Explaining the Size of the State in New Zealand, 1972-2014

Norman Gemmell, Derek Gill and Loc Nguyen

WORKING PAPER 10/2016
November 2016

Working Papers in Public Finance

VICTORIA
UNIVERSITY OF WELLINGTON
TE WHARE WĀNANGA
O TE ŪPOKO O TE IKA A MĀUI

Chair in Public Finance
Victoria Business School

The Working Papers in Public Finance series is published by the Victoria Business School to disseminate initial research on public finance topics, from economists, accountants, finance, law and tax specialists, to a wider audience. Any opinions and views expressed in these papers are those of the author(s). They should not be attributed to Victoria University of Wellington or the sponsors of the Chair in Public Finance.

Further enquiries to:
The Administrator
Chair in Public Finance
Victoria University of Wellington
PO Box 600
Wellington 6041
New Zealand

Phone: +64-4-463-9656
Email: cpf-info@vuw.ac.nz

Papers in the series can be downloaded from the following website:
http://www.victoria.ac.nz/cpf/working-papers

EXPLAINING THE SIZE OF THE STATE IN NEW ZEALAND, 1972-2014[†]

Norman Gemmell, Derek Gill and Loc Nguyen[*]

November 2016

ABSTRACT

Historical data on various measures of the economic size of the government sector in New Zealand suggest considerable short-term variability and hint at a number of possible longer-term trends. This paper follows up on that description by asking the question: how far can established models of government size help to 'explain' those changes in New Zealand since the early 1970s? Using public expenditure as our size metric, we specify three distinct econometric models each consistent with explanations offered in one of three separate strands in the international public finance, public choice and public administration literatures. We then nest those models to see whether any one model dominates or whether a more eclectic explanation finds support. Our empirical testing for the period 1972-2015 reveals that all three models offer some insight into changes in the size of government expenditure in New Zealand; indeed the best performing empirical model contains variables associated with each of the three literatures. The public choice approach seems to receive strongest support when a nested model is permitted. More generally, suitably capturing short-term dynamics turns out to be important for reliable estimation of longer-term trends in government expenditures.

[†] We are grateful to the NZIER Public Good fund for financial support of this research. Almost all of the data described in the paper are available at https://data1850.nz/. The full dataset is available from the authors in Excel spreadsheet format.

[*] Respectively *Chair in Public Finance*, Victoria Business School, Victoria University of Wellington; *Principal Economist & Head of Public Good*, New Zealand Institute of Economic Research (NZIER); and *Research Assistant*, NZIER.

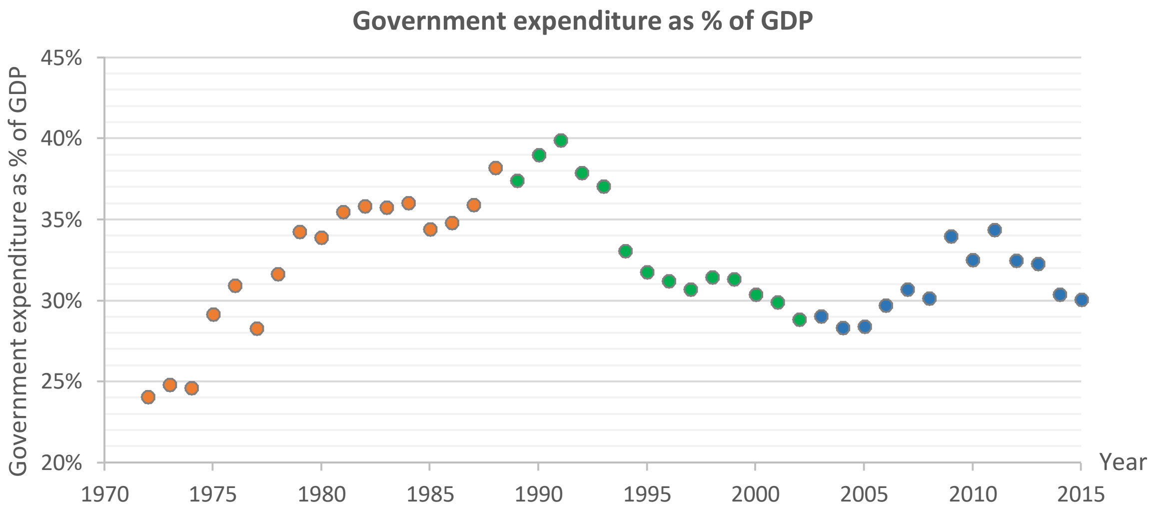
1. Introduction

In Gemmell et al. (2016) we provided some historical evidence on a variety of measures of the size of the government sector in New Zealand going back, in some cases, to the early 20th century. We found little support in the data for a general 'hollowing out' or shrinking of the state, though some changes following the 1980s reforms have persisted. While there is clear evidence that the state's role as a producer of marketed output has shrunk since the 1980s (and with that its role as employer), for a number of other measures the state's relative role has often remained broadly the same, or recovered towards earlier levels. Many measures indicated a high degree of short-term volatility or cyclicality in government size and hinted at a number of possible longer-term trends. The present paper follows up on that description by asking the question: how far can some traditional models of government size help to account for those observed time-series patterns in the size of government in New Zealand?

Focusing on public expenditure as our size metric, we specify three distinct econometric models each consistent with a separate strand in the public finance, public choice and public administration literatures. We then nest these models, effectively pitting them against each other to see whether any one model dominates or whether a more eclectic explanation emerges. Our empirical modelling covers the period from 1972 (when data on a wider set of relevant variables became available) to 2015, hence covering years either side of, and including, the well-known mid-'80s/early-'90s period of major fiscal and other reforms.

The government expenditure data which we seek to explain are shown in Figure 1 where, for illustration, we separate observations into three (colour-coded) periods: 1972-1988, 1989-2002, and 2003-15. These three periods broadly align with three 'public administration' phases identified by Boston and Eichbaum (2006). It can be seen that, while the precise turning points in the time-series are debatable, there appear to be at least three longer-term trends in the government expenditure to GDP ratio over the 40+ year period.

Figure 1 Government Expenditure 1972-2015



The rest of the paper is organised as follows. Section 2 sets out the three theoretical models that have figured prominently in previous empirical studies of the sources of government growth. Section 3 then discusses some methodological and data-related aspects of our modelling for New Zealand, while section 4 present the econometric results for our three models separately and in combination. Some conclusions are drawn in section 5.

2. Models of Public Sector Size

There are three distinct literatures that have previously put forward hypotheses regarding the determinants of the size of the government sector in developed democratic economies and that are potentially relevant to New Zealand. For convenience, we label these literatures as public finance, public choice, and public administration. The labels are used here more for ease of identification and comparison than for definitional purposes. Indeed, as we argue below, there can be 'observational equivalence' in the observable relationships among variables predicted by different models proposed in these different literature strands hence making some distinctions moot.

Public Finance Approaches

Standard public finance explanations of the size of government expenditures have generally treated these expenditures analogously to household expenditures on privately produced goods. Namely, public expenditures reflect household demand for publicly-provided goods taking the form of a conventional consumer demand function in which household income, relative prices and other household-specific characteristics determine demand for the good in question; see, for example, Gemmell (1990, 1993), Gemmell et al. (2004), Cullis and Jones, (2009; chapter 14). In the case of goods delivered via public expenditure (i.e. free at the point of consumption), prices are typically measured by some form of consumer 'tax price' or public expenditure deflator. Tax prices in this context represent the implicit tax burden (marginal or average) associated with public provision and are often proxied by an average or marginal tax rate. In such models the political mechanism by which consumer demand is translated into public action is usually left unspecified or is implicitly a median or 'decisive' voter choice; the latter mechanism being made more explicit in public choice approaches.

This public finance approach therefore gives rise to a fairly standard empirical consumer demand equation for publicly-provided output taking a linear or log-linear form, hence:

$$G = \alpha y^{\beta} N^{\gamma} {p_g}^{\varepsilon_g} {p_c}^{\varepsilon_c} \boldsymbol{Z}_{PF}^{\eta} \quad (1)$$

where G is real output of the publicly-provided good, y is real income per capita, N is population, p_g and p_c are respectively the prices of the publicly-provided good and a composite private good and $\mathbf{Z}_{PF}$ is a vector of household (or economy-wide) characteristics suggested by the public finance approach as likely to impact on demand for G. Sign expectations on parameters are: $\beta > 1$ (< 1) if demand for government-provided goods is, on average, income-elastic (-inelastic);[1]

[1] This public expenditure-income relationship is sometimes referred to as 'Wagner's Law', though the term has been applied to a variety of definitions (see Gemmell, 1990, Cullis and Jones, 2009, pp.436-7); hence we avoid its use here.

γ = 0 if G is a pure public good (hence additional consumers do not affect total demand), but γ > 0 if G is subject to congestion or has other private-good characteristics; while the price elasticities, ε_g and ε_c, are respectively negative and positive (private goods are necessarily net substitutes on average for government-provided goods in this two-good context).

Given difficulties accurately measuring implicit government output prices, p_g, and since data are more reliably available for public expenditures, E (= p_gG), rather than real output, G, equation (1) is typically translated into a public expenditure equation such that the expected sign on the parameter on p_g is positive if $\varepsilon_g > -1$. Thus:

$$E = \alpha y^{\beta} N^{\gamma} p_g{}^{(\varepsilon_g+1)} p_c{}^{\varepsilon_c} \boldsymbol{Z}_{PF}^{\eta} \tag{1'}$$

and the p_g term on the right-hand-side of (1') can be captured by available proxies, and/or re-written as a relative price, p_g/p_c.

A further strand of the public finance literature suggests that long-term trends in the relative price of public provision might be expected – 'Baumol's cost disease', named after Baumol and Bowen (1966) and Baumol (1967). These arguments are based on exogenous innovation and productivity trend differences across public/private sectors which affect relative labour costs of publicly-provided goods over time. This feeds through to relative output prices in a competitive economy, and might therefore be captured by the prices terms in (1').

The above discussion, and equation (1), relate to real government *output*, G. But transfers such as social welfare payments and unemployment benefit are also important components of total public expenditure. In the context of models of public expenditure we can think of those arising from consumer preferences for 'services' such as a social safety net, or for redistribution, giving rise to a similar equation to (1') for transfer expenditures. Over the long-run, as with publicly-provided goods, these demands for transfers might be transmitted via the electoral cycle with governments responding with provision. In practice, the institutional structure of many transfers takes the form of 'contingent benefits'; that is, their receipt is dependent on specific eligibility criteria, such as being on low income or unemployed. As a result, over the shorter term, expenditure on transfers can be expected to vary as economic conditions affect the numbers of those eligible; as reflected for example in the impact of economic recessions or shocks on unemployment and social welfare payments.

The discussion above suggests that, for the case of transfer expenditures, income and a tax price measure would continue to be relevant on the right-hand-side of (1'). Population effects on transfers, as captured by γ, are also relevant, reflecting the extent of sharing economies. For example, a value of $\gamma < 1$ indicates that a higher population is associated with a less than proportionate increase in transfers. Especially over the short-term, the impact of contingent benefits on expenditures might be captured by proxies for short-term economic shocks such as unemployment rates or benefit recipient numbers.[2]

[2] These effects might be adequately captured by short-term movements in GDP but since shocks to GDP are often associated with disproportionately large movements in social transfers and unemployment, some additional 'shock' variables are likely to be relevant to explanations of changes in public transfer expenditures.

