A replication of: "Does High Public Debt Consistently Stifle Economic Growth? A Critique of Reinhart and Rogof"

Thomas de Graaff

Friday, August 22, 2014

Abstract

We replicate Reinhart and Rogoff (2010a and 2010b) and find that coding errors, selective exclusion of available data, and unconventional weighting of summary statistics lead to serious errors that inaccurately represent the relationship between public debt and GDP growth among 20 advanced economies in the post-war period. Our finding is that when properly calculated, the average real GDP growth rate for countries carrying a public-debt-to-GDP ratio of over 90 percent is actually 2.2 percent, not ???0.1 percent as published in Reinhart and Rogoff. That is, contrary to RR, average GDP growth at public debt/GDP ratios over 90 percent is not dramatically different than when debt/GDP ratios are lower. We also show how the relationship between public debt and GDP growth varies significantly by time period and country. Overall, the evidence we review contradicts Reinhart and Rogoff's claim to have identified an important stylized fact, that public debt loads greater than 90 percent of GDP consistently reduce GDP growth

Introduction

In "Growth in Time of Debt" Reinhart and Rogoff (Reinhart and Rogoff (2010a) and Reinhart and Rogoff (2010b); from here on RR) propose a set of "stylized facts" concerning the relationship between public debt and GDP growth. RR's "main result is that whereas the link between growth and debt seems relatively weak at 'normal' debt levels, median growth rates for countries with public debt over roughly 90 percent of GDP are about one percent lower than otherwise; (mean) growth rates are several percent lower" (Reinhart and Rogoff (2010a), p. 573).

To build the case for a stylized fact, RR stresses the relevance of the relationship to a range of times and places and the robustness of the finding to modest adjustments of the econometric methods and categorizations. The RR methods are non-parametric and appealingly straightforward. RR organizes country-years in four groups by public debt/GDP ratios, 0–30 percent, 30–60 percent, 60–90 percent, and greater than 90 percent. They then compare average real GDP growth rates across the debt/GDP groupings. The straightforward non-parametric method highlights a nonlinear relationship, with effects appearing at levels of public debt around 90 percent of GDP. We present RR's key results on mean real GDP growth from Figure 2 of Reinhart and Rogoff (2010a) and Appendix Table 1 of Reinhart and Rogoff (2010b) in Table 1.

Ratio of public debt to gdp	below 30 percent	30 to 60 percent	60 to 90 percent	90 percent and above
Average real grp growth	4.1	2.8	2.8	-0.1

Table 1: Real GDP Growth as the Level of Public Debt Varies 20 advanced economies, 1946–2009

Figure 2 in Reinhart and Rogoff (2010a) and the first line of Appendix Table 1 in Reinhart and Rogoff (2010b) in fact do not match perfectly, but they do deliver a consistent message about growth in time of debt: real GDP growth is relatively stable around 3 to 4 percent until the ratio of public debt to GDP reaches 90

percent. At that point and beyond, average GDP growth drops sharply to zero or slightly negative.

A necessary condition for a stylized fact is accuracy. We replicate RR and finnd that coding errors, selective exclusion of available data, and unconventional weighting of summary statistics lead to serious errors that inaccurately represent the relationship between public debt and growth among these 20 advanced economies in the post-war period. Our most basic finding is that when properly calculated, the average real GDP growth rate for countries carrying a public debt-to-GDP ratio of over 90 percent is actually 2.2 percent, not -0.1 percent as RR claims. That is, contrary to RR, average GDP growth at public debt/GDP ratios over 90 percent is not dramatically different than when public debt/GDP ratios are lower.

We additionally refute the RR evidence for an "historical boundary" around public debt/GDP of 90 percent, above which growth is substantively and non-linearly reduced. In fact, there is a major non-linearity in the relationship between public debt and GDP growth, but that non-linearity is between the lowest two public debt/GDP categories, 0–30 percent and 30–60 percent, a range that is not relevant to current policy debate.

