Temporal Disaggregation of Time Series

Sax, Christoph and Steiner, Peter

December 2013

Temporal Disaggregation of Time Series

by Christoph Sax and Peter Steiner

Abstract Temporal disaggregation methods are used to disaggregate low frequency time series to higher frequency series, where either the sum, the average, the first or the last value of the resulting high frequency series is consistent with the low frequency series. Temporal disaggregation can be performed with or without one or more high frequency indicator series. The package **tempdisagg** is a collection of several methods for temporal disaggregation.

Introduction

Not having a time series at the desired frequency is a common problem for researchers and analysts. For example, instead of quarterly sales, they only have annual sales. Instead of a daily stock market index, they only have a weekly index. While there is no way to fully make up for the missing data, there are useful workarounds: with the help of one or more high frequency indicator series, the low frequency series may be *disaggregated* into a high frequency series. For example, quarterly exports could help disaggregating annual sales, and a foreign stock market index could help disaggregating the stock market index at home.

Even when there is no high frequency indicator series, one still may want to disaggregate a low frequency series. While the accuracy of the resulting high frequency series will be low, it may still be worth doing so. For example, estimating a vector-autoregressive model requires all variables to have the same frequency. Having one bad high frequency series could still be preferable to the switch to a lower frequency.

The package **tempdisagg** (Sax and Steiner, 2013) implements the following standard methods for temporal disaggregation: Denton, Denton-Cholette, Chow-Lin, Fernandez and Litterman. On the one hand, **Denton** (Denton, 1971) and **Denton-Cholette** (e.g. Dagum and Cholette, 2006) are primarily concerned with movement preservation, generating a series that is similar to the indicator series whether or not the indicator is correlated with the low frequency series. Alternatively, these methods can disaggregate a series without an indicator. On the other hand, Chow-Lin, Fernandez and Litterman use one or several indicators and perform a regression on the low frequency series. **Chow-Lin** (Chow and Lin, 1971) is suited for stationary or cointegrated series, while **Fernandez** (Fernández, 1981) and **Litterman** (Litterman, 1983) deal with non-cointegrated series.

All disaggregation methods ensure that either the sum, the average, the first or the last value of the resulting high frequency series is consistent with the low frequency series. They can deal with situations where the high frequency is an integer multiple of the low frequency (e.g. years to quarters, weeks to days), but not with irregular frequencies (e.g. weeks to months).

Temporal disaggregation methods are widely used in official statistics. For example, in France, Italy and other European countries, quarterly figures of Gross Domestic Product (GDP) are computed using disaggregation methods. Outside of R, there are several software packages to perform temporal disaggregation: Ecotrim by Barcellan et al. (2003); a Matlab extension by Quilis (2012); and a RATS extension by Doan (2008). An overview of the capabilities of the different software programs is given in Table 1.[1]

The first section discusses the standard methods for temporal disaggregation and summarizes them in a unifying framework. Section 2 discusses the working and implementation of the **tempdisagg** package. Section 3 presents an illustrative example.

A framework for disaggregation

The aim of temporal disaggregation is to find an unknown high frequency series y, whose sums, averages, first or last values are consistent with a known low frequency series y_l (The subscript l denotes low frequency variables). In order to estimate y, one or more other high frequency indicator variables can be used. We collect these high frequency series in a matrix X. For the ease of exposition and without loss of generality, the terms annual and quarterly will be used instead of low frequency and high frequency hereafter.

The diversity of temporal disaggregation methods can be narrowed by putting the methods in a two-step framework: First, a preliminary quarterly series p has to be determined; second, the

---

[1]Currently, there is no support for temporal disaggreation methods with a contemporaneous constraint (Di Fonzo, 1994).

Methods	Ecotrim	Matlab add-on	RATS add-on	tempdisagg
Chow-Lin (max. log)	◑	●	●	●
Chow-Lin (min. RSS)	●	●	○	●
Fernández	●	●	●	●
Litterman (max. log)	◑	●	●	●
Litterman (min. RSS)	●	●	○	●
Denton-Cholette	◐	◑	○	●
Denton	○	◐	○	●
Contemporaneous constraint	◐	◐	○	○

Table 1: Software packages (●/◐/◑/○: full/partial/erroneous/no implementation).