Equation (1′) therefore provides a possible testing equation for a public finance based model applied to New Zealand data, if suitable price proxies can be obtained and potential characteristics for inclusion in $\mathbf{Z}_{PF}$ can be identified. In practice, a further adjustment to equation (1′) employed in most empirical approaches is to convert expenditures, E, in (1′) into relative terms, such as by expressing as a ratio of GDP, Y (= yN).[3] One implication of this is that interpretation of the parameters on y and N on the right-hand-side of (1′) now differ (they become β-1 and γ-1) and can be positive or negative.

Public Choice Approaches

The pubic choice approach to public expenditure modelling encompasses a huge range of political economy related factors based on a variety of views of the workings of political systems, public bureaucracies, interest groups etc. These are discussed at length by, for example, Mueller (2003) and discussing them in detail is beyond the scope of the present paper. Nevertheless this diverse literature has produced a number of hypotheses and insights that in principle we might examine with New Zealand public expenditure data.

Before turning to these hypotheses, it is worth noting that Mueller (2003; pp. 501-6), Tanzi and Schuknecht (2000) and others have pointed to a number of stylised facts relating to the size of the public sector across a range of (mainly OECD) countries. In particular, they note (1) in almost all OECD countries, the state grew in relative size over the 20th century; (2) this growth was most rapid in the period after 1960; but (3) it had largely stopped (reaching a plateau or reversal) by the early 1980s. Our econometric modelling in section 5 allows us to examine more formally how far those patterns, particularly (3), are supported for New Zealand after 1972.

Longer-term data on total expenditure and tax revenue, in Gemmell et al (2016) do suggest some support for (1) and (2) above. The New Zealand government sector clearly grew relative to GDP over the 20th Century (see Gemmell et al., 2016), with that growth being most rapid from the end of WW2 in 1946 (rather than 1960) when the first Labour Administration was in its 4th term after a decade in office. These time-series for expenditure and tax are also suggestive of a flattening or reversal in the post-WW2 trend growth of the public sector but somewhat later in New Zealand, in the late 1980s or early 1990s, rather than the early 1980s.

On public choice based hypotheses for public sector growth, Mueller (2003) outlines various possible explanations which may be separated into ‘demand-side’ and ‘supply-side’ based. Demand-side explanations conceptualise the state as constrained by or the responding to the preferences of voters. This so-called Chicago style Public Choice[4] places citizen-voters above the state such that the government uses public expenditures and taxation to deliver on voters’ demands expressed through the democratic electoral process.

[3] Strictly, if E is defined in *nominal* terms, then the appropriate denominator is PY, where P is the GDP deflator. Real GDP, Y, is used as a denominator when public expenditure is defined in *real* (or ‘volume’) terms. While this ‘volume’ designation is appropriate when considering government *output*, it has no counterpart for government *transfers*; hence a nominal ratio of P_gG/PY is commonly used to ‘deflate’ public expenditure.

[4] See Becker (1983) for an analysis of political markets with pressure groups competing for political influence.

The motivation for those demands include the provision of public goods, elimination of externalities and a desire for redistribution of income or wealth.[5] In such models the key citizen-voters' preferences to whom the government responds, may be median voters (in multi-issue elections or electorates there will generally be more than one) or particular 'special interest groups' of voters whose issue-specific preferences are pivotal for electoral outcomes.

These demand-side approaches have been used, for example, to specify empirical tests of median voter models that look similar to equation (1') above, but where relevant income, and price terms, and voter characteristics, refer specifically to the hypothesised decisive voter. Tests for the public expenditure impact of voter preferences for redistribution in particular have relied, for example, on changes in the voting franchise towards a lower income median voter, as a source of identification (Meltzer and Richard, 1983). Alternatively differences in public expenditure outcomes between governments formed by political parties with left-of-centre versus right-of-centre platforms have been used as proxies for voter preferences favouring more or less redistribution.

Turning to 'supply-side' (so-called Virginia style Public Choice) explanations, these implicitly place the state above citizens, of which three are especially prominent. Firstly recognition of principal-agent problems between politicians and bureaucrats in public provision underlies Niskanen's (1968, 1971) analysis of bureaucracy which leads to arguments that the size of the state is influenced by the preferences of bureaucrats, their forms of remuneration etc. The source of increasing government expenditure in this approach relies on bureaucratic utility that is positively related to the size of their budgets and/or subordinate numbers.

Secondly, a separate supply-side strand has stressed the potential for voters to suffer from fiscal illusion due to less than fully transparent political processes for determining public expenditure and other fiscal outcomes. In the US, for example, it has been argued that indirect (sales) taxes are more transparent to taxpayers than income taxes because of institutional conventions. This arises because sales taxes are formally (and visibly, it is argued) added at point-of-sale transactions while pay-as-you-earn income taxes are not as visible to income taxpayers. Different conventions in some European countries have led to the opposite argument – that income taxes are more visible.[6]

Whichever tax type is more/less transparent in different institutional settings is less relevant here than the general argument that if voters under-estimate their tax liabilities with respect to a particular tax or taxes in general, they are less likely to react adversely to a tax-funded increase in public expenditures. Likewise, if they suffer from 'debt illusion' (unaware of the true personal cost of debt-funded expenditure increases), they are less likely to vote against governments that propose activities involving debt-financed expenditure increases.

These tax illusion arguments have been used to explain differences in stated preferences for public spending across voters; see Ballard and Gupta (2016) for a recent example. They have

[5] See, for example, Meltzer and Richard (1981) for a specific conceptual pubic choice model in which redistribution is the sole motivation for public spending. Herwartz and Theilen (2016) provide some recent OECD evidence.
[6] See Chetty et al. (2009) for an empirical challenge to this US sales tax salience argument, and Gemmell et al. (1999) for UK evidence on direct versus indirect tax illusion.

also been used to ‘explain’ increases in public expenditure over time, especially in the US. However, to be valid such an approach also requires institutional explanations for, and evidence of, *changes* in fiscal illusion if they are to be capable of explaining observed longer-term increases *and* decreases in public expenditure as well as short-term fluctuations.

Thirdly, a related supply-side argument concerns the role of the revenue elasticity of a tax (or ‘tax elasticity’). The tax elasticity is the proportionate change in tax revenue in association with a given proportionate increase in income.[7] In some public choice models of public expenditure this is a form of fiscal illusion. In particular, it is argued that where the tax elasticity exceeds one, this generates a more than proportionate increase in revenues as nominal incomes rise due to ‘automatic’ effects even when there are no discretionary changes in tax settings. To the extent that discretionary tax changes are more visible (and hence salient) to voter-taxpayers than those associated with real income growth or price inflation, the tax elasticity may provide a proxy for the extent of fiscal illusion associated with that tax.

Not surprisingly, this argument is especially applicable to personal income taxes where the typically progressive tax rate structure generates an elasticity greater than one. In New Zealand, for example, this is estimated to be around 1.3. While similar arguments apply in principal to other tax types, these are often argued to have revenue elasticities close to one, at least over the longer-term.[8]

A further public choice argument for a role for the tax elasticity in public expenditure trends is that expenditure changes are politically easier to propose and implement when the required revenue is available ‘automatically’ than when it must be raised via discretionary increases in tax rates etc., or when new debt-funding is required. This can, but need not, represent a fiscal illusion effect. That is, fully informed voter-taxpayers may willingly sanction increased public spending when it arises from revenues generated by growing incomes, but not when it requires an increase in taxes paid for given income levels.[9]

This argument suggests that measures of the revenue elasticity of income taxes could form a useful element of an empirical public choice model of public expenditure, though they would not distinguish fiscal illusion from non-illusion aspects. However, annual values of revenue elasticities are not generally available. Fortunately, as shown by Creedy and Gemmell (2004), the tax elasticity is directly related to the structure of marginal income tax rates and brackets, data on which are readily available for most income tax systems. For New Zealand annual

---

[7] The elasticity is sometimes defined with respect to (w.r.t.) an increase in the *tax base*, rather than income. To assess the response of revenue w.r.t. income then requires values for an additional elasticity: for the tax base w.r.t. income.

[8] Corporate tax revenues are known to be relatively volatile over the short-term, in part reflecting the volatile nature of corporate profits, though some corporate tax regimes may also tend to embody relatively elastic structures; see Creedy and Gemmell (2010).

[9] In the UK, for example, in the run-up to the 2010 election the Cameron-led Conservatives (subsequently elected to form the 2010-15 coalition government) explicitly made the case for public expenditure increases and tax funding under the slogan “sharing the proceeds of growth”, with the term ‘sharing’ referring to increases in gross incomes shared between increases in tax revenues paid (implicitly to fund public expenditure, and without tax *rate* increases) and increases in private disposable incomes. See Dorey (2009) for discussion.

estimates of a suitable income-weighted average of marginal income tax rates were constructed by Bandyopadhyay et al. (2011) for the period 1907-2009, and extended to 2015 in Gemmell et al. (2016).

Considering these various public choice hypotheses above suggests a possible, if somewhat eclectic, empirical model of the following form.

$$E = \alpha y_m^{\beta 1} \mathrm{Int}^{\beta 2} \mathrm{Redist}^{\beta 3} \mathrm{Party}^{\beta 4} \mathrm{AMTR}^{\beta 5} Z_{PC}^{\eta} \quad (2)$$

where y_m represents the income of the decisive (e.g. median) voter; *Int* measures the influence/size of specific interest groups that aim to impact on public expenditure levels or choices (e.g. lobby groups for welfare spending or environmental protection); and *Redist* measures the extent of redistributional preferences within the voting population. This might be partly captured by the variable, *Party*, which could take the simple form of a dummy variable for left/right-of-centre parties in government, or a more detailed party measure for period-specific governments. This latter case would identify how far observed government expenditure levels are specific to particular governments, rather than a broader left/right redistributional 'leaning'.

The variable, AMTR, captures the tax elasticity hypothesis as described above. The vector $\mathbf{Z}_{PC}$ captures other taxpayer, democratic or bureaucratic system characteristics associated with the public choice approach (such as voting system properties; e.g. 'first past the post (FPP)' and 'mixed member proportional (MMP)' systems).

Equation (2) does not explicitly include separate fiscal illusion variables/proxies though, as noted, the AMTR variable may capture some such effects. This might be tested by decomposing the AMTR into its discretionary and automatic components. *If* the latter are more subject to fiscal illusion than the former, a larger response of government spending to the automatic component would be *consistent with* a fiscal illusion argument. It would however also be consistent with fully informed taxpayers displaying a greater willingness for automatic revenue increases to fund increases public spending, as argued above (as opposed to preferring compensatory reductions in tax rates). Given these and other difficulties obtaining reliable, time-varying measures of fiscal illusion, we are there therefore sceptical of the ability of such empirical models to reliably identify such effects on spending.

Finally, while Niskanen's modelling of 'bureaucratic motivation' has proved helpful at the micro/management level in understanding possible incentives for public servants to influence the level and type of public spending, finding suitable proxies capable of testing robustly for those effects on spending at an *aggregate* level and based on time-series data, has proved problematic for previous investigations.[10] In the context of our dataset, the best we can hope to achieve is to identify how far public expenditure levels or changes are composed of spending

[10] See Cullis and Jones (2009, pp. 432-3) for a critique. A key problem with time-series based investigations is the difficulty of identification without specific institutional reforms to enable 'pre- versus post-' evaluation. Inman (1987) suggests that increases over time in California's use of state referendums to limit local property taxes may reflect attempts to limit bureaucrats' powers. Empirical identification relied on surveys of voters; see Inman (1987, p.744).

related primarily to bureaucratic activity which may be proxied, for example, by the extent of provision of government administration services within total spending.

