Queen’s Economics Department Working Paper No. 1290

Identifying Fiscal Policy (In)effectiveness from the Differential Counter-Cyclicality of Government Spending in the Interwar Period

Nicolas-Guillaume Martineau
Universite de Sherbrooke

Gregor W. Smith
Queen‘s University

Department of Economics
Queen’s University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

5-2014

Identifying Fiscal Policy (In)effectiveness from the Differential Counter-Cyclicality of Government Spending in the Interwar Period

Nicolas-Guillaume Martineau and Gregor W. Smith

April 2014

Abstract

Differences across decades in the counter-cyclical stance of fiscal policy can identify whether the growth in government spending affects output growth and so speeds recovery from a recession. We study government-spending reaction functions from the 1920s and 1930s for twenty countries. There are two main findings. First, surprisingly, government spending was less counter-cyclical in the 1930s than in the 1920s. Second, the growth of government spending did not have a significant effect on output growth, so that there is little evidence that this feature of fiscal policy played a stabilizing role in the interwar period.

Des différences dans le degré de contra-cyclicité de la politique budgétaire peuvent servir à identifier si la croissance des dépenses publiques a un effet sur la croissance de l'output, et accélère ainsi la reprise de l'activité suivant une récession. Nous étudions des fonctions de réaction des dépenses publiques pour les années 1920 et 1930, et ce pour vingt pays. Il en ressort deux résultats principaux. Premièrement, et fait surprenant, les dépenses publiques étaient moins contra-cycliques durant les années 1930 que 1920. Deuxièmement, la croissance des dépenses publiques n'a pas eu d'effet significatif sur la croissance de l'output, ce qui permet de douter que cet aspect de la politique budgétaire ait joué un rôle stabilisateur durant l'entre-guerre.

JEL classification: E62, N10

Keywords: fiscal policy, business-cycle history, Great Depression, interwar economy

Corresponding author: Nicolas-Guillaume Martineau, Department of Economics, Glendon College, York University; ngmartineau@gmail.com; (819) 919–0899

1. Introduction

Did the international experience of the interwar period show that counter-cyclical government spending hastened recovery from the Great Depression? Answering this question is an ambitious task, but we try to contribute to the answer by studying a large (for the time) panel of twenty countries from 1920 to 1939. An advantage of this breadth and of studying this time period is that there is great heterogeneity over time and countries both in the business cycle and in the stance of fiscal policy. This variation in cycles and policies should help us identify the role of fiscal policy. But a disadvantage of such a panel is that many macroeconomic measures—including consumption and real wages for example—are not available. So this approach is a complement to studies of individual countries, like the US and UK, where such series are available.

The limitations of the data, in frequency and coverage, may prevent us from reaching a precise answer about the efficacy of conter-cyclical government spending, but it is still of interest to know whether that is the conclusion. Of course, the answer and its precision depend on an identification scheme. This paper adopts a new one: the main identifying assumption is that government spending could have increased output in any country but was not tried to the same extent at every time. Identification relies on a difference over time in government spending reaction functions that capture the response of government spending to national income. We use this difference to estimate the effect of government spending on the growth of income in turn.

Of course, business cycles in output also were affected by other shocks besides fiscal policy changes. One example is the decline in US output during the Great Depression, transmitted to these countries as a decline in export demand. A second example is the stance of domestic monetary policy, and specifically the timing of leaving the gold standard. We control for these shocks in trying to isolate the effect of changes in government spending. In fact, the identification depends on such observable, exogenous shocks, because output and government spending are treated as endogenous to each economy. There is a significant impact of US output growth on the business cycles of other countries, so it is not the case that no statistical relationship can be detected because of the limitations of the historical data.

The paper measures a correlation between the growth of government spending and the growth of real output, controlling both for other influences on output and for the endogeneity of government spending. The idea is that this partial correlation should be present if government spending had a significant, macroeconomic effect, whatever the mechanism by which this occurred. We find a statistically significant difference in government-spending reaction functions between the 1920s and the 1903s, with policy becoming *less* counter-cyclical in the latter decade. But there is little evidence that government spending growth affected output growth.

2. Data and Design

The design of the study largely stems from the availability of data. We study twenty economies for the period from 1920 to 1939: Argentina, Australia, Austria, Belgium, Brazil, Canada, Czechoslovakia, Denmark, Finland, France, Germany, Hungary, Italy, Japan, the Netherlands, Norway, Portugal, Spain, Sweden, and Switzerland. The UK and US are included as influences on business cycles within this group, though we do not try to identify fiscal policy efficacy using their own policies.

We restrict the investigation to the interwar period for several reasons. First, data for the period before 1920 are even scarcer and of course would raise the issue of how to model wartime spending during 1914–1918. Second, data for the period after 1939 would raise the same issue for 1939–1945. The focus is on whether there is evidence for the macroeconomic effects of government spending solely for the interwar period, rather than for long spans or later episodes. One advantage of this approach is that some research suggests the impact of government-spending changes is greater when interest rates are low and unemployment is high, features which were characteristic of many countries in the 1930s.

The data are at annual frequency. We selected all countries for which we could find measures of national income (GDP or GNP) and government spending. For most countries the government spending measure comes from public accounts rather than national accounts and so includes some transfers. We also use measures of defence spending collected by the League of Nations. The appendix contains details of the sources.

The government spending series we adopt begin in 1924 for Austria, Germany, and Hungary. Periods of hyperinflation thus are omitted from this study. Though it would certainly be interesting to study the measurement and impact of real government spending during hyperinflations, we cannot pursue those questions given the available data.

Figure 1 shows real national income per capita for five countries, scaled so that 1920=100. The figure shows the countries with the most and least growth in the interwar period, as well as two intermediate cases including Canada. The great heterogeneity in growth rates is evident, as countries such as Portugal and Japan grew, albeit with cycles, while Austria stagnated. The very different paths for Spain and Canada during the 1930s show that there was not simply a single, world cycle.

[Figure 1 about here]

Figure 2 shows that there also was heterogeneity in real government spending per capita, where again we show the extreme cases and two intermediate ones. The role of the state expanded enormously in Japan while government spending also grew significantly in Sweden late in the 1930s. Real government spending per capita was relatively constant in Canada, while it declined in Czechoslovakia.

[Figure 2 about here]

Figure 3 shows government spending as a percentage share of national output for Canada, France, Italy, and Japan. This subset of countries shows that there is also much variation across countries in the time path of this ratio. The ordering of paths for output is not merely duplicated in the ordering of paths for government spending, a feature which should aid identification. Figure 3 shows that there were some sharp increases in government spending as a share of GDP, for example in Italy in 1936 and in France in 1939. There also were some sharp decreases, such as those in France and other European countries (including Czechoslovakia and Belgium) during the early 1920s. Notice also that the overall levels of these shares are in some cases very high; one probably cannot argue that the government sector was too small to have affected the economy.

[Figure 3 about here]

It is worth noting some macroeconomic variables that are *not* available except for very narrow panels of countries: consumption, hours, markups, average tax rates, real wages and employment, and country-weights in international trade. We also lack complete data on the composition of government spending. For example, the League of Nations yearbooks report transfers to persons only for 1925–1929. We also considered unemployment rates as an alternate business-cycle indicator, but found that they were available for many fewer country-year combinations than was true of output. Largely because of these missing macroeconomic data, we are not able to isolate the mechanism by which government spending may have affected the path of national income. For example, we do not know whether it affected the labour market via sticky wages, a wealth effect on labour supply, or a signal of future monetary expansion.

Data are available on tax revenue and hence on budget deficits. They are not included directly in the statistical model for several reasons. First, a benchmark model of the effects of tax timing—Ricardian equivalence—holds that the deficit may have any correlation with output. This absence of a prediction contrasts with the case of government spending, where a range of macroeconomic models predict a positive impact on output. Second, we generally do not have data on tax rates and their variation over time and across countries, which would be necessary for the careful study of the impact of this aspect of fiscal policy.

We thus focus on government spending, like much recent research on post–1945 data. However, we do examine whether the impact of government spending varies with the level of the government budget deficit.

Overall the design of the study is to include as many countries as possible, particularly given the low frequency of the available data. We then study the simultaneous determination of government spending and national income within this panel. The main aim is to investigate whether the heterogeneity in government-spending reactions during the interwar period can help us assess the impact of government spending on output. We next show how this heterogeneity can potentially provide identification.

3. Identification

Let t count years and i count countries. Let y_{it} denote the growth rate of real output and g_{it} denote the growth rate of real government spending. Let x_t denote a weakly exogenous variable, such as output growth in the US or the UK. We study mainly small and medium-sized economies and so take US and UK output growth as given. We also discuss alternative exogenous variables below.

Then consider a statistical model of domestic output growth as depending on the growth rate of government spending and on the exogenous variable:

$$y_{it} = \delta_{yi} + \beta g_{it} + \omega x_t + \epsilon_{yit}. \tag{1}$$

The parameter β measures the impact of the growth in government spending on the growth of output while the parameter ω measures the impact of the exogenous variable x_t. The shock to output growth, ϵ_{yit}, captures country-specific events such as the Japanese earthquake of 1923, the Credit Anstalt failure in Austria in 1931, or the French strikes of 1936.

Suppose that the reaction function for government spending is:

$$g_{it} = \delta_{gi} + \alpha y_{it} + \epsilon_{git}. \tag{2}$$

The parameter α measures the response of government spending to the domestic business cycle. A negative value, for example, describes a counter-cyclical policy. Notice that β, ω, and α are common across countries.

The intercepts δ_{yi} and δ_{gi} can differ across countries and across variables. The parameter δ_{yi} may reflect the trend in productivity or capital accumulation in country i. The corresponding intercept for government spending, δ_{gi}, may differ from that value. This difference can then capture the secular growth of the state in Japan during the interwar period, for example. By allowing for such growth in the output share of government spending we do not falsely confuse that with cyclical policy. Thus, if both intercepts are higher for Japan than for other countries we would not interpret that information as demonstrating the effects of fiscal policy.

The system (1)–(2) is in growth rates to provide stationarity. Combining growth rates and country-specific intercepts is a reasonable way to statistically model a short panel, here with only twenty annual observations. We omit lags, first to conserve degrees of freedom and second because a dynamic panel-data model with country-specific fixed effects would require further instruments for consistency, which again may be challenging. A third reason is that there are generally some missing observations in a historical panel like this one. For example, we do not have government spending data during the Spanish civil war or during some of the central European hyperinflations of the early 1920s. A model of growth rates, as opposed to one with higher-order dynamics, gives stationarity but minimizes the loss of evidence from these missing data. We thus leave exploring added dynamics for further research with larger data sets. In that case, one also could apply the identification scheme to innovations instead of to growth rates.

Estimating the parameters requires statistical assumptions. We begin with the natural assumptions that the shocks have mean zero in each country and are uncorrelated with the exogenous variable x_t. However, the shocks ϵ_{yit} can be correlated across groups of the twenty countries, or there can be country-specific shocks.