For the purposes of this discussion, we follow RR in assuming that causation runs from public debt to GDP growth. RR concludes, "At the very minimum, this would suggest that traditional debt management issues should be at the forefront of public policy concerns" (Reinhart and Rogoff 2010a, 578). In other work (see, for example, Reinhart and Rogoff (2011)), Reinhart and Rogoff acknowledge the potential for reverse causality, i.e., that weak economic growth may increase debt by reducing tax revenue and increasing public expenditures. (Reinhart and Rogoff 2010a; Reinhart and Rogoff 2010b) however, make clear that the implied direction of causation runs from public debt to GDP growth.

Publication, Citations, Public Impact, and Policy Relevance

According to Reinhart's and Rogoff's website[1] the findings reported in the two 2010 papers formed the basis for testimony before the Senate Budget Committee (Reinhart, February 9, 2010) and a Financial Times opinion piece "Why We Should Expect Low Growth amid Debt"

The key tables and figures have been reprinted in additional Reinhart and Rogoff publications and presentations of Centre for Economic Policy Research and the Peter G. Peterson Institute for International Economics. A Google Scholar search for the publication excluding pieces by the authors themselves finds more than 500 results.[2]

The key findings have also been widely cited in popular media. Reinhart's and Rogoff's website lists 76 high-profile features, including *The Economist, Wall Street Journal, New York Times, Washington Post*, Fox News, National Public Radio, and MSNBC, as well as many international publications and broadcasts.

Furthermore, (Reinhart and Rogoff 2010a) is the only evidence cited in the "Paul Ryan Budget" on the consequences of high public debt for economic growth. Representative Ryan's "Path to Prosperity" reports

> A well-known study completed by economists Ken Rogoff and Carmen Reinhart confirms this common-sense conclusion. The study found conclusive empirical evidence that gross debt (meaning all debt that a government owes, including debt held in government trust funds) exceeding 90 percent of the economy has a significant negative eect on economic growth. (Ryan 2012, 78)

RR have clearly exerted a major in uence in recent years on public policy debates over the management of government debt and fiscal policy more broadly. Their findings have provided significant support for the austerity agenda that has been ascendant in Europe and the United States since 2010.

[1] http://www.reinhartandrogoff.com/related-research/growth-in-a-time-of-debt-featured-in (visited 7 April 2013).

[2] A search on [Reinhart Rogoff "Growth in a Time of Debt" -author:rogoff -author:reinhart] yielded 538 Google Scholar results on 7 April 2013).

Replication

RR examines three data samples: 20 advanced economies over 1946–2009; the same 20 economies over roughly 200 years; and 20 emerging market economies 1970–2009. We replicate the results only from the first sample as these are the most relevant to current U.S. and European policy debates, and they require the least splicing of data from multiple sources. We focus exclusively on their results regarding means because these have generated the most widespread attention. On their website, Reinhart and Rogoff provide public access to country historical data for public debt and GDP growth in spreadsheets with complete source documentation.[3] However, the spreadsheets do not include guidance on the exact data series, years, and methods used in RR.

We were unable to replicate the RR results from the publicly available country spreadsheet data although our initial results from the publicly available data closely resemble the results we ultimately present as correct. Reinhart and Rogoff kindly provided us with the working spreadsheet from the RR analysis. With the working spreadsheet, we were able to approximate closely the published RR results. While using RR's working spreadsheet, we identified coding errors, selective exclusion of available data, and unconventional weighting of summary statistics.

Selective exclusion of available data and data gaps

RR designates 1946–2009 as the period of analysis of the post-war advanced economies with table notes indicating gaps or other unavailability of the data. In general, RR used data if they were available in the working spreadsheet. Most differences in period of coverage concern the starting year of the data. For example, the US series extends back to 1946. Outside the US, the series for some countries do not begin until the 1950's and that for Greece is unavailable before 1970. Nine countries are available from 1946, seventeen from 1951, and all countries but Greece enter the dataset by 1957. There are some gaps and oddities in the data. For example, public debt/GDP is unavailable for France for 1973–1978, real GDP growth is unavailable for Spain for 1959–1980, Austria experienced 27.3 and 18.9 percent real GDP growth in 1948 and 1949 (with both years in lower public-debt groups), and Portugal's debt/GDP jumps by 25 percentage points from 1999 to 2000 when the country's currency and the denomination of the series changed from the escudo to the euro. We largely accept the RR data on debt/GDP and real GDP growth as given and do not pursue the implications of data gaps.

