

A New Test of Ricardian Equivalence Using the Narrative Record on Tax Changes

Alfred A. Haug*

Abstract

This paper empirically tests the Ricardian equivalence hypothesis with a narrative measure of tax shocks. The present value, at the time of legislation, for tax increases motivated solely by concerns for improving the fiscal health of the government is used for the tests. These tax news represent a switch from debt to tax financing that should have no effects on the economy if Ricardian equivalence holds as a good approximation. Such a tax increase seems to have positive effects on real GDP in the post-1980:IV period. However, this is due to fiscal anticipation as many of the tax increases are implemented with substantial delays and distortionary taxes increase economic activity before taxes go up, which is caused by intertemporal substitution. Therefore, Ricardian equivalence is rejected.

JEL Classification: E62, C51

Keywords: Ricardian equivalence hypothesis; narrative record; exogenous tax changes; government budget deficits.

*Department of Economics, University of Otago, P.O. Box 56, Dunedin 9054, New Zealand. Telephone: +64 3 479 5636; E-mail address: Alfred.Haug@otago.ac.nz. The author thanks, without implicating, David H. Romer, anonymous referees, and participants at the 2015 NZESG in OZ Conference at the National Centre for Econometric Research at Queensland University of Technology for very helpful comments on an earlier version.

1. Introduction

The recent global financial crisis and the ensuing sovereign debt crises in Europe have put fiscal policy center stage. An important benchmark for fiscal policy is the Ricardian equivalence hypothesis. It states that it does not matter for the economy whether government expenditures are financed by current taxation or instead by issuing government bonds (Barro, 1974). Ricardian households are only concerned about the present value of their intertemporal tax liabilities, which are in turn determined by the present value of the stream of government expenditures and currently outstanding government debt. The timing of taxes does not matter.¹ Government bonds are simply seen as postponed taxes that will have to be paid at some future date. Hence, a switch from tax to bond financing has no effects on macroeconomic variables. Barro (1974) argued that Ricardian consumers react, keeping the stream of government spending fixed, to an increase in the government deficit by increasing savings by an equal amount. The tax cut leads to a dollar-for-dollar increase in bond holdings. Therefore, neither output nor interest rates change. The same neutrality argument holds for a switch from bond to tax financing. The Ricardian equivalence hypothesis is based on several restrictive assumptions, such as lump-sum taxes, perfect capital markets where forward-looking households do not face liquidity constraints and can borrow at the same interest rate as the government, altruistic operative bequests that link generations, and no uncertainty about the future tax incidence.

A large number of theoretical and empirical papers have studied Ricardian equivalence. This literature has been surveyed by Seater (1993) and more recently by Ricciuti (2003). On a theoretical level, relaxing the very strict assumptions necessary for Ricardian equivalence can lead to government bonds having either positive or negative net wealth effects for households, as argued by Barro (1974,1989), instead of government bonds not being considered net wealth at all under Ricardian equivalence. For example, Judd (1987) demonstrated this in a theoretical model with distortionary

¹Ricardian equivalence does not imply that all fiscal policy actions have no impact on the real economy: a change in the present value of government spending affects household net wealth and therefore household consumption and real output in the economy. Ricardian equivalence is a Modigliani-Miller theorem for the public sector.

taxes, finite lives, and adjustment costs. Wealth effects could cancel each other or be altogether negligible. Therefore, the fact that tax systems are generally not based on lump-sum taxation does not invalidate Ricardian equivalence, as, for example, transportation costs do not invalidate the assumption of perfect competition in many applications. The issue cannot be settled on theoretical grounds. An answer to whether Ricardian equivalence is a good approximation to reality has to come from empirical studies.² This means that the issue of how good of an approximation Ricardian equivalence is cannot be determine based on theory and empirical tests are called for. However, the two surveys show that the empirical evidence is not conclusive at all.

The goal of this paper is to shed new light on the controversial issue of whether the economy displays Ricardian equivalence features, which is of great relevance in particular to orienting theoretical models. I suggest a new test of Ricardian equivalence that relies on a narrative measure for a surprise switch from bond financing to taxation. Such a switch should have no effects on real GDP and interest rates if Ricardian equivalence is a good approximation. Furthermore, such a switch leaves the intertemporal stream of all tax liabilities unchanged because currently outstanding government bonds have to be paid for with future taxes. A current increase in taxes that reduces currently outstanding bonds implies an offsetting reduction in future taxes so that in present value terms overall tax liabilities are unchanged. I use for this purpose the present value of discretionary *changes* in tax revenue that were motivate by concerns about the inherited government budget deficit and were introduced for reasons unrelated to current macroeconomic fluctuations or government spending. Such a measure has been constructed for United States fiscal policy, based on the narrative record, by Romer and Romer (2010).³ They used historical documents to carefully identify types of legislated tax changes from the motivations provided by

²See also Evans (1991) and Bailey (1993) on further discussions of why Ricardian equivalence may be a reasonable approximation. Furthermore, Evans *et al.* (2012) recently showed that rational expectations are not necessary for Ricardian equivalence to hold and a certain adaptive learning rule instead can produce equivalence.

