DEPARTMENT OF FOOD AND RESOURCE ECONOMICS
UNIVERSITY OF COPENHAGEN

IFRO Working Paper

A Ray-Based Input Distance Function to Model Zero-Valued Output Quantities: Derivation and an Empirical Application

Juan José Price
Arne Henningsen

2022 / 03

IFRO Working Paper 2022 / 03

A Ray-Based Input Distance Function to Model Zero-Valued Output Quantities: Derivation and an Empirical Application

Authors: Juan José Price, Arne Henningsen

JEL-classification: C51, D22, D24

Published: June 2022

See the full series IFRO Working Paper here:
www.ifro.ku.dk/english/publications/ifro_series/working_papers/

Department of Food and Resource Economics (IFRO)
University of Copenhagen
Rolighedsvej 23
DK 1958 Frederiksberg DENMARK
www.ifro.ku.dk/english/

A Ray-Based Input Distance Function to Model Zero-Valued Output Quantities: Derivation and an Empirical Application*

Juan José Price[†] & Arne Henningsen[‡]

*We are especially grateful to the comments that we received from Mette Asmild, Peter Bogetoft, and Christopher O'Donnell on an earlier version of this paper. We also appreciate the comments made by participants of the North American Productivity Workshop, NAPW 2021, particularly those of Robin Sickles. Of course, the authors take full responsibility for any remaining errors. We thank Lucas Alexander Kock and Berit Fruelund Kjærside from the Danish Ministry of Culture, and Monika Bille Nielsen from Statistics Denmark, who provided the data. Juan José Price acknowledges financial support from Copenhagen Business School (CBS) and Macquarie University.

[†]Business School, Universidad Adolfo Ibáñez, juan.price@uai.cl

[‡]Dept. of Food and Resource Economics, University of Copenhagen, arne@ifro.ku.dk

Abstract

We derive and empirically apply an input-oriented distance function based on the stochastic ray production function suggested by Löthgren (1997, 2000). We show that the derived ray-based input distance function is suitable for modeling production technologies based on logarithmic functional forms (e.g., Cobb-Douglas and Translog) when control over inputs is greater than control over outputs and when some productive entities do not produce the entire set of outputs — two situations that are jointly present in various economic sectors. We also address a weakness of the stochastic ray function, namely its sensitivity to the outputs' ordering, by using a model-selection approach and a model-averaging approach. We estimate a ray-based Translog input distance function with a data set of Danish museums. These museums have more control over their inputs than over their outputs, and many of them do not produce the entire set of outputs that is considered in our analysis. Given the importance of monotonicity conditions in efficiency analysis, we demonstrate how to impose monotonicity on ray-based input distance functions. As part of the empirical analysis, we estimate technical efficiencies, distance elasticities of the inputs and outputs, and scale elasticities and establish how the production frontier is affected by some environmental variables that are of interest to the museum sector.

Keywords: Stochastic ray production frontier, distance function, input-oriented efficiency, zero output quantities, model averaging, monotonicity, museums.

JEL codes: C51 - D22 - D24

1 Introduction

In econometric efficiency and productivity analysis, multi-output production technologies are usually modeled by so-called distance functions (Färe and Primont, 1990; Kumbhakar and Lovell, 2000), with the most commonly used functional forms being the Cobb-Douglas and Translog specifications. However, these functional forms are unsuitable whenever some producers in a given data set do not produce all outputs so that logarithms of the quantities of the non-produced outputs are undefined. Replacing zero output quantities with arbitrarily small numbers has been suggested but this approach has proven problematic (Henningsen et al., 2015; N'guessan et al., 2017).

The Cobb-Douglas and Translog stochastic ray production frontier functions (Löthgren, 1997, 2000), which are in fact output-oriented distance functions (Henningsen et al., 2015), can be applied to data with zero output quantities. However, in many production processes, managers are expected to have more control over inputs than over outputs and, hence, an input-oriented distance function would be more appropriate for an empirical analysis than an output-oriented distance function.[1] In this paper, we propose a solution to this problem which involves deriving input-oriented Cobb-Douglas and Translog distance functions based on stochastic ray production frontier functions.

We assess these specifications empirically by analyzing input-oriented technical efficiency with data from state-recognized museums in Denmark. This data set appears to be especially suitable for empirically assessing the ray function because many of these museums only produce some of the six outputs that we consider in our analysis. For instance, some museums produce exhibitions, educational programs and engage in conservation but do not do research. Furthermore, these museums have less control over outputs and more control over inputs. Although these entities are not owned by the state, they still receive public funding and the goals in terms of some outputs, such as research and educational programs, are in part defined jointly with the granting and regulatory body, the Danish Agency for Culture and Palaces. Furthermore, the number of visitors is, to a great extent, a demand driven indicator, which in turn determines decisions in terms of the number of exhibitions and other events. Hence, using an input-oriented distance function is more appropriate than using an output-oriented distance function.

[1] See Kumbhakar et al. (2015) regarding the importance of choosing the correct orientation for the distance function.

As part of this empirical analysis, we address a weakness of the stochastic ray function, namely its sensitivity to the ordering of the outputs, by using both a model-selection approach and a model-averaging approach. Given the importance of monotonicity conditions in efficiency analysis (O'Donnell and Coelli, 2005; Henningsen and Henning, 2009), we derive the monotonicity conditions for the inputs and outputs and demonstrate how to impose them on ray-based input distance functions. Furthermore, we derive and estimate the distance elasticities of the inputs and outputs, the elasticity of scale, and how the production technology is affected by environmental variables that are of interest in this sector, both from managerial and regulatory perspectives.

In the following section, we derive and present the methodology and the specification of our model. Section 3 presents the data. Section 4 presents the results. The last section presents our conclusions and possible limitations.

2 Empirical model specification

2.1 Derivation of a ray-based input distance function

As mentioned in the introduction, traditional distance functions based on logarithmic functional forms cannot handle observations where one (or more) outputs have a quantity equal to zero. An alternative specification that solves this problem is the stochastic ray production function, which was originally proposed by Löthgren (1997). This functional form represents the vector of output quantities as polar coordinates rather than as Cartesian coordinates, i.e., the vector of output quantities is represented by its (Euclidean) length and its direction rather than by the individual output quantities.

A stochastic ray production frontier model is defined as:

$$\ln \left\| y \right\| = f^* (\ln x, \varphi (y), z) - u^* + v^*, \tag{1}$$

where $x = (x_1, x_2, \ldots, x_N)^\top \in \mathbb{R}^N_{>0}$ is a vector of N strictly positive input quantities, $y = (y_1, y_2, \ldots, y_M)^\top \in \mathbb{R}^M_{\geq 0}$ is a vector of M non-negative output quantities with at least one strictly positive output quantity $\exists i \in 1, \ldots, M : y_i > 0$, $\varphi(y) = \big(\varphi_1(y), \varphi_2(y), \ldots, \varphi_{M-1}(y)\big)^\top$ is a vector of the angles of the vector of output quantities y with $\varphi_i(y) = \arccos\left(y_i \middle/ \sqrt{\sum_{j=i}^{M} y_j^2} \right) \forall\, i = 1, \ldots, M-1$, $\left\| y \right\| = \sqrt{\sum_{j=1}^{M} y_j^2}$ is the (Euclidean) length of the vector of output quantities, $z = (z_1, z_2, \ldots, z_K)^\top \in \mathbb{R}^K$ is a vector of K 'environmental' variables, $u^* \geq 0$ is the inefficiency term, v^* is the noise term, and $f^*(\cdot)$ is a function that can take various functional forms, e.g., linear for a Cobb-Douglas stochastic ray production frontier or quadratic for a Translog stochastic ray production frontier. The stochastic ray production frontier (1) can be seen as a Shephard output distance function (Henningsen et al., 2015):

$$\ln D^o(x, y, z) = -u^* = f(\ln x, \varphi(y), z) + \ln \|y\| + v, \tag{2}$$

where $D^o(x, y, z) = e^{-u^*}$ with $0 \leq D^o(x, y, z) \leq 1$ is a Shephard output distance function, $f(\cdot) = -f^*(\cdot)$, and $v = -v^*$.

