Using macroeconomic forecasts to produce detailed forecasts of industry GDP and employment[1]

Paper presented to
52nd New Zealand Association of Economists Annual Conference
Wellington

David Grimmond
Infometrics Ltd

30 June 2011

[1] This paper was prepared with the financial assistance of the Department of Labour. Although I am very grateful for the financial assistance of the Department and also to comments received from Department staff, particularly Ram SriRamaratnam and Dirk Van Seventer, all opinions and the responsibility for any errors in this paper are my own.

Table of Contents

Authorship

This report has been prepared by David Grimmond.

Email: david.grimmond@infometrics.co.nz

1. INTRODUCTION

The purpose of this paper is to document the method of forecasting industry output and employment that captures historical inter-relationships between activity levels in each industry and also the inter-relationship of this activity with macro-economic conditions. The model development restricts the selection of exogenous variables to a subset of key-indicator variables that are forecast in the macro-economic forecast process. The result is the production of detailed industry level forecasts that are based on an internally consistent set of macroeconomic forecasts. The structure of the model ensures that activity levels are constrained to sum to national aggregates.

This paper focuses on forecasts of industry output and employment at the national level. However, Infometrics has extended the approach, with use of its regional database to generate a framework for forecasting employment and output in each territory authority area in New Zealand.

This paper was prepared with the financial assistance of the Department of Labour. Although I am very grateful for the financial assistance of the Department and also to comments received from Department staff, particularly Ram SriRamaratnam and Dirk Van Seventer, all opinions and the responsibility for any errors in this paper are my own.

2. Model structure

A mix of principal component and regression techniques is used to link key macroeconomic indicators (eg inflation, interest rates, unemployment, the exchange rate, business profitability etc) to prospects for individual industries. The approach produces forecasts for individual industries that account for the recent performance of the industry, the impact of key macroeconomic influences on performance in that industry, and is also constrained to ensure that the sum of production in all industries equals forecasts of overall economic activity. Forecasts of employment in individual industries are derived from the industry output forecasts and industry specific forecasts of labour productivity.

Principal component analysis is critical to the modelling approach. Principal component analysis allows a structured decomposition of a panel dataset into independent (orthogonal) clusters that have co-moved over time. This is important as macro-economic forces can have simultaneous (though not necessarily equal) impacts on output and employment in different industries. Using principal components allows inter-dependencies between different industries to be isolated and a reduction in the dimensionality of the analysis (thus removing noise from the model and focusing on the underlying economic drivers). The orthogonal properties of the principal components mean that one can estimate forecasting equations for each of the associated time components independently using standard single equation OLS regression techniques and then apply matrix multiplication to obtain forecasts for individual industries.

Figure 1: Conceptual illustration of model structure

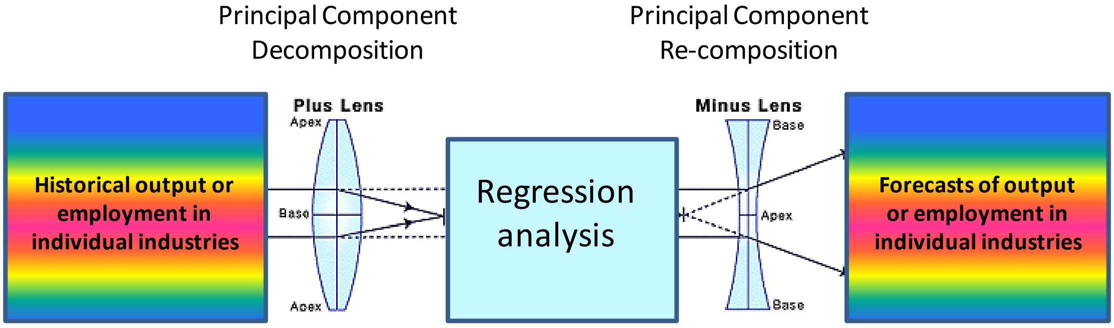


Figure 1 presents a schematic illustration of the model structure. Historical panel data of industry output and employment is analysed using principal component analysis in order to cluster historical co-movements. The derived principal components have two dimensions, a cross section correlation element and an inter-temporal time component that records common movements in the cluster over time. A key property of principal component analysis is an ability to order the components in order of their ability to explain historical variation in the panel of interest. This property can be used to reduce the dimensionality and scale of the forecast problem, for example from forecasting say 30 inter-dependent time series into one of forecasting 5 or fewer independent time components.

Thus the model structure becomes one of using principal component analysis to identify the critical time components. One then uses regression techniques to develop forecast models for each of the time components. Once one has derived forecasts of the time components one can then generate forecasts of the full panel by multiplying the forecast time components with their associated cross-sectional elements.

Model aggregates are constrained to sum to national aggregates by forecasting industry output in terms of shares of national output and selecting one industry as the residual (the notional 'Unallocated' industry). Employment forecasts are linked to output forecasts by modelling industry labour productivity, expressed as output per employee. Thus employment numbers are derived by dividing forecasts of output by forecasts of output per employees. This approach does not ensure that industry employment forecasts will necessarily sum to national forecasts of aggregate employment, but this can be forced on the model by applying some minor post-model reconciliation.

3. Data

Historical production data is sourced from the actual chain-volume series of gross domestic product by industry and is expressed in 1995/96 prices. The industries are based on the 31 national accounting industry definitions (see Table 1). The ownership of owner-occupied dwellings and the unallocated industries are notional in nature, with the former accounting for the service received by homeowners living in their own homes and the latter accounting for the collection of indirect taxes in the production process and the transfer of funds associated from financial services.

Table 1: National account industries

Industries
Agriculture
Fishing
Forestry and logging
Mining
Food, beverage and tobacco manufacturing
Textile and apparel manufacturing
Wood and paper product manufacturing
Printing, publishing and recorded media
Petroleum, chemical, plastic and rubber product manufacturing
Non-metallic mineral products manufacturing
Metal product manufacturing
Machinery and equipment manufacturing
Furniture and other manufacturing
Electricity, gas and water supply
Construction
Wholesale trade
Retail trade
Accommodation, restaurants and bars
Transport and storage
Communication services
Finance and insurance
Property services
Ownership of owner-occupied dwellings
Business services
Central government admin and defence
Local government administration
Education
Health and community services
Cultural and recreational services
Personal and other community services
Unallocated
Total gross domestic product

The analysis of industrial production uses production as a share of total national GDP, ie:

$$q_{i,t} \equiv \frac{Q_{i,t}}{\sum Q_{i,t}} \tag{1}$$

This normalises the analysis around the level of total GDP and converts the forecasting problem to one of determining the proportion of total GDP produced by each of the 31 production industries.

The employment data is sourced from Infometrics' Regional Industry Employment Model (RIEM) database. The RIEM is built on quarterly and annual LEED data extracted by special request from Statistics New Zealand at the territorial authority level. The employment measures use information from the annual series of LEED and census to provide estimates of self employment by industry. The inclusion of self employment helps to avoid a significant undercount of employment in certain industries such as agriculture and construction. The RIEM provides estimates of the number of people employed in 480 industries in each region and territorial authority for each quarter since March 1997. In the current exercise the focus is on national employment and in the 29 national accounting industries (ie the same as in Table 1 but excluding the two notional industries: ownership of owner-occupied dwellings and the unallocated industry).

The input into the forecast model is actually of industrial output per employee:

$$l_{i,t} \equiv \frac{L_{i,t}}{Q_{i,t}} \tag{2}$$

Forecasts of industrial output and employment are therefore conditional on having a forecast of overall GDP, ΣQ_i, producing forecasts of the individual industry's share of GDP, q_i, and industrial output per employee in each industry, l_i. Thus forecasts of employment (L) in industry i in period t can be defined as:

$$\hat{L}_{i,t} = \hat{l}_{i,t} \cdot \hat{q}_{i,t} \cdot \sum \hat{Q}_{i,t} \tag{3}$$

This approach does not ensure that industry employment forecasts will necessarily sum to national forecasts of aggregate employment, but this can be forced on the model by applying some minor post-model reconciliation so that the final estimate for industry i depends on its proportion of the summed industry estimates times the aggregate forecast of national employment. eg

$$L^f_{i,t} = \frac{L^e_{i,t}}{\sum L^e_{i,t}} \cdot L^f_{a,t} \tag{4}$$

4. PRINCIPAL COMPONENT DECOMPOSITION

Principal component analysis is a mathematical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of uncorrelated variables called principal components. The number of principal components is less than or equal to the number of original variables. This transformation is defined in such a way that the first principal component has as high a variance as possible (that is, accounts for as much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible under the constraint that it be orthogonal to (uncorrelated with) the preceding components.

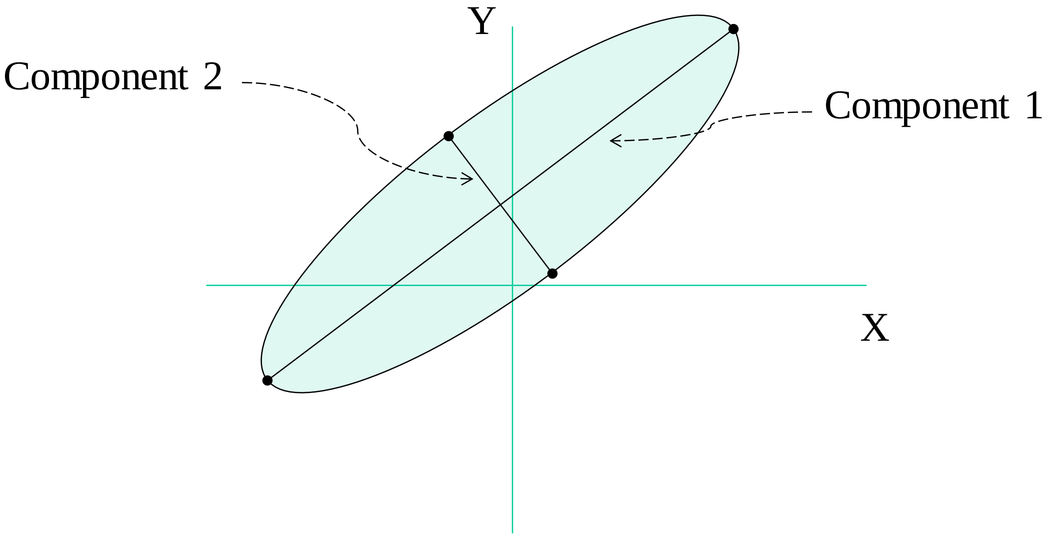


Figure 2

The main applications of principal component techniques are: (1) to *reduce* the number of variables and (2) to *detect structure* in the relationships between variables, that is to *classify variables*. We adopt a principal component approach because we consider there to be a strong chance that industrial production and employment will respond in a common way to external stimuli, and there may also be internal interactions between the industries (eg an increase in production in one industry will stimulate production levels in associated industries). The principal component approach uses eigenvector/eigenvalue matrix algebra techniques to distil correlations between activity in different industries.

The approach has the added advantage of reducing the number of forecast equations, in the case of industrial production from around 30 forecast equations to 5 or fewer. The historical database of industrial production uses 88 quarterly observations from June 1988 through to

March 2010 for each of 30 industries. The analysis is undertaken using industrial production as a share of total national GDP, ie:

$$q_{i,t} \equiv \frac{Q_{i,t}}{\sum Q_{i,t}} \tag{5}$$

This normalises the analysis around the level of total GDP and converts the forecasting problem to one of determining the proportion of total GDP produced by each of the 31 production industries. As the sum of these shares equal 1 by definition, it is inappropriate to include all of the industry shares as this would result in perfect colinearity in the resulting matrix. To resolve this issue the principal component analysis excludes the notional 'unallocated' industry, with forecasts of this industry derived as the residual item, ie:

$$\hat{q}_{u,t} = 1 - \sum_{i=1}^{u-1} \hat{q}_{i,t} \tag{6}$$

The net result is that the principal component analysis for industrial production is based on a 30 by 88 matrix of 30 industries and 88 time periods (quarters), which we denote as S.