More significant are RR's data exclusions with three other countries: Australia (1946– 1950), New Zealand (1946–1949), and Canada (1946–1950).[4] The exclusions for New Zealand are of particular significance. This is because all four of the excluded years were in the highest, 90 percent and above, public debt/GDP category. Real GDP growth rates in those years were 7.7, 10.9, -9.9, and 10.8 percent. After the exclusion of these years, New Zealand contributes only one year to the highest public debt/GDP category, 1951, with a real GDP growth rate of -7.6 percent. The exclusion of the missing years is alone responsible for a reduction of -0.3 percentage points of estimated real GDP growth in the highest public debt/GDP category. Further, RR's unconventional weighting method that we describe below amplifies the effect of the exclusion of years for New Zealand so that it has a very large effect on the RR results.

RR reports 96 country-years in the highest public debt/GDP category. Our corrected analysis finds 110 country-years in the highest, above-90-percent public debt/GDP, category. The difference is accounted for by the years that RR excluded: 5 years for Australia; 5 years for Canada; and 4 years of New Zealand. With the spreadsheet error discussed below, RR in fact estimated GDP growth in the highest public debt/GDP

[3] See http://www.reinhartandrogoff.com/data/browse-by-topic/topics/9/ and http://www.reinhartandrogoff.com/data/browse-by-topic/topics/16/

[4] All of these cases would contribute observations to the highest public debt/GDP category. In contrast to these exclusions, all of the data for the US, which contributes all of its four observations in the highest public debt/GDP category in these early years, are included. The US series includes the very large GDP decline associated with post-World War II demobilization discussed in detail in (Irons and Bivens 2010). In 1946, the US public debt/GDP ratio was 121.3 percent, and the economy contracted by 10.9 percent. In the 1946–2009 study period, the U.S. had exactly four years, 1946–1949, with a public debt/GDP ratio above 90 percent. Growth in these years was -10.9, -0.9, 4.4, and -0.5. (see Irons and Bivens 2010 for more detailed discussion.)

category with only 71 country-years of data: 25 years of Belgium were dropped in addition to the 14 already accounted for by the years that RR excluded.

Spreadsheet coding error

A coding error in the RR working spreadsheet entirely excludes five countries, Australia, Austria, Belgium, Canada, and Denmark, from the analysis.[5] The omitted countries are selected alphabetically and, hence, likely randomly with respect to economic relationships. This spreadsheet error, compounded with other errors, is responsible for a -0.3 percentage- point error in RR's published average real GDP growth in the highest public debt/GDP category. It also overstates growth in the lowest public debt/GDP category (0 to 30 percent) by 0.1 percentage point and understates growth in the second public debt/GDP category (30 to 60 percent) by -0.2 percentage point.

Unconventional weighting of summary statistics

RR adopts a non-standard weighting methodology for measuring average real GDP growth within their four public debt/GDP categories. After assigning each country-year to one of four public debt/GDP groups, RR calculates the average real GDP growth for each country within the group, that is, a single average value for the country for all the years it appeared in the category. For example, real GDP growth in the UK averaged 2:4 percent per year during the 19 years that the UK appeared in the highest public debt/GDP category while real GDP growth for the US averaged -1.0 percent per year during the 4 years that the US appeared in the highest category. The country averages within each group were then averaged, equally weighted by country, to calculate the average real GDP growth rate within each public debt/GDP grouping.