A similar argument can be made for redistribution-motivated spending as proposed by some of the public choice models outlined above. While an *ex ante* redistribution motive is hard to identify other than through survey-related methods, our data can reveal how far, *ex post*, public spending is composed of spending with a large redistributive dimension, such as social welfare spending and (arguably) health and education spending. As with bureaucracy, this cannot identify exogenous influences of redistributive motives on total spending but at least indicates how far outcomes are characterised by redistribution, bureaucratic services, etc. We explore these issues further in section 3.

In summary, although rigorously testing public choice arguments related to government size has proved difficult in the existing international literature, it nevertheless provides a useful source of broad propositions related to the political economy of the size of the state which can be explored empirically alongside alternative explanations to see if they find any support in the New Zealand case.

Public Administration Approaches

As we have seen, the public finance and public choice literatures tend to reference each other and have various elements in common. The public administration literature, on the other hand, is essentially separate and tends not to reference other approaches. It is also more descriptive of institutional arrangements rather than adopting a methodology of establishing hypotheses and testing those against available data using specific empirical/econometric methods. This makes it more difficult to compare with the other approaches and incorporate into econometric models. The literature on Public Administration and Public Management address questions of how the state organises itself and how the shape of the state might be changing but is less focused on explaining changes in the size of the state *per se*.

The literature emphasises a change on the shape of the state in a number of so-called 'Scando-Saxon' (Scandinavian and Anglo-saxon) countries of the OECD associated with the introduction of New Public Management (NPM), a practitioner driven eclectic movement drawing on managerialism and new institutional economics. New Zealand is regarded as the pioneer and 'poster child' of NPM and is commonly portrayed as going through at least two stages:

1. Old Public Administration – with origins in the progressive movement leading to the introduction of the Public Service Act in 1912
2. New Public Management – with its origins in the wide ranging NZ programme of reforms of the late 1980s reflected in the enactment of the State Owned Enterprises act 1986, the State Sector Act 1988, the Public Finance Act 1989 and subsequently the Public Audit Act 2002 and the Crown Entities Act 2004.

A number of authors point to a third post NPM stage although accounts differ about the timing and extent of the change. For example, Halligan (2007, p.221) suggests:

"*New Zealand has also experienced three generations of change. The 1990s has been recognized as being a second generation (Scott, 1997), with distinctive reform themes even if a preoccupation was dealing with consequences of the original model in a changing environment. However, despite this substantial interregnum, it is preferable to adhere to two models – the initial model (late 1980s–early 1990s) and the emergent revision of the original in the 2000s (Boston and Eichbaum, 2006; Gregory, 2006)".*

Thus Halligan's proposition is that there are 3 models and three waves of change in New Zealand:

1. Old Public Administration (OPA) – prior to the late 1980s;
2. New Public Management (NPM) – 1988/89 – 2002 – with 2 waves within the model – wave 1 – implementation and wave 2 mid 1990's adaptations such as the addition of a strategic management system;
3. Post New Public Management (PNPM) – from 2003 onwards.

The empirical base used to buttress claims for these phases in the public management literature is patchy and narrative accounts are often treated as evidence. Lodge and Gill (2011) look for systematic evidence of a clear move from NPM to Post NPM in New Zealand but "find limited evidence of such a shift, suggesting that the wider literature needs to move to a more careful methodological treatment of empirical patterns. … [T]his article points to rather more diverse empirical and messy patterns" (pp.141-2).

The question from the public management and administration literature we explore using the government expenditure dataset is: can we find evidence of the move from OPA to NPM to PNPM in the data? The key problem for operationalising this question is, of course, identifying the essence or core of NPM from a literature in which definitions and accounts differ.

Hood (1991) suggests NPM is linked to 4 other 'megatrends' of the time, of which the first ("attempts to slow down or reverse government growth in terms of overt public spending and staffing") is in principle directly testable with our data.[11] In a similar vein, Pollitt (1995) proposes eight core elements of New Public Management, of which the first (below) is directly related to the size of government budgets. The eight elements are:

1. Costing cutting, capping budgets;
2. Introducing market and quasi market mechanism (contracting out);
3. Disaggregation on large departments into arms-length government;
4. Separation of functions;
5. Decentralisation of authority within organisations;
6. Performance management;
7. Local not national determination of pay;
8. Increased emphasis on service and quality.

[11] The other three 'megatrends' are "the shift toward privatization and quasi-privatization and away from core government institutions"; "the development of automation … in the production and distribution of public services"; and "the development of a more international agenda, increasingly focused on general issues of public management, policy design, decision styles and intergovernmental cooperation".

Pollitt's list is silent about several facets of the state's role that might be assessed for New Zealand, such as the state's role in regulation, taxation, transfers spending, privatisation, macro-economic management etc. Rather, Pollitt's NPM categories are largely about how the state organises itself (2, 3, 4) or how public agencies conduct business within their organisations (5, 6, 7, and 8). However, the emphasis on cost cutting (1) suggests that NPM can be associated with a slowdown and reversal in the growth of government, especially relative to the rest of the economy.

Following the approach to testing of the public finance and public choice models we suggest a test of the NPM model of the general form:

$$E = a\text{NPM}^{b1}\text{PNPM}^{b2}Z_{PA}^{\eta} \tag{3}$$

where the NPM and PNPM variables are respectively slope dummy variables taking values of 1 for 1989-2015 and 2003-2015, and zero otherwise. The vector, $\mathbf{Z}_{PA}$, represents other variables associated with the public administration model, or controls such as cyclical factors required to reliably identify the two hypothesised trend changes from 1989 and 2003. It necessarily includes, for example, shift dummy variables for the two periods considered. By defining the two periods in this way (rather than 1989-*2002* and 2003-15), the trend in *E* over the 2003-15 period is given by the sum of the estimated parameters on the two slope dummies, b_1 and b_2.

Finally, the public administration model clearly captures the period during and following the major government-sector and other reforms from the mid-to-late 1980s. However the major tax and spending related reforms are often attributed essentially to the 1986-1996 period.[12] In our empirical analysis we therefore also consider how far this specific '*Reform*' period provides a better/worse description of the changes in trends otherwise ascribed to the NPM model.

3. Applying Regression Methods

This section first outlines the econometric approach that we follow (3.1), then provides some descriptive statistics on the variables used in (3.2). Regression results are reported in section 4.

3.1 Econometric approach

Before testing the models outlined in section 2 on New Zealand data a number of preliminary issues must be addressed.

Firstly, we have relatively consistent annual time-series data for 1972-2015 for our dependent variable, the ratio of government expenditure to GDP (as shown in Figure 1), and most of the independent variables of interest, giving over 40 annual observations. Most of those variables, with the exception of dummy variables, are non-stationary and I(1) – see Appendix Table A1.

We therefore use an Autoregressive Distributed Lag, ARDL(p, q), model parameterised in error-correction model (ECM) form to identify short- and long-run effects and return-to-equilibrium adjustment speeds. Consider the following general ARDL(p, q) specification for y_t:

[12] See, for example, figure 2 in Evans *et al.* (1996).

$$y_t = \sum_{j=1}^{p} \alpha_j y_{t-j} + \sum_{j=0}^{q} \beta_j X_{t-j} + \varepsilon_t \tag{4.1}$$

$$\Delta X_t = \zeta_1 \Delta X_{t-1} + \zeta_2 \Delta X_{t-2} + \dots + \zeta_s \Delta X_{t-s} + u_t \tag{4.2}$$

where the vector X_t in (4.1) includes the independent variables of interest. The α, β, and ζ are parameters to be estimated; where the ζs capture the autoregressive process in ΔX_t.

A number of parameterisations of (4.1) are possible, but it is convenient to express (4.1) in error correction (ECM) form:[13]

$$\Delta y_t = \phi(y_{t-1} - \beta X_t) + \sum_{j=1}^{p-1} \alpha_j^* \Delta y_{t-j} + \sum_{j=0}^{q-1} \beta_j^* \Delta X_{t-j} + \varepsilon_t \tag{5}$$

where $\phi = -(1 - \sum_{j=1}^{p} \alpha_j^*)$ captures the error correcting component, and $\beta = (\sum_{j=0}^{q} \beta_j^* / \phi)$ captures the long-run equilibrium relationships between *y* and *X*, with short run effects measured by β_j^* – the parameters associated with the ΔX variables in (5). The error correction term, ϕ, is a measure of the speed at which the model returns to equilibrium after a shock.

In a series of papers in the mid-1990s, Pesaran and associates (see, e.g. Pesaran and Smith, 1995; Pesaran, 1997; Pesaran and Shin, 1999) demonstrated that, under a number of reasonable conditions, estimates of the long-run parameter vector obtained from an ARDL representation of OLS regressions are consistent. Indeed they demonstrate that, where variables are I(1), parameter estimates are 'super-consistent'.[14]

Further, simulation results in Pesaran et al. (1999) demonstrate that even in small samples, standard *t*- and *F*-tests on long-run parameters from the ECM are valid, given suitable specification of the lag structures of dependent and independent variables.[15] In this case, right-hand-side variables can be regarded as 'weakly exogenous', allowing relationships with the dependent variable to be interpreted as causal.[16]

Following Pesaran and Shin (1999) we initially run ARDL(2, 2) or ARDL(≤2, ≤2) models, in the latter case selecting preferred lag structures based on the Schwarz Information Criterion (SIC). In most cases we find that one lag is sufficient and hence to maximise degrees of freedom with our relatively limited time-series, we often report results below for an ARDL(1, 1) or ARDL(≤1, ≤1) model. However, in all cases where specifications involving a second lag are preferred, these are reported.

A further issue concerns the functional form of the relationships tested. We have explored both linear and log-linear forms of regressions on equation (1'), (2) and (3). In general linear

[13] See Gemmell, Kneller and Sanz (2016) for further discussion.

[14] In addition to the presence of I(1) variables, independent variables should not be co-integrated among themselves but be co-integrated with the dependent variable.

[15] Potential endogeneity is dealt with by appropriate augmentation of the lag structure of the ARDL(p, q) model to an ARDL (p, m) model where m ≥ q; see Pesaran and Shin (1999).

[16] See, for example, Pesaran (1997, pp. 182–185), Pesaran and Shin (1999, pp. 381–387; 404–405). Pesaran and Shin further show that where serial correlation is a concern, 'appropriate modification of the orders of the ARDL model' (p.386) is sufficient to deal with both the serial correlation in the error process and regressor endogeneity.

regressions either marginally or substantively out-performed the log-linear form; the former are therefore typically reported below.

To examine empirically the role of a number of variables suggested by theory requires proxies. Those available represent their conceptual equivalents with varying degrees of accuracy and, as noted above, a given proxy may capture elements from more than one theoretical framework. As a result they should be interpreted with caution. The variables used in our analysis are listed below.