This specification of a stochastic-ray-based output distance function can be generalized to the following 'general' distance function:

$$\ln D(x, y, z) = f\left(\ln x, \varphi(y), \ln \|y\|, z\right) + v. \tag{3}$$

If function $\exp(f(\cdot))$ is linear homogeneous in output quantities y (e.g., $f(\ln x, \varphi(y), \ln \|y\|, z) = f(\ln x, \varphi(y), z) + \ln \|y\|$), function $D(x, y, z)$ can be seen as a Shephard output distance function. However, if function $\exp(f(\cdot))$ is linear homogeneous in input quantities x, function $D(x, y, z)$ can be seen as a Shephard input distance function. We can make this function linear homogeneous in input quantities x so that we get a Shephard input distance function, e.g., by modifying equation (3) to:

$$\ln D^i(x, y, z) = f\left(\ln \tilde{x}, \varphi(y), \ln \|y\|, z\right) + \ln x_N + v, \tag{4}$$

where $\tilde{x} = (\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_{N-1})^\top = (x_1/x_N, x_2/x_N, \ldots, x_{N-1}/x_N)^\top$ is a vector of normalized input quantities.

By replacing the logarithm of the Shephard input distance measure $D^i(x, y) \geq 1$, i.e., $\ln D^i(x, y) \geq 0$, with the inefficiency term $u \geq 0$ and a little re-arranging, we get:

$$-\ln x_N = f\left(\ln \tilde{x}, \varphi(y), \ln \|y\|, z\right) - u + v, \tag{5}$$

which we can estimate using stochastic frontier analysis.

Assuming a quadratic functional form for $f(\cdot)$, we get the following equation, which can be easily estimated as a stochastic frontier model:[2]

$$
\begin{aligned}
-\ln x_N = &\alpha_0 + \sum_{i=1}^{N-1} \alpha_i \ln(\tilde{x}_i) + \frac{1}{2}\sum_{i=1}^{N-1}\sum_{j=1}^{N-1} \alpha_{ij} \ln(\tilde{x}_i)\ln\left(\tilde{x}_j\right) \\
&+ \sum_{i=1}^{M-1} \beta_i \varphi_i(y) + \beta_M \ln\left\|y\right\| + \frac{1}{2}\sum_{i=1}^{M-1}\sum_{j=1}^{M-1} \beta_{ij}\varphi_i(y)\varphi_j(y) \\
&+ \sum_{i=1}^{M-1} \beta_{iM}\varphi_i(y)\ln\left\|y\right\| + \frac{1}{2}\beta_{MM}\left(\ln\left\|y\right\|\right)^2 \\
&+ \sum_{i=1}^{K} \delta_i z_i + \frac{1}{2}\sum_{i=1}^{K}\sum_{j=1}^{K} \delta_{ij} z_i z_j \\
&+ \sum_{i=1}^{N-1}\sum_{j=1}^{M-1} \psi_{ij} \ln(\tilde{x}_i)\varphi_j(y) + \sum_{i=1}^{N-1} \psi_{iM} \ln(\tilde{x}_i)\ln\left\|y\right\| \\
&+ \sum_{i=1}^{N-1}\sum_{j=1}^{K} \xi_{ij} \ln(\tilde{x}_i) z_j \\
&+ \sum_{i=1}^{M-1}\sum_{j=1}^{K} \zeta_{ij}\varphi_i(y) z_j + \sum_{i=1}^{K} \zeta_{Mi} \ln\left\|y\right\| z_i - u + v
\end{aligned} \tag{6}
$$

with $\alpha_{ij} = \alpha_{ji} \, \forall \, i,j = 1,\ldots,N-1$, $\beta_{ij} = \beta_{ji} \, \forall \, i,j = 1,\ldots,M-1$, and $\delta_{ij} = \delta_{ji} \, \forall \, i,j = 1,\ldots,K$.[3]

2.2 Distance elasticities

Assuming that the inefficiency term u and the noise term v are both independent from the explanatory variables ($\tilde{x}$, y, z), the distance elasticities of the N inputs and the M outputs are:

$$
\frac{\partial \ln D^i(x,y,z)}{\partial \ln x_i} = \alpha_i + \sum_{j=1}^{N} \alpha_{ij} \ln x_j + \sum_{j=1}^{M-1} \psi_{ij}\varphi_j(y) + \psi_{iM} \ln\left\|y\right\| \tag{7}
$$

[2]The Cobb-Douglas functional form is a special case with $\alpha_{ij} = 0 \, \forall \, i,j = 1,\ldots,N$, $\beta_{ij} = 0 \, \forall \, i,j = 1,\ldots M$, $\delta_{ij} = 0 \, \forall \, i,j = 1,\ldots,K$, $\psi_{ij} = 0 \, \forall \, i = 1,\ldots,N; j = 1,\ldots,M$, $\xi_{ij} = 0 \, \forall \, i = 1,\ldots,N; j = 1,\ldots,K$, and $\zeta_{ij} = 0 \, \forall \, i = 1,\ldots,M; j = 1,\ldots,K$. As mentioned in Section 4, we test the Translog specification defined in equation (6) against the corresponding Cobb-Douglas specification. The test rejects the Cobb-Douglas specification in favor of the Translog specification. Therefore, the following derivations are based on the Translog specification.

[3]In Appendix A, we present the derivation of equation (6).

$$+\sum_{j=1}^{K}\xi_{ij}z_j \;\forall\; i=1,\ldots,N$$

$$\begin{aligned}\frac{\partial \ln D^i(x,y,z)}{\partial \ln y_i} = & \sum_{j=1}^{M-1}\beta_j y_i \Omega_{ji} + \beta_M \frac{y_i^2}{\left\|y\right\|^2} + \sum_{j=1}^{M-1}\sum_{k=1}^{M-1}\beta_{jk}\varphi_k(y) y_i \Omega_{ji} \\ & + \sum_{j=1}^{M-1}\beta_{jM}\left(y_i\Omega_{ji}\ln\left\|y\right\| + \varphi_j(y)\frac{y_i^2}{\left\|y\right\|^2}\right) + \beta_{MM}\frac{\ln\left\|y\right\| y_i^2}{\left\|y\right\|^2} \\ & + \sum_{j=1}^{N}\sum_{k=1}^{M-1}\psi_{jk}\ln x_j y_i \Omega_{ki} + \sum_{j=1}^{N}\psi_{jM}\ln x_j \frac{y_i^2}{\left\|y\right\|^2} \\ & + \sum_{j=1}^{M-1}\sum_{k=1}^{K}\zeta_{jk}z_k y_i \Omega_{ji} + \sum_{j=1}^{K}\zeta_{Mj}z_j\frac{y_i^2}{\left\|y\right\|^2} \;\forall\; i=1,\ldots,M,\end{aligned} \tag{8}$$

respectively, with:

$$\Omega_{ji} \equiv \frac{\partial \varphi_j(y)}{\partial y_i} = \begin{cases} 0 & \text{if } i<j \vee \left(i=j \wedge y_j>0 \wedge \sum_{k=j+1}^{M} y_k = 0\right) \\ \text{undefined} & \text{if } i \geq j \wedge y_j = y_{j+1} = \ldots = y_M = 0 \\ \frac{y_j}{\|y\|_j^2} & \text{if } i>j \wedge \sum_{k\in[(j+1)..M]\setminus i} y_k = 0 \\ \frac{y_i y_j - I_{i=j}\|y\|_j^2}{\|y\|_j^2\,\|y\|_{j+1}} & \text{otherwise} \end{cases} \tag{9}$$

$$\forall\; i=1,\ldots,M, j=1,\ldots,M-1,$$

$$\left\|y\right\|_i \equiv \sqrt{\sum_{j=i}^{M} y_j^2}, \tag{10}$$

$\alpha_N = 1-\sum_{i=1}^{N-1}\alpha_i$, $\alpha_{Nj} = -\sum_{i=1}^{N-1}\alpha_{ij}\,\forall\, j=1,\ldots,N$, $\alpha_{iN} = -\sum_{j=1}^{N-1}\alpha_{ij}\,\forall\, i=1,\ldots,N$, $\psi_{Nj} = -\sum_{i=1}^{N-1}\psi_{ij}\,\forall\, j=1,\ldots,M$, $\xi_{Nj} = -\sum_{i=1}^{N-1}\xi_{ij}\,\forall\, j=1,\ldots,K$, and I_{cond} being an indicator function that takes the value one if the condition in the subscript *cond* is fulfilled and the value zero otherwise.

In order to quantity the effects of the environmental variables on the production technology, we derived the following semi-elasticities:[4]

$$\frac{\partial \ln D^i(x,y,z)}{\partial\, z_i} = \delta_i + \sum_{j=1}^{K}\delta_{ij}z_j + \sum_{j=1}^{N}\xi_{ji}\ln x_i + \sum_{j=1}^{M-1}\zeta_{ji}\varphi_j(y) + \zeta_{Mi}\ln\left\|y\right\| \tag{11}$$

[4]If z_i is in logarithm, equation (11) indicates elasticities rather than semi-elasticities.