RR does not indicate or discuss the decision to weight equally by country rather than by country-year. In fact, possible within-country serially correlated relationships could support an argument that not every additional country-year contributes proportionally additional information. Yet equal weighting of country averages entirely ignores the number of years that a country experienced a high level of public debt relative to GDP. Thus, the existence of serial correlation could mean that, with Greece and the UK, 19 years carrying a public debt/GDP load over 90 percent and averaging 2.9 percent and 2.4 percent GDP growth respectively do not each warrant 19 times the weight as New Zealand's single year at -7.6 percent GDP growth or five times the weight as the US's four years with an average of -2.0 percent GDP growth. But equal weighting by country gives a one-year episode as much weight as nearly two decades in the above 90 percent public debt/GDP range. RR needs to justify this methodology in detail. It otherwise appears arbitrary and unsupportable.

Table 2 presents average results by country for the above-90-percent public debt/GDP category for the alternative methods. (Table A-1 presents the full results for all debt/GDP categories.) The first three columns show the number of years that each country spent in the highest debt/GDP category. The Correct column reports the most available data for 1946–2009. The RR Exclusion column excludes available early years of data for Australia (1946–1950), Canada (1946–1950), and New Zealand (1946–1949). The RR Spreadsheet Error column rejects the spreadsheet error that omits all years for Australia, Austria, Belgium, Canada, and Denmark from the analysis. The Weights columns show the alternative weightings to compute average real GDP growth. The Country-Years weights column shows weights proportional to the number of country-years in the highest public debt/GDP category. The RR weights column shows the equal weighting by country used in RR. The GDP Growth columns show average real GDP growth for each country in the years in which it appeared in the highest debt/GDP category. The Correct GDP Growth column shows the average real GDP growth for all available country-years. The RR GDP Growth column shows the average real GDP growth used in RR with excluded years, spreadsheet errors, and a transcription error.

For example, Canada spent 5 years in the highest public debt/GDP category (4.5 percent of the 110 country-years in this category) and Canada's average real GDP growth during these 5 years was 3.0 percent per year. However the RR spreadsheet error and the RR years exclusion result in Canada not providing any data for the computation of the average for the highest debt/GDP category.

[5] RR averaged cells in lines 30 to 44 instead of lines 30 to 49.

In the case of New Zealand, instead of constituting 5 of 110 country-years at 2.6 percent growth, the country contributes -7.9 percent growth for a full 14.3 percent (one-seventh) of the RR's GDP growth estimate for the above 90 percent public debt/GDP grouping.[6]

110 country-years appear in the highest public debt/GDP category with only 10 countries ever appearing in the category. Three of these, Australia, Belgium, and Canada, were excluded from the analysis by spreadsheet error, leaving seven countries in the highest category in RR. The included countries are Greece (19 years in the highest category with average real GDP growth of 2.9 percent per year); Ireland (7 years with average growth of 2.4 percent); Italy (10 years with average growth of 1.0 percent); Japan (11 years with average growth of 0:7 percent); New Zealand (1 year with average growth of -7.6 percent), the UK (19 years with average growth of 2.4 percent), and the US (4 years with average growth of -2.0 percent). As we noted above, the exclusion of four years for New Zealand (only a 4.5 percent loss of country-years in the highest public debt/GDP category) has a major effect on the computed average in the highest public debt/GDP category. It reduces the average growth for New Zealand in the highest public debt/GDP category from 2.6 to -7.6 percent per year. The combined effect of excluding the years for New Zealand and equally weighting the countries (rather than weighting by country-years) reduces the measured average real GDP growth in the highest public debt category by a very substantial 1.9 percentage points.

Summary: years, spreadsheet, weighting, and transcription

Table 3 summarizes the errors in RR and their effect on the estimates of average real GDP growth in each public debt/GDP category. Some of the errors have strong interactive effects. Table 3 shows the effect of each possible interaction of the spreadsheet error, selective year exclusion, and country weighting.

