5

Alternative Decomposition Methods

5.1 Introduction

In this chapter, we examine three decomposition methods. These techniques are concerned with parametric models. The total change of the demographic function is decomposed into the contributions of the various factors included in the model.

The first method generally referred to as "Regression Decomposition" (RD) is extensively used in demography. It is based on linear relationships of a dependent variable in terms of independent variables in various subpopulations. As an application here we study the regression models for the achieved number of children for three regions of Mexico: central, north and south. The mean number of children is regressed in each region by four factors with varying regional levels of influence. The reasons for these disparities are studied by comparing the linear models between the different regions. The total change in this regression decomposition exercise is the result of adding the disparities in parameters and independent variables among the different regions.

The second method is called the "Purging Method". Similar to RD, first a relationship model is found. Here, the selected model is a multiplicative (or log-linear) model for cross-classified data. We employ Mexican migration cross-classified by educational achievement and region of origin: central, north and south. The educational level in Mexico differs by region of origin. It is, therefore, important to free the frequencies of the undesirable interaction between region of origin and the distribution of the educational attainment. The confounding factor is eliminated by "purging" the model. In other words, the model is transformed into a model without the undesired factor, which allows comparisons by educational attainment.

The last model, the "Delta Method", is as the two previous models based on a parametric function. The partial derivatives of the defined model with respect to all parameters are calculated. The total change in the defined model is found by addition of the partial derivatives.

Here, assuming that mortality follows a Gompertz trajectory we study the change in mortality explained by the contribution of each of the parameters.

5.2 Regression Decomposition

This section reviews the regression decomposition (RD) method. This methodology's goal is to explain the disparity between linear relationships of the different subpopulations. Two components explain the variation in these linear models. The first effect on the outcome of the dependent variable corresponds to group membership. This component is interpreted as the influence of the factors in the different subpopulations. The second component accounts for the difference in the characteristics of the two groups.

Among the researchers that have contributed to the development of regression decomposition are Coleman et al. (1972), Alwin and Hauser (1975), Sobel (1983), Clogg and Eliason (1986), Firebaugh (1989), Rodgers (1990), Al-Qudsi and Shah (1991), Leibowitz and Klerman (1995), Hanson et al. (1996), Bidani and Ravallion (1997), and Shen (1999).

The decomposition methods examined in Chapter 3 are based on arithmetic manipulations of differences between demographic variables. Similarly, regression decomposition is based on arithmetic manipulations of differences between the estimated relationships of the regressions.

The RD method is explained and illustrated by an example. Coleman et al. (1972) present a regression decomposition that accounts for the difference between two groups based on the relationship of interest. Suppose the two groups are $i = 1, 2$, and in each group there is a linear relation between a dependent variable y_i, and a set of K variables x_{ik}. The size of the groups is specified as n_i and the process generating the linear relation as

$$y_i = a_i + \sum_{k=1}^{K} b_{ik} x_{ik} + \epsilon_i \tag{5.1}$$

where a_i is an intercept, b_{ik} the kth parameter, and the stochastic disturbance ϵ_i has mean zero and variance σ_i^2.

Then the groups' means $\bar{y}_i$ are of the form

$$\bar{y}_i = \alpha_i + \sum_{k=1}^{K} \beta_{ik} \bar{x}_{ik} \tag{5.2}$$

where $\bar{x}_{ik}$ is the mean of the kth explanatory variable in the ith group. The intercept and parameter estimates are denoted by α_i and β_{ik}, respectively.

The disparity between the two groups is studied by comparison of the means, $\bar{y}_2 - \bar{y}_1$. This difference is decomposed as follows:

$$\begin{aligned} \bar{y}_2 - \bar{y}_1 &= (\alpha_2 - \alpha_1) + \sum_{k=1}^{K} \left(\frac{\bar{x}_{1k} + \bar{x}_{2k}}{2} \right) (\beta_{2k} - \beta_{1k}) \\ &\quad + \sum_{k=1}^{K} \left(\frac{\beta_{1k} + \beta_{2k}}{2} \right) (\bar{x}_{2k} - \bar{x}_{1k}) \,. \end{aligned} \tag{5.3}$$

The components of these expressions refer to two sources of change: a) the difference in parameters, that is, the difference in intercepts α_i, and slopes β_{ik}, which are the first two terms of (5.3), and b) the difference in the mean values of the independent variables $\bar{x}_{ik}$, the third component in (5.3). The component (a) may be interpreted as the influence of group membership on the dependent variable. The second component (b) is the difference in the characteristics between the two groups.

