

A sensitivity analysis of probabilistic sensitivity analysis in terms of the density function for the input variables

Peer-reviewed author version

De Mulder, Wim; MOLENBERGHS, Geert & VERBEKE, Geert (2017) A sensitivity analysis of probabilistic sensitivity analysis in terms of the density function for the input variables. In: JOURNAL OF STATISTICAL COMPUTATION AND SIMULATION, 87(7), p. 1429-1445.

DOI: 10.1080/00949655.2016.1270280

Handle: <http://hdl.handle.net/1942/26397>

To appear in the *Journal of Statistical Computation and Simulation*  
Vol. 00, No. 00, Month 20XX, 1–18

## A sensitivity analysis of probabilistic sensitivity analysis in terms of the density function for the input variables

Wim De Mulder<sup>a\*</sup>, Geert Molenberghs<sup>b,a</sup> and Geert Verbeke<sup>a,b</sup>

<sup>a</sup>*Leuven Biostatistics & Statistical Bioinformatics Centre (L-BioStat), Leuven, Belgium;*

<sup>b</sup>*Interuniversity Institute for Biostatistics and Statistical Bioinformatics (I-BioStat), Hasselt, Belgium*

(Received 00 Month 20XX; final version received 00 Month 20XX)

Probabilistic sensitivity analysis allows to incorporate background knowledge on the considered input variables more easily than many other existing sensitivity analysis techniques. Incorporation of such knowledge is performed by constructing a joint density function over the input domain. However, it rarely happens that available knowledge directly and uniquely translates into such a density function. A naturally arising question is then to what extent the choice of density function determines the values of the considered sensitivity measures. In this paper we perform simulation studies to address this question. Our empirical analysis suggests some guidelines, but also cautions to practitioners in the field of probabilistic sensitivity analysis.

**Keywords:** probabilistic sensitivity analysis; agent-based models; Gaussian process emulation; mean effect; sensitivity index

### 1. Introduction

Sensitivity analysis (SA) is an invaluable tool in studying certain relationships between input variables and output variables. A wide range of techniques have been developed in the SA domain, each of which measures a certain relationship between input and output in a different way [1]. *Probabilistic sensitivity analysis* is becoming increasingly popular, as it has some desirable features not possessed by many other SA techniques, as discussed in Section 3.1 below. Especially in situations where some information is available about the correctness or optimality of certain values of the input variables, the use of probabilistic SA is advantageous as this technique allows to take into account that one value is more plausible than another one. This is done by defining a density function over the input space. However, in virtually all practical applications the available information on the plausibility of the values belonging to the considered domain is insufficient to be transformed into a density function in a unique way. It is custom among practitioners to solve this ambiguous task by just taking *any* density function, often uniform or Gaussian, that matches or at least is not contradictory to the available information (see, e.g., [2]). However, this rests on the implicit assumption that the measures calculated with probabilistic SA are insensitive to the choice of density function among all density functions that are reasonable for the problem at hand. This might not be true and in a

---

\*Corresponding author. Email: wim.demulder@cs.kuleuven.be

worst case scenario the sensitivity of the calculated measures with respect to the choice of density function might be higher than the sensitivity of the output with respect to the input, making the results of the sensitivity analysis useless.

In this paper we empirically investigate the influence of the choice of density function on the mean effect and the sensitivity index, two important measures in probabilistic SA (see Sections 3.2 and 3.3), via two simulation studies where limited information about the density of the input values is available, conform to many practical situations. The first simulation study entails a benchmark function and the second one is a real world application in the context of agent-based models (agent-based models are shortly reviewed in Section 4.1).

## 2. Outline of the paper

Section 3 provides an introduction to probabilistic SA and describes the mean effect and the sensitivity index, the two SA measures that will be calculated for our simulation studies. In Section 4 we discuss the methods that will play a role in our simulation studies. First, we consider agent-based models, as one of our simulation studies performs a probabilistic sensitivity analysis of the parameters of an agent-based model that we have developed. Secondly, Gaussian process emulation, as agent-based models are very time-consuming and to reduce computation time to an acceptable level we approximated our agent-based model with a Gaussian process emulator, which is a member of the class of surrogate models [3]. Finally, Monte Carlo integration, since the process of probabilistic SA is in itself also very time-consuming due to the computation of integrals, and Monte Carlo integration is used to speed up this process. Sections 5 and 6 are the main sections of this paper, where we empirically investigate the influence of the choice of density function on the two considered probabilistic SA measures in two completely different simulation studies, namely one where the model is a benchmark function (Section 5) and one where the model is an agent-based model that simulates mechanisms in assortative mating (Section 6). Section 7 summarizes the main findings.

## 3. Probabilistic sensitivity analysis

This section is strongly based on the material on the MUCM toolkit website<sup>1</sup>.

### 3.1. Motivation for probabilistic sensitivity analysis

Sensitivity analysis began as a response to concerns about the consequences of mis-specifying the values of inputs to a computer model, also called simulator,  $f$  which is assumed here to map vectors to real numbers. Since the true input  $\tilde{\mathbf{x}}$  is typically unknown, the user makes use of heuristics to find a good estimate  $\hat{\mathbf{x}}$  of the true input. The value  $f(\hat{\mathbf{x}})$  can then be regarded as a good estimate of  $f(\tilde{\mathbf{x}})$  *provided that* the output does not change too quickly with varying inputs. That is, if the output is not sensitive to any input component. The output is said to be sensitive to a particular input component  $x_j$  if the output changes substantially when  $x_j$  is perturbed slightly. This led to the idea of measuring sensitivity by differentiation of the function. The measure of sensitivity to  $x_j$  is then the derivative  $\partial f(\mathbf{x})/\partial x_j$ , evaluated at  $\mathbf{x} = \hat{\mathbf{x}}$ .

---

<sup>1</sup><http://mucm.aston.ac.uk/toolkit/index.php?page=ThreadTopicSensitivityAnalysis.html>

SA based on derivatives has a number of deficiencies, however. The differential measures only the impact of an *infinitesimal* change in  $x_j$ , and for this reason this kind of SA is referred to as *local SA*. If the response of the output to  $x_j$  is far from linear, then perturbing  $x_j$  more than a tiny amount might have an effect that is not well represented by the derivative. Furthermore, the derivative is not invariant to the units of measurement. If, for instance, we choose to measure  $x_j$  in kilometers rather than meters, then  $\partial f(\mathbf{x})/\partial x_j$  will change by a factor of 1000, giving the false impression that the sensitivity has dramatically increased.

A simple alternative to derivative-based SA involves perturbing  $x_j$  from its nominal value  $\hat{x}_j$ , say to a value  $x'_j$ , with all other input components held at their nominal values. The resulting change in output is then regarded as a measure of sensitivity to  $x_j$ . This is known as *one-way SA*, because the inputs are varied only one at a time from their nominal values. One-way SA addresses some problems noted for local SA. First, we do not consider only infinitesimal perturbations. Secondly, nonlinearities in the response to  $x_j$  are automatically accounted for since these nonlinearities are encapsulated in the function evaluations. Thirdly, changing the units of measurement for  $x_j$  has no influence on the output and thus on the change of output neither. However, one-way SA has the severe limitation that it depends on how far we perturb the considered input component and in practical situations there is typically not a definite value for the magnitude of perturbation that can be considered as most relevant. Furthermore, one-way SA fails to quantify joint effects of perturbing more than one input at the same time. The effect of perturbing  $x_1$  and  $x_2$  together cannot be inferred from knowing the effects of perturbing them individually. An obvious solution is to detect interactions by perturbing more than one input component at the same time, known as *multi-way SA*. However, the problem remains that the sensitivity still depends on the magnitude of the perturbations of the considered input components.