Variable name	Abbrev.	Source & Comments
Gov. expenditure (as % GDP)	E (E/Y)	www.data1850.com
Population	Pop	www.data1850.com
GDP & GDP per capita	GDP, GDPpc	www.data1850.com
Unemployment rate	U	www.data1850.com ?
Relative gov./pte. price	Pg/Pc	StatsNZ. Pg/Pc based on industry deflators for government (= admin, educ., health) and pte. industry (= others), 1978-2011. Based on gov./pte. *consumption* deflator for 2011-15.[17]
Redistributional expenditure	Redist	Gov. exp. on social welfare incl. NZ Super.
Bureaucratic output	Bureau	Government non-marketed GDP (% total market sector GDP)
National party dummy variable	Na	Dummy represents 'right-of-centre' (National Party-led) governments
Public admin period dummies	NPM, PNPM	NPM = 1 (1990-2001) PNPM = 1 (2002-2015); zero otherwise
GFC dummy; 'Reform' dummy	GFC; Reform	'GFC period' years (=1, 2008-2012); 'Reform' period (=1, 1986-94), zero otherwise
Ave. marginal income tax rate	AMTR	Income-weighted average of income tax MTRs
AMTR components	AMTR(dt), AMTR(dw)	AMTR components (indices) for tax rate change (dt), and income weight change (dw)
Canterbury earthquake dummy	Eq	Dummy =1 for 2011-2015; zero otherwise

For example, the public finance (PF) approach suggests that aggregate or average income levels may be related to demands for public expenditure, while the public choice (PC) approach specifies a decisive voter who may, or may not, be the median income earner. In either case distinguishing between effects due to mean or median income in such empirical exercises is generally not possible. Hence we use total income and income per capita as proxies for income-related demand effects but this prevents distinct interpretations in terms of the underlying (PF or PC) models proposed

A similar argument applies to unemployment. We use the rate of unemployment to capture fluctuations in the demand for welfare-related and other expenditures that may not be adequately measured by fluctuations in incomes since such expenditures generally move disproportionately with incomes. However, this may reflect a conventional demand effect that arises from associated changes in the *distribution* of income when levels change in the short-run, rather than the income level itself, or capture the effects of voter preferences for redistribution as proposed by some public choice models.

Our average marginal tax rate, AMTR, is defined as the income-weighted average of marginal income tax rates faced by all taxpayers in a given year. That is:

[17] The two relative price series overlap for the years 1988-2011 during which they are highly correlated: r = 0.976.

$$AMTR = \sum_{j=1}^{N} w_j m_j \qquad (6)$$

where j = 1 …N refers to income tax bracket, j, m_j is the marginal tax rate in that bracket, and w_j is the share of total income in tax bracket j. The change in AMTR can then be decomposed into:

$$d(AMTR) = w_j \sum_{j=1}^{N} dm_j + m_j \sum_{j=1}^{N} dw_j \qquad (7)$$

The first term on the right-hand-side (RHS) captures the discretionary tax change component while the second term captures the effect of changing incomes (usually increases) via fiscal drag which affects the weights, dw_j, associated with each tax rate bracket.[18]

3.2 Descriptive Statistics

As Figure 1 above showed, government expenditure has fluctuated considerably as a ratio of GDP over both the short- and longer-terms since 1972. Table 1 shows that the ratio averaged around 32% within a range of 24-40%. A similar picture emerges for tax revenues though fluctuations were around a lower mean of 29%, reflecting the years of public sector deficits especially during the 1970s, '80s and after the global financial crisis in 2008.

Table 1 Descriptive Statistics for Selected Variables

Sample: 1972-2015	Obs.	Mean	St. Dev.	Min.	Max.
Government expenditure	44	0.323	0.038	0.241	0.399
Tax revenue	44	0.292	0.027	0.238	0.349
GDP per capita ($)	44	36,518	6,370	27,573	47,834
Population (m)	44	3.642	0.484	2.887	4.533
Relative tax price (from 1978)	38	0.880	0.084	0.728	1.022
Unemployment	44	0.049	0.027	0.009	0.107
AMTR	44	0.329	0.055	0.260	0.446
Education & Heath share	44	0.329	0.044	0.250	0.396
Soc. Sec. & Welfare share	44	0.332	0.041	0.232	0.403
Bureaucracy ratio (to 2013)	42	0.115	0.012	0.101	0.147

Table 1 confirms that there is substantial variation in other variables that will serve as RHS variables in regressions below. For example, unemployment varied from 1% to 10% over the period, relative public/private sector prices varied from 0.73 to over 1.0, while the public spending shares on education/health and social security/welfare both fluctuated around means of 33% but ranged from 23% to 40%. Together with various political and other dummy variables, these data are used in the next section to apply regression analysis to the three political economy models described earlier.

[18] Changes in tax thresholds affecting tax brackets represent tax rate changes for sub-sets of taxpayers within this framework.

4. Regression Results

We begin by examining the PF model using total government expenditure as a percent of GDP, E/Y, throughout and including income per capita (GDPpc), population (Pop), relative prices (P_g/P_c), unemployment (Unemp), and a dummy variable for the effects of the Christchurch earthquakes (Eq) on public expenditures (2011-2015) on the RHS. We investigated both linear and log-linear forms and generally found the former provided a better fit to the data. We therefore report those in Table 2. Since data on P_g/P_c are only available from 1978, to maximise the sample period, we set 1972-1977 values for P_g/P_c to its 1978 value. This effectively ensures that P_g/P_c cannot play any role in explaining any changes in E/Y during those six years hence, *ceteris paribus*, tending to *understate* the importance of any P_g/P_c effect on expenditures. Table 1 also reports comparable regressions over 1978-2014 and it can be seen (regressions (3) & (4)) that in general this does not change the interpretation from the regression, with a parameter estimate for P_g/P_c that is smaller and more robustly identified in (3).

4.1 Public finance model

The results in Table 2, regression (1) are consistent with predictions from a PF model. Long-run parameters suggest that increases in GDP per capita are positively related in the long-run to the ratio of government expenditure to GDP but not significantly so, consistent with a unit income elasticity of demand (for *expenditures*). The short-run, or impact, effect of increases in GDP per capita appear to be to modestly reduce E/Y. Regressions (3) and (4), for the shorter period, provide some evidence of significant (hence, income elastic) effects of GDP per capita on expenditures, though this is generally overturned in later results.

For given GDP per capita etc., higher population is associated with lower E/Y, consistent with substantial sharing economies or public good characteristics. In regression (1), for example, a one unit (= 1 million in this dataset) increase in population is associated with a long-run fall in the E/Y ratio of 0.21. Evaluated at means (E/Y: 0.32; Pop: 3.64 million), these estimates imply that a 1% increase in population reduces the E/Y ratio by around 2.4%.

Bearing in mind the earlier discussion regarding expected signs on P_g/P_c when using E, rather than G, in the dependent variable, parameter estimates in Table 2 imply that, when evaluated at means, a 1% increase in relative prices is associated with a roughly 1% increase in the E/Y ratio. That is, increases in prices are fully passed through in the long-run into increases in expenditures, P_gG, with no reduction in real government output, G. This may partly reflect a tendency for public sector budgeting for annual expenditure increases to be based on nominal values.

On unemployment effects, Table 2 indicates that a long-run effect on E/Y is only obtained when allowing for up to 1 lag but not when allowing for up to 2 lags. The regression diagnostics clearly favour the latter, suggesting that short-run fluctuations in E/Y are better captured by a longer lag structure (for E/Y) than the cyclically-sensitive unemployment variable. Short-run parameters confirm that, when 2 lags are permitted, changes in unemployment have a significant immediate positive effect on E/Y as expected.

Table 2 also provides evidence that, other things equal, the Canterbury earthquakes were associated with a short-run increase in government spending, but no (or a small) long-run impact. Of course, since our *Eq* dummy refers to the end of our time-series (2011-15), the 'long-run' in this case simply implies that, based on the final 5 years of data, there is evidence of an otherwise unexplained persistent positive effect on spending post-2010.

Table 2 Regression Results: Public Finance Model

	Reg (1)	Reg (2)	Reg (3)	Reg (4)
Sample:	1973-2015	1974-2015	1979-2015	1980-2015
(lags)	(p,q ≤ 1)	(p,q ≤ 2)	(p,q ≤ 1)	(p,q ≤ 2)
Long-run	Dependent variable: Change in Government Expenditure/GDP (Exp)			
RGDPpc	0.0084	0.0017	0.0067	0.0080
	(1.40)	(0.26)	(2.68)	(2.36)
	[0.172]	[0.798]	[0.012]	[0.027]
Pop	-0.213	-0.169	-0.206	-0.223
	(-2.78)	(-1.88)	(-6.26)	(-4.48}
	[0.009]	[0.069]	[0.000]	[0.000]
P_g/P_c	0.378	0.591	0.306	0.392
	(1.67)	(2.52)	(2.84)	(2.36)
	[0.104]	[0.017]	[0.008]	[0.027]
Unemp	0.631	-0.011	0.507	0.220
	(2.06)	(-0.03)	(3.57)	(0.89)
	[0.047]	[0.980]	[0.001]	[0.382]
Eq	-0.003	0.007	0.024	0.007
	(-0.13)	(0.33)	(2.48)	(0.44)
	[0.898]	[0.741]	[0.020]	[0.661]
Error-correction	-0.384	-0.334	-0.614	-0.437
	(-3.91)	(-3.48)	(-5.76)	(-3.49)
	[0.000]	[0.001]	[0.000]	[0.002]
Short-run				
Exp: D2		-0.329	-	-0.416
		(-2.69)		(-3.55)
RGDPpc: D1, D2	-0.011	-0.008	-.0117	-0.008 -0.005
	(-3.51)	(-2.26)	(-4.78)	(-3.17) (-1.86)
Pop	-0.218	-	-	-
	(-1.87)			
P_g/P_c	-	-	-	-
Unemp	-	0.727	-	0.555
		(2.20)		(2.49)
Eq	0.025	-	0.023	0.015
	(1.82)		(2.47)	(1.81)
Constant	0.182	0.129	0.341	0.226
	(3.35)	(2.39)	(5.37)	(2.80)
N	43	42	37	36
Adj-R^2	0.476	0.577	0.696	0.786
PSS Bounds F-test	5.86**	6.44**	11.64**	8.62**

Notes: (t-ratios) and [p-values] shown below parameter estimates. D1, D2 indicate first and second lag changes.
** Pesaran-Smith- Shin (PSS) bounds test (F-test) significant at 1% = reject H0: no long-run relationship.