As an application of this technique we look at the total number of children per person for different regions of Mexico. The mean number of children is regressed in each region by four factors, namely educational level, unions in life, age at birth of the first child and the sex of the respondent. These factors are used here as predictors of the achieved number of children. Their level of influence differs from region to region and factor to factor. The reasons for these disparities are examined by applying the *RD* method.

First we look at the outcomes of the regression models for the Mexican regions. These results are based on the National Retrospective Demographic Survey, or EDER (1998) in Spanish, carried out in Mexico in 1998. Table 5.1 includes the regression models for the achieved number of children for the three aforementioned regions of Mexico. Here it should be noted that this example is not meant to be an exhaustive study of Mexican fertility but rather it is an application of the *RD* technique. Again, the mean number of children has been regressed by four factors, namely educational level, unions in life, parental age at birth of the first child and the sex of the respondent. For each region there are two columns containing the coefficients, denoted as β, and means of these factors, $\bar{x}$. The first row is for the intercept α_i. In the analyzed survey the mean number of children is higher among Mexicans from central Mexico than for those from both north and south Mexico, 5.14, 4.50 and 4.60 respectively. Education level is the highest in the north and lowest in the south, as reflected in the average of this factor $\bar{x}$. But this factor negatively influences the achieved number of children, seen in the coefficients of the factor β. The negative influences of education on the mean number of children is more pronounced in the center than in the other two regions.

In a similar way other features come to light. For example, the age at first child is also the highest in the north and the lowest in the south of Mexico. Nevertheless, this factor exhibits only a minor influence on the dependent variable. To understand the disparities between characteristics, $\bar{x}$, and the influence of these characteristics in each region, β, we applied the *RD* method.

Table 5.2 presents the results of applying the *RD* method shown in equation (5.3) with respect to the differential in achieved number of children between Mexicans from the center and those from the north and south of the country. The columns of the difference in parameters $\Delta\beta$ evaluate the influence of the factors in the different regions,

$$\Delta\beta = (\alpha_2 - \alpha_1) + \sum_{k=1}^{K} \left(\frac{\bar{x}_{1k} + \bar{x}_{2k}}{2} \right) (\beta_{2k} - \beta_{1k}).$$

The columns of the change in the average values $\Delta\bar{x}$ represent the proportion of the change that can be explained by the characteristics of the populations in the north or south compared

Table 5.1: Number of children regressed by age at first child, education achievement, sex and the number of unions in life, for Mexicans from the center, north and south of the country.

Variable	Center		North		South	
Number of children	5.14		4.50		4.60	
	β	$\bar{x}$	β	$\bar{x}$	β	$\bar{x}$
Intercept	11.237***		11.366***		9.856***	
	(0.593)		(0.827)		(0.737)	
Education	-1.409***	1.298	-1.261***	1.540	-1.169***	1.225
	(0.083)		(0.123)		(0.107)	
Unions in life	0.354	1.100	0.267	1.110	0.279	1.120
	(0.218)		(0.251)		(0.242)	
Age at first child	-0.170***	22.930	-0.170***	23.450	-0.164***	22.630
	(0.016)		(0.022)		(0.020)	
Sex	-0.490***	1.530	-0.807***	1.540	-0.289	1.530
	(0.168)		(0.241)		(0.207)	
$\bar{\text{R}}$ square	0.287		0.275		0.294	

