this measure just means that the proposed indexation of the tax rates	*Onderdeel van dit beleidspakket is het laten passeren van de inflatiecorrectie op het einde van de eerste en tweede tariefschijf; de opbrengst hiervan wordt besteed onder meer aan verlaging van het tarief eerste schijf en genoemde verhoging van het arbeidskostenforfait.*	Miljoenennota 1994, p. 51
Aanpassing vermogensbelasting, o.a. verhoging ondernemingsvrijstelling en R&D-faciliteit loonbelasting	Change to property tax (increase enterprise exemption and research and	Exo (LR)	-305	Long term measures aimed to increase innovation in business.	*Voor de werkgelegenheid en voor een gezonde economische ontwikkeling is, aansluitend bij een goed functionerende arbeidsmarkt, een gezond en innovatief bedrijfsleven essentieel. In dit kader is een tweetal fiscale maatregelen getroffen. Ter bevordering van onderzoeks- en ontwikkelingsactiviteiten in het bedrijfsleven wordt in de eerste plaats een fiscale faciliteit in het leven geroepen*	Miljoenennota 1994, p. 12

	development faciltiies)				*voor loonkosten van R&D-uitgaven van (met name kleine en middelgrote) ondernemingen. In de tweede plaats wordt met het oog op de positie van met name het midden- en kleinbedrijf de druk van de vermogensbelasting verlaagd. Dit gebeurt door de ondernemingsvrijstelling in de vermogensbelasting te verruimen en door het percentage van de samenlooopregeling inkomstenbelasting/vermogensbelasting te verlagen.*	
Verhoging AVAR-percentage 1994 tot 104%	Increase AVAR-percentage	Exo (LR)	375	This is a measure related to the profit tax. Absence of any evidence	Evidence absent	Miljoenennota 1994, Bijlage 4, p. 115
Tariefafstapje Vpb van f 250 000 naar f 100 000 (1/7/1994)	Lower tarrif rate on the profit tax (1/7/1994)	Exo (LR)	-148	Fiscal infrastructure and globalizing economies cited, technical changes to rate structures. Therefore exogenous.	*Zo heeft in het licht van de voortgaande globalisering van economieën het vorige kabinet 390 miljoen structureel uitgetrokken voor maatregelen ter versteviging van de fiscale infrastructuur. Van de in dit verband voorgestelde maatregelen hebben de verhoging van de zelfstandigenaftrek en het inkorten van het tariefafstapje in de vennootschapsbelasting reeds in juli van dit jaar hun beslag gekregen.*	Miljoenennota 1994, Bijlage 4, p. 91
Verhoging zelfstandigenaftrek (1/7/1994)	Increase self-employed credit (1/7/1994)	Exo (LR)	-63	Same as above	See explanation above	Miljoenennota 1994, Bijlage 4, p. 91

V. New measure of tax shocks

A. Key figures of the newly constructed tax variables

Table 3 shows the yield of the narrative analysis for the period 1955 to 2015. I find a total of 322 legislated changes, where changes may mean a single tax change, or a tax package consisting of multiple tax changes. I only count a package as one change when 1) the package is mentioned as such consistently in policy documents, and 2) all tax changes in one package share one motivation.

Table 3—: Tax changes in the narrative dataset

Period	Total	Exogenous	Endogenous	Revenue impact (% of GDP)
1955-1965	8	7	1	1.2
1966-1975	41	31	10	7.6
1976-1985	34	28	6	6.7
1986-1995	65	59	6	4.2
1996-2005	113	109	3	3.3
2006-2015	61	46	15	2.6
All periods	322	289	41	25.6

Packages counted as one measure, provided they are mentioned consistently as such in the policy documents and share one policy motivation. The revenue impact as a % of GDP is the absolute value of a tax change relative to nominal GDP.

Of all legislated measures, I identify 289 as exogenous, leaving 41 to be endogenous. Clearly, the vast majority of tax policy is taken with one of the underlying exogenous motivations presented in Table 2 (about 86% of all measures). Dutch tax changes are by far mostly exogenous, with many policy objectives aimed at the supply side of the economy. The revenue impact on GDP in the decades 1966-1975 and 1976 were particularly high: 7.6% and 6.7% respectively. In the three decades that followed, tax changes as a percentage of GDP were considerably smaller.

Two decades stand out. Between 1966 and 1975, 10 changes classified as endogenous. This is a relatively high number compared to the three decades between 1976 and 2015, in which I find 15 endogenous changes in total. In the three decades that follow, I only find 18 countercyclical changes. De Kam (2015) provides an explanation of this observation. With both high inflation and unemployment in the 1970s, economists started to question the effectiveness of coutercyclical policy in Keynesian spirit. Not the demand side but supply side, policymakers believed, was at the roots of stagflation and should be addressed through increasing labour force participation by decreasing taxes on labour and

stimulating saving. On this, De Kam writes: "Since the 1980s recommendations of supply side economists are widely followed up by policymakers... The top rate of the income tax is reduced as well as the corporate tax rate." I find direct evidence of this claim in the narrative record, with a clear rise in exogenous 'structural' measures to improve economic infastructure, and a decline in countercyclical in the decades that followed. I present Table 3 graphically in Figure 1, where positive values as a percentage of GDP imply tax hikes, negative values are tax cuts.

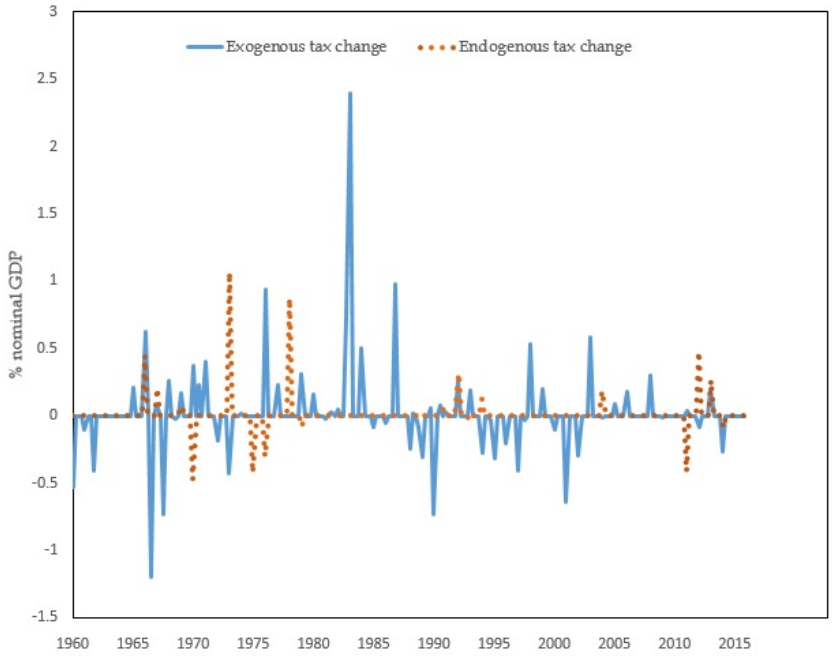


Figure 1. : Exogenous and endogenous tax changes

The other striking decade with respect to the number of endogenous tax measures is 2006 to 2015. Of the 61 legislated tax changes those years, about 15 (25%) are endogenous. Relative to other decades, this means a large share of endogenous measures. These were mainly taken in the second half of the decade and are a mix of both countercyclical measures to stimulate the economy (2011 and 2012) and deficit reducing policy (between 2012 to 2015). Other than in the 1980s and 1990s where policymakers were confronted with large deficits as a result of years of expenditure exceeding income, austerity measures aimed to improve public finance in the short-run as a response of worsening economic conditions and a desire to meet EMU obligations and contain the deficit below 3%.

Table 4 specifies tax measures by most important tax types: income and wage tax rate changes and tax free sums in the income and wage tax (IWTR), income and wage

deductible change[20] (IWTD), corporate tax rate change (CTR), corporate deductibles (CTD), consumption tax rate[21] (COTR) and all other measures.

Table 4—: Tax measures: specified by tax type

Period	Total	IWTR	IWTD	CTR	CTD	COTR	Other
1955-1965	8	4	2	0	0	2	0
1966-1975	41	11	2	4	1	20	2
1976-1985	34	7	5	1	4	12	5
1986-1995	65	10	6	7	2	18	22
1996-2005	113	22	15	3	12	14	47
2006-2015	61	8	13	1	9	8	23
All periods	322	62	43	16	28	74	99

Notes: IWTR is income/wage tax change (rate), IWTD is income/wage tax change (deductible), CTR is corporate tax rate, CTD is corporate deductible, COTR is consumption tax rate, other is all other tax measures

Several properties are worth mentioning. First, changes to the wage and income tax rates are an important and constant part of tax policy in all periods. Between 1955 and 1965, half of all tax measures were of this nature. This share declined in later periods, but was always a substantial part of Dutch tax policy. Second, deductibles in the income tax started to play a more important role in later decades. An explanation could be the evolution of the tax system as an instrument to increase labour supply, promote sustainable consumption and RD investment. Another is that the digital era facilitated a more sophisticated and 'fine-tuned' tax policy, effectively targeting specific (income) groups. Third, policymakers extended and amended consumption taxes the most: they make up 74 of 322 tax changes. In the early periods, this is mainly explained by the introduction and extension of excise duties, in the later periods this entailed moving product categories to different tax rates ('van verlaagd naar standaard') and amending the VAT rate.[22] Fourth, the 'other' category, where I placed all tax measures that did not fit into one of the main traditional categories, has grown through the years. This category consists of all other taxes (such as property and transfer tax) and illustrates well that policymakers explored other types of taxation than the more traditional tax categories.

[20]With this, I mean to say all 'aftrekposten'.

[21]This includes the revenue tax 'cascade system' (till 1968), VAT (1968 onward), excise duties, tariffs and several other consumption taxes.

[22]Since the introduction of the two rate VAT (4% and 12%), the VAT rate has been changed 9 times. As of 1 October 2012, its rates are 6% (essentials) and 21% (non-essentials).

B. *Processing the data to usable input*

Now that I have a dataset of all endogenous and exogenous tax changes, I construct a quarterly time series from 1960 to 2015. The resulting series is a change in projected revenues, normalised by GDP and expressed as a percentage of GDP. The series could therefore be seen as changes in the average tax rate.

I process the narrative dataset in the following way. Almost all tax changes take effect on 1 January. When they do not, measures are often introduced on the first day of a quarter, so 1 April, 1 July and 1 October. Following RR, I put tax changes that take effect in the second half of a quarter in the next: a change that takes effect on June 10 is put in quarter 3. Liability mutations for tax changes taking effect on dates other than January 1 are sometimes expressed for the remaining months.[23] When this happens, I scale the estimated amount to a full year, so estimations of all tax are for a change in liabilities in 12 months.

Assigning tax changes in the way of RR follows the logic that real effects of tax changes are observable in roughly the same quarter as the introduction of a tax change. This may be true for immediate behavioural responses to the news of a tax change. However, real effects to a tax change may plausibly only be observable months after the implementation of new discretionary tax policy. Consider a change in the height of a certain income tax deductible, such as expenditure on education. The revised height of deductibles takes effect on 1 January of a certain year and will be effective immediately. Households file their tax declarations, though, mostly in April of the *following* year. This means that the change in deductibles will only affect disposable household income months after the change took effect. Following this logic, it may prove useful to account for these 'real' implementation lags. I therefore experiment with lags of 2, 4 and 6 quarters, so that a tax change that takes effect on 1 January is matched with output growth 2 quarters later, and so on.

I account for the introduction of the euro in the following way. Policy documents on tax measures taking effect in 2002 or later (regardless of when they were written) mention the effect of tax measures in euros. I only use euro figures. I therefore convert GDP in the years up to and including 2001 to euro, by using the set conversion rate.[24] The GDP figures from the OECD database are all in euro.

In line with RR, I scale tax liabilities to be expressed as a percentage of nominal GDP. This way, the resulting percentage can be interpreted as a change in the average tax rate. I further express output as its natural logarithm in line with my basic estimation

[23]For example, the sources may say that a tax change which takes effect on June 1 is estimated to change government liabilities with a certain amount, over the remainder of the years, so 7 months.

[24]The conversion rate is 2.20371 guilder to one euro.

model.

C. Descriptive statistics of the tax variable

Following the assignment rules above, I put all observations in their respective quarters. The resulting dataset consists of tax changes for 64 quarters spanning 55 years. The exogenous tax variable has a mean of nearly zero (0.0016% of GDP). This implies that exogenous tax policy (in liability mutation) consisted of about as much tax hikes as cuts. The standard deviation of the tax series is 0.0026% of GDP. This makes the variation substantially smaller than that of the RR tax variable (0.24% of GDP). The variable of endogenous tax changes has a slightly larger mean (0.0104% of GDP) and a standard deviation of 0.0012.

I cannot test whether the exogenous series is contemporaneously uncorrelated with other macroeconomic elements, but it also instructive to see whether output contains information on tax movements. If the identification of exogenous tax changes has been successful, tax shocks are 'unforcastable' on the basis of past output movements.

Table 5 presents the results from the Granger Causality test using output and the exogenous tax series. I perform the test for 4, 8 and 12 quarter lags. P-values are high, 0.68 and over, and over 0.9 for 12 lags. The null hypothesis that output does not Granger-cause tax clearly cannot be rejected for the various lag lengths. The results are assuring: the Granger Causality test provides no ground for the assumption that output movements explain the exogenous tax series, and hence does not point in the direction of endogeneity in the model.

Table 5—: Granger Causality - Exogenous Tax Series

Lags	Test statistic	P-value
4	2.31	0.68
8	4.46	0.81
12	5.81	0.93

I also perform this test for my endogenous series for the same lengths. Table 6 now shows much lower P-values: 0.31 for 4 lags to 0.02 for 12 lags. There is no evidence that output Granger-causes the endogenous tax series for lag length 4 and 8. However, the lower P-value and the clear rejection of the null hypothesis at lag length 12 points to a substantially stronger causal relationship between output and tax for my endogenous series than is the case for my exogenous tax series.

Table 6—: Granger Causality - Endogenous Tax Series

Lags	Test statistic	P-value
4	4.79	0.31
8	10.18	0.25
12	22.94	0.02

The much lower P-values for the endogenous tax series and rejection of the null for 12 lag lengths supports the view that in my full series of tax changes, there is indeed concern for endogeneity. With a lag length of 12, output contains information on the movement of the tax series. This problems seems eliminated in my exogenous series, where p-values are far from significant. This gives reason to believe that identification of exogenous tax changes from the full set of discretionary policy has been successful.

VI. The effect of tax changes on output

A. Baseline specifications

Having identified all exogenous shocks from the full dataset, I am able to use a simple regression. I regress real output growth on a constant, the contemporaneous value and numerous lags of the measure of exogenous tax changes. Thus, I estimate:

$$\Delta Y_t = \alpha_0 + \sum_{i=0}^{M} b_i \Delta T_{t-i} + e_t \,. \tag{8}$$

where ΔY_t is the logarithm of real output and ΔT_{t-i} is the exogenous measure of tax changes. I follow RR in choosing 12 lags for the exogenous tax variable.

In Section VII, I experiment by adding control variables. There is one control variable, however, sufficiently important to consider in the baseline specification. I add lagged output to the equation in (9) for multiple reasons. First, output is likely to be partially determined by the previous period, so its lags will likely explain parts of its movement. Second, other factors influencing output are likely to be serially correlated and lagged output controls for these elements. Third, adding lags of output controls for a distinct feature of the democratic process. There is a chance that even exogenously identified discretionary tax policy, correlates with economic performance. RR give the example of politicians favouring tax cuts, who might be more likely to be elected when economic growth is low. In this case, tax cuts that are supposedly of the exogenous type are more likely to happen when output is below trend growth. Adding lagged output addresses this possibility.