Table 3A Regression Results: Public Choice Model

	Reg (5)	Reg (6)	Reg (7)	Reg (8)	Reg (9)
Sample:	1974-2015	1980-2015	1980-2015	1974-2015	1980-2015
(lags)	(p,q ≤ 2)	(p,q ≤ 2)	(p,q ≤ 2)	(p,q ≤ 2)	(p,q ≤ 1) †
Long-run	Dependent variable: Change in Government Expenditure/GDP (Exp)				
AMTR	0.797	0.728	0.721		
	(15.5)	(14.0)	(4.03)		
	[0.000]	[0.000]	[0.000]		
AMTR-dt				0.0023	0.0012
				(7.10}	(2.54}
				[0.000]	[0.017]
AMTR-dw				0.0023	0.0006
				(6.56}	(0.83}
				[0.000]	[0.415]
Na	-0.019	-0.016	-0.049	-0.018	-0.018
	(-3.78)	(-3.92)	(-2.48)	(-2.76)	(-1.60)
	[0.001]	[0.001]	[0.020]	[0.009]	[0.120]
Unemp	1.649	1.353		1.664	1.163
	(16.5)	(11.4)		(10.6)	(3.31)
	[0.000]	[0.000]		[0.000]	[0.003]
Eq	0.030	0.025	0.025	0.038	0.013
	(4.66)	(4.57)	(1.15)	(3.76)	(0.82)
	[0.000]	[0.000]	[0.262]	[0.001]	[0.419]
Error-correction	-0.929	-0.916	-0.298	-0.809	-0.437
	(-3.91)	(-6.90)	(-3.97)	(-5.92)	(-3.10)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.004]
Short-run					
Exp: D2	-	-	-0.332	-0.416	-
			(-2.44)	(-3.55)	
AMTR: D1, D2	-0.46	-0.37, -0.36	-0.24 -0.43	-0.001 (dt)	-
	(-4.3)	(-3.8) (-3.3)	(-2.0) (2.9)	(-3.46)	
				-0.002 (dw)	-
				(-1.95)	
Na: : D1, D2	-0.01 0.02	0.01	0.036	-	-
	(-1.8) (2.4)	(1.8)	(3.7)		
Unemp	-	-		-	0.821
					(3.95)
Eq	-	-	-	-	0.024
					(2.15)
Constant	-0.011	0.023	0.034	-0.229	0.064
	(-0.61)	(1.31)	(1.52)	(-3.42)	(0.75)
N	42	36	36	42	37
Adj-R^2	0.682	0.672	0.458	0.556	0.566
PSS Bounds F-test	18.0**	15.1**	8.61**	8.93**	3.96*

Notes: (t-ratios) and [p-values] are shown below parameter estimates. D1, D2 indicate first and second lag changes. ** [*] Pesaran-Smith-Shin (PSS) bounds test (F-test) significant at 1% [5%] = reject H0: no long-run relationship. † We report a (p, q ≤ 1), rather than (p, q ≤ 2), regression as the latter gave inferior results and only selected a second lag for one variable (Unemp).

Table 3B Regression Results: Public Choice Model

Sample: 1974-2015 Lags: p,q ≤ 1	Reg (10)+	Reg (11)+	Reg (12)	Reg (13)	Reg (14)+	Reg (15)
Long-run	Dependent variable: Change in Government Expenditure/GDP (Exp)					
AMTR	1.239 (7.56) [0.000]	1.098 (3.56) [0.001]	1.034 (9.89) [0.000]	1.258 (7.57) [0.000]	0.739 (3.12) [0.004]	0.981 (9.01) [0.000]
Log(GDPpc)	-	-	-	0.025 (0.95) [0.349]	0.012 (0.60 [0.555]	0.008 (0.36) [0.723]
Educ-Heath share(-1)	0.439 (3.15) [0.004]	0.318 (1.78) [0.086]	0.291 (2.75) [0.010]	0.341 (2.34) [0.026]	0.107 (0.68) [0.500]	0.221 (1.77) [0.088]
Soc-Welfare share(-1)	0.141 (1.32) [0.196]			0.138 (1.25) [0.221]		
Bureau(-1)		-0.161 (-0.25) [0.802]			0.459 (0.93) [0.359]	
Na	-0.023 (-3.43) [0.002]	-0.016 (-3.16) [0.004]	-0.016 (-3.41) [0.002]	-0.020 (-2.88) [0.007]	-0.013 (-3.08) [0.005]	-0.014 (-2.88) [0.007]
Unemp	2.404 (9.45) [0.000]	2.271 (4.95) [0.000]	2.185 (10.5) [0.000]	2.141 (8.67) [0.000]	1.659 (4.43) [0.000]	1.957 (8.60) [0.000]
Eq	0.026 (2.53) [0.017]	0.022 (2.46) [0.020]	0.020 (2.79) [0.009]	0.021 (1.91) [0.066]	0.016 (2.14) [0.041]	0.019 (2.63) [0.013]
Error-correction	-0.752 (-6.62) [0.000]	-0.826 (-4.27) [0.000]	-0.859 (-7.52) [0.000]	-0.713 (-6.65) [0.000]	-0.969 (-4.79) [0.000]	-0.804 (-7.21) [0.000]
Short-run						
AMTR	-0.50 (-5.1)	-0.53 (-4.3)	-0.51 (-4.9)	-0.44 (-4.7)	-0.40 (-3.1)	-0.47 (-4.7)
Log(GDPpc)				-0.25 (-2.7)	-0.27 (-2.2)	-0.22 (2.1)
Na:	-0.019 (-2.4)	-0.021 (-2.5)	-0.021 (2.8)	-0.026 (-3.4)	-0.025 (-2.8)	0.027 (3.5)
Educ-Heath share (EH)	-	-	-	-	-	0.20 (1.7)
Soc-Welfare share (SSW)	-0.29 (-4.0)			-0.28 (-2.8)		
N	42	41	42	42	41	42
Adj-R^2	0.751	0.704	0.712	0.791	0.730	0.748
PSS Bounds F-test	16.7**	13.9**	17.2**	14.4**	11.2**	14.3**

Notes: (t-ratios) and [p-values] are shown below parameter estimates. D1, D2 indicate first and second lag changes. ** Pesaran-Smith-Shin (PSS) bounds test (F-test) significant at 1% [5%] = reject H0: no long-run relationship. + 1974-2014 due to missing *Bureau* data.

Finally, the error correction terms suggests that adjustment to exogenous shocks (such as those associated with the RHS variables) occurs relatively quickly with around one-third to two-thirds of any disequilibrium corrected within one year.

4.2 Public choice model

Tables 3A and 3B reports results for our empirical public choice models. We begin in Table 3A (regressions (5-7)) by considering only the average marginal income tax rate (AMTR) as an exogenous measure of the revenue-elastic properties of the tax system, and two distributional proxies – a dummy variable for right-of-centre led governments (Na), and the unemployment rate. In addition, in regressions (8) and (9) we consider the AMTR decomposition into AMTR-dt and AMTR-dw.[19] We also include the Canterbury earthquake dummy, *Eq*, in all regressions since this represents a relevant exogenous shock to public spending and GDP regardless of the model being tested.

On the inclusion of the Na dummy, right-of-centre governments are also often regarded as favouring a smaller public sector, other things equal, so that a significant (negative) parameter estimate on this variable need not necessarily signal redistributive preferences. However since a large fraction of government expenditure is redistributive, any reductions in total government spending by such governments are likely to involve less redistributive spending. For example, education, health, social security & welfare represent around 66% of total government expenditure on average over 1971-2015.

Regressions (5)-(7), covering both the periods from 1973 and 1979, generally provide a good fit to the data. Adjusted-R^2 are around 0.67, and can be seen to provide strong support for a role for the AMTR in 'explaining' positive long-run expenditure growth, with a 1 percentage point increase in the AMTR associated with around a 0.7 percentage point increase in the E/Y ratio. The table also shows that the immediate (short-run) impact of an increase that the AMTR is to *reduce* expenditure by around 0.3-0.4 percentage points.

It is not clear why this would be the case, although larger increases in the AMTR (especially automatic increases) tend to be associated with faster growth in nominal GDP which, in the absence of immediate discretionary increases in government spending, would tend to reduce E/Y. Further, cyclical upturns in GDP which raise the AMTR would normally be associated with reduced social welfare-related spending, hence dampening the E/Y ratio. Results (not shown) for the AMTR remain strong when the contemporaneous/lagged values used in Table 3A are replaced by one-lag equivalents; that is using lags 1-3 instead of 0-2.

When the AMTR variable is decomposed into its two 'discretionary' (AMTR-dt) and 'automatic' (AMTR-dw) parts, regressions (8) and (9) – for the post-1972 and post-1978 periods respectively – indicate that the two effects take the expected positive signs and are statistically significant in 3 of the 4 cases. There is also no evidence here that the automatic effect is larger

[19] For the decomposition we use an Index where each of AMTR, AMTR-dt and AMTR-dw are set equal to 100 in 1962.Hence the size of parameter estimates on the two decomposition variables cannot be directly compared with those on AMTR.

than the discretionary effect, hence providing no support for a 'fiscal illusion' type argument. Indeed, for the shorter period, 1980-2015, the opposite is supported with discretionary increases (AMTR-dt) more likely to be associated with increased E/Y.

Table 3A, regression (7) shows the importance of the unemployment variable for the overall result with a much reduced adj-R^2 and more extensive lag structures on other variables now selected. It would seem therefore that, while increases in unemployment may be associated with increased redistributional expenditures (e.g. via social welfare payments), their primary role in these regressions is to accommodate the short-run fluctuations in those expenditures. Nevertheless, at least in this public choice model, the *Unemp* proxy appears to have long-run effects on E/Y implying that increases in spending induced by short-run shocks may have persistent effects on spending beyond the immediate shock.[20]

The dummy for National Party-led governments is significantly negative, suggesting that, other things equal, such governments do tend to have lower government expenditure/GDP ratios. This appears to be true in both the short- and long-run, with a consistent, somewhat larger (more negative) short-run estimated effect than for the long-run. According to this model, under National governments, government spending has been, on average, almost two percentage points of GDP lower than other (Labour-led) governments over 1972-2015. This is after taking account of cyclical and earthquake-related factors, as captured by *Unemp* and *Eq*.

Finally, the Canterbury earthquake dummy (=1, 2011-15) in Table 3A generally supports the view that this period has been associated with high public spending, other things equal. Of course the other major shock around this time was the global financial crisis (GFC), which may be partly captured by *Eq*. However, when we replace *Eq* with a GFC dummy for the years 2009-2012, while this is also positively signed it is not statistically significant. When both dummy variables are nested in the same regression, *Eq* clearly dominates (t-ratios = 4.08 for *Eq*; 1.22 for *GFC*).

Table 3B adds a wider set of proxies for redistributional aspects of public spending, allowing us to examine how far the public choice argument that redistributional dimensions dominate public spending increases. We test two spending components usually regarded as redistribution-related: social security & welfare spending (mainly transfers, including NZ Superannuation), and education & health spending. To minimise endogeneity concerns, we specify these as shares of total spending rather than as ratios to GDP, and run our ARDL model excluding contemporaneous values of the two share variables (i.e. using lags 1-2 instead of 0-1).[21]

As an, albeit limited, proxy for bureaucracy-related expenditures we include a '*Bureau'* proxy, calculated as the ratio of government non-marketed GDP to total GDP net of the government non-market component. This 'net' adjustment avoids the numerator also being

[20] Of course, to the extent that increases in unemployment in a downturn are followed by reductions in a subsequent upturn, then such persistent effects present in principle, would not be observed over complete cycles.
[21] We use a ($p, q \leq 1$) rather than ($p, q \leq 2$) here both to save on degrees of freedom with this larger variable set and because two lags rarely seem supported by AIC tests.

include in the denominator. Lastly, Table 3B regressions were run without (in 10-12), and with (in 13-15), a GDP per capita variable.[22] This serves both to ensure that any effects identified by our public choice variables in Table 3B are not due solely to the omission of GDP per capita, and because previous public choice studies have used GDP per capita as a proxy for median voter income.

The addition of the three expenditure component variables has no substantive effect on the interpretation of results for variables previously included in Table 3A, so we do not discuss these further here. Likewise, repeating regressions (10) – (12) by adding the GDP per capita variable in (13) – (15) suggests that the variable has little impact on other variables in the regressions and is statistically insignificantly different from zero. This is consistent with government expenditure having an income elasticity of approximately unity, though whether this represents the preference of a 'decisive voter' cannot be determined from these results.