*Probabilistic SA* avoids the arbitrary choice of a value for the magnitude of perturbation by considering a distribution for the inputs. A joint density function  $\omega$  over the input space is chosen by the user, such that the value  $\omega(\mathbf{x})$  in an input  $\mathbf{x}$  can be interpreted as a *weight* assigned to that input value. Typically one will choose  $\omega$  such that for an input value  $\mathbf{x}$  that is more realistic or more likely than another input value  $\mathbf{x}'$  it holds that  $\omega(\mathbf{x}) > \omega(\mathbf{x}')$ . Consequently, probabilistic SA does not rely on arbitrary choices of magnitudes of perturbation, since it takes into account *all* relevant perturbations. Several measures have been developed that then summarizes the corresponding changes in output by taking a *weighted average* of all these changes, where the weights are determined by  $\omega$ . In this paper we focus on two important probabilistic SA measures, namely the mean effect and the sensitivity index. The next two sections briefly review these two measures in a non general way, assuming two input variables and one output variable, as this applies to our case studies.

### 3.2. The mean effect

The mean effect is the weighted average of the simulator output, where the average is taken over the values of one of two input variables  $x$  and  $y$ , with the other variable fixed. Let us say that  $x$  is given a fixed value  $x_0$ . Then we define the corresponding mean effect  $M_1(x_0)$  as

$$M_1(x_0) = E[f(x, y)|x = x_0] = \int f(x_0, y) \omega_{2|1}(y|x_0) dy \quad (1)$$

Thus given a certain value  $x_0$  for  $x$ ,  $M_1(x_0)$  returns the weighted average of the simulator output over all input values for which  $x = x_0$ . The weights are provided by the conditional density  $\omega_{2|1}$ .

The interpretation of the mean effect is that it allows to analyze the influence of  $x$  to the simulator output. To that end, we consider repeatedly different values for  $x$  and for each of these values we compute the (weighted) average simulator outcome with respect to the other input variable. Another way to obtain insight into the influence of  $x$  on the output would be to choose a value  $y_0$  for  $y$  and then calculate the simulator output for varying values of  $x$ , while keeping  $y$  at  $y_0$ . However,  $M_1$  would then depend on the (often arbitrary) choice of value for  $y$ , resembling the one-way and multi-way sensitivity analysis approaches and their related deficiencies as outlined in Section 3.1. In probabilistic SA we avoid an arbitrary choice of a value for an input variable that is left out of consideration by computing the simulator outcome for *each* value, and then averaging these outcomes according to the weight attached to each of these values, as given by  $\omega_{2|1}$ . This is then repeated for different values of  $x$ , thereby providing insight into how the simulator outcome varies as  $x$  varies. The definition of  $M_2$  is, of course, completely similar.

### 3.3. The sensitivity index

The value provided by  $M_1$  or  $M_2$  depends on the value that we have assigned to the input variable of interest. For many purposes, it is helpful to have a single summary of how sensitive the output is to a given input variable, so that we have a definite answer to the question whether the input variable has a large effect on the output. The *sensitivity variance*  $V_1$  measures the magnitude of the influence of  $x$  to the output as follows:

$$V_1 = \int (M_1(x) - M)^2 \omega_1(x) dx \quad (2)$$

with  $M = E[f(x, y)]$ . To interpret this measure, we first notice that  $M$  is the weighted average of the simulator outcome. If we would be asked to provide an estimate for the simulator outcome in an unknown input point, then  $M$  would be our best estimate. Now suppose that the true value of  $x$  is known to be  $x_0$ , while the value of  $y$  is unknown. What would then be the best estimate of the simulator outcome? It would be the (weighted) average of all simulator outcomes over all inputs of the form  $(x_0, y)$  where  $y$  varies over all values of its domain. As we know from Section 3.2 this is precisely the value  $M_1(x_0)$ . The squared difference  $(M_1(x) - M)^2$ , which appears in the definition of the sensitivity variance (2), tells us how large the difference is between the estimate for the simulator outcome in case we know the value of  $x$  and in case we do not. The measure  $V_1$  then takes into account that  $x$  can potentially take *any* value in its domain by calculating this difference with respect to each value and then taking the weighted average of it. This description makes clear that if  $V_1$  is large, the estimate of the simulator outcome with or without having information about  $x$  is very different. In other words: the larger  $V_1$  is, the more important it is to have information about the true or most likely values of  $x$ . Simply speaking, the larger  $V_1$ , the more important the variable  $x$ . The definition of  $V_2$  is, again, completely similar.

The sensitivity variance  $V_1$  is in units that are the square of the units of the simulator output, while it is common to measure sensitivity by a dimensionless index. To that end,

the *sensitivity index*  $S_1$  has been introduced:

$$S_1 = V_1/V \quad (3)$$

with

$$V = \int (f(x, y) - M)^2 \omega(x, y) dx dy \quad (4)$$

The sensitivity index  $S_1$  also indicates the importance of  $x$ , as it is a kind of a normalized version of  $V_1$ . It is easily shown that  $0 \leq S_1 \leq 1$ . On the MUCM toolkit website, it is stated that, e.g.,  $S_1 = 0.5$  indicates that uncertainty about  $x$  accounts for 50% of the overall uncertainty in the outcome when no information about the true values of any of the input variables is available. Or, differently stated, that the uncertainty in the output would be reduced by 50% if we would know the most likely or true value of  $x$ . However, it is important to appreciate that, in general, it does not hold that  $S_1 + S_2 = 1$ . The reason is that each sensitivity index restricts attention to one input variable, so that the interaction between these variables is not taken into account. In general we have that  $S_1 + S_2 < 1$ . It is thus more correct to state that uncertainty in the output would be reduced by *at least*  $(100 \times S_1)\%$  if  $x$  would be known. Due to the correlation between  $x$  and  $y$ , knowing the true value of  $x$  will also provide some information about the true value of  $y$ , thereby also reducing the uncertainty about the true value of  $y$  which reduces in turn the uncertainty about the output further. The definition of  $S_2$  is analogous.

## 4. Methods

### 4.1. Agent-based models

An agent-based model (ABM) is a computer model that simulates the behavior and interactions of agents [4]. An agent is the representation of a main entity of the system under study, e.g. individuals in a study of marriage markets, or couples in a study of fertility decisions. Each agent is given certain characteristics that are relevant in the context of the study. For example, the characteristics of an agent representing an individual on the marriage market might be gender, age, and characteristics of a desirable partner. Theoretical and empirical considerations guide the assignment of values to each characteristic. Thus, e.g., one might assign partner preferences based on existing survey data. After specifying the agents, one has to decide on rules that govern the interaction between them. Such rules are also to be derived from theoretical and empirical background knowledge, although simplifications are often necessary in order to control the complexity of the model. A simple rule in case of modeling fertility decisions is that women adopt a particular birth control strategy when a certain share of the other women in their personal network have also adopted this strategy. Having specified the ABM, one let the agents interact in a simulation environment, thereby typically changing the assigned values of the characteristics. This is then repeated a number of times, until a predefined number of simulation steps is reached or a stable state is attained. If the resulting simulation output is close to empirical data, hypotheses to explain the real world behavior can be obtained in terms of the assigned characteristics of and interactions between the agents.