differences between the annual values of the preliminary series and the annual values of the observed series have to be distributed among the preliminary quarterly series. The sum of the preliminary quarterly series and the distributed annual residuals yields the final estimation of the quarterly series, $\hat{y}$. Formally,

$$\hat{y} = p + Du_l. \tag{1}$$

D is a $n \times n_l$ distribution matrix, with n and n_l denoting the number of quarterly and annual observations, respectively. u_l is a vector of length n_l and contains the differences between the annualized values of p and the actual annual values, y_l:

$$u_l \equiv y_l - Cp. \tag{2}$$

Multiplying the $n_l \times n$ conversion matrix, C, with a quarterly series performs *annualization*. With two years and eight quarters, and annual values representing the sum of the quarterly values (e.g. GDP), the conversion matrix, C, is constructed the following way:[2]

$$C = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}. \tag{3}$$

Equation (1) constitutes a unifying framework for all disaggregation methods. The methods differ in how they determine the preliminary series, p, and the distribution matrix, D. Table 2 summarizes the differences in the calculation of p and D. We will discuss them in turn.

Preliminary series

The methods of **Denton** and **Denton-Cholette** use a single indicator as their preliminary series:

$$p = X, \tag{4}$$

where X is a $n \times 1$ matrix. As a special case, a constant (e.g. a series consisting of only 1s in each quarter) can be embodied as an indicator, allowing for temporal disaggregation without high frequency indicator series.

The regression-based methods **Chow-Lin**, **Fernandez** and **Litterman** perform a Generalized Least Squares Regression (GLS) of the annual values, y_l, on the annualized quarterly indicator series, CX. In this case, X represents a $n \times m$ matrix, where m denotes the number of indicators (including a possible constant). For a given variance-covariance matrix, Σ, the GLS estimator, $\hat{\beta}$, is calculated in the standard way (the estimation of Σ is discussed below):

$$\hat{\beta}(\Sigma) = \left[X'C'(C\Sigma C')^{-1}CX \right]^{-1} X'C'(C\Sigma C')^{-1}y_l. \tag{5}$$

The critical assumption of the regression-based methods is that the linear relationship between the annual series CX and y_l also holds between the quarterly series X and y. Thus, the preliminary series is calculated as the fitted values of the GLS regression:

$$p = \hat{\beta}X. \tag{6}$$

---

[2]Generally, for annual values representing the sum of the quarterly values, C is constructed as $I_{n_l} \otimes [1, 1, 1, 1]$, the Kronecker product of an identity matrix of size n_l and a transposed vector of length n/n_l (4, in the present case). If instead of sums, annual values are averages of the quarterly values, the transposed vector becomes $[0.25, 0.25, 0.25, 0.25]$. If annual values are equal to the first or the last quarterly value, the vector is $[1, 0, 0, 0]$ or $[0, 0, 0, 1]$, respectively.

Methods	p	D	Σ
`denton`	X	$\Sigma C'(C\Sigma C')^{-1}$	Σ_{D}
`denton-cholette`	X	D_{DC}	
`chow-lin-maxlog, chow-lin-minrss-ecotrim, chow-lin-minrss-quilis`	$\hat{\beta}X$	$\Sigma C'(C\Sigma C')^{-1}$	$\Sigma_{\text{CL}}(\rho)$
`litterman-maxlog, litterman-minrss`	$\hat{\beta}X$	$\Sigma C'(C\Sigma C')^{-1}$	$\Sigma_{\text{L}}(\rho)$
`fernandez`	$\hat{\beta}X$	$\Sigma C'(C\Sigma C')^{-1}$	$\Sigma_{\text{L}}(0)$

Table 2: Methods for temporal disaggregation.

Distribution matrix

With the exception of Denton-Cholette, the distribution matrix of all temporal disaggregation methods is a function of the variance-covariance matrix, Σ:

$$D = \Sigma C'(C\Sigma C')^{-1}. \tag{7}$$

The **Denton** methods minimize the squared absolute or relative deviations from a (differenced) indicator series, where the parameter h defines the degree of differencing. For the *additive* Denton methods and for $h = 0$, the sum of the squared absolute deviations between the indicator and the final series is minimized. For $h = 1$, the deviations of first differences are minimized, for $h = 2$, the deviations of the differences of the first differences, and so forth. For the *proportional* Denton methods, deviations are measured in relative terms.