The errors have relatively small effects on measured average real GDP growth in the lower three public debt/GDP categories. GDP growth in the lowest public debt/GDP category is roughly 4 percent per year and in the next two categories is around 3 percent per year with or without correcting the errors.

In the over-90-percent public debt/GDP category, however, the effects of the errors are substantial. For example, the impact of the excluded years for New Zealand is greatly amplified when equal country weighting assigns 14.3 percent (1/7) of the weight for the average to the single year in which New Zealand is included in the above-90-percent public debt/GDP group. This one year is when GDP growth in New Zealand was -7.6 percent. The exclusion of years coupled with the country|as opposed to country-year|weighting alone accounts for almost -2 percentage points of under-measured GDP growth. The spreadsheet and transcription errors account for an additional -0.4 percentage point. In total, as we show in Table 3, actual average real growth in the high public debt category is +2.2 percent per year compared to the -0.1 percent per year published in RR. The actual gap between the highest and next highest debt/GDP categories is 1.0 percentage point (i.e., 3.2 percent less 2.2 percent). In other words, with their estimate that average GDP growth in the above-90-percent public debt/GDP group is -0.1 percent, RR overstates the gap by 2.3 percentage points or a factor of nearly two and a half.

[6] An apparent transcription error in transferring the country average from the country-specific sheets to the summary sheet reduced New Zealand's average growth in the highest public debt category from -7.6 to -7.9 percent per year. With only seven countries appearing in the highest public debt/GDP group, this transcription error reduces the estimate of average real GDP growth by another -0.1 percentage point.