ABMs are especially useful when substantial background knowledge about *individual* behavior is available, but where the resulting *macro level* behavior is not known or

not understood well. That is, in cases where knowledge of the parts does not imply understanding of the whole. As an example, it has been observed that even when all drivers try to drive fluently, vehicles are sometimes stopped by phantom traffic jams, i.e. jams without an obvious reason such as an accident or bottleneck [5, 6]. The social sciences are a prime example of a domain where ABMs have proven to be a major methodology to analyze dynamic systems [7, 8].

#### 4.2. *Gaussian process emulation*

Gaussian process emulation [9], or emulation for short, is a subclass of surrogate modeling [3], where the objective is to obtain a fast-running approximation to a complex, time-consuming model. The surrogate model in emulation is conveniently called the emulator, while the time-consuming model is often referred to as the simulator.

The emulator is determined via a Bayesian analysis as follows. In the first step the output of the simulator in any point is modeled as a Gaussian distribution. The mean of this distribution is modeled as a linear combination of user-chosen regression functions applied to the given input, while the variance is a constant and expresses the uncertainty about the true value of the simulator. Output values in different points are typically not independent, as output values corresponding to input values that are close to each other are also close to each other in most practical applications. This is taken into account by modeling the correlation between output values. The resulting prior model is a Gaussian process, i.e. any finite collection of output values is modeled as a multivariate normal distribution. In a second step, training data are obtained by running the simulator on selected input points. Selection of appropriate training data input points is often done via the Latin hypercube method [10]. In a third and final step the parameters of the emulator, consisting of the constant variance and the coefficients in the linear combination of the regression functions, are optimized, typically via maximum likelihood. Using Bayes' rule, the Gaussian distribution in each output value is updated to a Student's t-distribution, although for all practical applications this distribution is approximated by a Gaussian distribution. For each input vector  $\mathbf{x}$  the emulator then provides a distribution for the corresponding output value. The mean of this distribution is denoted as  $m(\mathbf{x})$  and is considered as approximation to the unknown (except for training data points) output of the simulator in  $\mathbf{x}$ . The variance of this distribution can be used to obtain a credible interval around  $m(\mathbf{x})$ . The larger this credible interval, the more likely that  $m(\mathbf{x})$  is not a good approximation to the simulator output in  $\mathbf{x}$ , and vice versa.

#### 4.3. *Monte Carlo integration*

To explain the basic principles of Monte Carlo integration, we assume that the integral  $I$  of a function  $F : \mathbb{R} \rightarrow \mathbb{R}$  over an interval  $[a, b]$  has to be determined numerically. It is basic knowledge [11] that for a large class of functions,  $I$  can be approximated by

$$\hat{I}_n = \sum_{i=1}^n F(x_i)(x_i - x_{i-1}) \quad (5)$$

where  $x_0 = a < x_1 < \dots < x_n = b$ , for sufficiently large  $n$ . We can take  $x_i - x_{i-1}$  as a constant, namely  $(b - a)/n$ . Then:

$$\hat{I}_n = \frac{b - a}{n} \sum_{i=1}^n F(x_i) \quad (6)$$

Thus we find that  $I$  can be approximated by the average of values of  $F$  in chosen input points multiplied by the length of the interval. Notice that the factors  $x_i - x_{i-1}$ , which are present in (5), have been removed in the above approximation and thus it is not required anymore that  $x_{i-1} < x_i$ . The basic Monte Carlo method chooses the input points  $x_i \in [a, b]$  according to the uniform distribution over  $[a, b]$ .

This method is easily extended to higher dimensions.

Some important advantages of Monte Carlo integration are outlined in [12]:

- The Monte Carlo technique typically outperforms other numerical approximation methods when the input domain is higher dimensional. The efficiency of Monte Carlo integration relative to other methods increases when the dimension of the problem increases. For example, quadrature formulae [13] become very complex while the Monte Carlo integration technique remains almost unchanged in more than one dimension.
- The convergence of Monte Carlo integration is independent of dimensionality regardless of the smoothness of the integrand.
- Monte Carlo integration is simple since only two basic operations are required, namely sampling and function evaluation.

All integrals in the simulation studies below have been calculated with the Monte Carlo method.

## 5. Simulation study of probabilistic sensitivity analysis for benchmark function

In our first simulation study the simulator is a real valued benchmark function with two input variables having one global optimum. In a typical application, the analytical expression of the function would not be known, and therefore we will not make use of information about sensitivity of the output on the inputs that can be derived from the analytical formulation. On the other hand, it is not uncommon to have a situation where inputs corresponding to the optimal value are at the same time also more likely to occur in the real world application at hand, and where information is available about the location of such 'optimal inputs' (see, for example, the next simulation study). Therefore, we incorporate the location of the optimum in our probabilistic sensitivity analysis, but not any other information.

### 5.1. Description of benchmark function

The benchmark function we consider is the so-called Leon function [14]:

$$f(x, y) = 100(y - x^2)^2 + (1 - x)^2 \quad (7)$$

with  $-1.2 \leq x, y \leq 1.2$ . A global minimum is located at  $(1, 1)$  with  $f(1, 1) = 0$ .



## 5.2. Description of density functions

We consider two classes of density functions for the input variables. The first class is the normal density function where the mean is located at (1,1), the point that corresponds to the unique global optimum. The standard deviation is given varying values, namely 0.5, 1, 1.5, 2, 2.5 and 3. The correlation is chosen either as 0 or 0.5. This gives  $6 \times 2 = 12$  Gaussian density functions. The second class is the uniform density function. Although this function does not take into account any available information, it is useful as a 'baseline density function', i.e. a kind of reference for comparison of our research findings. For example, although we hope that SA measures do not vary much with the choice of density function as long as the density function is compatible with the available information, we also hope that the measures *do* vary noticeably when changing such a density function into the uniform density function, since otherwise this would imply that probabilistic SA is not able to incorporate available information.

## 5.3. Results

Figures 1 and 2 show the results for  $M_1$  and  $M_2$ . Three important observations can be made. First, the difference between the measures for the Gaussian density functions appear to become larger for values of the considered input variable closer to the boundary. Except for  $M_1$  in case the correlation is 0, the lowest value of both measures over all Gaussian density functions is around 100 when the value of the considered input variable is -1.2, while the highest value in -1.2 is around 250, i.e. 2.5 times larger. Secondly, the measures differ notably depending on whether a Gaussian density or the uniform density is used. The difference is especially striking for  $M_1$ . Comparing the Gaussian densities with the uniform density, it is again observed that the difference is more noticeable near the end points of the interval. A third observation is that although the exact values of  $M_1$  and  $M_2$  depend on the chosen density function, the general *trend* of both measures for varying input values is rather independent of this choice. That is,  $M_1$  decreases until around  $x = 0$  and then increases again, except for the Gaussian density with standard deviation 0.5, while  $M_2$  sharply decreases until around  $y = 0.4$  and then increases slightly.

Figures 3 and 4 display  $S_1$  and  $S_2$ . It is clear that both  $S_1$  and  $S_2$  are remarkably larger in case the uniform density is used. Restricting attention to the Gaussian densities, the difference between the values of  $S_1$  is acceptable, with the smallest value of  $S_1$  being about 0.1 (density with standard deviation equal to 1 and correlation equal to 0) and the largest value of  $S_2$  being about 0.17 (density with standard deviation equal to 0.5 and correlation equal to 0.5). However, the difference between the values of  $S_2$  are much larger, with the smallest value being about 0.1 (density with standard deviation equal to 0.5 and correlation equal to 0), and the largest value being about 0.37 (density with standard deviation equal to 1 and correlation equal to 0.5).

