Using multilevel models to estimate seasonal factors from short span data

Craig H. McLaren and Xichuan (Mark) Zhang
Australian Bureau of Statistics,
Chandler St, GPO Box 10,
Belconnen, 2616,
ACT, Australia

1. Introduction

Consider a new survey replacing an existing old survey. Assume that the new survey measures the same underlying activity although the coverage and scope may be different. That is the underlying movements are the same but the time series may shift to a different level and different seasonal variations. To assess the impact of the changes, original estimates can be calculated using data collected from the old and new survey for one or more overlapping time periods. The number of overlapping time periods is typically short due to cost constraints of running two surveys simultaneously. Traditional seasonal adjustment methods cannot adequately calculate estimates for short time series. This paper shows how information from two parallel surveys, disaggregated over a number of lower level series, can be used in multi-level modelling to test the differences in the seasonal factors between the old and new survey data at an aggregate level for the overlapping time periods. Seasonal factors for the time series obtained from the new survey can also be obtained.

2. Modelling framework

Denote an original series at time t by y_{ijt} where for survey $i=0$ (old survey) or $i=1$ (new survey), data is available at level $j=1,\dots,n$. Consider the log additive decomposition for a time series

$$\log(y_{ijt}) = \log(T_{ijt}) + \log(S_{ijt}) + \log(I_{ijt})$$

where T_{ijt} is the trend component, S_{ijt} is the seasonal component, and I_{ijt} is the irregular component. Assume the trend follows a locally linear process $\log(T_{ijt}) = (a + a_{ij}) + (b + b_j)t$ within the overlapping period and a and b represent the fixed intercept and slope term for the trend, a_{ij} denotes the intercept random effect at the survey and group level, and b_j represents the random effect of the slope at group level. Assume $\log(S_{ijt}) = (\mathbf{c} + \mathbf{c}_{ij})'\mathbf{s}_t$ where $\mathbf{s}_t$ is a seasonal dummy variable, $\mathbf{c}$ is the common fixed seasonal factor vector, and $\mathbf{c}_{ij}$ is the seasonal random effect at the survey and group level. All random effects are normally distributed with an appropriate variance structure. This gives a multi level mixed random model

$$\log(y_{ijt}) = (a + bt + \mathbf{c}'\mathbf{s}_t) + (a_{ij} + b_j t + \mathbf{c}_{ij}\mathbf{s}_t) + \varepsilon_{ijt}$$

Here, $\mathbf{c} = (\mathrm{c}_{(1)},\dots,\mathrm{c}_{(p)})$ and $\sum_{k=1}^{p}\mathrm{c}_{(k)} = 0$ where p is the number of periods ($p=4$ for quarterly and $p=12$ for monthly). The estimated intercepts and slopes can then be used to produce a detrended series. The fixed seasonal factors of the old and new surveys with the random seasonal effects at group level and at survey within group can be used to provide statistical inferences to judge whether the seasonal factors have been changed. That is

$$\begin{aligned}\log(y^*_{ijt}) &= \log(y_{ijt}) - (\hat{a} + \hat{b}t + \hat{a}_{ij} + \hat{b}_j t) \\ &= \mathbf{c}_i'\mathbf{s}_t + (\mathbf{c}_j + \mathbf{c}_{ij})'\mathbf{s}_t + \varepsilon_{ijt}\end{aligned}$$

where $\mathbf{c}_i$ is the fixed seasonal effect for the old and new survey and $\mathbf{c}_j$ and $\mathbf{c}_{ij}$ are random seasonal effects at group level and at survey within group level respectively. The statistical influences of $\mathbf{c}_i$ are used to assess if the seasonal factors of the old and new surveys are different. McLaren and Zhang (2003) describe how this approach is used within the ABS framework.

3. Real world example

The Quarterly Business Indicator Survey (New) (ABS, 2001a) replaced the Survey of Employment and Earnings (Old) (ABS, 2001b). This data is used in the ABS National Accounts. Four parallel quarter estimates at the state by industry level were available from March to December 2001.The number of data points is small, 15 industries with 4 overlapping data points, but the impact on the multi-level model parameter estimates will be negligible (Maas and Cox, 2002). The industry estimates were used to estimate seasonal factors at the state level (Table 1). As the estimated seasonal factors between the old and new survey were not significantly different for the Australian Total, the old seasonal factors were used for the March 2002 National Account publication.

	NSW		* Victoria		Queensland		South Australia		Western Australia		* Tasmania		Northern Territory		* ACT		* Total x Ind	
2001	New	Old	New	Old	New	Old	New	Old	New	Old	New	Old	New	Old	New	Old	New	Old
Mar	0.984	0.981	0.983	0.985	0.924	0.972	0.998	0.969	0.986	0.979	0.988	0.974	1.047	1.004	0.986	0.962	0.977	0.980
Jun	1.004	0.994	0.999	0.993	0.979	0.992	0.985	0.992	1.024	1.015	1.023	1.019	1.001	0.995	1.021	1.043	0.999	0.996
Sep	0.983	0.995	0.984	0.992	1.019	0.991	0.999	1.008	0.962	0.993	0.967	0.965	0.966	0.987	0.967	0.990	0.988	0.993
Dec	1.029	1.030	1.034	1.031	1.081	1.045	1.015	1.030	1.029	1.013	1.022	1.042	0.987	1.014	1.024	1.005	1.036	1.031

Table 1: Comparison of filter based seasonal factors for the old and new survey for different states. * denotes that the seasonal factors were not significantly different between the old and new survey (at the 80% confidence level).

4. Conclusion

In the case where lower level group data are available, the use of a multi-level model with mixed effects provides an approach to estimate seasonal factors for short span time series at an aggregate level. Seasonally adjusted and trend estimates can then be derived for short span time series.

REFERENCES

Australian Bureau of Statistics (2001a). Business Indicators, Cat. 5676.0, Canberra, Australia.

Australian Bureau of Statistics (2001b). Wage and Salary Earners, Cat. 6248.0, Canberra, Australia.

Maas, C.J.M. and Cox, J.J. (2002). Robustness of multilevel parameter estimates against small sample sizes. Inter. Sociological Assoc., XVth ISA World Congress of Sociology, July 7-13, 2002.

McLaren, C.H. and Zhang, X. (2003). Estimation of seasonal factors for a short time span using multi-level modelling with mixed effects, Working Paper No. 03/1, Australian Bureau of Statistics.

RÉSUMÉ

Les changements à la méthodologie d'un aperçu peuvent effectuer sur des évaluations originales d'aperçu et des évaluations derives corrigées des variations saisonnières et de tendance. Des évaluations de parallèle du vieux et nouvel aperçu peuvent être obtenues pendant un nombre restreint de périodes de temps de recouvrement. Les méthodes saisonnières traditionnelles d'ajustement ne peuvent pas en juste proportion calculer des évaluations corrigées des variations saisonnières et de tendance pour des series chronologiques courtes. La théorie, et un exemple, de l'évaluation des facteurs saisonniers pour les envergures courtes des données de série chronologique sont donnés.