Source: Based on EDER (1998). The notations β and $\bar{x}$ corresponds to the parameter estimates and the means, respectively. *** p<0.001. The states of the central region are: Aguascalientes, Colima, Distrito Federal, Guanajuato, Hidalgo, Jalisco, México, Michoacán, Morelos, Nayarit, Puebla, Queretaro, San Luis Potosí, Tlaxcala and Zacatecas. The states of the north region are: Baja California, Baja California Sur, Coahuila, Chihuahua, Durango, Nuevo León, Sinaloa, Sonora and Tamauilipas. The states of the south region are: Campeche, Chiapas, Guerrero, Oaxaca, Quintana Roo, Tabasco, Veracruz and Yucatán.

to the center,

$$\Delta\bar{x} = \sum_{k=1}^{K} \left(\frac{\beta_{1k} + \beta_{2k}}{2} \right) (\bar{x}_{2k} - \bar{x}_{1k}) \, .$$

In both, $\Delta\beta$ and $\Delta\bar{x}$ the center is denoted by the subindex 1 and the north or the south by 2. In Table 5.1 both education and age at first child show an inverse influence on the total number of children born. The higher the level of these two variables, the lower the number of children born. These inverse relationships are not equally present in all Mexican regions. Table 5.2 shows the comparison between the northern and central regions and the southern and central regions of Mexico in the columns labeled North-Center and South-Center respectively. A negative value in the table indicates an advantage in favor of Mexicans from the central region, while a positive value indicates advantages for those from the northern or southern regions.

The factors of education and age at first child show negative values of $\Delta\bar{x}$ and T in the columns of the North-Center and positive values in the South-Center. On the other hand, the difference in parameters, $\Delta\beta$, is positive for both North-Center and South-Center. By combining these two results we reach the following conclusion. The central region shows the greatest inverse relationship between the factors of education and age at first child, and the achieved number of children. However, the northern region has achieved levels of the

Table 5.2: Regression decomposition of the differential in achieved number of children between Mexicans from the center and those from the north and south. Number of children is regressed on educational achievement, number of unions in life, age at first child and sex.

Factor	North-Center			South-Center		
	$\Delta\beta$	$\Delta\bar{x}$	T	$\Delta\beta$	$\Delta\bar{x}$	T
Intercept	0.129	0	0.129	-1.381	0	-1.381
Education	0.210	-0.323	-0.113	0.303	0.094	0.397
Unions in life	-0.096	0.003	-0.093	-0.083	0.006	-0.077
Age at first child	0	-0.088	-0.088	0.137	0.050	0.187
Sex	-0.487	-0.006	-0.493	0.308	0	0.308
Total contribution	-0.244	-0.414	-0.658	-0.716	0.150	-0.566

Source: Based on EDER (1998). The notations $\Delta\beta$ and $\Delta\bar{x}$ correspond to the differences in the parameter estimates and the means, respectively. The column T contains the addition of parameters and means differences. A negative entry indicates an advantage in favor of the people from the central region while a positive entry indicates an advantage for people from the north or the south.

characteristics under study seen in the $\Delta\bar{x}$ terms, that surpass the contribution of the greater influence of the factors in the central region, $\Delta\beta < \Delta\bar{x}$. As a result, the total contribution of the factors of education and age at first child is negative in the column T of the North-Center. Medina (2000) presents a thorough study on the Mexican educational level by region. The results shown in Table 5.1 and 5.2 accord with those of Medina's and they indicate the necessity for regional comparisons.

In the last row of Table 5.2 we find the total contribution of the difference in parameters and means. The total contribution of the difference in characteristics, $\Delta\bar{x}$, is the most important factor in the North-Center difference. The northern region has undesirable characteristics that impact on the increase in the number of children. For the South-Center disparities, $\Delta\bar{x}$ is a minor component due to similar characteristics between the southern and central regions. The overall influence of the factors, $\Delta\beta$, is the major component featuring in the South-Center difference.