Principal component analysis can then decompose the matrix S into 30 orthonormal components (ie thirty 88x1 vectors) with each component's associated eigenvector. A property of this decomposition is that matrix multiplication of the stacked principal component and eigenvectors yields the original matrix of industrial output shares, ie

$$S = PC * EV \tag{7}$$

where PC is an 88x30 matrix of principal components and EV is a 30x30 matrix of associated eigenvectors. Thus the matrix multiplication product of the decompositions yields the 30x88 matrix S as per the equation above. The principal components and associated eigenvectors are ordered in descending order. The implication is that one can generate estimates of S using fewer eigenvectors, eg for any k<30 one can estimate S by:

$$S \approx PC(88 \times k) * EV(k \times 30) \tag{8}$$

This property plus the linear independence of each principal component (orthonormality) are exploited to generate forecasts. As a high percentage of the variation in S can perhaps be explained by a small number of principal components it is usually possible to greatly simplify the forecasting process. Forecasts can be generated for the entire dataset by forecasting just a handful (k) of the principal components. The orthonormality of the principal components means that each principal component can be forecast independently, and it is the eigenvectors that allow the forecasts to capture co-movements in the series. In practice, time series forecasting processes are used to forecast values of PC. These forecast values are then multiplied against the eigenvectors to generate forecasts of the industry shares, ie the re-composition phase illustrated in Figure 1.

The principal component analysis for output per employee is similar, except that the dimension of the panel data matrix, S is 44 quarters by 29 industries. The shorter time period reflects the availability of LEED based employment data and concerns with the reliability of linking the RIEM employment data with earlier employment survey data. This is illustrated with reference to employment in the textile and apparel industry in Figure 3, where the properties of the quarter to quarter movements in employments are quite different prior to the availability of LEED employment data.

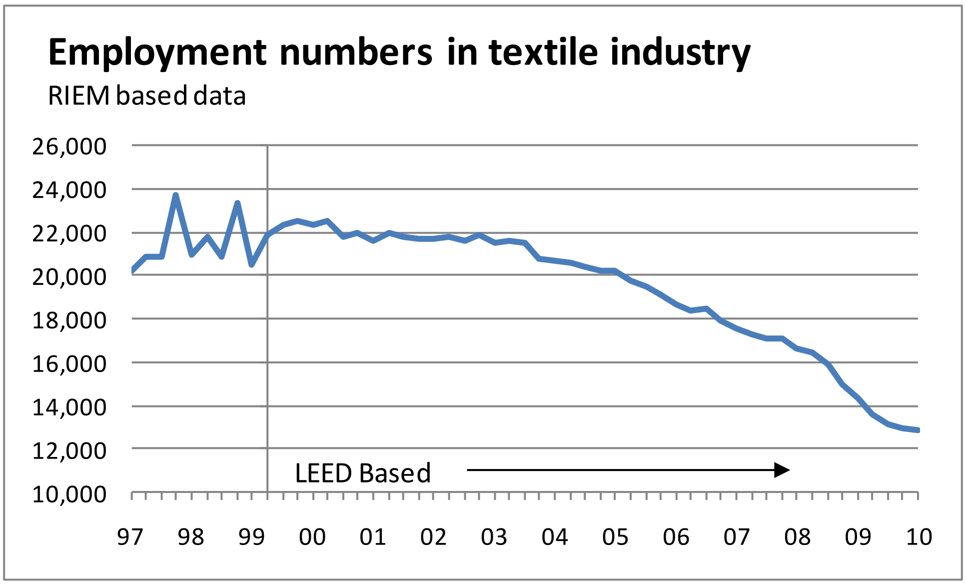


Figure 3

The use of 29 industries reflects the absence of employment in the two notional industries: ownership of owner-occupied dwellings and the unallocated industry. As the output per employee data is not share based all relevant industries need to be included in the principal component analysis.

Principal component results

The results of the principal component disaggregation for industry output shares are presented illustratively for the first five components in Figure 4. Similar graphs are produced for output per employees in Figure 5. The top graphs illustrate the eigenvectors, which can be interpreted as cross correlations between different spending items. The second tier of graphs illustrate the way that this weighted cluster has evolved over time, known as the time component or principal component.

In between the top and bottom tier of graphs is a measure of the amount of variation explained by each principal component, both in terms of additional and cumulative explanation. Thus in terms of the analysis of shares of GDP (Figure 4) 63.4% of the variation in industry shares of GDP during the 88 quarter period from June 1988 to March 2010 can be explained by the first principal component. An additional 18.4% of the variation is explained by the second component, so the first two components explain a cumulative 81.4% of the variation. Five components explain 95.2% of the variation, though the fifth component explained just 2.8% of this variation.

From a more descriptive perspective, one can interpret the first component as describing a trend since the mid-1990s of an expanding share of economic activity in New Zealand taking place in communications, finance, business services and to a lesser extent property and health services. These expansions have accompanied a downward trend in the importance of services from owner-occupied dwellings, and to a lesser extent primary production, manufacturing and government activity. One obtains this interpretation from multiplying the cross correlations in the top left graph by the associated time component in the lower left graph. The time component is dominated by a downward trend since the mid-1990s. Thus, industries with negative cross correlations, such as the communications industry, are industries that have experienced a trend expansion in share of GDP.

The other graphs of time components in Figure 4 indicate that the other components are less about secular trends and more about cyclical shifts in production between different industries. The second component appears to be primarily about cycles in agriculture and food processing production. The third seems to be about some offsetting forces between production in services such as finance, ownership of dwellings, and central government on one side with wholesale services, construction and manufacturing activity. It is difficult to discern any meaningful explanation in the variation described by the fourth component, however the fifth component appears to be mainly related to a large step increase in food processing production that took place in the September quarter in 2002.

Having decomposed the industry share of output and output per employee data into cross correlation and time components, the forecast problem reduces into one of establishing forecast equations for each of the time components. Once one has a forecast of the time components one can generate forecasts of the industry shares of output and output per employees using matrix multiplication as per equation 8. The time component graphs in Figure 4 and Figure 5 actually present forecasts of these time components. We turn to explaining the process for deriving these forecasts in the next section.

Figure 4: Principal component analysis of shares of GDP

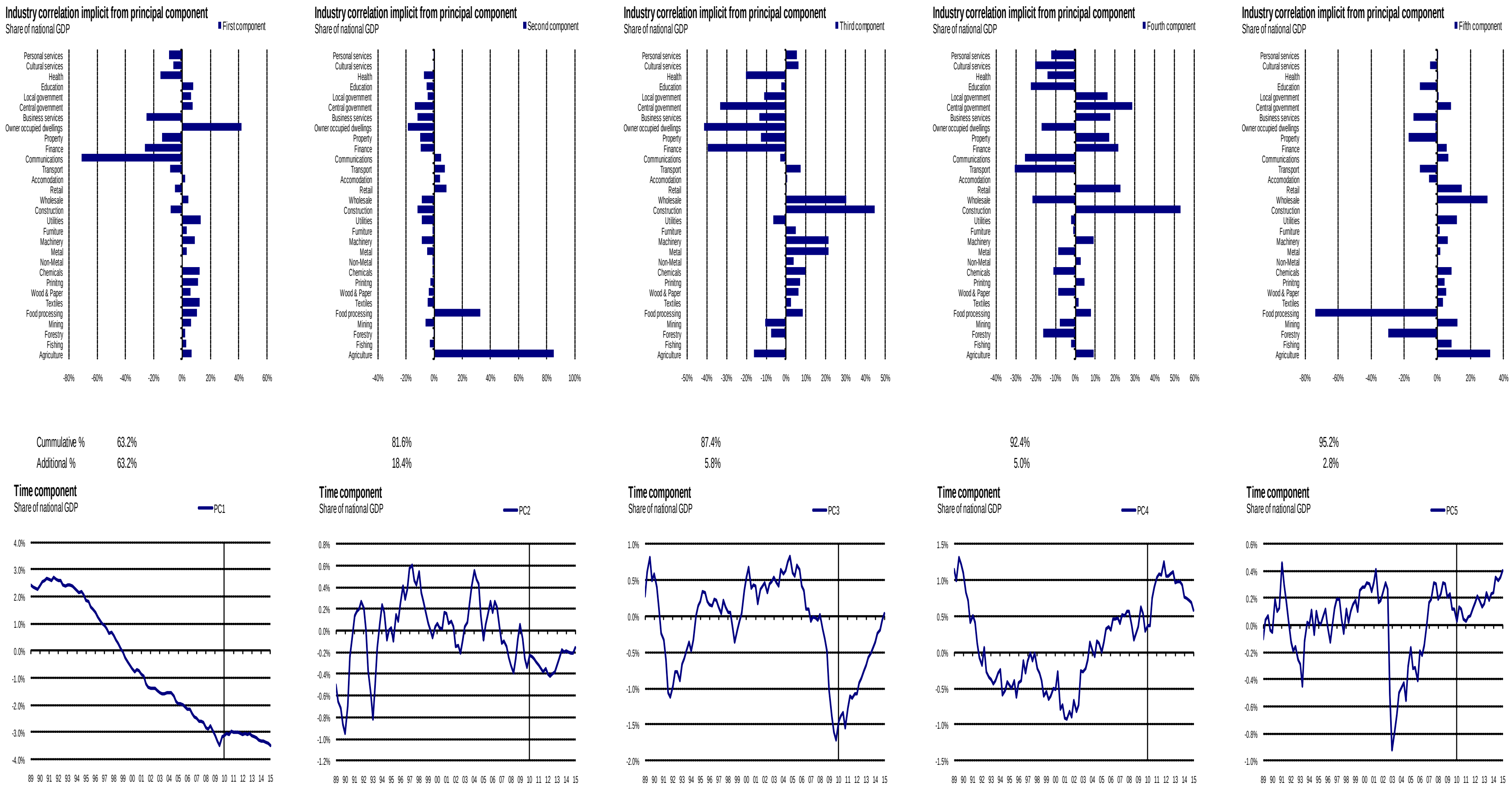
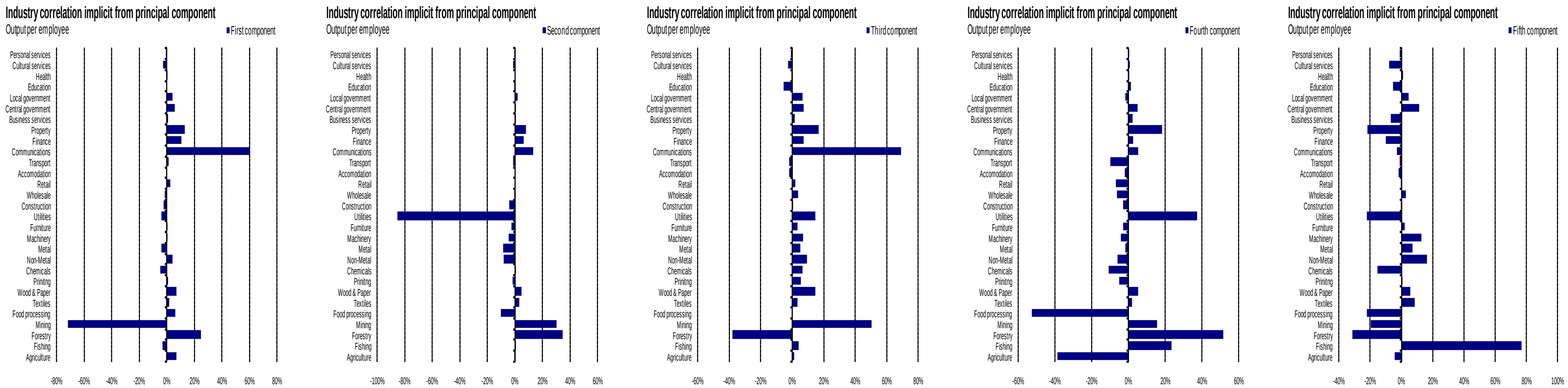
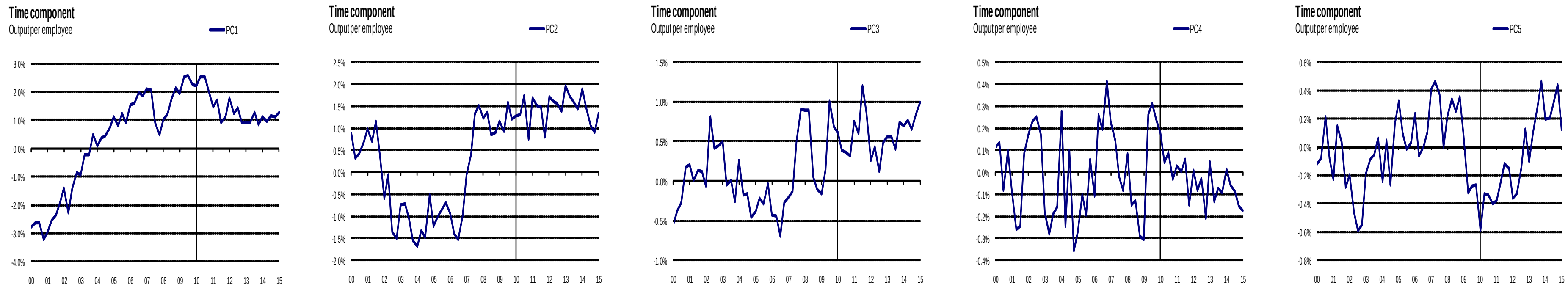


Figure 5: Principal component analysis of industrial output per employee measures



	First	Second	Third	Fourth	Fifth
Cummulative %	62.7%	84.1%	92.2%	95.4%	97.1%
Additional %	62.7%	21.4%	8.1%	3.2%	1.8%



5. ESTIMATION RESULTS

The forecast process becomes one of forecasting the time components illustrated in the bottom half of Figure 4 and Figure 5. An equilibrium correction model structure is used. The orthonormal property of the relationship between individual principal components suggests that there is little advantage from using vector estimation approaches so the forecast equations are estimated independently.