For the reasons given above, there is much reason to believe that equation (9) is able to produce estimates of the multiplier closer to the actual relationship between tax and output than equation (8). In keeping with RR, I estimate both equations (8) and (9). In the robustness checks of next section, I perform the tests with lagged output throughout. I add 11 lags of output; doing so allows me to keep the sample period. So, I estimate:

$$\Delta Y_t = \alpha_0 + \sum_{i=0}^{M} b_i \Delta T_{t-i} + \sum_{j=1}^{N} c_j \Delta Y_{t-j} + e_t \,, \tag{9}$$

where the coefficient, the measure of tax and output and the error term is the same as

before.

B. Baseline results

Table 7 gives the regression results of equation (8). The initial response—the impact multiplier—is sharp and negative with an estimated coefficient of 0.92%. A 1% increase (cut) in the implied average tax rate leads in the same quarter to a decline (increase) in GDP growth of about the same size. The coefficient is highly significant with t = -2.31, but in this limited setup without lagged output also the only significant coefficient. 7 out of 13 estimated coefficients are negative. I show the results with a confidence interval of 68%, corresponding to one standard error.[25] The R-squared is low with a value of 0.06, yet by only a few basis points compared to the statistic presented in Romer and Romer (2009). The Durbin-Watson statistic of 2.28 is not substantially away from 2, giving no reason to believe that the error terms are autocorrelated.

Table 7—: Regression results of the effect of tax on output *(no lagged output)*

	Measure of Exogenous Tax Changes (no controls)	Standard error	t	P	68% confidence interval	
Lag 0	-0.92	0.40	-2.31	0.02	-1.32	-0.53
Lag 1	0.50	0.40	1.25	0.21	0.10	0.91
Lag 2	0.44	0.40	1.1	0.27	0.04	0.85
Lag 3	-0.52	0.40	-1.29	0.20	-0.92	-0.12
Lag 4	0.35	0.42	0.84	0.40	-0.06	0.77
Lag 5	-0.19	0.41	-0.47	0.64	-0.60	0.22
Lag 6	-0.56	0.41	-1.36	0.18	-0.97	-0.15
Lag 7	0.27	0.41	0.65	0.52	-0.14	0.68
Lag 8	-0.09	0.41	-0.23	0.82	-0.51	0.32
Lag 9	-0.03	0.40	-0.06	0.95	-0.43	0.38
Lag 10	0.05	0.40	0.13	0.90	-0.35	0.45
Lag 11	0.12	0.40	0.29	0.77	-0.28	0.52
Lag 12	-0.37	0.39	-0.92	0.36	-0.76	0.03
Constant	0.01	0.00	6.73	0.00	0.01	0.01
R-squared	0.06					
D.W.	2.28					

[25]One-standard error bands are also employed by RR and in Cloyne (2013).

Figure 3 presents the implied path of real GDP (in logs) following a tax increase of 1% of GDP. Given the simple setup of the regression, the blue line in Figure 3 represents the sum of the coefficients (including the contemporaneous coefficient). I show the effect of tax on output with one-standard error bands represented by the grey lines. Immediately after the first quarter the coefficients for quarters 2 and 3 are positive, leading to a cumulative response of around zero between quarters 3 and 6. After quarter 6, the output response becomes negative again and stabilises around -0.6% of GDP. Tax hikes seem to have a substantial negative contemporaneous effect on real GDP growth, and a sustained negative effect after quarter 7.

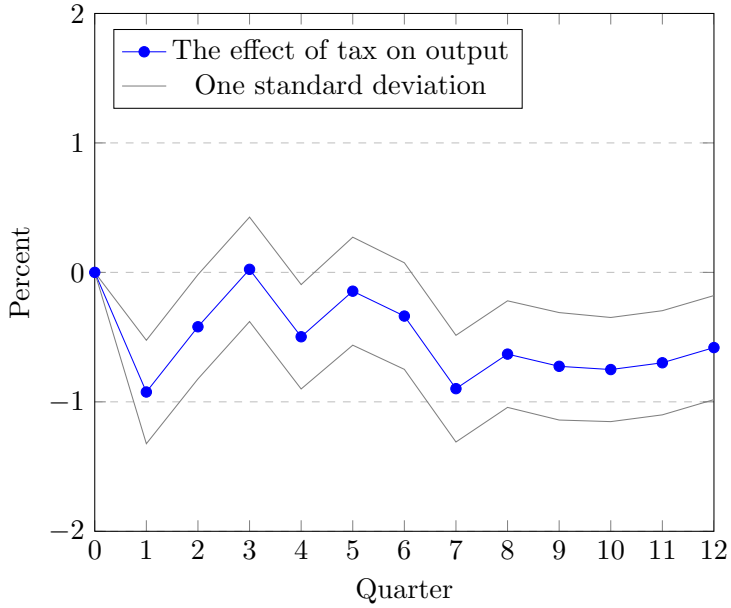


Figure 2. : Eq. (8) - The effect of tax on output (*no lagged output*)

I further estimate equation (9) with 11 lags of output and present the results in Table 8. Of the 13 estimated coefficients, 6 have a negative sign. The R-squared remains low and the D.W. statistic is now below but still close to 2.

The initial response of GDP to a tax change of 1% of GDP is about the same (-0.85%) as in the estimation without lagged output and is again significant at the 5% level with t = -2.15. Lag 2 is now also significant *and* positive with an estimated coefficient of 0.78 ($t = 1.94$). In this estimation with lagged output, I find both a strong and significant decline of output following a tax shock of 1% of GDP in the same quarter, and a strong and significant increase of output two quarters later.

Table 8—: Regression results of the effect of tax on output *(lagged output)*

	Measure of Exogenous Tax Changes (no controls)	Standard error	t	P	68% confidence interval	
Lag 0	-0.85	0.39	-2.15	0.03	-1.24	-0.45
Lag 1	0.31	0.40	0.78	0.44	-0.09	0.71
Lag 2	0.78	0.40	1.94	0.05	0.38	1.18
Lag 3	-0.33	0.40	-0.81	0.42	-0.73	0.08
Lag 4	0.36	0.40	0.88	0.38	-0.05	0.78
Lag 5	-0.11	0.41	-0.26	0.79	-0.52	0.30
Lag 6	-0.45	0.41	-1.11	0.27	-0.86	-0.04
Lag 7	0.12	0.41	0.29	0.77	-0.29	0.53
Lag 8	-0.22	0.41	-0.54	0.59	-0.63	0.19
Lag 9	0.12	0.40	0.31	0.76	-0.27	0.52
Lag 10	0.14	0.40	0.36	0.72	-0.25	0.54
Lag 11	0.08	0.39	0.20	0.85	-0.32	0.47
Lag 12	-0.14	0.38	-0.38	0.71	-0.53	0.24
Constant	0.00	0.00	2.37	0.02	0.00	0.01
R-squared	0.18					
D.W.	1.84					

Figure 4 graphs the sum of the coefficients. The solid line represents the results of the estimation with lagged output, the dashed line the previous estimation (8) without lagged output. I again show the implied path of real GDP following a tax hike of 1% of GDP with one-standard error bands.

The output response with lagged output follows roughly the same pattern as the response without output lags, but is 'elevated' after quarter 3 and by about 0.5% after quarter 5 compared to the results without lagged output. The cumulative output response for quarters 3, 4 and 6 is even slightly above zero. For the remaining quarters, the output response oscillates between 0 and -0.4%. In this estimation with lagged output, I find evidence of a strong and negative response of GDP to tax in the first 3 quarters, but do not find support that the multiplier is statistically different from 0 in later quarters.

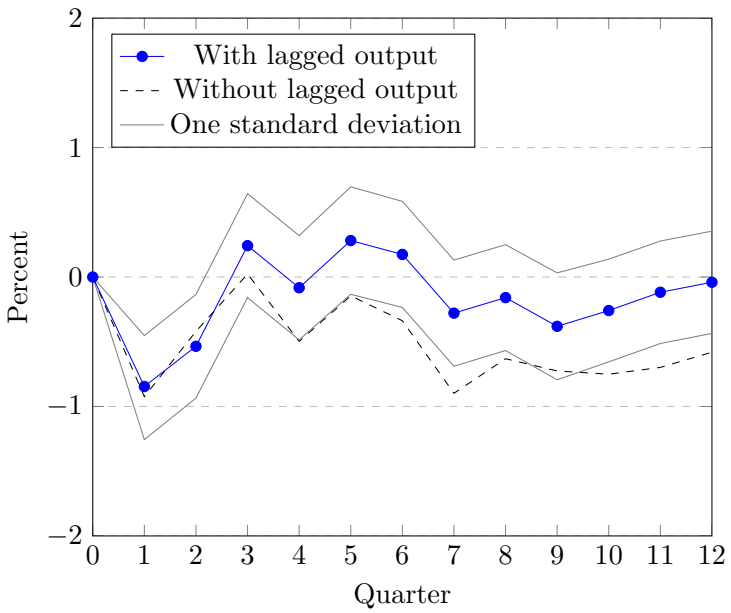


Figure 3. : Eq (9) - The effect of tax on output (*controlling for lagged output*)

C. Interpretation of the results

The estimation results from equations (8) and (9) enable me to draw several conclusions on the behaviour of real log GDP responding to a 1% increase of tax.

First, the contemporaneous response of GDP to a tax increase matters: using OLS the impact multiplier is large and significant. In equations (8) and (9), the initial response is also the largest. Figure 2 (equation 8) shows that the cumulative multiplier in quarter 7 is about as large as the impact multiplier (yet not significant). Figure 3 (equation 8 and 9) shows that output growth strongly negatively responds to a 1% change in tax as a percentage of GDP in the quarter that the tax change was introduced, and that the effect after 8 quarters (2 years) is slightly below zero (no lagged output) and around zero (lagged output).

It is fair to say, based on previous argumentation, that the specification with lagged output produces a better representation of the actual relationship between tax and output. A conclusion could therefore be that the effect of tax on output is strongly negative upon impact, is back at zero in quarter 3, and is not statistically different from zero in later quarters.

Second, the graphs of the estimation results of (8) and (9) show that the initial output growth decline is substantial, but so is its return to zero after quarter 2. At around quarter 3, the effects of the impact response have faded. This observation leads to the

claim that output growth adjusts to a tax shock in about 3 quarters.

It is questionable whether we should attach any significance to the observed output decline after quarter 5 in Figure 3. None of lags 5 through 12 are significant, on itself providing enough ground to dismiss the results for these quarters. However, the decline in the later quarters is mostly explained by lag 6, with -0.56 (t = -1.36) in the estimation results of equation (8) and -0.45 (t = -1.11) in equation (9).

What could be an explanation of a sudden decline 1.5 years after the the tax shock? A logical interpretation is that the estimation captures both short and long-run effects. In response to an increase in tax, households may reduce savings to keep consumption constant and firms may alter short-run business decisions. In the longer run, higher tax levels may distort the labour-leisure trade-off. In both equation (8) and (9), a tax increase seems to have a sustained negative effect on log GDP growth of about 0.4% after 1.5 years.

Compared to the results of RR, I find maximum implied multipliers with about a third in size (RR find a maximum multiplier of over 3). The graphical representation of the RR estimations do not exhibit the 'downward opening parabolic shape', with a multiplier (as is the case in this paper) that first declines, then returns to about zero but further declines after. Compared to RR, fewer lag coefficinents are significant in my estimations. An explanation could be the smaller variation in tax policy (see also Section V: C).

My results are in line with estimations using the SVAR method. I find a maximum contractionary response of about 0.9% of GDP, comparable to results from Blanchard and Perotti (2002) and Perotti (2012). This effect is higher than the impact multiplier of the Saffier II structural macromodel used by the The Netherlands Bureau for Economic Policy Analysis (impact of about -0.4%, maximum contractionary effect of -1.6%). The estimated effect is much higher than the impact multiplier of the Delfi-model (0.1%), but is comparable to its maximum cumulative multiplier (0.8%). Compared to the narrative studies on the US, the multipliers presented in this paper are on the lower end of the spectrum, but seem to represent about an average when I consider all studies presented in Table 1 of Section II: E. There are no empirical studies for the Netherlands to which I can compare my results.

Two features of the Dutch economy could explain why I estimate smaller output responses to a change in tax than similar narrative studies do on the US, UK and Germany. First, the economy of The Netherlands is exceptionally open. Trade as a percentage of GDP was 96% in 1960 and 154% in 2015.[26][27] As a comparison, trade was the size of

[26] Source: CBS.

[27] Much of the Dutch imports are re-exported and much of its exports were imported. Subtracting re-exportation would approximately cut the mentioned figures by half—yet it is still fair to consider the economy of the Netherlands as very open.

9% of GDP of the United States in 1960, and 28% in 2015.[28] Multipliers tend to be lower in countries more open to trade as a tax shock will spread into other countries through imports (Barrell et al., 2012; Ilzetzki et al., 2013). Second, the Dutch economy has relatively large automatic stabilisers. The automatic response of transfers and taxes offsets part of the initial shock. Dolls et al. (2012) provide ground for this claim, by showing how demand stabilisation is substantially larger in the Netherlands than in the US. The de facto fixed exchange rate regime in almost all years in the sample period partly offsets these suppressing effects.[29]

[28] I obtain trade-to-GDP figures from World Bank Open Data and figures on Dutch re-exportation from Statistics Netherlands.

[29] The guilder was factually a gold exchange standard in the sample period to 1971 under Bretton Woods. In 1983, the guilder was officially pegged to the Deutsche Mark. The currency was replaced by the euro in 1999 (scriptural) and 2002 (fully).

VII. Robustness

If the identified tax shocks are truly exogenous, the estimation results reflect the true response of output to a tax change and controlling for other elements would not be necessary. However, measurement error is always a possibility: I might have erroneously considered endogenous tax policy to be exogenous. In addition, the fact that I estimate several non-significant lag coefficients could give rise to the question as to whether the estimated responses are correlations by chance. I test the sensitivity of my findings by dropping outliers and check the robustness of my results by adding control variables.

A. Dummy for quarters with both exogenous and endogenous measures

I have pointed to the advantage of centralised taxation and a clear implementation date of new taxes earlier in this paper. That many tax changes take effect at the same moment also presents a downside: the introduction of exogenous tax changes may coincide with the introduction of endogenous tax changes. Sticking to my classification scheme, I obviously do not consider endogenous tax changes in my estimation. This neglects the possibility that endogenous measures may affect GDP contemporaneously with an exogenous tax change, the latter which is part of the estimation. By removing all endogenous tax changes that take effect in the same quarter as exogenous tax changes, I could introduce endogeneity in the model by omitting a necessary variable (the endogenous tax change). This might bias my results.

To control for this possibility I take a dummy variable (without lags) for the quarters with both exogenous and endogenous legislated tax measures (16 in total). I run equation (9), the baseline specification with lagged output along with the dummy variable. The dummy has a coefficient of -0.01 with t = -1.37. I include the estimation output with standard errors for the contemporaneous cofficient and its lags in Appendix B.

Figure 4 shows the result of the implied path of log GDP with the added dummy variable. Controlling for quarters with both endogenous and exogenous tax policy hardly changes the output response. The figure shows that my estimations are unlikely to be influenced by quarters in which both exogenous and endogenous tax changes took effect.

