B Application

The following application illustrates the testing approach in a realistic setting. Swiss GDP data is calculated by the Swiss Federal Statistical Office (SFSO). This agency produces data at annual frequency. The latest available data very often is older than one year and substantial revision occur infrequently. High frequency data of Swiss GDP is provided by the ministry of economic affairs' State Secretariat for Economic Affairs (SECO). This data is derived from the annual aggregate by various indicator estimations and disaggregation procedures. It is important to notice that "genuine" high frequency data for Swiss GDP does not exist. Therefore, it appears worthwhile considering alternative options. The next lines exemplify how the new disaggregation test can be used for identifying potential alternatives.

The aim of the analysis will be framed as follows. Out of a set of potential related series and given the CL model we are interested in the most suitable indicator variable that will be used to facilitate disaggregation.[2] Suitability is measured in terms of testing the disaggregation restriction. If there is more than one appropriate model / indicator variable we also consider the forecast properties.

B.1 The regression set-up and the related variables

The disaggregate model (1) is employed with $X_h = (c, x_{h,t}^{(i)}, t)$ representing the vector of exogenous variables. Here, c is a constant term and t is a time trend. The variable $x_{h,t}^{(i)}, i = 1, \dots, 9$ represents the related series of interest. We choose x_i to be the *cumulated sum* of business tendency survey variables calculated with the semantic shock identification method (Müller-Kademann and Köberl, 2008).[3] Upon cumulating we probably obtain a degree-one integrated time series with which GDP cointegrates under the CL assumptions.[4]

[2] To keep things simple and due to the data poor environment we employ just one such indicator variable.

[3] We are grateful to Müller-Kademann and Köberl for providing us with their data.

[4] In general, formal tests (Johansen, 1988; Johansen, 1992) could be used to justify the cointegration approach.

More specifically, the data source is a business tendency survey in the manufacturing industry. The two survey questions of interest are related to the firms' capacity utilisation. One asks whether the firm's technical capacities are currently too high, just right, or too low. The other inquires the degree of the capacity utilisation within the past three months in percentage points, where the firms can choose from a range of 50% to 110% in five percentage steps. From the latter we can calculate the percentage change in production capacity from t to $t+1$ and compare this to the judgement about capacity utilisation given by the firm in the previous period, that is in t.

We thus obtain the following combinations of expectations (rows) and corresponding realisations (columns).

Table 4: The basic structure of the data

		realisation		
		+	=	−
judgement	+	pp	pe	pm
	=	ep	ee	em
	−	mp	me	mm

The rows describe the judgement of the firms in t about their current technical capacity; '+' stands for 'too high', '=' for just right, '–' for too low. The columns list the possible outcomes in capacity utilisation changes. A '+' means that the level of capacity utilisation has been augmented between t and $t+1$, a '=' stands for an unchanged level and '–' means a lower level. On the basis of theses expectation–realisation combinations we obtain the nine different qualities, $i = 1, \ldots 9$, of $x_{h,t}^{(i)}$.

The related variables $x_{h,t}^{(i)}$ are eventually obtained as the cumulated sums of the expectation–realisation combinations over time. Let denote by $\chi_{h,t}^{(i)}$ the expectation–realisation combination "pp" as of table 4 at t for which $i = 1$, say. Then $x_{h,t}^{(1)} = \sum_{j=1}^{t} \chi_{h,j}^{(1)}$. The cumulation will most likely result in a degree-one integrated variable which ensures a balanced equation in model (1) and opens the scope for cointegration. Cointegration corresponds to the assumption $|\rho| < 1$. Cointegration could also be tested formally. Without cointegration forecasts will be very bad which might serve as an informal check in practical applications (see table 5).

This particular data (without cumulation over time) has been shown to generate very good forecasts for GDP growth in a quarterly time series model (Müller and Köberl, 2012). In Müller and Köberl (2012) the empirical model builds, however, on the historical SECO high frequency data and thus blurs the value added by the survey data since the forecasts also depend on the quality of SECO's disaggregation and nowcasting capabilities. The following disaggregation exercise, by contrast, will tell whether or not the survey data can also be used for straightforward disaggregation independent of SECO's input.