The *RD* formula (5.3) described above is not unique. Sobel (1983) used the parameters and independent variables from the second group as weights, and he obtained the following decomposition

$$\begin{aligned}\bar{y}_2 - \bar{y}_1 &= (\alpha_2 - \alpha_1) + \sum_{k=1}^{K} \bar{x}_{2k}(\beta_{2k} - \beta_{1k}) \\ &\quad + \sum_{k=1}^{K} \beta_{2k}\,(\bar{x}_{2k} - \bar{x}_{1k}) - \sum_{k=1}^{K} (\beta_{2k} - \beta_{1k})\,(\bar{x}_{2k} - \bar{x}_{1k})\,. \qquad (5.4)\end{aligned}$$

This decomposition then has three components because an interaction term has been added. The interaction component is used to balance the formula and is interpreted as a good (bad) election of the weight if its value is small (large). In some research, decompositions of the type seen in (5.4) are preferred over (5.3). Examples of these are found in the works of Al-Qudsi and Shah (1991) and Hanson et al. (1996). There the parameters and independent variables of one group are considered as weights. Al-Qudsi and Shah (1991) compared the economic progress between foreigners and nationals in Kuwait, taking the Kuwaiti population as the weights. Hanson et al. (1996) studied the trends in child support comparing several periods and taking the values seen in the initial period as the weights.

Another decomposition procedure is proposed by Clogg and Eliason (1986). These authors are concerned with higher-order moments of the independent variable x_i. Let a polynomial regression model of order K be

$$y_i = a_i + \sum_{k=1}^{K} b_{ik} x_i^k + \epsilon_i \tag{5.5}$$

where x_i^k is the kth power of the independent variable x_i. This decomposition is similar to (5.3), but now the means $\bar{x}_{ik}$ are substituted by the sample moments of the independent variable x_i, $\bar{x}_{ik} = \frac{\sum_{j=1}^{n_i} x_{ij}^k}{n_i}$.

The works by Firebaugh (1989) and Rodgers (1990) are focused on the decomposition of changes in aggregate measures of a characteristic of a population. These studies question the possibility of analogous decompositions such as those presented in Chapter 3. The associations are from rate and compositional components, as seen in Chapter 3, into cohort replacement which are a part of and within cohort change. The average cohort replacement is the movement that occurs every year towards younger generations when new birth cohorts are added and old ones disappear. The within cohort change is an alteration of a characteristic of individuals independent of their cohort membership.

Firebaugh (1989) assumes that an aggregate measure $\bar{y}_i$ is linearly related to time, t_i, and to a cohort component C_i. His suggestion for a decomposition of the difference $\bar{y}_2 - \bar{y}_1$ is one which involves the difference between years $t_2 - t_1$ and another denoting the difference between cohorts $C_2 - C_1$. In both cases, the parameters β of the initial time are used. The difference over years corresponds to changes within a cohort while the cohort replacement describes the difference in the growth rates at birth.

Rodgers (1990) questioned the possible interpretation of the components in the decomposition suggested by Firebaugh. Instead, Rodgers suggests searching for variables related to cohort or period effects in order to examine whether they explained the change.

The existent relationships between the independent variables also inspired the development of other decompositions. By using path analysis Alwin and Hauser (1975) divided the linearity expression of a few independent variables into direct and indirect components. The direct component is the effect of the independent variables on the dependent variable. The indirect part describes the effect of the independent variables mediated by other independent variables. Figure 5.1 shows a causal diagram for a linear recursive model. The variable Y is linearly related to X_1, X_2, X_3 and X_4, while the variables X_1, X_2 and X_3 also are interrelated. The corresponding model contains a direct effect of X_1, X_2, X_3 and X_4 on Y depicted in Figure 5.1 by the straight lines. The broken lines correspond to the indirect effect, these are the effects of X_1, X_2 and X_3 on Y which are mediated by the other variables.

