CEP 14-01

The Effect of Federal Government Size on Private Economic Performance in Canada: 1870–2011

J. Stephen Ferris and Marcel-Cristian Voia

Carleton University

March 2014; revised 12 September 2014

CARLETON ECONOMIC PAPERS

Department of Economics

1125 Colonel By Drive
Ottawa, Ontario, Canada
K1S 5B6

The effect of federal government size on private economic performance in Canada: 1870 – 2011

J. Stephen Ferris* Marcel C. Voia†

September 12, 2014

Abstract

This paper re-examines the relation between private economic performance and federal government size in Canada over the long 1870-2011 time period. The particular focus is on whether the effect of government size on private output has an inverted U shape with a tipping point. Its innovation is to use nonparametric techniques to assess whether the quadratic form most often employed is the appropriate parametric form for undertaking significance tests and whether that relationship is stable across the period. The empirical work does find a nonlinear relationship with a tipping point but finds the quadratic form applicable only to the early 1870-1936 time period. The latter period is more consistent with a linear relationship embodying a constant rather than increasing output cost. It follows that policy prescriptions based on the fear that further expansion in government size generates ever increasing cost become more problematic.

Key words: Government Size, nonlinear time series, nonparametric methods, tipping point, endogeneity correction.

Journal of Economic Literature classification: H21, H23, C22, C26.

*Corresponding Author: Economics Department and Centre for Monetary and Financial Economics (CMFE), Carleton University. Mailing address: Economics Department, Carleton University, Loeb Building 1125 Colonel By Drive, Ottawa, Ontario, K1S 5B6 Canada. Tel (613) 520-2600-; FAX: (613)-520-3906; email: steve.ferris@carleton.ca

†Economics Department and Centre for Monetary and Financial Economics (CMFE), Carleton University. Mailing address: Economics Department, Carleton University, Loeb Building 1125 Colonel By Drive, Ottawa, Ontario, K1S 5B6 Canada. Tel (613) 520-2600-3546; FAX: (613)-520-3906; email: marcel-cristian_voia@carleton.ca.

1 Introduction

In Canada, as elsewhere, a considerable literature exists on the relationship between government size and economic performance.[1] While much of that discussion relates historically to the question of how government size responds to changes in income and output–*Wagner's Law*–the strand of the literature we are interested in reverses causality to ask whether government complements or discourages private economic performance. A primary reason for the latter interest is because most developed economies have experienced long periods of growth in the size and scope of government so that periods of contraction trigger a concern that government size may have become excessive, unduly constraining private performance. More recently, in part in response to the development of endogenous growth theory, analysis has focused on the effect of government size in relation to growth (Barro, 1990; Armey, 1995). Here the consensus view is that larger size has a negative effect on the growth rate (at least in developed economies). For example, Afonso and Furceri (2010, p. 527) investigate the effect of government size and its volatility on economic growth in OECD and European Union countries and conclude that "both dimensions tend to hamper growth." [2] Similarly Bergh and Henrekson (2011, p.1) conclude that "most recent studies typically find a negative correlation between total government size and economic growth". Finally, Facchini and Melki (2013, p.2) survey sixty investigations of the relation between government size and economic outcomes and find that "66.6% of the studies find a negative effect from government size while only 8.3% find the opposite effect and 25.1% are inconclusive."

In this paper we re-examine this issue in relation to the size of the Canadian federal government over the long time period since Confederation (1870-2011). Federal size is particularly important since the federal government remains responsible for the most basic levels of individual and state security and because it is charged with counter-cyclical fiscal policy responsibilities. [3] Our analysis begins by asking that if government size did affect economic performance, what would be the appropriate performance measure that could be linked meaningfully to government size. Second we ask whether the expected effect of government size on performance would be linear. Arguing first that in the long run it is the level rather than the growth rate of economic performance that can be related meaningfully to government size and, second, that that relationship should be nonlinear with a tipping point, we test for the shape of that relationship

[1] Important contributions to the general literature on the effect of government size on output include: Landau, 1983; Kormendi and Maguire, 1985; Ram, 1986; Grossman, 1987. Contributions relating government size and output in Canada include Scully, 1989; Afxentiou and Serletis, 1991; Chao and Grubel, 1998; Petry, Imbeau, Crete and Clavet, 2000; Voia and Ferris, 2013. Chao and Grubel (following Scully's methodology) find the optimal long run size of all governments in Canada to be about 34% of GDP.

[2] Afonso and Furceri (2010) find that a one percent increase in government size decreases output growth by .12% for OECD countries and .13% for European Union countries.

[3] The latter is important because counter-cyclical Keynesian fiscal intervention becomes embodied in the long run relationship linking government size and output.

using the quadratic form. Robustness checks on the size and significance of the implied tipping point indicate the need to correct both for correlations arising among the covariates across time and for potential endogeneity arising between government size and private output. Although doing so confirms the quadratic form, the size of the confidence interval about the tipping point and the compatibility of the data with the cubic form lead us to adopt nonparametric modeling methods that generalize the nonlinear form in ways that do not require imposing symmetry. These investigations also lead to the discovery of a likely break in the form of the time series around 1937. This serves to reconcile the plausibility of the divergent forms suggested by the earlier parametric tests done over the entire period.

Our nonparametric method uses the spline-based method developed by Ma and Racine (2013), Ma, Racine and Yang (2011) and Nie and Racine (2012) to describe the forms of the unconstrained relationship arising in the data. They allow the unconstrained patterns of response to different control variables to be illustrated in a convenient graphical way and in a form that allows for the incorporation of endogenous regressors through the generation of instrumental variable (IV) nonparametric plots. The enhancement of the analysis of tipping points by surrounding the point estimate with an appropriate confidence interval allows assessment of whether or not a quadratic model estimate of optimal size is meaningful and thus relevant for policy analysis. To anticipate our final conclusion, a tipping point is discovered in Canada for the earliest (1870-1937) time period but not in the post 1937 time period.

2 Time series and endogeneity concerns with government size and economic performance

The time series issue posed by the long run relationship between government size and the alternative measures of economic performance can be seen in the following diagrams. In Figure 1 below we show government size, measured as the logarithm of the ratio of total non-interest federal government expenditures to GNP ($\ln GSize$), in relation to both the level of private economic performance, measured as the logarithm of private output per capita ($\ln PYPC$), and its rate of change, the growth rate of private output per capita ($PCGROWTH$).[4] As can be seen from the top right panel of Figure 1, $\ln PYPC$ has risen more or less continuously over the past century and a half in Canada. In contrast the bottom panel shows that the growth rate of per capita output, $PCGROWTH$, does not increase, varying more or less randomly about a constant mean of 1.88% per year. In econometric terms, the level of private economic activity is non-stationary or integrated of order one, I(1), while its rate of growth is stationary or integrated of

[4] Private output is defined as GNP minus total non-interest federal government expenditures.

order zero, I(0).[5] When we turn to examine the long run growth of government size, $\ln GSize$ (the left panel) it is immediately apparent that abstracting from the spikes associated with the two world wars, government size has increased continuously since 1870. Beginning from the low level of 3.5 percent of GNP in 1870, federal government size increased to over 20 percent of GNP by 1975, before falling back to 13.7 percent of GNP by 2011. In econometric terms, $\ln GSize$ is I(1) or non-stationary. [6]

– insert Figure 1 about here –

The significance of this time series issue is that when variables of different order are regressed together, the resulting coefficient estimates can be interpreted erroneously. For example, a regression that finds no relationship between an upward trending level of government size and a stationary growth rate may lead to the rejection of a meaningful relationship arising between the two levels. Similarly variables that trend either directly or inversely through time are often misinterpreted as being causally related. Finally, stationary variables that are related through time are often misinterpreted as implying a permanent rather than transient relationship between measures. This suggests that when putting together longer run time series in a hypothesis test, one should first look to their order of integration then, if relating I(1) variables, look for the presence of cointegration among the set of interrelated variables. In our case we begin by exploring the reasoning that would link together the two I(1) variables: government size and the level of private output per capita.[7]

The second significant econometric issue to be faced is endogeneity. That is, while our interest is on how government size affects private output, the literature investigating *Wagner's Law* argues that the increase in government size derives from an expansion in the scale and complexity of the private economy. It follows that the ability to interpret the correlation between government size and per capita output as a measure of governments effect on private output is somewhat problematic. To be more precise about any one of these causal routes, the analysis must control for the potential feedback that can come from induced changes to the other side. This we discuss at length in Section 4 below. Before turning to these empirical issues, however,

[5] The order of integration refers to the number of times a time series must be differenced before finding stationarity. The adjusted Dickey Fuller test statistic for $\ln PYPC$ is -0.091 (constant) and is -10.016 (constant) for $PCGROWTH$. The corresponding MacKinnon 1% critical value of -3.496 allows rejection of the hypothesis that the growth rate is nonstationary.

[6] The adjusted Dickey Fuller test statistic for$\ln GSize$ is -1.886 (constant) and is -6.99 (constant) for its rate of change, $D.\ln GSize$. The corresponding MacKinnon 1% critical value is -3.497.

[7] While the analysis could begin by linking the two first differences, doing so loses the information that could arise from a relationship between the two levels. Similarly because the business cycle is stationary over time, transitory changes in government size that reflect purely countercyclical intervention may dominate the fewer permanent changes in government size that are of interest to this analysis. By initially looking for cointegration among levels we get a cleaner measure of the long run relationships (with the cyclical effects remaining in the residuals). See Ferris (2014) for an expansion of this idea in relation to government size in New Zealand.

we first motivate our empirical hypotheses through an overview of public choice theory on the effect of government size on private economic performance.

3 Public choice and the effect of federal government size on the private economy

Broadly speaking, public choice analysis views increases in government size as producing two opposing effects on the output of the private sector. First in terms of generated benefits, initial levels of government spending are viewed as providing basic levels of security and protection that keep individuals safe from physical threats (through collective policing, national defense and diplomatic services) and safe from internal harm (through the administration of justice as codified in criminal and commercial law). Further expansion of these roles allow higher levels and better qualities of policing and legal services that in turn permit individuals greater predictability in their dealing with strangers thus helping to realize larger volumes of production, trade and welfare (Coase, 1960; Becker, 1983; Wittman, 1995). To the extent that communal services such as health, sanitation, social welfare, education, and research and development are provided directly or, as in Canada's case, funded indirectly by federal revenues, larger levels of government spending increase the quality of productive inputs and through this the output of the private sector (Dahlman, 1991; Thomson and Jensen, 2013). Even more directly, the federal government's provision and monitoring of transportation infrastructure provide inputs that complement private capital and enhance private output (Karras, 1997; Sturm, Kuper and de Haan, 1998). It follows that if the government undertakes projects in order of their social merit, the social marginal product of government's involvement in the private economy will begin positively and fall as government size increases.

The opposing channel of influence focuses on the costs of government intervention. The necessity of funding government activities means that as government size grows, ever larger levels of resources must be obtained from the private sector. Their acquisition through taxation–through higher tax levels and often differential tax rates–decreases private output by discouraging the supply of productive inputs and distorting the cost of their provision (Stuart, 1984; Usher, 1986). Thus as the size of government grows, the tax price of government services increases. This has led writers such as Grossman (1987), Scully (1989, 2000), Armey (1995) and Facchini and Melki (2013) to argue that on net, larger government size must encounter diminishing returns that at some point may reach the point where further increases in size reduce rather than increase private output.[8] The opposing effects of government size on private per capita output are illus-

[8] Theories directed at explaining why government size may become too large include Niskanen's (1968) theory of the expansionary incentives of the bureau, Meltzer and Richard's (1981) median voter theory, and Buchanan and Tullock's (1962) emphasis of the common pool problems arising in modern democracies.

trated in Figure 2 below. The level of government size that maximizes per capita private output is illustrated as and is referred to as the tipping point.