It is also important to evaluate  $S_1/S_2$ , since the ratio of both measures is an indication of the relative importance of both input variables. Figure 5 shows this ratio. The difference between its value for the Gaussian density function with standard deviation equal to 0.5, especially when the correlation equals 0.5, and the other density functions is very striking. While all values clearly indicate that the second input variable is much more important than the first one (in the sense that  $S_1/S_2 < 1$ ), the value in case of the Gaussian density with standard deviation 0.5 indicates the opposite.

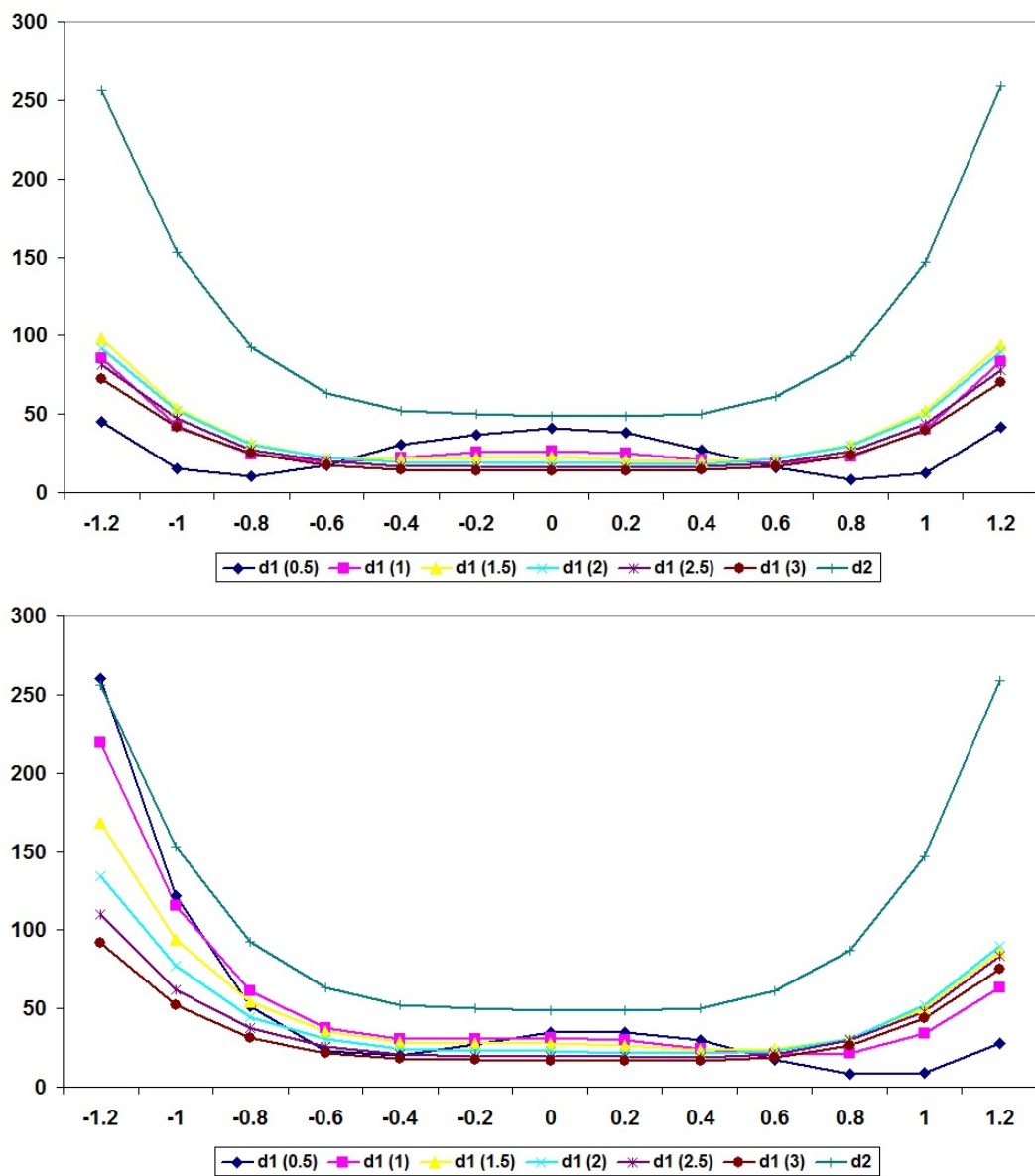


Figure 1.  $M_1$  for benchmark function (above: correlation = 0, below: correlation = 0.5),  $d_1$  refers to Gaussian density with standard deviation between parentheses,  $d_2$  refers to uniform density