The specific model structure of individual forecast equations is empirically based using a general-to-specific approach for selecting explanatory variables. This approach suits the forecast aim of the current model and the empirical underpinning of the clustered time components that constitute the objective variables of the forecast equations. Consistent with the modelling approach of David Hendry we "err on the side of profligacy" at the early stages of model development (Hendry, 2011, p10). However, variable parsimony is an objective of the variable selection process.

The potential list of explanatory variables tested are variables that are both forecast (or at least monitored) in the current forecast process within Infometrics and which were assessed to have some potential influence on supply or demand conditions in different industries. This approach reflects a philosophy that the objective of any model "is to get out more than you put in" (Ball, 2004, p160). The usefulness of a forecast model is reduced if it simply transfers the source of forecast uncertainty from the endogenous to the exogenous variables. The aim instead is to have the model build on the judgements that underpin a thorough and robust macroeconomic forecast process. The intended result is to produce industry output and employment forecasts that are consistent with a set of macroeconomic forecasts. The full list of variables experimented with is presented in Table 2. The table also lists the presence of variables in the final forecast equations for the various time components for industry share of GDP and output per employee.

In practical terms, the first step in model development was to filter the potential explanatory variables based on their correlation with the dependent variable. The Autometrics procedure in PcGive is then used to select the best candidates. This process was also used to select the appropriate lag structure, with up to four lags of explanatory variables examined. If necessary the models were further streamlined after examining the cross correlation matrix of the regressors, in particular to remove variables that were highly correlated with other explanatory variables.[2] Assuming that no issues are highlighted by the diagnostic tests, the final hurdle is that the resulting model is intuitive from economic grounds.

[2] This is evidence of multicolinearity, but potentially more important in terms of our forecasting focus, is that it signals the inclusion of spurious regressors that will potentially distort consequent forecasts.

Table 2: Explanatory variables investigated in model estimation

		Variable present in explanatory equations for:									
		GDP					Output per employee				
Code	**Description**	**NPC1**	**NPC2**	**NPC3**	**NPC4**	**NPC5**	**LPC1**	**LPC2**	**LPC3**	**LPC4**	**LPC5**
	Labour market										
QH	Productivity (real output/hr)	*	*					*			
CoE	Comp of Employess % tot income				*			*			
RWY	Real wage (income)										*
RWQ	Real wage (product)						*				
PR	Labour participation rate									*	
Emp	Employment	*									
UR	Unemployment rate	*								*	*
ER	Employment Rate					*	*				*
	External sector										
BoP	BOP current account balance/GDP		*			*	*		*		
ToT	Terms of trade		*			*		*			
TDE	Total domestic expenditure as % of GDP						*		*		*
	Business sector										
BusI	Bus investment as % GDP	*						*			*
posr	Private operating surplus as % of GDP	*		*		*			*		
	Financial										
TWI	Trade weighted index of New Zealand exchange rate	*		*	*					*	
90day	90-day bank bill rate										*
NZbond	10-year government bond rate, NZ						*		*		
Wbond	10-year government bond rate, World			*							
MCI	Monetary conditions index (calculated using 3xTWI:1x90 day)		*								
	Government sector										
GCR	Government consumption as % of expenditure on GDP		*	*	*						*
GIR	Government investment as % of expenditure on GDP		*		*	*			*		
	Climate										
SMD	Soil moisture deficit		*							*	
	Not used										
ULC	Unit labour costs (wage bill/real output)										
Hours	Average weekly hours (millions)										
Mcons	Imports/Consumption										
PrivI	Private investment (% of GDP)										

This might seem a bit late in the process to consider the economic meaningfulness of the equation, but this needs to be balanced against the fact that the variables tested were limited to those that were potentially meaningful. Also the models have a reduced form structure, given the forecast purpose of the models. This means that we are looking for meaningful correlations, and abstract from identifying the causal transmission paths.

As noted, an equilibrium correction model structure is used which has the general equation structure of a long run equation estimated in levels:

$$Y_t = \alpha_0 + \sum \alpha_i X_{i,t} + z_t \tag{9}$$

Where Y is the dependent variable, X_i are exogenous explanatory variables and z_t is the model error term. The associated short run equation is:

$$\Delta Y_t = \beta_0 + \beta_1 z_{t-1} + \sum \beta_j \, \Delta X_{j,t} + \mu_t \tag{10}$$

Where Δ signifies the variable is expressed in change format and z_{t-1} is the lagged error term from the long run equation. The logic behind this equation structure is that the long run error term is measuring aspects of the extent that the variable of interest has diverged from its long run sustainable, or equilibrium, level, and that this will influence the subsequent pattern of change. That is the forecast approach is premised on a relationship of reversion to a stable long run relationship as established in the long run equation (9).

Estimation results for the preferred forecasting equations are presented in the Appendix 1. All variables are seasonally adjusted and except for the balance of payments and the monetary conditions index, are expressed in natural logarithms.

The model results are generally acceptable with evidence of cointegrated long run relationships and no evidence of statistical issues with any of the diagnostic tests associated with the short run (change) equations. There is potentially one equation that will warrant further examination, the equation for lpc4, where we failed to find a cointegrated long run relationship based on the variables tested. There is also evidence of some coefficient instability in this equation (with respect to the magnitude of the coefficient for the unemployment rate) when was estimated over a shorter period (up to March 2008), see *Model forecast performance* below. However, to put this issue in context, the fourth principal component explains just 3.2% of the variation in industry output per employee over the estimation period.

The actual equations used for forecasting the time components are derived by substituting for

The actual forecast equations are derived by first rearranging equation (9):

$$z_t = Y_t - \left(\alpha_0 + \sum \alpha_i X_{i,t}\right) \tag{11}$$

Then by substituting this into equation 10, and then by rearranging (noting that $\Delta Y_t = Y_t - Y_{t-1}$) to produce a forecast equation of the form:

$$\widehat{Y}_t = (\beta_0 - \gamma\alpha_0) + (1+\gamma)Y_{t-1} - \gamma\left(\sum \alpha_i X_{i,t-1}\right) + \sum \beta_j \Delta X_{j,t} \tag{12}$$

The resulting forecasting equations are presented in Appendix 2.

6. MODEL PROPERTIES

In this section we examine some of the properties of the resulting forecast model of industry output and employment. Table 3 presents calculations of the implicit explanatory power (R^2) of the forecast models of industry shares of GDP and industry output per employee. These are calculated by weighting the R^2 of the short run equations for each of the time components by the percentage of variability explained by each component. Thus, the explanatory power is dominated by the performance of the forecast equations for the first time components, npc1 and lpc1. The net result of these calculations are implicit R2's of 0.4667 for the industry share of total GDP and 0.6519 for industry output per employee.

Table 3: Implicit in-sample model explanatory power

	Percentage of variability explained by component	Short run equation R^2
Industry share of total GDP		
NPC1	63.2%	0.5347
NPC2	18.4%	0.3303
NPC3	5.8%	0.5047
NPC4	5.0%	0.4652
NPC5	2.8%	0.5474
Implicit total	95.2%	0.4667
Industry output per employee		
LPC1	62.7%	0.7714
LPC2	21.4%	0.4440
LPC3	8.1%	0.5804
LPC4	3.2%	0.5006
LPC5	1.8%	0.5793
Implicit total	97.1%	0.6519

This result suggests that the model only explains about half of the observed variation in industry shares of GDA and two-thirds of the variation in industry output per employee. This would suggest that caution is warranted in expectations of the potential forecast performance of the model. Below we examine:

- model stability by re-estimating the model over a shorter time period (up to March 2008)
- ex-post forecast performance by examining the model's ability to "forecast" industry and employment outcomes in the eight quarters to March 2010.
- The sensitivity of model forecasts to one-standard-deviation impulses in exogenous variables.

Model forecast performance

The explanatory equations underpinning the forecast model, as described in the *Estimation results* section, were re-estimated over a shorter time period, ending in the March quarter 2008. The model estimated over the earlier period is then used to forecast outcomes in the following eight quarters until March 2010. This provides an indication of the forecast performance of the model based on full information (ie it is based on actual, rather than forecasts of the exogenous variables). This exercise highlighted two potential issues with the model structure:

1. There is some evidence of coefficient instability in the change equation for the fourth principal component of industry output per employee (lpc4). The coefficient for the unemployment rate variable ($DLURSA_{-1}$) is 0.0228 (SE=0.006) when estimated up to March 2008, but declines to 0.0091 (SE=0.004) when estimated up to March 2010. Although the variable is significant in both equations, the magnitude of the impact is statistically smaller in the enlarged sample period.
2. The model has problems forecasting the third principal component for industry share of national output (npc3). Although changing the estimation period did not produce any discernible or statistically significant differences in the model structure, the resulting forecasts of the npc3 variable were consistently greater than actual outcomes, and statistically so in three of the eight forecast quarters.

Although these issues warrant further investigation, their relative importance should also be considered. The npc3 variable relates to just 5.8% of the past variation in industry output shares and lpc4 relates to just 3.2% of the variation in output per employee between industries.

Also as Hendry and Clements (2001) note, forecast accuracy measures should relate to the relevant loss function, ie what is the purpose of the forecasts. Here the time component forecasts are just intermediate variables; of themselves they are reasonably meaningless, except in their role in assisting the process of forecasting industry level GDP and employment. Therefore it is perhaps more useful to examine the forecast performance relative to these final factors.

The performance of the model in forecasting industry GDP and employment in the eight quarters to March 2010 are presented in Table 4. The columns labelled % of industry present the average error in each industry expressed as a percent of actual outcomes (ie $(\hat{Y}-Y)/Y$). The forecast performance for GDP appears to be better than for employment, with the root mean square percent error for GDP of 5.8%, compared with 8.0% for employment. The forecasts erred on the positive side for GDP (average error of 1.3%) and on the negative side for employment (average error of -1.4%).[3] As a percent of industry output, the industry

[3] The forecasts were of shares of output and employment, the national totals were therefore a given known fact in this exercise. A positive average error simply means that the count of positive errors outnumbered the count of negative errors, and vice versa for a negative average.

GDP forecast errors ranged from -5.7% (forestry and logging) to 11.5% (textile and apparel manufacturing). For employment the industry forecast errors ranged from -20.6% of industry employment (fishing) to 14,4% (food, beverage and tobacco manufacturing).

The '% of national' columns put these industry forecasts in context by dividing the errors by national output and employment. This can be viewed as providing an indication of where errors are more material from a national perspective. For example the 21% error in forecasts of employment in fishing needs to be taken in the context that employment in fishing comprised just 0.2% of national employment during this period.

In terms of national output and employment, the forecast errors for industry output ranged from -0.3% of national output (communication services) to 0.4% (construction). The employment forecast errors ranged from -0.5% of national employment for business services to 0.7% for agriculture.

The sum of absolute errors in the bottom row of the table probably provides the best summary of the model's forecast performance. These indicate modelling errors of 2.6% for industry output and 4.9% for employment.

In both regards the forecast performance of the model appears to perform to a higher standard than was possible suggested by our calculations of implicit R2 reported in Table 3.