³Their measure of discretionary tax news shocks has been used by several others, such as Favero and Giavazzi (2012), and Mertens and Ravn (2012). Furthermore, the narrative approach was also used by Ramey and Shapiro (1998) and Ramey (2011) to study the effects of large U.S. military build-ups.

lawmakers at the time of passage, such as the U.S. Congress, Economic Report of the President, Budget, Senate, Social Security, and Treasury Secretary reports, records, and bulletins. They call this type of a fiscal surprise shock (news) an exogenous (as opposed to endogenous) "deficit-driven" tax revenue increase motivated by concerns about inherited budget deficits and not caused by current spending.⁴

Section 2 describes aspects of the data relevant for my study and Section 3 presents results for the full post-WWII sample and for two sub-samples, one prior to and the other post 1980:IV. While Ricardian equivalence apparently seems to hold over the full sample, I show that the regressions over the full sample are not structurally stable. Once I account for a break in 1980:IV, the prior sample has too few observations to allow reliable inference. But the post-1980:IV sample delivers empirical results that show that tax increases that are motivated by dealing with inherited budget deficits have statically significant effects on economic growth. Ricardian equivalence is therefore not supported by the data. This is a result relevant for current fiscal consolidation policies in various countries. The conclusion discusses further implication of my findings.

2. Exogenous Tax News Motivated by Deficit Reduction: The Data

Romer and Romer (2010) constructed from scratch in their seminal paper two "news" series for discretionary tax changes that are unrelated to the business cycle. One is for tax changes motivated by concerns about long-run economic growth. The other is for tax changes motivated by concerns about the government budget deficit that are unrelated to the current business cycle. For each of the two series they construct two measures: one recording the amount of the tax change relative to GDP when it takes effect; the other recording the present value of the tax change when it

⁴A closely related study, using Romer and Romer's methodology, was presented by Cloyne (2013) for the United Kingdom. However, the United Kingdom has no precedents of "fiscal consolidations," dealing with inherited budget deficits that are unrelated to the business cycle, as would be necessary for testing Ricardian equivalence. Comparable data are therefore not available for the United Kingdom.

is legislated, again as a ratio to GDP. The three-year Treasury bond rate is used for discounting future tax revenue changes to the time when a measure was passed into law.⁵

Romer and Romer (2010) used in their empirical analysis several measures of exogenous tax news. One is the sum of the deficit driven tax changes and the tax changes motivated by log-run economic growth, at the time of their implementation. A second and third measure are the two news components in separation: deficit-driven tax changes and long-run tax changes, both at the time of their implementation. A fourth measure is the present value of the stream of both exogenous tax changes at the time when they were legislated, instead of actual values at the time of their implementation as in the first measure. Romer and Romer did not consider separately the present values of the deficit-driven tax changes.⁶ I use in this paper the present values of the nominal deficit-driven tax changes at the time when they are legislated, as percent of nominal GDP.

Most of the deficit-driven (exogenous) tax changes were motivated by concerns over the long-run solvency of the Social Security system. For example, the Social Security Amendments of 1977 and 1983 were major tax increases that did not simultaneously increase benefits. The largest deficit-driven tax increases not related to Social Security were those in the Tax Equity and Fiscal Responsibility Tax Act of 1982 and the Omnibus Budget Reconciliation Acts of 1990 and 1993. There are ten (nine) quarters with deficit-driven exogenous tax changes in the sample period from 1945:I to 2007:IV (1950:I to 2007:IV). This is a considerably larger number than the four quarterly non-zero observations for the war dummy variable in Ramey (2011). Figure 1 depicts the present values of deficit-driven tax changes as percent of GDP. The values range from 0.023% in 1997:III to 1.153% in 1977:IV. In comparison, the non-zero defense build-up dates in Ramey (2011, Table II, pp. 26-27), modelled as dummies with values of either zero or one, represent the following shares of GDP: 63.06% in 1950:III, 0.33% in 1965:I, 6.36% in 1980:I, and 0.94% in 2001:III. Two of

⁵Footnote 25 in Romer and Romer (2010, p. 793) provides more details.

⁶Romer and Romer (2009a) on the other hand studied tax cuts driven by long-run economic considerations in order to test the hypothesis that such tax reductions lead to decreases in future government spending.

these events show much larger portions of GDP than values in my sample, though not the remaining ones. In addition, none of the deficit-driven dates overlap with the dates for the present values of long-run exogenous tax changes of Romer and Romer (2010).

When Ricardian equivalence does not hold, the permanent income hypothesis predicts that consumers react to news about current or future tax increases (decreases) by immediately reducing (increasing) consumption. This initial reaction could then possibly lead to additional dynamics. However, there should be no additional reaction associated with the time when tax liabilities actually take effect if they are not implemented at the time of being passed into law. The correct approach under the permanent income hypothesis for dealing with tax news is therefore to use the present value of tax news at the time when they are announced and to relate output movements to current and lagged values of a measure of news about tax changes. This scenario holds with lump-sum taxes. On the other hand, in the presence of distortionary taxes when Ricardian equivalence does not hold, intertemporal substitution predicts an increase in consumption (or investment, depending on the type of tax used) when the implementation of a tax increase is delayed. In other words, spending is moved forward in time to take advantage of lower taxes before the tax increase takes effect. Leeper *et al.* (2008) referred to this as fiscal foresight or anticipation. In this case, the model needs to account for both the date when the tax change is announced (legislated) and the dates when it is implemented.