– insert Figure 2 about here –

4 Testing for the nonlinear effect of government size and the search for a tipping point

4.1 The baseline parametric model

The traditional test for whether government exerts an overall positive or negative effect on private output uses the quadratic form to capture the predicted nonlinear effect of size on private output and to evaluate whether current size is above or below the tipping point (in Canada, for example, see Chao and Grubel, 1998). Hence we begin by using as our baseline version of this parametric model:

$$\ln\left(\frac{y_t}{N_t}\right) = \beta_0 + \beta_1 \ln GSize_t + \eta_2 \ln GSize_t^2 + \beta_3 \ln Agric_t + \beta_4 \ln Young + \beta_5 \ln USiip_t + u_t \quad (1)$$

where the dependent variable, $\ln(\frac{y_t}{N_t})$, represents the logarithm of private output per capita, $\ln GSize_t$ and $\ln GSize_t^2$ represent the logarithms of federal government size and size squared while $\ln Agric_t$, the logarithm of the share of agriculture in the labour force, $\ln Young_t$, the logarithm of the percentage of the population age 16 and younger, and $\ln USiip_t$, the logarithm of the U.S. index of industrial production control for other nongovernmental influences on private output.[9] Table 1 provides descriptive statistics for these variables and detail on their sources. The key point to be noted is that all variables except for one of the political variables (to be introduced later) are I(1).

Equation (1) tests for an inverted-U shape to the effect of government size on output and hence whether there is a tipping point. The level of $\ln(\frac{y_t}{N_t})$ at which the curve reaches a maximum/minimum can be solved for as $(\frac{\beta_1}{-2\beta_2})$ and, if a maximum, is known as the tipping point.

[9] The absence of long run data on factor inputs such as capital leads to reliance on proxies such as these for long period analysis. The rational for their use include: the share of employment in agriculture captures (inversely) the growing industrial base of production and urbanization of the Canadian economy; the percentage young captures the effect of demographics, while the very close integration of the Canadian and US economies means that the US index of industrial production captures common improvements in productivity arising across the two countries. The small size of the Canadian economy relative to the US means that we can treat *lnUSiip* as exogenous to the performance of the Canadian economy.

In our case the value of per capita income $(\frac{y_t}{N_t})^*$ that corresponds to the tipping point δ is

$$\delta = \exp\left(\frac{\beta_1}{-2\beta_2}\right) \tag{2}$$

with $\beta_1 > 0$ and $\beta_2 < 0$. The sign restrictions imply that a maximum for $\ln(\frac{y_t}{N_t})$ is reached at a positive level of government size but are not imposed at the estimation stage. Our ultimate objective is to derive a meaning measure of δ and its confidence interval (see the Appendix B for details about confidence set identification).

In column (1) of Table 2 below we present the OLS estimate of equation (1) for the entire 1870 - 2011 time period. The results suggest that the set of variables used to explain private Canadian output per capita performs well, perhaps even too well, with the high R^2 and low Durbin-Watson statistic showing classic signs of a spurious regression. The large absolute size of the adjusted Dickey Fuller (ADF) statistic does, however, allow us to reject the hypothesis of a unit root. This implies that the I(1) variables in column (1) are cointegrated enabling the interpretation of the estimated equation as a long run equilibrium relationship arising among these variables. Unfortunately, although the OLS coefficient estimates and their t-statistics are consistent with our priors, particularly the suggestion that the relationship between government size and per capita output has an inverted U-shape, only the coefficient estimates are super-consistent with the standard errors likely biased downwards because of correlations arising among the covariates over time. This becomes apparent when we correct for intertemporal correlations among the covariates by presenting the dynamic OLS equation estimates (DOLS) in column (2). There it can be seen that the inverted U-shaped relationship between $\ln GSize$ and $\ln PYPC$ found in column (1) was fragile, with both $\ln GSize$ and $\ln GSize^2$ now losing their significance.

There is however a second source of bias in the estimates of the coefficients of $\ln GSize$ and $\ln GSize^2$ as determinants of the variation in private per capita output. That is as discussed earlier, endogeneity in the form of two way causality in the relationship between $\ln GSize$ and $\ln PYPC$ is likely present in the data. When we test for this possibility by running a Granger causality test, we find the probability that $\ln GSize$ does not Granger cause $\ln PYPC$ and vice versa is zero in both cases (for all tests using two or more lags). Correcting for endogeneity means finding instruments that are correlated with government size but not with the error term in equation (1) so that the instruments influence private output only through their effect on government size. Here we rely on previous studies of (federal) government size in Canada (see for example, Winer and Ferris, 2008 and Ferris, Park and Winer, 2011) that have found two political variables and the flow of immigration into Canada to be significant determinants of government size. These variables: effective participation by the public in the electoral process, $\ln Effectivepart$; the percentage of seats won by the winning political party, $\ln Seats$; and the ratio of immigration to population, $\ln Imratio$ were then used as instruments in the two stage

least squares (2SLS) equations.

In columns (3) and (4) of Table 2 we present the 2SLS estimates, first uncorrected for time persistence and then combined with the DOLS correction. In column (3) the effect of instrumenting can be seen as allowing the quadratic effect of government size on per capita output to reappear in the data. Because the sign and significance of the control variables remain largely unchanged across all of these equations, the 2SLS estimates do provide greater confidence that government size has had a positive but diminishing effect on private output. The coefficient estimates suggest that the effect peaks at about 8.5% of GNP. As importantly, the test for endogeneity now suggests that after instrumenting in this manner, the probability that endogeneity remains in the relationship between $\ln GSize$, $\ln GSize^2$ and $\ln PYPC$ is close to zero. In column (4) the 2SLS equation instruments government size and adjusts dynamically to account for intertemporal correlations among the right hand side variables. As expected the intertemporal correction does increase the standard errors somewhat, however the probability that the slope coefficient estimates are equal to zero never rises above five percent. The correction for persistence also results in the estimate of the tipping point falling somewhat, now to 6.85 percent. As was the case with the earlier DOLS adjustment to the OLS estimate in column (2), the intertemporal corrections have little effect on the coefficient estimates and standard errors of two of the control variables ($\ln Agric$ and $\ln USiip$) but do reduce to insignificance the effect of $LnYoung$, the percentage of the population 16 and below.

While the performance of the quadratic form in columns (3) and (4) and the significance of the signs on $\ln GSize$ and $\ln GSize^2$ support the hypothesis of an inverted U-shaped relationship for government size, the relatively small size of the tipping point estimates and the change that arises between columns (3) and (4) raise the question of how much confidence can be placed on either estimate. Somewhat more generally, because the tipping point is the ratio of two regression parameters a number of statistical issues arise when trying to establish an appropriate confidence interval. To do so we adopt two strategies. First, the Delta method uses a truncated Taylor series expansion and asymptotic theory to establish a Wald-type confidence interval.[10] Adopting the 95% confidence criterion establishes Delta upper and lower bounds for column (4) that are, respectively, -0.0513767 and 13.73884. The weakness of the Delta method is that as the estimate of approaches zero, the tipping point becomes weakly identified. An alternative method that does not require strong identification and hence is robust to mistakes in modelling the form of the relation uses Fieller (1954). Using the ninety five percent confidence criterion, the Fieller method in our case produces upper and lower bounds about the 6.85 estimate of the tipping point of 1.3732 and 14.2883.[11] In both cases the confidence interval established is quite large and, in the Delta case, its size is sufficiently big that we cannot conclude with confidence

[10] See Appendix B for particulars in the construction of the Delta and Fieller confidence intervals.

[11] Bolduc, Khalaf and Yelou (2010) show that asymptotically Fieller's solution and the Delta method give similar results when the denominator is far from zero.

that the tipping point is positive. These measures of uncertainty in the coefficient estimate suggest a relatively flat surface in the function linking the logarithms of government size and per capita output.

A further sign that the quadratic form may not fully capture the nonlinear relationship arising between government size and private output is shown in column (5) where we replace the quadratic relationship in (1) with a cubic one. The results mean that the cubic form is also not inconsistent with the data and with an estimated tipping point that at the lower end of the confidence intervals established above (at 3.21 percent). In this case the confidence intervals about the estimated tipping point are somewhat tighter, with the Delta method generating lower and upper bounds of 1.348 and 4.436 and the Fieller method producing not dissimilar lower and upper bounds of 1.699 and 4.721.

5 Nonparametric analysis

Our tests for the hypothesized relationship between government size and private output per capita shown in Figure 2 have thus far used two parametric models that confirm nonlinearity but do not allow a very precise estimate of the tipping point. Because that investigation does not suggest the superiority of one particular parametric form, we propose using nonparametric methods. As in the earlier case of the parametric analysis we estimate conditional nonparametric models without correcting for endogeneity. This allows a discussion of its general advantages before we correct for endogeneity. The advantage of using nonparametric methods is that in cases where an inverted-U shape is suggested, a tipping point can be re-estimated after relaxing symmetry. That is, while the theoretic relationship between government size and private per capita output shown in Figure 2 is not necessarily symmetric about the tipping point, the quadratic model used to test for a tipping point imposes symmetry in its structure. Thus our method checks whether the latter assumption is overly restrictive, whether it affects the estimate of the tipping point and particularly its significance. We also check for time instabilities nonparametrically, by breaking our sample into two time periods. Since in our case conditional analysis is required, this allows us to analyze the information-content of the controls. Controlling for endogeneity in the nonparametric regression insures that the confoundedness arising between government size and private output is eliminated and a more unbiased relationship is estimated.

To implement the nonparametric spline-based method of Ma, Racine and Yang (2011), Nie and Racine (2012) and Ma and Racine (2013) we assume that the conditional mean depends on government size and the controls adopted earlier and follows a non-linear, unknown function approximated by the best-fit B-splines that allows for heteroskedasticity of an unknown form.

That is, we assume

$$\ln(\frac{y_t}{N_t}) = f(\ln GSize_t, Controls_t) + \sigma(GSize, Controls_t)w_t \tag{3}$$

where σ and w_t are unknown and estimation is conducted assuming exogenous and possibly endogenous covariates. In controlling for endogeneity we use the same instruments as those employed in the earlier parametric analysis. Appendix C provides a more technical description of the method used and generates the graphical representations of the fitted function together with the partial effects associated with each covariate. It also provides point-wise confidence bands. To control for endogeneity in the nonparametric conditional representation we estimate a nonparametric IV following a method developed by Horowitz (2011)that uses regression spline methodology. As a robustness check we use Darolles, Fan, Florens and Renault (2011) whose method is modified for regression splines (because Darolles et. al. use local constant kernel weighting). The instruments used are again the same as those used in the parametric regressions. Appendix D provides a technical description of the implementation of the nonparametric IV method. Because the nonparametric estimations do not account for the time series properties of the data nor do they impose stationarity, nonparametric assumptions are not necessarily weaker than the parametric assumptions used earlier (aside from relaxing the shape restriction). For this reason, the estimated curves are not viewed as a formal test of the fit of (3). Rather we view them as alternatives that can be used to suggest tractable parametric specifications and/or alternative ways of estimating. Moreover, in the absence of consensus on the fit of the quadratic model, relaxing the shape restriction can itself be informative. To provide a comparison with the quadratic specification we look for tipping points in the nonparametric forms, for asymmetries about them and for time instabilities. Given that linear and other finite-dimensional parametric models make strong assumptions about the population being modeled, the nonparametric models have the additional advantage of not relying on any functional assumptions, which reduces the potential bias associated with any given parametric form.

Nonparametric estimation of (3) produces two sets of partial regression surfaces that we report in Figures 3 and 4. In Figure 3 we present the model without instrumental variables along with their confidence intervals.

– insert Figure 3 about here –

Figure 4 presents the model with confidence intervals when instrumental variables are used.