B.2 Variable selection, estimation and testing

The independent variable is official Swiss real GDP data. Due to changes in the definition of industry classifications in 1998 the survey data is available in a consistent form for 1999 – 2012 only. The effective sample size is thus 1999 – 2010 as we set aside two observations for (pseudo) ex-ante forecasting. Hence, we are in a data poor environment with little statistical guidance available for choosing the best model. Traditionally, only forecast comparisons, coefficient tests and criteria for model fit might be used. On top of these, we are also able to test for the appropriateness of the disaggregation procedure.

For selecting the most suitable model we estimate the CL model, perform t-test on the related variable, calculate the R^2, mean squared forecast error (MSFE) and the p-value for the new likelihood ratio test statistic. In addition, we also give the estimates of ρ. Values of ρ not too close to one (smaller than 0.9 say) offer some comfort with respect to validity of the implicit covariance-stationarity and cointegration assumptions. The table 5 reports the results.

The estimation results tell a very plausible story. For example, for all "negative" surprises

Table 5: Disaggregation of annual Swiss GDP to quarterly values

$y_t = c + \beta_i x^{(i)}_{h,t} + t + u_t,$
$u_t = \rho u_{t-1} + \varepsilon_t$
$E(\varepsilon_t) = 0,\ Var(\varepsilon_t) = \sigma^2$
effective sample: 1999q1 – 2010q4
forecasts: 2011q1 – 2012q4

Related variable[a]	$\hat{\rho}$	$\hat{\beta}_i$	t-stat.	R^2	p-value[c]	MSFE
mm	0.79	44.53	0.51	0.92	4.18	1.25
me	0.79	102.5	2.01	0.94	3.80	0.92
mp	0.84	243.7	2.21	0.94	11.59	17780.00
em	0.87	-50.4	-3.24	0.95	6.20	0.63
ee	0.84	50.95	3.84	0.96	9.99	48.76
ep	0.86	60.1	3.75	0.96	5.08	48.47
surprise indicator[b]	0.7	-128.7	-6.61	0.98	2.90	1.07
pe	0.69	-105.3	-3.46	0.96	4.18	29.22
pp	0.79	-36.62	-1.41	0.93	3.69	24.99

[a] The related variables are calculated as percentages of firms which have expected an improvement ("p"), no change ("e"), or a decrease ("m") in their business' performance and experience an improvement ("p"), no change ("e"), or a decrease ("m"). Thus nine combinations of expectations and realisations are observed (Müller-Kademann and Köberl, 2008). All percentages enter the regression as cumulated sums.
[b] The surprise indicator is defined as the combination of positive expectations and negative outcome ("pm").
[c] The p-value (in percent) is given for the $\chi^2(1)$ test statistic on (11), see p. 14.

(em, pm, pe) the estimated β coefficient is also negative. In other words, the sequence of cumulated negative impulses reported by the surveyed firms is negatively associated with real GDP. There is one more negative β coefficient (pp) but it is insignificant as is the coefficient for mm which is also a category best described as shock adjustment. All estimated ρ are well below one indicating covariance stationarity. The model fit seems to be very satisfactory with no R^2 being smaller than 92%. However, since there are only eleven observations but three explanatory variables, the large R^2 should not come as a surprise.