For the additive Denton method with $h = 1$, the variance-covariance matrix has the following structure:

$$\Sigma_{\text{D}} = (\Delta'\Delta)^{-1} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 2 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & \cdots & n \end{bmatrix}, \tag{8}$$

where Δ is a $n \times n$ difference matrix with 1 on its main diagonal, -1 on its first subdiagonal and 0 elsewhere. For $h = 2$, $\Delta'\Delta$ is multiplied by Δ' from the left and Δ from the right side. For $h = 0$, it is the identity matrix of size n.

Denton-Cholette is a modification of the original approach and removes the spurious transient movement at the beginning of the resulting series. While generally preferable, the calculation of the distribution matrix, D_{DC}, does not fit into the simple framework (see Dagum and Cholette, 2006, pp. 136, for an extensive description).

Chow-Lin assumes that the quarterly residuals follow an autoregressive process of order 1 (AR1), i.e., $u_t = \rho u_{t-1} + \epsilon_t$, where ϵ is $\text{WN}(0, \sigma_\epsilon)$ (with WN denoting White Noise) and $|\rho| < 1$. The resulting covariance matrix has the following form:

$$\Sigma_{\text{CL}}(\rho) = \frac{\sigma_\epsilon^2}{1 - \rho^2} \cdot \begin{bmatrix} 1 & \rho & \cdots & \rho^{n-1} \\ \rho & 1 & \cdots & \rho^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \cdots & 1 \end{bmatrix}. \tag{9}$$

The estimation of Σ_{CL} thus requires the estimation of an AR1 parameter ρ, which will be discussed in the next section. The variance, σ_ϵ^2, cancels out and does not affect the calculation of neither D nor $\hat{\beta}$.

The remaining methods deal with cases when the quarterly indicators and the annual series are not cointegrated. **Fernandez** and **Litterman** assume that the quarterly residuals follow a non-stationary process, i.e. $u_t = u_{t-1} + v_t$, where v is an AR1 ($v_t = \rho v_{t-1} + \epsilon_t$, where ϵ is $\text{WN}(0, \sigma_\epsilon)$). Fernandez is a special case of Litterman, where $\rho = 0$, and, therefore, u follows a random walk. The variance-covariance matrix can be calculated as follows:

$$\Sigma_{\text{L}}(\rho) = \sigma_\epsilon^2 \left[\Delta' H(\rho)' H(\rho) \Delta \right]^{-1}, \tag{10}$$

where Δ is the same $n \times n$ difference matrix as in the Denton case; $H(\rho)$ is a $n \times n$ matrix with 1 on its main diagonal, $-\rho$ on its first subdiagonal and 0 elsewhere. For the special case of Fernandez, with

$\rho = 0$, the resulting covariance matrix has the following form:

$$\Sigma_L(0) = \sigma_\epsilon^2 \cdot (\Delta'\Delta)^{-1} = \sigma_\epsilon^2 \cdot \Sigma_D. \qquad (11)$$

Estimating the autoregressive parameter

There are several ways to estimate the autoregressive parameter ρ in the Chow-Lin and Litterman methods. An iterative procedure has been proposed by Chow and Lin (1971). It infers the parameter from the observed autocorrelation of the low frequency residuals, u_l.

In a different approach, Bournay and Laroque (1979, p. 23) suggest the maximization of the likelihood of the GLS-regression:

$$L(\rho, \sigma_\epsilon^2, \beta) = \frac{\exp\left[-\frac{1}{2}u_l' \left(C\Sigma C'\right)^{-1} u_l\right]}{(2\pi)^{n_l/2} \cdot \left[\det\left(C\Sigma C'\right)\right]^{1/2}}, \qquad (12)$$

where u_l is given by Eq. (2) and (6). $\hat{\beta}$ turns out to be the GLS estimator from Eq. (5). The maximum likelihood estimator of the autoregressive parameter, $\hat{\rho}$, is a consistent estimator of the true value, thus it has been chosen as the default estimator. However, in some cases, $\hat{\rho}$ turns out to be negative even if the true ρ is positive. Thus, by default, **tempdisagg** constrains the optimization space for ρ to positive values.

A final approach is the minimization of the weighted residual sum of squares, as it has been suggested by Barbone et al. (1981):

$$RSS(\rho, \sigma_\epsilon^2, \beta) = u_l' \left(C\Sigma C'\right)^{-1} u_l. \qquad (13)$$

Contrary to the maximum likelihood approach, σ_ϵ^2 does not cancel out. The results are thus sensitive to the specification of Σ, with different implementations leading to different but inconsistent estimations of ρ.