– insert Figure 4 about here –

Here the 'partial' surfaces correspond to the estimated output plotted as a function of one predictor with the remaining predictors (the covariates in (3) that do not appear on the axes of the reported figures) held constant at their median values.

A comparison of the results when endogeneity is and is not controlled for results in the same overall pattern of findings as appeared with the earlier parametric regressions. In particular, when endogeneity is not addressed (as in Fig 3) we cannot identify an inverted U-shaped nonlinear relationship between private per capita output and government size. However when endogeneity is controlled for (as in Figure 4) the results match the general form of the inverted U-shape obtained with the parametric model (when we controlled for endogeneity and used the DOLS correction for the long run relationship). The nonparametric result in Figure 4 also indicates that the underlying relationship is not perfectly symmetric about the tipping point.

In addition to describing the form of these curves, the partial surfaces can be used to find the points at which the derivative of the nonlinear surface approaches zero. These points can be used as diagnostic checks on the parametric estimates of the tipping points derived earlier. To obtain these points we use the instrumental regression function derived from the structural econometric model

$$E[Y - \phi(Z, X)|W] = 0, \tag{4}$$

where Y is $\frac{y_t}{N_t}$, $\phi(Z, X)$ is an instrumental regression function that involves endogenous variable Z (GSize) and exogenous variables X and instrumnets W to compute the partial derivative of a nonparametric estimation. Here the Xs consist of ($\ln Agric_t$, the logarithm of the share of agriculture in the labour force, $\ln Young_t$, the logarithm of the percentage of the population age 16 and younger, and $\ln USiip_t$, the logarithm of the U.S. index of industrial production control for other nongovernmental influences on private output). The Ws consist of the instruments used in the parametric regression: effective participation by the public in the electoral process, $\ln Effectivepart$; the percentage of seats won by the winning political party, $\ln Seats$; and the ratio of immigration to population, $\ln Imratio$). To estimate this relationship we follow the approach of Florens and Racine (2012) where the derivative function ϕ is the solution of an ill-posed inverse problem computed using Landweber-Fridman regularization. The derivative is then used to obtain the tipping point nonparametrically. In Figure 5 we present the derivative function for the model with instrumental variables.

– insert Figure 5 about here –

The derivative plot shows that we have two values of zero one at $Z = 1.4$ and another one at $Z = 3.3$, which confirms the visible maxima and minima in Figure 4.

5.1 The Quadratic versus Cubic form: The role of structural change in the data

In this section we question whether the form of the relationship between government size and per capita output changed over the analyzed period. Consequently we test for structural breaks

using two methods: one that assumes a known break point (see the Chow Test in Table 3) and the other that allows for nonstationary data with an unknown break point (see the Zivot-Andrews test in Table 4). The results of other tests for a structural break are also presented in Table 3. With all these tests we identify a break in the data in 1936. This appears as the deep trough in $\ln PYPC$ in the top panel of Figure 1 (representing the depth of the Great Depression).

Using 1936 as the break point, we divided the data into two sub-periods and re-estimated the parametric and the nonparametric models. In both cases similar period-specific results were found. In particular, an inverted U-shaped curve in the relationship between private output and government size was found for the early (1870-1936) time period while in the subsequent period no such curve could be identified. Consequently the first period of our analysis appears to be the one that has driven the results found for the combined periods.

To see this point more clearly consider the results in Table 5. Column (1) of that table presents the 2SLS estimates with the DOLS correction for 1870 - 1936 using the previous quadratic model in $\ln GSize$. In column (2) the same 2SLS estimate with the DOLS correction is estimated but only for the later 1937 -2011 time period. Finally, in column (3) the same data as in column (2) is used (for 1937-2011) to estimate a linear rather than quadratic model in $LnGSize$. The results show a striking difference between the two periods. In early 1870-1936 period a tipping point can be clearly identified for the private output as a function of government size that is similar in value to that found for the full sample, 6.67%). For the later 1937-2011 time period no tipping point can be identified. It follows that the results for the full sample are likely driven by the first period of analysis. Because of the earlier concern with the use of a specific parametric form, the nonparametric analysis presented in Section 5 was redone for the two separate time periods. Again while no specific time series assumptions are imposed in this analysis, the results confirm descriptively the findings described above. The early period is consistent with a nonlinear relationship between government size and private output that peaks whereas the later time period is not. This is apparent in Figures 6 and 7.

– insert Figure 6 about here –

– insert Figure 7 about here –

Inspection of the diagrams illustrates similar differences arising across the two sub-periods for the effects of the percentage of labour force in agriculture and the effect of a younger labour force on private output. While in the first sub-period the fall in agricultural participation affected private output negatively, the effect was reversed in the second sub-period. For the participation rate of younger labour force we found an effect similar to that of the full sample only for 1870-1936. In the second period the effect becomes insignificant. In Figure 8 we present the derivative function for the model with instrumental variables for the 1870-1936 period:

– insert Figure 8 about here –

The derivative plot shows that we have a value of zero at $Z = 1.6$, which again confirms the visble maxima in Figure 5.

The importance of finding a break in the quadratic form is that it helps resolve the earlier concern that the data could not distinguish clearly between the quadratic and cubic forms estimated over the longer time period. The latter alternative was important because it raised the possibility that further increases in government size could increase rather than decrease private output per capita. Finding the presence of a break in the early quadratic relation however makes it apparent that the addition of a linear trend between government size and output in the post 1937 time period to the quadratic shape arising earlier (when government size was considerably smaller) has produced the appearance of a tailing off of the inverted U shape in the later period (and larger size) that is characteristic of the cubic shape estimated in column (5) of Table 2. That is, the cubic form appears to fit over the full period only because the quadratic relationship between government size and performance ended in the period just before WWII. This interpretation is supported by the complementary finding that the power of the single cubic relationship breaks down when re-estimated over the two sub-periods. In terms of policy significance, finding that the current time period is better approximated by a linear trend rather than the upper tail of a quadratic or cubic function means that the cost of expanding government has remained constant rather than increasing or decreasing in size. While this finding does not overturn the point that the post WWII expansion in government size in Canada (i.e., increasing beyond the estimated tipping point) has come at a real cost in terms of foregone private output, it does diminish the fear that further increases accelerate those costs. Moreover, the presence of a cost need not imply that the output loss was not worthwhile. Rather the post-war expansion of government services to accomplish such redistributive objectives as achieving greater income equality and promoting equity of opportunity may well have been socially beneficial. By re-emphasizing the point that more government means less of everything else, the analysis reinforces the policy obligation to consider and justify net benefit before encouraging further expansion.

6 Conclusion

In this paper we have argued that an inverted U-shaped relationship for the effect of government size on private economic performance holds for Canada with two special caveats. First because of the time series properties of government size and economic performance, the long term link with government size should be to the level rather than growth rate of private output. Second, the appearance of a positive followed by a negative effect of government size on the level of private performance can be found for Canada only for the early 1870 to 1936 time period. For

the later time period the effect of size on output has been linear, consistent with constant rather than increasing cost.

Our identification strategy combines both parametric and nonparametric specifications, with nonparametric estimation used as a robustness check on the quadratic relationship traditionally estimated in the literature. Our main finding is that regardless of the statistical model specified, similar results were always obtained. Finding a tipping point within the neighborhood of a closed confidence set for our nonparametric test replicated the significance of our earlier parametrically estimated tipping points. The dual identification strategy also illustrated the robustness of this result to many of the possible misspecifications typically associated with parametric models.

In terms of the policy implications of our analysis, finding a break in the nonlinear relationship between government size and private economic performance in the period immediately following the Great Depression has significance. In particular, if the quadratic estimates had remained stable over the entire period, then the pattern of continuing increases in government size would have produced ever larger losses in terms of foregone per capita output. Such reasoning has formed the basis for a number of government policies directed explicitly at restraining or reversing the growth of government size. Finding that the quadratic form of that relationship is no longer relevant in the current time period suggests some caution in this regard.[12] While our findings do not challenge the primary conclusion– that more government comes at a cost–the finding of a linear relationship in the current period does suggest that government has become successful in holding down or reversing the earlier tendency for cost to increase. Thus while our finding may be somewhat reassuring with respect to the nature of government costs and the potential cost of larger government size, nothing in the analysis reverses the need for economists to be vigilant when evaluating both new and old programs. In particular the ability to recognize a cost (benefit) in any further expansion (contraction) in size only increases the obligation to weigh appropriately the concomitant gain (loss) in equity and/or efficiency that will appear the other side of the trade-off. Evaluating the latter remains the most difficult challenge facing public policy analysis.

[12]Note that a reversal of the traditional presumption that government size only grows has been evident in some OECD countries for some time. See, for example, Borcherding, Ferris and Garzoni (2004).

7 References

Afonso, A. and D. Furceri 2010 "Government size, composition, volatility and economic growth" *European Journal of Political Economy,* 26: 517-532.

Afxentiou, P.C. and A. Serletis 1991 "A Time-Series Analysis of the relationship between Government Expenditure and GDP in Canada" *Public Finance Quarterly,* 19(3): 316-333.

Andrews, D., Zivot, E. 1992. "Further evidence on the Great Crash, the oil price shock, and the unit-root hypothesis" *Journal of Business and Economic Statistics,* 10: 25 -70.

Armey, D. 1995 *The Freedom Revolution.* Regnery Publishing: Washington.

Barro, R.J. 1990 "Government spending in a simple model of endogenous growth" *Journal of Political Economy,* 98: 103-125.

Bergh, A. and M. Henrekson 2011 "Government Size and Growth: A Survey and Interpretation of the Evidence" *IFN Working Paper 858,* Stockholm: Research Institute of Industrial Economics.

Bernard, J, Gavin M., Khalaf L. and M. Voia 2013 "The Environmental Kuznets Curve: Tipping Points, Uncertainty and Weak Identification" *CREATE Working Paper 2011-4*, 34 pages. forthcoming in *Environmental and Resource Economics.*

Blundell, R., X. Chen and D. Kristensen 2007 "Semi-nonparametric IV estimation of shape invariant Engel curves", *Econometrica,* 75: 1613-575.

Bolduc, D., L. Khalaf and C. Yelou 2010 "Identification robust confidence set methods for inference on parameter ratios with application to discrete choice"*Journal of Econometrics,* 157(2): 317-327.

Borcherding, T., J.S. Ferris and A. Garzoni 2004 Changes in the Real Size of Government Since 1970", in Juergen Backhaus and Richard Wagner (eds.), *Kluwer Handbook in Public Finance,* Kluwer Academic Press, Norwell, Ma., 77-108.

Chao, J. C.P. and H. Grubel 1998 "Optimal Levels of Spending and Taxation in Canada" in H. Grubel ed. *How to Use the Fiscal Surplus.* Vancouver: Fraser Institute, 53-68.

Chow, Gregory C. 1960 "Tests of Equality between Sets of Coefficients in Two Linear Regressions" *Econometrica,* 28 (3): 591- 605.

Darolles, S. Y. Fan, J.P. Florens and E. Renault 2011 "Nonparametric Instrumental Regression" *Econometrica*, 79: 1541-1565.

De Witte, K. and W. Moesen 2010 "Sizing the Government" *Public Choice,* 145: 39-55.

Facchini F. and M. Melki 2013 "Optimal Government Size and Economic Growth in France (1871-2008): An Explanation by the State and Market Failures" *European Journal of Political Economy,* 31:1-14.

Ferris, J.S., Park, S.B. and S.L. Winer 2008 "Studying the Role of Political Competition in the Evolution of Government Size over Long Horizons" *Public Choice,* 137(1-2): 369-401.

Ferris, J.S. 2012 "Government Size, Government Debt and Economic Performance with particular application to New Zealand" *Economic Record ,* (forthcoming 2014).