In the system (1)-(2) there obviously would be simultaneity bias in OLS estimation of either equation. Before elaborating on a consistent estimator, we should comment on the exclusion restrictions in the system. First, x_t is excluded from the g-equation (2): foreign output growth affects domestic output growth but does not directly affect government-spending growth. The idea is that the former effect stems from trade or financial linkages or from a common component in productivity shocks across countries, whereas it is difficult to see a reason why domestic government spending should respond to the foreign business cycle. Second, while x shifts the y-equation (1) there is no corresponding shift variable exclusively in the g-equation (2). It is challenging to measure the exogenous component of elections, for example, and it seems unlikely that they would not also affect output. For this same reason any correlation between the two shocks is admissible; we do not assume $\text{cov}(\epsilon_{yit}, \epsilon_{git}) = 0$.

An informative way to study identification is to solve the system to give the reduced form:

$$\begin{aligned} y_{it} &= \frac{\delta_{yi} + \beta\delta_{gi}}{1-\beta\alpha} + \frac{\omega}{1-\beta\alpha} x_t + \frac{\beta}{1-\beta\alpha}\epsilon_{git} + \frac{1}{1-\beta\alpha}\epsilon_{yit} \\ g_{it} &= \frac{\delta_{gi} + \alpha\delta_{yi}}{1-\beta\alpha} + \alpha_i \frac{\omega}{1-\beta\alpha} x_t + \big(1 + \alpha\frac{\beta}{1-\beta\alpha}\big)\epsilon_{git} + \frac{\alpha}{1-\beta\alpha}\epsilon_{yit}. \end{aligned} \tag{3}$$

The composite intercepts and error terms play no role in our identification so we focus on the slopes, which we denote b_{yi} and b_{gi}, with:

$$\begin{aligned} b_{yi} &= \frac{\omega}{1-\beta\alpha} \\ b_{gi} &= \frac{\alpha\omega}{1-\beta\alpha}. \end{aligned} \tag{4}$$

It is easy to see that $\alpha = b_{gi}/b_{yi}$ and thus to see the textbook result that only α is identified from the two reduced-form slopes. The exclusion of x_t from the reaction function (2) allows α to be identified, but one cannot measure β, the impact of the government-spending growth rates on the output growth rates y_{it}.

The reduced-form equations (3) also show that the correlation between y and g does not tell one about the effectiveness of fiscal policy. First, any correlation between the shocks ϵ_{yit} and ϵ_{git} is possible. Second, even if the shocks are uncorrelated, a positive correlation between output growth, y_{it}, and government-spending growth, g_{it}, could reflect either (a) $\beta = 0$ and $\alpha > 0$ as growth allowed an expansion of the role of the state, or (b) $\beta > 0$ and $\alpha < 0$, with large enough shocks ϵ_{git}, where government spending growth does increase output growth.

Our identification of β relies on the heterogeneity of the reaction function over time, combined with the assumption of a common fiscal-policy impact β and a common, non-zero weight ω on the exogenous variable. The impact of changes in government spending or changes in foreign output growth on the domestic business cycle is the same over time and across countries. But the reaction function parameter, α, differs across time periods. This is the differential counter-cyclicality referred to in the title of the paper.

To see identification it is enough to consider two disjoint time periods, labelled r and s, with different reaction-function parameters α_r and α_s. There now are four reduced-form slopes and four parameters, β, ω, α_r, and α_s, so this necessary condition for identification is satisfied. The four reduced-form slopes are:

$$\begin{aligned} b_{yr} &= \frac{\omega}{1-\beta\alpha_r} \\ b_{gr} &= \frac{\alpha_r\omega}{1-\beta\alpha_r}, \end{aligned} \tag{5}$$

with similar equations in time period s. It is straightforward to show that:

$$\begin{aligned} \alpha_r &= \frac{b_{gr}}{b_{yr}} \\ \alpha_s &= \frac{b_{gs}}{b_{ys}} \end{aligned} \tag{6}$$

It is almost as straightforward to show that:

$$\begin{aligned}\beta &= \frac{b_{yr} - b_{ys}}{b_{gr} - b_{gs}} \\ \omega &= \frac{b_{gr}b_{ys} - b_{yr}b_{gs}}{b_{gs} - b_{gr}},\end{aligned} \tag{7}$$

so that the parameters are just identified.

We can estimate β either from the reduced form or by instrumental variables. Let d_i be a dummy variable for country i. Let d_r and d_s be dummy variables that select the specific time periods r and s. Then the population moment conditions are:

$$\mathrm{E}(y_{it} - \delta_{yi} - \beta g_{it} - \omega x_t | d_i, d_r x_t, d_s x_t) = 0, \tag{8}$$

which simply require that the y-equation residuals have mean zero and are uncorrelated with x_t, period by period. The instruments d_i obviously identify δ_{yi}. To show that the other two moment conditions identify β and ω, ignore the intercepts and write the conditions (8) as:

$$\begin{aligned}\mathrm{E}y_{it}d_r x_t - \beta \mathrm{E}g_{it}d_r x_t - \omega \mathrm{E}d_r x_t^2 = 0 \\ \mathrm{E}y_{it}d_s x_t - \beta \mathrm{E}g_{it}d_s x_t - \omega \mathrm{E}d_s x_t^2 = 0.\end{aligned} \tag{9}$$

These conditions combine to give:

$$\begin{aligned}\beta &= \frac{\mathrm{E}y_{it}d_r x_t/\mathrm{E}d_r x_t^2 - \mathrm{E}y_{it}d_s x_t/\mathrm{E}d_s x_t^2}{\mathrm{E}g_{it}d_r x_t/\mathrm{E}d_r x_t^2 - \mathrm{E}g_{it}d_s x_t/\mathrm{E}d_s x_t^2} \\ &= \frac{b_{yr} - b_{ys}}{b_{gr} - b_{gs}},\end{aligned} \tag{10}$$

which reproduces our finding (7). This rewriting thus shows that the moment conditions (8) use all the information on β from the solved, reduced-form system: having as many moments (in other words, time periods with different values of α) as parameters is necessary and sufficient for identifying β.

Equation (8) shows that the instruments are cross-products or interaction terms: $d_r x_t, d_s x_t$. That also shows why we cannot simply replace x_t with a set of time effects (year-specific dummy variables) common to all countries. In that case the two cross-products would simply reproduce the original vector of time effects, adding no new instrument.

Figure 4 gives the intuition behind our identification. The upward-sloping line shows the effect of government spending growth on output growth, with slope β common to each time period and country. The two downward-sloping lines show reaction functions, one for period r and one for period s. A single, observable shock x_t shifts the upward-sloping,

y-equation line in the same way for both time periods. That shift traces out the reaction functions and so identifies their slopes α_r and α_s. But connecting the dots after such a shock, in other words pooling data from the two time periods, also identifies the slope β, as figure 4 shows.

[Figure 4 about here]

A reduced form with slopes that vary over time also could result from a structure in which the stance of fiscal policy α was constant but its impact or efficacy, β varied. But interpreting the reduced form as we have done is consistent with the way economists generally have tried to draw lessons from the macroeconomic experience of the Great Depression. For example, some economists and historians implicitly argue that counter-cyclical policy (a) was tried at some times but not others and (b) could have been effective at all times. We try to study the implications of this perspective.

It also is important to note that our approach does not require that governments pursued public works counter-cyclically with stabilization in mind. Any component of government spending whose cyclicality differed over time can provide identification. For example, increased defence spending in Japan after 1931, in Germany after 1933, or in Italy associated with the invasion of Abyssinia in 1936 may provide identifying information.

The impact parameter β also can be identified if α varied not over time but instead over countries, with some countries (such as Sweden perhaps) pursuing a more counter-cyclical policy than others. For example, Japan's fiscal expansion in 1931 while Takahashi was finance minister contrasts with the fiscal retrenchment in Germany under Chancellor Brüning at the same time, a difference that again may aid identification. The key feature is that these differences must be partly in the systematic component of government spending growth, linked to output growth by the parameter α; they cannot simply be differences in shocks ϵ_{git}. The working paper version of this paper (Martineau and Smith, 2012) explored this possibility. However, we did not find statistical evidence of systematic variation across countries, so we do not stress that potential source of identification here.

To briefly formalize the identification requirements, the first requirement is that α varies across time or countries. Denote by n the cardinality (*i.e.* number of distinct elements) of a vector of parameters. Then we need:

$$n(\alpha_{it}) \geq 2. \tag{11}$$

At least two distinct values of α are necessary for identification. Of course, further distinct values of α_{it} provide over-identification. We do not show that result formally because it is a straightforward extension of the earlier algebra. Section 5 below provides evidence on this condition.

Now suppose that there are H variables in x_t. The second requirement is that:

$$n(\beta_{it}) + n(\omega_{it}) \leq n(\alpha_{it}) \times H. \tag{12}$$

This is simply the usual method-of-moments requirement that the number of moments in conditions (8) must be greater than or equal to the number of parameters. For example, when there are two reaction-function parameters, $n(\alpha_{it}) = 2$, and a single exogenous variable, $H = 1$, one can identify β and ω. Thus whenever $n(\alpha_{it}) > 2$ or $H > 1$ one can either relax some restrictions on the constancy of β_{it} and ω_{it} over time and countries or else gain efficiency and test the over-identifying restrictions.

Notice from the second requirement (12) that a break over time in α in country i would allow identification country-by-country, in other words estimation of country-specific values β_i and ω_i. But country-specific values of α_i that vary over time—say from the 1920s to the 1930s—will be very imprecisely estimated with nine or ten annual observations on growth rates per decade, leading to the syndrome of weak instruments. Then estimates of β_i also will be very imprecise. To avoid this syndrome, we estimate a common β across countries. In practice we find a single, significant break in α over time for countries studied as a group, so that $n(\alpha_{it}) = 2$, which limits the diversity in β and ω we can consider.

In sum, if the government-spending reaction varied significantly across countries or over time then there is scope for estimating the impact of government-spending growth on output growth. Identification comes from differences in the systematic component of fiscal policy, not from trying to isolate properties of exogenous changes in government spending. The next section pauses to compare our method to those in recent research on fiscal policy's macroeconomic effects. It also gives references to previous research on fiscal policy specifically in the interwar period.

4. Research Context

This section briefly outlines research on the macroeconomic effects of changes in government spending. One aim is to explain why we do not use either a DSGE model or a structural VAR model with unrestricted lags, which are the standard approaches. A second aim is to give the reader some benchmarks to which to compare our empirical findings below.

4.1 Government Spending in Recessions

Isolating the effects of changes in government spending can be challenging even with contemporary, quarterly data. Even for the US, where a long span of data is available, there is ongoing debate about whether conclusions about the effect of fiscal policy depend on specific, wartime episodes. Auerbach, Gale, and Harris (2010), Hall (2009), and Ramey (2011) review methods and findings from VAR methods and DSGE models as well as from microeconomic data.