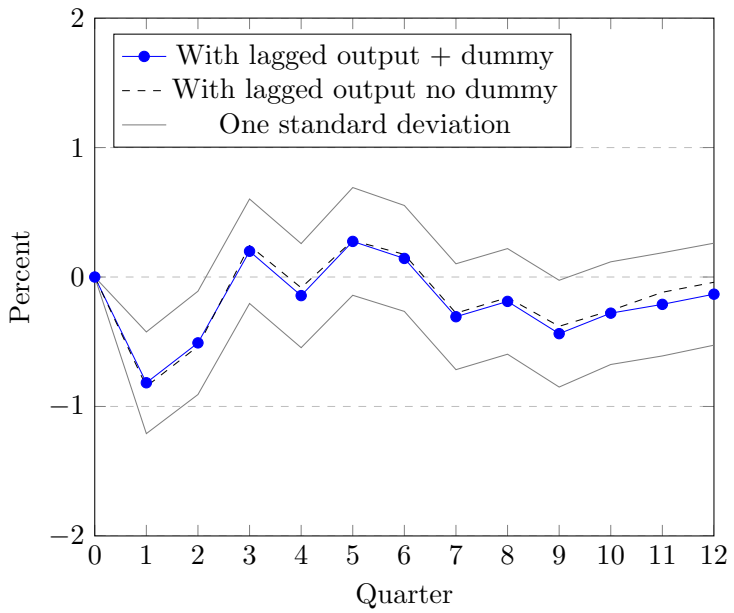


Figure 4. : The effect of tax on output (*lagged output and mixed tax change dummy*)

B. The effect of tax on output for specific time periods

I further run estimation (9) for subsets of my sample to see if the behaviour of output to a tax shock is robust over time. I choose three time periods: 1960-1980, 1981-1999 and 2000-2015.

The following reasoning constitutes my choice. Between 1960 and 1980, fiscal policy followed a predominantly Keynesian doctrine aiming to dampen business cycle fluctuations (De Kam, 2015). I find a considerable number of endogenous tax changes for these years which are not part of my estimation. The period between 1981 and 1999 was characterised by a supply side oriented doctrine, with many tax changes aiming to raise output structurally in the longer run. The first part of the 2000-2015 time frame is marked by a strong supply-side doctrine, coupled with multiple deficit reducing endogenous tax changes later in the period. While I only estimate the effect of exogenous tax changes, it is still interesting to see if output responses in different decades vary, each characterised by different tax policy.

I graph the results in Figure 5 and present the regression output in Appendix C. A word of caution is in order: using parts of my sample (thereby using fewer observations) greatly reduces the significance of the estimated coefficients. For example, now only the contemporaneous value (the impact multiplier) for 1981-1999 is significant.

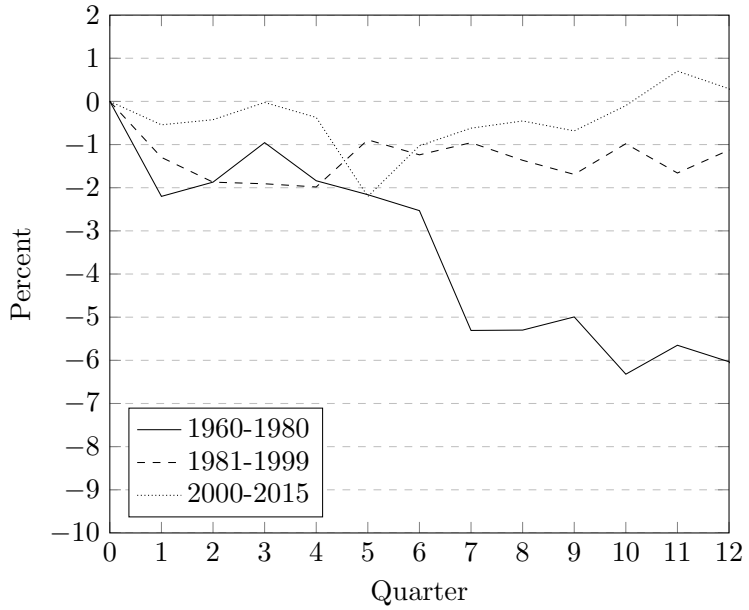


Figure 5. : The effect of tax on output for individual periods

Figure 5 shows that all multipliers on impact are negative, but that their sizes vary widely. Immediate output responses for 1960-1980 and 1981-1999 exceed 2% and 1% respectively, while the impact effect of a 1% tax increase of GDP for the period 2000-2015 is about 0.5%. The surge in the multiplier between quarter 1 and 3 in the estimation of (9) with the full period is observed in both time periods 1960-1980 and 2000-2015. The maximum contractionary effect of 1981-1999 and 2000-2015 is about 2% of GDP, while the implied output effect for 1960-1980 drops sharply after quarter 6 to a whopping -6% in quarter 10. The effect in quarter 6 (-2.78%) is especially strong and significant (t=-2.20). The standard errors (see Appendix C) are implausibly large (smallest 1.13 and largest 1.28). No economic logic can explain this observation.

Given the many non-significant coefficients and the limited number of observations it is impossible to draw strong conclusions. Nevertheless, takeaways from this exercise could be that 1) impact multipliers for the subset time periods are negative, 2) a tendency to make good on parts of the initial sharp decline is observed in quarter 3 for both periods 1960-1980 and 2000-2015, and 3) disregarding the anomalies of 1960-1980, the estimates beyond quarter 5 point to a sustained multiplier between -0.5% and -1.0% approximately, 4) the earliest period characterised by demand-side measures produced strongest output effects, and 5) the period 2000-2015 mainly characterised by supply-side and deficit-reducing measures produced the smallest response of output to a change in taxes. This is also the period in which tax changes as a percentage of GDP were smallest.

C. *Alternative tax measures*

Large tax shocks are legitimate observations to consider. However, it would be instructive to see whether they drive the main results. I experiment by removing outliers from the sample and estimate equation (9) with lagged output. I remove the three largest observations, which range from about 1% of GDP to about 2.5% of GDP in 1976q1, 1983q1 and 1986q3. A summary is given by Table 9 and the estimation results are displayed in Appendix D.

Table 9—: Effect of alternative tax measures

Specification	Initial response of GDP to a tax increase of 1% of GDP (standard error)	Maximum contractionary impact on GDP of a tax increase of 1% of GDP (standard error)
Without upper outliers	-1.28%	-1.41%
(1960:I-2015:I)	(0.64)	(0.64)
Without lower outliers	-0.88	-0.88
(1960:I-2015:I)	(0.39)	(0.39)
Full tax measure	-0.38	-0.38
(1960:I-2015:I)	(0.37)	(0.37)

Removing the upper outliers causes the estimated coefficient to be more negative. The impact multiplier is now -1.28% (previously -0.92%) and the maximum contractionary impact on GDP of a tax increase of 1% of GDP is -1.41% (previously also -0.92%). The impact multiplier is highly significant with t = -2.41. Apparently, the three large tax packages have a smaller effect on GDP growth than the average response to the other tax shocks. Anticipation effects can plausibly account for this observation. For example, the large tax hike in 1983 (2.5% of GDP) was accompanied by vast media attention.[30] Stronger behavioural responses in anticipation of the introduction of the tax package could lead to a lower observed response in the time period studied.

I perform the opposite test by removing all minimum outliers: I remove all observations smaller than 0.01% of GDP (21 quarters in total).[31] This does not change my baseline result: the impact multiplier is now -0.89 with t = -2.28 (also the maximum contractionary impact on GDP growth).

[30]See for example the digitally accessible copy of Leidse Courant of 25 April 1983, page 5, at http://alturl.com/oqgap.

[31]The reader might remember that tax changes smaller than 0.01% of GDP were in principle not accounted for in the dataset. However, as I often encounter multiple tax changes in the same quarter that partly off-set each other in size, it might still occur that the size of the observations is smaller than 0.01% of GDP.

I further estimate (9) using the full measure of tax changes of both exogenous and endogenous tax changes. The analysis in Section II has shown that estimating the effect of tax on output by including endogenous tax measures is likely to underestimate the tax multiplier. The result shown in Table 9 is just what one would expect: the estimated implied multipliers are substantially lower than the exogenous baseline result with a contemporaneous effect on GDP of -0.38%. This is also the maximum contractionary effect. Standard errors are large (about 0.37 for each coefficient) and none of the coefficients are significant.

D. Controlling for other shocks to output

I add several control variables to equation (9) and present the estimation results in Appendix E. For all control variables, I take the contemporaneous value and 12 lags.

The most obvious control variable to consider is government spending. In identifying suitable tax changes for the estimation of the output effects to tax changes, I have isolated all exogenous tax changes possibly correlating with changes in government spending. The tax changes in my estimation should therefore be uncorrelated with changes in spending. However, especially in smaller samples, variables could correlate by chance. Additionally, government spending can have strong output effects and correlation with my tax measure could bias the estimation results.

There are no quarterly figures available for spending by the Dutch government prior to 1993. I therefore intrapolate yearly data. This inherently reduces the suitability of the measure as a control variable, but is nevertheless the measure closest to actual quarterly changes in government spending. I use euro figures on changes in government spending, before subtraction of depreciation and minus interest payments.

Table 10 shows that adding government spending as a control variable in the estimation has little effect on the initial response of GDP to a tax increase of 1% of GDP. The implied impact multiplier is -0.77%, which is also the maximum contractionary effect on GDP growth. My exogenous tax measure does not seem to be correlated with my series of spending changes.

I further control for monetary policy by adding the policy interest rate to equation (9). The policy interest rate plausibly correlates with output and reflects multiple elements influencing GDP. For example, interest rates may be higher in times of high inflation or after an increase in spending by the government. I take the prevailing interest rate at the end of the quarter, which means that to the first quarter of the year I assign the interest rate of 31 March, to the second quarter the rate on 30 June, and so forth.

Table 10—: Effect of adding a control variable

Control variable	Initial response of GDP to a tax increase of 1% of GDP (standard error)	Maximum contractionary impact on GDP of a tax increase of 1% of GDP (standard error)
Government spending	-0.77%	-0.77%
(1960:I-2015:I)	(0.41)	(0.41)
Policy interest rate	-0.58	-1.24
(1960:I-2015:I)	(0.34)	(0.34)
Consumer price index	-0.87	-1.17
(1963:I-2015:I)	(0.36)	(0.32)
Relative price of oil	-0.96	-0.96
(1960:I-2007:IV)	(0.44)	(0.44)

Adding interest rate reduces the initial response of output to a tax change: it is now -0.58% (t= 1.68). The maximum contractionary response is obtained after 12 quarters and is -1.24%. I add another monetary variable to the equation by controlling for inflation. I use a measure of the consumer price index, assigning CPI to the quarters in the same way as I have done for the policy interest rate. The estimated impact multiplier is roughly similar (and highly significant), the total response of tax to output is now -1.17%.

Finally, I control for supply shocks by including the relative price of oil. I use the RR variable with a time period 1960:I to 2007:I. The impact multiplier is almost similar to the estimation of the baseline and the coefficient is significant. The total effect of output to tax is similar to the impact multiplier.

All in all, my results are mostly robust. Initial responses of GDP to a tax increase of 1% of GDP with these control variables range from -0.58% to -0.96%. For all control variables except government spending, maximum contractionary effects on GDP are slightly higher than without these control variables. It is worth mentioning that lags for these later quarters are mostly non-significant, while impact multipliers are mostly (overwhelmingly) significant.

E. Implementation lags

Finally, I experiment with several implementation lags to account for the possibility that real effects of tax changes are not observable in the same quarter as the law change, but only several quarters later. I perform the estimation with a lag of 2, 4 and 6 quarters and present the main results in Table 11 and the full results in Appendix F.

Table 11—: Implementation lags

Lags (quarters)	Initial response of GDP to a tax increase of 1% of GDP (standard error)	Maximum impact on GDP of a tax increase of 1% of GDP (standard error)
2	0.87% (0.40)	0.87% (0.40)
4	0.04 (0.40)	-1.10 (0.38)
6	-0.21 (0.39)	-1.33 (0.38)

Assigning the tax changes to different quarters changes my results compared to the baseline. By using two lags, I observe a large positive impact of a tax increase on GDP. The effect of the tax increase on output remains positive throughout. By using four and six quarter lags I obtain maximum contractionary effects that are larger than the baseline estimation. The initial response of GDP to a tax increase of 1% of GDP is considerably smaller: slightly positive with four lags (0.04%) and negative for six lags (-0.21%). None of the estimated values with four and six quarter lags are significant at the 5% level.

Two observations are striking. First, using 2 quarter lags produces positive output responses to a tax change. It is hard to give a single explanation why the estimation with an implementation lag of two quarters gives a sharp positive result. Because of this positive initial response of GDP to tax, none of the effects in later quarters is negative, as is the case with the results from the baseline specifications. What is obvious, though, is that this output movement is also observed in the baseline specification results, about 2 quarters after the introduction of a tax change. This may point to the possibility that the real effect of a tax change is in fact present in the same quarter as the introduction of a tax change. In other words, this suggests that the way I have assigned tax changes to the quarters in the baseline is actually correct. It is also worth mentioning that two quarter implementation lags possibly present the least realistic scenario, as the implementation lag is unlikely to take two quarters.

Second, the sharp initial responses of the baseline result are not present in the estimation with four and six quarter implementation lags, although the maximum impact on GDP is comparable yet somewhat larger than the baseline (-1.10% and -1.33% for four and six lags respectively). Implied multipliers with implementation lags 4 and 6 are therefore roughly equal to my baseline specification, although the coefficients for the higher lags in the estimation are mostly non-significant.

The results in Table 11 should be interpreted with some caution. I did not disentangle between different types of tax and I have not accounted for the fact that although most tax changes take effect in January of a year, there are also many tax changes taking effect around July. This could potentially influence the results of this robustness check.

VIII. Lessons and policy implications

What could policymakers take away from the analysis and results presented in this thesis? In the introduction, I wrote that knowledge on the tax multiplier improves the precision of tax policy instruments. Knowing the effects of tax on output, I argued, policymakers would be better able to design tax policy in accordance with policy objectives. I draw three lessons and two policy implications from the analysis in this thesis. I end this section with a simple counterfactual analysis of the tax increases in the Netherlands between 2011 and 2017.

Lesson 1: *Tax hikes contract output, tax cuts raise output*

My results suggest that a tax hike reduces output growth, while a tax cut produces the opposite effect. A tax cut mutating revenue by one euro reduces output by about as much. Taking my baseline result, I have estimated that 1% increase in tax as a percentage of GDP leads to an output decline of 0.9%. Given that I have about as many liability mutations (in size) with a positive value (tax increases) as a negative value (tax cuts), the result of this paper could also be interpreted to mean that for any 1% decrease in tax as a percentage of GDP, output rises with 0.9%.

Lesson 2: *Empirical support for Saffier II and Delfi models, but no consensus on timing*

Given the absence of other empirical estimations for the Netherlands, the results of this paper are best compared to the multipliers in macromodels. The Saffier II model of The Netherlands Bureau for Economic Policy Analysis and the Delfi model produce implied total multipliers that are roughly similar to the multipliers estimated in this study. My results are not in congruence with the models, however, on the timing. The structural (Saffier II) and DSGE (Delfi) model have small impact multipliers, that rise over multiple years to be about as large as the maximum responses estimated in this study.

Lesson 3: *No multiplier equals the average*

The estimations in this paper are based on historic realisations of the effect of tax on output in the Netherlands. They therefore present a guiding tool for policymakers, but should not be interpreted as universal tax multipliers for the Dutch economy. It should be clear from the analysis in the first part of this thesis that multipliers may vary across levels of economic activity, also within an economy.