Fieller E. C. 1954 "Some Problems in Interval Estimation" *Journal of the Royal Statistical Society,* B 16: 175-185.

Florens, J.P. and J.S. Racine 2012, Nonparametric Instrumental Derivatives, *Working Paper*

Ghosh Roy, A. 2009 "Evidence on economic growth and government size" *Applied Economics,* 41: 607-614.

Gillespie, W. I., 1991 *Tax, Borrow and Spend: Financing Federal Spending in Canada, 1867 - 1990.* Carleton University Press, Ottawa.

Grossman, P. 1987 "The Optimal Size of Government" *Public Choice,* 53:131-147.

Horowitz, J.L. 2011 Applied Nonparametric Instrumental Variables Estimation, *Econometrica*, 79, 347-394.

Holmes, J.M. and P. A. Hutton 1990 "On the causal relationship between government expenditures and national income" *Review of Economics and Statistics,* 72(1): 87-95.

Kau, J.B. and P.H. Rubin 1981 "The Size of Government" *Public Choice,* 37: 261-274.

Kormendi, R. and P. Maguire 1985 "Macroeconomic determinants of growth: Cross-country evidence" *Journal of Monetary Economics,* 16(4): 141-163.

Kneebone, R.D. 1992 "Centralization and the size of government in Canada" *Applied Economics,* 24: 1293-1300.

Leacy, F.H., Urquhart, M.C., and K.A.H. Buckley 1983 *Historical Statistics of Canada.* Ottawa: Supply and Services Canada.

Landau, D. 1983 "Government expenditures and economic growth: A cross-country study" *Southern Economic Journal,* 49: 783-92.

Ma S., Racine J. S. and L. Yang 2011 "Additive Regression Splines with Irrelevant Categorical and Continuous Regressors" *Working paper, McMaster University and Michigan State University.*

Ma, S. and J.S. Racine 2013 "Additive Regression Splines with Irrelevant Categorical and Continuous Regressors" *Statistica Sinica,* 23: 515-541.

MacKinnon, J. G. 2010 "Critical Values for Cointegration Tests" *Queens Economics Department Working Paper* No. 127.

Mourmouras, I.A., and Lee, J.E. 1999 "Government spending on infrastructure in an endogenous growth model with finite horizons" *Journal of Economics and Business,* 51: 395-407.

Nie, Z. and J.S. Racine 2012 "The crs package: Nonparametric Regression Splines for Continuous and Categorical Predictors," *The R Journal,* 4: 48-56.

Oehlert G.W. 1992 "A note on the delta method" *American Statistician,* 46: 2729.

Petry, F., Imbeau L., Crete J. and M. Clavet 2000 "Explaining the evolution of Government size in the Canadian Provinces" *Public Finance Review,* 28(1): 26-47.

Ram, R. 1986 "Government size and economic growth: A new framework and some evidence from cross-section and time-series", *American Economic Review,* 76: 191-203.

Romero-Avila, D. and R. Strauch 2008 "Public finances and long-term growth in Europe: Evidence from a panel data analysis" *European Journal of Political Economy,* 24: 172-191.

Scully, G.W. 1989 "The size of the state, economic growth and the efficient utilization of national resources" *Public Choice,* 63: 149-164.

Singh, B. and B.S. Sahni 1984 Causality between public expenditure and national income *Review of Economics and Statistics,* 66: 630-644.

Urquhart, M.C. 1993 *Canadian Historical Statistics and Gross National Product: 1870-1926.* McGill-Queen's University Press.

Voia, M. and J.S. Ferris 2013 "Do Canadian business cycle peaks predict Federal election calls?" *European Journal of Political Economy,* 29(1): 102-118.

Winer, S.L. and J.S. Ferris 2008 "Searching for Keynesianism" *European Journal of Political Economy,* 24(2): 294-316.

8 Appendices

8.1 Appendix A: Economic and Political Variables

AGRIC = percentage of the labour force in agriculture 1870-1925 M.C. Urquhart (1993) *Gross National Product, Derivation of Estimates,* p.24; 1926-1975 Cansim D31251/D31252; 1975-2011 Cansim II v2710106/v2710104

GNP = gross national product in current dollars. 1870-1926: Urquhart (1993: 24-25) (in millions); 1927-1938: Leacy et al. (1983: 130); 19391960 Canadian Economic Observer (Table 1.4), CANSIM D11073 = GNP at market prices. 1961-2011 CANSIM I D16466 = CANSIM II V499724 (aggregated from quarterly data).

GOV = total federal government expenditure net of interest payments.1870-1989: Gillespie (1991: 284-286); 1990-1996: Public Accounts of Canada 1996-97: 1997-2000: Federal Government Public Accounts, Table 3 Budgetary Revenues Department of Finance web site, September 2001. To this we add the return on government investment (ROI) originally subtracted by Gillespie for his own purposes. Expenditure is net of interest paid to the private sector. Data on ROI: 1870 to 1915: Public Accounts (1917: 64); 1915-1967: Dominion Government Revenue and Expenditure: Details of Adjustments 1915-1967 Table W-1; 1916-17 to 1966-67: Securing Economic Renewal - The Fiscal Plan, Feb 10, 1988, Table XI; 1987-88 to 1996-97: Public Accounts 1996, Table 2.2. Interest on the Debt (ID) was subtracted out (with adjustment for interest paid to the Bank of Canada (BCI) ultimately returned to the government). Data on ID: 1870-1926: Leacy et al. (1983: Series H19-34): Federal Government budgetary expenditures, classified by function, 1867-1975; 1926-1995: Cansim D11166. 1996-2000: Cansim D18445. Finally, data for BCI: copied by hand from the Annual Reports of The Bank of Canada, Statement of Income and Expense, Annually, 1935-2000. Net Income paid to the Receiver General (for the Consolidated Revenue Acct). Note: all government data are converted from fiscal to calendar years, and allows for a change in the definition of the fiscal year in 1906/07, as described in Gillespie (1991: Appendix C).

GSIZE = non-interest federal government, direct public expenditure, calculated as: GOV/GNP.

LnGSize = Log(GSIZE); **Ln GSizesq** = LnGSize*LnGSize; LnGSizecb = LnGSizesq*LnGSize.

Immigration 1870 1953 O.J. Firestone Canadas Economic Development 1867-1953 Table 83, Population, Families, Births, Deaths (in thousands); 1954-1995; Cansim D27; 1996-2011 Cansim II v16.

Imratio = Immigration/POP

P = GNP deflator before 1927 and GDP deflator after (1986 = 100). 1870-1926: Urquhart, (1993), 24-25; 1927-1995 (1986=100): Cansim data label D14476; 1996-2006 Cansim D140668. All indexes converted to 1986 = 100 basis.

POP = Canadian population size, 1870 1926, M.C. Urquhart (1993), Gross Nation Product of

Canada, pp.24,25; 1927-1955 Cansim data label D31248; 1956-2011 Cansim II Table 051-0005.
PYPC = (GNP GOV)/(P*POP); **LNPYPC** = Log(PYPC).
USiip = US Index of Industrial Production: 1870-1918: Table A15 NBER Nutter 1868-1929; 1919 2011 (INDPRO) Board of Governors of the Federal Reserve System, G.17 Industrial Production and Capacity Utilization, Seasonally Adjusted, Monthly, Index 2007=100. Note: Average of 12 months.
Young = percentage of the population 16/17 and younger; 1870 -1920, Lacey et al (1983) Interpolated from Census figures Table A28 -45 sum of columns 29,30,31, and 32, all divided by 28; 1921-1970 Cansim C892547; 1971-2011 Cansim II v466965.
Effective_part(icipation) = percentage of the population voting in the federal election = fraction of the population registered to votes times voter turnout Source: Elections Canada web site,
$www.elections.ca/pastelections/AHistoryoftheVoteinCanada$: $Appendix$ **Seats** = percentage of seats won by the winning party. Electoral Results by Election, 1867 to Date. Web site of the Parliament of Canada: *www.parl.gc.ca.*
Winningmargin = the percentage of the votes won the winning party minus the percentage won by its closest rival multiplied by years from the closets election, Electoral results by Party, 1867 to Date. Web site of the Parliament of Canada: *www.parl.gc.ca.*

8.2 Appendix B.1: Identification of the parametric tipping point - using quadratics of $\ln GSize$

We follow a similar justification for confidence set identification as in Bernard, Gavin, Khalaf and Voia (2013). If β_1 and β_2 in equation (1) are well identified, consistent and asymptotically normal estimates [denoted $\hat{\beta}_1$ and $\hat{\beta}_2$] with a tractable variance/covariance matrix [denoted $\hat{\Sigma}_\beta$] are readily available. Plugging $\hat{\beta}_1$ and $\hat{\beta}_2$ in equation (2) yields a consistent estimate for δ denoted $\hat{\delta}$. Furthermore, given $\hat{\Sigma}_\beta$, a standard error [denoted $\hat{\Sigma}_\delta^{1/2}$] can easily be obtained for inference on δ via the Delta method. The usual t-statistic

$$\mathtt{t}_D\left(\delta_0\right) = (\hat{\delta} - \delta_0)/\hat{\Sigma}_\delta^{1/2}$$

associated with the Null Hypothesis that $\delta = \delta_0$ where δ_0 is any known value, which yields usual significance tests as well as the standard confidence interval for inference on δ, is thus easy to derive. Assuming the considered estimator, $\hat{\delta} = \exp(\hat{\beta}_1/(-2\hat{\beta}_2))$ where $\hat{\beta}_1$ and $\hat{\beta}_2$ is consistent and asymptotically normal, the so-called Delta method produces the following $1-\alpha$

level confidence interval for δ:

$$\lim DCS\left(\delta;\alpha\right)=\left[\hat{\delta}\pm z_{\alpha/2}\hat{\Sigma}_{\delta}^{1/2}\right],\ \hat{\Sigma}_{\delta}=\hat{G}'\hat{\Sigma}_{12}\hat{G},\ \hat{G}=\left[\begin{array}{c}\frac{\partial\hat{\delta}}{\partial\hat{\beta}_1}\\ \frac{\partial\hat{\delta}}{\partial\hat{\beta}_2}\end{array}\right]=\left[\begin{array}{c}\frac{\exp(\hat{\beta}_1/(-2\hat{\beta}_2))}{-2\hat{\beta}_2}\\ \frac{(\hat{\beta}_1/2)\exp(\hat{\beta}_1/(-2\hat{\beta}_2))}{\hat{\beta}_2^2}\end{array}\right] \tag{5}$$

where $z_{\alpha/2}$ refers to the two-tailed α-level standard normal cut-off point, and

$$\hat{\Sigma}_{12}=\left[\begin{array}{cc}\hat{v}_1 & \hat{v}_{12}\\ \hat{v}_{12} & \hat{v}_2\end{array}\right]$$

refers to the subset of the variance/covariance matrix of the estimates $\hat{\beta}_1$ and $\hat{\beta}_2$. The Delta Confidence Set $\left[\hat{\delta}\pm z_{\alpha/2}\hat{\Sigma}_{\delta}^{1/2}\right]$ collects the values of δ_0 for which the $\mathtt{t}$-statistic

$$\mathtt{t}_D\left(\delta_0\right)=\left(\hat{\delta}-\delta_0\right)/\hat{\Sigma}_{\delta}^{1/2}$$

associated with

$$\mathcal{H}_D(\delta_0):\delta-\delta_0$$

is not significant at the α level. Said differently, if we solve the inequality

$$\left|\hat{\delta}-\delta_0\right|/\hat{\Sigma}_{\delta}^{1/2}<z_{\alpha/2} \tag{6}$$