On redistributional expenditures, Table 3B results suggest that increases in the expenditure share of education and health are associated with a persistent long-run (subsequent) increase in total spending of around 0.3 or 0.4 percent of GDP. Estimated effects for social security and welfare (SSW), though also positive, are much smaller and statistically less robust. This SSW result may partly arise from the fact that much SSW expenditure is composed of transfers that are cyclical in nature and hence less likely to have persistent effects on total expenditure levels. On bureaucracy effects, Table 3B results fail to identify significant associations with total spending levels; indeed when GDP per capita is excluded from the regression the estimated sign is negative.[23]

4.3 Public administration model

As noted in section 2, the multifaceted nature of the 'new public management' (NPM) reforms in New Zealand (a) make them difficult to test empirically in this context, and (b) are generally related to reforms of internal public administrative practices rather than public sector size *per se*. Nevertheless, since reducing the size, or improving the efficiency, of the government budget was one motivating factor behind the New Zealand NPM reforms, it is interesting to ask whether these reform episodes had a demonstrable effect on our government expenditure size measure.

Following the discussion in section 2, we use two dummy variables to represent the two 'phases' of public management reform: the initial NPM period of 1989-2002, and the 'post-NPM' (PMPM) period from 2003 onwards. We set the NPM dummy = 1, 1989-2015; and PMPM dummy = 1, 2003-15, with equivalent period slope dummies. Hence when both dummies are included in regressions the *ceteris paribus* NPM effect on expenditure for 1989-2002 is

[22] Log(GDPpc) was used here because regressions generally failed to converge when GDPpc was used instead.
[23] When only *Bureau* (of the three spending categories) is included in the regression, it takes a significant positive sign. However when nested with either of the other two spending categories, we cannot reject the hypothesis that its effect is zero, while this hypothesis is rejected for the other spending categories.

identified, with PNPM identifying the *additional* 2003-15 impacts.[24] We have also examined alternative start dates (up to 1 year either side of 1989 and 2002) and found that results are generally unchanged by these re-definitions.

We again examine both the shorter (from 1978 and longer (from 1972) time-series, with lags reducing the sample period by a further 1 or 2 years. In this case, due to the dominant role of dummy variables (where lagged values cannot be included due to perfect collinearity) we have to choose a lag structure rather than allow endogenous selection of lags. We examine both ARDL (2, 2, 0) and ARDL (1, 1, 0) models where the zero refers to the dummy variables. Variable addition tests are used to choose between the two ARDL structures.

Regressions are shown in Table 4. At first sight – see regressions (16) and (17) – the public administration model appears consistent with the data: the NPM period is associated with a declining expenditure/GDP ratio of almost 1 percentage point per year (e.g. -0.009 for 1979-2015), that is arrested in the PNPM period from 2003 (+0.010 for 2003-2015). Adding an unemployment variable in regression (18) to capture cyclical effects is strongly supported by the data with a regression fit now around 0.56.

However, regressions (19) – (21) allow for a longer lag structure to capture short-run adjustment processes and this appears to undermine the role of the public administration variables. Firstly, regressions (19) and (20) examine how far a model with *only* expenditure and unemployment, with 2 lags, can explain the data. This reveals that, for 1980-2015 for example, 62% of the changes in expenditure can be explained simply by allowing for the autoregressive process in *Exp* and the cyclically-related *Unemp* variable. The addition of the NPM and PNPM variables in regression (21) yield insignificant estimated coefficients on the NPM and PNPM variables, a poorer fit overall, and more marginal support for the presence of a long-run relationship (F = 4.4).

Of course, the fact that period *t* changes in expenditure levels (as a % of GDP) are closely related to past changes (in *t-1* and *t-2*) begs the question of what determines those past expenditure changes. They could, for example, be influenced by the public administration reform characteristics described earlier, or indeed public finance and/or public choice variables. However, whereas the previously examined public finance/choice regression models demonstrated that *adding* relevant variables to the autoregressive process improved model performance, this appears not to be the case for the public administration model in Table 4. And, if the new NPM regime introduced around 1989 had any longer-term effects this should in principle be discernible separately from short-run adjustment effects.

[24] We also examine a 'fiscal reform' dummy (1986-1996) but find it performs no better than the NPM/PNPM variables. Indeed, replacing the NMP and PNPM variables in regression (21) yields a significantly *positive* parameter estimate (0.045; t = 3.92) on the reform dummy. Adding a slope dummy for the reform period the NPM/PNPM variables continue to perform better.

Table 4 Regression Results: Public Administration Model

Sample:	1974-2015	1979-2015	1979-2015	1974-2015	1980-2015	1974-2015
	Reg (16)	Reg (17)	Reg (18)	Reg (19)	Reg (20)	Reg (21)
Lags:	p,q = 1, 0	p,q = 1, 0	p,q = 1, 1, 0	p,q = 2, 2	p,q = 2, 2	p,q = 2, 2, 0
Long-run	Dependent variable: Change in Government Expenditure/GDP (Exp)					
NPM (slope)	-0.009	-0.009	-0.005	-	-	-0.003
	(-1.67)	(-4.20)	(-3.83)			(-0.89)
	[0.103]	[0.000]	[0.001]			[0.379]
PNPM (slope)	0.008	0.010	0.005	-	-	0.001
	(0.88)	(2.17)	(2.57)			(0.19)
	[0.386]	[0.003]	[0.015]			[0.852]
Unemp			0.618	-1.35	-1.55	0.218
			(2.44)	(-1.33)	(-1.54)	(0.27)
			[0.021]	[0.193]	[0.133]	[0.786]
Error-correction	-0.185	-0.435	-0.618	-0.171	-0.143	-0.294
	(-2.21)	(-2.77)	(-3.77)	(-2.35)	(-2.40)	(-2.64)
	[0.033]	[0.009]	[0.001]	[0.024]	[0.023]	[0.013]
Short-run						
Exp D2	-	-	-	-0.40	-0.55	-0.336
				(-2.86)	(-4.11)	(-2.2)
Unemp D1, D2			0.76	1.31 0.36	1.26 0.51	1.17 0.24
			(3.83)	(5.3) (21.3)	(6.8) (2.2)	(4.3) (0.78)
N	43	37	37	42	36	42
Adj-R^2	0.164	0.219	0.557	0.509	0.624	0.491
PSS Bounds F-test	2.65	3.01	4.33*	10.2**	12.3**	4.40*

Notes: (t-ratios) and [p-values] are shown below parameter estimates. D1, D2 indicate first and second lag changes. **[*] Pesaran-Smith-Shin (PSS) bounds test (F-test) significant at 1% [5%] = reject H0: no long-run relationship.

It is tempting to conclude that the (possibly substantive) reforms to New Zealand's public sector management over this period, have not impacted on the trend in the public expenditure to GDP ratio, when due allowance is made for short-term fluctuations, despite the superficial evidence in Figure 1. This is not to deny the possibility of major changes in New Zealand's public sector management system, nor that this had discernible effects on management performance. But, if so, these are not strongly evident in the outcome for one of the underlying motivations for reform; namely control of, or reductions in, public expenditures.[25]

4.4 Combining empirical models

As noted earlier, our various motivating models to explain levels changes in public expenditure, involved several empirical proxy variables in common as well as a number of

[25] We also examined regressions such as (21) but in which our '*Bureau'* variable formed the dependent variable — since public management reforms may have been reflected only in the size of public *non-market* sector GDP. However the NPM and PNPM variables also performed poorly, failing statistical significance tests at the 10% level or higher.

distinct variables. Results in Table 2 (for a public finance model) and Tables 3A and 3B (for public choice models) each provide evidence consistent with an explanatory role for their hypothesised relationships, at least in a statistical sense. This raises the question of whether one model is unambiguously preferred to others, or whether a more eclectic model drawing on element of public finance/choice/administration provides a better 'fit'.

With sufficient data and degrees of freedom an ideal approach would be to nest all three models in a single equation, thereafter following a 'general-to-specific' approach to identify the 'best' model specification.[26] However since we have a limited time-series of up to 43 annual observations, together with around 15 independent variables (and their lagged values), there are insufficient degrees of freedom to pursue such an approach reliably here.

Instead, given the limited support (so far) for the public administration model, our approach is first to consider the public finance model in Table 2 and add to it variables from the public choice model in Table 3B. To check robustness, we then repeat this process but beginning with the public choice model. Finally we consider whether adding public administration variables to the resulting model improves its performance.

To save space below we do not report all of the general-to-specific pathways that we follow to identify the best fitting models. Table 5 records some of the key results obtained from the variable addition/deletion procedure. We begin with the ARDL (≤1, ≤1) model in Table 2, regression (1), repeated as regression (5.1) in Table 5.[27] Subsequently, with a suitably reduced model, we consider whether the introduction of two-period lags is warranted; see regressions (5.7) and (5.9).

With the error-correction parameterisation in Table 5, simple variable addition tests based on conventional *F*-tests (where $F = t^2$ for a single added variables) on the long-run or short-run parameters of each added variable are not appropriate. This is because variable addition/deletion may change the lag structures more generally selected by the AIC. Hence we use the regression *F*-statistic, or equivalently, changes in the adjusted-R^2 to compare nested and non-nested models in Table 5.

Regression (5.2) shows that adding the AMTR variable from the public choice model in (5.1) is strongly supported: the adjusted-R^2 increases by around 0.2 and all included variables (with the exception of P_g/P_c) have statistically significant effects on expenditures in the long-run and/or short-run.

[26] See, for example, Campos et al. (2005). In their Introduction, they note one useful approach as: "1. Ascertain that the general statistical model is congruent. 2. Eliminate a variable (or variables) that satisfies the selection (i.e., simplification) criteria. 3. Check that the simplified model remains congruent. 4. Continue steps 2 and 3 until none of the remaining variables can be eliminated.

[27] Results in previous tables suggest that one lag or less is most commonly selected by the AIC. This allows us to save on degrees of freedom. A similar process to that shown in Table 5 but using a p, q = 1 model yields similar results.