When Ricardian equivalence holds, households incorporate the intertemporal government budget constraint into their permanent income. Ricardian consumers are forward looking. The tax news measure relevant to them is the present value of what Romer and Romer (2010) call the exogenous "deficit-driven tax changes." This measure is constructed in such a way that it is orthogonal to all other information available in a given period.⁷ It is therefore ideally suited for testing Ricardian equivalence. At the time when the Romer and Romer (2010) tax shocks are announced (not when they are implemented), there are no associated changes to government

⁷See Romer and Romer (2010), Favero and Giavazzi (2012), and Mertens and Ravn (2012) on empirical support.

spending announced, so that there are no government spending shocks at that time that are associated with the specific tax news.⁸ This is exactly the scenario needed for testing the Ricardian equivalence hypothesis, i.e., keeping the expected stream of government spending fixed when the tax news shock hits. If Ricardian equivalence is a reasonable approximation, news about deficit-driven tax increases represent simply a switch from debt to tax financing and this news variable should therefore have no statistically significant effects on real GDP.

Would Romer and Romer's (2010) tax cuts motivated by concerns about long-run growth be suitable for testing Ricardian equivalence? The long-run tax-change variables were constructed so that they are orthogonal to the current business cycle and current government spending. However, the stated intention of these tax shocks was explicitly to either decrease the size of the government (e.g., the 1981 Reagan tax cuts) or to spur productivity growth and increase efficiency. This subcomponent of tax changes, which stimulates long-run growth, is therefore not valid for testing Ricardian equivalence because there are associated changes in *future* government spending or in *future* income that both affect permanent income. In order to illustrate the Ricardian case, it is useful to look at a simplified version of the intertemporal government budget constraint (see, e.g., Hamilton and Flavin, 1986, Hakkio and Rush, 1991, and Haug, 1996):

$$\sum_{s=1}^{\infty} (1+r)^{-s} RE_{t+s} = B_t + \sum_{s=1}^{\infty} (1+r)^{-s} GE_{t+s} - \lim_{s \rightarrow \infty} (1+r)^{-s} B_{t+s},$$

where RE_t is real tax revenue in period t , including revenue from monetizing deficits by printing money (which is treated as an inflation tax). B_t denotes real government debt, r is the real interest rate, and GE_t stands for real government expenditures on purchases of goods and services plus transfer payments. It is assumed that the government does not follow a Ponzi scheme so that intertemporal government budget

⁸In a regression context this means that such narrative variables are, as long as orthogonality holds, uncorrelated with other included and other omitted regression variables and their regression coefficient estimates are unbiased. In this case, the only effect of omitted variables is to increase the residual variance. It is therefore possible to analyze the effects of exogenous tax changes on economic activity, such as real GDP growth, without specifying an economic model that includes other fiscal and monetary policy variables.

balance holds:

$$\lim_{s \rightarrow \infty} (1 + r)^{-s} B_{t+s} = 0.$$

Ricardian households do not change their consumption and savings plans as long as the present discounted values of expected future tax revenues and of expected future government spending do not change. The future path of government debt is irrelevant because all debt has to be ultimately financed by taxes, as implied by the above government budget constraint, provided the limit term goes to zero. Deficit-driven tax news fit this scenario under our null hypothesis of Ricardian equivalence, because the tax increases are explicitly targeting a reduction in inherited deficits and therefore lead to less national debt because deficits are financed by issuing debt. Less national debt means an offsetting reduction (in present value terms) in implied future taxes. Hence, a deficit reduction now is a switch from future to current taxation. In other words, the two different time paths of taxes (before and after the deficit-driven tax news) have the same present discounted values.

The stated purpose of deficit-driven tax increases is to swap debt for taxes without changing the stream of government spending. In comparison, long-run tax cuts are likely to affect the stream of future government spending under our null hypothesis because the stated intention of long-run tax cuts is to reduce the stream of future government spending (reduce the government size) or reduce the tax burden through stimulating future long-run economic growth and thus increasing the tax base.⁹ Therefore, the answer to the above question is no: long-run tax cuts are not suitable for testing Ricardian equivalence.

The data cover the period from the first quarter in 1945 to the fourth quarter in 2007. Following Romer and Romer (2010), I start regressions in 1950:I in order to allow for sufficient lags drawing on observations from earlier quarters. The data for the narrative measure of exogenous deficit-driven tax changes (as percent of GDP), the real GDP growth rate, the natural logarithm of real GDP, and the three-month Treasury bill rate are all from Romer and Romer's (2010) data file. The data are available at <https://www.aeaweb.org/articles.php?doi=10.1257/aer.100.3>, along with

⁹It is straightforward to define the above intertemporal government budget constraint in terms of ratios to GDP to illustrate this effect (see, e.g., Hakkio and Rush, 1991).