for δ_0, we get $\left[\hat{\delta}\pm z_{\alpha/2}\hat{\Sigma}_{\delta}^{1/2}\right]$. In statistics, solving for δ_0 in (6) is known as "inverting the $\mathtt{t}_D\left(\delta_0\right)$ test", where inverting a test with respect to a parameter means collecting all values (here δ_0) not rejected by this test at the α level. If the Delta method is applied when estimating weakly-identified parameters, denominator close to zero, it produces very tight confidence intervals that are focused on "wrong" values, which results in a poor coverage. Also, the estimated intervals will fail to cover the true parameter value, and given the tightness of the interval estimate, this will go unnoticed. These problems can be avoided if one applies Fieller based confidence set estimation method. The Fieller method inverts an alternative $\mathtt{t}$-statistic

$$\mathtt{t}_F\left(d_0\right)=(\hat{g}_1-d_0\hat{g}_2)/[(\hat{v}_1+d_0^2\hat{v}_2-2d_0\hat{v}_{12})^{1/2}]$$

associated with

$$\mathcal{H}_F(d_0):g_1-d_0g_2=0.$$

Inverting $\mathtt{t}_F\left(d_0\right)$ requires solving for the set of d_0 values that are not rejected at level α using $\mathtt{t}_F\left(d_0\right)$ and a standard normal two-tailed cut-off $\mathtt{z}_{\alpha/2}$. In other words, we need to collect the d_0

values such that $|\mathtt{t}_F(d_0)| \leq z_{\alpha/2}$ or alternatively such that

$$(\hat{g}_1 - d_0\hat{g}_2)^2 \leq \mathtt{z}^2_{\alpha/2}(\hat{v}_1 + d_0^2\hat{v}_2 - 2d_0\hat{v}_{12}) \tag{7}$$

which is a second degree inequality in d_0. The resulting solution denoted $\lim FCS(d;\alpha)$ is either a bounded interval, an unbounded interval, or the entire real line, where the unbounded solutions occurs when the denominator is close to zero. $\mathcal{H}_F(d_0)$ can be linked to $\mathcal{H}_D(\delta_0)$ for $d_0 = \log(\delta_0)$. Because $\lim FCS(d;\alpha)$ is obtained as described, taking the exponential of its limits provides the desired confidence set for δ. This corresponds to projecting the region implied by $\lim FCS(d;\alpha)$. We can see that by replacing d_0 by $\log(\delta_0)$ in (7) does not distort the inequality nor its statistical foundations. Solving for δ_0 in the resulting inequality would numerically coincide with applying the exponential to the limits of $\lim FCS(d;\alpha)$. The $\mathtt{t}_D(\delta_0)$ test and $\lim DCS(\delta;\alpha)$ interval will have α and $(1-\alpha)$ as effective levels, if $\mathtt{t}_D(\delta_0) \overset{asy}{\sim} N(0,1)$. The $\mathtt{t}_F(d_0)$ test and $\lim FCS(d;\alpha)$ interval will also achieve level control if $\mathtt{t}_F(d_0) \overset{asy}{\sim} N(0,1)$. Because $\mathcal{H}_D(\delta_0)$ requires $g_2 \neq 0$, the Gaussian approximation fails for $\mathtt{t}_D(\delta_0)$. In contrast, and because $\mathcal{H}_F(d_0)$ does not require restricting the parameter space of β_2, nor of β_1 for that matter

$$\hat{\delta} \text{ is asymptotically normal} \Rightarrow \mathtt{t}_F(d_0) \overset{asy}{\sim} N(0,1).$$

It follows that the $\lim FCS(d;\alpha)$ will achieve level control whether b_2 is zero or not. The Fieller method requires solving inequality (7) for d_0, which may be rexeprresd as

$$Ad_0^2 + 2Bd_0 + C \leq 0 \tag{8}$$

$$A = \hat{g}_2^2 - z^2_{\alpha/2}\hat{v}_2, \quad B = -\hat{g}_1\hat{g}_2 + z^2_{\alpha/2}\hat{v}_{12}, \quad C = \hat{g}_1^2 - z^2_{\alpha/2}\hat{v}_1. \tag{9}$$

Except for a set of measure zero, $A \neq 0$. Similarly, except for a set of measure zero, $\Delta = B^2 - AC \neq 0$. Real roots

$$d_{01} = \frac{-B - \sqrt{\Delta}}{A}, \quad d_{02} = \frac{-B + \sqrt{\Delta}}{A}$$

exist if and only if $\Delta > 0$, so

$$\lim FCS(d;\alpha) = \begin{cases} [d_{01}, \quad d_{02}] & if \quad A > 0 \\ (-\infty, \quad d_{01}] \cup [d_{02}, \quad +\infty) & if \quad A < 0 \end{cases}. \tag{10}$$

Bolduc, Khalaf and Yelou (2010) further show that: (i) if $\Delta < 0$, then $A < 0$ and $\lim FCS(d;\alpha) = \mathbb{R}$; (ii) $\lim FCS(d;\alpha)$ contains the point estimate $\hat{g}_1/\hat{g}_2$ and thus cannot be empty, and (iii) asymptotically, Fieller's solution and the *Delta* method give similar results when the former leads to an interval, *i.e.* when the denominator is far from zero. Taking the exponential of the

limits of $\lim FCS\,(d;\alpha)$ provides a confidence set for $\exp(d)$.

8.3 Appendix B.2: Identification of the parametric tipping point - using cubics for $\ln GSize$

For the cubic model in $\ln GSize$:

$$\ln\left(\frac{y_t}{N_t}\right) = \beta_0+\beta_1 \ln GSize_t+\beta_2 \ln GSize_t^2+\beta_3 \ln GSize_t^3+\beta_4 \ln Agric_t+\beta_5 \ln Young_t+\beta_6 \ln USiip_t+u_t \tag{11}$$

With parameters β_1, β_2 and β_3, where β_3 is not weakly identified, we can construct the confidence sets using the Delta method for both possibile tipping points (min and max). Consider that the first tipping point is a maximum and the second tipping point is a minimum as identified in our regression. The two tipping points are identified by solving:

$$\frac{\partial \ln\left(\frac{y_t}{N_t}\right)}{\partial \ln GSize} = 0. \tag{12}$$

Solving the above equation gives:

$$\ln GSize_{max;min} = \frac{-\beta_2 \mp \sqrt{\beta_2^2 - 3\beta_1\beta_3}}{3\beta_3}. \tag{13}$$

The tipping points for maximum and minimum can be then obtained by taking the exponential of $\ln GSize_{min;max}$, which is a monotone transformation of the transformed log measure obtaiend from the regression:

$$GSize_{max;min} = e^{\frac{-\beta_2 \mp \sqrt{\beta_2^2 - 3\beta_1\beta_3}}{3\beta_3}}. \tag{14}$$

Consider $\theta = \{\beta_1, \beta_2, \beta_3\}$ is the set of the three parameters used to identify the tipping points. The variances for the two tipping points are obtained using Delta method as foolws:

$$Var(GSize_{max;min}) = D\Sigma D' = \left(\frac{\partial GSize_{min;max}}{\partial \theta}\right)\Sigma\left(\frac{\partial GSize_{min;max}}{\partial \theta}\right)' \tag{15}$$

where

$$\frac{\partial GSize_{max;min}}{\partial \theta} = \left(\frac{\partial GSize_{mamx;min}}{\partial \beta_1}\ \frac{\partial GSize_{max;min}}{\partial \beta_2}\ \frac{\partial GSize_{max;min}}{\partial \beta_3}\right)$$

and Σ is defined as:

$$\Sigma = \begin{pmatrix} \sigma^2_{\beta_1} & \sigma_{\beta_1\beta_2} & \sigma_{\beta_1\beta_3} \\ \sigma_{\beta_2\beta_1} & \sigma^2_{\beta_2} & \sigma_{\beta_2\beta_3} \\ \sigma_{\beta_3\beta_1} & \sigma_{\beta_3\beta_2} & \sigma^2_{\beta_3} \end{pmatrix}.$$

For our specific case where the first tipping point is a maximum, $\frac{\partial GSize_{max}}{\partial \beta_1}$, $\frac{\partial GSize_{max}}{\partial \beta_2}$ and $\frac{\partial GSize_{max}}{\partial \beta_3}$ are:

$$\frac{\partial GSize_{max}}{\partial \beta_1} = e^{\frac{-\beta_2-\sqrt{\beta_2^2-3\beta_1\beta_3}}{3\beta_3}} \left(\frac{(\beta_2^2 - 3\beta_1\beta_3)^{-0.5}}{2} \right),$$

$$\frac{\partial GSize_{max}}{\partial \beta_2} = e^{\frac{-\beta_2-\sqrt{\beta_2^2-3\beta_1\beta_3}}{3\beta_3}} \left(\frac{-1 - \beta_2(\beta_2^2 - 3\beta_1\beta_3)^{-0.5}}{3\beta_3} \right),$$

and

$$\frac{\partial GSize_{max}}{\partial \beta_3} = e^{\frac{-\beta_2-\sqrt{\beta_2^2-3\beta_1\beta_3}}{3\beta_3}} \left(\frac{\beta_2 + 3\beta_1\beta_3(\beta_2^2 - 3\beta_1\beta_3)^{-0.5} + (\beta_2^2 - 3\beta_1\beta_3)^{0.5}}{3\beta_3^2} \right).$$

When the second tipping point is a minimum, $\frac{\partial GSize_{min}}{\partial \beta_1}$, $\frac{\partial GSize_{min}}{\partial \beta_2}$ and $\frac{\partial GSize_{min}}{\partial \beta_3}$ are:

$$\frac{\partial GSize_{min}}{\partial \beta_1} = -e^{\frac{-\beta_2+\sqrt{\beta_2^2-3\beta_1\beta_3}}{3\beta_3}} \left(\frac{(\beta_2^2 - 3\beta_1\beta_3)^{-0.5}}{2} \right),$$

$$\frac{\partial GSize_{min}}{\partial \beta_2} = e^{\frac{-\beta_2+\sqrt{\beta_2^2-3\beta_1\beta_3}}{3\beta_3}} \left(\frac{-1 + \beta_2(\beta_2^2 - 3\beta_1\beta_3)^{-0.5}}{3\beta_3} \right),$$

and

$$\frac{\partial GSize_{min}}{\partial \beta_3} = e^{\frac{-\beta_2+\sqrt{\beta_2^2-3\beta_1\beta_3}}{3\beta_3}} \left(\frac{\beta_2 - 3\beta_1\beta_3(\beta_2^2 - 3\beta_1\beta_3)^{-0.5} - (\beta_2^2 - 3\beta_1\beta_3)^{0.5}}{3\beta_3^2} \right).$$

Therefore $Var(GSize_{max,min})$ can be expressed as:

$$Var(GSize_{max,min}) = \left(\frac{\partial GSize_{max,min}}{\partial \beta_1} \right)^2 \sigma_{\beta_1}^2 + 2 \left(\frac{\partial GSize_{max,min}}{\partial \beta_1} \right) \left(\frac{\partial GSize_{max,min}}{\partial \beta_2} \right) \sigma_{\beta_1\beta_2} +$$

$$+2 \left(\frac{\partial GSize_{max,min}}{\partial \beta_1} \right) \left(\frac{\partial GSize_{max,min}}{\partial \beta_3} \right) \sigma_{\beta_1\beta_3} + \left(\frac{\partial GSize_{max,min}}{\partial \beta_2} \right)^2 \sigma_{\beta_2}^2$$

$$+2 \left(\frac{\partial GSize_{max,in}}{\partial \beta_2} \right) \left(\frac{\partial GSize_{max,min}}{\partial \beta_3} \right) \sigma_{\beta_2\beta_3} + \left(\frac{\partial GSize_{max,min}}{\partial \beta_3} \right)^2 \sigma_{\beta_3}^2$$

With the help of Delta method the following confidence interval $1-\alpha$ level confidence interval for $GSize_{max,min}$:

$$DCS\left(GSize_{max,min};\alpha\right) = \left[\widehat{GSize_{max,min}} \pm z_{\alpha/2}\sqrt{Var(\widehat{GSize_{max;min}})} \right] \tag{16}$$

where $z_{\alpha/2}$ refers to the two-tailed α-level standard normal cut-off point.