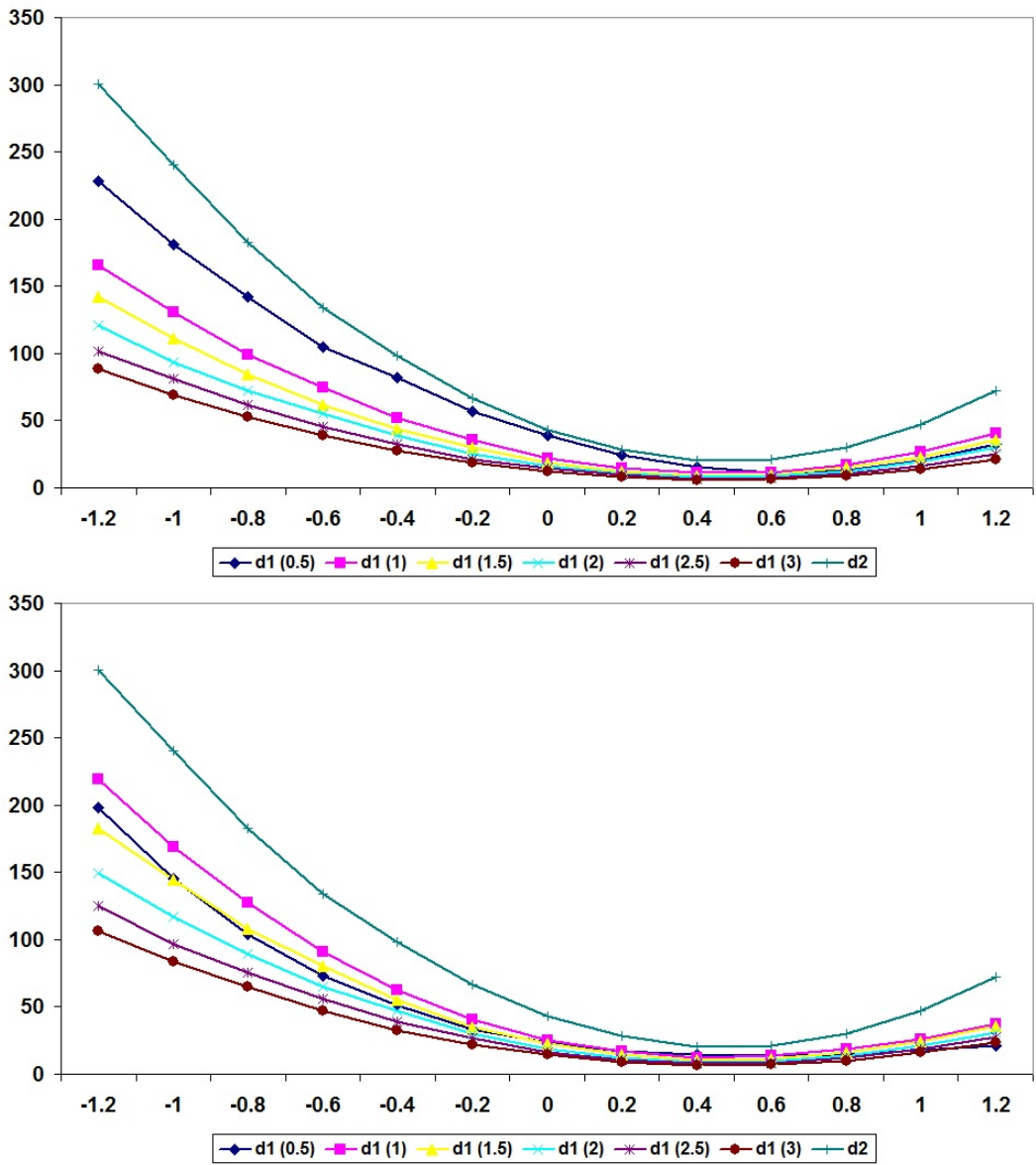


Figure 2.  $M2$  for benchmark function (above: correlation = 0, below: correlation = 0.5),  $d_1$  refers to Gaussian density with standard deviation between parentheses,  $d_2$  refers to uniform density

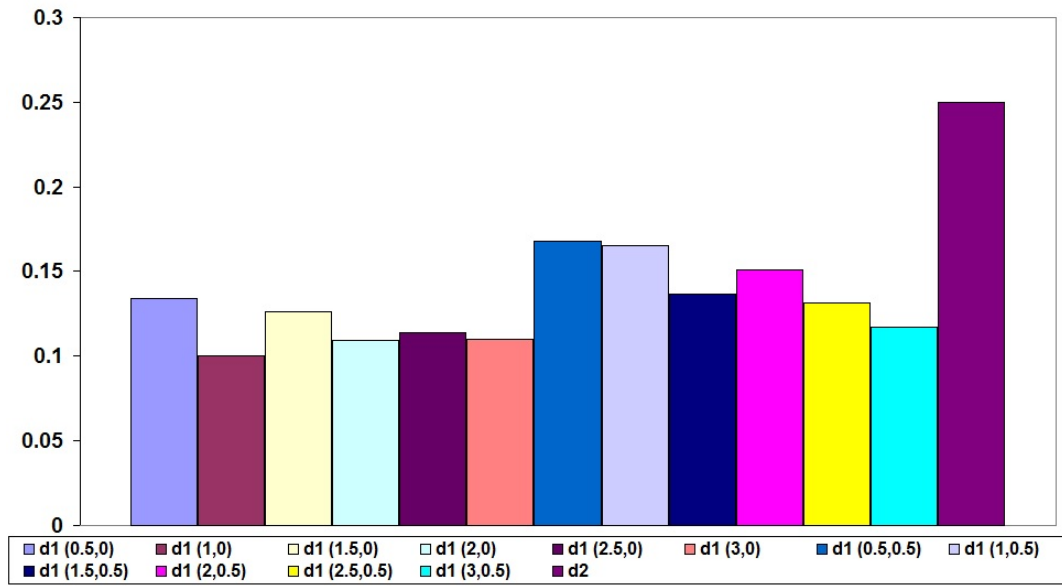


Figure 3.  $S_1$  for benchmark function,  $d_1$  refers to Gaussian density with standard deviation and correlation resp. between parentheses,  $d_2$  refers to uniform density

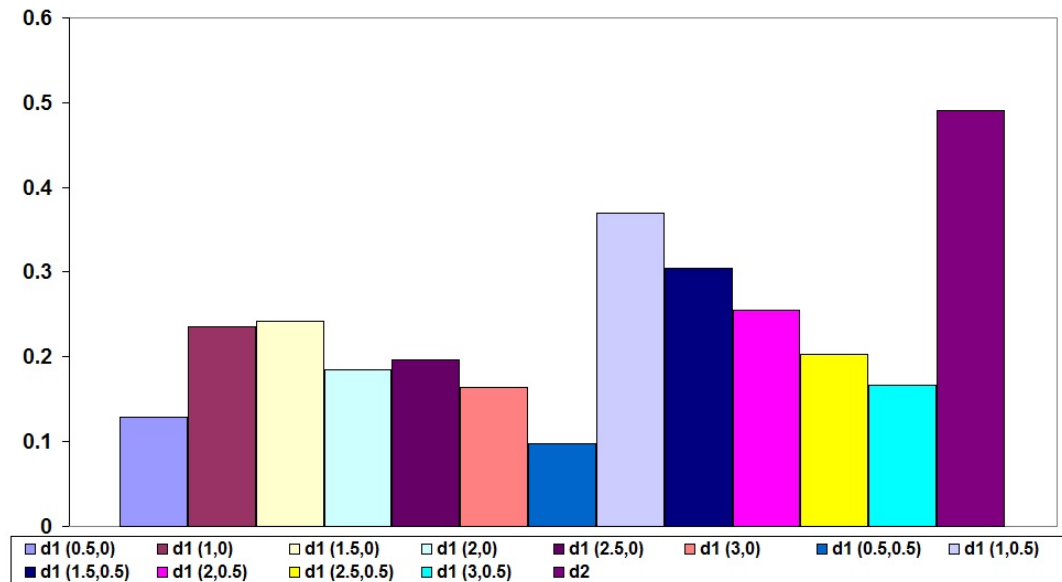


Figure 4.  $S_2$  for benchmark function,  $d_1$  refers to Gaussian density with standard deviation and correlation resp. between parentheses,  $d_2$  refers to uniform density

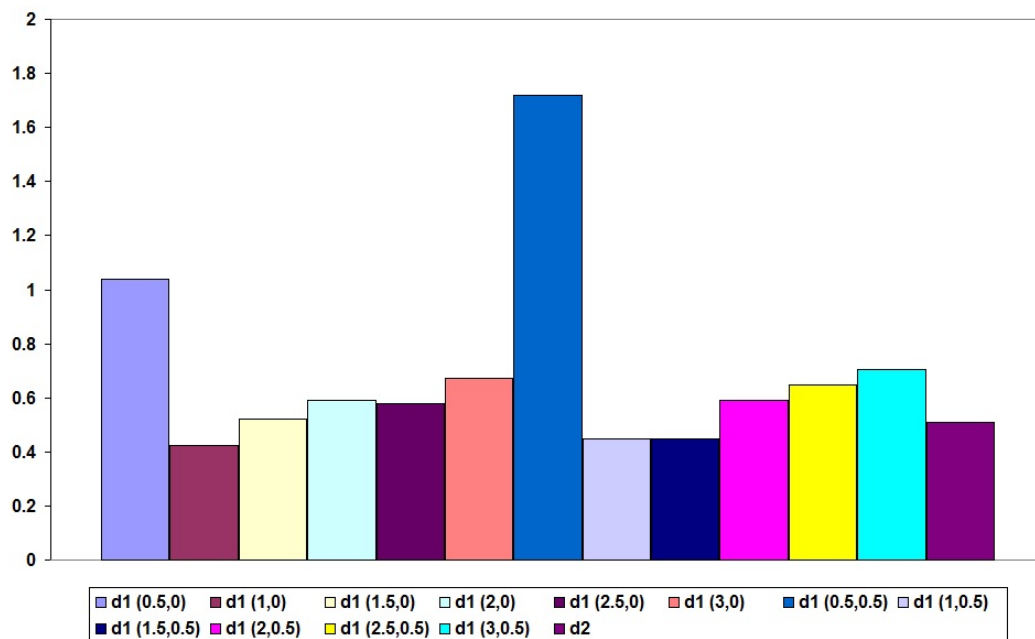


Figure 5.  $S1/S2$  for benchmark function,  $d_1$  refers to Gaussian density with standard deviation and correlation resp. between parentheses,  $d_2$  refers to uniform density