The tempdisagg package

The selection of a temporal disaggregation model is similar to the selection of a linear regression model. Thus, td, the main function of the package, closely mirrors the working of the lm function (package **stats**), including taking advantage of the formula interface.[3]

```
td(formula, conversion = "sum", to = "quarterly", method = "chow-lin-maxlog",
   truncated.rho = 0, fixed.rho = 0.5, criterion = "proportional", h = 1,
   start = NULL, end = NULL, ...)
```

The left hand side of the formula denotes the low frequency series, the right hand side the indicators. If no indicator is specified, the right hand side must be set equal to 1. The variables can be entered as time series objects of class "ts" or as standard vectors or matrices. If entered as "ts" objects, the resulting series will be "ts" objects as well.

The conversion argument indicates whether the low frequency values are sums, averages, first or last values of the high frequency values ("sum" (default), "average", "first" or "last", respectively). The method argument indicates the method of temporal disaggregation, as shown in Table 2 (see ?td for a complete listing of methods). The to argument indicates the high frequency destination as a character string ("quarterly" (default) or "monthly") or as a scalar (e.g. 2, 7, for year-semester or week-day conversion). It is only required if no indicator series is specified (Denton methods), or if standard vectors are used instead of time series objects. Finally, you can set an optional start or end date. This is identical to pre-processing the input series with window.

td returns an object of class "td". The function predict computes the disaggregated high frequency series, $\hat{y}$. If the high frequency indicator series are longer than the low frequency series, the resulting series will be extrapolated.

The implementation of **tempdisagg** follows the same notation and modular structure as the exposure in the previous section. Internally, td uses the optimize function (package **stats**) to solve the one-dimensional optimization problem at the core of the Chow-Lin and Litterman methods. For GLS estimation, td uses an efficient and nummerically stable algorithm that is based on the qr-decomposition (Paige, 1979).

---

[3]There is no data argument in td, however. Because td is working with series of different length and frequencies, it is not possible to combine them in a single "data.frame".

An example

Suppose we have an annual series and want to create quarterly values that sum up to the annual values. Panel 1 of Fig. 1 depicts annual sales of the pharmaceutical and chemical industry in Switzerland, sales.a, from which we want to create a quarterly series. The following example demonstrates the basic use of **tempdisagg**. It can also be run by demo(tempdisagg).

The most simple method is "denton-cholette" without an indicator series. It performs a simple interpolation that meets the temporal additivity constraint. In R, this can be done the following way:

```
> library(tempdisagg)
> data(swisspharma)
> m1 <- td(sales.a ~ 1, to = "quarterly", method = "denton-cholette")
> predict(m1)
```

td produces an object of class "td". The formula, sales.a ~ 1, indicates that our low frequency variable, sales.a, will be disaggregated with a constant, 1 (see ?formula for the handling of the intercept in the formula interface). The resulting quarterly values of sales can be extracted with the predict function. As there is no additional information on quarterly movements, the resulting series is very smooth (Panel 2 of Fig. 1).

While this purely mathematical approach is easy to perform and does not need any other data series, the economic value of the resulting series may be limited. There might be a related quarterly series that follows a similar movement than sales. For example, we may use quarterly exports of pharmaceutical and chemical products, exports.q (Panel 3 of Fig. 1):

```
> m2 <- td(sales.a ~ 0 + exports.q, method = "denton-cholette")
```

Because we cannot use more than one indicator with the "denton-cholette" (or "denton") method, the intercept must be specified as missing in the formula (0). Contrary to the first example, the to argument is redundant, because the destination frequency can be interfered from the time series properties of exports.q. Applying the predict function to the resulting model leads to a much more interesting series, as shown in Panel 4 of Fig. 1. As the indicator series is longer than the annual series, there is an extrapolation period, in which quarterly sales are forecasted.

With an indicator, the "denton-cholette" method simply transfers the movement of the indicator to the resulting series. Even if in fact there were no correlation between the two series, there would be a strong similarity between the indicator and the resulting series. In contrast, regression based methods transfer the movement only if the indicator series and the resulting series are actually correlated on the annual level. For example, a Chow-Lin regression of the same problem as above can be performed the following way:

```
> m3 <- td(sales.a ~ exports.q)
```