$\forall\ i = 1, \ldots, K.$

2.3 Monotonicity conditions

Monotonicity conditions derived from microeconomic theory imply that the Shephard input distance functions (technical inefficiency) should be non-decreasing in input quantities and non-increasing in output quantities, i.e., $\partial D^i(x,y,z)/\partial x_i \geq 0\ \forall i = 1,\ldots,N$ and $\partial D^i(x,y,z)/\partial y_i \leq 0\ \forall i = 1,\ldots,M$, respectively. As $D^i(x,y,z) > 0$ and $x_i > 0\,\forall\, i = 1,\ldots,N$, the monotonicity condition $\partial D^i(x,y,z)/\partial x_i \geq 0$ is equivalent to the condition $(\partial D^i(x,y,z)/\partial x_i)\,(x_i/D^i(x,y,z)) = \partial \ln D^i(x,y,z)/\partial \ln x_i \geq 0$. Hence, we can use the right-hand side of equation (7) as the monotonicity condition regarding the input quantities. However, as we allow output quantities to be zero, we cannot use the right-hand side of equation (8) as the monotonicity condition regarding the output quantities because if an output quantity is zero (i.e., $y_i = 0$), the right-hand side of equation (8) is zero and, thus, non-positive even if the monotonicity condition regarding this output is violated, i.e., $\partial D^i(x,y,z)/\partial y_i > 0$. In order to have monotonicity conditions that can also be applied in case of zero output quantities, we use the semi-elasticity of the outputs (i.e., $\partial \ln D^i(x,y,z)/\partial y_i \leq 0\ \forall i = 1,\ldots,M$) as the monotonicity conditions regarding the output quantities. Using only the estimated coefficients and replacing the non-estimated coefficients by the respective homogeneity restrictions (e.g., $\alpha_N = 1 - \sum_{i=1}^{M-1} \alpha_i$) and the respective symmetry normalizations (e.g., $\alpha_{ij} = \alpha_{ji}\,\forall\, i > j$), we obtain the following monotonicity conditions for $x_i; i = 1,\ldots,N-1$, x_N, and $y_i; i = 1,\ldots,M$:

$$0 \leq \alpha_i + \sum_{j=1}^{i} \alpha_{ji} \ln \tilde{x}_j + \sum_{j=i+1}^{N-1} \alpha_{ij} \ln \tilde{x}_j \tag{12}$$

$$+ \sum_{j=1}^{M-1} \psi_{ij}\varphi_j(y) + \psi_{iM} \ln \left\|y\right\| + \sum_{j=1}^{K} \xi_{ij} z_j\ \forall\, i = 1,\ldots,N-1,$$

$$-1 \leq -\sum_{k=1}^{N-1} \alpha_k - \sum_{j=1}^{N-1}\sum_{k=j}^{N-1} \alpha_{jk}\left(\ln \tilde{x}_j + I_{j\neq k} \ln \tilde{x}_k\right) \tag{13}$$

$$-\sum_{j=1}^{M-1}\sum_{k=1}^{N-1} \psi_{kj}\varphi_j(y) - \sum_{k=1}^{N-1} \psi_{kM} \ln \left\|y\right\| - \sum_{j=1}^{K}\sum_{k=1}^{N-1} \xi_{kj} z_j, and$$

$$0 \leq -\sum_{j=1}^{M-1} \beta_j \Omega_{ji} - \beta_M \frac{y_i}{\left\|y\right\|^2} - \sum_{j=1}^{M-1}\sum_{k=j}^{M-1} \beta_{jk}\left(\varphi_k(y)\,\Omega_{ji} + I_{j\neq k}\varphi_j(y)\,\Omega_{ki}\right) \tag{14}$$

$$
\begin{aligned}
&-\sum_{j=1}^{M-1}\beta_{jM}\left(\Omega_{ji}\ln\|y\|+\varphi_j(y)\frac{y_i}{\|y\|^2}\right)-\beta_{MM}\frac{\ln\|y\|\,y_i}{\|y\|^2}\\
&-\sum_{j=1}^{N-1}\sum_{k=1}^{M-1}\psi_{jk}\ln\tilde{x}_j\Omega_{ki}-\sum_{j=1}^{N-1}\psi_{jM}\ln\tilde{x}_j\frac{y_i}{\|y\|^2}\\
&-\sum_{j=1}^{M-1}\sum_{k=1}^{K}\zeta_{jk}z_k\Omega_{ji}-\sum_{j=1}^{K}\zeta_{Mj}\frac{z_jy_i}{\|y\|^2}\ \forall\, i=1,\dots,M,
\end{aligned}
$$

respectively.

As all these monotonicity conditions are linear in the estimated coefficients, we can represent them in matrix form as $R\theta \geq r$, where $\theta = (\alpha_0, \alpha_1, \dots, \alpha_{N-1},\ \alpha_{11}, \dots, \alpha_{1,N-1},\ \alpha_{22}, \dots, \alpha_{2,N-1},\ \dots, \alpha_{N-1,N-1},\ \beta_1, \dots, \beta_M,\ \beta_{11}, \dots, \beta_{1M},\ \beta_{22}, \dots, \beta_{2M},\ \dots, \beta_{MM},\ \delta_1, \dots, \delta_K,\ \delta_{11}, \dots, \delta_{1K},\ \delta_{22}, \dots, \delta_{2K},\ \dots, \delta_{KK},\ \psi_{11}, \dots, \psi_{1M},\ \psi_{21}, \dots, \psi_{2M},\ \dots, \psi_{N-1,M},\ \xi_{11}, \dots, \xi_{1K},\ \xi_{21}, \dots, \xi_{2K},\ \dots, \xi_{N-1,K},\ \zeta_{11}, \dots, \zeta_{1K},\ \zeta_{21}, \dots, \zeta_{2K},\ \dots, \zeta_{MK})$ is a vector of all estimated coefficients, $R = [R_{i,\theta}]$ is a matrix with one row i for each of the $N+M$ monotonicity conditions and one column for each of the coefficients in θ, and r is a vector with one element for each of the $N+M$ monotonicity restrictions. Each element of the matrix R and of the vector r is defined in Appendix B. Given that all these monotonicity conditions are linear in the estimated coefficients, we can use the method suggested by Henningsen and Henning (2009) to impose the monotonicity conditions.[5]

3 Data

We estimate our model with data from state-recognized Danish museums[6]. The data set includes information from 93 museums over a six-year period: 2012 and 2014-2018[7]. This is an unbalanced panel database with 528 observations. The following inputs, outputs, and environmental variables are considered:

Inputs

[5] As the method proposed by Henningsen and Henning (2009) requires monotonicity conditions that are linear in the estimated coefficients, we cannot use the monotonicity conditions $\partial D^i(x,y,z)/\partial x_i = (D^i(x,y,z)/x_i)\ (\partial D^i(x,y,z)/\partial x_i) \geq 0\ \forall i = 1,\dots,N$ and $\partial D^i(x,y,z)/\partial y_i\ (D^i(x,y,z)/y_i)\ (\partial D^i(x,y,z)/\partial y_i) \leq 0\ \forall i = 1,\dots,M$, because they are non-linear in the estimated coefficients (given that $D^i(x,y,z)$ depends on the estimated coefficients).

[6] State-recognized museums are not owned by the state but receive public funding and are obliged to preserve, register, research and exhibit their collections. Complying with this wide set of responsibilities entails delivering multiple cultural and educational services so that evaluating their technical efficiency in a multi-output setting is warranted.

[7] Data for 2013 is not available.

- *Scientific labor*: This includes conservationists, researchers, and other scientific staff (full time equivalents).
- *Non-scientific labor*: This category consists of management, administrative and maintenance staff (full time equivalents).
- *Capital*: The sum of running and maintenance costs is used as proxy for the capital stock. This appears to make economic sense as larger collections require more space for both storage and exhibition, which in turn requires additional operating and maintenance costs[8]. These costs have been expressed in real terms (Danish Kroner of 2014).

Outputs

- *Visitors*: Number of (physical) visitors.
- *Research*: Number of scientific articles published.
- *Exhibitions*: Number of temporary exhibitions (i.e., excluding the permanent exhibition).
- *Education*: Number of primary school classes on educational visits to the museum. It seems better, from a supply perspective, to consider the number of school classes regardless of the number of students in each group, as there are fixed costs involved in each visit.
- *Events*: Number of events other than exhibitions (e.g., workshops, conferences, book presentations) that take place on the premises of the museums.
- *Conservation*: This output corresponds to the expenditure on conservation activities, expressed in real terms (Danish Kroner of 2014).