## 6. Simulation study of probabilistic sensitivity analysis for agent-based model

### 6.1. Description of our agent-based model

We have developed an ABM that simulates the mechanisms that might have linked the reversal of gender inequality in higher education with changes in patterns of educational assortative mating in Belgium [15]. Educational assortative mating refers to the sorting of individuals into relations based on educational attainment. Empirical evidence shows a reversal of gender inequality in educational attainment from the 1970s on. That is, while men have almost always received more education than women before that turning point, this imbalance has turned around in large parts of the world. In many countries, women now outperform men in participation and success in higher education [16]. Research indicates that this reversal affects patterns of assortative mating [17, 18].

We do not discuss the specification of the characteristics of and the interaction between the agents, as this has been described extensively in our previous work. What is important for the purpose of our current work is that our ABM comprises two parameters, denoted as  $w_s^f$  and  $w_s^m$ , that determine the importance that female and male agents attach to the education of prospective partners. The higher their values, the more agents prefer partners with similar educational attainment, and the more willing they become to marry similarly educated agents. The allowable values of these parameters are the interval  $[0,2]$ . The simulation outcome is the fraction of hypogamic couples, i.e. couples in which the woman has a higher educational attainment than her partner. For given values  $x$  and  $y$  of  $w_s^f$  and  $w_s^m$  resp. we denote this outcome as  $f(x, y)$ .

In previous work optimal values for the two parameters were determined by using a genetic algorithm that finds the parameter values for which the corresponding simulation

output is closest to the empirical value of the fraction of hypogamic couples. Interestingly, it was found that there is not a finite number of optimal values, but that the optimal values of  $w_s^f$  and  $w_s^m$  lie along a straight line  $L$  given by

$$w_s^m = -1.2034 w_s^f + 2.4360 \quad (8)$$

as described in [19].

Since our ABM is very time consuming, we approximate the ABM by an emulator. The details of the construction of the emulator are discussed in [19].

## 6.2. Description of density functions

We consider three classes of density functions.

To construct the first class we start by considering a function  $w$  that is constant on the line segment  $S$  that arises by taking the intersection of the line (8) and the square  $[0, 2] \times [0, 2]$ , and decreases linearly with increasing distance  $d$  from  $L$  for points within  $[0, 2] \times [0, 2]$  and that is zero outside this square. Since points outside  $[0, 2] \times [0, 2]$  will be given weight 0, it is appropriate to give the point that belongs to  $[0, 2] \times [0, 2]$  and that is furthest away from  $L$  weight 0. To find this point we notice the following. Since  $S$  lies above the diagonal from  $(0,2)$  to  $(2,0)$  it is clear that this point lies in the triangle formed by the 3 points  $(0,0)$ ,  $(0,2)$ ,  $(2,0)$ . All points on lines parallel to  $L$  are at the same distance of  $L$ . Thus the point that is sought is on a line parallel to  $L$  and belongs to the considered triangle. Moving such a parallel line away from  $L$  increases the distance to  $L$ . The line that is parallel to  $L$  and that goes through  $(0,0)$  is the parallel line that is furthest away from  $L$  while still containing a point from  $[0, 2] \times [0, 2]$ . The sought point is thus  $(0,0)$  and the distance from it to  $L$  is  $d((0,0), L) \equiv \gamma \approx 1.557$ .

We then define  $w$  as

$$\begin{aligned} w(x, y) &= -d((x, y), L) + \gamma && \text{if } (x, y) \in [0, 2] \times [0, 2] \\ &= 0 && \text{otherwise} \end{aligned}$$

A joint density function  $\omega$  is obtained by normalizing  $w$  as follows:

$$\omega(x, y) = 1/c w(x, y) \quad (9)$$

with  $c$  such that

$$c \int_0^2 \int_0^2 w(x_0, y_0) dx_0 dy_0 = 1$$

For the second class we start from a function  $w$  that is constant on  $S$  and that decreases exponentially with increasing distance from  $S$ :

$$\begin{aligned} w(x, y) &= \exp(-\beta d(x, y)) - \exp(-\beta \gamma) && \text{if } (x, y) \in [0, 2] \times [0, 2] \\ &= 0 && \text{otherwise} \end{aligned}$$

where  $\beta > 0$  is a parameter. In our simulation study we assign values 0.5, 1, 1.5, 2 and 2.5 to  $\beta$ . The term involving  $\gamma$  is to ensure that  $w(0,0) = 0$ , as in the case of the first

class of density functions. A density function is then obtained from  $w$  in the same way as for the first class.

The third density function is, again, the uniform density.

### 6.3. Results

Figures 6 and 7 show  $M_1$  and  $M_2$ . The values for  $M_1$  and  $M_2$  now appear to be rather independent of the choice of density function. Differences are again larger as values move towards the endpoints of the interval.

Figure 8 shows that the differences between the values of  $S_1$  are moderate for the several density functions, although the difference between the smallest value of about 0.38 and the largest value of about 0.61 is far from negligible. All values of  $S_2$  are rather similar, as seen from Figure 9, except the one for the uniform density function which is significantly larger than the other values. Figure 10 shows the ratio of  $S_1$  and  $S_2$ . Although the values vary from about 1.6 (uniform density) to 2.7 (second class of density functions with  $\beta = 1.5$ ) they all point to the same conclusion, namely that the first input variable, i.e.  $w_s^f$ , is much more important than the second input variable, i.e.  $w_s^m$ .

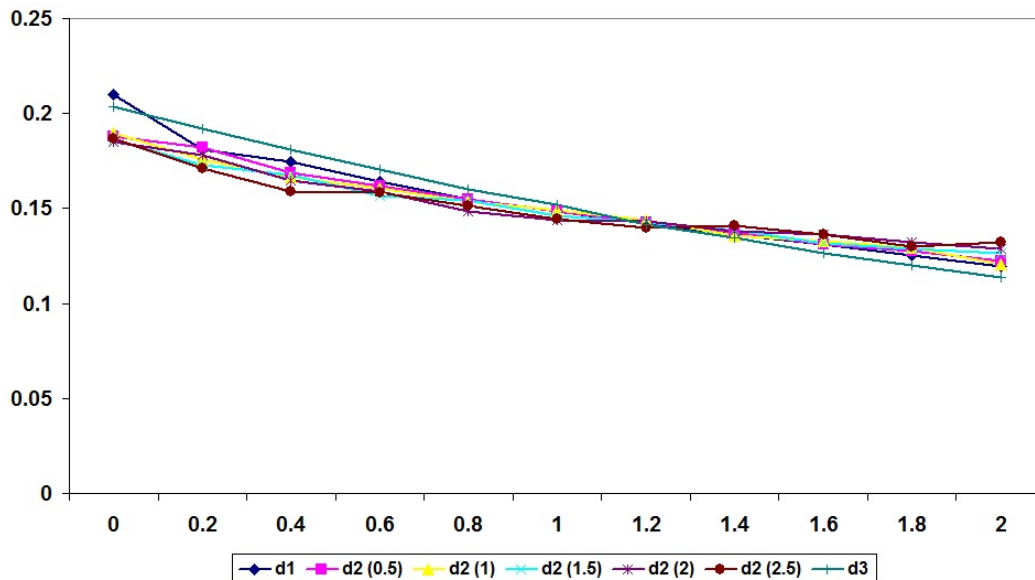


Figure 6.  $M_1$  for ABM,  $d_i$ ,  $i = 1, 2, 3$  refers to  $i$ th class of density functions with value for  $\beta$  between parentheses for the second class