Table 4

Forecast errors of GDP and employment in eight quarters to March 2010

Industry	GDP		Employment	
	% of industry	% of national	% of industry	% of national
Agriculture	-4.6%	-0.2%	12.3%	0.7%
Fishing	-0.2%	0.0%	-20.6%	0.0%
Forestry and logging	-5.7%	-0.1%	-2.8%	0.0%
Mining	-5.1%	-0.1%	-17.9%	0.0%
Food, beverage and tobacco manufacturing	1.7%	0.1%	14.4%	0.4%
Textile and apparel manufacturing	11.5%	0.1%	5.2%	0.0%
Wood and paper product manufacturing	0.8%	0.0%	-5.9%	-0.1%
Printing, publishing and recorded media	8.8%	0.1%	4.9%	0.0%
Petroleum, chemical, plastic and rubber products	2.7%	0.0%	0.0%	0.0%
Non-metallic mineral products manufacturing	5.2%	0.0%	0.5%	0.0%
Metal product manufacturing	6.2%	0.1%	-1.2%	0.0%
Machinery and equipment manufacturing	9.2%	0.2%	2.1%	0.0%
Furniture and other manufacturing	8.1%	0.0%	3.5%	0.0%
Electricity, gas and water supply	0.3%	0.0%	-5.8%	0.0%
Construction	9.3%	0.4%	5.6%	0.4%
Wholesale trade	0.8%	0.1%	-2.4%	-0.1%
Retail trade	0.7%	0.0%	2.6%	0.3%
Accommodation, restaurants and bars	0.2%	0.0%	3.8%	0.2%
Transport and storage	-2.5%	-0.1%	-1.0%	0.0%
Communication services	-4.5%	-0.3%	-9.3%	-0.1%
Finance and insurance	-1.8%	-0.1%	-4.1%	-0.1%
Property services	0.2%	0.0%	-4.6%	-0.1%
Business services	-1.5%	-0.1%	-4.6%	-0.5%
Central government admin and defence	-0.1%	0.0%	-5.4%	-0.3%
Local government administration	0.2%	0.0%	-3.2%	-0.1%
Education	2.9%	0.0%	7.1%	0.2%
Health and community services	-1.7%	-0.1%	-3.5%	-0.2%
Cultural and recreational services	-2.9%	-0.2%	-3.9%	-0.4%
Personal and other community services	-3.6%	-0.1%	-5.6%	-0.2%
Ownership of owner-occupied dwellings	-3.1%	0.0%		
Unallocated	8.7%	0.2%		
Summary statistics				
Average	1.3%		-1.4%	
Root Mean Square Percent Error	5.8%	0.1%	8.0%	0.2%
Largest positive impact	11.5%	0.4%	14.4%	0.7%
Largest negative impact	-5.7%	-0.3%	-20.6%	-0.5%
Range of impacts	17.2%	0.7%	35.0%	1.2%
Sum		0.0%		0.0%
Sum of absolute errors		2.6%		4.9%

Sensitivity tests

A common mantra of David Hendry's is that one needs to be careful using observed forecast performance for model assessment, for example from Hendy and Clements (2001) p16:

> *Forecasting success is not a good index for model selection: the observation of forecast failure merely denotes that something has changed relative to the previous state, with no logically valid implications for the model of that state. Some failures are due to bad models, and some successes occur despite serious misspecification.*

Thus although the ex post forecast performance just discussed do not necessarily highlight any fatal problems with the model, these results are perhaps less interesting than having a better understanding of the sensitivity of forecasts to exogenous forces. To do this numerous impulse tests were undertaken. Each time an impulse equalling one standard deviation in an exogenous variable (or a combination of variables) is imposed in the March quarter 2011. The deviations in coming quarters are then compared with a set of baseline forecasts. The exogenous variables are thus identical to those in the baseline forecasts in every quarter except for the impulse quarter. The implication is that although there are short run deviations, the model forecasts eventually return to the baseline forecasts.

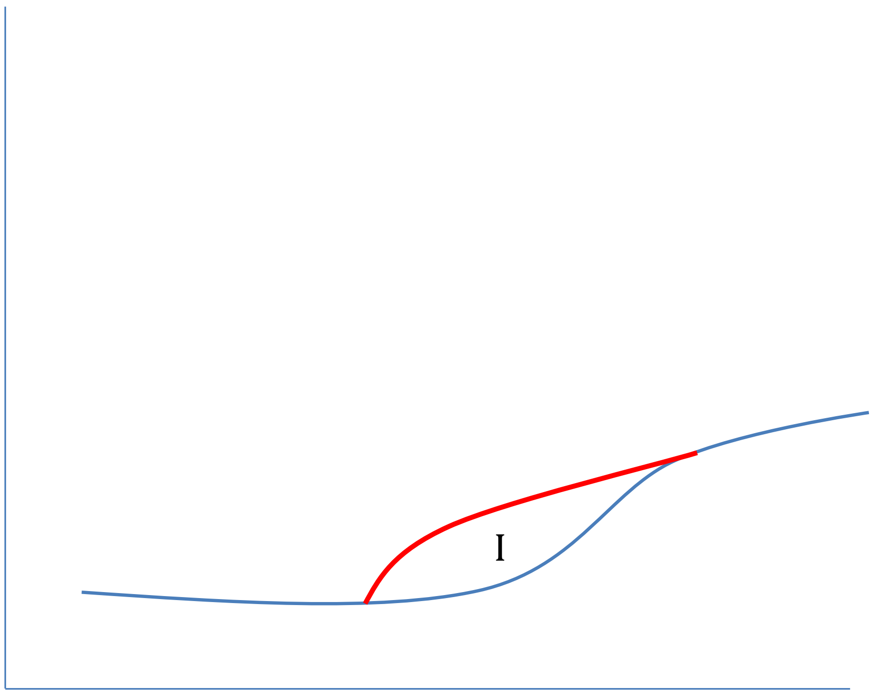


Figure 6: Schematic illustration of impulse

This is illustrated schematically in Figure 6. If the impulse produces a divergence in the outcome for a variable of interest this occurs for a finite period. The analysis that follows relates to measures of the impulse impact, which is equivalent to an estimate of the area of the gap between the curves, labelled I in Figure 6.

A summary of the results of the impulse analysis is presented in Table 5. The results presented are for positive shocks (ie increases in the exogenous variables)[4]. This provides a summary of the scale of the relative impacts. More detail on the impacts on output and employment in specific industries are presented in Table 6 and Table 7.

Table 5

Summary of impulse impact on industry employment and GDP

Root mean square percent deviations

Code	Description	Size of 1 SD impulse	Employment	GDP
GDP	Real production GDP	2.4%	2.3%	2.3%
	Labour market			
QH	Productivity (real output/hr)	2.0%	1.1%	1.1%
CoE	Comp of Employess % tot income	2.1%	1.3%	1.0%
RW	Real wage	1.5%	0.1%	0.0%
PR	Labour participation rate	0.9%	0.5%	0.0%
UR	Unemployment rate	17.7%	1.4%	0.5%
ER	Employment Rate	1.7%	2.7%	0.1%
	External sector			
BoP	BOP current account balance/GDP	233.0%	0.7%	0.6%
ToT	Terms of trade	5.0%	2.4%	2.2%
TDE	Total domestic expenditure as % of GDP	2.3%	2.0%	0.0%
	Business sector			
BusI	Bus investment % GDP	9.4%	0.8%	0.2%
posr	Private operating surplus as % of GDP	8.1%	2.5%	2.7%
	Financial			
TWI	Trade weighted index of New Zealand exchange rate	9.1%	1.1%	1.1%
90day	90-day bank bill rate	29.0%	1.8%	0.0%
NZbond	10-year government bond rate, NZ	14.9%	1.9%	0.0%
Wbond	10-year government bond rate, World	15.1%	1.1%	1.1%
IR	All interest rates		2.8%	1.1%
MCI	Monetary conditions index (calculated using 3xTWI:1x90 day)		1.8%	1.1%
	Government sector			
GCR	Government consumption as % of expenditure on GDP	3.9%	1.1%	1.2%
GIR	Government investment as % of expenditure on GDP	17.9%	1.6%	1.2%
	Climate			
SMD	Soil moisture deficit	58.3%	0.2%	0.2%

The first column in Table 5 provides an indication of what constitutes a standard deviation impulse. Note that the scale of these numbers reflects the average size of the variable as well as the amount of variability. For example, the balance of payments has varied from +2% to -14% of GDP in the period since 1990. Normalising the impulses around one standard

[4] Impulse impacts are generally symmetric, reflecting the linear functional structure of the explanatory equations.

deviation means that one can interpret the root mean square deviations in the two right-hand columns as providing a measure of the relative sensitivity of the model to the various exogenous variables.

The only shock that impacts on the absolute level of activity in the model is changes to national GDP forecasts, and this typically has a proportional impact on industry employment and output. Though, as the results in Table 6 and Table 7 indicate some industries are more sensitive to changes overall levels of economic activity than others.

According to the impulse results a change in private operating surplus typically signals the largest reallocation of output between industries, with output in manufacturing appearing most sensitive to business profitability. The next biggest influence on industry output comes from the terms of trade, with an increase signalling increases in fishing, mining and manufacturing activity, except for food processing, which declines with agriculture production, perhaps implying that the historical phenomenon of a re-stocking farm response to positive terms of trade shock still continues in New Zealand.

These changes also feed through into the allocation of employment. The model also implies that employment by industry is also sensitive to general labour market indicators (eg the employment rate), domestic spending levels (as proxied by total domestic expenditure), and financial conditions.

These employment results perhaps highlight a potential limitation with the model structure. The reliance on empirical evidence for selecting exogenous variables means that alternatives of some similar variables are included in separate places in the model, eg different interest rate measures, different labour market proxies. This potentially puts an onus on using a set of forecast inputs that result from a robust forecast process.

Table 6

Accumulated percentage divergence in employment resulting from postive one standard deviation impulse in exogenous variables in March 2011