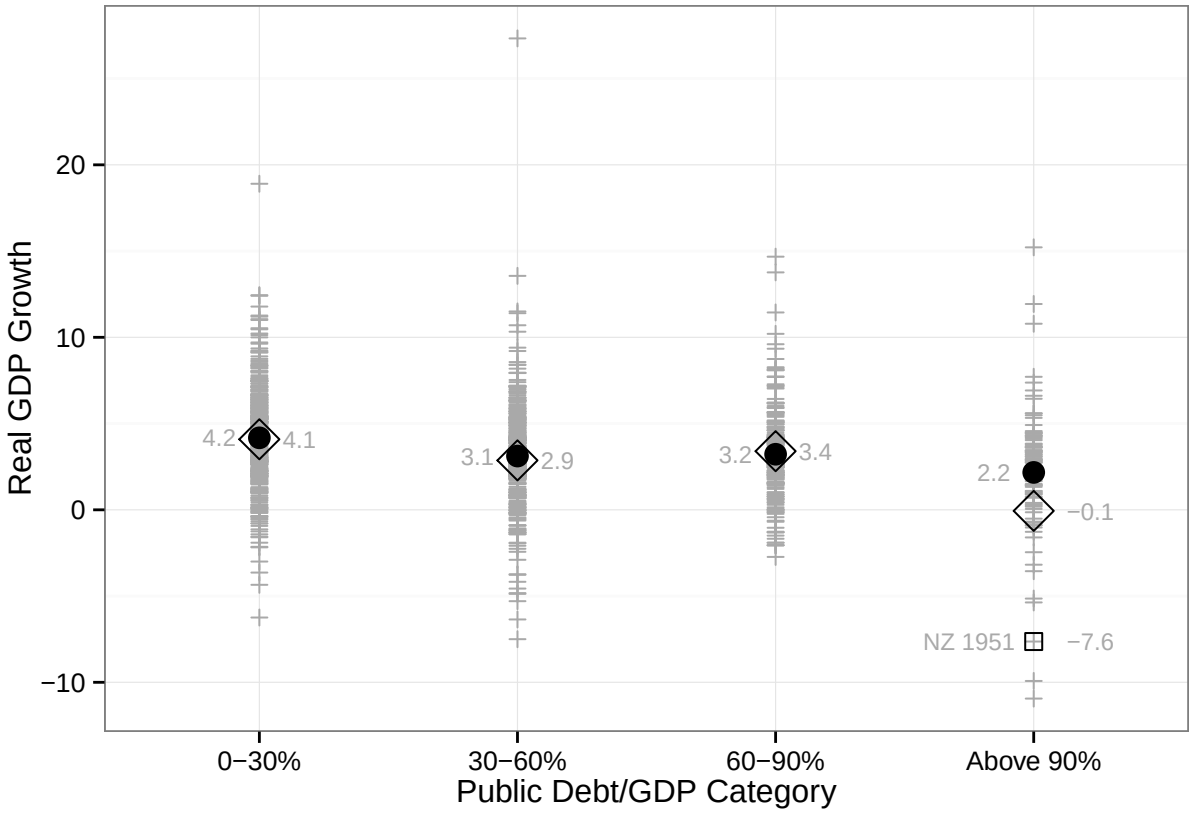
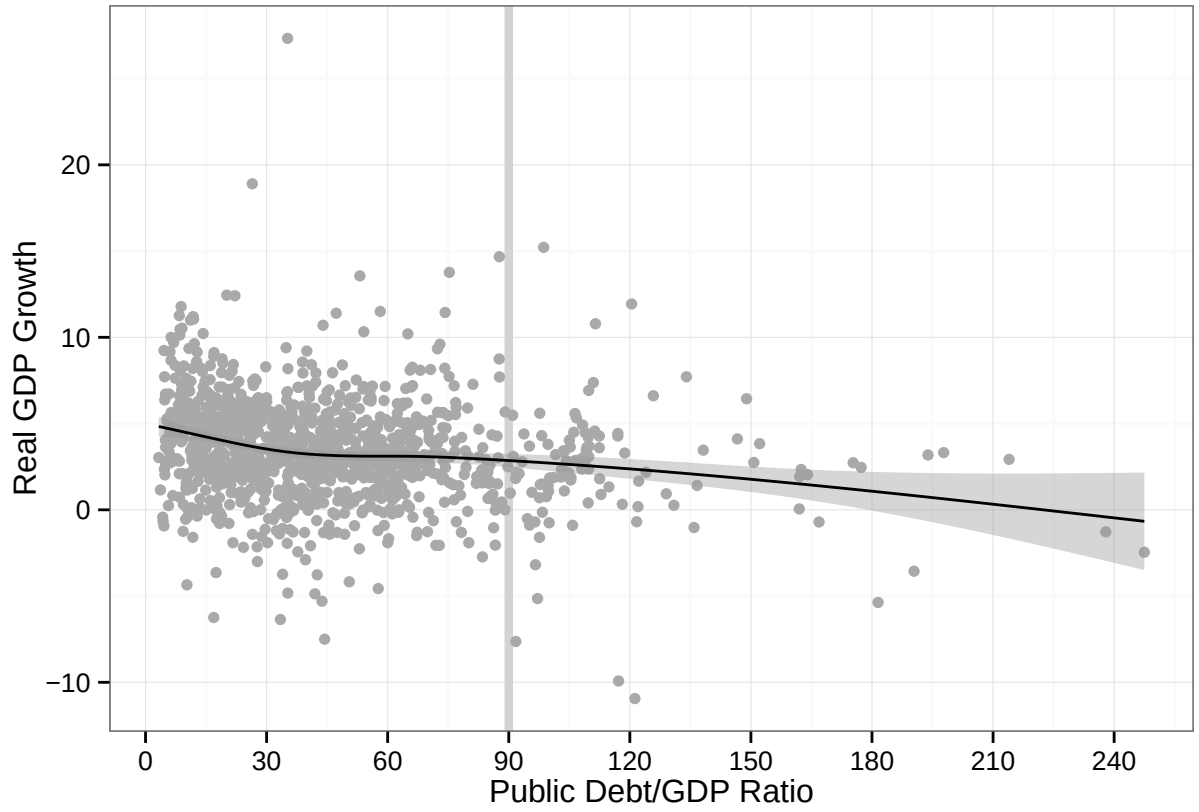


Figure 1 presents all of the country-year data, as continuous real GDP growth rates plotted against public debt/GDP categories. RR mean growth estimates are indicated by diamonds with the corrected growth estimates indicated by filled circles. The substantial error in the RR estimates of mean real GDP growth in the 90 percent public debt/GDP category is evident in the plot as is the relatively inconsequential errors in the lower three categories. The plot also shows large variation in real GDP growth in each public debt/GDP category. Finally, the plot includes an empty square as the data point for New Zealand in 1951, which alone accounts for one-seventh of RR's result for the highest public debt/GDP category.