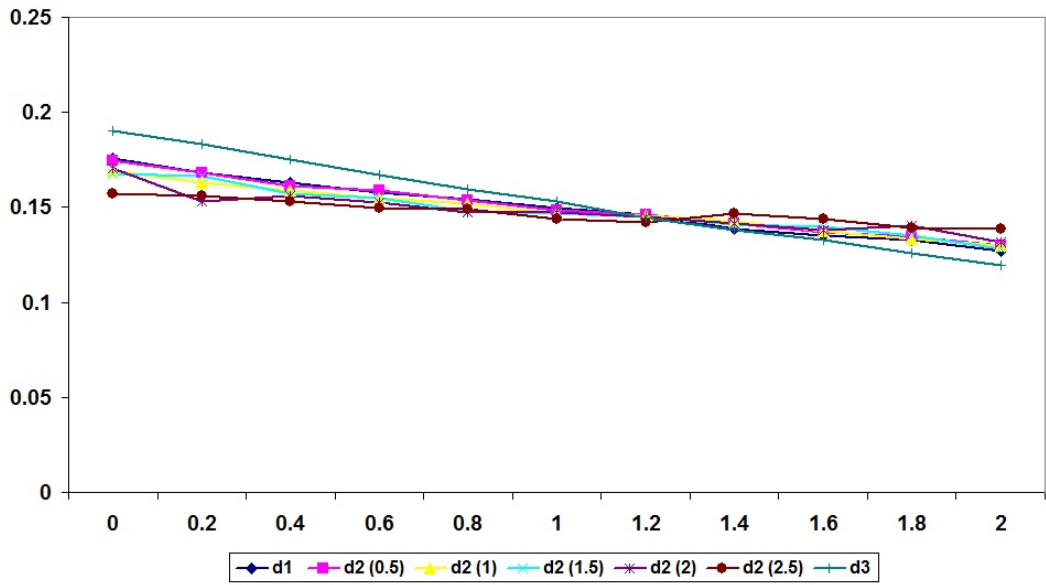


Figure 7.  $M_2$  for ABM,  $d_i, i = 1, 2, 3$  refers to  $i$ th class of density functions with value for  $\beta$  between parentheses for the second class

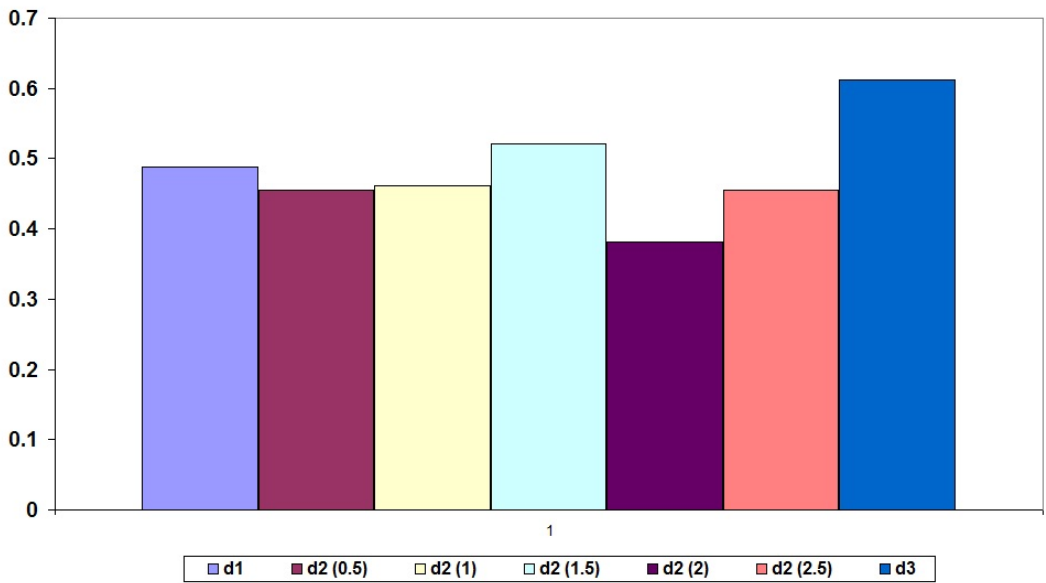


Figure 8.  $S_1$  for ABM,  $d_i, i = 1, 2, 3$  refers to  $i$ th class of density functions with value for  $\beta$  between parentheses for the second class



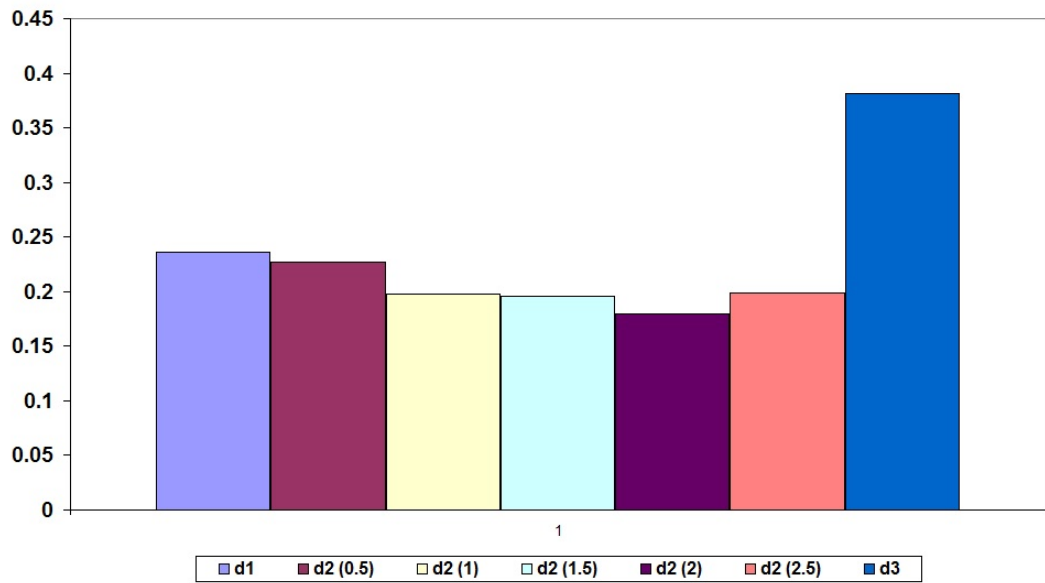


Figure 9.  $S_2$  for ABM,  $d_i, i = 1, 2, 3$  refers to  $i$ th class of density functions with value for  $\beta$  between parentheses for the second class