DSGE modelling is challenging for the interwar period because we lack data on measures such as hours, real wages, markups, tax rates, and consumption for the panel of countries. Such measures are needed to reliably calibrate the models and to measure shocks. For example, Cole and Ohanian (2011) find one can construct Solow residuals only for five countries in our panel (plus the UK and US) because of the lack of employment data.

In structural VAR models, identification of government-spending shocks can come from (a) timing (*e.g.* the restriction that discretionary spending does not respond to the cycle within a quarter), (b) sign restrictions, or (c) the narrative approach (including using military spending). By studying countries with long spans of both quarterly and annual data, Born and Müller (2012) show that timing restrictions may well hold for government spending on goods and services even in annual data. There remain two obstacles to applying a VAR with a timing restriction for the interwar period, though. First, government spending for most countries is defined broadly to include transfers, which may well respond to the state of the cycle within the year. Second, there are too few time-series observations to allow reliable inferences that depend on country-specific dynamics.

Hall (2009) reviews multiplier findings from SVAR studies for the US. He shows that it is difficult to be precise about the impacts because there was not much variation in government spending except during World War II and the Korean War. Barro and Redlick (2011) study a long time series for the US to exploit this sort of variation. In a similar vein, we exploit the variation across decades during the interwar period. Their equation for output is similar to ours—output growth explained in part by growth in government purchases—but they also can include measures of expected growth in government spending and of tax rate changes; we do not have those series for these twenty countries. Barro and Redlick also argue that non-defence spending likely responds to the state of the economy; we use the variation in this response across decades to give identification.

A broad theme of much empirical research on the macroeconomic effects of government spending is that the impact of policy changes may be a variable that depends on such features as (a) the composition of spending (*e.g.* on infrastructure or transfers), (b) the financing (*e.g.* the timing of distorting taxes or the extent to which government spending signals future monetization), (c) the expected persistence of the policy. Unfortunately, most of these issues also cannot easily be studied for the interwar period because of data limitations. However, we do study (in section 8) whether the impact of government spending changes was larger when accompanied by a budget deficit.

This research also suggests that the impact of policy changes depends on (d) the state of the business cycle and (e) the presence of a liquidity trap. For example, Auerbach and Gorodnichenko (2012) find for the US that government spending multipliers are much larger in recessions than in expansions. Parker (2011) also argues that one would expect the effect of a shock to government spending to be larger in a recession and also larger in a liquidity trap. Our study focuses on a period that includes the Great Depression, when many countries of course experienced both very high unemployment and very low interest rates. Finally, in the cross-country dimension, Ilzetski, Mendoza, and Végh (2013) argue that the efficacy of fiscal policy has recently depended on (f) the exchange-rate regime. We also control for cross-country differences in monetary (exchange-rate) policy.

4.2 Government Spending in the Interwar Period

The effects of government spending during the interwar period have been studied most extensively for the US and the UK. Crafts and Fearon (2010) survey research on the 1930s economy, mainly on those two countries. The upper panel of figure 5 shows the levels of real output per capita for the UK (in black) and US (in gray) for the interwar period, scaled so that they begin at 100 in 1920. The differential experiences of these two countries during the Great Depression are well known, and reflected in the figure. By comparison with the business cycle in the UK, the depression in the US was much more severe and with a later trough (in 1933 rather than 1931 for the UK). The vertical scale is the same as in figure 1, which thus highlights the even greater heterogeneity of the experiences of the twenty countries we study.

[Figure 5 about here]

Fishback (2010) provides a comprehensive review of US monetary and fiscal policy in the 1930s and outlines applications of DSGE models and VAR methods. Among recent studies, Eggertsson (2008) argues that expected future policy changed in 1933, in the form of a change in inflation expectations. He suggests that one of the triggers for the change in expectations was the large increase in government spending in 1933. Overall for the interwar period, though, Fishback reaches the same conclusion as Romer (1992) or Brown (1956, p 863) that

> fiscal policy, then, seems to have been an unsuccessful recovery device in the 'thirties not because it did not work, but because it was not tried.

Middleton (2010) describes fiscal policy in the UK in the 1930s and reaches a similar conclusion. In contrast, Crafts and Mills (2012) use information from news about defence spending in the UK to argue that government-spending shocks were large (especially those associated with 1930s rearmament) but nevertheless that the fiscal multiplier was small.

However, some recent research for the US argues that increases in government spending did significantly affect the path of output in the very late 1930s and early 1940s. Gordon and Krenn (2010) conclude from a VAR model that government spending shocks played a large role in ending the US Great Depression after 1939 but before 1942. For the 1941–1945 period, McGrattan and Ohanian (2010) use a neoclassical DSGE model to measure the effects (for example in factor markets) of large government-spending shocks and conclude they had significant effects on output.

The lower panel of figure 5 shows government spending as a share of output in the UK (in black) and US (in gray). For the UK that share was very stable until 1938. For the US, the share rose gradually during the early 1930s, then levelled off, with a well-known dip in 1937 as some New Deal initiatives were reversed. The fiscal expansions in Italy or Japan in the 1930s were much more dramatic than that in the US. Our approach is based on the possibility that the extent to which counter-cyclical fiscal policy was tried varied across time. Of course, variation across time in the government spending share of output (as shown in figure 3) does not automatically imply variation in the activist response to output, but we measure that latter variation explicitly in the next section.

Formal, statistical analysis of the role of fiscal policy in individual countries (other than the US and UK) during the interwar period usually is infeasible because of the absence of data at greater than annual frequency. But there are some noteworthy exceptions to this rule, including studies that fit VARs. Ritschl (2002) reviews research on fiscal policy in Germany in the 1930s and also assesses the macroeconomic data, which are more detailed than for most countries in the panel we study. He estimates a VAR model and concludes that government budget deficits had no significant effect on the path of German output; rearmament simply crowded out private spending. Cha (2003) estimates a VAR using a range of monthly data for Japan from 1929 to 1936 and concludes that fiscal shocks did play a large role in its relatively rapid recovery from the Depression. Almunia, Bénétrix, Eichengreen, O'Rourke, and Rua (2010) study an international panel of twenty seven countries from 1925 to 1939 using an SVAR with a timing restriction. They find large effects of government spending shocks and thus conclude that fiscal policy may well have been effective had it been implemented more widely.

Histories of the Depression in small, open economies sometimes comment on the stance and possible impact of fiscal policy. Feinstein, Temin, and Toniolo (2008, pp 194–199) provide a guide to these histories. Examples include the works by Safarian (1970) for

Canada or Montgomery (1938) and Myrdal (1939) for Sweden. Safarian's (1970) book on the Great Depression in Canada developed a theme that has been echoed by other research; he attributed much of the downturn and recovery to events in the US and UK. Jonung (1979) and Grytten (2006) describe the post-war consensus view that Swedish fiscal policy was effective in reducing the scale of the Depression in Sweden. But they also concur that economists have been correct to question this consensus since the 1970s. Jonung noted that formal econometric analysis of the sources of recovery was lacking.

Even some contemporary observers doubted whether the Swedish *krispolitik* had a significant effect on recovery. Myrdal (1939) argued that counter-cyclical policy in Sweden was largely a failure in part "because this policy was carried out only half-heartedly" (p 183) with public works programs "of much smaller scope than would have been desirable" (p 184) and in part because of the "adverse reaction of business confidence, which has too often restricted or even possibly reversed its stimulating effects" (p 187). The share of government spending in GDP in Sweden rose from 8.2% in 1930 to 12.4% in 1933, though some other countries experienced sharper fiscal expansions.

Montgomery (1938), in what was surely one of the earliest books on recovery from the Depression, attributed Sweden's recovery to allowing the krona exchange rate to float (leaving the gold standard with the UK) and to foreign recovery, particularly in the UK. On the effect of fiscal policy changes, however, Montgomery (1938, p 67) also noted:

> This question can best be answered by the use of the comparative method. For the purposes of an enquiry of this kind we find the ideal type when those countries which are to be compared agree in other respects but differ in that particular point the significance of which we intend to investigate. We can hardly expect to come across such ideal types in the world of reality ...

But we may come across types as close to ideal as possible in this group of countries, if the degree of counter-cyclicality of government spending varied systematically over time.

Our potential contributions are based on using this variation over time in government-spending reaction to assess efficacy (rather than measuring shocks) and on conditioning on business cycles in the US and UK to provide instruments. We do not apply a VAR given the short span of annual data. To see reason for skepticism about that method in these data, recall the Bartlett standard error for autocorrelation coefficients, with T observations, which is $1/\sqrt{T}$. With nineteen observations per country this formula implies that only autocorrelations above 0.45 will yield t-statistics above 1.96. (With the fourteen observations used by Almunia *et al* only autocorrelations above 0.52 would be so significant.) Thus we study the short panel using growth rates, with country-specific intercepts (and we also test for residual autocorrelation).

5. Differences in Government-Spending Reaction across Decades

As section 3 outlined, identifying β, the impact of government-spending growth on income growth, requires variation in the reaction-function coefficient α, either over time or across countries. To test for such variation, we ran the reduced-form regressions:

$$g_{it} = b_{0i} + b_g x_t + e_{it}, \tag{13}$$

where e_{it} is an OLS residual. According to the model's reduced form (3), the slope b_g is a composite of α, β, and ω. Thus, finding that this slope varies across decades is necessary for measuring variation in α, which in turn is necessary for identification.

We examined variation in the reduced-form slope coefficient across decades, allowing b_g to take on different values for the 1920s and 1930s, while taking the same value for all countries within each decade. We measured x_t first as US output growth, y_{ut}, and second as a vector including UK output growth, y_{kt}. In both cases the F-statistic testing the restriction to a common slope across decades had a p-value of 0.00, readily rejecting the restriction. These highly significant, first-stage F-statistics show that we have a strong instrument. In both cases the coefficient b_g rises from the 1920s to the 1930s, suggesting that there was *less* counter-cyclicality in government spending in the 1930s. We also allowed for variation across *both* countries and decades, and tested the restriction to a common reduced-form coefficient across countries that varied only by decade. This restriction was readily accepted, as the test statistic had a p-value of 0.83.

The reduced form does not identify α, so we also estimated the moment conditions that do so. Define d_i as a dummy variable for country i, d_{1920s} as a dummy variable for the 1920s, and d_{1930s} as a dummy variable for the 1930s. We estimated the g-equation:

$$g_{it} = \delta_{gi} + \alpha y_{it} + \epsilon_{git}, \tag{14}$$

with instruments $z_{it} = \{d_i, d_{1920s} y_{ut}, d_{1930s} y_{ut}\}$, to estimate a common reaction for government spending for the 1920s that differs from the reaction for the 1930s. We found a significant difference between the two coefficients. That difference is sufficient to identify β, but it also is of independent interest. Notably, when we estimate the reaction function (14) pooled across countries we find a large, significant *increase* in $\hat{\alpha}$ from the 1920s to the 1930s, so that on average government spending indeed became less counter-cyclical. The estimate (with standard error in brackets) rises from -1.33 (0.55) in the 1920s to 0.49 (0.41) in the 1930s. The difference between these two values of α is 1.83 (0.62), a shift which provides identifying information for β.