We have now introduced the regression decomposition methodology. A common feature of the methods discussed above is that these statistical inferences are used when the information is derived from samples rather than from whole populations. Therefore, together with the

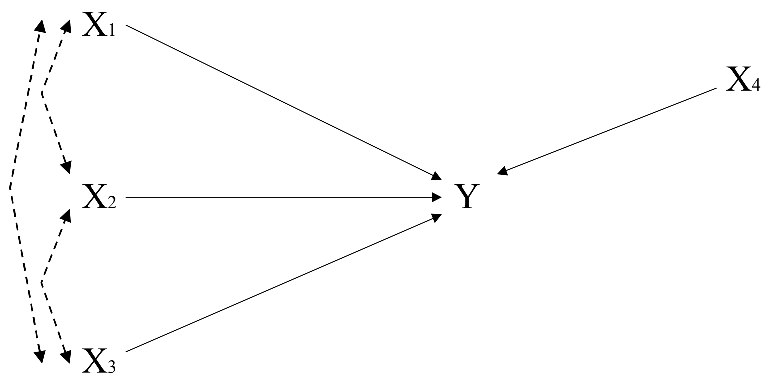

Figure 5.1: Causal diagram for a linear recursive model.

calculations of the regression decomposition a test of statistical hypotheses should also be considered. Since this is beyond the focus of this study, we refer to the study by Sobel (1983), which shows how standard errors and test statistical hypotheses can be computed.

5.3 The Purging Method

Several authors have made contributions to multiplicative models of rates, standardization and components analysis, see Schoen (1970), Teachman (1977), Clogg (1978), Willekens (1982), Willekens and Shah (1984), Clogg and Eliason (1988), Liao (1989) and Xie (1989). In this section the development of this method is presented based on the works of the authors mentioned above.

Schoen (1970) recommended using the geometric mean of the age-specific mortality rates as an average of relative indexes. The author points out the properties which are desirable for a mortality index:

1. Each population should have a real unique number.
2. It should respect proportionality. If mortality is two times higher in the second population, then the index should also be twice as high in the second population.
3. The index should not be affected by confounding components.
4. The index should reflect the nature of the underlying mortality function.

Teachman (1977) further analyzed these properties and showed the relationship between Schoen's index and log-linear models. Log-linear models fulfill all the properties of a mortality index and are useful when making compositional controlled comparisons.

In data, which is cross-classified by different characteristics, we find differences in the rate of occurrence of some event due to group compositions. Let the three-way contingency table be classified by a composition variable C, a group variable G, and a dependent variable D (or occurrence of the event). In the following example concerning Mexico, C is the region of origin, G is educational attainment and D is the dependent variable of Mexican migration.

Let the categories C, G and D be indexed as i, j and k, respectively. In a three-way contingency table the expected frequencies in the cells (i, j, k), are denoted as F_{ijk}, and the cell-specific proportions that fall in the event k can be defined as

$$v_{ij(k)} = \frac{F_{ijk}}{F_{ij\cdot}}, \tag{5.6}$$

where the dot in $F_{ij\cdot}$ corresponds to the summation over the replaced subscript, $F_{ij\cdot} = \sum_k F_{ijk}$. The overall proportion of occurrence in the jth group is given in terms of the frequencies as

$$v_{\cdot j(k)} = \frac{F_{\cdot jk}}{F_{\cdot j\cdot}}. \tag{5.7}$$

Following Goodman's (1970) work, a general multiplicative model in the three-way contingency table CGD may be written as

$$F_{ijk} = \eta\tau_i^C\tau_j^G\tau_k^D\tau_{ij}^{CG}\tau_{ik}^{CD}\tau_{jk}^{GD}\tau_{ijk}^{CGD}, \tag{5.8}$$

where η is the scale factor and the τ parameters denote various kinds of main effects and interactions. The estimates of the proportions of occurrence of the event are obtained by substituting (5.6) or (5.7) by the frequencies obtained from the model parameters in (5.8). Willekens and Shah (1984) show that instead of estimating frequencies and then obtaining proportions it could be possible to calculate directly the model parameter for proportions. They complement their formal proof of the two estimation procedures with applications to migration proportions.

If we assume a multiplicative log-linear model for frequency data of contingency tables, as found in (5.8), the "purged" method suggested by Clogg (1978) can be applied. His suggestion is to remove the confounding effect τ_{ij}^{CG} by looking at the ratios of $F_{ijk}^* = \frac{F_{ijk}}{\tau_{ij}^{CG}}$. The "purged", adjusted or standardized proportions for the confounding effect τ_{ij}^{CG}, are defined as

$$v_{ij(k)}^{**} = \frac{F_{ijk}^*}{F_{ij\cdot}^*}. \tag{5.9}$$

The final adjustments to the "purged" proportions in (5.9) concern rescaling which is carried out to ensure that the sum of the "purged" frequencies is equal to the observed frequencies.