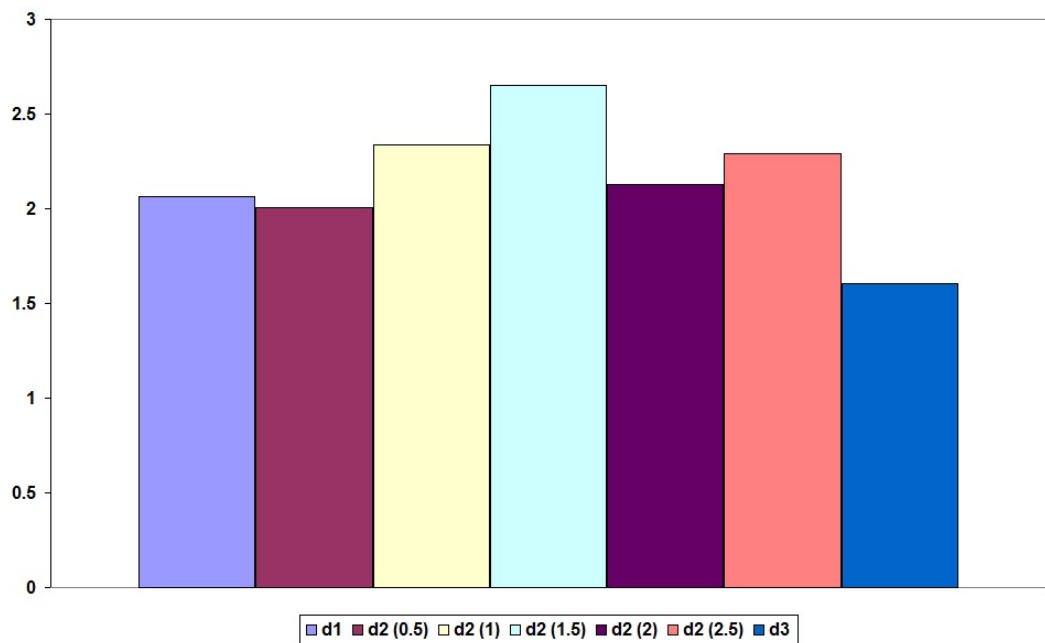


Figure 10.  $S_1/S_2$  for ABM,  $d_i, i = 1, 2, 3$  refers to  $i$ th class of density functions with value for  $\beta$  between parentheses for the second class

## 7. Summary

In this paper we empirically analyzed, via two simulation studies, the influence of the choice of density function for the input variables on the mean effect and on the sensitivity

index, two often used sensitivity measures in probabilistic sensitivity analysis. Although further research is obviously required to increase understanding of the mechanisms involved, some preliminary conclusions can be drawn from our empirical analysis which can serve as hypotheses to be tested in further research, or as guidelines and cautions to practitioners in the field of probabilistic sensitivity analysis.

First, if background knowledge is available that tells something on which values are more likely or more correct than other values, this information should be incorporated in the density function. Ignoring this principle, e.g. by considering the uniform density for the sake of convenience, might produce significantly different values for the considered sensitivity measures which are, therefore, probably not trustworthy.

Secondly, it is important that the obtained values for the sensitivity measures are not interpreted too strictly. That is, in case of the mean effect focus should be on the general trend of the related graphs and in case of the sensitivity index attention should be directed towards the relative importance of the input variables. An analysis in terms of the exact values of the measures, e.g. 'the first input variable is 2.2 times more important than the second input variable', is meaningless, given that it is impossible to construct a unique density function that is in a one-to-one correspondence with available information for virtually all practical applications.

Thirdly, and related to the foregoing remark, our empirical analysis suggests that a general analysis (i.e. not in terms of exact values of the sensitivity measures) is rather insensitive to the choice of density function, as long as a certain level of background information is incorporated in the construction of the density function. In our simulation studies all conclusions concerning the trend of the mean effect and the relative importance of the input variables were very similar in case information on the optimality of the input variables was taken into account in the density function, except for  $M_1$  and the  $S_1/S_2$  ratio for the benchmark function where a Gaussian density with standard deviation 0.5 was used. Thus, depending on the application, it might be important to know, at least, whether the standard deviation is 'small' or 'large'.

## Acknowledgement

André Grow (Centre for Sociological Research, Leuven, Belgium) is greatly acknowledged for providing the software related to his agent-based model on educational assortative mating and for his insightful explanation of the underlying principles of his model.

## References

- [1] Kleijnen JPC. Sensitivity analysis and related analyses: A review of some statistical techniques. *Journal of Statistical Computation and Simulation* 1997; 57:111–142.
- [2] Bijak J, Hilton J, Silverman E, Cao VD. Reforging the wedding ring: exploring a semi-artificial model of population for the United Kingdom with Gaussian process emulators. *Demographic Research* 2013; 29:729–766.
- [3] Koziel S, Leifsson L. *Surrogate-based modeling and optimization*. Springer New York; 2013.
- [4] de Marchi S, Page SE. Agent-Based Models. *Annual Review of Political Science* 2014; 17:1–20.
- [5] Kerner BS, Konhäuser P. Cluster effect in initially homogeneous traffic flow. *Physical Review E* 1993; 48:R2335–R2338.
- [6] Kuehne R. Freeway control using a dynamic traffic flow model and vehicle reidentification techniques. *Transportation Research Record* 1991; 1320:251–259.

- [7] Gilbert N. Agent-based models: quantitative applications in the social sciences. SAGE Publications; 2007.
- [8] Epstein JM. Generative Social Science: Studies in Agent-Based Computational Modeling. Princeton University Press; 2006.
- [9] O'Hagan A. Bayesian analysis of computer code outputs: A tutorial. *Reliability Engineering & System Safety* 2006; 91:1290–1300.
- [10] Moon H, Dean A, Santner T. Algorithms for generating maximin Latin hypercube and orthogonal designs. *Journal of Statistical Theory and Practice* 2010; 5:81–98.
- [11] Davis PJ, Rheinboldt W, Rabinowitz P. *Methods of numerical integration*. Academic Press; 1984.
- [12] Mishra M, Gupta N. Monte Carlo integration technique for the analysis of electromagnetic scattering from conducting surfaces. *Progress in Electromagnetic Research* 2008; 79:91–106.
- [13] Golub G, Meurant G. *Matrices, moments and quadrature with applications*. Princeton University Press; 2009.
- [14] Lavi A, Vogel TP. *Recent advances in optimization techniques*. John Wiley & Sons; 1966.
- [15] Grow A, Van Bavel J. Assortative mating and the reversal of gender inequality in education in Europe: an agent-based model. *PLoS ONE* 2015; 10.
- [16] Schofer E. The worldwide expansion of higher education in the twentieth century. *American Sociological Review* 2005; 70:898–920.
- [17] Esteve A, García-Román J, Permanyer I. The gender-gap reversal in education and its effect on union formation: The end of hypergamy? *Population and Development Review* 2012; 38: 535–546.
- [18] Van Bavel J. The reversal of gender inequality in education, union formation and fertility in Europe. *Vienna Yearbook of Population Research* 2012; 10:127–154.
- [19] De Mulder W, Grow A, Molenberghs G, Verbeke G. Application of statistical emulation to an agent-based model: assortative mating and the reversal of gender inequality in education in Belgium. In: Friedl H, Wagner H, editors. *30th International Workshop on Statistical Modelling*; 2015 July 6-10; Linz, Austria. 2015. p. 139–144.