When it comes to selecting a model for actual disaggregation, the model fit cannot serve as a way to discriminate between related series. The t-values can be used as arguments for deselecting mm and pp. The t-value for me is just marginally larger than 2, so further information seems desirable. The comparison of forecasting capabilities easily reduces the set of potential related variables further down to me, em and the surprise indicator. Among these three the surprise indicator generates the worst forecasts while at the same time featuring the largest R^2 and

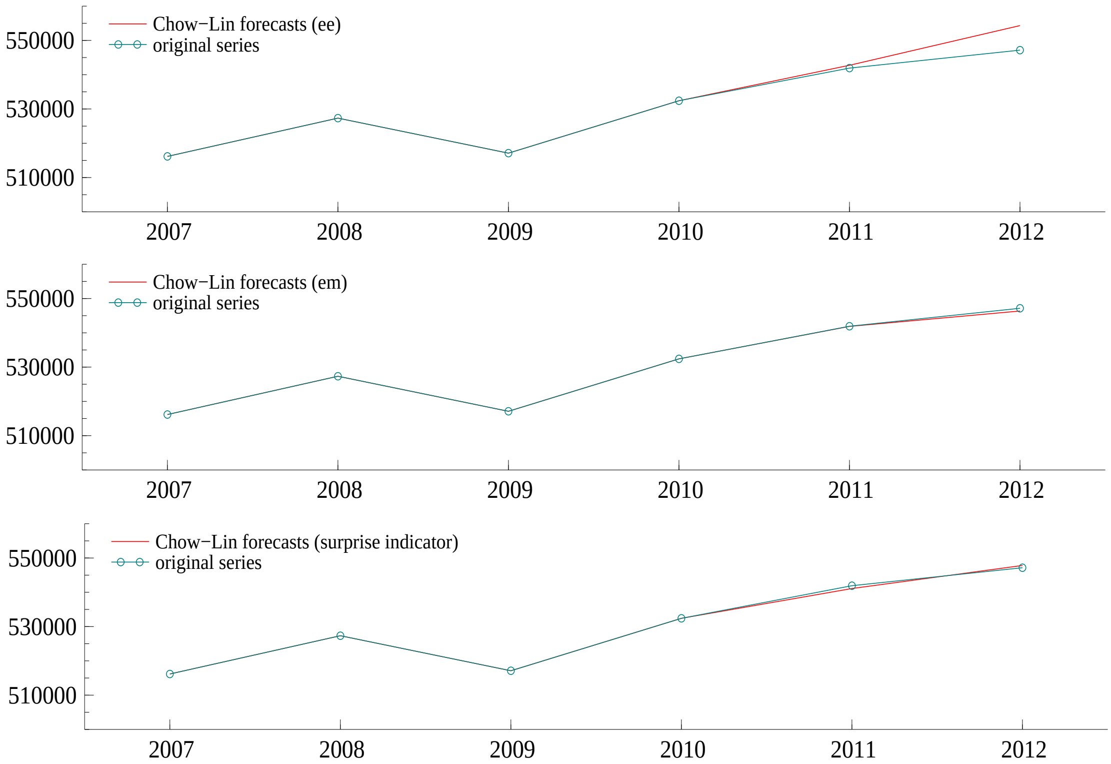
t-value on β. The variable em does best in terms of forecasting while me is somewhere in between.

The final decision can now be made on the basis of the disaggregation test. Of these three variables only em results in a CL model that is externally (forecasts, model fit, t-value) as well as internally valid. The latter property can be inferred from accepting the null hypothesis of consistent aggregate and disaggregate data generating processes. The test statistic implies rejection for me and pm at the five percent level while it cannot be rejected for me. Interestingly, the univariate time series model that is used for the forecast experiment reported in Müller and Köberl (2012) also features the surprise indicator (pm) and me as the only relevant variables out of the nine possible considered in table 5.

B.3 The results

Pictures 4 and 5 on pages 6 and 7 respectively illustrate the outcomes. The first, figure 4, concerns model selection by showing the different forecast properties of selected potential series. The two top panels compare the forecasts for Swiss GDP for the years 2011 and 2012 by the series ee and em. According to the test statistics the aggregation restriction is valid for both series with p-levels of 9.99% and 6.2% respectively. Therefore, em is preferred over ee. The bottom panel shows the forecast on the basis of the surprise indicator (pm). Overall, the surprise indicator is inferior to em because the aggregation restriction is rejected. In terms of forecasting em and pm are very close.