RATS software code and a detailed description of the construction of narrative variables (Romer and Romer, 2009b).

The data were retrieved on 29 October 2013. Regressions were carried out with the software RATS, version 6.35, based on the original code of Romer and Romer (2010), except for the CUSUM of squares tests that were done in EViews, version 8.1. In order to check the accuracy of the code, regressions were repeated with EViews.

3. Ricardian Equivalence Test Results: OLS and VAR

3.1 Full Sample Period: OLS Without and With Controls for Output Growth

The first step in the analysis of Romer and Romer (2010, equation (6), p. 780) is to study the effect of an exogenous tax change on real output growth. Instead of using total exogenous tax changes, I select the present value of deficit-driven tax changes only:

$$\Delta Y_t = \alpha + \sum_{i=0}^{12} \beta_i \Delta T_{t-i} + \varepsilon_t, \quad (1)$$

where ΔY_t is the log change (growth rate) of real GDP, ΔT_t is the present value of the stream of tax changes legislated at time t , and ε_t is a mean-zero Gaussian error that is orthogonal to all ΔT_{t-i} by construction of the narrative tax changes. The sum $\sum_{i=0}^j \beta_i$ measures the cumulative fiscal multiplier from period 0 to period j .

The permanent income hypothesis predicts for tax increases a negative effect on consumption as expected after-tax real income falls and government bonds are treated as net wealth under Ricardian non-equivalence so that overall household wealth falls. This causes additional dynamics that lower real GDP so that $\beta_i < 0$ for some or all i . On the other hand, intertemporal substitution predicts, in the presence of distortionary taxes and Ricardian non-equivalence, an increase in consumption (or investment, depending on the type of tax used) before the implementation of a tax increase (assuming there is a delay; hence $\beta_i > 0$ for periods i till the implementation date) and a decrease after the tax takes effect (with $\beta_i < 0$ for periods i after the

implementation date). In this case, the proper model is one that accounts for both the date when the tax change is legislated and the date(s) when it is implemented. In contrast, Ricardian equivalence predicts that the effects on real GDP from such a deficit-driven tax increase should not be statistically significantly different from zero (all $\beta_i = 0$).

Figure 2 presents the cumulative effect of a 1% of GDP increase in deficit-driven tax revenue (solid line) on real GDP, estimated with ordinary least squares (OLS). These are the cumulative multipliers for exogenous deficit-driven tax increases. The effect on real GDP is negative in the first period and positive thereafter, reaching a peak of 1.98% five periods after the tax shock and then declining. Are the effects of deficit driven tax increases statistically significantly different from zero? The answer depends on the width of the confidence bands used in Figure 2. The usual two standard-error confidence band, which is for normally distributed coefficient estimates approximately a 95% confidence band, indicates no statistical significance at any horizon in Figure 2. Using instead the one standard-error confidence band, which is equivalent to an approximate 68% confidence band, gives a different picture. Now, the responses of real GDP are statistically significantly different from zero at horizons three to eight after the tax shock. The responses are positive so that the tax increase causes an increase in real GDP at these horizons.

The next step in Romer and Romer (2010) is to include in equation (1) the lags of GDP growth:

$$\Delta Y_t = \alpha + \sum_{i=0}^{12} \beta_i \Delta T_{t-i} + \sum_{j=1}^{11} \gamma_j \Delta Y_{t-j} + \varepsilon_t. \quad (2)$$

The lags of output growth control for the dynamics of GDP and a "multitude of other influences" (Romer and Romer, 2010, p. 781). Figure 3 reports the effects of a 1% of GDP increase in deficit-driven taxes on real GDP when lagged output growth is controlled for in the regression. Overall the effects (cumulative multipliers) are similar in magnitude to those without controls, though the peak effect is lower. The two and one standard-error confidence bands are calculated from bootstraps with 10,000 replications, following Romer and Romer (2010). The effect in the first period

is again negative but remains statistically insignificant for the 95% (two standard-error) confidence band. The 68% (one standard-error) confidence band indicates statistical significance at the first horizon, although just barely so. The only other horizon where there is possibly statistical significance is the fourth one, but again it is very much a borderline case.

Romer and Romer (2010) report 68% confidence bands. However, Ramey (2011, footnote 11, p. 11) pointed out that there is "no formal justification" provided in the literature for using 68% confidence bands instead of the usual 95% bands. Furthermore, it is neither justified on theoretical grounds to use 68% bands.

The empirical evidence over the full sample period would seem to favor the Ricardian equivalence hypothesis because tax increases intended to reduce inherited budget deficits have no statistically significant effects on real output at usual confidence levels. But, the crucial question is whether the regressions in equations (1) and (2) are structurally stable over the sample period.

The cumulative sum (CUSUM) of squares test is based on squared recursive residuals that are calculated from sequentially increasing samples, starting with the first thirteen observations in our case up to the full sample size. This test was proposed by Brown *et al.* (1975). The test provides an indication of parameter or variance instability in the case of unknown break dates. Figures 4a and 4b depict the values of the CUSUM of squares statistic along with the upper and lower 5% significance lines. Test statistic values outside the area between the two significance lines mark the time periods when structural change likely occurred. For both Figures 4a and 4b, calculated from equations (1) and (2), the time interval of structural change includes the year 1980. I therefore follow Romer and Romer (2010), among many others, and split the sample into a period prior to 1980:IV and post-1980:IV.