Environmental variables

[8]Information about the capital stock of museums is generally not available. This has been pointed out in the existing literature for other countries (O'Hagan, 1998; Feldstein, 1991; Frey and Meier, 2006; Grampp, 1989; Peacock and Godfrey, 1974) and it is also the case in Denmark. To cope with this problem, Bishop and Brand (2003) use the same proxy as we use, whereas del Barrio et al. (2009) and del Barrio-Tellado and Herrero-Prieto (2019) use the area of museums in square meters. We chose the first of these proxies because information about the area of the museums is not available in our data set.

In order to control for factors that might affect the production frontier (i.e., the minimum input quantities museums require to produce a certain level of output quantities), we include two environmental variables:

- *Number of visit sites*: Some museums manage two or more visit sites (up to twenty). Therefore, it is expected that these museums require more inputs for every given level of outputs than those that manage only one site. For example, museums that manage several (small) sites must have guards and receptionists at each visit site and, thus, they are expected to need more inputs than museums that produce the same outputs but manage only one (large) visit site. We use the natural logarithm of the number of visit sites as an environmental variable.[9]

- *Special responsibilities*: Some museums have responsibilities that go beyond those specified by the legislation. Examples include taking care of the archaeological sites located in their municipality and collaborating on archaeological investigations. It is expected that these museums require more inputs for every given level of outputs under analysis than those that do not have special responsibilities. We consider a dummy variable that takes the value one if museums have special responsibilities and zero otherwise.

Table 1 presents descriptive statistics of the input quantities, output quantities, and environmental variables. In order to make our results invariant to units of measurement, we mean-scale all output quantities.

4 Results

We estimate the ray-based Translog input distance function defined in equation (6) as stochastic frontier model[10] by using the R software (R Core Team, 2020) and its add-on packages "sfaR" (Dakpo et al., 2021) and "quadprog" (Turlach and Weingessel, 2011). Given that one of our 'environmental' variables (z_2: special responsibilities) is a dummy

[9]The purpose of using the natural logarithm is to reduce the occurrence of extreme values and to reduce the skewness of the distribution of this variable.

[10] Although ignoring the panel structure of the data set likely results in a violation of the i.i.d. assumption, we conduct a 'pooled' estimation because, in our data set, the variation of the variables between years is very low and the number of time periods in the data set is so small that individual effects at the museum level capture almost all variation in the variables, which means there is not enough variation to estimate the parameters of the distance function. Furthermore, we would like our efficiency scores to be based on the benchmarking of museums against each other, but a *true fixed effects* or a *true random effects* stochastic frontier model would remove the variation between the museums and, thus, prevent benchmarking between the museums.

Table 1: Summary Statistics

Variable	$n(=0)$	Mean	St. Dev.	Min	Median	Max
capital [1,000 real Danish Kroner]	0	2,608.98	4,443.81	107.97	1,247.19	65,117.11
scientific labour [full time equivalents]	0	7.76	7.76	1	4.8	59
non-scientific labour [full time equivalents]	0	18.98	26.46	0.40	10.10	180.00
visitors [number of physical visitors]	0	88,677.27	125,810.90	2,473	39,884	835,606
conservation [1,000 real Danish Kroner]	43	197.43	267.47	0.00	98.81	2,340.55
research [number of articles]	118	5.25	7.94	0	2	59
exhibitions [number of exhibitions]	15	5.52	4.57	0	4	45
education [number of group visits]	0	170.95	251.23	2	107	4,461
events [number of events]	0	246.82	361.67	1	118	2,576
visit sites [number of visit sites]	0	2.69	3.33	1	1	20
responsibilities [1 = special responsibilities]	367	0.30	0.46	0	0	1

variable, we make one adjustment to equation (6). As the quadratic term z_2^2 of the dummy variable for special responsibilities equals its non-squared term z_2, i.e., $z_2^2 = z_2$, we exclude the term z_2^2 in order to avoid perfect multicollinearity. This corresponds to imposing the restriction $\delta_{22} = 0$.[11]

As the estimation results of ray-based production or distance functions depend on the ordering of the outputs, we estimate the model with all 720 possible orderings of the outputs[12] and select the ordering that gives the best fit in terms of the log-likelihood value as suggested by Henningsen et al. (2017). This is the so-called model-selection method. Additionally, we use the model-averaging method and obtain weighted averages of estimation results (e.g., distance elasticities, elasticities of scale, and technical efficiency estimates) for all possible orderings of outputs as suggested by Huang and Lai (2012) for stochastic frontier models in general and as suggested by Tsionas et al. (2021) for the stochastic ray production frontiers.

In order to include all observations in all orderings of outputs (so that the obtained log-likelihood values are comparable), we apply a small trick for the orderings of outputs where there are observations for which the quantities of the last two (or more) outputs are zero. This is because whenever the quantities of outputs j to M are zero for $j \leq M-1$ (i.e., $y_i = 0 \; \forall \; i = j, \ldots, M$ with $j \leq M-1$), the angles $\theta_i; i = j, \ldots, M-1$ are undefined. In these (few) cases,[13] we calculate the undefined angles by assuming that $y_i = \kappa \; \forall \; i = j, \ldots, M$, where κ is an arbitrarily small strictly positive value. Under this assumption, we get $\theta_i = \arccos\left(1/\sqrt{2}\right) = \pi/4 \; \forall \; i = j, \ldots, M-1$, i.e., a 45-degree angle, independent of the value of κ.[14] This problem does not occur when using the ordering of outputs that gives the best fit to the model. Hence, we use this trick only when

[11] We keep the quadratic term z_1^2, the interaction term $z_1 z_2$, and also the interaction terms between the environmental variables and the other explanatory variables in the model.

[12] As the ordering of the last two outputs should, theoretically, not affect the results (Henningsen et al., 2017), it should be sufficient to estimate the model with 360 different orderings of the outputs. As numerical inaccuracies could result in different estimates for the two models that only differ in terms of the last two outputs, we estimate the model with all 720 different orderings of the outputs. However, our estimates for models that only differ in terms of the last two outputs are always very close to each other or are even virtually identical.

[13] In none of the 720 orderings of outputs are more than two of the last outputs zero. When the last two outputs are 'exhibitions' and 'research', there are 7 observations for which the last two outputs are zero. When the last two outputs are 'conservation' and 'research', there are 5 observations for which the last two outputs are zero. When the last two outputs are 'conservation' and 'exhibition', there are 2 observations for which the last two outputs are zero. In all other orderings, we do not need to apply the trick.

[14] It is worth pointing out that in our approach, the values of $\varphi_i(y)$ and, thus, the estimation results do not depend on the size of the arbitrarily small value of κ. Hence, this is not a naive solution such as the one that replaces zero values with an arbitrarily small number.

searching for the ordering of outputs that gives the highest log-likelihood value and when applying the model averaging method.

After estimating the models, we determine the extent to which they fulfill the monotonicity conditions by calculating the right-hand sides of equations (12), (13), and (14). These conditions are violated for at least some of the observations regardless of which ordering of outputs is used. For instance, when using the ordering of outputs that gives the best fit, monotonicity conditions regarding the inputs are violated at 8%, 1%, and 10% of the observations while monotonicity conditions regarding the outputs are violated at 4%, 52%, 9%, 16%, 38%, and 16% of the observations reported in the order of inputs and outputs used in Table 1. We proceed by imposing the monotonicity conditions using a procedure that is based on Henningsen and Henning (2009), which has been adapted to the ray-based Translog input distance function as explained in Appendix C.

We select the ordering of the outputs that results in the highest log-likelihood value of the restricted model.[15] This ordering is the one that is used in Table 1. The results of the unrestricted estimation and the coefficients of the restricted estimation (with monotonicity conditions imposed) for this ordering of outputs are presented in Table 2.

Given that we use a Translog functional form, most coefficients presented in Table 2 do not have a relevant interpretation. In order to obtain estimation results with relevant real-world meanings, we calculate distance elasticities of the inputs and outputs, the elasticity of scale, the effects of the 'environmental' variables on the production frontier, and the technical efficiency estimates (see Figures 5, 6, 7, and 8 in Appendix D).

Figure 1 presents scatter plots between the log-likelihood value (indicating the fit of the model) and four important estimation results (i.e., mean and median technical efficiency estimates and mean and median elasticity of scale) for each of the 720 possible orderings of the outputs. The estimation results depend on the ordering of the outputs to a moderate extent. If one disregards the 50% of the orderings that give the worst fit, the estimation results depend even less on the ordering of the outputs. In our empirical application, the model-averaged estimates (indicated by the horizontal red lines in Figure 1) are rather close to the estimates based on the ordering of the outputs that gives the best fit (indicated by the dot furthest to the right in each panel of Figure 1).

