Prevalence and Frequency - Technical Appendix

Ben Matthews

2019-06-12

Implementation Standardization and decomposition are used to calculate the overall contributions of age-structure, prevalence and frequency to change in the aggregate convicted offending rate between 1989 and 2011. To calculate standardized rates across agestructure, prevalence and frequency, equations described in Das Gupta (1993, eqs.3.12-3.17) are used to apply standardization and decomposition for three vector components, with the three vectors being the age-specific prevalence, frequency and age-structure. For a three-factor decomposition, we can describe the convicted offending rate in a particular year as 1. R = F(*alpha*, *beta*, *gamma*). where R is the convicted offending rate in the chosen year. This is expressed as a function of three vectors *alpha*, *beta* and *gamma*: the age-specific proportion of the total population, age-specific prevalence and age-specific frequency respectively. It is worth noting that this equation is is equivalent to 2. R=_sum_i _alpha_i , _beta_i , _gamma_i Where _alpha_i is the age-specific proportion of the total population, _beta_i, is the age-specific prevalence rate and _gamma_i is the age-specific frequency rate. Calculations used to standardize and decompose convicted offending rates will therefore give both the overall contributions of these three factors as well as their age-specific contributions. With the vector notation we can describe the convicted offending rate in the first comparison year (1) as 3. R1 = F(A, B, C) and the convicted offending rate in the second comparison year (2) as 4. R2 = F(a, b, c). From equations 3.12-3.17 in Das Gupta (1993) we see that: the _beta___gamma_-standardized rate (that is, the rate standardized by prevalence and frequency) in year 1 is Q(A), and in year 2 is Q(a); the _alpha___gamma_-standardized rate (the rate standardized by age-structure and frequency) in year 1 is Q(B), and in year 2 is Q(b); the _alpha___beta_-standardized rate (the rate standardized by age-structure and prevalence) in year 1 is Q(C), and in year 2 is Q(c). Consequently we can calculate the effect of age-structure (the *alpha*-effect) by calculating Q(A)-Q(a), the prevalence effect (*beta*-effect) by Q(B)-Q(b) and the frequency effect (*gamma*-effect) by calculating Q(C)-Q(c), where104 5. Q(A) = F(A, b , c)+F(A, B,C) 3 + F(A, b ,C)+F(A,B, c) 6 ,

6. Q(B) = F(a ,B, c)+F(A,B,C) 3

- F(a ,B,C)+F(A ,B, c) 6 ,

7. Q(C) = F(a ,b ,C)+F(A, B,C) 3

- F(a ,B,C)+F(A ,b ,C) 6 and

8. Q(a) = F(a ,b ,c)+F(a,B,C) 3

- F(a ,b ,C)+F(a ,B, c) 6 ,

9. Q(b) = F(a ,b ,c)+F(A , b ,C) 3

- F(a ,b ,C)+F(A,b, c) 6 ,

10. Q(c) = F(a ,b ,c)+F(A ,B, c) 3

- F(a ,B, c)+F(A, b , c) 6 For example, to calculate the effect of change in age structure between 1989 and 2011 on the overall convictions rate, first requires calculating the prevalence and frequency standardized rate for 1989 (_beta___gamma_-standardized rate, Q(A)) for 1989. To this, the 1989 age-structure (that is, the proportion of the population made up by each year of age) is multiplied by the age-specific 2011 prevalence rates and 2011 frequency rates. This figure is then added to the product of the 1989 age-structure, 1989 prevalence and 1989 frequency, all divided by three. This is the figure represented by the first part of equation 5; F(A, b , c)+F(A, B,C) 3 . Then, the 1989 age-structure is multiplied by the 2011 prevalence and 1989 frequency, and then added to the 1989 age-structure multiplied by the 1989 prevalence and 2011 frequency, all divided by six. This is the figure represented by the second part of equation 5; F(A, b ,C)+F(A,B, c) 6 . Together, these two equations produce the overall convictions rate for 1989 standardized by the prevalence and frequency rates for 1989 and 2011 (Q(A)), with the age-structure for 1989 held constant across these two equations. Following the same procedure but holding constant 2011 overall conviction rates produces the figure Q(a). The difference between Q(A) and Q(a) provides the effect of changes in age-structure on overall conviction rates between 1989 and 2011. Results of standardization and decomposition can then be verified by checking that the difference in the convicted offending rates in the two comparison years is equal to the sum of the age, prevalence and frequency effects. Put another way, results should indicate that R2-R1 = *alpha*-effect+ *beta*-effect+ *gamma*-effect. Results of standardization and decomposition will first be presented to show the overall contributions of prevalence, frequency and age-structure to aggregate convicted offending rates. Following this age-specific contributions will be presented visually to show where in the age-distribution change has occurred. Standardization and decomposition are calculated for men and women separately, and their results compared to analyse potential differences in the impact of these three factors across sex.

  Levitt and Marvell and Moody

One feature of standardization and decomposition as calculated using these equations is that results represent averages across the two years being compared (Das Gupta 1993). This means that calculating standardization and decomposition for the start and end of the period covered by SOI will only give average effects of prevalence, frequency and age-structure, and will not be able to account for nonlinearity in these trends. To account for non-linearity in these trends standardization and decomposition are repeated for different periods identified in the data using shaded contour plots. This also helps to hedge against potential ahistoricism (LaFree 1999) – assuming consistent effects over time – by shifting focus away from overall effects of prevalence, frequency and age-structure to examining the relationship between prevalence, frequency and population in different periods. Comparing results between 1989 and 2011 and between different periods within the SOI can demonstrate whether the effects of age-structure, prevalence and frequency have been consistent over time. Even with this approach, it should be noted that the results are sensitive to the choice of years selected. However, informing the selection of years with results from data visualization can help to eliminate potential bias from arbitrary selection of comparison years.