3.2 The Prior- and Post-1980:IV Periods: OLS Without and With Controls for Output Growth

Figure 5a shows for equation (2) the cumulative effect of a 1% of GDP increase in deficit-driven tax revenue on GDP, in the sub-sample period 1950:I to 1980:IV. The effect on real GDP is not statistically significant whether the bootstrapped 95% or

68% confidence bands are used, except for the first period. The confidence bands are quite wide indicating that the coefficients are not estimated very precisely. This is not surprising because there are only three news shocks (tax changes) in this sample period.

Figure 5b presents for equation (2) cumulative effects of a 1% of GDP increase in deficit-driven tax revenue on GDP in the post-1980:IV period from 1981:I to 2007:IV. The response to a tax increase is positive and statistically significant from period four to nine after the impact for the bootstrapped 95% confidence band. The impact reaches a peak of 4.72%, i.e., the multiplier peaks at a positive value of 4.72.

3.3 The Prior- and Post-1980:IV Period: A VAR With Tax Shocks, Real GDP and an Interest Rate

An alternative model specification to single equations is a VAR. The three-variable VAR consists of the narrative deficit-driven present value of the tax shock as a percent of GDP, the log of real GDP and the three-month Treasury bill rate in percent. Following Romer and Romer (2010), I use 12 lags on each variable and order them for the Cholesky decomposition with the tax shock first, followed by real GDP and then the Treasury bill rate last. This ordering allows for tax changes to affect output contemporaneously but output shocks and interest rate shocks have no contemporaneous, i.e., within the same quarter, effects on tax changes. The interest rate responds contemporaneously to all shocks. There are no restrictions on lagged responses of variables to shocks. The appeal of the VAR approach is that it allows for rich endogenous dynamics that follow unexpected fiscal policy changes.

Mertens and Ravn (2014) reconciled the different estimates of tax multipliers in narrative approaches, such as in Romer and Romer (2010), with those derived from structural VAR models as pioneered by Blanchard and Perotti (2002) and applied by many others, among them Favero and Giavazzi (2012), who combined the narrative approach with the VAR. Mertens and Ravn (2014) showed that the lower estimated tax multipliers in structural VARs can be explained by output elasticities of tax revenues, plugged into such models, that are lower than supported by empirical evidence, or by not accounting for measurement errors of narrative measures

when they are embedded in "proxy" VARs. They found support for tax multipliers of the size as calculated from models used by Romer and Romer (2010) with narrative measures.

Figure 6a shows the cumulative impulse response of real GDP to a deficit-driven tax shock in the sample prior to 1980:IV. It is negative and statistically significant initially, in the first period, then becomes positive afterwards but is no longer statistically significant at 95% bootstrapped confidence intervals. The bootstrapped 68% confidence band paints a similar picture with somewhat borderline significance in periods seven and eight. The band is quite wide and, consistent with the single equation regression results, reveals imprecise estimation.

Figure 6b is for the post-1980:IV period and the cumulative effect of a 1% of GDP deficit-driven tax increase. The effects on real GDP are positive throughout the impulse horizon considered. At the 95% bootstrapped confidence band, the tax increase leads to an increase in real GDP that is statically significant from period three to eight and peaks at a value of 3.11 in period six after the tax shock. Again, the results are similar to those with the single equation regressions from before.

3.4 The Role of Fiscal Anticipation in the Post-1980:IV Period

One possible explanation of the apparent positive fiscal multipliers in response to a tax increase is the intertemporal substitution effect due to lags in implementing tax increases for the post-1980:IV period. The effect of the announcement of a tax increase in future quarters (instead of an immediate increase at the time of the announcement) will lead to increases in spending before the tax increase takes effect, whether through direct effects or through increased work hours when payroll taxes are involved. For this argument to hold, taxes are not assumed to be lump-sum but distortionary. In order to explore this possibility, I separate out deficit-driven tax changes that are implemented with a delay of one or more quarters and therefore become anticipated in quarters after the announcement. The remainder of deficit-driven tax change announcements are news or shocks and unanticipated. Following the same procedure as explained in Mertens and Ravn (2012) but applied to deficit-driven tax changes instead, I run separate regression for unanticipated and anticipated deficit-

driven tax changes.

Figure 7a reports the cumulative multipliers for an unanticipated deficit-driven tax shock in the post 1980:IV period. As one would expect based on standard theory, the effects on real GDP are negative at all horizons. The largest negative effect occurs in the third quarter after the impact and the multiplier reaches a value of -1.67, however, all multiplier estimates are insignificant at the usual 5% level. Estimates are too imprecise to draw further conclusions on the size of the tax multiplier when tax shocks are unanticipated. This is due to the small sample size with only five observations. More observations on tax shocks would be needed to reduce the width of the confidence bands and to calculate reliable multiplier estimates.