We also examined government-spending reactions that differed across countries. For example, we tested the null hypothesis that $b_{gi} = b_g$, a common value for all twenty countries. We could not reject this hypothesis. In other words, the cross-country variation in government-spending reaction functions is not sufficient to provide strong instruments and reliably identify β.

Out of historical interest, we also detoured to test whether the reaction-function coefficient appeared to take on a country-specific value α_i for a single country while all other countries had a common value. According to these estimates, Finland and Portugal had government-spending growth rates which were significantly more pro-cyclical than those of the remaining countries, while the Netherlands and Spain had more counter-cyclical paths. Some assessments by contemporary observers and historians fit with this econometric evidence. Montgomery (1938) noted that Finland's fiscal policy stance was more pro-cyclical than Sweden's during the 1930s. Fonseca Marinheiro (2006) observed that fiscal policy in Portugal was also quite pro-cyclical due to the fiscal austerity measures administered by Salazar, who was finance minister with full powers after 1928. For the Netherlands, van Zanden (1998, pp 112–113) described how public works spending and unemployment benefits were introduced through the government capital budget leading to "*de jure* financial orthodoxy (the ordinary budget as presented to parliament was balanced) and *de facto* large deficits."

We next present results from the statistical model outlined in section 3. We base identification on cross-decade differences in the government-spending reaction parameter, α, and treat output growth in the US and UK as exogenous. Then we consider controlling for exchange-rate changes, allowing β to vary across sub-samples, and estimating the impact of the growth in defence spending.

6. Government Spending Impact 1921–1939

Our benchmark, statistical model conditions on output growth in the US, y_{ut}, and output growth in the UK, y_{kt}, which thus compose x_t for all countries. Including these as exogenous variables is designed to capture the effect of the business cycles in the US and UK on the demand for exports from the countries in the panel. It may also capture an effect of the US and UK business cycles that operated through financial linkages. These foreign variables also can capture the presence of a common, world shock. A good example is the banking crisis of 1931 that occurred in Austria, Germany, the US, and other countries. We lack a good financial crisis indicator for each country, but these variables may reflect that common shock. Finally, this feature of the statistical model also may reflect a common technology shock that we cannot observe directly. For example, Betts, Bordo, and Redish (1996) find evidence of a large, real-side shock, common to both the US and Canada in the 1930s.

We estimated the output-growth equation:

$$y_{it} = \delta_{yi} + \beta g_{it} + \omega x_t + \epsilon_{yit}, \tag{15}$$

with instruments $z_{it} = \{d_i, d_{1920s}x_t, d_{1930s}x_t\}$. Any correlation between ϵ_{yit} and ϵ_{git} is possible, so there are no cross-equation restrictions between the y-equations (15) and the g-equations (14) which thus can be estimated separately with no loss of efficiency. Estimation is by continuously-updated GMM and standard errors are robust to heteroskedasticity.

Section 5 found evidence of a single break in α across decades, and the instruments reflect that first-stage evidence. The two values of α provide $H \times 2$ restrictions, where H is the number of variables in x. Thus if $x_t = y_{ut}$ then we can identify two slope parameters that are common across countries, β and ω_u, while if $x_t = \{y_{ut}, y_{kt}\}$ then we can identify three slope parameters, β, ω_u, and ω_k, and conduct a test based on a single over-identifying restriction.

Notice that we do not report results that use instruments $d_i d_{1920s} x_t$ and $d_i d_{1930s} x_t$, where the model's restrictions hold separately by each country and decade. Adopting that instrument set would operationalize the idea that there are twenty country-specific breaks in α across the two decades. But tests for such breaks obviously would lack power with only nineteen observations per country and thus would yield weak instruments. Instead, we base estimation only on the property that the pooled reaction function varied by decade, a variation which section 5 showed provides a strong instrument.

Table 1 shows results. In the first row $x_t = y_{ut}$ and both β and ω are just-identified. The estimated impact of government-spending growth is positive but statistically insignificant at conventional levels of significance. But this is not because nothing is significant; the impact of US output growth on domestic output growth, measured by $\hat{\omega}_u$, is positive and significant at the 1% level. Naturally, a significant x-variable is necessary for identifying β, and table 1 shows that y_{ut} qualifies. The second row of table 1 shows results when we also allow for the impact of output growth in the UK. The central results on the lack of impact of government-spending policy and the significant impact of US output growth are unchanged. And we do not find a significant effect of UK output growth once we control for US output growth. We also inspected the twenty residual series visually for evidence of autocorrelation. There is no such evidence, though perhaps no test will be very informative with this span of annual data.

Ideally, we also would allow for coefficients on US output growth that vary across countries. That variation might reflect geography, comparative advantage, or other causes of differential trade links to the US. It also could capture differences in exposure to the financial crisis of the early 1930s. But identifying such variation along with β would require more instruments, for example from significant variation in the reaction-function coefficients, α_i, across countries, something which the first-stage tests did not support.

One possible explanation for a small value of $\hat{\beta}$ in table 1 is that there is measurement error in the historical data, leading to a familiar attenuation bias. But note that table 1 does find a statistically significant role for US output growth, so we do not find that the data are simply too few and too noisy to detect any correlations. Also, there may be less measurement error in government spending (from public accounts) than in national income. The appendix outlines the sources for both series.

So far, then, there is information in the variation across decades in government-spending reaction. The change is in the direction of *less* counter-cyclicality during the 1930s. Exploiting this change to identify the effectiveness parameter β leads to the conclusion that it is insignificantly different from zero. These findings of course also depend on both the conditioning variables x_t. We next consider additional x-variables and an alternative measure of government spending.

7. Conditioning on Exchange-Rate Changes

We next control for an indicator of monetary policy. Choudri and Kochin (1980) reported on the insulating properties of floating exchange rates for European countries such as Spain during the Great Depression. Eichengreen and Sachs (1985), Eichengreen (1992), and Bernanke (1995) have argued persuasively that the event of leaving the gold standard can be treated as exogenous and that it hastened recovery in a broad cross-section of countries. Eichengreen (2004) provided a comprehensive review and interpretation of research on the international Great Depression that also supports this idea. While their studies focus on the 1930s, there also was great heterogeneity in monetary policy in the 1920s. For example, the depth of the UK depression in the early 1920s often has been attributed to the Bank of England's restrictive monetary policy aimed at restoring the gold standard at pre-war parity. Given this research, we next control for exchange-rate changes, viewed as an indicator of monetary policy for the interwar period, by including them as an x-variable when we try to measure the impact of changes in government spending.

We first included the growth rate of the local-currency price of gold. Thus large values reflect loose monetary policy. This variable may capture the well-known effect that reflation of the domestic economy was the means by which monetary policy accelerated recovery. Section 2 noted that government spending series for Austria, Germany, and Hungary begin in 1924, after their hyperinflations. We thus omitted these extreme observations in testing for this effect and used only the depreciations of the new currencies (the shilling, reichsmark, and pengő respectively) in these countries. Figure 6 shows some examples of the time series in levels, with large, cumulative depreciations in Portugal, Brazil, and Japan.

[Figure 6 about here]

We studied this rate of depreciation (a) as the sole regressor, with a common coefficient across countries and (b) with US output growth also included. The rate of depreciation was statistically insignificant (at conventional levels) in each of these cases, as was the growth rate of government spending. And when US output growth was included it remained highly significant, essentially reproducing table 1.

How can one reconcile this evidence with the well-known statistical findings of Bernanke (1995)? First, we study 1921–1939 whereas he focused on the recovery from the Great Depression during 1931–1935. Second, he grouped countries according to whether they were on or off the gold standard, whereas we measure the rate of depreciation against gold for each country. He reported a significant effect of leaving the gold standard on a range of economic indicators. We study real GDP growth, whereas his indicators included manufacturing production and employment. We also have data on slightly different groups of countries.

To examine the effect of these differences we also estimated the output equations (a) with the on/off gold indicator used by Bernanke (rather than the rate of depreciation), (b) for 1931–1935, and (c) with dummy variables for each year, as in his study. We found a positive correlation between being off gold during the early 1930s and output growth, but not a statistically significant one. The results were similar when we used the on/off gold indicator for the entire 1921–1939 time period.

Given Eichengreen's (1992) conclusions about exogeneity, being on or off the gold standard seems a better candidate for an exogenous variable than the scale of depreciation, which could well respond to business-cycle conditions. But this indicator too is statistically insignificant for output growth for the entire interwar period. Perhaps that is not surprising, for the 1920s of course included years in which a number of countries, including Argentina, Finland, France, Italy, and Sweden, were off the gold standard yet following monetary contractions to try to restore their gold-standard parities. Pooling those episodes with the subsequent departures from the gold standard in the early 1930s leads to very little correlation between the exchange-rate regime and output growth.

A second, traditional way to measure exchange rates is relative to other currencies. This indicator measures relative reflation but also may provide information on a mechanism for recovery: depreciation that promoted net exports. We alternately included the rate of depreciation of each domestic currency against the US dollar or against sterling. Both these candidates for inclusion as exogenous variables were statistically insignificant also. Finding x-variables to be statistically significant is necessary for identification here, but we do not find that for exchange-rate depreciations. Omitting them thus does not explain the insignificant values of $\hat{\beta}$ found in table 1.

We conclude our discussion of controlling for depreciations with two important qualifications. First, finding a fool-proof indicator of the exchange-rate regime is not easy even in postwar data. Rose (2011) notes the lack of consensus on how to classify these regimes and reviews present-day classification methods and their impact. An obvious example of the pitfalls in such indicators arises for Austria, Czechoslovakia, Germany, and Hungary, which remained officially on the gold standard in the 1930s but applied a wide range of exchange controls. Their nominal exchange rates thus do not provide accurate measures of domestic reflation. Of course, exchange controls also were widespread in the early 1920s.

Second, the statistical insignificance of the exchange-rate indicators in this panel does not mean that adhering to, then leaving, the gold standard was not a central cause of the interwar business cycle. Our statistical model is compatible with the possibility that monetary policies in the US and the UK were important contributors to the business cycle in the world economy, as stressed by Feinstein, Temin, and Toniolo (2008). They report on the simultaneous monetary contractions in the US and Germany in 1931 for example. But we cannot isolate an additional role for country-specific exchange rate changes in these other economies. In the case of fiscal policy, in contrast, few scholars have argued that government-spending shocks were important in the UK and US business cycles, so our finding that we cannot find a role for them in these twenty countries reinforces that conclusion.

8. Sorting the Effect of Government Spending Growth

So far we have pooled all observations to allow identification and estimate the average effect of government-spending growth as precisely as possible. But this pooling may obscure some heterogeneity in β, the parameter that measures the impact of government-spending growth on output growth. We next explore possible heterogeneity along several dimensions in which we have reliable data.

Imagine that β is not a constant but depends on some variable v_{it}. Unfortunately, we cannot usually add an interaction term $v_{it}g_{it}$ to the y-equation, for v_{it} sometimes will be endogenous. Thus, we simply sort the observations into groups based on the value of several criterion variables v_{it} to see if that reveals variation in $\hat{\beta}$.