Impulse variables

Industry	GDP	QH	CoE	RW	PR	UR	ER	BoP	ToT	TDE	BusI	Posr	TWI	90day	NZbond	Wbond	IR	MCI	GCR	GIR	SMD
Agriculture	2.4%	1.0%	0.3%	0.0%	1.5%	0.3%	-3.7%	0.3%	-1.5%	0.2%	0.1%	-1.5%	-1.7%	0.2%	2.8%	-0.5%	2.5%	-1.5%	-0.5%	-1.2%	-0.7%
Fishing	0.7%	-3.4%	-1.6%	0.5%	-1.4%	-6.8%	4.5%	-2.8%	10.0%	-9.8%	2.3%	-1.2%	1.1%	-8.8%	-2.1%	-0.1%	-11.1%	-7.8%	0.4%	6.1%	0.7%
Forestry and logging	2.3%	-0.2%	0.6%	0.0%	-0.7%	1.0%	-3.6%	1.7%	-5.4%	1.1%	1.8%	0.7%	-1.5%	0.8%	2.7%	-1.2%	2.3%	-0.7%	-0.3%	-2.8%	0.1%
Mining	1.9%	-1.0%	0.8%	0.1%	-0.3%	0.4%	8.5%	-0.9%	1.6%	0.5%	1.1%	-6.4%	-0.8%	0.1%	-6.3%	-1.6%	-7.8%	-0.7%	1.2%	0.1%	0.2%
Food, beverage and tobacco manufacturing	2.3%	0.3%	-1.1%	0.0%	1.0%	1.1%	-1.5%	0.9%	-2.6%	1.2%	-0.8%	0.9%	-0.8%	1.1%	1.2%	0.3%	2.6%	0.3%	-0.6%	-2.2%	-0.3%
Textile and apparel manufacturing	0.2%	-3.2%	1.8%	0.0%	-0.3%	0.3%	-0.5%	-0.7%	1.4%	-1.6%	1.5%	0.9%	2.7%	-1.4%	0.4%	0.7%	-0.3%	1.3%	0.2%	1.4%	0.4%
Wood and paper product manufacturing	2.2%	-0.5%	0.1%	0.0%	-0.3%	-0.3%	-1.7%	-0.5%	0.3%	-0.7%	0.6%	-0.7%	0.1%	-0.6%	1.3%	0.6%	1.3%	-0.6%	-0.8%	-0.8%	0.1%
Printing, publishing and recorded media	1.4%	-1.4%	0.5%	0.0%	0.0%	0.4%	0.1%	-0.4%	1.0%	-0.4%	0.2%	1.6%	1.8%	-0.4%	0.0%	1.1%	0.7%	1.4%	-0.4%	0.5%	0.1%
Petroleum, chemical, plastic and rubber products	1.4%	-1.4%	-1.3%	0.0%	0.1%	1.5%	2.4%	-0.5%	1.1%	0.9%	0.5%	2.6%	-0.2%	0.8%	-1.8%	1.4%	0.5%	0.5%	-2.2%	-0.4%	0.1%
Non-metallic mineral products manufacturing	2.5%	-0.1%	-0.4%	0.0%	-0.1%	-1.1%	-0.7%	-0.3%	0.8%	-1.4%	-0.6%	1.1%	0.9%	-1.2%	0.5%	1.0%	0.3%	-0.3%	-0.3%	0.0%	0.1%
Metal product manufacturing	2.3%	-0.4%	-2.9%	0.0%	-0.1%	-0.7%	2.7%	-0.3%	1.5%	-1.0%	-1.1%	5.2%	0.2%	-1.0%	-2.0%	2.5%	-0.4%	-0.7%	-2.6%	-0.5%	0.2%
Machinery and equipment manufacturing	2.1%	-0.7%	-0.5%	0.0%	0.0%	-1.2%	0.6%	-0.5%	1.5%	-1.8%	-0.5%	2.5%	1.5%	-1.7%	-0.5%	1.7%	-0.4%	-0.2%	-0.7%	0.3%	0.2%
Furniture and other manufacturing	1.6%	-1.2%	-1.3%	0.0%	0.0%	0.1%	0.4%	-0.4%	1.3%	-0.7%	-0.3%	4.1%	1.0%	-0.6%	-0.3%	2.1%	1.2%	0.4%	-1.9%	-0.1%	0.1%
Electricity, gas and water supply	1.9%	-1.0%	-3.2%	0.0%	-0.4%	0.5%	1.0%	-0.4%	2.9%	0.1%	-1.8%	-1.7%	0.6%	0.1%	-0.7%	-0.5%	-1.2%	0.7%	0.7%	0.8%	0.2%
Construction	2.7%	0.1%	0.0%	0.0%	0.0%	-0.4%	2.8%	-0.1%	1.0%	-0.2%	-1.0%	3.5%	1.7%	-0.3%	-2.1%	1.4%	-0.9%	1.5%	-0.1%	1.1%	0.1%
Wholesale trade	2.5%	-0.1%	-0.3%	0.0%	0.0%	-0.4%	1.2%	-0.3%	0.6%	-0.5%	0.1%	1.1%	-0.2%	-0.5%	-0.9%	0.6%	-0.8%	-0.7%	-0.8%	0.1%	0.0%
Retail trade	2.6%	0.3%	0.9%	0.0%	0.3%	-0.3%	-2.0%	-0.2%	-0.1%	-0.2%	0.2%	-0.8%	0.0%	-0.2%	1.5%	0.0%	1.2%	-0.2%	0.2%	0.0%	-0.1%
Accommodation, restaurants and bars	2.4%	0.1%	0.3%	0.0%	0.0%	0.3%	0.8%	0.5%	-1.0%	0.3%	0.2%	2.0%	0.0%	0.3%	-0.6%	0.0%	-0.3%	0.2%	-0.2%	0.2%	-0.1%
Transport and storage	2.7%	0.4%	-1.1%	0.0%	0.1%	-0.3%	-0.1%	0.2%	-0.5%	-0.2%	-0.1%	1.3%	-1.2%	-0.2%	0.0%	0.2%	0.1%	-1.4%	-1.0%	-0.7%	-0.1%
Communication services	3.6%	1.4%	0.2%	0.1%	-0.2%	-1.0%	-5.6%	-0.5%	-0.4%	-0.2%	0.1%	-3.4%	-1.2%	-0.2%	4.2%	-0.1%	3.9%	-1.4%	-0.4%	-1.8%	0.0%
Finance and insurance	2.9%	0.4%	1.2%	0.0%	-0.2%	-0.2%	-1.5%	-0.1%	-0.1%	0.1%	0.3%	-2.6%	0.1%	0.1%	1.1%	-0.9%	0.3%	0.2%	1.1%	0.2%	0.0%
Property services	2.7%	0.2%	1.2%	0.0%	-0.4%	0.2%	-1.9%	-0.2%	-0.6%	0.5%	0.4%	-2.0%	0.5%	0.4%	1.4%	-0.3%	1.5%	1.0%	0.6%	-0.5%	0.1%
Business services	1.9%	-0.8%	0.1%	0.0%	-0.3%	1.0%	-0.1%	-0.1%	0.1%	0.7%	0.5%	-2.7%	0.1%	0.7%	0.1%	-1.0%	-0.3%	0.8%	0.8%	-0.1%	0.1%
Central government admin and defence	2.8%	0.3%	0.6%	0.0%	-0.3%	-1.1%	-1.6%	-0.2%	-0.2%	-1.2%	0.1%	-1.7%	0.3%	-1.0%	1.2%	-0.3%	-0.1%	-0.7%	0.5%	-0.4%	0.1%
Local government administration	2.3%	-0.4%	1.6%	0.0%	-0.2%	-0.4%	-0.8%	-0.3%	0.5%	-0.6%	0.3%	-4.2%	0.9%	-0.5%	0.6%	-1.4%	-1.4%	0.3%	2.2%	0.8%	0.1%
Education	2.1%	-0.6%	1.7%	0.0%	-0.3%	1.1%	1.0%	0.5%	0.5%	1.0%	0.1%	0.4%	1.8%	0.9%	-0.7%	-1.2%	-1.0%	2.7%	2.3%	2.9%	0.1%
Health and community services	2.3%	-0.3%	-0.7%	0.0%	-0.2%	-0.3%	0.1%	0.1%	-0.4%	-0.5%	0.2%	0.2%	-0.6%	-0.4%	0.0%	-0.1%	-0.6%	-1.1%	-0.4%	-0.5%	0.1%
Cultural and recreational services	2.8%	0.3%	-0.7%	0.0%	-0.3%	0.7%	2.9%	0.2%	0.3%	1.1%	-0.3%	-0.1%	-0.5%	1.0%	-2.2%	-0.6%	-1.8%	0.5%	0.4%	0.6%	0.0%
Personal and other community services	2.9%	0.5%	-2.1%	0.0%	-0.2%	-0.3%	1.5%	0.2%	-0.3%	0.0%	-0.2%	2.3%	-1.7%	0.0%	-1.1%	0.6%	-0.5%	-1.7%	-1.7%	-0.9%	0.0%
Total	2.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Summary statistics																					
Average	2.2%	-0.4%	-0.2%	0.0%	-0.1%	-0.2%	0.2%	-0.2%	0.5%	-0.5%	0.1%	0.1%	0.2%	-0.4%	-0.1%	0.2%	-0.4%	-0.3%	-0.2%	0.1%	0.1%
Root Mean Square Percent Deviation	2.3%	1.1%	1.3%	0.1%	0.5%	1.4%	2.7%	0.7%	2.4%	2.0%	0.8%	2.5%	1.1%	1.8%	1.9%	1.1%	2.8%	1.8%	1.1%	1.6%	0.2%
Largest positive impact	3.6%	1.4%	1.8%	0.5%	1.5%	1.5%	8.5%	1.7%	10.0%	1.2%	2.3%	5.2%	2.7%	1.1%	4.2%	2.5%	3.9%	2.7%	2.3%	6.1%	0.7%
Largest negative impact (or smallest)	0.2%	-3.4%	-3.2%	0.0%	-1.4%	-6.8%	-5.6%	-2.8%	-5.4%	-9.8%	-1.8%	-6.4%	-1.7%	-8.8%	-6.3%	-1.6%	-11.1%	-7.8%	-2.6%	-2.8%	-0.7%
Range of impacts	3.4%	4.8%	5.0%	0.5%	2.9%	8.2%	14.1%	4.5%	15.3%	11.0%	4.0%	11.6%	4.4%	9.9%	10.5%	4.1%	14.9%	10.5%	4.9%	8.8%	1.4%

Table 7

Accumulated percentage divergence in real GDP resulting from postive one standard deviation impulse in exogenous variables in March 2011