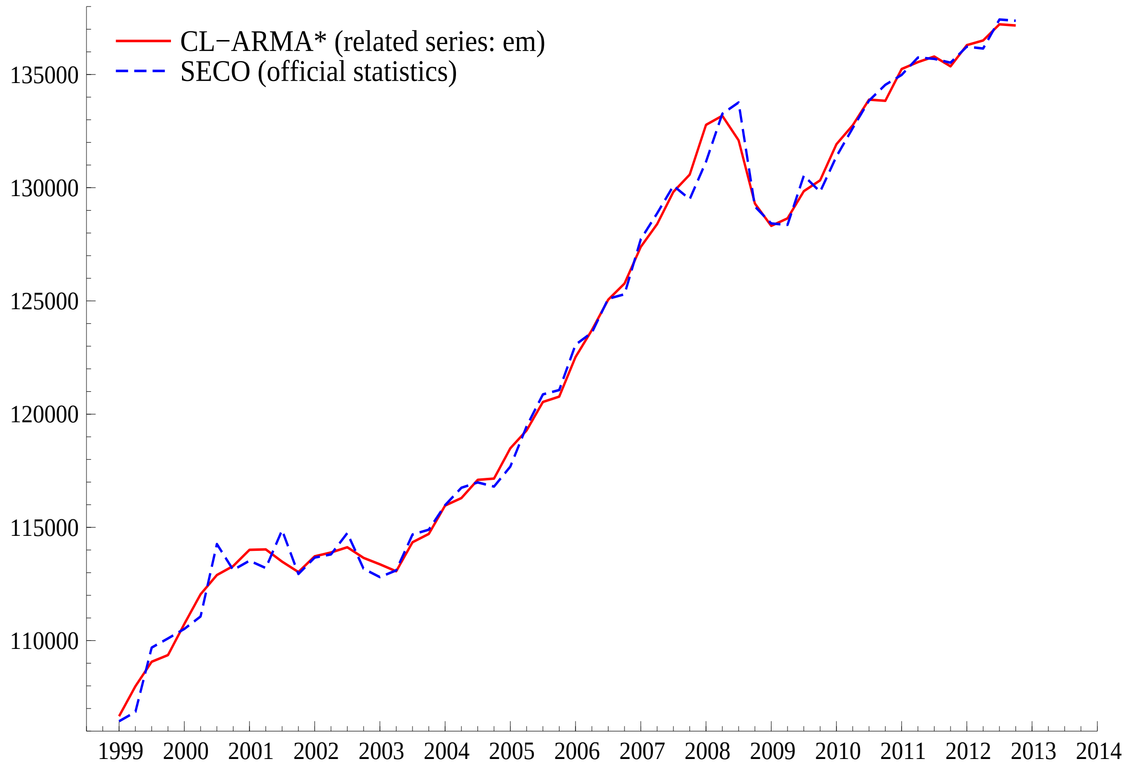
Comparing the official SECO data to the disaggregation by em reveals that the SECO data appears spikier and somewhat lagging. For example, using the survey data puts the local peak marking the onset of the financial crises to 2008q2 while according to SECO GDP kept expanding one more quarter. Similarly, the consecutive trough came in 2009, first quarter according to the survey data while GDP started to grow again in 2009q2 only if judged by SECO data. Both



Vertical axis: mill. Swiss Francs, horizontal axis: year. Disaggregation and forecast based on CL model (see equation (1)).

Figure 4: Forecast comparison and temporal disaggregation of Swiss real GDP by the Chow-Lin procedure using survey data

quarterly data series are disaggregation estimates. The survey data approach, however, can be justified by internal validation on the basis of testing the aggregation restriction.



Vertical axis: mill. Swiss Francs, horizontal axis: year and quarter. Disaggregation based on CL model (see (1)) and on SECO methods: `http://www.seco.admin.ch/themen/00374/00456/00458/index.html?lang=en`.

Figure 5: Temporal disaggregation of Swiss real GDP by the Chow-Lin using survey data and by SECO