The first candidate as a sorting criterion is the value of the government budget balance, as a percentage of national income. The idea of course is that the impact of a change in government spending might depend on the financing method, perhaps with loan-financed changes having a larger effect than tax-financed ones.

Once again the interwar period is an interesting time to study, because there was large variation in budget balances over time and countries, which should be informative. For example, Belgium, Czechoslovakia, France, and Italy in the early 1920s all ran budget deficits that often exceeded 10% of GDP. Other European countries, like Denmark and Switzerland, remained near budget balance throughout the interwar period. A number of countries, such as Canada, swung from surplus to deficit during the 1930s. Japan ran surpluses throughout the period.

We sort by observations (not countries) according to the value of the budget surplus or deficit. Then we re-estimate the y-equation within each group of observations. Table 2 contains the results. The upper two rows divide the observations into two groups, while the lower three rows divide them into three groups again ranked by budget balances. The x-variable is US output growth, with a common coefficient within each group of observations.

The point estimate $\hat{\beta}$ is positive in each group of observations, but it remains statistically insignificant at conventional levels of significance. When we divide the sample in half, the point estimate is larger for the sample with larger budget surpluses. When we divide the sample in thirds the point estimate is largest for the sample with the intermediate surpluses. Meanwhile, the statistical significance of US output growth also varies somewhat across the sub-samples. Overall though, this first sorting does not alter the conclusions about β.

A second way to sort observations is according to the share of government spending in GDP. One might expect that the impact of a change in the growth rate of GDP would be smaller in countries or years in which government spending is a relatively small proportion of national income. Again, there is considerable variation in this ratio in these historical data. Figure 3 shows the government spending share varying widely across countries and over time. When we sort observations by this ratio and estimate the y-equation on sub-samples (not shown), we find a highly significant role for US output growth for each sub-sample, but generally no significant effect of government-spending growth.

Finally, we also sorted the observations according to the exchange-rate regime and specifically by each of (a) whether the gold standard was in place or not, (b) the rate of depreciation against gold, and (c) the rate of depreciation against the US dollar. This sorting is designed in part to detect the traditional, Mundell-Fleming view that government spending is more effective under a fixed nominal exchange rate than under a floating one.

However, once again none of these ways of sorting the observations gives a sub-sample with a statistically significant coefficient $\hat{\beta}$ on government spending growth. In fact, the point estimate is negative for low rates of nominal depreciation and positive for high rates of depreciation.

9. Defence Spending

We also study whether the growth of defence spending influenced real output growth during the interwar period. This investigation should be informative for two reasons. First, there are sharp spikes up in defence spending in a range of years and countries, such as Japan during 1932–1934, Italy during 1935–1936 and Australia, Canada, Portugal, and other countries in 1939, as well as spikes down in Argentina during 1937 and France during 1936. This variation should help identify the impact, if any, on GDP growth. Second, the impact parameter, β, may be larger for defence spending than for other categories of spending or transfers, for theory suggests it may be the most stimulative via its welfare-reducing effect on labour supply. Admittedly, focusing on defence spending does not answer the question of whether changes in government spending overall affected the path of output, but it may tell us about a large component of the public-sector budget and whether it did matter to overall growth or could have done so.

We draw on data on defence spending from the League of Nations yearbooks, which yield 313 annual observations. We adopt the same statistical model as for total government spending and so allow for defence spending to respond to the business cycle. The results are very similar to those for total spending in sections 5 and 6 and so are not shown separately. There is a significant correlation between US output growth and output growth in the panel and there is evidence of differences in α, the government-spending reaction parameter for defence spending, across decades. But the impact coefficient $\hat{\beta}$ is statistically insignificant and its point estimate is negative. In addition, we also find no significant role for defence spending when we treat its growth rate as exogenous, and so estimate by ordinary least squares.

10. Conclusion

Was counter-cyclical government spending undertaken in the interwar period? Was it effective? Answers to these questions may be sensitive to the identification scheme chosen. This paper adopts a new identification scheme that seems to capture the spirit of many informal assessments of the role of this type of fiscal policy in the interwar period. First, it was tried to a greater extent at some times than in others so the reaction-function parameter, α, varied over time. Second, it could have been effective at any time so the impact parameter, β, was the same across decades.

With limited time-series data for the interwar period, it is challenging to measure country-specific shocks as a means to assess fiscal policy's effectiveness. One also cannot rely on measures of hours or consumption that one would naturally study in assessing more recent fiscal policy. Instead, we look for a correlation between the growth of government spending and the growth of real output, controlling both for other influences on output and for the endogeneity of government spending. The idea is that this partial correlation should be present if government spending had a significant effect, whatever the mechanism by which this occurred. This approach makes use of the great variation across countries and time in the growth rates of output and government spending. It also relies on an exogenous variable, and we find that US output growth was statistically significant for this panel of economies.

There are two main findings. First, the government-spending reaction function varied over time, with government spending becoming less counter-cyclical in the 1930s than in the 1920s. Second, there was no statistically significant impact of government-spending growth on output growth. The 95% confidence interval for β is (-0.22,0.36). The upper limit of this range is comparable to the point estimate of the impact of the US business cycle on domestic output growth.

Given the limited data we are modest about the findings. So far, though, the case that these policies mattered to output growth cannot be proved using this identification and annual macroeconomic data for a panel of twenty countries with great variation in their macroeconomic histories. Finding an insignificant effect of government spending growth on output growth for the interwar period is especially noteworthy because this period included many years with high unemployment, very low interest rates, and large changes in defence spending, circumstances that recent research suggests should lead to a macroeconomic impact for changes in government spending. But for these countries on average we cannot reject the Treasury view.

11. Appendix: Data Sources

11.1 Countries and Currencies

The twenty countries are Argentina, Australia, Austria, Belgium, Brazil, Canada, Czechoslovakia, Denmark, Finland, France, Germany, Hungary, Italy, Japan, Netherlands, Norway, Portugal, Spain, Sweden, and Switzerland. The United Kingdom and the United States are the reference countries.

Five countries changed currencies or had multiple currencies during the period: Argentina the peso papel and the peso oro; Austria the schilling and the crown; Belgium the franc and the belga; Germany the reichsmark and papiermark; and Hungary the pengő and korona. We take these changes into account in constructing exchange rates.

11.2 Nominal Output

Annual, nominal output from 1920 to 1939 come from Michael Bordo's financial crises dataset http://econweb.rutgers.edu/bordo/Financial_Crises_Database.zip, which also gives original sources, except for data for Austria (1924–1937), Czechosolvakia (1920–1937), and Hungary (1925–1939), which come from Mitchell (1998). Output is measured as GDP for Argentina, Australia, Brazil, Czechoslovakia, Denmark, and Norway and as GNP for other countries.

11.3 Population

Annual population series are from Bordo (nd) except for data for Czechoslovakia from Mitchell (2003).

11.4 Government Expenditures

Estimates of current government spending on goods and services on a national accounts basis for the interwar period are scarce. Data are available from Liesner (1989) only for Australia, Canada, Germany, Italy, Japan, and Sweden, and these series do not all begin in 1920.

For the full set of countries the measure of government spending is broader and includes some transfers. Such total government expenditure data are obtained from Mitchell (1998, 2003). These data are for central governments. The following states were unitary or highly centralised: Belgium, Brazil (especially after 1930), Czechoslovakia, Denmark, France, Finland, Germany (especially after 1933), Hungary, Italy, Japan, the Netherlands, Norway, Portugal, Spain (until 1932), and Sweden. For more decentralised or federal states there is the possibility of changes in local government expenditure offsetting those at the central level. Such states were Argentina, Australia, Austria (though with no state tax revenue), Brazil (1920–30), Canada, Germany (1920–33), Spain (1932–1939), and Switzerland.

Due to the data not being expressed in calendar years for all countries, and due to differences in the reporting of data for fiscal years, two types of moving averages were computed to approximate calendar-year expenditures from fiscal-year expenditures, wherever applicable. For Australia, Denmark, Hungary, Italy, Norway, and Sweden the data were reported for the fiscal year ending in the given calendar year (*e.g.* 1920 data actually corresponding to 1919–20 fiscal-year data), so the moving average for that year was calculated using the fiscal year that followed (*e.g.* the average of 1919–20 and 1920–21). In contrast, for Canada, Germany, and Japan the data were reported for the fiscal year ending in the next calendar year (*e.g.* 1920 data actually corresponding to 1920–21 fiscal-year data), so the moving average for that year was calculated using the fiscal year that preceded it. Finally, some countries (France, Portugal, and Sweden) changed fiscal years and so required weights that varied over time in order to approximate calendar-year data from fiscal-year data. The details are available from the authors.

11.5 Deflators

Real output and government expenditures are estimated by dividing the nominal series by cost-of-living indexes from Bordo except for data for Czechoslovakia, Hungary, and Spain which are from Mitchell (1998). For Austria and Germany, prices are in terms of gold after currency stabilizations in 1922 and 1923 respectively. To avoid hyperinflationary periods we include Austria from 1923 and Germany and Hungary from 1924.

11.6 Government Revenue

Government revenues are from Mitchell (1998, 2003). Exact definitions of revenues differ from country to country, with some total revenue figures being only for taxation. After being adjusted from a fiscal-year to a calendar-year basis (as described for expenditures above), they are used with total government spending to calculate nominal government budgetary balance in each year. The ratio to nominal GDP is then used as the budget balance share.

11.7 Exchange Rates

Exchange rates are measured first in national currency per USD, from the League of Nations (1926–1940). They are converted to gold values by multiplication by the USD value in terms of gold from Officer and Williamson (2010).

11.8 Unemployment Rates

Unemployment rates are from Mitchell (1998) or Maddison (2003), which present different measures. But there are no data for Argentina, Brazil, Czechoslovakia, Hungary, Portugal, or Spain, and those for France, Italy, and Japan are for limited years. The total number of country-year observations was 199, compared to the 369 observations on real output.

11.9 Defence Spending

Annual defence spending is from the League of Nations (1924–1940) yearbooks, then deflated by the national cost-of-living index.