Implication 1: *Tax policy a potent tool for stabilisation*

Output responses to a change in taxes are substantial and significantly different

from zero. This finding implies that tax policy can serve as an instrument for economic stabilisation. The estimation results suggest that output responses to tax changes are considerable on average, not only at times of low economic activity.

Implication 2: *Higher taxation reduces the tax base, lower taxation expands it*

A tax hike erodes the tax base. A lower output as a result of a increase in the average level of taxation will also lead to a reduction in future tax revenue. Tax increases to reduce a budget deficit thus come against considerable costs. On the other hand, the considerable effects of a tax cut increase output and therefore the tax base.

Counterfactual: *Cumulative output loss between 2011-2017*

How policymakers could interpret (and work) with the multiplier estimated in this study is best shown by presenting a counterfactual analysis of Dutch tax policy in recent years. Between 2011 and 2017, the Dutch government introduced austerity measures and cut government spending in an effort to reduce the budget deficit. Deficit-reducing measures between 2011 and 2017 amounted to 7 billion euro a year on average (about 1% of GDP per year), totalling 46 billion euro over this period (Suyker, 2016). This package increased the effective average tax burden by about 16 billion euro, as shown by Table 12.

Table 12—: Deficit reducing measures (in billion euros), leading to a mutation of the tax burden: 2011-2017

	2011	2012	2013	2014	2015	2016	2017
Coalition agreement Rutte-Verhagen	1	1	4	5	6	6	6
Budget agreement 2013		1	9	7	6	7	7
Coalition agreement Rutte-Asscher			1	1	4	5	5
Housing and social agreement			1	1	0	0	0
6 billion package (including budget agreement 2014)			-2	3	2	2	1
Budget 2015					-1	-1	-1
5 billion package						-4	-4
Budget 2017							2
Total EMU tax burden mutation	1	2	13	17	18	14	16

The above table shows the cumulative effect of deficit reducing measures that led to an increase of the tax burden for the period 2011 to 2017. Some of the packages lowered taxes (such as the tax relief package of the budget 2016 and additional measures in 2017), but the net effect is a vast increase of the implied average tax burden.

To show the effect of the multiplier in deficit-reducing tax changes, I perform a rudimentary back-of-the-envelope calculation. I compute the structural output loss with a tax multiplier of 0.9 over the period 2011-2017 and present the results in Table 13.

Table 13—: Cumulative output loss

Year	Cumulative output loss (in bln euro)
2011	0.9
2012	1.8
2013	11.7
2014	15.3
2015	16.2
2016	12.6
2017	13.5
Total	72

The cumulative output loss is obtained by multiplying the total EMU tax burden mutation with the estimated multiplier. I assume a constant multiplier of 0.9 over the years. The cumulative output loss over 2011-2017 then equals 72 billion euro. With an average tax burden of 37.8% on every euro,[32] this means a loss in government revenue worth 27.2 billion euro, or about 4.2% of GDP in 2011 figures.

This quick calculation obviously does not do justice to the complexity of the circumstances in which these deficit-reducing measures were considered. It is therefore ill-suited to provide grounds for an analysis on the need or desirability of these measures. It could, however, contribute to explaining the movement of the budget deficit which—despite the extensive efforts to reduce it—improved from -4.3% in 2012Q3 to -2.4% in 2013Q4, and declined to -3% in the two quarters after.

[32]Figure over 2016. Source: CBS.

IX. Conclusion

This paper investigates the effect of tax changes on output. I find that much of Dutch tax policy between 1955 and 2015 can be categorised as either exogenous or endogenous to output. Exogenous tax changes are not influenced by output and unrelated to other factors affecting output, so they can be used to estimate output effects to a tax change. Out of 322 observed tax changes, I identify 289 as exogenous.

The estimation results suggest a decline in output growth as a result of a tax increase. Following a tax increase of 1% of GDP, output growth decreases by -0.92% in the specification without lagged output, and by -0.85% in the specification with lagged output. The estimated contemporaneous responses can be interpreted as impact multipliers and are highly significant. By removing upper outliers, I find a maximum contractionary impact on GDP as large as -1.41%.

This thesis ties in with comparable work using the narrative method on the US, UK and Germany. My estimated output responses are smaller. As the analysis has shown, the open Dutch economy and the presence of powerful automatic stabilisers provide an explanation for these differences. There is a lack of other empirical work on the Dutch tax multiplier, which makes it impossible to compare results. My results are in line with estimations of multiplier in other countries using the SVAR method. I find a maximum contractionary response of about 0.9% of GDP, comparable to results from Blanchard and Perotti (2002) and Perotti (2012). This effect is higher than the impact multiplier of the Saffier II structural macromodel used by the The Netherlands Bureau for Economic Policy Analysis (impact of about -0.4%, maximum contractionary effect of -1.6%). The estimated impact effect is higher than the impact multiplier of the Delfi-model (0.1%), but is comparable to its maximum multiplier (0.8%). Compared to the narrative studies on the US, my multiplier are on the lower end of the spectrum, but seem to represent about an average when I consider other notable studies to date.

The output decline upon impact is significant and sharp. In the specification with lagged output, the output response returns to zero and this movement is significant too. This result suggests a short-run adjustment to a tax shock after approximately 6 months. In later quarters, the output response is below zero, but estimations are not significant. All in all, output effects seem to have faded 2 years after the tax shock, but there is no reason to believe that estimates for these later quarters are statistically different from zero.

While the implied impact multipliers of the baseline estimations are large and significant, estimations are not very precise. In line with the literature, I have shown the estimation results with wide confidence intervals. Data on Dutch tax policy, moreover,

provides substantially less variation than tax changes in comparable studies for other countries. Although this is not problematic, it could be an explanation of the many non-significant coefficients. It is worth noting that extending the scope of the study by including observations before 1960 is unlikely to add variation to the tax measure. While I find many legislated tax changes, tax changes as a percentage of nominal GDP are mostly small. In other words, the Dutch tax system is extensive and is caters to specific policy issues, but is relative to the size of the economy not large.

This study recognises that some tax changes are useful for the estimation of the effect of tax on output, and others are not. I have closely followed the identification approach of the literature on narrative studies. It is therefore unlikely that classifying exogenous tax changes in a different way will result in much improvement or additional variety to the data. The fact that most tax changes take effect on the same date (on 1 January), means that they turn up as one and the same observation in a quarter. This feature of the Dutch legislative process is a natural limitation to the number of observations and variation in the data.

The companion paper to this thesis documents legislated tax changes in 60 consecutive postwar years. Although it is the result of an effort to record all legislated tax policy, it does not claim to be exhaustive. I am by default limited to available documents. Thanks to the rich sources of extensive digital archives I was able to recover many legitimate observations, at least of a sufficient number to estimate the effect of tax on output. However, correctly assigning the timing to each has proven to be a challenging exercise. Tax proposals may receive more attention in government documents than their final implementation. It follows that narrower inspection of the primary sources could reduce possible measurement error in the dataset.

My results are mostly robust. Controlling for government spending, interest rate, inflation and supply shocks alter implied impact and total multipliers, but not considerably. On average, the response of GDP to a tax change is equal to the liability mutation that follows from the tax change.

Additional research on this topic could help to gain more insights in the effects of specific taxes on output. The newly constructed dataset could prove useful for that purpose. For example, what output effects are likely to follow from a change in income tax, or corporate tax? Which taxes are likely to produce the largest output effects? Knowledge on this could prove useful to policymakers: when policy should be countercyclical, which taxes are specifically effective in raising output in the short-run? In addition, the narrative analysis points to an even more extensive budgetary policy of the Dutch government. Analysis of the multiplier in the Netherlands could surely extend to the macroeconomic effects of public spending changes.

Another interesting avenue to take would be to experiment further with the way tax changes are matched with quarters. In this study, I have kept close to the methodology of RR in assigning tax changes to the quarter they take effect. Although there is much reason to suggest that real effects of tax changes are observable in the same quarter as the quarter in which a tax policy takes effect, features of the Dutch legislative process and revenue collection make it likely that output is affected (mostly) in the following year. Disentangling between tax changes and assigning lags to each based on the most likely implementation lags could add to the precision of the estimation of the effect of tax on output.

References

Acconcia, A., Corsetti, G., & Simonelli, S. (2014). Mafia and Public Spending: Evidence on the Fiscal Multiplier from a Quasi-experiment. *American Economic Review*, *104*(7), 2185-2209.

Auerbach, A. J., & Gorodnichenko, Y. (2012a, February). Fiscal Multipliers in Recession and Expansion. In *Fiscal Policy after the Financial Crisis* (p. 63-98). University of Chicago Press.

Auerbach, A. J., & Gorodnichenko, Y. (2012b, May). Measuring the Output Responses to Fiscal Policy. *American Economic Journal: Economic Policy*, *4*(2), 1-27.

Barrell, R., Holland, D., & Hurst, I. (2012). Fiscal Consolidation: Part 2. Fiscal Multipliers and Fiscal Consolidation. *OECD Economics Department Working Papers*, *933*.

Barro, R. (1974). Are Government Bonds Net Wealth? *Journal of Political Economy*, *82*(6), 1095-1117.

Barro, R. (1981). Output Effects of Government Purchases. *Journal of Political Economy*, *89*(6), 1086-1121.

Barro, R. (2007). *Macroeconomics: A Modern Approach*. Cengage Learning.

Barro, R. (2013). *Multipliers*. University Lecture (Harvard University).

Barro, R., & Redlick, C. (2011). Macroeconomic Effects From Government Purchases and Taxes. *The Quarterly Journal of Economics*, *126*(1), 51.

Batini, N., Eyraud, L., Forni, L., & Weber, A. (2014). Fiscal Multipliers: size, determinants and use in macroeconomic projections. *Technical Guidance note*.

Baxter, M., & King, R. G. (1993). Fiscal Policy in General Equilibrium. *The American Economic Review*, *83*(3), 315-334.

Beetsma, R. (2008, January). A Survey of the Effects of Discretionary Fiscal Policy. *Amsterdam School of Economics Research Institute*, *2*.

Bhattarai, K., & Trzeciakiewicz, D. (2017). Macroeconomic Impacts of Fiscal Policy Shocks in the UK: A DSGE analysis. *Economic Modelling*, *61*, 321–338.

Blanchard, O., & Leigh, D. (2013, May). Growth Forecast Errors and Fiscal Multipliers. *American Economic Review*, *103*(3), 117-20.

Blanchard, O., & Perotti, R. (2002). An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output. *The Quarterly Journal of Economics*, *117(4)*, 1329-1368.

Blanchard, O., & Quah, D. (1989, September). The Dynamic Effects of Aggregate Demand and Supply Disturbances. *American Economic Review*, *79*(4), 655-673.

Blanchard, O., & Watson, M. (1986). Are Business Cycles All Alike? In *The American Business Cycle: Continuity and Change* (p. 123-180). National Bureau of Economic

Research, Inc.

Bovens, M. A. P., Brandsma, G., Thesingh, D., & Wever, T. (2010). Aan het pluche gekleefd? Aard en achtergrond van het aftreden van individuele bewindslieden 1946-2009. *B en M: Tijdschrift voor Beleid, Politiek en Maatschappij*, *37*(4), 319–340.

Burnside, C., Eichenbaum, M., & Fisher, J. (2004). Fiscal shocks and their consequences. *Journal of Economic Theory*, *115*(1), 89-117.

Caldara, D., & Kamps, C. (2008). *What are the Effects of Fiscal Policy Shocks? A VAR-based Comparative Analysis* (Tech. Rep.).

Caldara, D., & Kamps, C. (2017). The Analytics of SVARs: A Unified Framework to Measure Fiscal Multipliers. *The Review of Economic Studies Limited*.

can severe fiscal contractions be expansionary? tales of two small european countries. (n.d.).

CBO. (2010, August). *Estimated Impact of the American Recovery and Reinvestment Act on Employment and Economic Output from April 2010 Through June 2010 Report* (Report). Congressional Budget Office.

Chinn, M. (2013). Fiscal Multipliers. In S. N. Durlauf & L. E. Blume (Eds.), *The New Palgrave Dictionary of Economics.* Basingstoke: Palgrave Macmillan.

Chodorow-Reich, G. (2017, July). *Geographic Cross-Sectional Fiscal Spending Multipliers: What Have We Learned?* (Working Paper No. 23577). National Bureau of Economic Research.

Chodorow-Reich, G., Feiveson, L., Liscow, Z., & Woolston, W. G. (2012, April). Does State Fiscal Relief during Recessions Increase Employment? Evidence from the American Recovery and Reinvestment Act. *American Economic Journal: Economic Policy*, *4*(3), 118-45.

Christ, C. (1994). The Cowles Commission's contributions to econometrics at Chicago, 1939-1955. *Journal of Economic Literature*, *32*(1), 30-59.

Chung, H., & Leeper, E. M. (2007, September). *What Has Financed Government Debt?* (Working Paper No. 13425). National Bureau of Economic Research.

Clemens, J., & Miran, S. (2012). Fiscal Policy Multipliers on Subnational Government Spending. *American Economic Journal: Economic Policy*, *4*(2), 46-68.

Cloyne, J. (2013, June). Discretionary Tax Changes and the Macroeconomy: New Narrative Evidence from the United Kingdom. *American Economic Review*, *103*(4), 1507-28.

Cogan, Cwik, Taylor, & Wieland. (2010). New Keynesian versus old Keynesian government spending multipliers. *Journal of Economic Dynamics and Control*, *34*(3), 281 - 295.

Commission of the European Communities. (2008, November). *Communication from the Commission to the Council: A European Recovery Plan* (Report). Commission of the European Communities.

Coudret, N. (2013). *Fiscal Multipliers in the Netherlands* (Master's thesis).

CPB. (2006, February). *De Nederlandse Collectieve Uitgaven in Historisch Perspectief* (Tech. Rep. No. 109). Centraal Planbureau/Netherlands Bureau for Economic Policy Analysis.

CPB. (2010, December). *Saffier II: 1 model voor de Nederlandse Economie, in 2 hoedanigheden, voor 3 toepassingen* (Tech. Rep. No. 217). Centraal Planbureau/Netherlands Bureau for Economic Policy Analysis.

CPB. (2016, June). *De Politieke Economie van de Hervormingen in de Loon- en Inkomstenbelasting* (Tech. Rep.). Centraal Planbureau/Netherlands Bureau for Economic Policy Analysis.

De Grauwe, P., & Ji, Y. (2013, October). *The Legacy of Austerity in the Eurozone* (CEPS Papers No. 8472). Centre for European Policy Studies.

Devries, P., Guajardo, J., Leigh, D., & Pescatori, A. (2011). A New Action-Based Dataset of Fiscal Consolidation. *IMF Working Papers*, *No. 11/128*.

DNB. (2011). *DELFI: DNB's Macroeconomic Policy Model of the Netherlands, DNB Occasional Studies* (Tech. Rep.).

Dolls, M., Fuest, C., & Peichl, A. (2012). Automatic stabilizers and economic crisis: US vs. Europe. *Journal of Public Economics*, *96*(3–4), 279 - 294.

Farhi, E., & Werning, I. (2016). Fiscal multipliers: Liquidity traps and currency unions. *Handbook of Macroeconomics*, *2*, 2417–2492.

Favero, C., & Giavazzi, F. (2009, August). *How Large are the Effects of Tax Changes?* (Working Paper No. 15303). National Bureau of Economic Research.

Fernández-Villaverde, J. (2010, March). The econometrics of DSGE models. *SERIEs*, *1*(1), 3–49.

Friedman, M., & Schwartz, A. (1963). *A Monetary History of the United States, 1867-1960.* Princeton University Press.

Gechert, S., & Will, H. (2012). *Fiscal Multipliers: A Meta Regression Analysis* (Tech. Rep.).

