Supplementary material to "Orthogonal Gaussian process models": Expanded results for Section 4.2

This section can be considered additional information for section 4.2 in the paper "Orthogonal Gaussian process models."

	n	MSPE	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_4$	$\hat{\beta}_6$	$\hat{\beta}_8$
OGP	20	75.0 [42.0]	77.8 [1.2]	72.1 [2.8]	0.2 [3.1]	-0.1 [2.5]	-15.2 [3.0]
	40	10.5 [3.3]	77.6 [0.4]	71.4 [0.8]	0.0 [0.7]	0.4 [0.9]	-15.6 [0.8]
	80	3.2 [0.9]	77.7 [0.2]	71.5 [0.3]	0.0 [0.3]	0.1 [0.3]	-15.5 [0.3]
	160	0.8 [0.3]	77.7 [0.1]	71.5 [0.1]	-0.0 [0.1]	0.1 [0.1]	-15.6 [0.1]
UK	20	155.7 [65.8]	80.2 [3.5]	72.9 [4.9]	0.4 [4.7]	-0.0 [4.3]	-15.4 [5.8]
	40	16.8 [15.3]	84.7 [3.6]	72.5 [4.1]	0.2 [1.4]	0.7 [1.5]	-16.4 [2.5]
	80	3.6 [1.0]	87.6 [3.6]	73.6 [3.4]	-0.0 [1.3]	0.6 [1.5]	-17.5 [2.6]
	160	0.8 [0.2]	92.6 [2.5]	75.1 [3.1]	0.2 [1.4]	0.1 [1.4]	-18.3 [1.9]
LS	20	155.4 [53.3]	78.1 [1.7]	72.6 [3.9]	-0.1 [5.0]	-0.4 [5.0]	-16.2 [5.8]
	40	18.4 [19.9]	77.9 [1.1]	71.5 [2.4]	1.1 [3.1]	1.0 [3.3]	-15.6 [2.9]
	80	3.6 [1.0]	77.9 [1.1]	71.8 [1.6]	-0.3 [2.1]	0.3 [2.2]	-15.7 [2.2]
	160	0.8 [0.3]	77.8 [0.6]	71.6 [1.2]	-0.1 [1.4]	-0.2 [1.4]	-16.1 [1.6]

Table 1: Table of the results from the numerical example in Section 4. The values in the $\hat{\beta}_i$ columns are the sample mean of 50 simulations with the sample standard deviation in square brackets. The MSPE column contains the the sample mean of 50 simulations with the sample standard deviation in square brackets (MSPE approximated with 1000 uniformly drawn points).

The results of the numerical simulation are in Table 1 which shows both the sample mean and standard deviation across the 50 replicates. Not all the $\hat{\beta}_i$s were shown for the sake of brevity. Measured by the mean squared prediction error, it appears that all methods have similar prediction accuracies at sample sizes 80 and 160, but the orthogonal approach seems slightly to be better for the smaller sample sizes. The similarity might be due to the fact that these models are similar with the exception of $g(x)$. As expected based on our discussion, the universal kriging method cannot identify the mean parameters well. Using universal kriging,

$\hat{\beta}_2$ has only a small reduction in standard deviation as the sample size is increased to 160, indicating a lack of convergence. When $n = 160$, the standard deviation of the estimate of $\hat{\beta}_2$ using orthogonal Gaussian processes is 10 times smaller than the estimate using universal kriging. The results using universal kriging are significantly worse for $\hat{\beta}_1$, which appears to be centered at different values depending on the sample size. While the least squares estimates appear to be converging to reasonable values, their convergence is slow relative to the orthogonal Gaussian process model. From a modeling standpoint, this may be because the least squares estimates do not incorporate the smoothness of the response.