At this juncture we present an application of the purging method. The cross-tabulated data consists of Mexican migration by educational achievement and region of origin. The data is based on the EDER (1998) survey. The observed frequencies of the cross-tabulated data are listed in Table 5.3. The observed proportions of migration by educational attainment in the whole country are 2.04% for persons without any education, 4.11% for persons with some elementary education, 4.23% for persons with secondary education and 3.37% for persons with post-secondary education.

The northern and central regions of Mexico have higher levels of educational attainment than the southern region. It is, therefore, particularly important to free the frequencies of this undesirable interaction. In Table 5.4 the ratios have been adjusted for the confounding influence of the interaction between region of origin and the distribution of the educational attainment. The adjusted ratio for the whole country is now 1.81% for those with no education,

Table 5.3: Observed Mexican migration by education achievement and region of origin.

	No education			Elementary			Secondary			Higher educ.		
	NM	M	T	NM	M	T	NM	M	T	NM	M	T
South	119	1	120	344	5	349	55	4	59	111	3	114
Center	223	6	229	601	34	635	180	4	184	221	8	229
North	43	1	44	269	13	282	105	7	112	127	5	132
Total	385	8	393	1214	52	1266	340	15	355	459	16	475
Ratio %		2.036			4.108			4.226			3.369	

Source: Based on EDER (1998). The notations NM and M correspond to not-migrated and migrated respectively. The states of the Central region are: Aguascalientes, Colima, Distrito Federal, Guanajuato, Hidalgo, Jalisco, México, Michoacán, Morelos, Nayarit, Puebla, Queretaro, San Luis Potosí, Tlaxcala and Zacatecas. The states of the North region are: Baja California, Baja California Sur, Coahuila, Chihuahua, Durango, Nuevo León, Sinaloa, Sonora and Tamauilipas. The states of the South region are: Campeche, Chiapas, Guerrero, Oaxaca, Quintana Roo, Tabasco, Veracruz and Yucatán.

Table 5.4: Mexican migration adjusted by using Clogg's method for purging the confounding influence of the interaction between region of origin and educational achievement.

	No education			Elementary			Secondary			Higher educ.		
	NM	M	T	NM	M	T	NM	M	T	NM	M	T
South	159	1	160	559	8	567	30	2	32	111	3	114
Center	132	4	136	383	22	405	259	6	265	221	8	229
North	95	2	97	280	14	294	54	4	58	127	5	132
Total	386	7	393	1222	44	1266	343	12	355	459	16	475
Ratio %		1.807			3.424			3.251			3.369	

Source: Based on EDER (1998). The notations NM and M are for not-migrated and migrated, respectively.

3.42% for those with elementary education, 3.25% for those with secondary education and 3.37% for those with higher education. As expected we see lower ratios for those with elementary and secondary education, the ratio for the latter group is even below that of the higher educated. The disparities between the educational levels of the different Mexican regions have a confounding effect on migration, which must be controlled. One option for controlling such an interaction is to use the purging method.

Other purging methods can be applied by dividing one or several of the parameter interactions τ from the expected frequencies in equation (5.8), as shown by Clogg and Eliason (1988), Shah (1988), Liao (1989) and Xie (1989).

Liao (1989) proposed a decomposition using purged rates. His model can be extended to several factors by solving a system of simultaneous linear formulas and obtaining the rate differences after the confounding factor has been purged.

The methodology of controlling for confounding factors through purging is mainly used in cases where models of the type seen in (5.8) can be estimated. As mentioned by Willekens

(1982) and Clogg and Eliason (1988), apart from the purged method it is important also to have indicators of the precision of the adjusted rates.