Gorodnichenko, Y., Mendoza, E. G., & Tesar, L. L. (2012, June). The Finnish Great Depression: From Russia with Love. *American Economic Review*, *102*(4), 1619-44.

Gottschalk, J. (2001). *An Introduction into the SVAR Methodology: Identification, Interpretation and Limitations of SVAR models* (Kiel Working Papers No. 1072). Kiel Institute for the World Economy (IfW).

Gruijters, A., & Smulders, J. (1998). *Dertien Kabinetten, Twaalf Ongelukken?* De Nederlandsche Bank.

Hall, R. (1980, January). Labor Supply and Aggregate Fluctuations. *Carnegie-Rochester Conference Series on Public Policy*, *12*(1), 7-33.

Hall, R. (2009, November). *By How Much Does GDP Rise if the Government Buys More Output?* (NBER Working Papers No. 15496). National Bureau of Economic Research, Inc.

Hamilton, J. D. (1985). Historical Causes of Postwar Oil Shocks and Recessions. *The Energy Journal*, *6*(1), 97-116.

Hansen, B. (1969). Jan Tinbergen: An Appraisal of His Contributions to Economics. *The Swedish Journal of Economics*, *71*(4), 325-336.

Hayo, B., & Uhl, M. (2014). The Macroeconomic Effects of Legislated Tax Changes in Germany. *Oxford Economic Papers*, *66*(2), 397.

Hayo, B., & Uhl, M. (2015, 10). Regional Effects of Federal Tax Shocks. *Southern Economic Journal*, *82*(2), 343–360.

Hoover, K. D., & Perez, S. J. (1994). Post hoc ergo propter once more an evaluation of 'does monetary policy matter?' in the spirit of james tobin. *Journal of Monetary Economics*, *34*(1), 47 - 74.

Ilzetzki, E., Mendoza, E. G., & Végh, C. A. (2013). How big (small?) are fiscal multipliers? *Journal of Monetary Economics*, *60*(2), 239-254.

IMF. (2010, October). *World Economic Outlook, "Will It Hurt? Macroeconomic Effects of Fiscal Consolidation* (Tech. Rep. No. Chapter 3). Author.

introductory econometrics: A modern approach. (n.d.).

Jacobs, B. (2007). De Zalmnorm en het Begrotingsbeleid 1994-2007. *Tijdschrift voor Openbare Financiën*, *1*.

Kam, C. (1988). *Tax Reform in a Welfare State: The case of The Netherlands* (Unpublished doctoral dissertation). Rijksuniversiteit Leiden.

Kam, C. (2015). *Het land van Beloften: Opbouw, crisis toekomst van de verzorgingsstaat.* atlas contact.

Korinek, A. (2015, October). *Thoughts on DSGE Macroeconomics: Matching the Moment, But Missing the Point?* (This paper was prepared for the 2015 Conference "A Just Society" honoring Joseph Stiglitz's 50 years of teaching.)

Leeper, E. M., Traum, N., & Walker, T. B. (n.d.). clearing up the fiscal multiplier morass.

Lucas, R. (1976). Econometric Policy Evaluation: A critique. *Carnegie-Rochester Conference Series on Public Policy*, *1*(1), 19-46.

Lukkezen, J. (2013, September). *Vraag en Antwoord over de Begrotingsmultiplier: CPB Achtergronddocument* (Tech. Rep.). Centraal Planbureau/Netherlands Bureau for Economic Policy Analysis.

Lukkezen, J., Jacobs, B., & Kool, C. (2016, August). *Macro-economics of Balance-Sheet Problems and the Liquidity Trap: CPB Achtergronddcoument* (Tech. Rep.). Centraal Planbureau/Netherlands Bureau for Economic Policy Analysis.

Mertens, K., & Ravn, M. O. (2014). A Reconciliation of SVAR and Narrative Estimates of Tax Multipliers. *Journal of Monetary Economics*, *68, Supplement*, S1 - S19.

Ministry of Finance. (2010, January). *Stability Programme of The Netherlands* (Report). Ministry of Finance.

Mittnik, S., & Semmler, W. (2012). Regime dependence of the Fiscal Multiplier. *Journal of Economic Behavior Organization*, *83*(3), 502 - 522. (The Great Recession: motivation for re-thinking paradigms in macroeconomic modeling)

Mountford, A., & Uhlig, H. (2009). What are the Effects of Fiscal Policy Shocks? *Journal of Applied Econometrics*, *24*(6), 960–992.

Nakamura, E., & Steinsson, J. (2014, March). Fiscal Stimulus in a Monetary Union: Evidence from US Regions. *American Economic Review*, *104*(3), 753-92.

Perotti, R. (2008, July). In Search of the Transmission Mechanism of Fiscal Policy. In *NBER Macroeconomics Annual 2007, Volume 22* (p. 169-226). National Bureau of Economic Research, Inc.

Perotti, R. (2012, May). The Effects of Tax Shocks on Output: Not So Large, but Not Small Either. *American Economic Journal: Economic Policy*, *4*(2), 214-37.

Platenga, J. (1999). Een Afwijkend Patroon. Honderd jaar vrouwenarbeid. *R. van der Bie, & P. Dehing (red.), Nationaal goed. Feiten en cijfers over onze samenleving (ca.)*, *1999*, 91–105.

Poterba, J. M. (1986). Explaining the Yield Spread between Taxable and Tax-exempt Bonds: The Role of Expected Tax Policy [Book]. In *Studies in state and local public finance* (p. 5-52). University of Chicago Press.

Ramey, V. (2011). Identifying Government Spending Shocks: It's all in the Timing. *The Quarterly Journal of Economics*, *126*(1), 1.

Ramey, V. (2016, February). *Macroeconomic Shocks and Their Propagation* (NBER Working Papers No. 21978). National Bureau of Economic Research, Inc.

Ramey, V., & Shapiro, M. (1998). Costly Capital Reallocation and the Effects of Government Spending. *Carnegie-Rochester Conference Series on Public Policy*, *48*(1), 145-194.

Ramey, V., & Zubairy, S. (2014, November). *Government Spending Multipliers in Good Times and in Bad: Evidence from U.S. Historical Data* (Working Paper No. 20719). National Bureau of Economic Research.

Reichling, F., & Whalen, C. (2015, February). *The Fiscal Multiplier and Economic Policy Analysis in the United States: Working Paper 2015-02* (Working Papers No. 49925). Congressional Budget Office.

Riera-Crichton, D., Vegh, C. A., & Vuletin, G. (2016). Tax multipliers: Pitfalls in measurement and identification. *Journal of Monetary Economics*, *79*, 30 - 48.

Roeger, W., & in 't Veld, J. (2009, January). *Fiscal Policy with Credit Constrained*

Households (European Economy - Economic Papers 2008 - 2015 No. 357). Directorate General Economic and Financial Affairs (DG ECFIN), European Commission.

Romer, C., & Romer, D. (1989). Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz. *NBER macroeconomics annual*, *4*, 121–170.

Romer, C., & Romer, D. (2004, September). A New Measure of Monetary Shocks: Derivation and Implications. *American Economic Review*, *94*(4), 1055-1084.

Romer, C., & Romer, D. (2009). A Narrative Analysis of Postwar Tax Changes. *Unpublished paper, University of California, Berkeley (June)*.

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

Semmler, W. (2013). The Macroeconomics of Austerity in the European Union. *Social Research*, *80*(3), 883-914.

Serrato, J. C. S., & Wingender, P. (2016, July). *Estimating Local Fiscal Multipliers* (Working Paper No. 22425). National Bureau of Economic Research.

Shoag, D. (2013, May). Using State Pension Shocks to Estimate Fiscal Multipliers since the Great Recession. *American Economic Review*, *103*(3), 121-24.

Sims, C. A. (1980). Macroeconomics and Reality. *Econometrica*, *48*(1), 1-48.

Stock, J. H., & Watson, M. W. (n.d.). vector autoregressions.

Suyker, W. (2016, September). *Tekortreducerende Maatregelen 2011-2017 - Versie MEV 2017* (Tech. Rep.). Centraal Planbureau/Netherlands Bureau for Economic Policy Analysis.

Visser, M. (1993). De Lijst en Haar Trekker; Enkele kanttekeningen bij het 'Wiegel-effect'. *Liberaal Reveil*, *34*(5), 206–208.

Woodford, M. (2011, January). Simple Analytics of the Government Expenditure Multiplier. *American Economic Journal: Macroeconomics*, *3*(1), 1-35.

Zubairy, S. (2014). On Fiscal Multipliers: Estimates from a Medium Scale DSGE Model. *International Economic Review*, *55*(1), 169–195.

Appendices

Appendix A: A brief modern history of Dutch tax policy

1955-1965: Reduced need for high postwar tax levels

In the first decade after the Second World War, tax levels in the Netherlands were high: rebuilding the country was a very costly affair and much of the repair of buildings and infrastructure and payments for social security was coordinated by the state. Several new taxes had been introduced and especially rates on "taxes in the personal sphere" (BM 1955) were raised. The financial-economic situation in the Netherlands in the first postwar years had been, to use the words of the Minister of Finance Johan van de Kieft in the introductory note of the Budget Memorandum 1955, "[a situation of] national impoverishment and disorganization". The main objective of the government had been, as stated in the same text, "to stimulate production and to reduce consumption expenditure". In addition, "government action must be considered against a background of high inflation and an impaired productive capacity". This had an immediate effect on taxes: "...concerning taxes, this manifested in severe tax pressure across the board", the minister wrote.

Together with the reiteration that "taxation on itself is never an objective",[33] the years from 1955 to 1965 were characterised by reducing tax rates on income, capital and profit tax and several cost increasing taxes. The decrease in government revenue as a result of the reduction in several tax rates in 1955 were relatively large, exceeding 400 million guilders (1% of GDP). In 1956 a smaller reduction followed amounting to about 50 million guilders (0.01% of GDP). There were no substantial changes in taxation between 1957 to 1959. Another reduction in income and wage taxes followed in 1960 of 250 million (0.05% of GDP). This last package was specifically aimed at reducing "average tax pressure for those not in marital union", a first step taken—which I could find, at least—in reducing the institutional differences between married couples and singles through the tax system.

Other tax changes that easily qualify as exogenous, are all changes that address the position of "the married working female". This policy responds to structural issues on the labour market and reflects the necessity to accommodate the growing number of women with a job in the 1960s. Married women in the Netherlands did not have legal capacity and it was common practice to terminate a job contract upon marriage. A parliamentary motion ('motie-Tendeloo') in 1955 ended this for women in government jobs and marked an important step in the emancipation of (working) women (Platenga, 1999). A year later a new law granted equal legal status to married men and women, allowing married women

[33] As stated in the Budget Memorandum 1956.

to keep their jobs. In 1962, just six years after law change, the cabinet even proposed a rate reduction on the income and wage tax when women performed paid work "outside the family household". There is ample evidence of the desire to reduce the differences in fiscal treatment of married couples where the husband alone performs paid labour, and the situation in which both married partners have a job. Effectively, this led to a tax alleviation on labour and income of 55 million guilders (<0.1% GDP).

In this decade, I find evidence of 14 tax changes. 12 were aimed at reducing the average tax rate, which I all classify as exogenous. The tax rate reductions follow after a period of sustained high tax levels that were an immediate consequence of the war. One tax change concerns a tax escalator, which I drop, and there is one tax decrease that I classify as endogenous.

I do not find evidence of countercyclical tax measures in the first decade of the dataset. The one tax change that I do classify as endogenous (taking effect in 1965), is a reduction in the wage and income taxes. There is no mention in the sources that the objective of the rate reduction is to raise output in the short term, which would be an endogenous motivation. The government attempts to address the progressiveness of the tax system: rising wages caused workers to 'progress' in tax bands, effectively resulting in heavier taxation on wages and income. The recovering world economy is mentioned in the sources to be a main driver of rising wages. The mentioned "urgency to reduce tax rates" effectively follows from a rise in output.

1966-1975: Extension and modernization of the tax system

The policy sources give account of 70 legislated tax changes in the period between 1966 and 1975. Where in the period 1955-1965 the vast majority of tax changes were related to the alleviation of tax pressure through a reduction in rates for the income and labour tax, I find more diverse motivations in the following decade. First, I find a continuing trend of lowering average rates (per 1 July 1967, the rates of the income and labour tax were reduced once more, leading to a decrease in tax revenue for the government of 710 million guilders, or 0.8% of GDP). Second, several new taxes were introduced (such as the VAT and the new profit tax in 1969). Third, the existing tax system was extended by several specific taxes (such as the change in the rate structure on mineral oils and petroleum).

Typical for this decade is that the government starts to approach tax policy as part of a broader policy programme, and clearly connects some tax changes to spending related measures. Taxes were used to finance expansionary budget policy, and are labelled 'coverage measures' ('dekkingsmaatregelen') in the policy documents. In 1968, taxes on specific consumption goods ('opcenten') were increased, totalling 131 million (0.1% of

GDP), to offset a spending increase in the same year. The government finds mostly less distortionary taxes to keep a spending change budget neutral, and raises taxes on soft drinks, beer and wine, confectionery and mopeds. Typical excises on consumption goods that serve to finance government expenditure are those on tobacco, alcoholic beverages, soft drinks, petroleum for end-users and motorised vehicles. It is interesting to see that taxes in the 1960s were commonly only temporarily raised to finance another tax or spending increase: the 47$^{\text{th}}$ percentage point of the corporate tax was extended by one year, in 1967, as part of a package to cover for several spending increases. As the coverage measures that I find between 1966 and 1975 are taken to offset another factor affecting output (a spending change), I classify these tax changes as endogenous.

The expansion of the customs union of the European Economic Community (EEC) was reflected in several measures to harmonise taxes between Member States. Increasing tariffs, for example, led to a gain of 130 million in 1968 (<0.1% GDP). Additionally, the harmonisation of excise duties between Belgium, the Netherlands and Luxembourg[34] resulted in a decrease in tax revenue (140 million or <0.1% GDP). Following multilateral talks for the reduction of trade tariffs led to a loss in tariff revenue. A notable example is the fourth phase of implementation of the Kennedy-round in 1971,[35] to lower several tariff rates. While arguing that EEC and GATT changes imposed by external bodies would carry too far (the Netherlands could in theory vote against such changes), it is safe to argue that these changes are not correlated by (a changing) output but are part of ongoing long-term efforts to reduce tariff rates and facilitate trade. Changes to taxes as result of harmonisation and tariff rate reduction are therefore all classified as exogenous.

Compared to the earlier decade, policy objectives of tax changes between 1966 and 1975 were more precisely formulated. Several taxes were aimed to address issues specific to certain (income) groups and to provide incentives. Clearly, taxation evolved from a system primarily setup as a source of government income to a potent instrument for redistribution and inducing certain behaviour. Good examples are specific measures for elderly to reduce taxes on old age pensions ('optrekking inkomensaftrek voor bejaarden') and multiple facilities (effectively leading to tax credits) for the self-employed to stimulate retirement saving in 1973.[36] Recurring tax changes in this time period are the extensions of save-as-you-earn schemes, where (payroll) workers could save a fraction of their gross income, making saving especially attractive.[37] All of these 'specific' policies targeting

[34]These countries are part of a politico-economic union founded in 1944, that originated as a customs union. Agreements on the free movement of capital and labour between these countries preceded agreements within the framework of the EEC.

[35]The Kennedy-round is a round of negotiations within the framework of The General Agreement on Tariffs and Trade (GATT) taking place between 1964 and 1967.

[36]The similarities between the debate on the need to reduce institutional differences between workers on a payroll and the self-employed in 1973 and 2017 are striking.