As "chow-lin-maxlog" is the default method, it does not need to be specified. Like with the corresponding lm method, summary produces an overview of the regression:

```
> summary(m3)

Call:
td(formula = sales.a ~ exports.q)

Residuals:
    Min     1Q  Median     3Q     Max
-77.892  -7.711  -4.628   9.647  36.448

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.241e+01  1.493e+00   8.311 1.06e-09 ***
exports.q   1.339e-02  1.672e-04  80.111  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

'chow-lin-maxlog' disaggregation with 'sum' conversion
36 low-freq. obs. converted to 146 high-freq. obs.
Adjusted R-squared: 0.9946  AR1-Parameter:     0 (truncated)
```

There is indeed a strong correlation between exports and sales, as it has been assumed in the "denton-cholette" example above. The coefficient of exports.q is highly significant, and the very
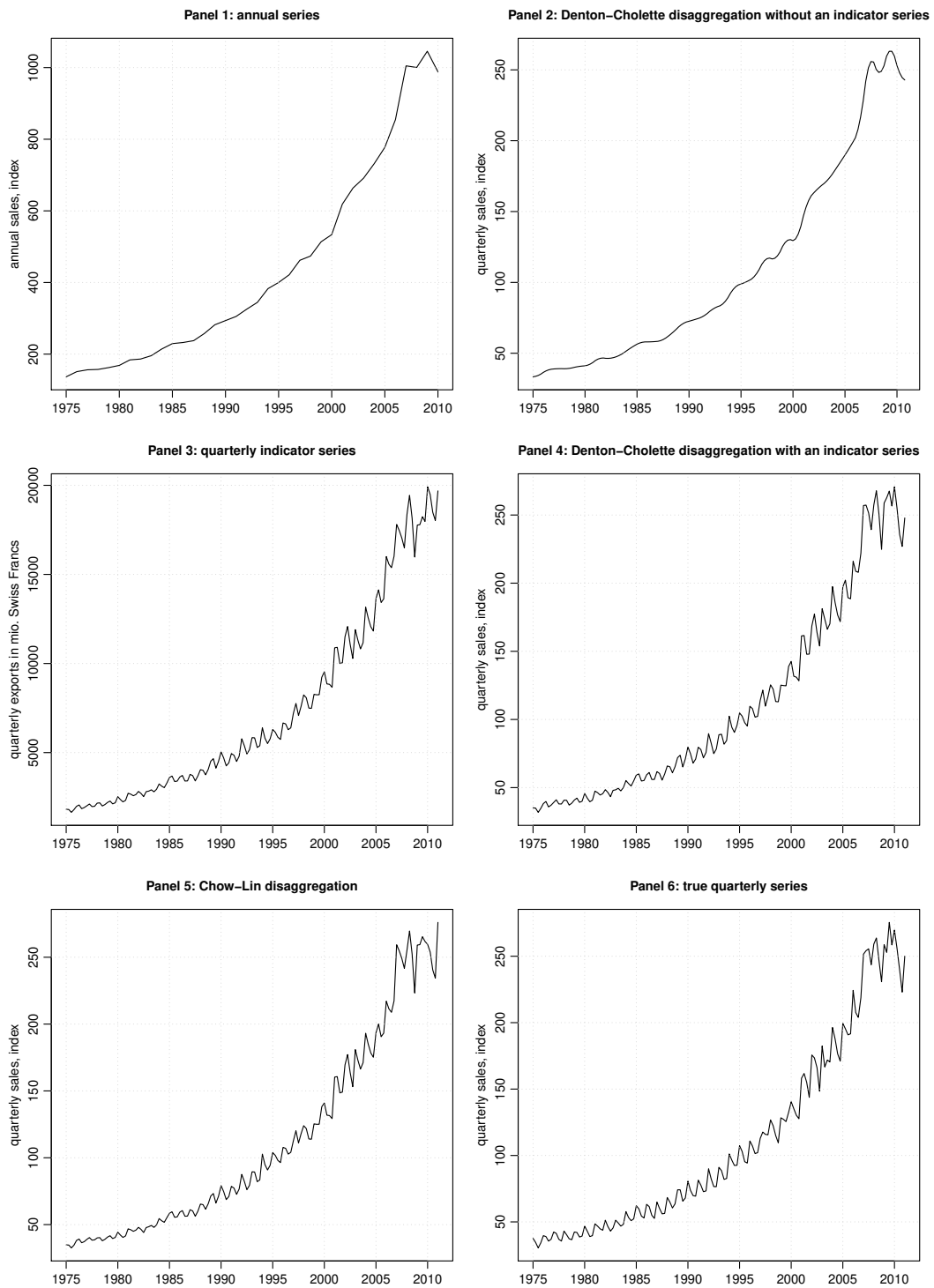
Figure 1: Disaggregating an annual series to quarterly series with no or one indicator series.