8.4 Appendix C: B-splines

The method from Ma and Racine (2013) uses a B-spline function for $f(.)$, which is a linear combination of B-splines of degree m defined as follows

$$\mathcal{B}(x) = \sum_{c=0}^{N+m} b_c B_{c,m}(x), \quad x \in [k_0, k_{N+1}]$$

where b_c are denoted "control points", $k_0, ..., k_{N+1}$ are known as a knot sequence [an individual term in this sequence is known as a knot],

$$B_{c,0}(x) = \left\{ \begin{array}{cc} 1 & k_c \leq x < k_{c+1} \\ 0 & otherwise \end{array} \right\}$$

which is referred to as the 'intercept', and

$$\begin{aligned} B_{c,j+1}(x) &= a_{c,j+1}(x) B_{c,j}(x) + [1 - a_{c+1,j+1}(x)] B_{c+1,j}(x), \\ a_{c,j+1}(x) &= \left\{ \begin{array}{cc} \frac{x-k_c}{k_{c+j}-k_c} & k_{c+j} \neq k_c \\ 0 & otherwise \end{array} \right\}. \end{aligned}$$

The unknown function $f(.)$ is estimated by least squares as

$$\hat{\mathcal{B}}(GovSize_{it}; covariates_t) = argmin_{\mathcal{B}(.)} \sum_{i=1}^{n} \sum_{t=1}^{T} \left[\frac{Y}{N_t} - \mathcal{B}(covariates_t) \right]^2.$$

Explicitly, this requires the estimation of the control points b_c. If covariates are considered endogenous and instruments provided, 2SLS is also possible. Underlying best fit parameters are selected by cross-validation; see Racine and Yang (2011) for further details. Further description of this R-package is available at: http://cran.r-project.org/web/packages/crs/crs.pdf.

8.5 Appendix D: nonparametric IV

Given the outcome Y (private output), the instrumental regression function $\phi(GSize, Controls)$ is identified through an instrumental variable but is not assumed to be known up to finitely many parameters.

$$Y = \phi(Gsize, X) + U; \tag{17}$$

with

$$E(U|W) = 0,$$

for endogeneous (Gsize), exogeneous (X), all instruments (W) and ϕ a function that satisfies regularity conditions but is otherwise unknown. The equation above define the structural econometric model:

$$E[Y - \phi(Gsize, X)|W] = 0. \tag{18}$$

We can estimate the function ϕ, by using Horowitz (2011) . In particular if we denote by $\hat{\phi}_j$, with $j = 1, ..., J_n$ the generalized Fourier coefficients of $\hat{\phi}$, we have:

$$\hat{\phi} = \sum_{j=1}^{J_n} \hat{\phi}_j \psi_j(x), \tag{19}$$

where $\psi_j(x)$ can be trigonometric functions, orthogonal polynomials, or splines. If we define by $\hat{\Phi} = (\hat{\phi}_1, ..., \hat{\phi}_J)$

$$\hat{\Phi} = (W_n' X_n)^{-1} W_n Y_n, \tag{20}$$

$\hat{\Phi}$ has the form of an IV estimator for a linear model in which the matrix of variables is X_n and the matrix of instruments is W_n . When the number of obsrevation n is small, $\hat{\phi}_j$ can have higher variation and therefore more unstable. To stabilize $\hat{\phi}_j$, Blundell, Chen and Kristensen (2007) is used as they provide an analytical solution to this problem that it is easy to compute.

9 Tables

Table 1. Descriptive Statistics and Data Sources

Variable Name	Mean	Maximum	Minimum	Std. Deviation	ADF Stat	ADF of first difference
LNPYPC	8.67	10.08	7.33	0.834	-0.091	-10.02***
lnGSize	2.26	3.77	1.25	0.603	-1.84	-7.47***
lnGSizesq	5.49	14.19	1.57	2.77	-1.88	-8.28***
lnAgric	2.78	4.07	0.54	1.17	2.02	-7.67***
lnYoung	3.55	3.88	2.99	0.232	0.93	-3.87***
lnUSiip	5.43	7.74	2.43	1.59	-1.44	-12.07***
lnImratio	-0.297	1.66	-2.67	0.85	-2.58*	-10.06***
lnEffective-part	3.38	3.96	2.1	0.58	-2.22	-12.14***
lnSeats	4.07	4.36	3.7	0.16	-4.91***	
lnWinningmargin	-2.38	-0.99	-3.41	0.54	-2.11	-11.74***

*** (**) [*] significantly different from zero at 1%(5%)[10%] using the MacKinnon (1996) critical values -3.48 (-2.88) [-2.58].

Table 2. The Effect of Government Size on Private Output Per Capita (full sample)

Regression type	*LNPYPC* OLS	*LNPYPC* DOLS	*LNPYPC* 2SLS	*LNPYPC* 2SLS; DOLS	*LNPYPC* 2SLS; DOLS
LnGSize	0.148	0.00003	1.279*	0.830**	10.59**
	1.64	0.001	1.84	2.03	2.46
LnGSizesq	-0.059	-0.033	-0.304**	-0.222**	-4.59**
	-3.24	-0.86	-2.05	-2.53	-2.21
LnGSizecb					0.614**
					2.29
LnAgric	-0.226	-0.260***	-0.190***	-0.231***	-0.479***
	-12.78	-10.71	-5.53	-9.33	-4.11
LnYoung	-0.233	-0.025	-0.293***	0.055	0.956**
	-6.48	-0.35	-3.12	0.61	2.35
LnUSiip	0.364	0.384***	0.362***	0.412***	0.439***
	26.7	19.68	12.65	19.31	17.4
CONSTANT	7.1	6.39***	6.17***	5.16***	-2.076
	53.93	20	8.44	9.7	-0.58
No. of Observations	142	137	142	137	137
Adj R2	0.995		0.986	0.997	0.995
Durbin-Watson (5, 142)	0.496				
ADF^M	-4.95**	-6.07***	-4.34*	-6.12***	-5.93***
Test of endogenous regressors: P-value			0.048	0.014	0.0009
Estimated Tipping point			8.14%	6.49%	6.11%

Second row of results represent the t-stat for each susequent table.
* (**) [***] significantly different from zero at 10, (5), and [1] percent
ADF^M MacKinnon (2010) critical cointegration values at $5\% **(1\% ***) : -4.82(-5.43)$ The DOLS versions regressions use the first difference of each covariate together with their two leads and lags. The 2SLS versions use as instruments for government size the two political variables: the logarithm of effective electoral participation and percentage of seats held by the winning political party, together with the immigration rate.

Table 3. Structural Change Tests for the Proposed Conditional Model of ln(GDP-Gov):
H_o: no Structural Change

Test	Value	P-Value
Chow Test [K, N-2*K]	4.7938	0.0002
Fisher Test [N2,(N1-K)]	8.7768	0.000
Wald Test	31.4181	0.0000
Likelihood Ratio Test	28.3827	0.0000
Lagrange Multiplier Test	25.7261	0.0001

Note: Test performed for N_1: 1st Period Obs = 67 and N_2: 2nd Period Obs = 75

Table 4. Zivot-Andrews unit root test allowing for breaks in ln(GDP-Gov)

Test	Minimum t-statistics at year 1936	Critical Value at 5 %
Zivot-Andrews	-3.924	-4.11

Note: Break identified at year 1936 (obs 67)

Table 5. The Effect of Government Size on Private Output Per Capita:
Case I: 1870 - 1936 and Case II: 1937-2011

Model	2SLS; DOLS: *LNPYPC*	2SLS; DOLS: *LNPYPC*	2SLS; DOLS: *LNPYPC*
Subsample	1870 - 1936	1937 - 2011 (1)	1937 - 2011 (2)
LnGSize	5.187***	0.149	-0.179***
	8.20	0.31	-3.65
LnGSizesq	-1.366***	-0.049	
	-8.20	-0.60	
LnAgric	0.581**	-0.234***	-1.189
	1.99	-4.43	-4.12
LnYoung	2.485***	-0.143	-0.36
	-6.48	-1.35	-0.68
LnUSiip	0.764***	0.391***	0.484***
	6.87	7.19	8.20
CONSTANT	-8.55***	6.51***	6.162***
	-3.07	17.37	17.95
No. of Observations	64	73	73
$AdjR^2$	0.995	0.999	0.998
Estimated Tipping point	6.67%	0%	0%

* (**) [***] significantly different from zero at 10, (5), and [1] percent.

The DOLS versions regressions use the first difference of each covariate together with their two leads and lags.

The 2SLS versions use as instruments for government size two political variables: the logarithm of effective electoral participation and percentage of seats held by the winning political party, together

with the immigration rate.

(1) - quadratic model in *Ln GSize*; (2) - linear model in *Ln GSize*.

Figure 1: The Logarithms of Federal Government Size, Private per capita Output and its Growth Rate: Canada 1870 2011