[37]These save-as-your-earn schemes, 'spaarloonregelingen', were amended at several occasions in later decades,

certain income groups do not correlate with output but rather reflect policymakers desire for redistribution and to induce certain behaviour. The measures are all classified as exogenous and serve as legitimate observations to estimate the effect of tax on output.

The tax system underwent substantial modernization between 1966 and 1975. The VAT was introduced in 1969, inspired by the French TVA and following the First and Second VAT Directive of the European Commission, replacing the tax cascade system of taxing turnover at every stage in the supply chain. The new VAT system effectively led to an *increase* of tax revenue for the government of structurally 110 million guilders (<0.1% GDP), but a transition period aditionally benefited the treasury, leading to gains in tax revenue in 1969 of another 238 million guilders (0.2% of GDP). In the same year, the government introduced a new law on corporate tax. Structural changes were smaller yet also entailed an increase in the de facto average tax rate: government revenues were raised by another 55 million (<0.1% GDP).

As in the earliest decade of the dataset, the majority of tax measures between 1966 and 1975 are of a structural rather than countercyclical nature. Between 1966 and 1970 I do not encounter counctercyclical measures.

After the termination of the Bretton Woods system in 1970, the economy of the United States weakened and suffered from high inflation, directly affecting the Netherlands through trade. Inflation surpassed surpassed 8% in 1972, but the economy performed relatively well nevertheless. This took its political toll: the five-party Cabinet was divided on the need to fight inflation and the budgetary route to take; austerity or expansionary budget policy. The Cabinet fell in the same year over these policy issues. Global inflationary pressure was exacerbated by the OAPEC[38] oil embargo. The 1973 oil crisis severely weakened the Dutch economy, and inflation surged to 10%. Three factors were of immediate concern: a high budget deficit, a strong wage-price-spiral[39] and an output decline. Perhaps the complexity and size of the economic situation led policymakers to pursue a rather inconsistent tax policy. In 1974, in the midst of the oil crisis, the cabinet raised several taxes amounting to 540 million (0.3% of GDP), to only a year later provide a tax reduction of 1000 million (0.6% of GDP). The Budget Memorandum labels this last set of tax measure 'a stimulus package'. There is therefore not much doubt that this policy responded to the economic conditions, and so I classify the package as endogenous. De Kam (1988) further describes how in this period government policy contributed to both high collective tax pressure and a worsening budget deficit.

On this large tax cut Wim Duisenberg, the minister of finance, writes in the

and phased from 2012 onwards.

[38] Abbreviation for Organization of Arab Petroleum Exporting Countries.

[39] The wage-price spiral is a cause-and-effect relationship between higher prices and wages. When confronted with higher prices, workers demand higher wages, in effect leading to higher prices. Many European countries faced this phenomenon in the early 1970s.

introduction of BM 1975: "The Netherlands, with its strong external position, partly but not only because of our natural gas ownership, belongs to the few countries in the industrialised world that can and must afford an expansive compensatory spending policy. The domestic situation, with an ever-increasing and stubborn level of unemployment, also forces it to do so". It is worth noting that I do not encounter many other such strong calls for expansionary fiscal policy in the policy documents.

In the meantime, as De Kam (1988) points out, the costs of the welfare state continued to rise sharply to a point where they were deemed unsustainable, but other mostly non-tax revenues such as income from the exploitation of natural gas masked budgetary derailing.

1976-1985: A top-heavy welfare state: the high collective tax burden as an obstacle to economic growth

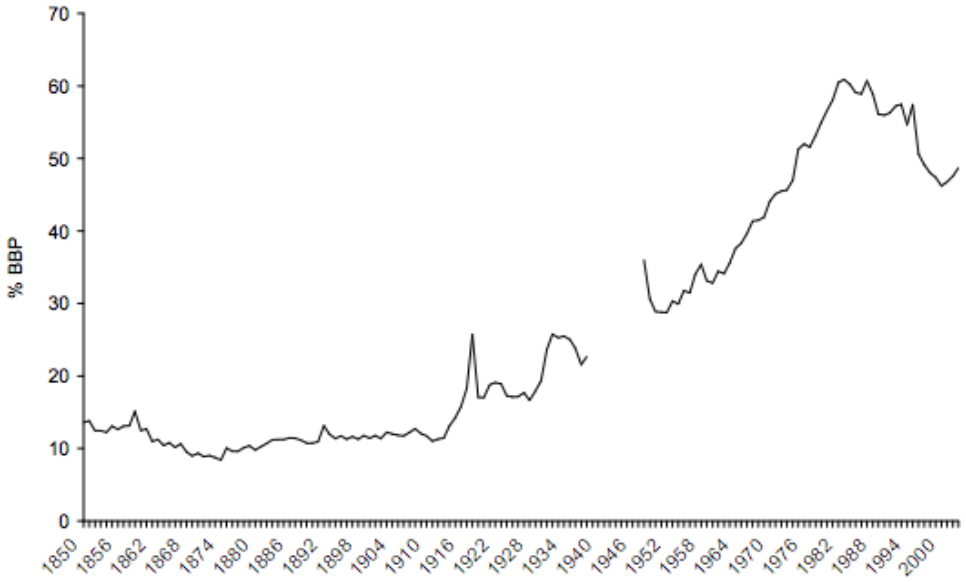
"For many years, the financial deficit and the collective tax burden have undergone a sharp deterioration, so that both have now achieved an unprecedented level," Minister of Finance Onno Ruding opens BM 1984 without much ambiguity. "Non-insignificant policy interventions carried out in previous years and in 1983 were not able to provide sufficient counterbalance against the disappointing economic situation and the tendency towards uncontrollable collective expenditure... In order to achieve the goals of the deficit, and taking into account the development of collective burden pressure in 1984", he continued, "the Cabinet considers tax cuts in the fiscal environment up to an amount of 2955 million (2.5% of GDP) on a cash basis unavoidable."

The 1983 Cabinet faced a build-up of collective expenditures that had been rising since 1952, but which had taken a sharp turn for the worse since 1971 (up from 45% to 59.8%). Cabinet-Van Agt I and II attempted to curb spending between 1977 and 1982 but failed: Minister of Finance Frans Andriessen proposed a comprehensive set of austerity measures in 1980 but did not gather enough support in parliament, causing him to resign (Bovens, Brandsma, Thesingh, & Wever, 2010).

Collective expenditure was therefore allowed to rise further and reached an all-time high of over 60% of GDP in 1983. Compared to 1955, collective expenditure had more than doubled. Figure 6 gives a historic account of the collective expenditure from 1850 to 2008, with clearly the sharpest increase in the postwar years.

There are multiple drivers of the sharp increase in collective expenditure over the years. CPB (2006) cites a changing demography, labour force participation, productivity increases (in both the private and collective sector), increase in social benefits and a changing view on tasks and instruments of the government. Over this period the average tax rate had risen substantially along with expenditure, but not proportionally as non-tax

Figure 6. : Historic overview of collective expenditure 1800-2008 (CPB, 2006)



revenue gained importance as a source of government income (De Kam, 1988). At any rate, the Cabinet in 1983 did not perceive yet another increase of tax rates to be a viable option, and stressed the need to both reduce its expenditure (mainly spending cuts) as well as its tax income (effectively reducing collective tax pressure).

The year 1982 served as a turning point for the uncontrollable and undressed rise of collective expenditure. The cabinet under Prime Minister Lubbers proposed a package totalling 13 billion. With 2.5% of GDP, this makes it far and away the largest package in the year considered. Measures worth 3 billion took effect in 1982 and 10 billion in 1983. These tax changes occurred at a time when the economy was in the deepest recession since the 1950s and unemployment rose to 7.9% in 1982 (up from 0.7% in 1971).[40] The economy is clearly in a bad state, but I consider the tax changes in 1982 and 1983 as exogenous nevertheless: the sources clearly mention that the packages are part of the effort to reduce collective expenditure. Moreover, in 1984, the government puts austerity measures in direct relation with the budget deficit, and less with the current economic situation. I therefore classify this package amounting to nearly 3 billion guilders as exogenous.

Many substantial measures in 1984 and 1985 were aimed at reducing the tax pressure on firms and households, sometimes to promote economic growth in the short term, at other times to alleviate the tax pressure resulting from a long-term increase

[40]Source: Statistics Netherlands.

of average tax pressure. In 1984, the Cabinet proposes a package aimed at stimulating economic activity of 700 million (0.2% of GDP). Finance Minister Ruding writes in the Budget Memorandum of 1984: "The government attaches great importance to reducing the direct burden on firms as a tool for promoting market recovery. It is desirable to improve the position of business by providing a direct and substantial financial impetus to achieve the intended effect in the short and medium term." As the objective is to raise output in the short run, this package falls in the endogenous category.

I find evidence that several measures in 1985 serve to alleviate the tax burden in the longer-term. Good examples are the expansion of the exemption for entrepreneurs (mutation of -15 million in tax revenue, <0.01% GDP), expansion of the dividend exemption (-35 million, <0.01% GDP) and the increase of the tax credit for self-employed (-15 million, <0.01% GDP). Given the long-term consideration, these tax changes are all exogenous. Another exogenous example is the introduction of the "two-earner-law" in 1985. The law entailed a revision of the structure of payroll and income taxes in 1985 and was a response to the criticism of differences in taxation of married and unmarried couples (leading to differences in 'tax free sums'). CPB (2016) describes how much of the law addressed the inequality due to the growing labour force participation of women (which had grown steadily since 1960s but accelerated in the 1980s). This is a long-term consideration and therefore serves as a legitimate exogenous tax change.

In 1982, the Wassenaar Agreement between employers' organisations and labour unions marked an important step towards restraining the severe wage growth. Under what was dubbed a 'no-nonsense policy', political parties CDA and VVD introduced sizeable social security cuts and curbed civil service and teacher salaries.

1986-1995: A delicate choice: reducing the budget deficit or curbing expenditure?

In 1986 Prime Minister Lubbers ran for a second term under the slogan "Let Lubbers finish his job", directly appealing to the public for a continuation of the path to more austerity (Visser, 1993). He was re-elected and Onno Ruding remained as minister of finance.

Although the economy had recovered from a deep recession halfway the 1980s, the combination of the high budget deficit and collective expenditure resurfaced as one of the most urgent challenges of the government. In 1986, this became painfully clear when the government was confronted with a sharp drop in its natural gas revenues, partially due to a drastic decline in global energy prices. On this, the minister of finance writes in the Budget Memorandum of 1987: "The most distinctive element of the 1987 budget is the dramatic decline in the natural gas income of the government, by 12.6 billion, from 21.2

to 8.6 billion... [This accounts for] 8% of total government income. Although this problem is now recognized and acknowledged that it makes measures necessary, many clearly do not realize that the problem is so detrimental that it requires drastic measures in 1987...". The measures for this year have been mainly on the spending side, the only tax measure that I found for 1987 is a continuation of increased excise duties on gas revenues to make up some (476 million, <0.01% GDP) of the losses.

In 1988, the government proposed a new set of tax measures to once again reduce collective tax pressure. Ruding writes in the Budget Memorandum of 1988 in relation to a reduction in wage and income tax rates: "The wish for further burden reduction underlines the importance of controlling and reducing the share of collective expenditure in national income". The package is substantial and reduces government tax revenue by 1300 million (0.3% of GDP). All tax changes in the 1970s through 1990s that are reliably related to the reduction of this inherited collective burden are classified as exogenous (they do not respond to current economic circumstances).

Reducing collective tax pressure was clearly an effort spanning decades. In 1994, Minister of Finance Gerrit Zalm writes: "The government contributes to making public finance healthy again. In addition to a further reduction of finance deficit, a considerable amount has been deducted for collective burden relief". As a compromise to the left-wing coalition partner PvdA, he lowers the rate of the first range in the income tax, relieving the tax burden with 1725 million (0.3% of GDP). The government also lowers corporate tax from 40% to 35% per 1 October 1988, leading to a loss in government revenue of 420 million (<0.01% GDP), in order to achieve "better alignment with rates abroad". These structural measures also fall in the exogenous category.

The government lowers the VAT rate with 1.5 percentage point in 1989, leading to a decrease in government revenue of 2040 million (0.4% of GDP). Multiple motivations are cited in the Budget Memorandum: a reduction of consumer prices is supposed to reduce inflationary pressure, purchasing power is retained and a 'stabilisation of collective pressure' is within reach. In addition, the government believes that lower VAT would lead to higher employment in the formal sector, reducing the number of people working in the informal economy. Weighing these motivations and given the fact that economic growth was not substantially away from normal, it is likely that the motivation of this tax change was exogenous.

Another reduction in the VAT per 1 October 1992 classifies as an endogenous tax change: its motivation was to fight inflation. Although inflation does not always have to be related to output or other factors that influence output, this measure comes at a time when economic growth is slowing down. To prevent further inflation and even more economic decline, the VAT is lowered on quite short notice. There is ample evidence that

this measure is countercyclical and should therefore be placed in the endogenous category.

In the years between 1989 and 1995 I find many (mostly smaller) tax changes to reduce the government deficit. Examples include the increase of several excise duties, moving products and services to the higher VAT rate and a reduction in fiscal benefits for child care. It occurs that the government in this decade has become increasingly creative in finding ways to cover for spending increases and to contribute to a reduction of the budget deficit: some 15 out of 74 tax changes cause a change in tax liabilities under 100 million (<0.01% GDP). Many of these smaller changes classify as exogenous.

In the 1990s, taxation was increasingly employed as an instrument to address environmental concerns. The increase of diesel with 6.7 cents (in 1990) is used primarily "for environmental policy", on which the government plans to spend nearly 1400 million guilders (0.3% of GDP). In 1994, the government introduces a minor consumption tax on an "environmental basis" (130 million revenue increase, or <0.01% GDP). Yet, it is significant as it marks the start of a process of greening the tax system in the late 1990s and early 2000s. Rates on waste and groundwater are additionally increased. On this, Minister of Finance Wim Kok writes in the Budget Memorandum of 1994: "In order to reduce CO emissions in accordance with policy objectives, the use of price instruments is desirable. With other environmental concerns, the use of price instruments can help to achieve the goals. With the environmental impact on consumption - which is expanded with the foundations of waste and groundwater - a step has already been taken in this direction." Tax changes that reflect environmental concerns are often exogenous measures, unless they are taken to finance a spending increase, as is the case for the diesel price increase in 1990.

The years between 1986 and 1995 also entailed the restructuring of several tax schemes, with the income tax 'Oort-operation' (after the chairman of the committee tasked with writing a proposal on a revised tax) the most substantial one. CPB (2016) cites three main motives for a large income tax revision in 1990: a simplification (the last revision was in 1964), a more competitive income tax (the top rate was 72%) and a more equal treatment of married and unmarried couples in light of the increased labour force participation of women. A year later, in 1991, the excise duties scheme underwent a major revision. The Oort-operation meant a loss of government revenue of structurally 4120 million (0.7% of GDP), the excise revision 30 million (<0.01% GDP), and are due to their structural nature both classified as exogenous.

During the term of Lubbers III (1989-1994) it became more apparent that the labour market exhibited severe and structural problems: employment and the number of social benefits remain high and even increase in 1993 and 1994 (Gruijters & Smulders, 1998). Additionally, the global economic situation worsens. There are also concerns

about a large number of discouraged workers and employees on disability schemes (hidden unemployment). I find several policies that aim to fight these issues. In 1994, the share of income that is untaxed ('arbeidskostenforfait') is increased, with the motivation (in the BM of the same year) that by doing so, "work becomes financially attractive, and [it will bring about] an improved connection between supply and demand in the lower segment of the labour market." Also, in 1995, the government lowers wage costs by increasing the premium free share of income ('franchise'), and writes in the BM of the same year: "The international recession has compounded the structural problems in the labour market. This implies that the problem of the shortage of jobs is addressed in priority. Especially employment growth in the lowest segment of the labour market is important because of the large number of inactive workers." Supply side policies to increase employment (more part-time jobs, stimulating labour force participation of women and migrant workers) in the long run are common during these years. I classify these long-run, structural considerations for an improved labour market and to alleviate distortions on labour supply as exogenous."

