Reproducing ONS's "Figure 11" from their published regression coefficients

The regression models, which include up to 221 terms, can be written as follows,

$$\ln(pay\ rate) = f(variables) + male * g(age) \quad (1)$$

Here the variable "*male*" takes the value 1 for men or the value 0 for women.

The function *g(age)* depends only on the variable "age" and is a quadratic,

$$g(age) = A + B * age + C * age^2 \quad (2)$$

The function $f(variables)$ is independent of the variable "*male*" (though it can depend upon "*age*"). From (1) we get,

$$pay\ rate = e^{f+mal\ *g} \quad (3)$$

i.e., *male pay rate* = e^{f+g} (4)

and, *female pay rate* = e^{f} (5)

Hence, *gender pay rate gap* = $\frac{e^{f+g}-e^{f}}{e^{f}} = e^{g} - 1$ (6)

Note that this gender gap has been defined as a percentage of the female pay rate (in the denominator) to be consistent with ONS's "Figure 11" (though inconsistent with usage elsewhere in ONS reports).

The utility of Equ.(6) is that the gender pay rate gap can be evaluated from *g* alone. It is not necessary to evaluate the (possibly very large number of) terms occurring in *f*. The regression coefficients are,

	A	**B**	**C**
Model 4	-0.237	0.0143	-0.000123
Model 6	-0.176	0.0108	-0.0000782

Hence the gender pay rate gap derived from the regression is…

Age	**Model 4**	**Model 6**
20	0%	0.9%
25	4.46%	4.62%
35	11.9%	11.2%
45	17.1%	16.37%
55	19.4%	19.89%
65	18.9%	21.60%