Figure 7b graphs the responses of real GDP to an anticipated change in taxes in the post-1980:IV period. If intertemporal substitution was happening, one would expect positive effects of announced tax increases on real GDP. Indeed, Figure 7b shows positive cumulative multipliers from three quarter onwards and the multipliers are statistically significantly different from zero in quarters five to twelve after the impact.¹⁰ The cumulative multiplier reaches a maximum value of 2.53 in quarter ten. The estimates are relatively precise, given that I have a sample with 12 observations on anticipated tax shocks in the post-1980:IV period. These results are consistent with intertemporal substitution due to fiscal anticipation or foresight as argued by Leeper *et al.* (2008). Furthermore, Ricardian equivalence is hence not supported by data using the narrative record on tax changes.

4. Conclusion

This paper contributes to the empirical literature on fiscal policy by using a part of a narrative measures developed by Romer and Romer (2010) in order to provide new empirical evidence on the validity of the Ricardian equivalence hypothesis. Ricardian equivalence may be a good approximation to reality and the fact that taxes are generally not lump-sum does not necessarily invalidate it, as shown for example by Judd (1987) in a theoretical model with distortionary taxes, finite lives, and adjust-

¹⁰Using instead a VAR as in Section 3.3 leads to similar results.

ment costs. The data set of Romer and Romer (2010) identified two sub-categories of tax shocks: measures to address inherited budget deficits and long-run measures to improve economic performance. Romer and Romer (2009a, 2010) did not study the effects of the deficit-driven sub-category that I examine in this paper. By doing so, I shed new light on an unresolved empirical question: how good of an approximation is Ricardian equivalence for fiscal policy? Ricardian equivalence questions that aggregate demand can be stimulated with tax cuts, all else the same, if forward-looking taxpayers take into account the implications of today's tax cuts, financed by borrowing, on taxes in the future. The Ricardian argument applies in reverse as well: if governments consolidate their budgets by increasing taxes, all else the same, in order to reduce government debt, which lowers future taxes, no negative aggregate demand effects follow this action because the intertemporal tax burden is unchanged.

The empirical results do not support the Ricardian equivalence hypothesis because the effects of a tax increase, keeping government spending fixed, have a statistically significant influence on real GDP. The results are robust to the specification of the model either as a single equation or as a vector-autoregressive system. While Ricardian equivalence seemed to be supported at first sight over the full post-WWII sample period, I find that this result does not hold up when I split the sample in 1980:IV. The post-1980:IV sample allows for precise inference that shows that increasing taxes solely motivated by concerns over the health of the fiscal position of the government has a statistically significant positive effect on real GDP for anticipated tax increases, as predicted by the theory of fiscal anticipation of Leeper *et al.* (2008) when tax increases are implemented with delays of several quarters. Intertemporal substitution in the presence of distortionary taxes explains these effects. The increase in economic activity ahead of the implementation of a tax increase therefore comes at the expense of lower economic activity after the implementation of the tax increase. It may well be that the effects of deviations from Ricardian equivalence are overall small, as argued by Evans (1993). Assessing the quantitative effects after the implementation of anticipated taxes is not a trivial task that I leave for future research.

References

- Bailey, Martin J. (1993). Note on Ricardian Equivalence. *Journal of Public Economics* 51, 437-226.
- Barro, Robert J. (1989). The Ricardian Approach to Budget Deficits. *Journal of Economic Perspectives* 3, 37-54.
- Barro, Robert J. (1974). Are Government Bonds Net Wealth? *Journal of Political Economy* 82, 1095-117.
- Blanchard, Olivier, and Roberto Perotti (2002). An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output. *Quarterly Journal of Economics* 117, 1329-1368.
- Brown, R. L., J. Durbin, and J. M. Evans (1975). Techniques for Testing the Constancy of Regression Relationships Over Time. *Journal of the Royal Statistical Society, Series B*, 37, 149–192.
- Cloyne, James (2013). Discretionary Tax Changes and the Macroeconomy: New Narrative Evidence from the United Kingdom. *American Economic Review* 103, 1507–1528.
- Evans, George W., Seppo Honkapohja and Kaushik Mitra (2012). Does Ricardian Equivalence Hold When Expectations Are Not Rational? *Journal of Money, Credit, and Banking* 44, 1259-1283.
- Evans, Paul (1993). Consumers are not Ricardian: Evidence from Nineteen Countries. *Economic Inquiry* 31, 534–548.
- Evans, Paul (1991). Is Ricardian Equivalence a Good Approximation? *Economic Inquiry* 29, 626–644.
- Favero, Carlo, and Francesco Giavazzi (2012). Measuring Tax Multipliers: The Narrative Method in Fiscal VARs. *American Economic Journal: Economic Policy* 4, 69-94.
- Hakkio, Craig S., and Mark Rush (1991). Is the Budget Deficit 'Too Large'? *Economic Inquiry* 29, 429-445.
- Hamilton, James D. and Marjorie A. Flavin (1986). On the Limitations of Government Borrowing: A Framework for Testing. *American Economic Review* 76, 808-819.
- Haug, Alfred A. (1996). Blanchard's Model of Consumption: An Empirical Study. *Journal of Business and Economic Statistics* 14, 169-177.
- Judd, Kenneth (1987). Debt and Distortionary Taxation in a Simple Perfect Foresight Model. *Journal of Monetary Economics* 20, 51-72.