In 1994 Minister of Finance Gerrit Zalm made a fundamental change to the way in which a cabinet could design its tax and budget policy. He lifted the dependency of spending on revenue in a cabinet's term (known as the 'Zalmnorm'). Political parties at the start of a cabinet's term bound themselves to a spending ceiling under assumptions of developments in the economy. Under this new norm, cabinets do not have to impose austerity measures when revenues are lower than expected, but neither can they increase spending with windfall revenues. The policy between 1983 and 1993 had been on several occasions procyclical as lower revenues plausibly coincided with lower economic growth. The Zalmnorm promised more public finance stability (Jacobs, 2007).

1996-2005: The tax system as a powerful tool to address specific issues

When I study the set of tax measures from 1996 and later years, it becomes clear that the tax system over the years had grown to much more than a source of government income. Greening of the tax system took off in the second half of the 1990s: the need to protect the environment and measures to contribute to environmental sustainability even received attention in several introductory texts of the Budget Memorandum. Tax rate increases on motorised vehicles (since the 1960s a constant source of government income and increased every few years) were put in direct relation with the effort 'to preserve the environment' (BM 1998).

I find evidence of the classic merit and demerit good distinction in the tax system.[41]

[41] Merit goods are defined as goods of which the government wants to stimulate its consumption and production, usually through subsidies. Demerit goods are those of which the government wants to limit the production and consumption, often by levying excises. Merit goods are said to have exceeding marginal social benefits over marginal private benefits, and demerit goods have lower marginal social benefits than marginal private benefits.

Towards the end of the 1990s, I see several measures relating to favourable fiscal treatment of activities related to schooling and research and development. Labour market policy does not just seem to focus on inducing the inactive to join the work force (issues relating to the poverty trap grew along with the expansion of the welfare state), but more on human capital. On this, in BM 1998: "Investments in human capital contribute significantly to the knowledge intensity and competitiveness of business. In addition, these investments have a significant impact on employment and the level and extent of the participation of older workers. Through continuous education, the employee retains his labour market and resilience." Government subsidisation in this sense takes place through fiscal deductibles for both business and individuals: fiscal facilities made it possible to deduct costs related to, for example, RD from tax statements.[42] Most of these changes are easily identified as aiming to structurally raise output in the long run and are therefore exogenous to output. Many of the incentive measures are minor (with an impact of 15 to 20 million on government revenue, both <0.01% GDP) and only serve as a tax change that is likely to have an output effect when taken together in ackages.

In the years 2000 to 2005, most of the tax changes were presented in packages, making it easy to identify them. In 2005, for example, tax measures (42 in total) were labelled either 'labour market and income policy', 'economic infrastructure', 'environment and mobility' and 'other measures'. This does not mean that the categorisation of these measures is sufficient to correctly identify them as exogenous or endogenous: the motivation is still key, though tax measures under the same label usually share the policy objective. For example, in relation to measures contributing to improving economic infrastructure, the Cabinet writes in the BM 2005 that "improving Dutch competitiveness is one of the cornerstones of this cabinet". There is little doubt that the tax changes aim to raise output in the long run and should therefore be considered exogenous.

In 2001 the income tax was substantially modernised. The former laws relating to the income and wage tax dated from 1964, and there was a strong desire to "accommodate the tax regime to modern day needs" (BM 2001). The new 'IB 2001' was characterised by a wider tax base and a reduction of the rates in the wage and income tax. Minister of Finance Zalm writes in the introduction of the Budget Memorandum: "Resolving old issues also creates new problems. The tension in the labour market rises and the risk of overheating is present. We need to be wise with the introduction of new measures to keep these issues in check". Given the peak economy, the government might have favoured a different timing to introduce the new tax, which effectively alleviated a tax burden with 6145 million (0.6% of GDP). A set of laws introduced along with the new

[42]Other good examples of tax policy aimed at promoting certain behaviour or activity are the extension of the tax reduction for education (1998), increase in the tax credit for the self-employed (1999) and the 'enterprise package for the 21st century' (2000).

income tax limited effects to income and redistribution. This additional package dampens macroeconomic effects somewhat, as it led to an increase in revenue of 1569 million (0.2% of GDP).

The year 2003 presents a good example of how restructuring the tax system may coincide with the desire to reduce a budget deficit in the short run. Minister of Finance Hans Hoogervorst proposed some 24 measures to mainly address issues of complexity and to terminate costly and non-effective taxes.[43] These measures together amounted to an increase in government tax revenue of 2700 million euro (0.5% of GDP). Without these measures, the minister notes, "there would be an EMU deficit of 1.2% GDP in 2003, which would diminish the outlook that the government is able to repay its debt within one generation".

In 2003, at the start of the second cabinet under Prime Minister Jan Peter Balkenende, economic growth was at the lowest since 1982, but labour market shortages and a relatively stable budget deficit make politicians reluctant to act. Later that year, economic growth had halted and the budget deficit increased swiftly. Tax policy for 2004 was mixed and consisted of measures that both raised and decreased government tax revenue. This was partially due to the Strategic Deal that was part of the Coalition Agreement, and effectively entailed several improvements to the tax system, raising revenue by 300 million (<0.01% GDP). The government also increased multiple tax credits, reducing the gap of those receiving benefits and the working, which was especially persistent for the elderly.[44] The problem thus accumulated over a longer time and could therefore be considered as exogenous. An additional set of tax measures ('aanvullende besluitvorming') reduced the budget deficit by 931 million (0.2% of GDP). Although battling costs of ageing are cited as one of the policy objectives, the tax changes mainly responded to the need of meeting the deficit obligations under the EMU. This set of tax measures is therefore classified as endogenous.

2006-2015: Structural measures amid a worsening public deficit

The caretaker cabinet under Prime Minister Balkenende (Balkenende III) took office in 2006 after the cabinet crisis of its predecessor and subsequent resignation. When the cabinet was installed the economy had almost completely recovered. This had in turn reduced the government deficit to about 1.8%.[45] Much of its tasks in 2006 was a continuation of economic policy of the previous cabinet: this translated in the implementation of

[43]Good examples are cancelling facilities related to savings salary (where a part of gross income could be used for savings), terminating the SPAK (a measure to equalise wage costs between young and older workers, and abolishing tax credits related to life annuity.

[44]Halfway the 1990s the labour force participation of people between 55 and 65 years old hit a low of 25%, source: Statistics Netherlands.

[45]Source: Statistics Netherlands.

new health care law, several packages to address issues related to economic infrastructure, environment and mobility, and the long-awaited reduction of the corporate tax rate. The new health care law entailed a change from a dual system with private and state funded insurance to private insurance, raising effective monthly premiums for lower income earners. To offset this income effect, the implementation was accompanied by a compensation package amounting to a revenue decrease of 957 million euro (0.2% of GDP.

Balkenende IV took office in February 2007 and proposed 17 tax changes in the Budget Memorandum 2008. The 8 measures in the package "sustainable living environment" (increase in tax revenue of 1010 million or 0.2% of GDP) and 8 tax changes in "social cohesion" (increase of 1259 million or 0.2% of GDP) classify as exogenous packages: the minister cites "the need to contribute to sustainability and a better living environment" for the first, and "sustainability of public finances in light of an ageing population" for the second set of changes.

As was the case during the 1990s, the government devotes multiple tax changes to increasing the labour supply of specific population groups (for both the intensive and extensive margin). In relation to the social cohesion package in BM 2008, the minister writes: "The adjustments to the tax credits are mainly aimed at promoting employment and the sustainability of public finances in the face of ageing. In particular, the employment rate of specific groups such as women, the elderly and people with lower incomes is very important. These people are not only stimulated to participate in the workforce, but are also stimulated to work longer hours. For example, the income-dependent reduction of the labour discount and the additional combination discount make an important contribution to the participation."

The Cabinet presented the Budget Memorandum for 2009 on Tuesday 16 September 2008 in The Hague, just one day after Lehman Brothers filed for Chapter 11 bankruptcy protection in New York City. It is hardly surprising, though painful nevertheless, that the policy proposed in the government's main policy document is oblivious to the unprecedented economic events that would draw the global economy in a recession not witnessed since the Great Depression in the 1930s.

Not to say that there were no hints of economic decline picked up by policymakers in the course of 2008. Finance Minister Wouter Bos, in the introduction of BM 2008, states: "A year [2008] characterised by declining economic growth, rising inflation and uncertain outlook for the global economy. In this context, the cabinet has chosen to make a strong commitment to further strengthening its well-founded position as the Netherlands: purchasing power stimulus and debt relief, promoting participation, administrative relief, innovation and an increasing financial surplus are dominant aspects of this." However, the Tax Plan for 2008 is rather small, with just a few million in revenue mutations.

Income policy accounts for the most substantial changes, with a reduction of 372 million in government revenues (<0.01% GDP). Changes mainly consist of an extension for parental leave and a bonus for continuing paid work beyond the age of 62 ('doorwerkbonus'). Parts of the revenue mutations are offset by a redistribution package with a revenue increase of 219 million (<0.01% GDP), consisting of rate changes and tax credits in the income tax.

Income policy accounts for the most substantial changes, with a reduction of 372 in 'average tax burden' for an extended parental leave and a bonus for continuing paid work beyond the age of 62 ('doorwerkbonus'), which is partly offset by a redistribution package of an revenue increase of 219 million (<0.01% GDP), consisting of some rates changes and tax credits in the income tax. The long-term considerations (ageing population, labour force participation and public finance) make it possible to identify all tax policy for 2008 as exogenous.

In March 2009 the Cabinet decided on a large fiscal stimulus[46] as a response to the deep economic recession, amounting to nearly 3000 million euro (0.5% of GDP, both spending increases and tax decreases). The Prime Minister writes about a "targeted stimulation for 2009 and 2010 in particular, with an emphasis on short-term employment... There will be a stimulus of about 3 billion euro in 2009 and about 3 billion euro in 2010 for the labour market (especially to fight youth unemployment)". Amid these direct measures to increase employment in the short-run (an obvious endogenous measure), the package also entailed multiple measures of a structural nature, on "education and knowledge, a sustainable economy, infrastructure and housing, and improved credit loans to business". Tax policy is mixed. As the measures come in one package and nearly all take effect on 1 July 2009, it is hard to effectively disentangle output responses of endogenous and exogenous measures. I therefore choose to classify the full package as endogenous.

Unlike 2009, the year 2010 presents only tax measures that contribute to sustainability of the economy and public finance in the long-run. "Doing nothing is not an option", Minister of Finance Bos writes in the introduction of BM 2010, "difficult decisions will be inevitable. They do not have to be made in one year... They must be preceded by fundamental political reconsiderations across the width of public policy and a social debate about how we can, though with fewer resources, make the Netherlands cleaner in 2020, smarter, stronger, and solid. This BM marks the start of this debate." I find no countercyclical tax measures in this year (like the spring stimulus package). Rather, tax measures aim to improve economic infrastructure and address long-run issues. The Cabinet presents 7 measures on entrepreneurship (alleviating tax burden by 803 million or 0.13% of GDP), which are the result of elaborate consultations with business and address technical issues

[46]The parliamentary 'Workgroup Gerritse' presented in February of 2009 a report on "possible policy responses to the turbulent macroeconomic developments of the previous months".

related to corporate taxation. The consultations started before the economic downturn in 2008, which makes it safe to classify them as exogenous. Other measures in that year include 9 greening tax changes, a fiscal simplification law and several changes to VAT levying on paint jobs and cleaners following ECOFIN decisions.

At the presentation of BM 2011, in September 2010, the Minister of Finance Jan Kees de Jager states that "the crisis is not fully resolved, yet the new economic developments give reasons for hope". He proposes two packages: a first under the flag of innovative entrepreneurship and a second labelled purchasing power package ('koopkrachtpakket'). I label them as endogenous for three reasons: first, they both mean a substantial alleviation of tax burden (totalling 2000 million or 0.3% of GDP), second, their implementation is put in direct relation to the economic crisis ("if recovering after a historic severe crisis, companies need financial space to revitalise their household accounts", BM 2011) and third, the purchasing power package seems to stand on its own, there are no other measures that would have effected redistribution substantially.

Policymakers realised that a recession in combination with expansionary fiscal policy would ultimately lead to an increasing government deficit. In the BM 2012, "healthy public finances" was mentioned as the prime priority of economic policy in that year. I find a very mixed tax policy of endogenous and exogenous tax changes. 22 simplification measures (totalling a revenue increase of 222 million, <0.01% GDP), a substantial research and development deductible (decrease of 399 million, <0.01% GDP) and a "vitality package" to increase labour force participation (revenue increase of 608 million, 0.01% GDP) are among the most important exogenous changes. By far the most prominent endogenous measure is the VAT 'high rate' increase with 2 percentage points, from 19% to 21% per 1 October 2012. The yield, structurally more than 4000 million euro 0.6% of GDP, "will be used partly in 2012 and 2013 to improve the EMU balance". This clearly signals the Cabinet's desire to reduce the deficit in the short run. In the third quarter of 2012, when the last chapters of the BM for the succeeding year were completed by a caretaker cabinet, the government deficit had dropped with 3 bp to 4.4 percent.[47]

The largest package in 2012 (revenue increase of 1750 million or 0.3% of GDP) changed the general tax credit and rates on the income and wage tax. "A large part of the 2013 deficit follows from the 2013 budget agreement and is used to get our public finances in order," Minister of Finance Jeroen Dijsselbloem writes when explaining the policy objective of the package. Another revenue increase follows from the assurance tax hike totalling 1222 million (0.2% of GDP), which I do classify as exogenous: the tax change was proposed to financially cover a mobility package ('wet herziening woon-werkverkeer'). The mobility package was first amended and then cancelled, but the assurance tax increase

[47]Source: Statistics Netherlands.

went ahead as planned, making it a legitimate exogenous observation.

Tax policy in 2014 shows well how after a year of substantial deficit reducing tax policy the tide can change: the Cabinet introduced an income policy package, effectively alleviating the tax burden on labour with 1800 million euro (0.3% of GDP). The extensive deliberations[48] between coalition partners VVD (right-wing liberal) and PvdA (labour) justify an exogenous classification. This package was accompanied by several deficit reducing measures, such as changes to the fiscal facilities on severance pay ('stamrecht'), leading to a revenue increase of over 2000 million euro (0.3% of GDP). Finally, for the year 2015 I do not find any exogenous measures. The most notable tax change was a change in the percentage of the first income category, reducing the tax burden by 539 million in 2015 (<0.01% GDP). This change was motivated by a lagging economic recovery and unemployment rates higher than deemed acceptable by policymakers.

[48]See the narrative document for a more elaborate explanation of how the set of tax decreases came about.