References

Almunia, Miguel, Agustín S. Bénétrix, Barry Eichengreen, Kevin H. O'Rourke, and Gisela Rua (2010) "From Great Depression to Great Credit Crisis: Similarities, Differences, and Lessons," *Economic Policy* 25: 219–265

Auerbach, Alan J., William G. Gale, and Benjamin H. Harris (2010) "Activist Fiscal Policy," *Journal of Economic Perspectives* 24(4): 141–164

Auerbach, Alan and Yuriy Gorodnichenko (2012) "Measuring the Output Responses to Fiscal Policy," *American Economic Journal: Economic Policy* 4: 1–27

Barro, Robert J. and Charles J. Redlick (2011) "Macroeconomic Effects from Government Purchases and Taxes," *Quarterly Journal of Economics* 126: 51–102

Bernanke, Ben S. (1995) "The Macroeconomics of the Great Depression: A Comparative Approach," *Journal of Money, Credit and Banking* 27: 1–28

Betts, Caroline M., Michael D. Bordo, and Angela Redish (1996) "A Small Open Economy in Depression: Lessons from Canada in the 1930s," *Canadian Journal of Economics* 29: 1–36

Bordo, Michael D. (nd) Financial Crises Database.
`http://econweb.rutgers.edu/bordo/Financial_Crises_Database.zip`

Born, Benjamin and Gernot J. Müller (2012) "Government Spending Shocks in Quarterly and Annual Time Series," *Journal of Money, Credit and Banking* 44: 507–517

Brown, E. Cary (1956) "Fiscal Policy in the Thirties: A Reappraisal," *American Economic Review* 46: 857–879

Cha, Myung Soo (2003) "Did Takahashi Korekiyo Rescue Japan from the Great Depression?" *Journal of Economic History* 63: 127–144

Choudhri, Ehsan U. and Levis A. Kochin (1980) "The Exchange Rate and the International Transmission of Business Cycle Disturbances: Some Evidence from the Great Depression," *Journal of Money, Credit and Banking* 12: 565–74

Crafts, Nicholas and Peter Fearon (2010) "Lessons from the 1930s Great Depression," *Oxford Review of Economic Policy* 26(3): 285–317

Crafts, Nicholas and Terence C. Mills (2012) "Rearmament to the Rescue? New Estimates of the Impact of 'Keynesian' policies in 1930s Britain," mimeo, Department of Economics, University of Warwick

Cole, Harold L. and Lee E. Ohanian (2011) "Cartelization Policies and the International Great Depression," mimeo, Department of Economics, UCLA

Eggertsson, Gauti B. (2008) "Great Expectations and the End of the Depression," *American Economic Review* 98: 1476–1516

Eichengreen, Barry (1992) *Golden Fetters: The Gold Standard and the Great Depression, 1919–1939.* (New York: Oxford University Press)

Eichengreen, Barry (2004) "Viewpoint: Understanding the Great Depression," *Canadian Journal of Economics* 37: 1–27

Eichengreen, Barry and Jeffrey Sachs (1985) "Exchange Rates and Economic Recovery in the 1930s," *Journal of Economic History* 45: 925–46

Feinstein, Charles H., Peter Temin, and Gianni Toniolo (2008) *The World Economy between the World Wars* (New York: Oxford University Press)

Fishback, Price V. (2010) "U.S. Monetary and Fiscal Policy in the 1930s," *Oxford Review of Economic Policy* 26: 385–413

Fonseca Marinheiro, Carlos (2006) "The Sustainability of Portuguese Fiscal Policy from a Historical Perspective," *Empirica* 33: 155–179

Gordon, Robert J., and Robert Krenn (2010) "The End of the Great Depression, 1939–41: Policy Contributions and Fiscal Multipliers," NBER Working Paper No. 16380

Grytten, Ola Honningdal (2006) "Why Was the Great Depression not so Great in the Nordic Countries?" mimeo, Department of Economics, Norwegian School of Economics and Business Administration

Hall, Robert E. (2009) "By How Much Does GDP Rise if the Government Buys More Output?" *Brookings Papers on Economic Activity* Fall: 183–231

Ilzetski, Ethan, Enrique G. Mendoza, and Carlos Végh (2013) "How Big (Small?) are Fiscal Multipliers?" *Journal of Monetary Economics* 60: 239–254

Jonung, Lars (1979) "Knut Wicksell's Norm of Price Stabilization and Swedish Monetary Policy in the 1930s," *Journal of Monetary Economics* 5: 459–496

League of Nations (1924–1940) *Armaments Year-book, 1924–1940.* Geneva: League of Nations.

Liesner, Thelma (1989) *One Hundred Years of Economic Statistics: United Kingdom, United States of America, Australia, Canada, France, Germany, Italy, Japan, Sweden,* 2nd ed. (London: The Economist Publications)

Maddison, Angus (2003) *The World Economy: Historical Statistics* (Paris: OECD)

Martineau, Nicolas-Guillaume and Gregor W. Smith (2012) "Identifying Fiscal Policy (In)effectiveness from the Differential Adoption of Keynesianism in the Interwar Period," Queen's University, Department of Economics, Working Paper 1290

McGrattan, Ellen R. and Lee E. Ohanian (2010) "Does Neoclassical Theory Account for the Effects of Big Fiscal Shocks? Evidence from World War II," *International Economic Review* 51: 509–532

Middleton, Roger (2010) "British Monetary and Fiscal Policy in the 1930s," *Oxford Review of Economic Policy* 26(3): 414–441

Mitchell, Brian R. (1998) *International Historical Statistics, 1750–1993*, 4th ed. (London: Macmillan)

Mitchell, Brian R. (2003) *International Historical Statistics: Europe 1750–2000*, 5th ed. (Basingstoke: Palgrave Macmillan)

Montgomery, Arthur (1938) *How Sweden Overcame the Depression, 1930–1933* (Stockholm: A. Bonniers boktryckeri)

Myrdal, Gunnar (1939) "Fiscal Policy in the Business Cycle," *American Economic Review (P)* 29: 183–193

Officer, Lawrence H. and Samuel H. Williamson (2010) "The Price of Gold, 1257 – Present," *MeasuringWorth* `http://www.measuringworth.com/gold/`

Parker, Jonathan A. (2011) "On Measuring the Effects of Fiscal Policy in Recessions," *Journal of Economic Literature* 49: 703–718

Ramey, Valerie A. (2011) "Can Government Purchases Stimulate the Economy?" *Journal of Economic Literature* 49: 673–685

Ritschl, Albrecht (2002) "Deficit Spending in the Nazi Recovery, 1933–1938: A Critical Reassessment," *Journal of the Japanese and International Economies* 16: 559–582

Romer, Christina D. (1992) "What Ended the Great Depression?" *Journal of Economic History* 52 (December): 757–784

Rose, Andrew K. (2011) "Exchange Rate Regimes in the Modern Era: Fixed, Floating, and Flaky," *Journal of Economic Literature* 49: 652–672

Safarian, A. Edward (1970) *The Canadian Economy in the Great Depression* (Toronto: McLelland and Stewart)

van Zanden, Jan L. (1998) *The Economic History of the Netherlands 1914–1995: A Small Open Economy in the 'Long' Twentieth Century* (London: Routledge)

[Lead footnote]

Martineau: Department of Economics, Glendon College, York University; martineau@glendon.yorku.ca. Smith: Department of Economics, Queen's University, Canada; smithgw@econ.queensu.ca. We thank the Social Sciences and Humanities Research Council of Canada and the Bank of Canada research fellowship programme for support of this research. We thank Steve Ambler, Hafedh Bouakez, Patrick Coe, Lee Ohanian, Christoph Priesmeier, and seminar participants at the CEA and AEA meetings, the Bank of Canada, the T2M conference, McMaster University, the Federal Reserve Bank of Philadelphia, the International Institute of Public Finance, and the Canadian Macroeconomics Study Group for very helpful comments. We thank two referees of this journal for detailed, constructive comments. The opinions are the authors' alone and are not those of the Bank of Canada.

Table 1: Government Spending Impact 1921–1939

$$y_{it} = \delta_{yi} + \beta g_{it} + \omega x_t + \epsilon_{yit}$$

$$z_{it} = \{d_i, d_{1920s} x_t, d_{1930s} x_t\}$$

x_t	$\hat{\beta}$ (se)	$\hat{\omega}_u$ (se)	$\hat{\omega}_k$ (se)	$J(df)$ (p)
y_{ut}	0.07 (0.15)	0.29*** (0.04)		
$\{y_{ut}, y_{kt}\}$	0.02 (0.10)	0.28*** (0.03)	0.09 (0.11)	0.79(1) (0.38)

Notes: There are $I \times T$=350 observations; y is real output growth, g is real government spending growth; y_u is growth in the US, y_k is growth in the UK. d_i=1 for country i; d_{1920s}=1 for the 1920s; d_{1930s}=1 for the 1930s. Standard errors are robust to heteroskedasticity. The J-statistic tests the over-identifying restriction. *** denotes a p-value less than 0.01.

Table 2: Sorting the Government Spending Impact by Budget Balance

$$y_{it} = \delta_{yi} + \beta g_{it} + \omega_u y_{ut} + \epsilon_{yit}$$

$$z_{it} = \{d_i, d_{1920s} y_{ut}, d_{1930s} y_{ut}\}$$

Surplus Range	$\hat{\beta}$ (se)	$\hat{\omega}_u$ (se)
-29% to -1.6%	0.07	0.36***
	(0.14)	(0.04)
-1.6% to 12.6%	0.45	0.20**
	(0.66)	(0.10)
-29% to -2.8%	-0.03	0.32***
	(0.15)	(0.05)
-2.8% to -0.6%	0.93	0.18
	(2.71)	(0.52)
-0.6% to 12.6%	0.25	0.18*
	(0.53)	(0.10)

Notes: There are $I \times T$=350 observations; y is real output growth; g is real government growth; y_u is growth in the US. The budget balance is a share of GDP. ***, **, and * denote p-values less than 0.01, 0.05, and 0.10 respectively.

Figure 1: Real Output per capita 1920–1939

Source: Levels of real national output per capita, scaled so that 1920=100, use nominal output, population, and cost-of-living indexes from Bordo (nd) except for Austria, Czechoslovakia, and Hungary where the sources are Mitchell (1998, 2003).