5.4 The Delta Method

The last alternative method here examined is the Delta method. It has been applied by Wilmoth (1988), Pullum, Tedrow and Herting (1989), Foster (1990), Gray (1991) and Pletcher et al. (2000) among others.

The Delta method is based on the chain rule for differentiation of functions of many variables. If v is a function of several variables, $v = f(x_1, x_2, ..., x_n)$, then the total differential of v is given by

$$dv = \sum_{i}^{n} \left(\frac{\partial v}{\partial x_i} \right) dx_i, \tag{5.10}$$

where $\frac{\partial v}{\partial x_i}$ is the differential of the function v with respect to the variable x_i, and dx_i is the differential of the variable x_i.

Pullum et al. (1989) used a finite approximation of (5.10),

$$\Delta v = \sum_{i}^{n} \left(\frac{\partial v}{\partial x_i} \right) \Delta x_i, \tag{5.11}$$

where Δv and Δx_i are finite differences. A similar link between the delta method and discrete decomposition techniques can be found in the work of Gray (1991).

Pullum and colleagues developed a procedure to allocate the changes of the mean fertility to changes in specific parities or groups of parities. The mean parity of a cohort, presented in equation (4.27) as the cohort TFR^c, can also be defined as a function of the parity progression ratios as

$$TFR^c = P_0 + P_0P_1 + P_0P_1P_2 + ... = \sum_{i=0}^{\infty} \prod_{j=0}^{i} P_j, \tag{5.12}$$

where P_j is the parity progression ratio from parity j to parity $j+1$, that is, the number of women of parity $j+1$ divided by women in parity j. Another way of looking at the P_j is the births of order $j+1$ divided by the births of order j. The total change in (5.12) is obtained by taking the changes in every parity. By using the Delta method we get

$$\Delta TFR^c = \sum_{i} \left(\frac{\partial TFR^c}{\partial P_i} \right) \Delta P_i. \tag{5.13}$$

For example, for parity 1 the contribution to the total change is

$$\begin{aligned} \left(\frac{\partial TFR^c}{\partial P_1} \right) \Delta P_1 &= P_0 \Delta P_1 + P_0 P_2 \Delta P_1 + ... \\ &= \left[\frac{TFR^c - P_0}{P_1} \right] \Delta P_1. \end{aligned} \tag{5.14}$$

Pullum and colleagues concluded that the impact of the different parities on the total change of the TFR^c diminishes for higher parities.

Another example of the Delta method is found in the work of Wilmoth (1988). He uses the method for analyzing mortality surfaces $\mu(a,t)$. Instead of examining the level of mortality, the author defines three directions of change for the $\mu(a,t)$: age, time and cohort. Using equation (5.10) the decomposition of the change in the $\mu(a,t)$ is

$$d\mu(a,t) = \frac{\partial\mu(a,t)}{\partial a}da + \frac{\partial\mu(a,t)}{\partial t}dt + \frac{\partial\mu(a+c,t+c)}{\partial c}dc, \tag{5.15}$$

where the three terms on the right hand side correspond to the change with respect to age, time and cohort respectively.

Foster (1990) and Pletcher et al. (2000) studied the derivatives of parametric schedules for demographic measures in fertility and mortality, respectively. Parametric schedules for demographic events take into account the strong regularities in vital rates of human populations. The change in the aggregate demographic measures are expressed as vectors of parameters.

As an application of the Delta method we look at the total change in the force of mortality. If mortality follows a shifting Gompertz trajectory then the force of mortality at age a and time t is

$$\mu(a,t) = \alpha(t)e^{\beta(t)a}, \tag{5.16}$$

where $\alpha(t)$ is the base level mortality and $\beta(t)$ the rate of increase with age, both at time t. Then it follows from equation (5.10) that the change in the age-specific mortality curve can be decomposed into

$$\begin{aligned}\Delta\mu(a,t) &= \frac{\partial\alpha e^{\beta a}}{\partial\alpha}\Delta\alpha + \frac{\partial\alpha e^{\beta a}}{\partial\beta}\Delta\beta \\ &= e^{\beta a}\Delta\alpha + \alpha e^{\beta a}a\Delta\beta. \end{aligned} \tag{5.17}$$

The values of α and β can be estimated from the intercept and slope of a regression line fitted to the logarithm of the age-specific death rates from age 30 to 95 years. That is the life span when mortality approximately follows a Gompertz trajectory. Table 5.5 shows the parameters estimate of the Gompertz trajectories for males and females in Hungary and Japan, for the years 1989 and 1999. In both countries, the male base levels of mortality, $\alpha(t)$, are higher than for females, while the values for the female rates of change over age, $\beta(t)$, are higher than the male rates.