Appendix B: The effect of tax on output with dummy variable

Table 14—: The effect of tax on output *(dummy for mixed quarters)*

	Measure of Exogenous Tax Changes	Standard error	t	P	68% confidence interval	
L0	-0.82	0.39	-2.08	0.04	-1.59	-0.04
L1	0.31	0.40	0.77	0.44	-0.48	1.10
L2	0.71	0.40	1.75	0.08	-0.09	1.50
L3	-0.34	0.40	-0.85	0.40	-1.14	0.45
L4	0.42	0.42	1.01	0.32	-0.40	1.24
L5	-0.13	0.41	-0.32	0.75	-0.94	0.68
L6	-0.45	0.41	-1.10	0.27	-1.26	0.36
L7	0.12	0.41	0.29	0.77	-0.69	0.92
L8	-0.25	0.41	-0.60	0.55	-1.06	0.56
L9	0.16	0.40	0.40	0.69	-0.62	0.94
L10	0.07	0.40	0.17	0.87	-0.72	0.85
L11	0.08	0.39	0.20	0.84	-0.70	0.86
L12	-0.16	0.38	-0.43	0.67	-0.92	0.59
Constant	0.00	0.00	2.58	0.01	0.00	0.01

Appendix C: The effect of tax on output for single periods

Table 15—: The effect of tax on output between 1960 and 1980

	Measure of Exogenous Tax Changes	Standard error	t	P	68% confidence interval	
L0	-2.20	1.17	-1.88	0.07	-3.38	-1.02
L1	0.33	1.21	0.27	0.79	-0.89	1.55
L2	0.92	1.22	0.75	0.46	-0.31	2.14
L3	-0.88	1.21	-0.73	0.47	-2.10	0.33
L4	-0.32	1.27	-0.25	0.80	-1.60	0.95
L5	-0.37	1.24	-0.30	0.77	-1.62	0.88
L6	-2.78	1.26	-2.20	0.03	-4.04	-1.51
L7	0.01	1.27	0.01	1.00	-1.27	1.29
L8	0.31	1.28	0.24	0.81	-0.98	1.59
L9	-1.33	1.19	-1.11	0.27	-2.53	-0.13
L10	0.67	1.19	0.56	0.58	-0.52	1.87
L11	-0.38	1.18	-0.32	0.75	-1.57	0.80
L12	-1.31	1.13	-1.16	0.25	-2.45	-0.17
Constant	0.01	0.01	1.80	0.08	0.01	0.02

Table 16—: The effect of tax on output between 1981 and 2000

	Measure of Exogenous Tax Changes	Standard error	t	P	68% confidence interval	
L0	-1.30	0.39	-3.32	0.00	-1.69	-0.90
L1	-0.57	0.45	-1.27	0.21	-1.03	-0.12
L2	-0.04	0.44	-0.09	0.93	-0.48	0.40
L3	-0.07	0.42	-0.18	0.86	-0.50	0.35
L4	1.09	0.30	3.69	0.00	0.79	1.39
L5	-0.35	0.34	-1.02	0.32	-0.69	0.00
L6	0.28	0.33	0.85	0.40	-0.05	0.62
L7	-0.41	0.32	-1.27	0.21	-0.73	-0.09
L8	-0.33	0.33	-0.98	0.33	-0.66	0.01
L9	0.71	0.27	2.59	0.01	0.43	0.99
L10	-0.68	0.30	-2.28	0.03	-0.98	-0.38
L11	0.53	0.31	1.69	0.10	0.22	0.85
L12	0.02	0.30	0.08	0.94	-0.28	0.33
Constant	0.00	0.00	0.67	0.50	0.00	0.01

Table 17—: The effect of tax on output between 2000 and 2015

	Measure of Exogenous Tax Changes	Standard error	t	P	68% confidence interval	
L0	-0.54	1.65	-0.33	0.75	-2.22	1.14
L1	0.12	1.62	0.07	0.94	-1.53	1.77
L2	0.40	1.60	0.25	0.81	-1.23	2.03
L3	-0.35	1.56	-0.22	0.83	-1.93	1.23
L4	-1.83	1.49	-1.23	0.23	-3.35	-0.32
L5	1.17	1.38	0.85	0.40	-0.23	2.58
L6	0.41	1.43	0.28	0.78	-1.05	1.86
L7	0.17	1.43	0.12	0.91	-1.29	1.62
L8	-0.23	0.96	-0.24	0.81	-1.20	0.75
L9	0.59	0.96	0.61	0.55	-0.40	1.57
L10	0.80	0.97	0.82	0.42	-0.19	1.78
L11	-0.41	0.99	-0.41	0.68	-1.42	0.60
L12	-1.27	0.93	-1.37	0.18	-2.21	-0.33
Constant	0.00	0.00	0.96	0.35	0.00	0.00

Appendix D: Controlling for outliers

Table 18—: The effect of tax on output without upper outliers

	Measure of Exogenous Tax Changes	Standard error	t	P	68% confidence interval	
L0	-1.29	0.64	-2.02	0.05	-1.92	-0.65
L1	-0.13	0.64	-0.20	0.84	-0.76	0.51
L2	0.33	0.64	0.51	0.61	-0.31	0.96
L3	0.20	0.64	0.32	0.75	-0.43	0.84
L4	-0.05	0.65	-0.08	0.93	-0.70	0.59
L5	0.58	0.64	0.91	0.37	-0.06	1.22
L6	-0.48	0.64	-0.75	0.45	-1.12	0.16
L7	0.29	0.64	0.45	0.65	-0.35	0.93
L8	-0.26	0.64	-0.40	0.69	-0.90	0.38
L9	-0.12	0.63	-0.19	0.85	-0.75	0.51
L10	0.39	0.63	0.62	0.54	-0.24	1.02
L11	-0.39	0.63	-0.61	0.54	-1.02	0.24
L12	0.49	0.62	0.79	0.43	-0.13	1.10
Constant	0.00	0.00	2.41	0.02	0.00	0.01

Table 19—: The effect of tax on output without downward outliers

	Measure of Exogenous Tax Changes	Standard error	t	P	68% confidence interval	
L0	-0.85	0.39	-2.15	0.03	-1.24	-0.45
L1	0.31	0.40	0.78	0.44	-0.09	0.71
L2	0.78	0.40	1.94	0.05	0.38	1.18
L3	-0.33	0.40	-0.81	0.42	-0.73	0.08
L4	0.36	0.42	0.88	0.38	-0.05	0.78
L5	-0.11	0.41	-0.26	0.79	-0.52	0.30
L6	-0.45	0.41	-1.11	0.27	-0.86	-0.04
L7	0.12	0.41	0.29	0.77	-0.29	0.53
L8	-0.22	0.41	-0.54	0.59	-0.63	0.19
L9	0.12	0.40	0.31	0.76	-0.27	0.52
L10	0.14	0.40	0.36	0.72	-0.25	0.54
L11	0.08	0.39	0.20	0.85	-0.32	0.47
L12	-0.14	0.38	-0.38	0.71	-0.53	0.24
Constant	0.00	0.00	2.37	0.02	0.00	0.01

Table 20—: The effect of tax on output with the full tax measure

	Measure of Exogenous Tax Changes	Standard error	t	P	68% confidence interval	
L0	-0.38	0.37	-1.03	0.31	-0.75	-0.01
L1	0.02	0.37	0.06	0.95	-0.35	0.40
L2	0.58	0.38	1.53	0.13	0.20	0.96
L3	-0.08	0.37	-0.21	0.83	-0.45	0.30
L4	0.02	0.38	0.06	0.95	-0.36	0.40
L5	0.33	0.38	0.86	0.39	-0.05	0.70
L6	-0.35	0.38	-0.92	0.36	-0.73	0.03
L7	0.16	0.37	0.42	0.68	-0.22	0.53
L8	-0.31	0.38	-0.82	0.41	-0.69	0.07
L9	0.08	0.37	0.21	0.83	-0.29	0.45
L10	-0.12	0.37	-0.32	0.75	-0.49	0.25
L11	0.51	0.37	1.40	0.16	0.15	0.88
L12	-0.29	0.36	-0.79	0.43	-0.65	0.07
Constant	0.00	0.00	2.13	0.03	0.00	0.01

Appendix E: Other control variables

Table 21—: The effect of tax on output *(government spending control)*

	Measure of Exogenous Tax Changes	Standard error	t	P	68% confidence interval	
L0	-0.77	0.41	-1.88	0.06	-1.18	-0.36
L1	0.56	0.42	1.34	0.18	0.14	0.98
L2	0.85	0.42	2.01	0.05	0.43	1.27
L3	-0.08	0.43	-0.18	0.85	-0.50	0.35
L4	0.38	0.44	0.86	0.39	-0.06	0.82
L5	-0.17	0.44	-0.38	0.70	-0.61	0.27
L6	-0.41	0.44	-0.94	0.35	-0.85	0.03
L7	0.18	0.44	0.41	0.69	-0.26	0.61
L8	-0.13	0.41	-0.33	0.75	-0.54	0.28
L9	0.21	0.40	0.52	0.61	-0.19	0.60
L10	0.22	0.40	0.55	0.58	-0.18	0.62
L11	0.19	0.40	0.48	0.63	-0.20	0.59
L12	0.04	0.39	0.09	0.93	-0.35	0.43
Constant	0.03	0.01	2.37	0.02	0.01	0.06

Table 22—: The effect of tax on output *(policy interest rate control)*

	Measure of Exogenous Tax Changes	Standard error	t	P	68% confidence interval	
L0	-0.58	0.34	-1.68	0.09	-1.26	0.10
L1	0.22	0.35	0.62	0.54	-0.47	0.90
L2	0.39	0.35	1.11	0.27	-0.30	1.08
L3	-0.37	0.35	-1.05	0.30	-1.06	0.32
L4	0.29	0.36	0.80	0.43	-0.42	1.00
L5	-0.51	0.36	-1.42	0.16	-1.22	0.20
L6	-0.35	0.36	-0.99	0.32	-1.06	0.35
L7	0.02	0.36	0.06	0.95	-0.69	0.73
L8	-0.19	0.36	-0.54	0.59	-0.91	0.52
L9	0.18	0.35	0.51	0.61	-0.51	0.86
L10	-0.32	0.35	-0.93	0.35	-1.00	0.36
L11	0.07	0.35	0.21	0.83	-0.61	0.76
L12	-0.08	0.34	-0.23	0.82	-0.75	0.60
Constant	0.01	0.00	1.77	0.08	0.00	0.02

Table 23—: The effect of tax on output *(consumer price index control)*

	Measure of Exogenous Tax Changes	Standard error	t	P	68% confidence interval	
L0	-0.87	0.36	-2.43	0.02	-1.58	-0.16
L1	0.36	0.36	0.98	0.33	-0.36	1.08
L2	0.57	0.36	1.57	0.12	-0.15	1.29
L3	-0.51	0.36	-1.40	0.16	-1.22	0.21
L4	0.10	0.37	0.26	0.80	-0.64	0.83
L5	-0.39	0.37	-1.05	0.29	-1.12	0.34
L6	-0.42	0.38	-1.12	0.26	-1.17	0.32
L7	0.20	0.38	0.52	0.61	-0.55	0.94
L8	-0.10	0.38	-0.26	0.80	-0.85	0.66
L9	0.24	0.37	0.66	0.51	-0.49	0.97
L10	-0.25	0.36	-0.69	0.49	-0.97	0.47
L11	0.25	0.36	0.70	0.49	-0.46	0.96
L12	-0.69	0.35	-1.96	0.05	-1.39	0.01
Constant	0.00	0.00	1.63	0.11	0.00	0.01

Table 24—: The effect of tax on output *(relative price of oil control)*

	Measure of Exogenous Tax Changes	Standard error	t	P	68% confidence interval	
L0	-0.96	0.44	-2.21	0.03	-1.82	-0.10
L1	0.27	0.44	0.60	0.55	-0.61	1.14
L2	0.76	0.45	1.71	0.09	-0.12	1.65
L3	-0.23	0.45	-0.52	0.60	-1.12	0.65
L4	0.50	0.45	1.11	0.27	-0.39	1.39
L5	-0.07	0.45	-0.17	0.87	-0.95	0.81
L6	-0.44	0.45	-0.99	0.32	-1.32	0.44
L7	0.00	0.45	0.00	1.00	-0.88	0.88
L8	-0.31	0.44	-0.71	0.48	-1.19	0.56
L9	0.06	0.42	0.14	0.89	-0.78	0.90
L10	0.12	0.42	0.28	0.78	-0.71	0.95
L11	0.11	0.42	0.25	0.80	-0.73	0.94
L12	-0.04	0.41	-0.09	0.93	-0.85	0.78
Constant	0.01	0.00	1.77	0.08	0.00	0.02

Appendix F: Implementation lags

Table 25—: Implementation lag of two quarters

	Measure of Exogenous Tax Changes	Standard error	t	P	68% confidence interval	
L0	0.87	0.40	2.18	0.03	0.47	1.27
L1	-0.28	0.40	-0.70	0.49	-0.69	0.12
L2	0.12	0.40	0.31	0.76	-0.28	0.53
L3	-0.01	0.40	-0.01	0.99	-0.41	0.39
L4	-0.42	0.42	-1.02	0.31	-0.84	-0.01
L5	0.11	0.41	0.26	0.80	-0.30	0.52
L6	-0.10	0.41	-0.25	0.80	-0.52	0.31
L7	0.08	0.41	0.19	0.85	-0.33	0.49
L8	0.19	0.41	0.47	0.64	-0.22	0.60
L9	0.05	0.40	0.12	0.91	-0.35	0.44
L10	-0.14	0.39	-0.35	0.73	-0.53	0.25
L11	0.00	0.39	0.00	1.00	-0.39	0.39
L12	-0.01	0.39	-0.01	0.99	-0.39	0.38
Constant	0.00	0.00	2.19	0.03	0.00	0.01

Table 26—: Implementation lag of four quarters

	Measure of Exogenous Tax Changes	Standard error	t	P	68% confidence interval	
L0	0.04	0.40	0.10	0.92	-0.36	0.44
L1	-0.03	0.40	-0.06	0.95	-0.42	0.37
L2	-0.20	0.40	-0.50	0.62	-0.60	0.20
L3	-0.01	0.40	-0.03	0.98	-0.41	0.39
L4	-0.22	0.41	-0.54	0.59	-0.64	0.19
L5	0.11	0.41	0.27	0.79	-0.30	0.52
L6	0.01	0.41	0.03	0.97	-0.39	0.42
L7	0.11	0.41	0.27	0.79	-0.30	0.52
L8	0.02	0.40	0.04	0.97	-0.39	0.42
L9	-0.01	0.39	-0.03	0.97	-0.40	0.37
L10	-0.04	0.39	-0.09	0.93	-0.42	0.35
L11	-0.15	0.39	-0.40	0.69	-0.54	0.23
L12	-0.72	0.38	-1.89	0.06	-1.11	-0.34
Constant	0.00	0.00	2.62	0.01	0.00	0.01

Table 27—: Implementation lag of six quarters

	Measure of Exogenous Tax Changes	Standard error	t	P	68% confidence interval	
L0	-0.21	0.39	-0.54	0.59	-0.60	0.18
L1	0.00	0.40	-0.01	1.00	-0.40	0.39
L2	-0.22	0.40	-0.55	0.58	-0.62	0.18
L3	0.10	0.40	0.25	0.80	-0.29	0.50
L4	-0.02	0.41	-0.04	0.97	-0.42	0.39
L5	0.12	0.41	0.30	0.76	-0.28	0.53
L6	0.00	0.40	-0.01	0.99	-0.40	0.40
L7	-0.02	0.40	-0.05	0.96	-0.42	0.38
L8	0.02	0.40	0.06	0.95	-0.37	0.42
L9	-0.16	0.39	-0.41	0.68	-0.54	0.22
L10	-0.74	0.39	-1.92	0.06	-1.13	-0.36
L11	0.03	0.39	0.08	0.93	-0.35	0.42
L12	-0.24	0.38	-0.61	0.54	-0.62	0.15
Constant	0.00	0.00	2.70	0.01	0.00	0.01