Impulse variables

Industry	GDP	QH	CoE	RW	PR	UR	ER	BoP	ToT	TDE	BusI	Posr	TWI	90day	NZbond	Wbond	IR	MCI	GCR	GIR	SMD
Agriculture	2.3%	0.9%	0.3%	0.0%	0.0%	0.1%	0.0%	0.3%	-1.4%	0.0%	0.0%	-1.3%	0.2%	0.0%	0.0%	-0.5%	-0.5%	0.2%	-0.5%	-0.5%	-0.8%
Fishing	0.6%	-3.5%	-2.0%	0.0%	0.0%	1.3%	0.3%	-2.6%	10.1%	0.0%	0.6%	0.0%	-0.5%	0.0%	0.0%	-0.1%	-0.1%	-0.5%	0.1%	0.1%	0.9%
Forestry and logging	2.2%	-0.3%	-2.4%	0.0%	0.0%	0.1%	0.1%	1.1%	-4.0%	0.0%	0.1%	-2.9%	-2.4%	0.0%	0.0%	-1.1%	-1.1%	-2.4%	-0.3%	-0.3%	0.0%
Mining	1.9%	-1.0%	-1.1%	0.0%	0.0%	0.4%	0.0%	-0.6%	2.4%	0.0%	0.2%	-3.9%	-1.1%	0.0%	0.0%	-1.6%	-1.6%	-1.1%	1.2%	1.2%	0.3%
Food, beverage and tobacco manufacturing	2.2%	0.2%	0.3%	0.0%	0.0%	0.2%	0.0%	0.8%	-3.2%	0.0%	0.1%	0.7%	0.5%	0.0%	0.0%	0.3%	0.3%	0.5%	-0.6%	-0.6%	-0.3%
Textile and apparel manufacturing	0.1%	-3.3%	0.4%	0.0%	0.0%	1.6%	0.0%	-0.5%	2.0%	0.0%	0.7%	1.9%	2.4%	0.0%	0.0%	0.8%	0.8%	2.4%	0.2%	0.2%	0.4%
Wood and paper product manufacturing	2.1%	-0.6%	-0.8%	0.0%	0.0%	0.2%	0.0%	-0.2%	0.8%	0.0%	0.1%	1.5%	-0.3%	0.0%	0.0%	0.6%	0.6%	-0.3%	-0.9%	-0.9%	0.1%
Printing, publishing and recorded media	1.3%	-1.5%	0.7%	0.0%	0.0%	0.8%	0.0%	-0.2%	1.0%	0.0%	0.3%	2.9%	1.9%	0.0%	0.0%	1.1%	1.1%	1.9%	-0.4%	-0.4%	0.1%
Petroleum, chemical, plastic and rubber products	1.3%	-1.4%	-1.5%	0.0%	0.0%	0.8%	0.0%	-0.3%	1.3%	0.0%	0.4%	3.7%	-0.1%	0.0%	0.0%	1.5%	1.5%	-0.1%	-2.2%	-2.2%	0.1%
Non-metallic mineral products manufacturing	2.4%	-0.2%	0.6%	0.0%	0.0%	0.0%	0.0%	-0.1%	0.3%	0.0%	0.0%	2.4%	0.9%	0.0%	0.0%	1.0%	1.0%	0.9%	-0.3%	-0.3%	0.1%
Metal product manufacturing	2.2%	-0.5%	-0.9%	0.0%	0.0%	0.2%	0.0%	-0.2%	0.7%	0.0%	0.1%	6.4%	0.2%	0.0%	0.0%	2.5%	2.5%	0.2%	-2.6%	-2.6%	0.2%
Machinery and equipment manufacturing	2.0%	-0.8%	0.7%	0.0%	0.0%	0.3%	0.0%	-0.2%	1.0%	0.0%	0.1%	4.3%	1.5%	0.0%	0.0%	1.7%	1.7%	1.5%	-0.8%	-0.8%	0.2%
Furniture and other manufacturing	1.5%	-1.3%	-0.4%	0.0%	0.0%	0.6%	0.0%	-0.2%	0.9%	0.0%	0.3%	5.4%	1.1%	0.0%	0.0%	2.1%	2.1%	1.1%	-2.0%	-2.0%	0.1%
Electricity, gas and water supply	1.8%	-1.0%	-0.2%	0.0%	0.0%	0.4%	0.0%	-0.4%	1.4%	0.0%	0.2%	-1.3%	0.2%	0.0%	0.0%	-0.5%	-0.5%	0.2%	0.7%	0.7%	0.2%
Construction	2.6%	0.0%	1.6%	0.0%	0.0%	-0.1%	0.0%	-0.1%	0.3%	0.0%	0.0%	3.6%	1.8%	0.0%	0.0%	1.4%	1.4%	1.8%	-0.1%	-0.1%	0.1%
Wholesale trade	2.4%	-0.1%	-0.4%	0.0%	0.0%	0.0%	0.0%	-0.2%	0.7%	0.0%	0.0%	1.7%	-0.1%	0.0%	0.0%	0.7%	0.7%	-0.1%	-0.8%	-0.8%	0.1%
Retail trade	2.5%	0.2%	0.5%	0.0%	0.0%	-0.1%	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%	0.4%	0.0%	0.0%	0.0%	0.0%	0.4%	0.2%	0.2%	-0.1%
Accommodation, restaurants and bars	2.3%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.2%	-0.8%	0.0%	0.0%	0.2%	0.1%	0.0%	0.0%	0.1%	0.1%	0.1%	-0.2%	-0.2%	-0.1%
Transport and storage	2.6%	0.3%	-1.0%	0.0%	0.0%	-0.1%	0.0%	0.1%	-0.6%	0.0%	-0.1%	0.6%	-1.0%	0.0%	0.0%	0.3%	0.3%	-1.0%	-1.0%	-1.0%	-0.1%
Communication services	3.5%	1.3%	-0.6%	0.0%	0.0%	-0.8%	0.0%	0.0%	0.0%	0.0%	-0.3%	-0.2%	-1.4%	0.0%	0.0%	-0.1%	-0.1%	-1.4%	-0.4%	-0.4%	0.0%
Finance and insurance	2.8%	0.3%	0.4%	0.0%	0.0%	-0.2%	0.0%	-0.1%	0.3%	0.0%	-0.1%	-2.1%	-0.1%	0.0%	0.0%	-0.8%	-0.8%	-0.1%	1.1%	1.1%	0.1%
Property services	2.6%	0.1%	0.4%	0.0%	0.0%	-0.1%	0.0%	0.1%	-0.2%	0.0%	-0.1%	-0.7%	0.1%	0.0%	0.0%	-0.3%	-0.3%	0.1%	0.6%	0.6%	0.1%
Business services	1.8%	-0.9%	-0.4%	0.0%	0.0%	0.4%	0.0%	-0.1%	0.3%	0.0%	0.2%	-2.4%	-0.2%	0.0%	0.0%	-1.0%	-1.0%	-0.2%	0.9%	0.9%	0.1%
Central government admin and defence	2.7%	0.2%	0.3%	0.0%	0.0%	-0.2%	0.0%	0.0%	0.0%	0.0%	-0.1%	-0.6%	0.0%	0.0%	0.0%	-0.2%	-0.2%	0.0%	0.5%	0.5%	0.1%
Local government administration	2.3%	-0.4%	1.1%	0.0%	0.0%	0.1%	0.0%	-0.2%	0.8%	0.0%	0.1%	-3.4%	0.7%	0.0%	0.0%	-1.4%	-1.4%	0.7%	2.2%	2.2%	0.2%
Education	2.0%	-0.7%	1.7%	0.0%	0.0%	0.3%	0.0%	-0.1%	0.5%	0.0%	0.1%	-3.0%	1.5%	0.0%	0.0%	-1.2%	-1.2%	1.5%	2.4%	2.4%	0.1%
Health and community services	2.2%	-0.4%	-1.0%	0.0%	0.0%	0.2%	0.0%	0.1%	-0.2%	0.0%	0.1%	-0.3%	-0.8%	0.0%	0.0%	-0.1%	-0.1%	-0.8%	-0.4%	-0.4%	0.1%
Cultural and recreational services	2.7%	0.2%	-0.4%	0.0%	0.0%	-0.2%	0.0%	0.0%	0.2%	0.0%	-0.1%	-1.4%	-0.7%	0.0%	0.0%	-0.6%	-0.6%	-0.7%	0.4%	0.4%	0.1%
Personal and other community services	2.8%	0.5%	-1.9%	0.0%	0.0%	-0.3%	0.0%	0.1%	-0.4%	0.0%	-0.1%	1.5%	-1.9%	0.0%	0.0%	0.6%	0.6%	-1.9%	-1.7%	-1.7%	0.0%
Ownership of owner-occupied dwellings	3.1%	0.8%	-1.3%	0.0%	0.0%	-0.5%	0.0%	0.0%	0.1%	0.0%	-0.2%	1.6%	-1.5%	0.0%	0.0%	0.6%	0.6%	-1.5%	-1.4%	-1.4%	0.0%
Unallocated	1.8%	-0.6%	0.5%	0.0%	0.0%	0.4%	0.0%	0.0%	-0.1%	0.0%	0.2%	5.2%	1.6%	0.0%	0.0%	2.1%	2.1%	1.6%	-1.7%	-1.7%	-0.1%
Total	2.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Summary statistics																					
Average	2.2%	-0.4%	-0.2%	0.0%	0.0%	0.2%	0.0%	-0.1%	0.5%	0.0%	0.1%	0.6%	0.1%	0.0%	0.0%	0.3%	0.3%	0.1%	-0.3%	-0.3%	0.1%
Root Mean Square Percent Deviation	2.3%	1.1%	1.0%	0.0%	0.0%	0.5%	0.1%	0.6%	2.2%	0.0%	0.2%	2.7%	1.1%	0.0%	0.0%	1.1%	1.1%	1.1%	1.2%	1.2%	0.2%
Largest positive impact	3.5%	1.3%	1.7%	0.0%	0.0%	1.6%	0.3%	1.1%	10.1%	0.0%	0.7%	6.4%	2.4%	0.0%	0.0%	2.5%	2.5%	2.4%	2.4%	2.4%	0.9%
Largest negative impact (or smallest)	0.1%	-3.5%	-2.4%	0.0%	0.0%	-0.8%	0.0%	-2.6%	-4.0%	0.0%	-0.3%	-3.9%	-2.4%	0.0%	0.0%	-1.6%	-1.6%	-2.4%	-2.6%	-2.6%	-0.8%
Range of impacts	3.4%	4.8%	4.0%	0.0%	0.0%	2.3%	0.3%	3.7%	14.1%	0.0%	1.0%	10.3%	4.8%	0.0%	0.0%	4.1%	4.1%	4.8%	4.9%	4.9%	1.6%

References

Ball P (2004) *Critical Mass*, Arrow Books: Reading

Doornik J A anf Hendry D f (2007) *PcGive 12 Empirical Econometric Modelling,* Timberlake Consultants Ltd

Hendry D F and Clements M P (2001) *Economic Forecasting: Some Lessons from Recent Research*, European Central Bank Working Paper No 82

Hendry D F (2011) *Empirical Economic Model Discovery and Theory Evaluation*, University of Oxford, Department of Economics, Discussion paper Number 529

Appendix 1: Regression Estimation Results

Industry production share principal component 1 (npc1)

Level equation

The estimation sample is: 1990(3) - 2010(1)

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-0.24083	0.00488	-49.3	0	0.9701
LQHSA	-0.24537	0.005578	-44	0	0.9627
LPOSRSA_2	-0.01924	0.00362	-5.31	0	0.2735
dmar94	0.014747	0.003102	4.75	0	0.2316

sigma	0.00307408	RSS	0.00070874634
R^2	0.976955	F(3,75) =	1060 [0.000]**
log-likelihood	346.952	DW	1.75
no. of observations	79	no. of parameters	4
mean(NPC1SA)	-0.00277589	var(NPC1SA)	0.000389309

AR 1-5 test: F(5,70) = 0.33608 [0.8894]
ARCH 1-4 test: F(4,67) = 0.27982 [0.8901]
Normality test: Chi^2(2) = 2.5765 [0.2758]
Hetero test: F(5,69) = 0.60645 [0.6952]
Hetero-X test: F(6,68) = 0.56489 [0.7568]
RESET test: F(1,74) = 0.10294 [0.7492]

Cointegration test: -4.5159 (Critical value -3.81)

Change equation

The estimation sample is: 1991(2) - 2010(1)

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-0.00055	9.73E-05	-5.6	0	0.3159
z1npc1_1	0.056989	0.03055	1.87	0.0665	0.0487
DLQHSA_1	-0.0177	0.005115	-3.46	0.0009	0.1497
DLBUSISA_3	0.003285	0.00159	2.07	0.0427	0.059
DLEMPSA_4	-0.03949	0.01243	-3.18	0.0022	0.1292
DLURSA_2	0.004393	0.00167	2.63	0.0105	0.0924
DLTWISA_2	0.008775	0.002257	3.89	0.0002	0.1819
ddec09	0.002683	0.00076	3.53	0.0008	0.1548

sigma	0.000659482	RSS	2.95743018e-005
R^2	0.534736	F(7,68) =	11.16 [0.000]**
log-likelihood	453.016	DW	1.88
no. of observations	76	no. of parameters	8
mean(DNPC1SA)	-0.000755402	var(DNPC1SA)	8.36375e-007

AR 1-5 test: F(5,63) = 0.29825 [0.9121]
ARCH 1-4 test: F(4,60) = 1.9299 [0.1171]
Normality test: Chi^2(2) = 1.7249 [0.4221]
Hetero test: F(13,54) = 0.31541 [0.9870]
Hetero-X test: F(28,39) = 0.76732 [0.7657]
RESET test: F(1,67) = 2.3796 [0.1276]

Industry production share principal component 2 (npc2)

Level equation

The estimation sample is: 1990(3) - 2010(1)

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	0.173894	0.04056	4.29	0.0001	0.2057
BOPSA_1	0.030362	0.01245	2.44	0.0172	0.0773
MCISA_2	4.95E-06	7.13E-07	6.95	0	0.4049
LQHSA	0.030585	0.007947	3.85	0.0003	0.1726
LTOTSA	-0.02842	0.004924	-5.77	0	0.3193
LGCRSA_2	-0.01983	0.004886	-4.06	0.0001	0.1884
LGIRSA_2	-0.00557	0.001607	-3.46	0.0009	0.1445
LSMDSA_3	-0.00095	0.000593	-1.6	0.1133	0.0349

sigma	0.00175167	RSS	0.000217852921
R^2	0.618052	F(7,71) =	16.41 [0.000]**
log-likelihood	393.549	DW	0.862
no. of observations	79	no. of parameters	8
mean(NPC2SA)	0.000574277	var(NPC2SA)	7.21991e-006

AR 1-5 test:	F(5,66)	=	10.674 [0.0000]**
ARCH 1-4 test:	F(4,63)	=	2.7068 [0.0380]*
Normality test:	Chi^2(2)	=	3.0973 [0.2125]
Hetero test:	F(14,56)	=	1.1170 [0.3641]
Hetero-X test:	F(35,35)	=	0.84036 [0.6951]
RESET test:	F(1,70)	=	0.030538 [0.8618]

Cointegration test: -4.6562 (Critical value -4.42)

Change equation

The estimation sample is: 1994(1) - 2010(1)

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-6.29E-05	0.000131	-0.481	0.632	0.0038
z2npc2_1	-0.27761	0.08675	-3.2	0.0022	0.1437
DMCISA_2	2.94E-06	7.90E-07	3.73	0.0004	0.1854
DLSMDSA_3	-0.00073	0.000271	-2.69	0.0092	0.1061

sigma	0.00105223	RSS	6.75379775e-005
R^2	0.330329	F(3,61) =	10.03 [0.000]**
log-likelihood	355.528	DW	1.62
no. of observations	65	no. of parameters	4
mean(DNPC2SA)	-7.37319e-005	var(DNPC2SA)	1.55158e-006

AR 1-5 test:	F(5,56)	=	1.1185 [0.3612]
ARCH 1-4 test:	F(4,53)	=	2.4512 [0.0572]
Normality test:	Chi^2(2)	=	0.22020 [0.8957]
Hetero test:	F(6,54)	=	0.52533 [0.7866]
Hetero-X test:	F(9,51)	=	0.57879 [0.8082]
RESET test:	F(1,60)	=	0.25597 [0.6148]

Industry production share principal component 3 (npc3)

Level equation

The estimation sample is: 1991(1) - 2010(1)