high adjusted R^2 points to a strong relationship between the two variables. The coefficients are the result of a GLS regression between the annual series. The AR1 parameter, ρ, was estimated to be negative; in order to avoid the undesirable side-effects of a negative ρ, it has been truncated to 0 (This feature can be turned off). Again, with the predict function, we can extract the resulting quarterly series of sales (Panel 5 of Fig. 1). Like all regression based methods, "chow-lin-maxlog" can also be used with more than one indicator series:

```
> m4 <- td(formula = sales.a ~ exports.q + imports.q)
```

In our example, we actually know the true data on quarterly sales, so we can compare the estimated values to the true values. With an indicator series, both the Denton method and Chow-Lin produce a series that is close to the true series (Panel 6 of Fig. 1). This is, of course, due to fact that in this example, exports are a good indicator for sales. If the indicator is less close to the series of interest, the resulting series will be less close to the true series.

Summary

tempdisagg implements the standard methods for temporal disaggregation. It offers a way to disaggregate a low frequency time series into a higher frequency series, while either the sum, the average, the first or the last value of the resulting high frequency series is consistent with the low frequency series. Temporal disaggregation can be performed with or without the help of one or more high frequency indicators. If good indicators are at hand, the resulting series may be close to the true series.

Bibliography

L. Barbone, G. Bodo, and I. Visco. Costi e profitti nell'industria in senso stretto: Un'analisi su serie trimestrali, 1970–1980. *Bolletino della Banca d'Italia*, pages 467–510, 1981. [p83]

R. Barcellan, T. Di Fonzo, D. Raffaele, V. Staplehurst, and D. Buono. *Ecotrim: A Program for Temporal Disaggregation of Time Series*, 2003. URL https://circabc.europa.eu/w/browse/c6049bc0-c633-4cab-9811-b476ffe08370. Version 1.01. [p80]

J. Bournay and G. Laroque. Réflexions sur la méthode d'élaboration des comptes trimestriels. *Annales de l'INSÉÉ*, 36:3–30, 1979. [p83]

G. C. Chow and A.-L. Lin. Best linear unbiased interpolation, distribution, and extrapolation of time series by related series. *The Review of Economics and Statistics*, 53(4):372–375, Nov. 1971. [p80, 83]

E. B. Dagum and P. A. Cholette. *Benchmarking, Temporal Distribution, and Reconciliation Methods for Time Series*. Lecture Notes in Statistics. Springer-Verlag, New York, 2006. [p80, 82]

F. T. Denton. Adjustment of monthly or quarterly series to annual totals: An approach based on quadratic minimization. *Journal of the American Statistical Association*, 66:99–102, Mar. 1971. [p80]

T. Di Fonzo. Temporal disaggregation of a system of time series when the aggregate is known: Optimal vs. adjustment methods. In *Workshop on Quarterly National Accounts*, pages 63–77, Paris, Dec. 1994. Eurostat. [p80]

T. Doan. *Disaggregate: A General Procedure for Interpolation*, 2008. URL www.estima.com/procs_perl/disaggregate.src. RATS library version Apr. 07, 2008. [p80]

R. B. Fernández. A methodological note on the estimation of time series. *The Review of Economics and Statistics*, 63(3):471–476, 1981. [p80]

R. B. Litterman. A random walk, Markov model for the distribution of time series. *Journal of Business & Economic Statistics*, 1(2):169–173, 1983. [p80]

C. C. Paige. Fast numerically stable computations for generalized linear least squares problems. *SIAM Journal on Numerical Analysis*, 16(1):165–171, 1979. [p83]

E. M. Quilis. *Temporal Disaggregation Library*, 2012. URL www.mathworks.com/matlabcentral/fileexchange/24438-. Matlab library version May 08, 2012. [p80]

C. Sax and P. Steiner. *tempdisagg: Methods for Temporal Disaggregation and Interpolation of Time Series*, 2013. URL http://CRAN.R-project.org/package=tempdisagg. R package version 0.22. [p80]

Christoph Sax
State Secretariat of Economic Affaires, Bern
University of Basel
Switzerland
christoph.sax@gmail.com

Peter Steiner
Federal Finance Administration, Bern
University of Bern
Switzerland
pete.steiner@gmx.net