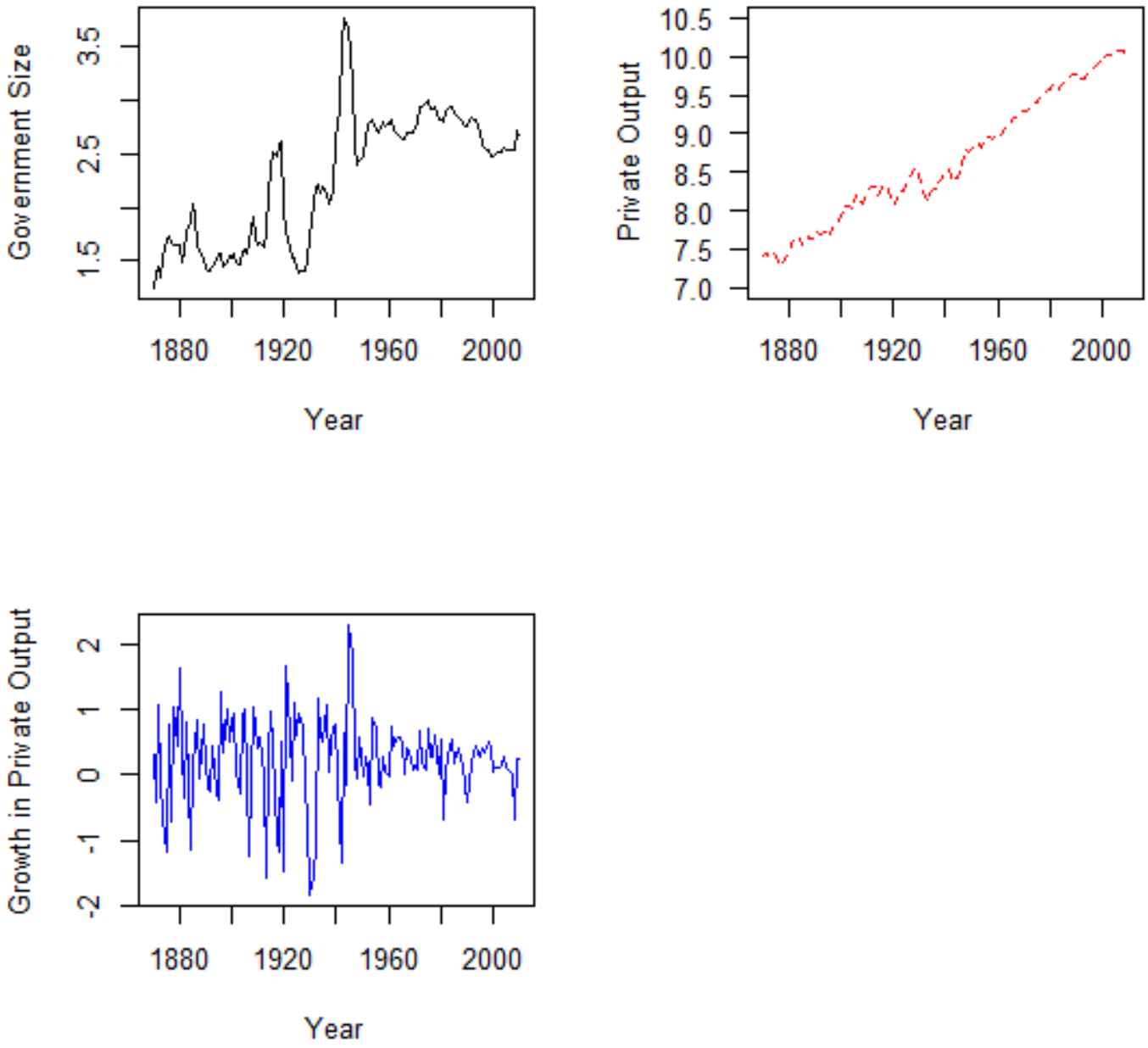

Figure 2: The Effects of Government Size and Tipping Point

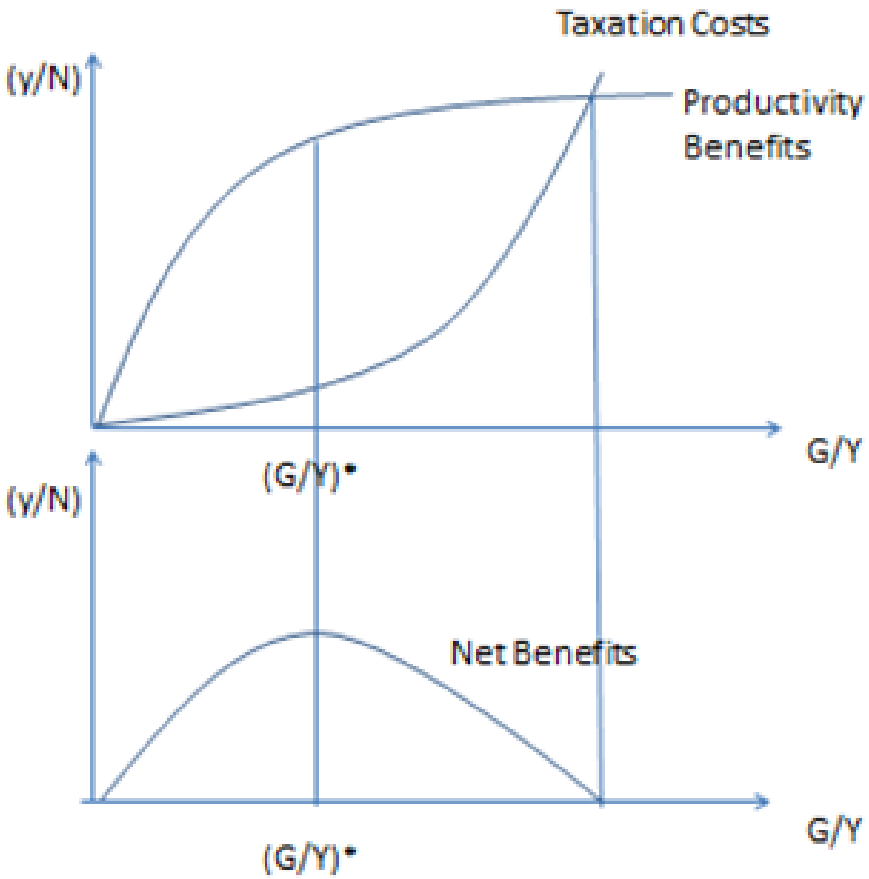

Figure 2 The Effects of Government Size and Tipping Point

Figure 3: ln(GDP-Gov) = f(GSize,X), no instruments, full period

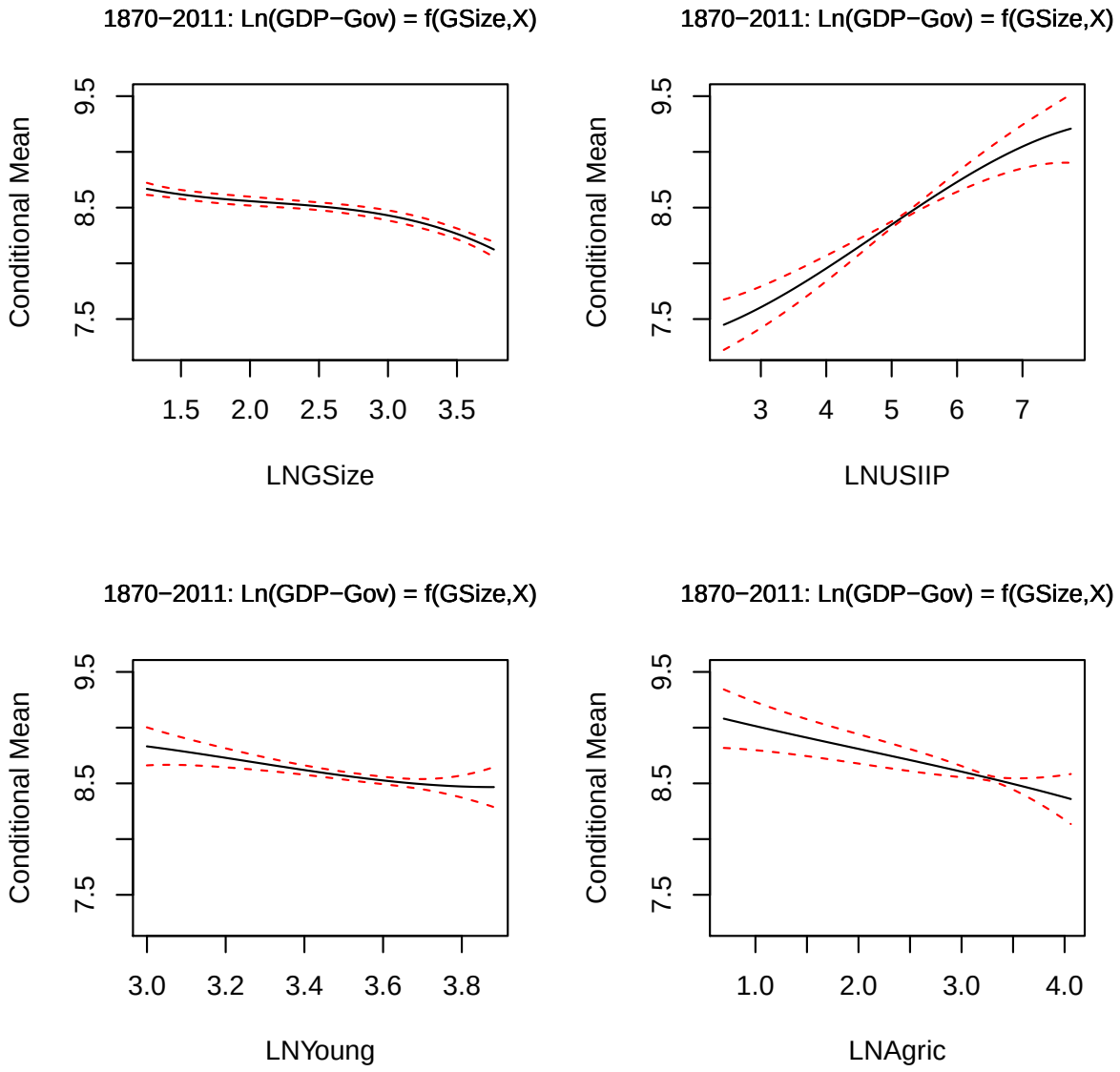

Figure 4: ln(GDP-Gov) = f(GSize,X), with instruments, full period