- Leeper, Eric M., Walker, Todd B., and Shu-Chun S. Yang (2008). Fiscal Foresight: Analytics and Econometrics. Working Paper No. 14028, National Bureau of Economic Research.
- Mertens, Karel, and Morten O. Ravn (2014). A Reconciliation of SVAR and Narrative Estimates of Tax Multipliers. *Journal of Monetary Economics* 68, S1-S19.
- Mertens, Karel, and Morten O. Ravn (2012). Empirical Evidence on the Aggregate Effects of Anticipated and Unanticipated US Tax Policy Shocks. *American Economic Journal: Economic Policy* 4, 145-181.
- Ramey, Valery A. (2011). Identifying Government Spending Shocks: It's all in the Timing. *Quarterly Journal of Economics* 126, 1-50.
- Ramey, Valerie A., and Matthew D. Shapiro (1998). Costly Capital Reallocation and the Effects of Government Spending. *Carnegie-Rochester Conference on Public Policy* 48, 145-194.
- Ricciuti, Roberto (2003). Assessing Ricardian Equivalence. *Journal of Economic Surveys* 17, 55-78.
- Romer, Christina D., and David H. Romer (2010). The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Policy Shocks. *American Economic Review* 100, 763-801.
- Romer, Christina D., and David H. Romer (2009a). Do Tax Cuts Starve the Beast? The Effects of Tax Changes on Government Spending. *Brookings Papers on Economic Activity*, 139-200.
- Romer, Christina D., and David H. Romer (2009b). A Narrative Analysis of Postwar Tax Changes. <https://www.aeaweb.org//articles.php?doi=10.1257/aer.100.3>.
- Seater, John J. (1993). Ricardian Equivalence. *Journal of Economic Literature* 31, 142-190.

FIGURE 1. Present values at time of passage for deficit-driven tax increases

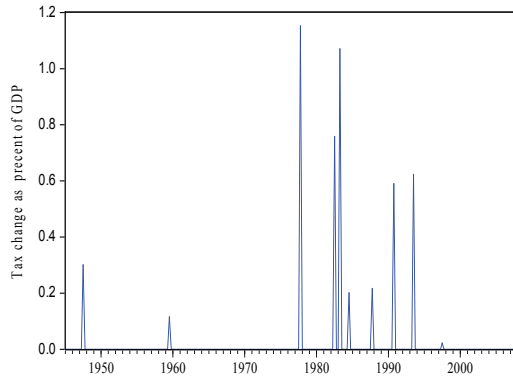


FIGURE 2. Estimated impact of an exogenous deficit-driven tax increase of 1% of GDP on GDP, as in equation (1) with no controls for lagged GDP growth and one and two standard-error confidence bands

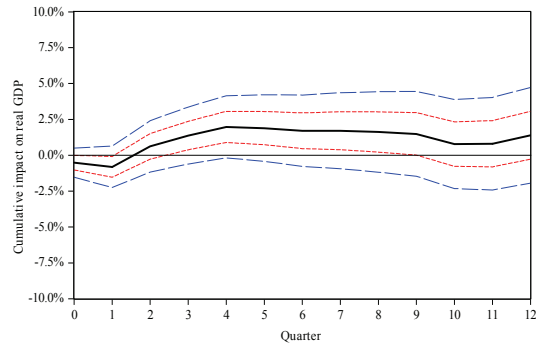


FIGURE 3. Estimated impact of an exogenous deficit-driven tax increase of 1% of GDP on GDP, as in equation (2) with controls for lagged GDP growth and one and two standard-error confidence bands (bootstrapped)

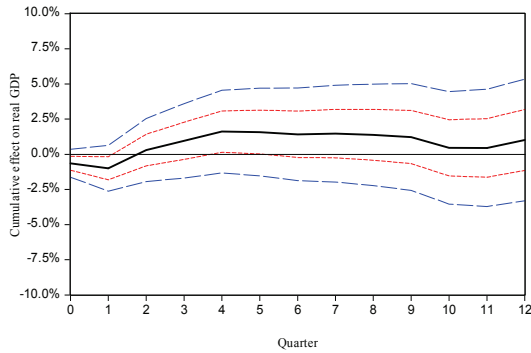


FIGURE 4a. Structural change test of the GDP growth regression with no controls as in equation (1)

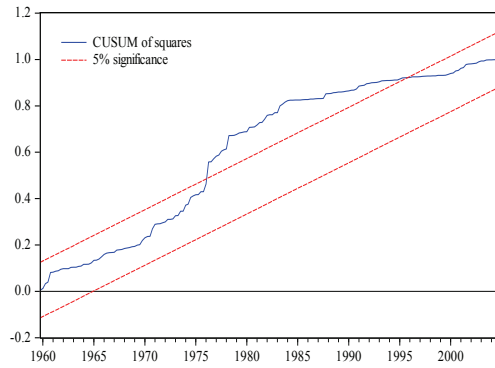


FIGURE 4b. Structural change test of the GDP growth regression with controls as in equation (2)