[15]As the number of estimated coefficients is the same for all orderings of the outputs, selecting the ordering of the outputs that gives the highest log-likelihood value is equivalent to using the Akaike information criterion (AIC) or the Bayesian information criterion (BIC) to select the ordering that gives the best fit to our data.

Table 2: Estimation results: unrestricted and restricted estimates

Coef.	Estimate	S.E.	Restricted	Coef.	Estimate	S.E.	Restricted
α_0	3.6032***	0.8286	3.2826	β_{66}	-0.0882	0.0552	-0.1408
α_1	-0.0386	0.2214	0.2341	ψ_{11}	0.1437	0.1157	0.0144
α_2	0.6640**	0.2112	0.5252	ψ_{12}	-0.1017	0.0750	-0.0206
β_1	-0.9030	0.8261	-1.5916	ψ_{13}	0.0211	0.0653	-0.0407
β_2	-1.7180***	0.5207	-0.9470	ψ_{14}	-0.0607	0.0628	-0.0310
β_3	-0.8821	0.4783	-0.8465	ψ_{15}	0.2496***	0.0610	0.0740
β_4	-0.9488	0.4978	-0.5256	ψ_{16}	0.0064	0.0391	-0.0023
β_5	-0.4071	0.3744	-0.2079	ψ_{21}	-0.0418	0.1212	0.0156
β_6	-0.8938***	0.2493	-1.1746	ψ_{22}	-0.0426	0.0705	-0.0032
δ_1	0.7211*	0.2820	0.4557	ψ_{23}	-0.0885	0.0624	-0.0067
δ_2	-2.2628***	0.5002	-0.6835	ψ_{24}	0.1115	0.0633	0.0371
α_{11}	-0.0734	0.0381	-0.0508	ψ_{25}	-0.0149	0.0649	-0.0308
α_{12}	0.0611	0.0331	0.0278	ψ_{26}	0.0167	0.0396	-0.0066
α_{22}	0.1717***	0.0455	0.0890	δ_{11}	-0.1430*	0.0727	-0.1554
β_{11}	0.6188	0.5872	1.7856	δ_{12}	-0.2756***	0.0740	-0.2978
β_{12}	0.3395	0.2530	-0.0488	ξ_{11}	-0.1014*	0.0465	-0.0596
β_{13}	-0.1134	0.2287	-0.0611	ξ_{12}	0.1918*	0.0778	0.1732
β_{14}	0.3249	0.2534	0.0523	ξ_{21}	0.0571	0.0388	0.0158
β_{15}	0.2831	0.2027	0.0555	ξ_{22}	-0.3986***	0.0757	-0.3006
β_{16}	0.2269	0.1300	0.3080	ζ_{11}	-0.6834***	0.1508	-0.3542
β_{22}	0.5025*	0.2447	0.6827	ζ_{12}	1.4213***	0.2780	0.6247
β_{23}	0.2947*	0.1353	0.0590	ζ_{21}	0.1180	0.0647	0.0880
β_{24}	-0.0093	0.1171	-0.0068	ζ_{22}	0.2147	0.1367	-0.0404
β_{25}	0.1398	0.1102	0.0343	ζ_{31}	0.0637	0.0627	0.0280
β_{26}	-0.0661	0.0781	0.0113	ζ_{32}	-0.1254	0.1355	-0.1020
β_{33}	0.9841***	0.2269	0.9535	ζ_{41}	0.0711	0.0765	0.0037
β_{34}	-0.3160**	0.1113	-0.1310	ζ_{42}	0.2244	0.1454	0.1323
β_{35}	-0.1964	0.1033	-0.0192	ζ_{51}	0.0296	0.0727	-0.0166
β_{36}	0.1712*	0.0666	0.0975	ζ_{52}	0.0832	0.1372	0.0086
β_{44}	0.8215***	0.2204	0.5828	ζ_{61}	0.0454	0.0442	0.0983
β_{45}	0.1825	0.1058	0.0151	ζ_{62}	-0.0148	0.0875	-0.0855
β_{46}	-0.1135	0.0738	0.0073	$\log(\sigma_u^2)$	-2.6487***	0.3871	-2.0672***
β_{55}	-0.0242	0.2088	0.1562	$\log(\sigma_v^2)$	-2.8951***	0.1754	-3.0046***
β_{56}	-0.0021	0.0630	0.0224				

Notes: Asterisks indicate significance levels, where: *** $\leq$ 0.001, ** $\leq$ 0.01, * $\leq$ 0.05. Column 'Restricted' indicates the restricted estimates. The restricted estimates of the model coefficients, i.e., all coefficients except for the variance parameters $\log(\sigma_u^2)$ and $\log(\sigma_v^2)$, are obtained in the second-step minimum-distance estimation and, thus, do not have standard errors or significance levels. The restricted estimate of the intercept is adjusted by the intercept of the third-step stochastic-frontier estimation. The variance parameters $\log(\sigma_u^2)$ and $\log(\sigma_v^2)$ of the restricted estimation are obtained in the third-step stochastic-frontier estimation.

This result not only holds for mean and median values of the estimated measures (as presented in Figure 1) but also for the values of these measures at the level of individual observations (see Figures 9, 10, 11, and 12 in Appendix E). However, this result should not be generalized as it could be different in other empirical applications.

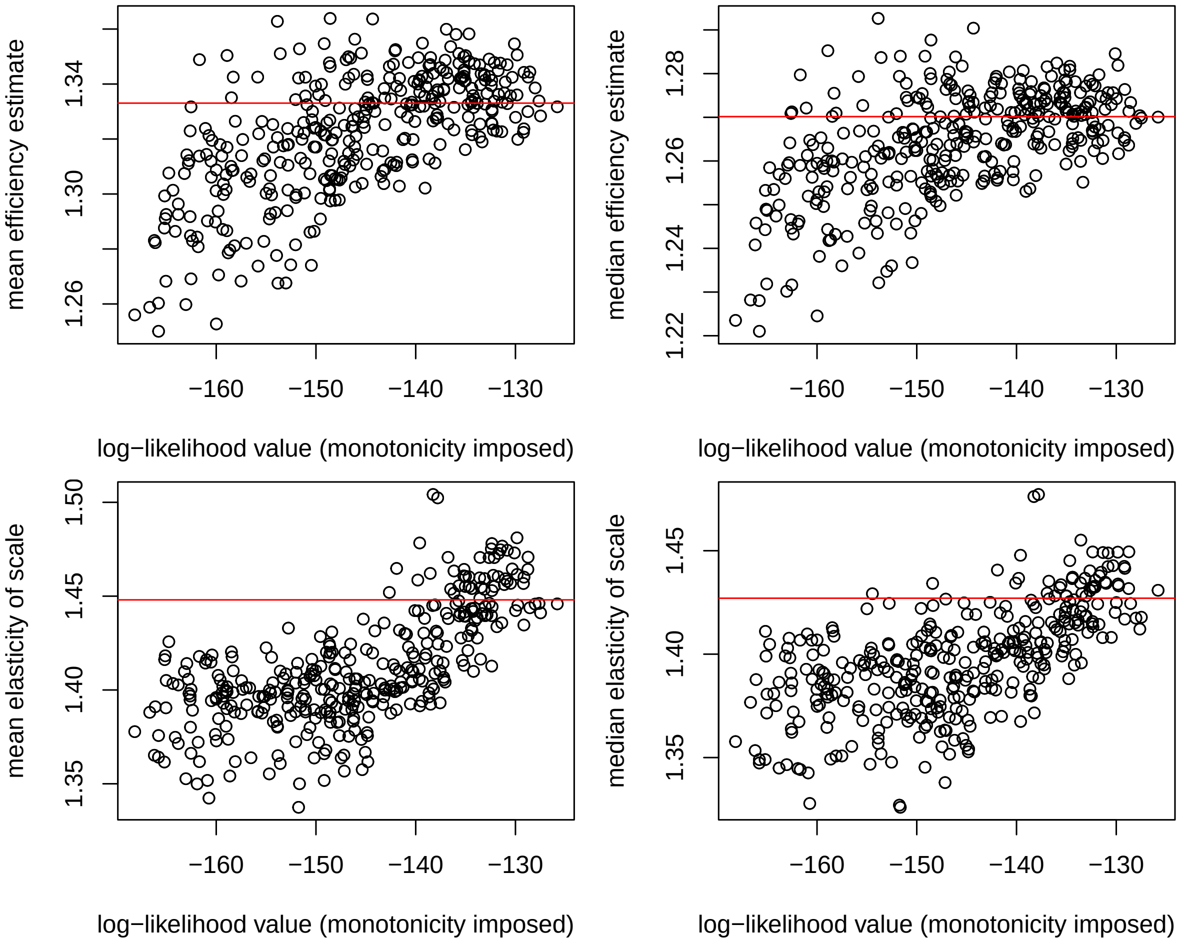

Figure 1: Relationships between the fit of the model and the mean and median technical efficiency estimates and the mean and median elasticities of scale for all possible orderings of outputs (the horizontal red line indicates the model-averaged estimate; models with a better fit, i.e., dots on the right-hand side of each panel, have a higher weight than models with a worse fit, i.e., dots on the left-hand side of each panel)