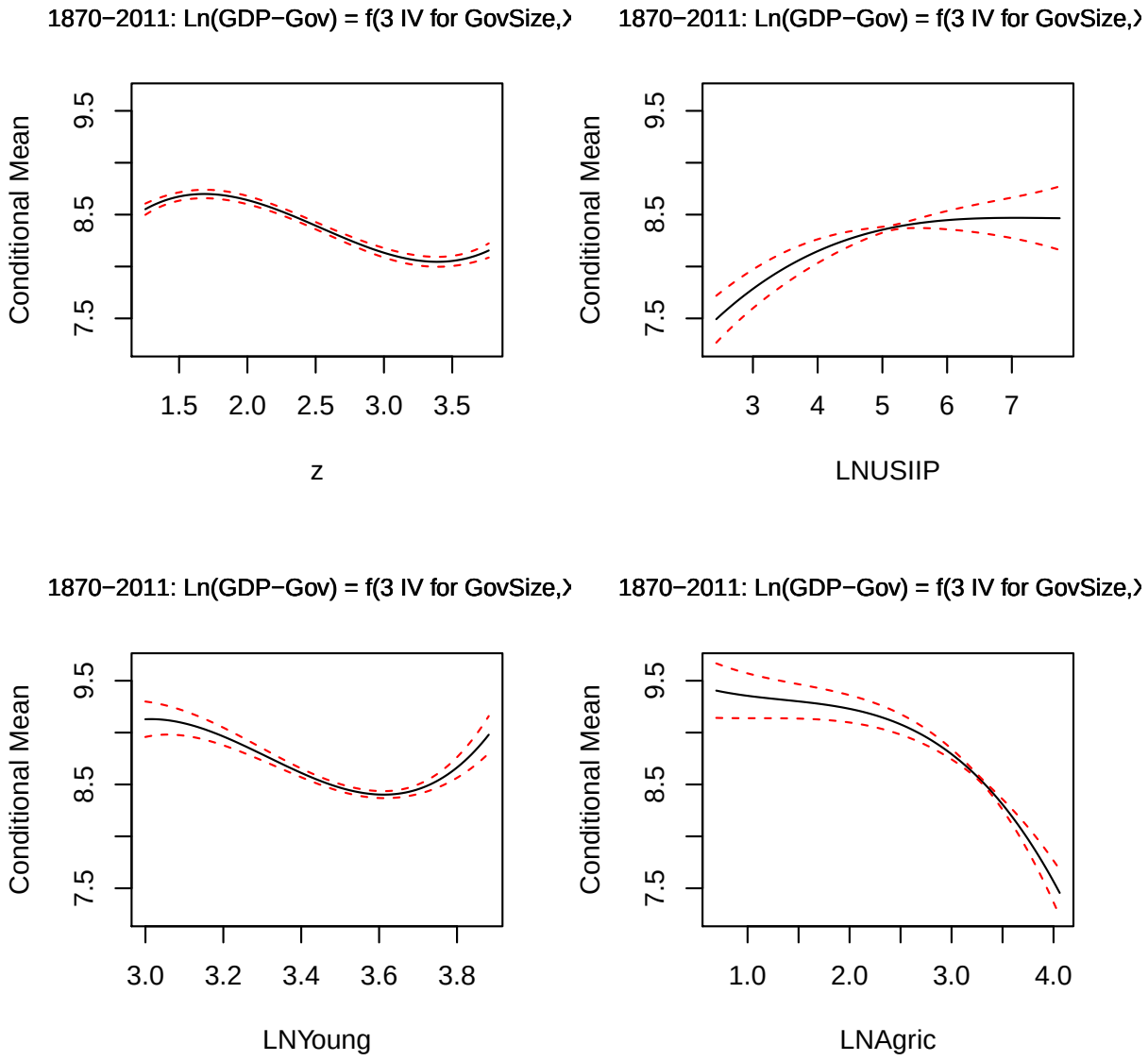

Figure 5: Derivative = f(Z=$\widehat{GSize}$), 1870-2011

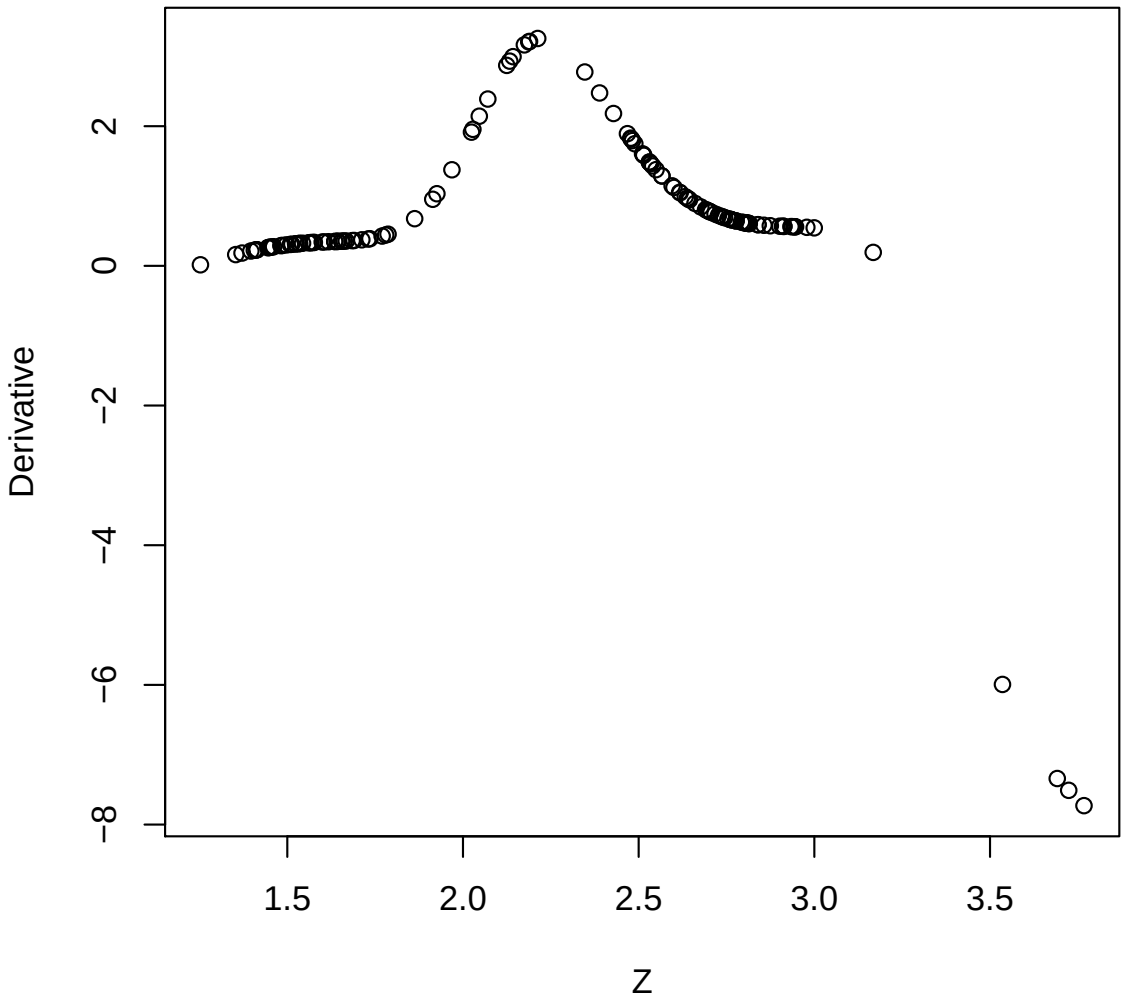

Figure 6: ln(GDP-Gov) = f(GSize,X), with instruments, 1870-1936

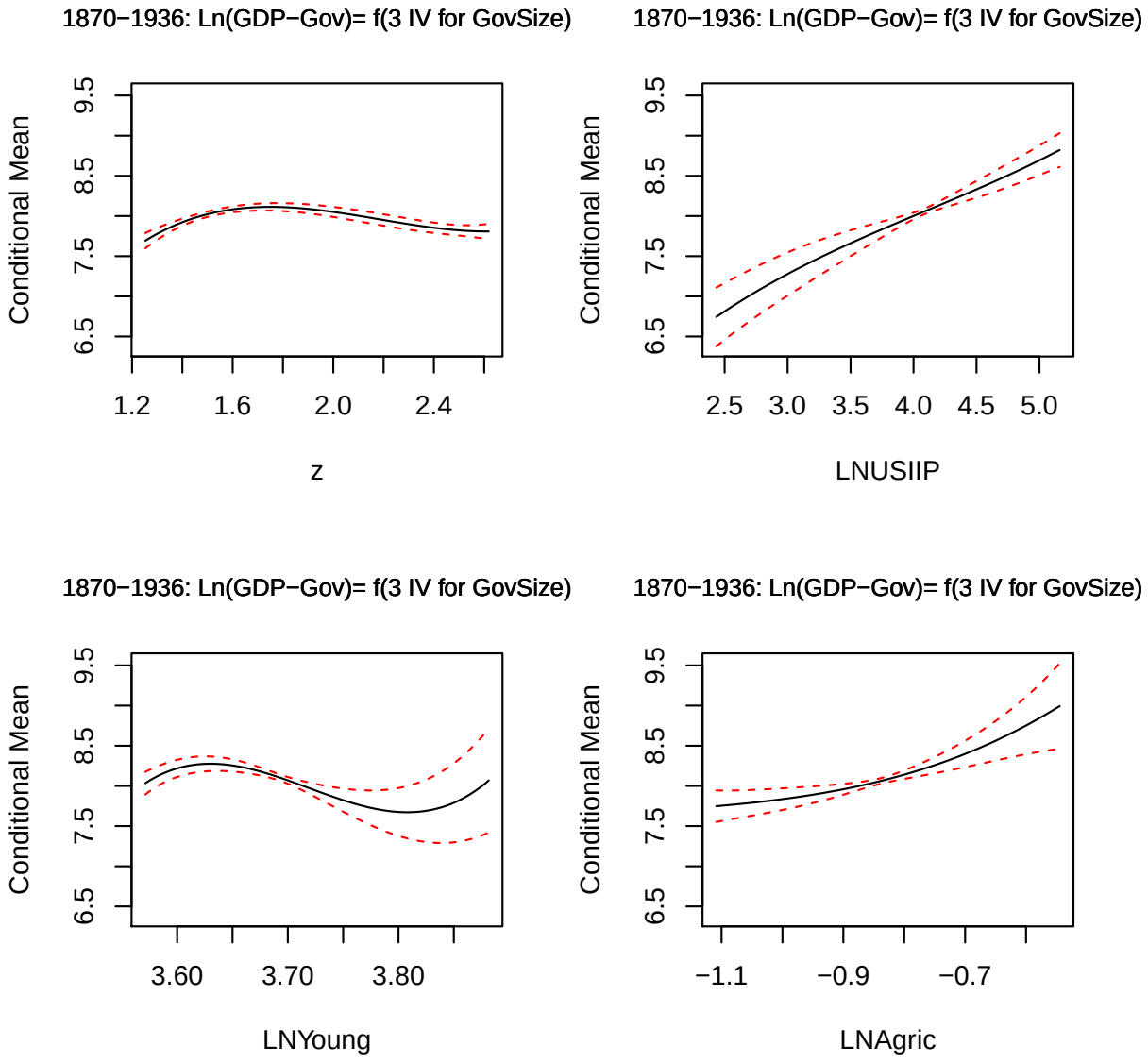

Figure 7: ln(GDP-Gov) = f(GSize,X), with instruments, 1937-2011

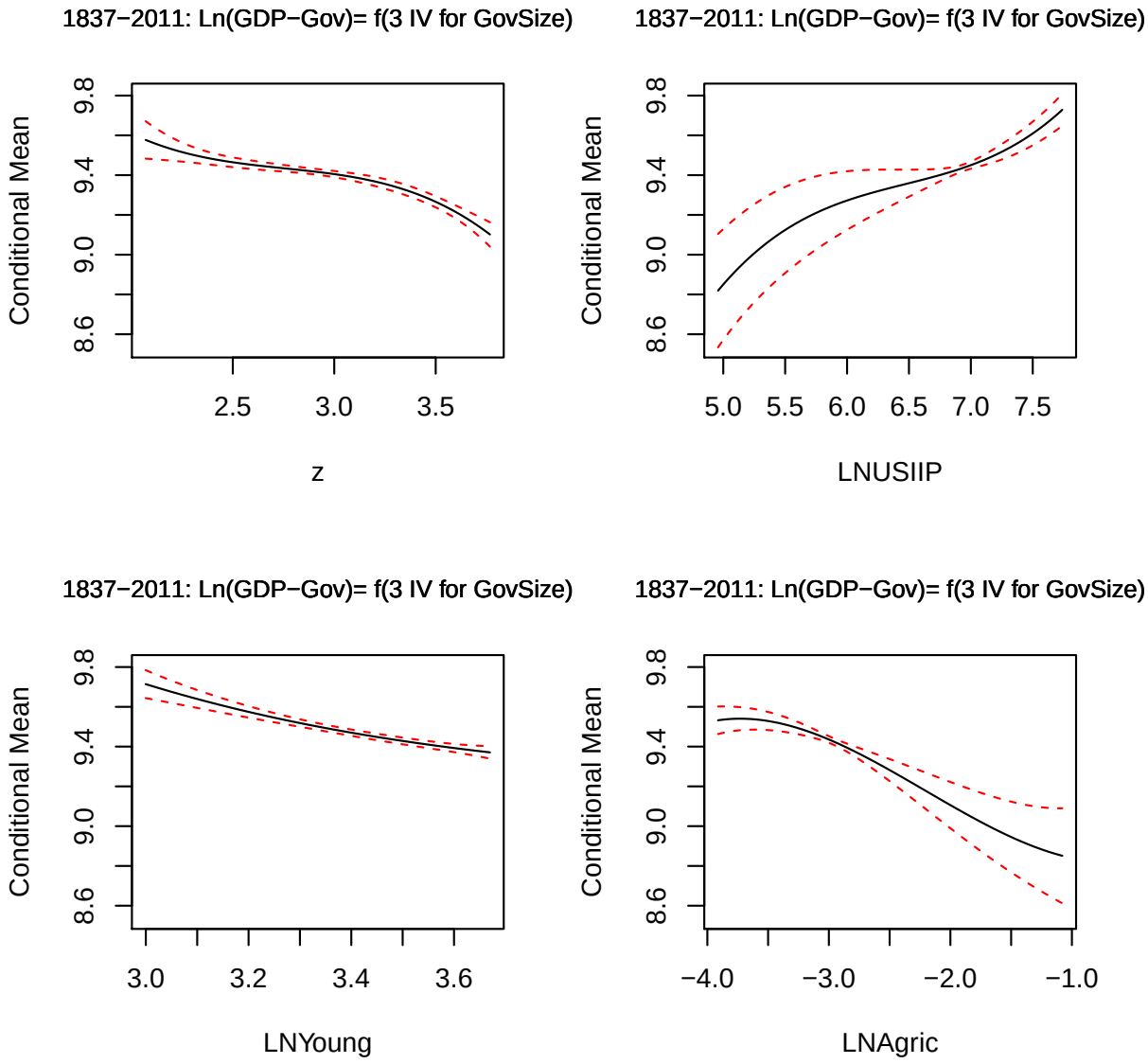

Figure 8: Derivative =f(Z=$\widehat{GSize}$), 1870-1936