Table 5 Regression Results: Public Finance Model

Sample: 1973-2015	Reg (5.1)	Reg (5.2)	Reg (5.3)	Reg (5.4)	Reg (5.5)	Reg (5.6)	Reg (5.7)	Reg (5.8)	Reg (5.9)
Lags:	(p,q ≤ 1)	(p,q ≤ 1)	(p,q ≤ 1)	(p,q ≤ 1)	(p,q ≤ 1)	(p,q = 2)	(p,q ≤ 2)	(p,q = 1, 1, 0)$	(p,q ≤ 2)
Long-run	Dependent variable: Change in Government Expenditure/GDP (Exp)								
RGDPpc	0.0084	0.0062+	-0.0014	-0.001	0.0042+	-0.0072	-	-	
	(1.40)	(2.60)	(-0.38)	(-0.29)	(0.85)	(-0.46)			
Pop	-0.213	-0.105+	0.024+	0.015+	-0.055	0.114	-	-	
	(-2.78)	(-3.47)	(0.44)	(0.03)	(-0.73)	(0.48)			
P_g/P_c	0.378	0.130	0.057	-	-	-	-	-	
	(1.67)	(1.35)	(0.65)						
Unemp	0.631	1.64	1.718	2.062	2.192	2.128	2.293	2.341	1.697
	(2.06)	(10.0)	(13.1)	(9,94)	(7.92)	(4.60)	(10.95)	(8.60)	(8.96)
Eq	-0.003	0.009+	0.021	0.020	0.025	0.039	0.016	- †	0.014
	(-0.13)	(1.07)	(2.81)	(2.84)	(2.17)	(2.12)	(2.00)		(1.72)
AMTR (‡ = -1)		0.628	0.848+	0.945+	1.151+	1.348+	1.081+	1.050+	‡ 0.994+
		(6.68)	(8.65)	(8.86)	(7.12)	(3.71)	(7.96)	(7.31)	(6.49)
Na			-0.017	-0.018+	-0.015+	-0.046+	-0.022+	-0.026+	-0.022
			(-1.78)	(-2.05)	(-1.44)	(-1.38)	(-3.88)	(-3.39)	(-3.53)
Educ-Health share (-1)				0.241	0.392	0.434+	0.435+	0.624+	0.353
				(2.08)	(2.34)	(1.86)	(3.66)	(3.05)	(2.52)
Soc. Sec. & Welfare					0.232+	0.305	0.154+	0.343	0.077+
share (-1)					(1.62)	(1.20)	(0.51)	(1.64)	(0.75)
NPM (slope)								-0.003	
								(-1.97)	
PNPM (slope)								0.007	
								(2.55)	
Error-correction	-0.384	-0.731	-0.901	-0.916	-0.679	-0.774	-0.818	-0.799	-0.784++
	(-3.91)	(-7.03)	(-7.84)	(-8.08)	(-6.14)	(-2.12)	(-8.44)	(-7.76)	(-6.62)
N	43	43	43	43	43	41	41	41	41
Adjusted-R^2	0.476	0.689	0.682	0.738	0.793	0.818	0.839	0.836	0.772
PSS Bounds F-test	5.86**	11.84**	11.24**	13.74**	12.77**	3.63*	26.45**	18.42**	18.41**

Notes: (t-ratios) are shown below parameter estimates. Short-run parameters are omitted to save space. ** Pesaran-Smith- Shin (PSS) bounds test (F-test) significant at 1% = reject H0: no long-run relationship. + Statistically significant short-run parameter. † When *Eq* is included in this regression it is insignificant (*t*-ratio = 0.59), the adj-R^2 is lower and parameter standard errors for NPM and PNPM are larger than in (5.8). ++ Second lag of Exp. is also significant (at 10.5%) in this regression. $ The lag structure here allows for a second lag on Educ-Health expenditure share, supported by a *t*-test.

Regression (5.3) reveals that adding a right-of-centre government dummy (Na) weakens the role of P_g/P_c further but makes little difference to the overall fit of the regression. Adding further public choice variables in the form of the two 'redistributive spending' variables (Education-Health share and the Social Security and Welfare share) can also be seen to improve the model fit in (5.4) and (5.5) and appear to be supported. (P_g/P_c is omitted from the regressions shown; it was associated with poorer diagnostics when included).

Table 5 also suggests that when those public choice variables are added, the previously observed statistically significant effects associated with GDP per capita and/or population disappear. It should not necessarily be inferred that these variables are unimportant for spending. Rather it may be indicative that expenditure and GDP are similarly affected (via a unit elastic relationship), while population increases have few if any congestion impacts on public expenditures. Nevertheless, for model selection purposes, these variables may be regarded as irrelevant; when they are omitted from subsequent regressions (compare (5.6) and (5.7)), the regression fit improves somewhat.

The results so far indicate that a public choice based model, augmented by unemployment and earthquake variables, performs fairly well, 'explaining' around 84% of the changes in E/Y in (5.7).[28] The unemployment variable could, of course, also be consistent with a public choice interpretation such as a voter or political preference for redistribution whereby expenditures are designed to responds to cyclically-induced distributional changes.

In regression (5.8) we add the public administration slope dummy variables, NPM and PNPM (and shift dummies not shown). Inclusion of *Eq* is not supported and hence is omitted in (5.8) – that is, the PNPM public administration dummies provide a better explanation of trends in those later years than the *Eq* dummy. In (5.8) the parameters for NPM and PNPM take the expected signs (negative and positive respectively) and are statistically significant. However the regression fit is not improved by their addition.

One interpretation of this regression is that, after accounting for other factors (including testing for Christchurch earthquake effects), modest changes in the trend of *E/Y* are still observed in association with the hypothesised NPM and PNPM periods. Finally regression (5.9), in light of weak exogeneity tests discussed below, substitutes lags 1-3 for AMTR for the lags 0-2 used in (5.7). Though this fits the data less well as assessed by the adjusted-R^2, it may represent a better description of causal impacts on *E/Y*; see below.

Testing for Exogeneity

We have sought to deal with possible endogeneity problems by, where possible and subject to degrees of freedom limitations, using lagged values of our independent variables in the ARDL specification. As noted earlier we also exclude contemporaneous values of our independent variables in some cases. Nevertheless, despite confirmation of long-run relationships (as shown

[28] This, and subsequent, regressions also allow for a longer-lag structure which is supported by the regression diagnostics.

by the PSS tests, for example) causal inferences from regressions such as those in Table 5 cannot be drawn unless the RHS variables can be regarded as at least weakly exogenous.

As a check for weak exogeneity – that is, whether the RHS variables are 'long-run forcing' for the cointegrating (long-run) parameters – we use the weak exogeneity test due to Johansen (1992) and Boswijk (1995). Specifically, following the approach of Calderón *et al.* (2015) and Gemmell *et al.* (2016a), we test whether changes in the RHS variables are statistically unrelated to the error correction terms from regressions (5.7) to (5.9) in Table 5.

Weak exogeneity is nevertheless consistent with each RHS variable reacting to its own lagged changes, lagged changes of other variables, or lagged changes in E/Y. Based on Calderón *et al.* (2015), we can test for weak exogeneity by estimating marginal models for each of our RHS variables thought likely to be endogenous, using a variable addition test to assess the statistical significance of the relevant error correction terms obtained from Table 5. If the RHS variables are confirmed as weakly exogenous, results in regressions (5.7)-(5.9) may be interpreted as causal effects.

Formally, we test the following marginal models:

$$\Delta x_t = \sum_j \beta_j \Delta \boldsymbol{X}_{t-j} + \sum_j \alpha_j \Delta \boldsymbol{y}_{t-j} + \delta \xi_t(\hat{\beta}) + \varepsilon_t \tag{8}$$

where x_t represents each element of the vector $\boldsymbol{X}_t$ of RHS variables; $\xi_t(\hat{\beta}) = (Y_{t-1} - \hat{\beta}\boldsymbol{X}_{t-1})$ is the estimated error correction (ECM) term, and ε_t is a random error term. The null hypothesis of weak exogeneity involves testing $\delta = 0$, as a *t*-test on the δ for each variable. Rejection of the null implies rejection of weak exogeneity.

Results of these tests for the potentially endogenous variables of the unemployment rate, AMTR and the two expenditure share variables are given in Table 6.[29] Based on regressions (5.7) and (5.8), this would suggest that we can accept the hypothesis of weak exogeneity except for the AMTR variable where the relevant *t*-ratio exceeds its critical value at any reasonable confidence level. There is some evidence of endogenous unemployment in (5.8) with a *t*-ratio of 2.0. This latter result is less of a concern since unemployment is known to be a lagging indicator of the economic cycle such that contemporaneous co-movement of this measure with E/Y is likely to reflect prior cyclical changes in GDP. However, the endogeneity of AMTR suggests more caution in interpreting its effects as causal.

For this reason, the AMTR variable in (5.7) which uses periods t to t-2 values, is replaced in (5.9) with period t-1 to t-3 values. It can be seen that this substantially changes the *t*-test results for AMTR (from 7.76 to 1.45) such that the hypothesis of weak exogeneity can readily be accepted. As Table 5 shows, this lagged AMTR variable continues to have a robust relationship with the expenditure/GDP ratio in both the short- and the long-run in (5.9). The estimated long-run parameter is also little affected (0.994 compared with 1.081). Furthermore, when the NPM and PNPM dummy variables are added to the specification in (5.9), both slope dummies take

[29] It seems reasonable in this context to treat the earthquake and political administration trend variables as exogenous *a priori*. Note also that predetermined variables (such as lagged values of the expenditure shares) need not be exogenous where there is a high degree of persistence in the data.

the expected signs (negative and positive respectively) but neither is statistically significant (t = -0.97; 1.39).

Table 6 Weak Exogeneity Tests: t-ratios (absolute values)

	Unemployment	AMTR	Educ.-Health share	Soc. Sec. & Welfare share
Regression (5.7)	1.64	7.76	0.54	1.14
Regression (5.8)	2.06	8.22	1.37	0.77
Regression (5.9)	0.53	1.45	0.59	0.75

5. Conclusions

Evidence on the economic size of the New Zealand government sector, using various size measures, suggests both short-term variability and a number of possible longer-term trends; see Gemmell et al. (2016). Focusing on one such measure – the ratio of public expenditure to GDP – this paper addressed the question: how far can established models of government size help to 'explain' observed changes in the size of government in New Zealand since the early 1970s?

Drawing on the existing conceptual literature, we began by specifying three distinct econometric models each consistent with explanations offered in one of three separate strands in the international public finance, public choice and public administration literatures. We then considered how far our understanding of public expenditure changes and trends could be improved by nesting these models, and using a general-to-specific approach to see whether any one model dominates or whether a more eclectic explanation finds support.

Our empirical modelling, for the period 1972-2015, suggested that all three models offer some insight into changes in the size of government expenditure in New Zealand. Indeed, the best performing empirical model contains variables associated with each of the three literatures. However public choice based variables appeared to find greatest empirical support in a nested model, which served to weaken public finance based variables. Testing the public administration (PA) model was limited to testing the hypothesis that trends in the data were consistent with three hypothesised time periods when different public management models are alleged to apply. Though the PA model on its own received only limited support (it could be out-performed by a simple autoregressive model with unemployment), there was some evidence that acknowledging the different PA trend periods alongside PF and PC influences was supported.

Perhaps one of the strongest findings is that suitably capturing short-term fluctuations (for example, via the inclusion of an unemployment variable, or suitable lag structures) turn out to be important for reliable estimation of longer-term trends in government expenditures.

What does our modelling suggest were the important influences on changes in government expenditure over the 40-plus year period? Firstly, it is possible to identify what did *not* affect spending; namely there is little support for the view that demand for government-provided goods or transfers grow more than proportionately as incomes increase, or 'Wagner's Law'. Neither does it seem that faster population growth generates more than, or less than, proportionate increases in spending.

Secondly, the prominent roles of the autoregressive and error correction processes confirm (a) that past levels of spending strongly constrain future levels; that is, changes in spending demonstrate inertia; and (b) short-term shocks to spending tend to self-correct back to long-run values. Hence, extravagant expansions, or contractions, to spending by some governments tend to be short-lived.

Thirdly, National-led governments seem to follow through on their reputation for cutting public spending – with the exception of the current (2008-) government. This latter result probably reflects the higher spending after the GFC and Christchurch earthquakes and the limited data available so far after those events. Though regressions may not be able to identify it, the recent data on E/Y does suggest large falls in government spending in recent years, especially since 2011. However, whether this is in excess of declines expected based on other controls, these regressions are probably unable to test reliably.

Fourthly, higher (lower) marginal income tax rates seem to be a good predictor of a larger (smaller) public spending/GDP ratio. This may partly reflect government planning revenue changes in advance of spending changes but the evidence suggests the strong link is with the income tax AMTR rather than tax revenues in general.