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-0.04152	0.02051	-2.02	0.0467	0.0546
LTWISA_4	0.005744	0.002991	1.92	0.0588	0.0494
LWBSA_1	0.010199	0.001197	8.52	0	0.5056
LPOSRSA	0.049342	0.004754	10.4	0	0.6027
LGCRSA_3	-4.24E-02	9.21E-03	-4.61	0	0.2301
dsep92	0.006285	0.002525	2.49	0.0151	0.0803

sigma 0.00231746 RSS 0.000381313061
R^2 0.864385 F(5,71) = 90.51 [0.000]**
log-likelihood 361.046 DW 1.03
no. of observations 77 no. of parameters 6
mean(NPC3SA) -0.000411679 var(NPC3SA) 3.65159e-005

AR 1-5 test: F(5,66) = 5.3444 [0.0003]**
ARCH 1-4 test: F(4,63) = 1.8648 [0.1277]
Normality test: Chi^2(2) = 4.2386 [0.1201]
Hetero test: F(9,61) = 1.7554 [0.0958]
Hetero-X test: F(15,55) = 1.8480 [0.0506]
RESET test: F(1,70) = 8.3091 [0.0052]**

Cointegration test: -4.1877 (Critical value -4.13)

Change equation

The estimation sample is: 1991(2) - 2010(1)

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-6.79E-05	0.000142	-0.478	0.6343	0.0032
z1npc3_1	-0.2369	0.06405	-3.7	0.0004	0.1615
DLTWISA_2	0.016079	0.003826	4.2	0.0001	0.1992
DLWBSA	0.007614	0.002061	3.69	0.0004	0.1613
DLPOSRSA	0.02106	0.004074	5.17	0	0.2735

sigma 0.00121324 RSS 0.000104508571
R^2 0.504658 F(4,71) = 18.08 [0.000]**
log-likelihood 405.046 DW 1.51
no. of observations 76 no. of parameters 5
mean(DNPC3SA) -0.00010184 var(DNPC3SA) 2.77609e-006

AR 1-5 test: F(5,66) = 1.5782 [0.1784]
ARCH 1-4 test: F(4,63) = 0.31725 [0.8654]
Normality test: Chi^2(2) = 3.9047 [0.1419]
Hetero test: F(8,62) = 0.94628 [0.4858]
Hetero-X test: F(14,56) = 0.90995 [0.5534]
RESET test: F(1,70) = 0.10595 [0.7458]

Industry production share principal component 4 (npc4)

Level equation

The estimation sample is: 1991(3) - 2010(1)

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-0.27465	0.03364	-8.16	0	0.4914
LCOESA_4	0.068722	0.009076	7.57	0	0.4538
LTWISA_2	0.016086	0.001918	8.39	0	0.5048
LGCRSA_1	0.02769	0.004224	6.56	0	0.3838
LGIRSA_4	0.003955	0.001107	3.57	0.0006	0.1561
Trend	-6.07E-05	1.88E-05	-3.24	0.0018	0.1319

sigma 0.00150913 RSS 0.00015714652
R^2 0.884577 F(5,69) = 105.8 [0.000]**
log-likelihood 383.923 DW 0.916
no. of observations 75 no. of parameters 6
mean(NPC4SA) -0.00152973 var(NPC4SA) 1.81531e-005

AR 1-5 test: F(5,64) = 6.1195 [0.0001]**
ARCH 1-4 test: F(4,61) = 4.1621 [0.0048]**
Normality test: Chi^2(2) = 0.083586 [0.9591]
Hetero test: F(10,58) = 2.7049 [0.0086]**
Hetero-X test: F(20,48) = 1.5412 [0.1107]
RESET test: F(1,68) = 0.017113 [0.8963]

Cointegration test: -4.6795 (Critical value -4.43)

Change equation

The estimation sample is: 1991(4) - 2010(1)

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-3.83E-05	0.000147	-0.261	0.795	0.001
z1npc4_1	-0.43215	0.09951	-4.34	0	0.2171
DLCOESA_4	0.060835	0.01916	3.18	0.0023	0.1291
DLTWISA_2	0.010688	0.003874	2.76	0.0074	0.1007
DLGCRSA_1	0.02096	0.004412	4.75	0	0.2492
DLGIRSA_4	0.003012	0.001468	2.05	0.044	0.0583

sigma 0.00123251 RSS 0.000103298001
R^2 0.465188 F(5,68) = 11.83 [0.000]**
log-likelihood 393.831 DW 1.96
no. of observations 74 no. of parameters 6
mean(DNPC4SA) 5.9838e-005 var(DNPC4SA) 2.61011e-006

AR 1-5 test: F(5,63) = 0.90857 [0.4813]
ARCH 1-4 test: F(4,60) = 0.71433 [0.5854]
Normality test: Chi^2(2) = 1.7532 [0.4162]
Hetero test: F(10,57) = 0.62964 [0.7823]
Hetero-X test: F(20,47) = 0.43218 [0.9780]
RESET test: F(1,67) = 0.70604 [0.4037]

Industry production share principal component 5 (npc5)

Level equation

The estimation sample is: 1991(1) - 2010(1)

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-0.18503	0.02955	-6.26	0	0.3526
lev2002	-0.00817	0.000777	-10.5	0	0.6056
BOPSA_4	-0.03835	0.0104	-3.69	0.0004	0.159
LTOTSA_2	0.029375	0.004344	6.76	0	0.3884
LGIRSA_4	0.005428	0.001193	4.55	0	0.2234

sigma 0.00168638 RSS 0.000204759041
R^2 0.671287 F(4,72) = 36.76 [0.000]**
log-likelihood 384.985 DW 0.836
no. of observations 77 no. of parameters 5
mean(NPC5SA) -1.50854e-005 var(NPC5SA) 8.08976e-006

AR 1-5 test: F(5,67) = 6.9978 [0.0000]**
ARCH 1-4 test: F(4,64) = 0.83770 [0.5063]
Normality test: Chi^2(2) = 4.5565 [0.1025]
Hetero test: F(7,64) = 0.31493 [0.9447]
Hetero-X test: F(13,58) = 0.35538 [0.9782]
RESET test: F(1,71) = 3.0682 [0.0842]

Cointegration test: -4.6344 (Critical value -4.13)

Change equation

The estimation sample is: 1991(2) - 2010(1)

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	4.25E-06	0.000124	0.0343	0.9727	0
z2npc5_1	-2.53E-01	0.07755	-3.26	0.0017	0.1302
DLPOSRSA_2	0.010935	0.003363	3.25	0.0018	0.1296
DLERSA_1	0.053729	0.01912	2.81	0.0064	0.1001
dsep02	-0.00774	0.001061	-7.29	0	0.4284

sigma 0.00105212 RSS 7.8594034e-005
R^2 0.547424 F(4,71) = 21.47 [0.000]**
log-likelihood 415.875 DW 1.84
no. of observations 76 no. of parameters 5
mean(DNPC5SA) -2.07419e-005 var(DNPC5SA) 2.28499e-006

AR 1-5 test: F(5,66) = 0.34022 [0.8867]
ARCH 1-4 test: F(4,63) = 0.56924 [0.6859]
Normality test: Chi^2(2) = 2.7565 [0.2520]
Hetero test: F(7,63) = 0.44148 [0.8723]
Hetero-X test: F(10,60) = 0.45833 [0.9100]
RESET test: F(1,70) = 0.31997 [0.5734]

Industry output per employee principal component 1 (lpc1)

Level equation

The estimation sample is: 1999(2) - 2010(1)

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	0.295316	0.01981	14.9	0	0.8443
LNZBSA	-0.04137	0.0138	-3	0.0046	0.1798
LERSA_4	0.48379	0.03444	14	0	0.828

sigma	0.00580789	RSS	0.00138299544
R^2	0.90968	F(2,41) =	206.5 [0.000]**
log-likelihood	165.656	DW	0.932
no. of observations	44	no. of parameters	3
mean(lpc1sa)	2.27273e-008	var(lpc1sa)	0.000348003

AR 1-3 test: F(3,38) = 4.7451 [0.0066]**
ARCH 1-3 test: F(3,35) = 0.61585 [0.6093]
Normality test: Chi^2(2) = 0.082070 [0.9598]
Hetero test: F(4,36) = 1.3043 [0.2869]
Hetero-X test: F(5,35) = 1.0178 [0.4221]
RESET test: F(1,40) = 0.89790 [0.3490]

Cointegration test: -3.5452 (Critical value -3.45)

Change equation

The estimation sample is: 1999(3) - 2010(1)

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	0.000896	0.000365	2.45	0.0191	0.1433
z1lpc1_1	-0.24121	0.06694	-3.6	0.0009	0.2651
DBOPSA_2	-9.73E-02	0.02375	-4.1	0.0002	0.3178
DLRWQSA	0.119876	0.032	3.75	0.0006	0.2805
DLTDESA	1.89E-01	3.01E-02	6.29	0	0.5238
dsep07	-0.01248	0.002334	-5.35	0	0.4425
djun09	0.010087	0.002686	3.75	0.0006	0.2814

sigma	0.0022839	RSS	0.000187782477
R^2	0.77143	F(6,36) =	20.25 [0.000]**
log-likelihood	204.326	DW	2.24
no. of observations	43	no. of parameters	7
mean(Dlpc1sa)	0.0011667	var(Dlpc1sa)	1.91059e-005

AR 1-3 test: F(3,33) = 0.79772 [0.5040]
ARCH 1-3 test: F(3,30) = 0.67273 [0.5755]
Normality test: Chi^2(2) = 1.2354 [0.5392]
Hetero test: F(10,25) = 2.3884 [0.0377]*
Hetero-X test: not enough observations
RESET test: F(1,35) = 0.019877 [0.8887]

Industry output per employee principal component 2 (lpc2)

Level equation

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	0.523329	0.103	5.08	0	0.3981
dtui07	0.015892	0.002694	5.9	0	0.4716
LBUSISA	-0.02451	0.00935	-2.62	0.0124	0.1498
LCOESA_1	-0.17754	0.02207	-8.05	0	0.624
LTOTSA_1	3.61E-02	1.78E-02	2.03	0.0495	0.0954

sigma 0.00459422 RSS 0.000823167331
R^2 0.830843 F(4,39) = 47.89 [0.000]**
log-likelihood 177.071 DW 1.25
no. of observations 44 no. of parameters 5
mean(lpc2sa) -5.22727e-008 var(lpc2sa) 0.000110597

AR 1-3 test: F(3,36) = 2.5907 [0.0678]
ARCH 1-3 test: F(3,33) = 0.42568 [0.7359]
Normality test: Chi^2(2) = 0.46971 [0.7907]
Hetero test: F(7,31) = 0.68743 [0.6817]
Hetero-X test: F(13,25) = 0.74606 [0.7041]
RESET test: F(1,38) = 14.240 [0.0005]**

Cointegration test: -4.1947 (Critical value -4.13)

Change equation

The estimation sample is: 1999(3) - 2010(1)

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	0.001352	0.000643	2.1	0.042	0.1018
z3lpc2_1	-0.35696	0.145	-2.46	0.0184	0.1345
DLQHSA	-0.09644	0.03941	-2.45	0.019	0.1331
DLCOESA_1	-3.27E-01	0.07427	-4.4	0.0001	0.3315

sigma 0.00390743 RSS 0.000595452438
R^2 0.443959 F(3,39) = 10.38 [0.000]**
log-likelihood 179.515 DW 1.64
no. of observations 43 no. of parameters 4
mean(Dlpc2sa) 0.000169653 var(Dlpc2sa) 2.49041e-005

AR 1-3 test: F(3,36) = 0.57990 [0.6320]
ARCH 1-3 test: F(3,33) = 0.29776 [0.8267]
Normality test: Chi^2(2) = 1.6645 [0.4351]
Hetero test: F(6,32) = 0.92343 [0.4913]
Hetero-X test: F(9,29) = 0.56917 [0.8108]
RESET test: F(1,38) = 2.1785 [0.1482]

Industry output per employee principal component 3 (lpc3)

Level equation

The estimation sample is: 1999(2) - 2010(1)

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	1.01E-01	2.41E-02	4.2	0.0001	0.3117
dtui07	0.011441	0.001926	5.94	0	0.475
BOPSA	0.064593	0.02069	3.12	0.0034	0.2
LPOSRSA	0.050681	0.01295	3.91	0.0004	0.282
LGIRSA_1	0.009659	0.003333	2.9	0.0061	0.1772

sigma 0.00339549 RSS 0.000449645576
R^2 0.600787 F(4,39) = 14.67 [0.000]**
log-likelihood 190.374 DW 1.13
no. of observations 44 no. of parameters 5
mean(lpc3sa) 1.13636e-008 var(lpc3sa) 2.55984e-005