Figure 2 presents the model-averaged efficiency estimates and elasticities of scale while the corresponding distance elasticities of the inputs, outputs, and 'environmental' variables are presented in Figures 13, 14, and 15, respectively, in Appendix F. Average technical efficiency is 1.33 (median: 1.27), which indicates that the museums use, on average, 33% more of each input than required to produce the given output quantities. The average elasticity of scale is 1.44 (median: 1.43), indicating considerable productivity gains from getting larger. While almost all museums with less than 70 employees (full-time equivalents, FTE, including both scientific and non-scientific staff) and less than 300,000 visitors per year operate under substantial increasing returns to scale, most large museums experience decreasing returns to scale (see Figure 3). Hence, the most productive size is roughly around 70 employees (FTE) and 300,000 visitors per year.

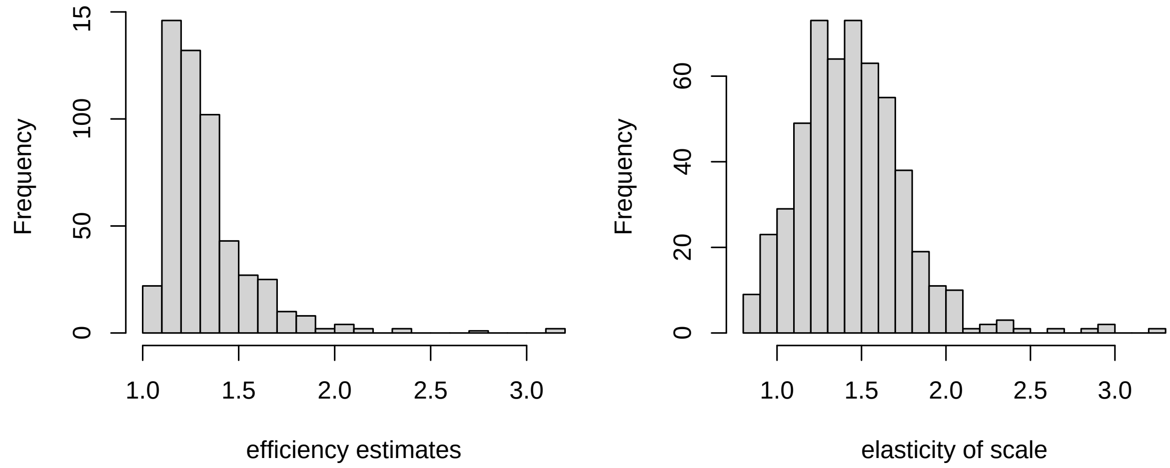

Figure 2: Model-averaged efficiency estimates and elasticities of scale

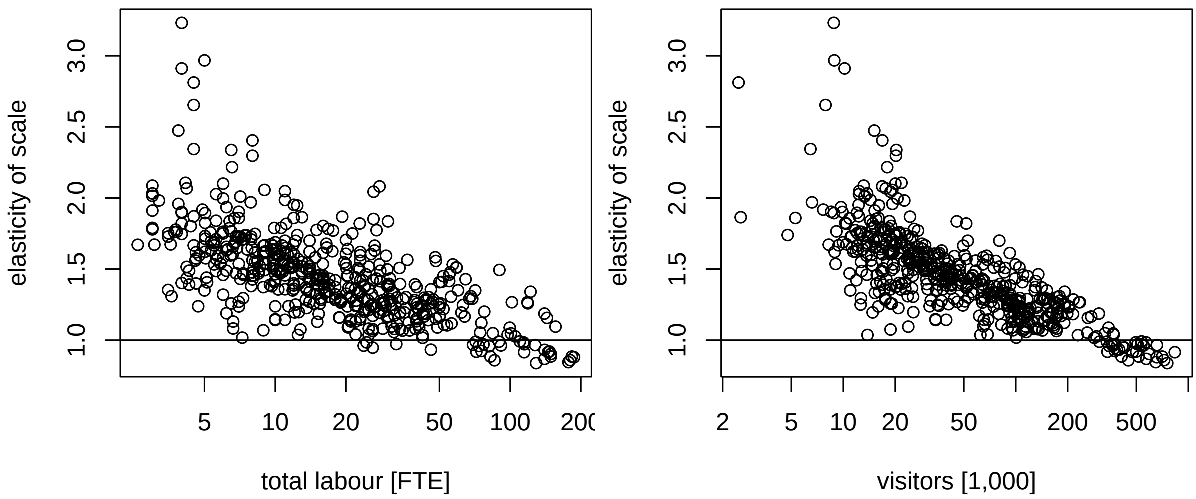

Figure 3: Relationship between the size of the museum and model-averaged elasticities of scale

Figure 4 presents the model-averaged distance elasticities and semi-elasticities of the two 'environmental' variables for museums with different numbers of visit sites. For museums with one to three visit sites, changing the number of visit sites can be positively or negatively related to productivity. However, for almost all museums with more than three visit sites, increasing the number of visit sites is related to a decrease in productivity. Given that museums with more than three visit sites usually have distance elasticities of visit sites between -0.2 and -0.4, increasing the number of visit sites by 10%, e.g., from 10 to 11 visit sites, requires an increase in all inputs of two to four percent to produce the same output quantities as with the current number of visit sites.[16] Hence, extending a museum to more than three visit sites seems to be associated with considerable additional costs.

The right panel of Figure 4 indicates that having additional responsibilities can be positively or negatively related to productivity for museums with just one visit site while having special responsibilities requires substantially higher input use for virtually all museums with more than one visit site. As only around 9 percent of museums with only one visit site have special responsibilities while around 71 percent of museums with more than one visit site have special responsibilities, we can conclude that the majority of museums that have special responsibilities need to substantially increase their input quantities (up to doubling the quantities as implied by a semi-elasticity of −1) to fulfill these additional tasks.

5 Conclusions

In this paper, we have derived an input-oriented distance function based on the stochastic ray production frontier, which is suitable for modeling production technologies based on logarithmic functional forms when control over inputs is greater than control over outputs and when some productive entities do not produce the entire set of outputs under analysis.

[16] The interpretation of distance elasticities of 'environmental' variables requires some thoughts: a distance elasticity of $\partial \ln D^i(x,y,z)/\partial \ln z_i = \Gamma_i$ indicates that increasing the 'environmental' variable z_i by one percent changes the distance measure $D^i(x,y,z)$ by Γ_i percent. As we want to assess the effect of 'environmental' variables on the technology rather than on the (in)efficiency level, our interpretations have to assume a constant distance measure $D^i(x,y,z)$, e.g., $D^i(x,y,z) = 1$, which means that we are looking at the production frontier. As the input distance function is linearly homogeneous in input quantities, i.e., $D^i(k \cdot x,y,z) = k \cdot D^i(x,y,z) \,\forall\, k > 0$, an increase in the distance measure $D^i(x,y,z)$ of one percent corresponds to an increase in all input quantities x of one percent. Similarly, in order to reverse a change in the distance measure $D^i(x,y,z)$ of Γ_i percent due to an increase in the 'environmental' variable z_i of one percent, all input quantities x need to change by $-\Gamma_i$ percent so that the distance measure $D^i(x,y,z)$ once again has the same value as it had before the change in the 'environmental' variable z_i. Hence, an increase in the 'environmental' variable z_i of one percent is related to a change in all inputs of $-\Gamma_i$ percent when holding the (in)efficiency level constant. Thus, if a distance elasticity of an 'environmental' variable is positive, i.e., $\partial \ln D^i(x,y,z)/\partial \ln z_i = \Gamma_i > 0$, an increase in the 'environmental' variable z_i of one percent is related to a decrease in all inputs of Γ_i percent and, thus, to an increase in productivity when holding the (in)efficiency level constant. A more formal (although perhaps less intuitive) derivation of this result can be obtained with the implicit function theorem holding the distance measure $D^i(k \cdot x,y,z)$ constant while allowing the input quantities to change proportionally by a factor $k > 0$. Hence, the relative change in all input quantities due to a change in the 'environmental' variable z_i evaluated at the original input quantities, i.e., $k = 1$, can be obtained by:

$$\frac{\partial \ln k}{\partial \ln z_i} = -\frac{\frac{\partial \ln D^i(k \cdot x,y,z)}{\partial \ln z_i}}{\frac{\partial \ln D^i(k \cdot x,y,z)}{\partial \ln k}} = -\frac{\partial \ln D^i(x,y,z)}{\partial \ln z_i} = -\Gamma_i.$$