Finally, larger shares of redistribution-related public spending appear to foreshadow a larger total public spending/GDP ratio and *vice versa*. This may indicate a tendency for government to treat redistributional spending as the 'marginal spend' when cut-backs or expansions in total spending are pursued (perhaps not surprising since it represent around two-thirds of total spending on average). However, it is also consistent with a redistributional motivation (rather than, say, public good provision) being a strong driving force behind longer-term public sector expansion. This seems likely to become potentially more important in future as population ageing (not formally included in our models) generate greater demands, *ceteris paribus*, for higher superannuation and health spending. Of course this could be counteracted by cuts in other spending types, though our evidence suggests that in the past this has generally been insufficient to prevent a rise on total spending to GDP.

References

Ballard, C.L and Gupta, S. (2016) *Perceptions and Realities of Average Tax Rates in the Federal Income Tax: Evidence from Michigan*. Unpublished manuscript (August), Michigan State University.

Bandyopadhyay, D., Barro. R., Couchman, J., Gemmell. N., Liao, G. and McAlister, F. (2012) Average marginal income tax rates in New Zealand, 1907-2009. *Working Papers in Public Finance*, No. WP02/2012. Wellington: Victoria University of Wellington.

Becker, G. (1983) A theory of competition among pressure groups for political influence. *Quarterly Journal of Economics*. 98, 3, 371-400.

Boston, J. and Eichbaum, C. (2006) State Sector Reform and Renewal in New Zealand: Lessons for Governance', in G.E. Caiden and T. Su (eds.) *The Repositioning of Public Governance: Global Experience and Challenges*. Taipei: Taiwan National University.

Boswijk, H. (1995) Efficient inference on cointegration parameters in structural error correction models. *Journal of Econometrics*, 69, 133-158.

Calderón, C., Moral-Benito, E. and Servén, L. (2015) Is infrastructure capital productive? A dynamic heterogeneous approach. *Journal of Applied Econometrics*, 30, 177-198.

Campos, J., Ericsson, N.R. and Hendry, D.F (eds.), (2005) *General-to Specific Modelling*. Cheltenham: Edward Elgar Publishing, Cheltenham.

Chetty, R., Looney, A. and Kroft, K. (2009) Salience and taxation: theory and evidence. *American Economic Review*, 99, 1145-1477.

Creedy, J. and Gemmell, N. (2004) 'The income tax revenue elasticities of income and consumption taxes in the UK. *Fiscal Studies*, 25, 55-77.

Creedy, J. and Gemmell, N. (2010) *Modelling Corporation Tax Revenue.* Cheltenham, UK: Edward Elgar.

Cullis, J. and Jones, P. (2009) *Public Finance and Public Choice*. Third Edition. Oxford: Oxford University Press.

Dorey, P. (2009) 'Sharing the proceeds of growth': Conservative economic policy under David Cameron. *The Political Quarterly*, 80, 259-269.

Dunleavy, P.J. (1989) 'The United Kingdom: paradoxes of an ungrounded statism', in F.G. Castles (ed.) *The Comparative History of Public Policy*, pp.242-291. Cambridge UK: Polity

Dunsire, A. and Hood, C.C. (1983) *Cutback Management in Public Bureaucracies*. Cambridge, UK: Cambridge University Press.

Evans, L., Grimes, A., Wilkinson, B. and Teece, D. (1996) Economic reform in New Zealand 1984-95: the pursuit of efficiency. *Journal of Economic Literature*, 34, 1856-1902.

Gemmell, N. (1990) Wagner's law, relative prices and the size of the public sector. *The Manchester School*, 58, 361-377.

Gemmell, N. (1993) *Public Sector Growth: Theories and International Evidence*. (Editor and Contributor). Aldershot, UK: Edward Elgar.

Gemmell, N., Morrissey, O. and Pinar, A. (1999) Fiscal illusion and the demand for government expenditures in the UK: a time-series analysis. *European Journal of Political Economy*, 15, 687-704.

Gemmell, N., Morrissey, O. and Pinar, A. (2004) Tax perceptions and preferences over tax structure in the UK. *Economic Journal*, 114, F117-F138.

Gemmell, N., Kneller, R. and Sanz, I. (2016a) Does the composition of government expenditure matter for long-run GDP levels? *Oxford Bulletin of Economics & Statistics*. 78, 4, on-line at DOI: 10.1111/obes.12121.

Gemmell, N., Gill, D. and Nguyen, L. (2016) The changing size of the state in New Zealand. *Working Papers in Public Finance*, No. WP04/2016. Wellington: Victoria University of Wellington.

Gemmell, N., Gill, D. and Nguyen, L. (2016a) The myth of the shrinking state in New Zealand, 1900-2015. *Policy Quarterly*, 12, 2-10.

Gregory, R. (2006) 'Theoretical Faith and Practical Works: De-Autonomizing and Joining-Up in the New Zealand State Sector', in T. Christensen and P. Laegreid (eds.) *Autonomy and Regulation: Coping with Agencies in the Modern State*. Cheltenham: Edward Elgar, pp. 137–61.

Halligan, J. (2007) Reintegrating government in third generation reforms of Australia and New Zealand. *Public Policy and Administration,* 22, 217-238.

Herwartz, H. and Theilen, B. (2016) Ideology and redistribution through public spending. European Journal of Political Economy, on-line at http://dx.doi.org/10.1016/j.ejpoleco.2016.11.002

Hood, C.C. (1991) Public management for all seasons. *Public Administration*, 69, 3-19.

Inman, R. (1987) Markets, governments and the "new" political economy, in A.J. Auerbach and M. Feldstein (eds) *Handbook of Public Economics*, pp.647-777. Amsterdam: North-Holland.

Johansen, S. (1992) Cointegration in partial systems and the efficiency of single-equation analysis. *Journal of Econometrics*, 52, 389-402.

Lodge, M. and Gill, D. (2011) Towards a new era of administrative reform? The myth of post-NPM in New Zealand. *Governance: An International Journal of Policy, Administration, and Institutions*, 24, 141–166.

Meltzer, A.H. and Richards, S.F. (1981) A rational theory of the size of government. *Journal of Political Economy*, 89, 914-927.

Meltzer, A. H. and Richard, S. F. (1983) Tests of a rational theory of the size of government. *Public Choice*, 41, 403–418.

Mueller, D. (2003) *Public Choice III*. Cambridge, UK and New York, US: Cambridge University Press.

Niskanen, W.A. (1968) The peculiar economics of bureaucracy. *American Economic Review.* 58, 293-305.

Niskanen, W.A. (1971) *Bureaucracy and Representative Government*. New Brunswick, US: Transaction Publishers.

Pesaran, M.H. (1997) The role of economic theory in modelling the long run. *Economic Journal*, 107, 178-191.

Pesaran, M.H. and Shin, Y. (1999) An autoregressive distributed lag modelling approach to cointegration analysis, in S. Strom (ed.), *Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium*, Cambridge University Press, Cambridge, chapter11, pp. 371-413.

Pesaran, M.H., Shin, Y. and Smith, R.P. (1999) Pooled mean group estimation of dynamic heterogeneous panels. *Journal of the American Statistical Association*, 94, 621-634.

Pollitt, C. (1995) Justification by works of faith: evaluating the new public management. *Evaluation*, 1, 133-154.

Scott, G., Ball, I. and Dale, T. (1997) New Zealand's public sector management reform: implications for the United States. *Journal of Policy Analysis and Management*, 16, 357-381.

Tanzi, V. and Schuknecht, L. (2000) *Public Spending in the 20th Century. A Global Perspective*. Cambridge, UK and New York, US: Cambridge University Press.

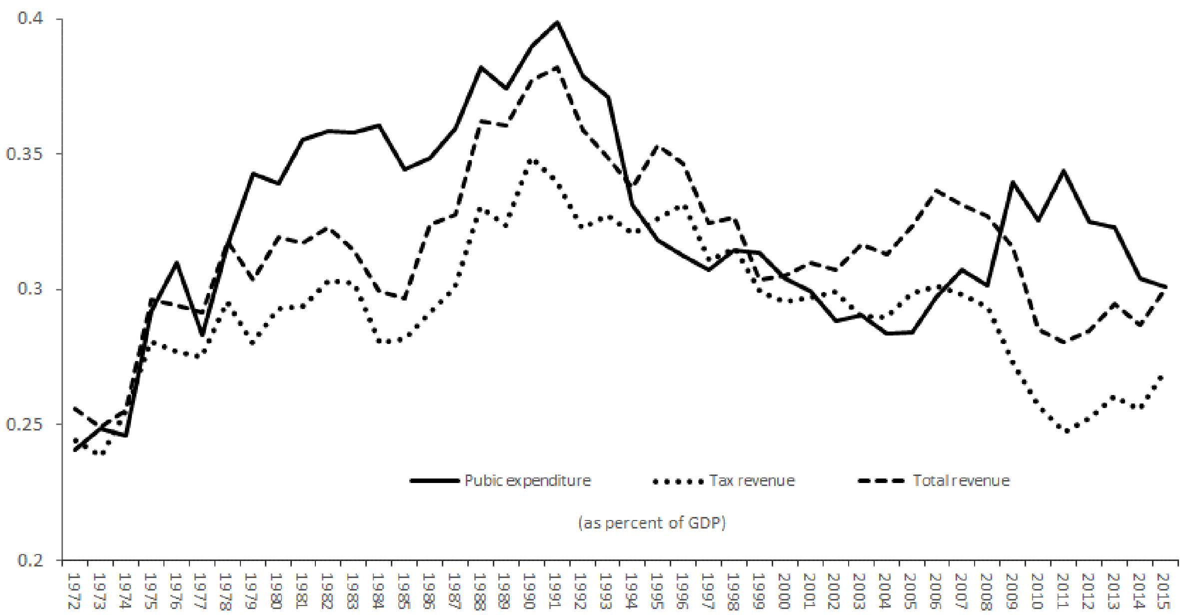
APPENDIX

Appendix Table A1 Unit Root (Augmented Dickey-Fuller) Tests for Regression Variables

Variable	DF statistic*	Variable	DF statistic*
Exp	-1.292	GDPpc	0.332
AMTR	-1.772	Log(GDPpc)	-0.139
Na	-1.549	Pop	3.046
Social-Welfare-share	-2.600	P_g/P_c	-0.914
Education-Health share	-0.759	Unemp	-2.028
Bureau	-1.721		

Note: * 5% (1%) DF critical values are -2.95 (-3.63), n = 43. First differencing yields stationary variables in all case except Pop which appears to be I(2).

Appendix Figure A1 Government Expenditure, Tax Revenue and Total Revenue, 1972-2015



About the Authors

Norman Gemmell is Professor of Public Finance at Victoria Business School, Victoria University of Wellington, New Zealand.
Email: norman.gemmell@vuw.ac.nz

Derek Gill is a Principal Economist at NZIER and a Research Associate of the IPGS at Victoria University of Wellington, New Zealand.
Email: derek.gill@nzier.org.nz

Loc Nguyen is a Research Assistant at NZIER and a former Honours student in the School of Economics at Victoria University of Wellington.

VICTORIA UNIVERSITY OF WELLINGTON TE WHARE WĀNANGA O TE ŪPOKO O TE IKA A MĀUI

Chair in Public Finance
Victoria Business School

Working Papers in Public Finance