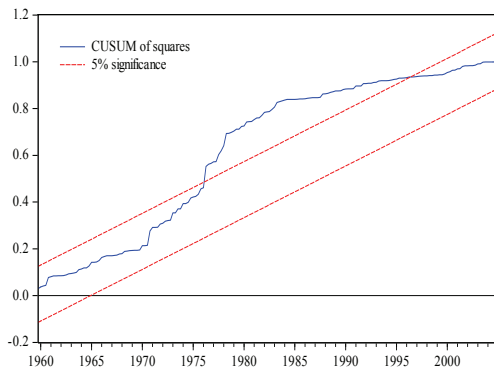


FIGURE 5a. Prior-1980:IV estimated impact of an exogenous deficit-driven tax increase of 1% of GDP on GDP, as in equation (2) with controls for lagged GDP growth and one and two standard-error confidence bands (bootstrapped)

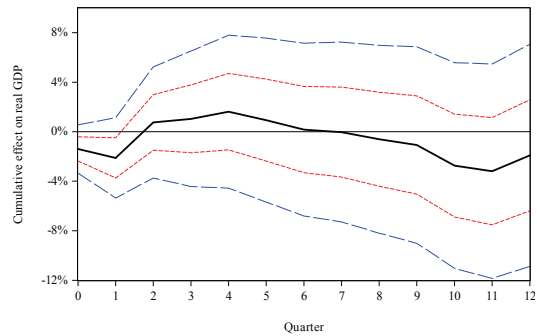


FIGURE 5b. Post-1980:IV estimated impact of an exogenous deficit-driven tax increase of 1% of GDP on GDP, as in equation (2) with controls for lagged GDP growth and one and two standard-error confidence bands (bootstrapped)

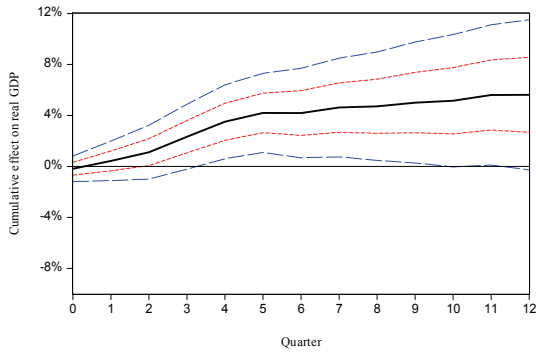


FIGURE 6a. Prior-1980:IV estimated impact of an exogenous deficit-driven tax increase of 1% of GDP on GDP, in a VAR with 3 variables and one and two standard-error confidence bands (bootstrapped)

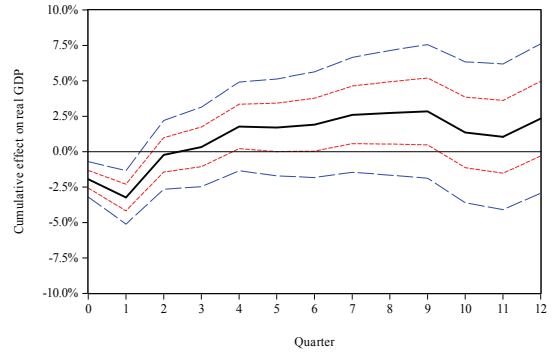


FIGURE 6b. Post-1980:IV estimated impact of an exogenous deficit-driven tax increase of 1% of GDP on GDP, in a VAR with 3 variables and one and two standard-error confidence bands (bootstrapped)

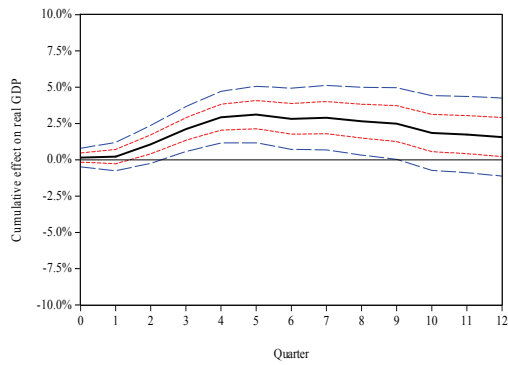


FIGURE 7a. Post-1980:IV estimated impact of an unanticipated exogenous deficit-driven tax increase of 1% of GDP on GDP, with one and two standard-error confidence bands

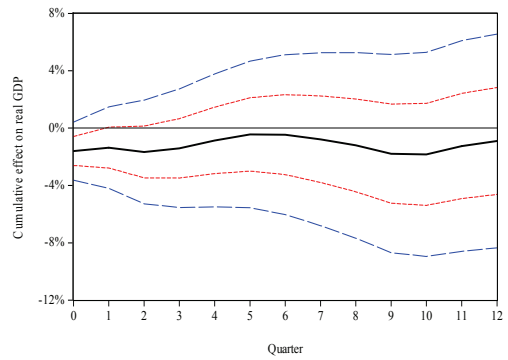


FIGURE 7b. Post-1980:IV estimated impact of an anticipated exogenous deficit-driven tax increase of 1% of GDP on GDP, with one and two standard-error confidence bands