The above-mentioned parameters are used in Table 5.5 for calculating the force of mortality at age 75 and the decomposition of the change as formulated in (5.17). Both Hungary and Japan experienced a decline in the force of mortality at age 75. The change is more pronounced in Japan than in Hungary. The parameters seem to influence the dynamic of the force of mortality in these two countries in different directions. In Hungary, we find a reduction in the base level of mortality with an almost equal increase in the rate of change over age. Japan has experienced the opposite change; a minor increase in the base level of mortality and a decline in the rate of change over age. In other words, improvements for those aged 30 have occurred in Hungary, while mortality rates in Japan have mainly changed in the older groups. In both countries females bear higher reductions than males.

Table 5.5: Gompertz trajectories calculated for males and females in Hungary and Japan, for 1989 and 1999. The force of mortality at age 75 is calculated together with the Delta decomposition of the change in the period.

	Hungary				Japan			
	Males		Females		Males		Females	
t	1989	1999	1989	1999	1989	1999	1989	1999
$\alpha(t)$ (%)	0.027	0.023	0.006	0.004	0.003	0.004	0.001	0.001
$\beta(t)$	0.078	0.079	0.092	0.095	0.097	0.095	0.102	0.098
$\mu(75, t)$	0.091	0.089	0.056	0.054	0.049	0.044	0.029	0.023
$\Delta\mu$ (%)	-0.018		-0.021		-0.050		-0.066	
$\frac{\partial\mu}{\partial\alpha}\Delta\alpha$ (%)	-0.124		-0.166		0.032		0.013	
$\frac{\partial\mu}{\partial\beta}\Delta\beta$ %	0.105		0.143		-0.083		-0.081	
$\Delta\mu = \frac{\partial\mu}{\partial\alpha}\Delta\alpha + \frac{\partial\mu}{\partial\beta}\Delta\beta$ (%)	-0.019		-0.023		-0.051		-0.068	

Source: Author's calculations based on equation (5.17). Data derived from the Human Mortality Database (2002).

5.5 Conclusion

Three decomposition methods are presented in this chapter. The regression decomposition and the purging methods are established techniques in demography, while the Delta method is relatively new in the field. These methods are concerned with parametric models. Regression decomposition and the Delta method separate the contribution of the parameters in the total change, while the purging method eliminates the confounding factors.

Much of the demographic information is derived from samples rather than from whole populations. In these cases the statistical inferences of the linear regressions and the log-linear models in the purging method should be accompanied by indicators of the precision of these models or standard errors. Together with the calculations of the regression decomposition and the purging methods, tests of statistical hypotheses should also be considered. Compared with the simple methods presented in Chapters 3 and 4, these standard errors and tests of statistical hypotheses are drawbacks in the present techniques.

The Delta method is more related to the proposed direct vs. compositional decomposition in Part III. The likeness comes from observing both methods cases that are continuously changing. The disadvantage of the Delta method compared to the other two in this chapter is the assumption of an initial relational model.

With this chapter we conclude Part II of the book. We have seen the evolution of a wide range of decomposition methods during the past years. Chapter 3 presented general decompositions based on arithmetic manipulations of a difference in demographic variables. Chapter 4 introduced the decomposition methods of particular demographic measures. In Chapter 5 we illustrated decompositions that are applied in parametric representations of demographic variables. These chapters dedicated to the previous decomposition methods are our framework for presenting direct vs. compositional decomposition. Part III of the book is focused on this decomposition method for changes over time.