

Appendix 3: Simulations under alternative assumptions

While injury and death risk can reasonably be treated as being constant over successive years of exposure for the same individual illness risk cannot. The risk of contracting an occupational illness in any year probably increases with the years of exposure for an individual worker. We have no information about how much risk rises with years of exposure. If the gradient in risk against years of exposure was linear, then the risk level at the mean duration of time in that occupation would also be the average risk over the career.

Under the assumption of linearity, we therefore sought to establish how sensitive the values that had been derived by our formula were to variations in the gradient in risk. In addition we determined the sensitivity of our results to the levels of risk, the initial size of the occupation, the rate of growth of the occupation, and the mean duration in the occupation.

We did this by running a series of simulations as follows: we varied the number of workers in the occupation over three levels, from the lowest number of workers, to the mean, to the highest number in the data; we varied the growth rates in the number of workers in an occupation from the lowest in the data, to the mean growth rate, to the highest growth rate in the data; we varied the mean duration in an occupation over three levels from the lowest duration in the data, to the mean, to the highest; we varied the linear gradient in risk over time from being constant over all years of exposure, to starting at 50 percent of `illmprob` and ending in 150 percent of `illmprob`, and to starting at 0 and ending at 200 percent of `illmprob`; we varied the values of `illmprob` from the lowest risk level in the data, to the mean, to the highest value in the data set.

Overall we simulated 241 sets of parameters. Each simulation was run over 1000 repetitions. The calendar time each simulation represented was always twice the mean duration in the occupation. From the census year counts of the number of workers in an occupation we knew the approximate annual entry and exit flows into the occupation, and we averaged the inter-census year flows to yield the counts in the census years.

Each exposure-year was represented by a binomial process in which the workers were exposed to a risk that varied with their years in the occupation. For example, in the simulation starting with zero risk in the first year of exposure the new entrants in an occupation would face zero risk but the workers who had been in the occupation one year would face a slightly higher risk and workers at the end of their careers would face twice the mean risk. Once a simulated worker “contracted” an illness the illness was assumed to be permanent. The total number of ill workers in a single run of the simulation were then accumulated. This total was put in to our formula along with the 1990 occupation counts. In this way we created a distribution of realized values for *illmprob* for each of the 241 set-ups.

These simulations showed how much, over the 1000 repetitions, sampling variation could cause deviations between the observed *illmprob* and the mean rate. The standard deviations of these sampling distributions were always less than 10 percent of the mean value of *illmprob*. Thus, even in the smallest occupation with the lowest growth rate, simple chance variation would only cause miniscule differences between the true risk and the sample value.

The simulation parameters and results in terms of means and standard deviations for three selected occupations are shown in Table C1 below. They were chosen to represent a small, a medium, and a large occupation. The mid-career illmprob values came from applying our approximation formula to the data in LFS *Accident and Ill-Health Trailer Questionnaire*. The mean value of the sampling distributions is always close to the mid-career value of illmprob. This result is not remarkable because the mean of 1,000 samples should be close to the expected value. However, the small standard deviations over all parameter combinations were not a foregone conclusion.

Appendix Table C1

Simulation Results for Three Representative Occupations for Male Manual Workers

	Large Occupation Carpenters and Joiners	Medium Occupation Chefs, Cooks	Small Occupation Sewers and Embroiders
1980 census code	104.1 & 105.1	063.1 & 064.0	102.3
Parameter Settings			
<i>n</i> in 1980	18807	6098	781
<i>n</i> in 1990	14216	14216	855
illmprob	0.014349	0.009893	0.021422
Mean duration	10	8	6
Simulation Results			
	Mean, Standard Dev.	Mean, Standard Dev.	Mean, Standard Dev.
risk constant	0.01434, 0.00031434	0.0099064, 0.00039796	0.021454, 0.0020824
risk from 0 to 2 × illmprob	0.014344, 0.00030479	0.0098865, 0.00041782	0.021419, 0.0021014
risk from			

0.5 × illmprob to 1.5 × illmprob	0.014353, 0.00031169	0.0098785, 0.00040809	0.021485, 0.0020528
--	-------------------------	--------------------------	---------------------

The Figure C1 below shows one of the sampling distribution of Illmprob over 1000 simulations for the occupation carpenters. The sampling distribution is clearly approximately normal. This normality result was not sensitive to any variation of the parameters within the range of the values in the data or the assumptions of the gradient in risk.

The program used for the simulations was RESAMPLING STATS. Samples of the coding for particular occupations are available from the authors.

Figure C2

Sampling Distribution of IIMprob over 1000 Simulations for the Occupation

Carpenters