AR 1-3 test: F(3,36) = 3.7310 [0.0196]*
ARCH 1-3 test: F(3,33) = 0.38444 [0.7649]
Normality test: Chi^2(2) = 0.64954 [0.7227]
Hetero test: F(7,31) = 1.0220 [0.4357]
Hetero-X test: F(13,25) = 1.1048 [0.3991]
RESET test: F(1,38) = 0.23068 [0.6338]
z1lpc3 [1999(2) - 2010(1)] saved to pc sa ec data.xls

Cointegration test: -4.3294 (Critical value -4.13)

Change equation

The estimation sample is: 1999(3) - 2010(1)

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	4.68E-04	3.70E-04	1.26	0.2143	0.0414
z1lpc3_1	-0.40246	0.1224	-3.29	0.0022	0.2262
DBOPSA_2	0.07612	0.02379	3.2	0.0028	0.2167
DLNZBSA_1	0.022772	0.007092	3.21	0.0027	0.2179
DLTDESA	-9.44E-02	0.03053	-3.09	0.0038	0.2052
DLGIRSA_3	-0.00884	0.003813	-2.32	0.0261	0.1267

sigma 0.00241519 RSS 0.000215826493
R^2 0.580408 F(5,37) = 10.24 [0.000]**
log-likelihood 201.334 DW 1.51
no. of observations 43 no. of parameters 6
mean(Dlpc3sa) 0.000407517 var(Dlpc3sa) 1.19621e-005

AR 1-3 test: F(3,34) = 2.0524 [0.1250]
ARCH 1-3 test: F(3,31) = 0.17339 [0.9136]
Normality test: Chi^2(2) = 1.1782 [0.5548]
Hetero test: F(10,26) = 0.62537 [0.7787]
Hetero-X test: F(20,16) = 0.60138 [0.8598]
RESET test: F(1,36) = 0.28609 [0.5960]

Industry output per employee principal component 4 (lpc4)

Level equation

The estimation sample is: 1999(2) - 2010(1)

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-0.21021	0.07415	-2.84	0.0071	0.1639
LPRSA_3	0.057983	0.01947	2.98	0.0049	0.1778
LTWISA_2	-0.00816	0.003056	-2.67	0.0108	0.1481

sigma	0.00189164	RSS	0.000146710841
R^2	0.191979	F(2,41) =	4.871 [0.013]*
log-likelihood	215.014	DW	1.45
no. of observations	44	no. of parameters	3
mean(lpc4sa)	1.31818e-009	var(lpc4sa)	4.12655e-006

AR 1-3 test: F(3,38) = 2.8478 [0.0502]
ARCH 1-3 test: F(3,35) = 0.64701 [0.5901]
Normality test: Chi^2(2) = 1.4306 [0.4890]
Hetero test: F(4,36) = 0.62465 [0.6479]
Hetero-X test: F(5,35) = 0.79571 [0.5602]
RESET test: F(1,40) = 1.4024 [0.2433]

Cointegration test: -3.0846 (Critical value -3.45)

Change equation

The estimation sample is: 1999(3) - 2010(1)

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-3.92E-05	0.000261	-0.15	0.8815	0.0006
z1lpc4_1	-0.57564	0.1435	-4.01	0.0003	0.2974
DLSMDSA_2	-1.28E-03	0.000588	-2.18	0.0353	0.1114
DLPRSA_4	1.03E-01	5.01E-02	2.06	0.046	0.1007
DLURSA_1	0.009066	0.004459	2.03	0.0491	0.0981

sigma	0.00166994	RSS	0.000105970972
R^2	0.500616	F(4,38) =	9.523 [0.000]**
log-likelihood	216.627	DW	1.82
no. of observations	43	no. of parameters	5
mean(Dlpc4sa)	6.56798e-005	var(Dlpc4sa)	4.93496e-006

AR 1-3 test: F(3,35) = 0.60612 [0.6155]
ARCH 1-3 test: F(3,32) = 0.90042 [0.4517]
Normality test: Chi^2(2) = 0.11606 [0.9436]
Hetero test: F(8,29) = 0.56177 [0.8000]
Hetero-X test: F(14,23) = 1.1200 [0.3923]
RESET test: F(1,37) = 0.12602 [0.7246]

Industry output per employee principal component 5 (lpc5)

Level equation

The estimation sample is: 1999(2) - 2010(1)

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-0.02662	0.006509	-4.09	0.0002	0.2949
LURSA_1	0.00863	0.0027	3.2	0.0027	0.2034
L90DSA_1	0.006584	0.001555	4.23	0.0001	0.3094
LTDESA	0.070425	0.01636	4.3	0.0001	0.3165

sigma	0.00202789	RSS	0.000164492947
R^2	0.512438	F(3,40) =	14.01 [0.000]**
log-likelihood	212.497	DW	1.46
no. of observations	44	no. of parameters	4
mean(lpc5sa)	-3.34091e-009	var(lpc5sa)	7.66769e-006

AR 1-3 test: F(3,37) = 1.2550 [0.3039]
ARCH 1-3 test: F(3,34) = 2.3134 [0.0934]
Normality test: Chi^2(2) = 1.3919 [0.4986]
Hetero test: F(6,33) = 0.33609 [0.9129]
Hetero-X test: F(9,30) = 0.70426 [0.7001]
RESET test: F(1,39) = 0.22439 [0.6384]

Cointegration test: -4.7704 (Critical value -3.81)

Change equation

The estimation sample is: 1999(3) - 2010(1)

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-0.0006	0.000279	-2.15	0.0378	0.1114
z1lpc5_1	-0.63923	0.134	-4.77	0	0.381
DLBUSISA_4	-1.21E-02	0.00524	-2.31	0.0263	0.1264
DLRWYSA	1.25E-01	4.03E-02	3.09	0.0038	0.2055
DLERSA	0.109733	0.04113	2.67	0.0113	0.1613
DLGCRSA	2.39E-02	0.008021	2.98	0.0051	0.1934

sigma	0.00163118	RSS	9.84480765e-005
R^2	0.579312	F(5,37) =	10.19 [0.000]**
log-likelihood	218.21	DW	1.88
no. of observations	43	no. of parameters	6
mean(Dlpc5sa)	-0.000188021	var(Dlpc5sa)	5.44225e-006

AR 1-3 test: F(3,34) = 0.11012 [0.9536]
ARCH 1-3 test: F(3,31) = 0.068807 [0.9761]
Normality test: Chi^2(2) = 1.9375 [0.3795]
Hetero test: F(10,26) = 0.71058 [0.7068]
Hetero-X test: F(20,16) = 0.93010 [0.5669]
RESET test: F(1,36) = 1.6957 [0.2011]

All regressions undertaken using PcGive 12. The variables used are listed in the table below. All data is seasonally adjusted (denoted by the suffix SA) and expressed in natural logarithms (denoted by prefix L) except for the balance of payments, which is expressed as percent of GDP and has both positive and negative values. The prefix DL denotes log changes ($ln(x_t)$-$ln(X_{t-1})$).

		Variable present in explanatory equations for:									
		GDP					Output per employee				
Code	**Description**	**NPC1**	**NPC2**	**NPC3**	**NPC4**	**NPC5**	**LPC1**	**LPC2**	**LPC3**	**LPC4**	**LPC5**
	Labour market										
QH	Productivity (real output/hr)	*	*					*			
CoE	Comp of Employess % tot income				*			*			
RWY	Real wage (income)										*
RWQ	Real wage (product)						*				
PR	Labour participation rate									*	
Emp	Employment	*									
UR	Unemployment rate	*								*	*
ER	Employment Rate					*	*				*
	External sector										
BoP	BOP current account balance/GDP		*			*	*		*		
ToT	Terms of trade		*			*		*			
TDE	Total domestic expenditure as % of GDP						*		*		*
	Business sector										
BusI	Bus investment as % GDP	*						*			*
posr	Private operating surplus as % of GDP	*		*		*			*		
	Financial										
TWI	Trade weighted index of New Zealand exchange rate	*		*	*					*	
90day	90-day bank bill rate										*
NZbond	10-year government bond rate, NZ						*		*		
Wbond	10-year government bond rate, World			*							
MCI	Monetary conditions index (calculated using 3xTWI:1x90 day)		*								
	Government sector										
GCR	Government consumption as % of expenditure on GDP		*	*	*						*
GIR	Government investment as % of expenditure on GDP		*		*	*			*		
	Climate										
SMD	Soil moisture deficit		*							*	

Appendix 2: Time Component Forecast Equations

Industry share of output time component forecasting equations

$$\begin{aligned} npc1 = {} & 0.0132 - 1.0570npc1_{-1} + 0.0140LQHSA_{-1} + 0.0011LPOSRSA_{-3} \\ & - 0.0177DLQHSA_{-1} + 0.0033DLBUSISA_{-3} - 0.0395DLEMPSA_{-4} \\ & + 0.0044DLURSA_{-2} + 0.0088DLTWISA_{-2} \end{aligned}$$

$$\begin{aligned} npc2 = {} & 0.0048 + 0.7224npc2_{-1} + 0.0084BOPSA_{-2} + 0.0000014MCISA_{-3} \\ & + 0.0085LQHSA_{-1} - 0.0079LTOTSA_{-1} - 0.0055LGCRSA_{-3} \\ & - 0.0015LGIRSA_{-3} - 0.0003LSMDSA_{-4} + 0.000003DMCISA_{-2} \\ & - 0.0007DLSMDSA_{-3} \end{aligned}$$

$$\begin{aligned} npc3 = {} & -0.0099 + 0.7631npc3_{-1} + 0.0014LTWISA_{-5} + 0.0024LWBSA_{-2} \\ & + 0.0117LPOSRSA_{-1} - 0.0101LGCRSA_{-4} + 0.0161DLTWISA_{-2} \\ & + 0.0076DLLWBSA + 0.0211DLPOSRSA \end{aligned}$$

$$\begin{aligned} npc4 = {} & -0.0119 + 0.5678npc4_{-1} + 0.0297LCOESA_{-5} + 0.0070LTWISA_{-3} \\ & + 0.0120LGCRSA_{-2} + 0.0017LGIRSA_{-5} + 0.0608DLCOESA_{-4} \\ & + 0.0107DLTWISA_{-2} + 0.0210DLGCRSA_{-1} + 0.0030DLGIRSA_{-4} \end{aligned}$$

$$\begin{aligned} npc5 = {} & -0.0488 + 0.7472npc5_{-1} - 0.0097BOPSA_{-5} + 0.0074LTOTSA_{-3} \\ & + 0.0014LGIRSA_{-5} + 0.0109DLPOSRSA_{-2} + 0.0537DLERSA_{-1} \end{aligned}$$

Industry output per employee time component forecasting equations

$$\begin{aligned} lpc1 = {} & 0.0721 + 0.7588lpc1_{-1} - 0.0100LNZBSA_{-1} + 0.1167LERSA_{-5} \\ & - 0.0973DBOPSA_{-2} + 0.1199DLRWQSA + 0.1893DLTDESA \end{aligned}$$

$$\begin{aligned} lpc2 = {} & 0.1938 + 0.6430lpc2_{-1} - 0.0088LBUSISA_{-1} - 0.0634\mathrm{LCOESA}_{-2} \\ & + 0.0129\mathrm{LTOTSA}_{-2} - 0.0964\mathrm{DLQHSA} - 0.3267\mathrm{DLCOESA}_{-1} \end{aligned}$$

$$lpc3 = 0.0457 + 0.05975lpc3_{-1} + 0.0260BOPSA_{-1} + 0.0204LPOSRSA_{-1} + 0.0039LGIRSA_{-2} + 0.0761DBOPSA_{-2} + 0.0228DLNZBSA_{-1} - 0.0944DLTDESA - 0.0088DLGIRSA_{-3}$$

$$lpc4 = -0.121 + 0.4244lpc4_{-1} + 0.0334LPRSA_{-4} - 0.0047LTWISA_{-3} - 0.0013DLSMDSA_{-2} + 0.1034DLPRSA_{-4} + 0.0091DLURSA_{-1}$$

$$lpc5 = -0.0173 + 0.3608lpc5_{-1} + 0.0055LURSA_{-2} + 0.0042L90DSA_{-2} + 0.0450LTDESA_{-1} - 0.0121DLBUSISA_{-4} + 0.1246DLRWYSA + 0.1097DLERSA + 0.0239DLGCRSA$$