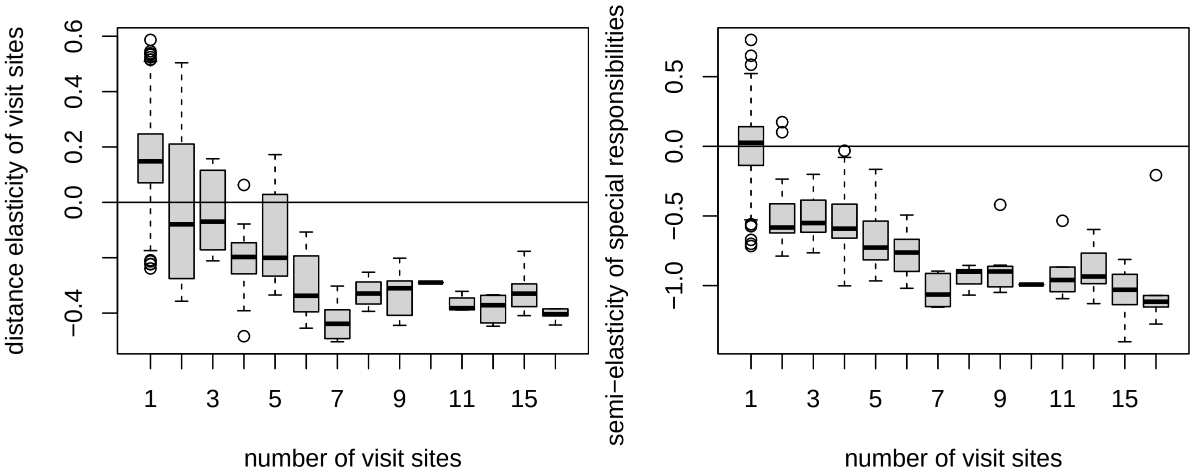

Figure 4: Relationship between the number of visit sites and model-averaged distance elasticities and semi-elasticities of the environmental variables

We have applied this model empirically to a data set of state-recognized Danish museums. As part of this empirical analysis, we have estimated the technical efficiencies, distance elasticities, and elasticity of scale and determined that some 'environmental' variables that are of interest to the museum sector are related to the required level of inputs. One has to be cautious when interpreting our results as causal effects because unobserved variables that affect the 'production technology' of the museums are certainly present and these may be correlated to some of our explanatory variables.[17] However, the main contribution of this paper is the development of ray-based input distance functions and the illustration that this specification is a useful tool for empirical analyses when inefficiency should be measured in an input-oriented way.

As part of this analysis, we also demonstrate how to impose monotonicity on ray-based input distance functions. Finally, we address a weakness of the stochastic ray function, namely its sensitivity to the ordering of the outputs. We follow two approaches. Firstly, we estimate the model with all possible ordering of outputs and select the ordering that outperforms the others based on the log-likelihood value. Secondly, we use model averaging over all possible orderings of the outputs. In our empirical application, the results based on the model with the highest log-likelihood value are very similar to our results based on model-averaging. However, whether this finding can be generalized or is just a coincidence in our empirical application is a question for future research.

[17]Time-invariant unobserved variables could be controlled for by estimating panel data models but—as previously mentioned in footnote 10—our data set does not allow us to do this.

References

Bishop, P. and Brand, S. (2003). The efficiency of museums: a stochastic frontier production function approach. *Applied Economics*, 35(17):1853–1858.

Dakpo, K. H., Desjeux, Y., and Latruffe, L. (2021). *sfaR: Stochastic Frontier Analysis using R*. R package version 0.1.0.

del Barrio, M. J., Herrero, L. C., and Ángel Sanz, J. (2009). Measuring the efficiency of heritage institutions: A case study of a regional system of museums in spain. *Journal of Cultural Heritage*, 10(2):258 – 268.

del Barrio-Tellado, M. J. and Herrero-Prieto, L. C. (2019). Modelling museum efficiency in producing inter-reliant outputs. *Journal of Cultural Economics*, 43(3):485–512.

Feldstein, M. (1991). *The Economics of Art Museums*. University of Chicago Press.

Frey, B. and Meier, S. (2006). The economics of museums. In Ginsburgh, V. and Throsby, D., editors, *Handbook of the Economics of Art and Culture: Volume 1*, volume 1, chapter 29, pages 1017–1047. Elsevier, 1 edition.

Färe, R. and Primont, D. (1990). A distance function approach to multioutput technologies. *Southern Economic Journal*, 56(4):879–891.

Grampp, W. D. (1989). *Pricing the priceless: art, artists, and economics / William D. Grampp*. Basic Books New York.

Henningsen, A., Bělín, M., and Henningsen, G. (2017). New insights into the stochastic ray production frontier. *Economics Letters*, 156:18 – 21.

Henningsen, A. and Henning, C. H. C. A. (2009). Imposing regional monotonicity on translog stochastic production frontiers with a simple three-step procedure. *Journal of Productivity Analysis*, 32(3):217.

Henningsen, G., Henningsen, A., and Jensen, U. (2015). A monte carlo study on multiple output stochastic frontiers: a comparison of two approaches. *Journal of Productivity Analysis*, 44(3):309–320.

Huang, C. and Lai, H.-p. (2012). Estimation of stochastic frontier models based on multimodel inference. *Journal of Productivity Analysis*, 38(3):273–284.

Kumbhakar, S., Wang, H.-J., and Horncastle, A. P. (2015). *A Practitioner's Guide to Stochastic Frontier Analysis Using Stata*. Cambridge University Press.

Kumbhakar, S. C. and Lovell, C. A. K. (2000). *Stochastic Frontier Analysis*. Cambridge University Press.

Löthgren, M. (1997). Generalized stochastic frontier production models. *Economics Letters*, 57(3):255 – 259.

Löthgren, M. (2000). Specification and estimation of stochastic multiple-output production and technical inefficiency. *Applied Economics*, 32(12):1533–1540.

N'guessan, Y. G., Featherstone, A., Odeh, O., and Upendram, S. (2017). Choice of the empirical definition of zero in the translog multiproduct cost functional form. *Applied Economics Letters*, 24(15):1112–1120.

O'Donnell, C. and Coelli, T. (2005). A bayesian approach to imposing curvature on distance functions. *Journal of Econometrics*, 126(2):493–523.

O'Hagan, J. W. (1998). Art museums: Collections, deaccessioning and donations. *Journal of Cultural Economics*, 22(2/3):197–207.

Peacock, A. and Godfrey, C. (1974). The economics of museums and galleries. In Towse, R., editor, *Cultural Economics: The Arts, The Heritage and the Media Industries*. Edward Elgar.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Tsionas, M., Izzeldin, M., Henningsen, A., and Paravalos, E. (2021). Addressing endogeneity when estimating stochastic ray production frontiers: A Bayesian approach. *Empirical Economics*, forthcoming.

Turlach, B. A. and Weingessel, A. (2011). *quadprog: Functions to solve Quadratic Programming Problems.* R package version 1.5-4.