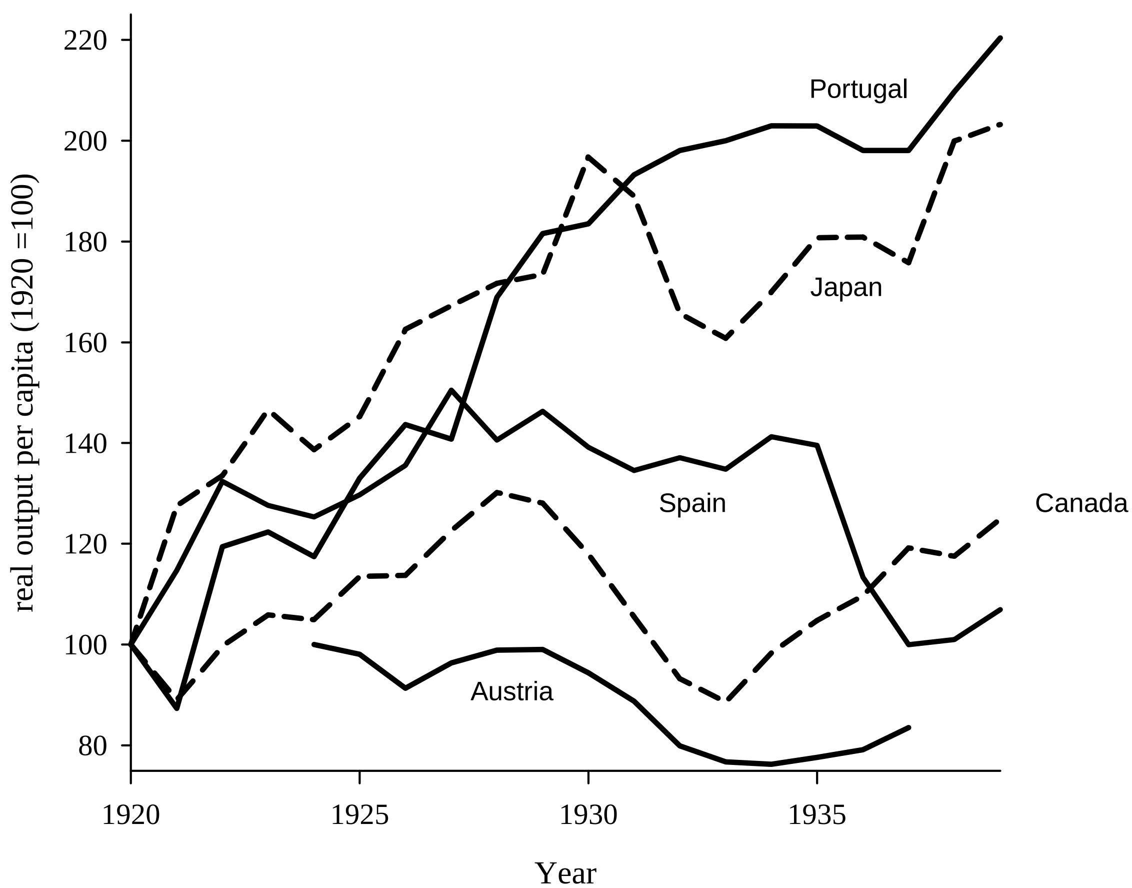
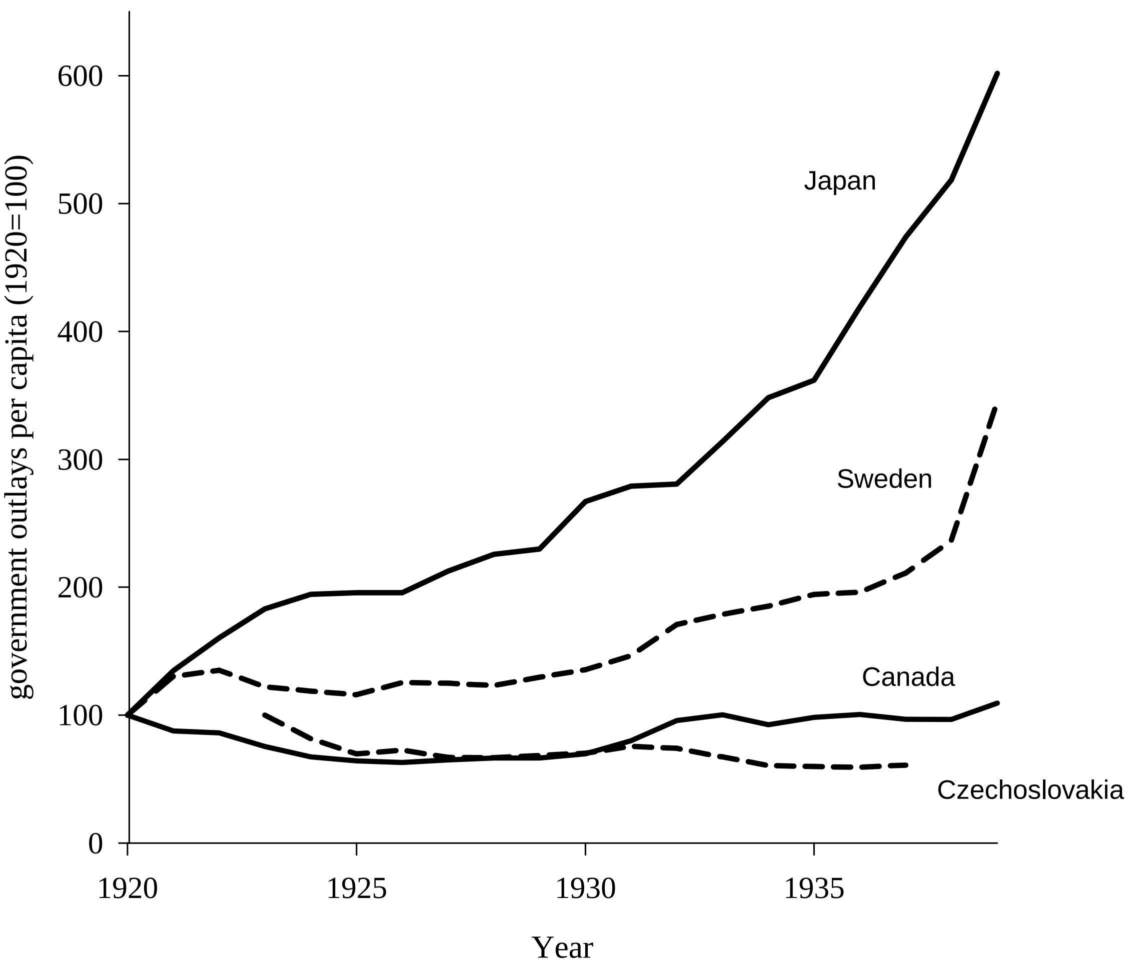

real output per capita (1920 =100)
220
200
180
160
140
120
100
80
1920
1925
1930
1935
Year
Portugal
Japan
Spain
Canada
Austria

Figure 2: Real Government Spending per capita 1920–1939
Source: Levels of government spending are from Mitchell (1998, 2003) averaged from a fiscal-year to a calendar-year basis as described in section 11.4. Population and cost-of-living indexes are as in figure 1.



government outlays per capita (1920=100)
600
500
400
300
200
100
0
1920
1925
1930
1935
Year
Japan
Sweden
Canada
Czechoslovakia

Figure 3: Government Spending Shares of Output

Source: Levels of nominal government spending and national output are from Bordo (nd) except for Austria, Czechoslovakia, and Hungary where they are from Mitchell (1998, 2003). Government spending is adjusted from a fiscal-year to a calendar-year basis as described in section 11.4.

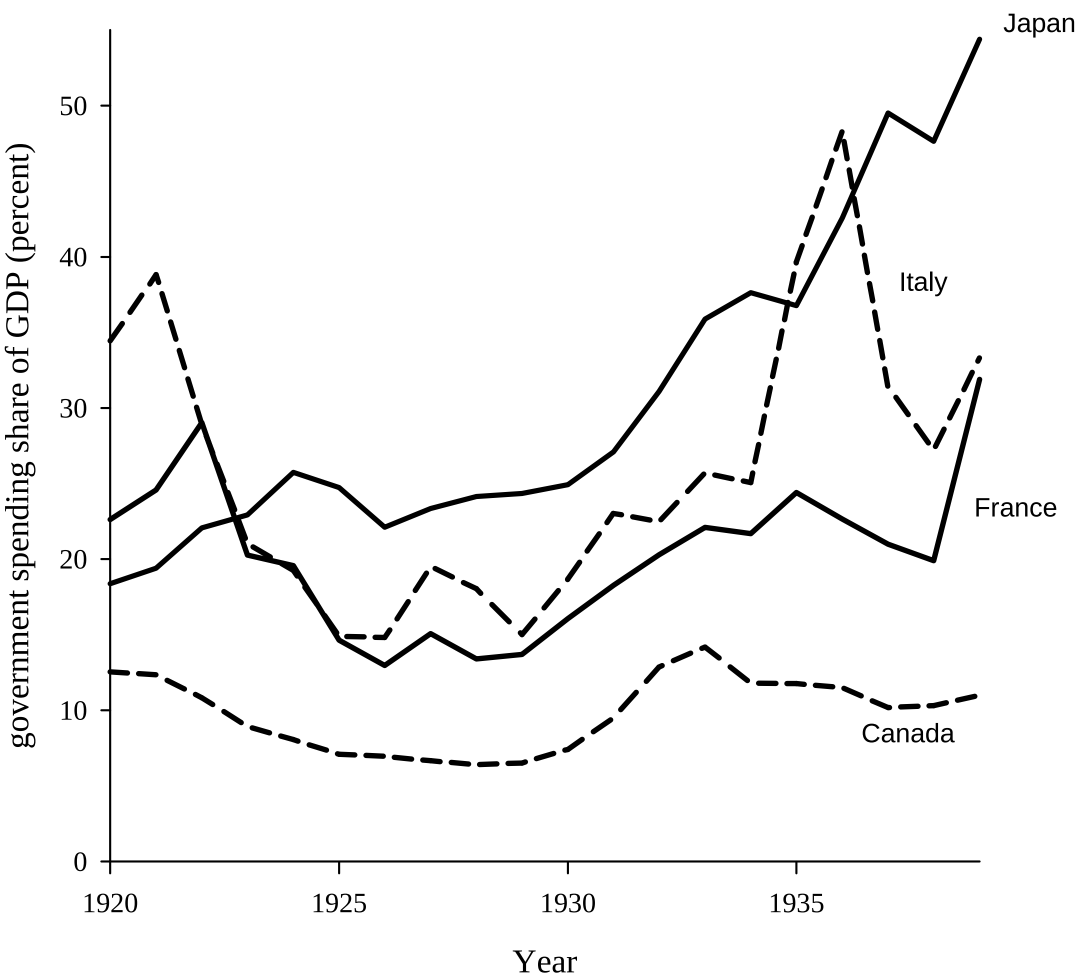
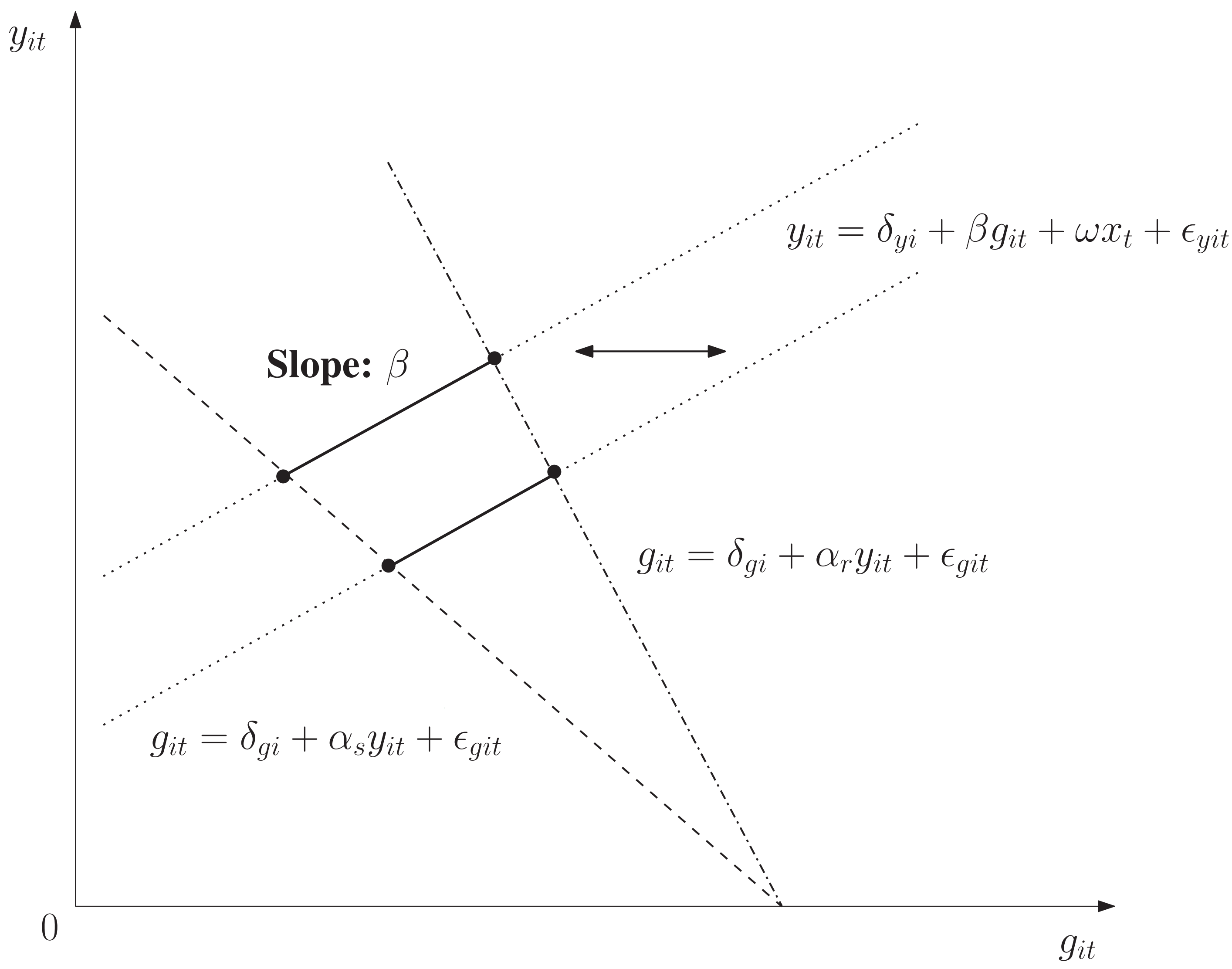