Non-linearity at the "historical boundary"?

Our revised results also provide an opportunity to re-examine non-linearity in the relationship between public debt and growth. RR asserts, "The nonlinear response of growth to debt as debt grows towards historical boundaries is reminiscent of the 'debt intolerance' phenomenon developed in Reinhart, Rogoff, and Savastano (2003)" (Reinhart and Rogoff 2010a, 577).

The corrected means within each public debt/GDP category cast doubt on the identification of a nonlinear response that was an important component of RR's findings. We explore the question in several ways. First, we add an additional public debt/GDP category, extending by an additional 30 percentage points of public debt/GDP ratio|that is, we add 90–120 percent and greater-than-120 percent categories. Figure 2 shows the results of the extension. Far from appearing to be a break, average real GDP growth in the category of public debt/GDP between 90 and 120 percent is 2.4 percent, reasonably close to the 3.2 percent GDP growth in the 60–90 percent category. GDP growth in the new category between 120 and 150 percent is lower at 1.6 percent but does not fall off a nonlinear cliff. Equally significant, as Figure 2 shows, variation in real GDP growth within each public debt/GDP category is large.



In Figure 3, we present a scatterplot of all of the country-years with continuous real GDP growth plotted against public debt/GDP ratio and include a locally fitted regression function.[7] No particular boundary or non-linearity is evident in either dimension around public debt/GDP of 90 percent. The data thin out gradually between 70 and 120 percent as is visible from the points in the scatterplot and the widening 95 percent confidence interval for mean growth. More generally, the wide range of GDP growth at various publicdebt levels is evident.

Finally, the scatterplot does suggest a non-linearity in the relationship, but that occurs in the change in the public debt/GDP ratio from 0 to 30 percent. This contradicts RR's claim that "it is evident that there is no obvious link between debt and growth until public debt reaches a threshold of 90 percent" (Reinhart and Rogoff 2010a, 575). Figure 4, which is a close-up of Figure 3 shows more clearly that average growth declines sharply with public debt/GDP between 0 and 30 percent; at 0 percent debt/GDP, average growth is almost 5 percent and by 30 percent it has declined to slightly more than 3 percent. The relationship between average GDP growth and public debt/GDP is relatively flat over a wide domain of debt/GDP values. Between public debt/GDP ratios of 38 percent and 117 percent, we cannot reject a null hypothesis that average real GDP growth is 3 percent.

In Table 4, we present regression analysis of real GDP growth by public debt/GDP category. The first row in both columns confirms significantly and substantively higher growth rates in the lowest 0{30 percent public debt/GDP category relative to other public debt/GDP categories.8 The results show modest differences among the other categories. In the first column, average GDP growth in the category of public debt/GDP above 90 percent is lower by about 1 percentage point than GDP growth in the 30–60 percent and 60–90 percent public debt/GDP categories. In the second column, average GDP growth in the category of public debt/GDP above 120 percent is substantially lower than GDP growth in the 30–60 percent and 60–90 percent debt/GDP categories. However, in the second column, which includes the new above-120-percent public

[7]The locally smoothed regression function is estimated with the general additive model with integrated smoothness estimation using the `mgcv` package in R. The smoothing parameter is selected with the default cross-validation

debt/GDP category, differences in average GDP growth in the categories 30–60 percent, 60–90 percent, and 90–120 percent cannot be statistically distinguished. An F-test on the hypothesis that, relative to the 30–60 category, the 60–90 difference and the 90–120 differences are both zero cannot be rejected (p-value = 0.11). To summarize, the regression results show that there is a non-linearity in the relationship between GDP growth and public debt between public debt levels of 0 to 30 percent of GDP. The results also indicate that average GDP growth tails off somewhat when the public debt/GDP ratio increases towards 120 percent, but there is no sharp turning point.