Japan
Italy
France
Canada
government spending share of GDP (percent)
0
10
20
30
40
50
1920
1925
1930
1935
Year

Figure 4: Fiscal Policy Responses and Effectiveness

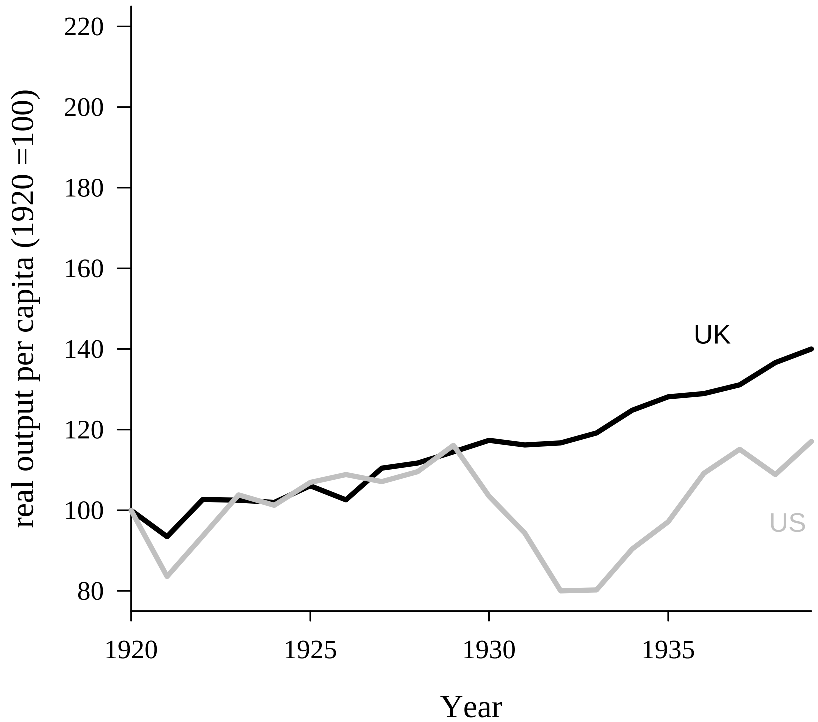
Time periods r and s identifying α_r, α_s, and β



y_{it}
$y_{it} = \delta_{yi} + \beta g_{it} + \omega x_t + \epsilon_{yit}$
Slope: β
$g_{it} = \delta_{gi} + \alpha_r y_{it} + \epsilon_{git}$
$g_{it} = \delta_{gi} + \alpha_s y_{it} + \epsilon_{git}$
0
g_{it}

Figure 5: US and UK History 1920–1939

The upper panel shows real output per capita. The lower panel shows the government spending share of output. Source: Levels of real output per capita, scaled so that 1920=100 use nominal output, population, and cost-of-living indexes from Bordo (nd). Nominal government spending is from Mitchell (1998).


220
200
180
160
140
120
100
80
real output per capita (1920 =100)
UK
US
1920
1925
1930
1935
Year

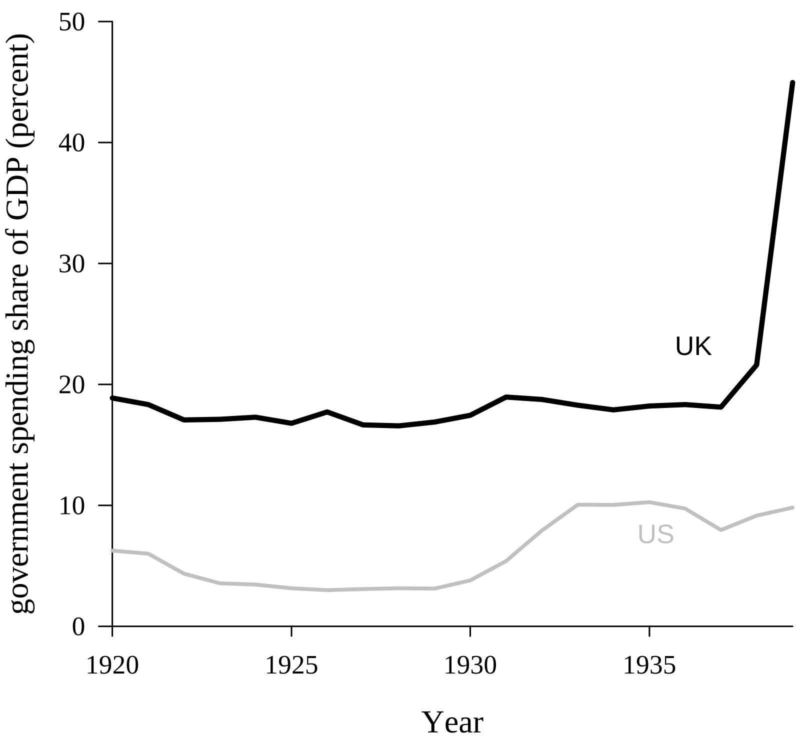
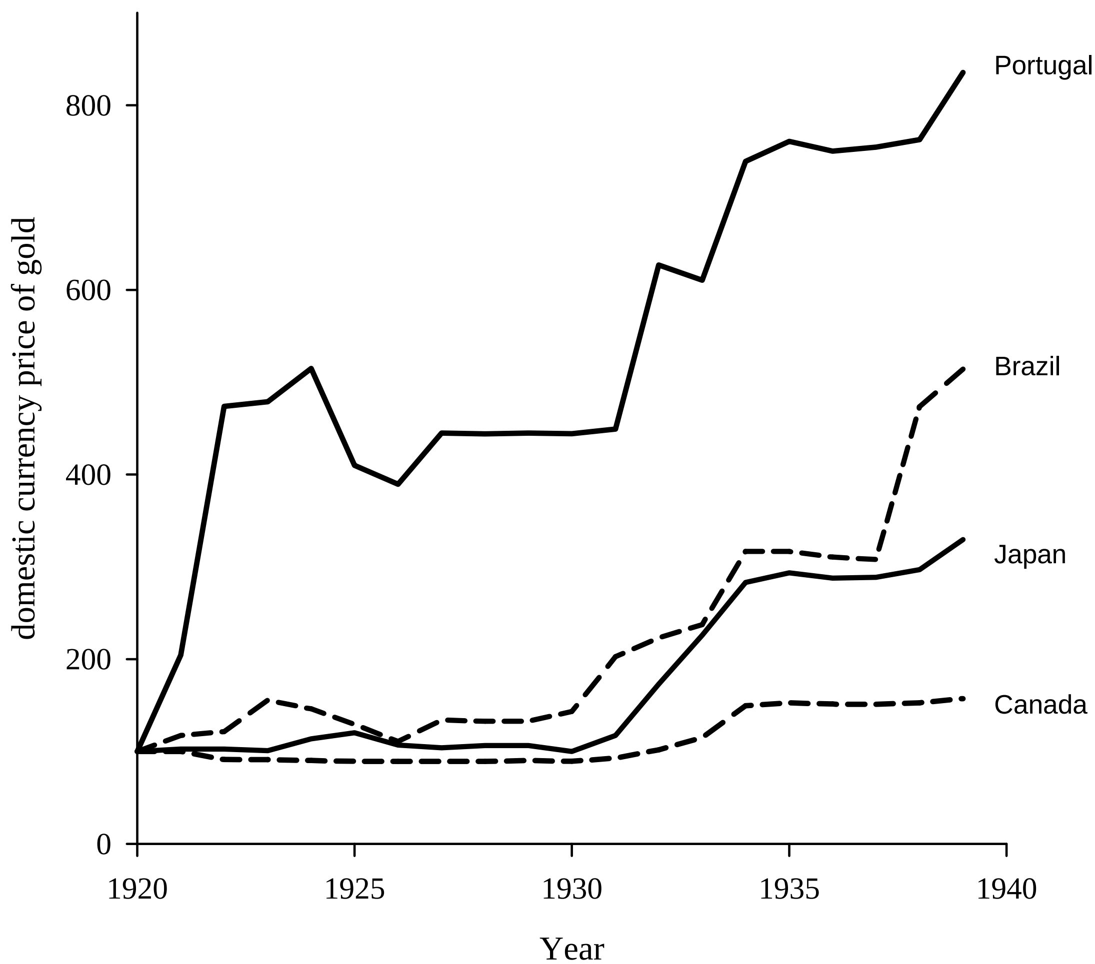

50
40
30
20
10
0
government spending share of GDP (percent)
UK
US
1920
1925
1930
1935
Year

Figure 6: Gold Prices/Exchange Rates

Source: Local-currency prices of gold scaled so that 1920=100 use exchange rates in national currency per USD from the League of Nations (1926–1940). They are converted to gold values using the gold price series from Officer and Williamson (2010).


Portugal
Brazil
Japan
Canada
800
600
400
200
0
domestic currency price of gold
1920
1925
1930
1935
1940
Year