Thus, the non-linearity in the relationship between public debt levels and GDP growth is not around a public debt/GDP ratio of 90 percent where RR have identified it. That is, the non-linearity is not in the domain of public debt/GDP values that is currently the focus of policy debate in the US and Europe.

Different results by period

We further explore the historical specificity of the result by examining average real GDP growth by public debt category for subsampled periods of the data. Table 5 presents results for 1950–2009, 1960–2009, 1970–2009, 1980–2009, 1990–2009, and 2000–2009. We see that the high GDP growth in the lowest public debt/GDP category erodes substantially in the shorter more recent periods. Thus, in the lowest, 0–30-percent public debt/GDP, GDP growth of 4.1 percent per year in the 1950–2009 sample declines to only 2.5 percent per year in the 1980–2009 sample. Growth in the middle two public debt/GDP categories also decelerates noticeably, with the average dropping by more than a percentage point in the

samples limited to later years. In contrast, average growth in the highest debt/GDP category is quite stable across all samples of years, remaining within 0.3 percentage points of 2 percent per year throughout. In recent years, real GDP growth in the highest, above-90-percent public debt/GDP category has outperformed that in the next highest category. These patterns suggest two important conclusions: (1) even the apparent non-linearity between the lowest-debt country-years and higher-debt country-years is an historically specific pattern, not a robust result across the full time period; and (2) the relationship between public debt and GDP growth is weaker in more recent years relative to the earlier years of the sample.

Conclusion

The influence of RR's findings comes from its straightforward, intuitive use of data to construct a stylized fact characterizing the relationship between public debt and GDP growth for a range of national economies. However, this laudable effort at clarity notwithstanding, RR has made significant errors in reaching the conclusion that countries facing public debt to GDP ratios above 90 percent will experience a major decline in GDP growth.[8] The key identified errors in RR, including spreadsheet errors, omission of available data, weighting, and transcription, reduced the measured average GDP growth of countries in the high public debt category. The full extent of those errors transforms the reality of modestly diminished average GDP growth rates for countries carrying high levels of public debt into a false image that high public debt ratios inevitably entail sharp declines in GDP growth. Moreover, as we show, there is a wide range of GDP growth performances at every level of public debt among the 20 advanced economies that RR survey.

RR's incorrect stylized fact has contributed substantially to ensuring that "traditional debt management issues should be at the forefront of public policy concerns" (Reinhart and Rogoff (2010a), p. 578). Specifically, RR's findings have served as an intellectual bulwark in support of austerity politics. The fact that RR's findings are wrong should therefore lead us to reassess the austerity agenda itself in both Europe and the United States.

[8] For econometricians a lesson from the problems in RR is the advantages of reproducible code relative to working spreadsheets. We are grateful to Reinhart and Rogoff for sharing the working spreadsheet, and we will make our simplified version of the spreadsheet and R code that reproduces RR and corrected results available on our website.

References

Irons, John, and Josh Bivens. 2010. "Government Debt and Economic Growth." *Economic Policy Institute Briefing Paper* 271.

Reinhart, Carmen M, and Kenneth S Rogoff. 2010a. "Growth in a Time of Debt." *American Economic Review* 100 (2): 573–8.

———. 2010b. "Growth in a Time of Debt." National Bureau of Economic Research.

———. 2011. "A Decade of Debt." National Bureau of Economic Research.

Reinhart, Carmen M, Kenneth S Rogoff, and Miguel A Savastano. 2003. "Debt Intolerance." National Bureau of Economic Research.

Ryan, Paul. 2012. "The Path to Prosperity: A Blueprint for American Renewal." House Budget Committee.