A Derivation of the ray-based Translog input distance function

Using the notation of Henningsen et al. (2017), a Translog stochastic ray production frontier is defined as:

$$\begin{aligned}\ln\left\|y\right\| = \alpha_0^* &+ \sum_{i=1}^{N}\alpha_i^*\ln x_i + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_{ij}^*\ln x_i\ln x_j \\ &+ \sum_{i=1}^{M-1}\beta_i^*\varphi_i(y) + \frac{1}{2}\sum_{i=1}^{M-1}\sum_{j=1}^{M-1}\beta_{ij}^*\varphi_i(y)\varphi_j(y) \\ &+ \sum_{i=1}^{K}\delta_i^* z_i + \frac{1}{2}\sum_{i=1}^{K}\sum_{j=1}^{K}\delta_{ij}^* z_i z_j \\ &+ \sum_{i=1}^{N}\sum_{j=1}^{M-1}\psi_{ij}^*\ln x_j\,\varphi_i(y) + \sum_{i=1}^{N}\sum_{j=1}^{K}\xi_{ij}^*\ln x_i z_j \\ &+ \sum_{i=1}^{M-1}\sum_{j=1}^{K}\zeta_{ij}^*\varphi_i(y) z_j - u^* + v^*\end{aligned} \tag{15}$$

with the normalizations $\alpha_{ij}^* = \alpha_{ji}^*\,\forall\, i,j = 1,\ldots,N$, $\beta_{ij}^* = \beta_{ji}^*\,\forall\, i,j = 1,\ldots,M-1$, and $\delta_{ij}^* = \delta_{ji}^*\,\forall\, i,j = 1,\ldots,K$. Function (15) can be seen as a Shephard output distance function (Henningsen et al., 2015):

$$\begin{aligned}\ln D^o(x,y,z) = -u^* = \alpha_0 &+ \sum_{i=1}^{N}\alpha_i\ln x_i + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_{ij}\ln x_i\ln x_j \\ &+ \sum_{i=1}^{M-1}\beta_i\varphi_i(y) + \ln\left\|y\right\| + \frac{1}{2}\sum_{i=1}^{M-1}\sum_{j=1}^{M-1}\beta_{ij}\varphi_i(y)\varphi_j(y) \\ &+ \sum_{i=1}^{K}\delta_i z_i + \frac{1}{2}\sum_{i=1}^{K}\sum_{j=1}^{K}\delta_{ij} z_i z_j \\ &+ \sum_{i=1}^{N}\sum_{j=1}^{M-1}\psi_{ij}\ln x_i\,\varphi_j(y) + \sum_{i=1}^{N}\sum_{j=1}^{K}\xi_{ij}\ln x_i z_j \\ &+ \sum_{i=1}^{M-1}\sum_{j=1}^{K}\zeta_{ij}\varphi_i(y) z_j + v,\end{aligned} \tag{16}$$

where $D^o(x,y,z) = e^{-u^*}$ with $0 \le D^o(x,y,z) \le 1$ is a Shephard output distance function and $\alpha_0 = -\alpha_0^*$, $\alpha_i = -\alpha_i^*\,\forall\, i = 1,\ldots,N$, $\alpha_{ij} = -\alpha_{ij}^*\,\forall\, i,j = 1,\ldots,N$, $\beta_i = -\beta_i^*\,\forall\, i = 1,\ldots M$,

$\beta_{ij} = -\beta^*_{ij} \,\forall\, i,j = 1,\dots M$, $\delta_i = -\delta^*_i \,\forall\, i = 1,\dots,K$, $\delta_{ij} = -\delta^*_{ij} \,\forall\, i,j = 1,\dots,K$, $\psi_{ij} = -\psi^*_{ij} \,\forall\, i = 1,\dots,N; j = 1,\dots,M$, $\xi_{ij} = -\xi^*_{ij} \,\forall\, i = 1,\dots,N; j = 1,\dots,K$, $\zeta_{ij} = -\zeta^*_{ij} \,\forall\, i = 1,\dots,M; j = 1,\dots,K$, and $v = -v^*$.

This specification of a stochastic-ray-based output distance function can be generalized to the following 'full' Translog distance function:

$$\begin{aligned}
\ln D(x,y,z) = {} & \alpha_0 - \sum_{i=1}^{N} \alpha_i \ln x_i + \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_{ij} \ln x_i \ln x_j \\
& + \sum_{i=1}^{M-1} \beta_i \varphi_i(y) + \beta_M \ln \|y\| + \frac{1}{2} \sum_{i=1}^{M-1} \sum_{j=1}^{M-1} \beta_{ij} \varphi_i(y) \varphi_j(y) \\
& + \sum_{i=1}^{M-1} \beta_{iM} \varphi_i(y) \ln \|y\| + \frac{1}{2} \beta_{MM} \left(\ln \|y\| \right)^2 \\
& + \sum_{i=1}^{K} \delta_i z_i + \frac{1}{2} \sum_{i=1}^{K} \sum_{j=1}^{K} \delta_{ij} z_i z_j \\
& + \sum_{i=1}^{N} \sum_{j=1}^{M-1} \psi_{ij} \ln x_i \, \varphi_j(y) + \sum_{i=1}^{N} \psi_{iM} \ln x_i \ln \|y\| \\
& + \sum_{i=1}^{N} \sum_{j=1}^{K} \xi_{ij} \ln x_i z_j \\
& + \sum_{i=1}^{M-1} \sum_{j=1}^{K} \zeta_{ij} \varphi_i(y) z_j + \sum_{i=1}^{K} \zeta_{Mi} \ln \|y\| z_i + v
\end{aligned} \tag{17}$$

where linear homogeneity in output quantities (as required by a Shephard output distance function) requires $\beta_M = 1$, $\beta_{iM} = 0 \,\forall\, i = 1,\dots,M$, $\psi_{Mi} = 0 \,\forall\, i = 1,\dots,N$, $\zeta_{Mi} = 0 \,\forall\, i = 1,\dots,K$.

If restrictions $\sum_{i=1}^{N} \alpha_i = 1$, $\sum_{i=1}^{N} \alpha_{ij} = 0 \,\forall\, j = 1,\dots,N \Leftrightarrow \sum_{j=1}^{N} \alpha_{ij} = 0 \,\forall\, i = 1,\dots,N$, $\sum_{i=1}^{N} \psi_{ij} = 0 \,\forall\, j = 1,\dots,M$, and $\sum_{i=1}^{N} \xi_{ij} = 0 \,\forall\, j = 1,\dots,K$ are fulfilled, function (17) is linearly homogeneous in input quantities and, thus, it can be used as a specification for a Shephard input distance function. We can impose these restrictions by replacing α_N by $1 - \sum_{i=1}^{N-1} \alpha_i$, replacing all $\alpha_{Nj}, j = 1,\dots,N$ by $-\sum_{i=1}^{N-1} \alpha_{ij}$, replacing all $\alpha_{iN}, i = 1,\dots,N$ by $-\sum_{j=1}^{N-1} \alpha_{ij}$, replacing all $\psi_{Nj}, j = 1,\dots,M$ by $-\sum_{i=1}^{N-1} \psi_{ij}$, and replacing all $\xi_{Nj}, j = 1,\dots,K$ by $-\sum_{i=1}^{N-1} \xi_{ij}$ so that we get a Shephard input distance function:

$$\ln D^i(x,y,z) = \alpha_0 + \sum_{i=1}^{N-1} \alpha_i \ln(\tilde{x}_i) + \ln x_N + \frac{1}{2}\sum_{i=1}^{N-1}\sum_{j=1}^{N-1} \alpha_{ij} \ln(\tilde{x}_i) \ln\left(\tilde{x}_j\right) \tag{18}$$

$$+ \sum_{i=1}^{M-1} \beta_i \varphi_i(y) + \beta_M \ln\left\|y\right\| + \frac{1}{2}\sum_{i=1}^{M-1}\sum_{j=1}^{M-1} \beta_{ij}\varphi_i(y)\varphi_j(y)$$

$$+ \sum_{m=1}^{M-1} \beta_{iM}\varphi_i(y) \ln\left\|y\right\| + \frac{1}{2}\beta_{MM}\left(\ln\left\|y\right\|\right)^2$$

$$+ \sum_{i=1}^{K} \delta_i z_i + \frac{1}{2}\sum_{i=1}^{K}\sum_{j=1}^{K} \delta_{ij} z_i z_j$$

$$+ \sum_{i=1}^{N-1}\sum_{j=1}^{M-1} \psi_{ij} \ln(\tilde{x}_i)\varphi_j(y) + \sum_{i=1}^{N-1} \psi_{iM} \ln(\tilde{x}_i) \ln\left\|y\right\|$$

$$+ \sum_{i=1}^{N-1}\sum_{j=1}^{K} \xi_{ij} \ln(\tilde{x}_i) z_j$$

$$+ \sum_{i=1}^{M-1}\sum_{j=1}^{K} \zeta_{ij}\varphi_i(y) z_j + \sum_{i=1}^{K} \zeta_{Mi} \ln\left\|y\right\| z_i + v.$$

By replacing the logarithm of the Shephard input distance measure $D^i(x,y,z) \geq 1$, i.e., $\ln D^i(x,y,z) \geq 0$, by $u \geq 0$ and a little re-arranging, we get:

$$-\ln x_N = \alpha_0 + \sum_{i=1}^{N-1} \alpha_i \ln(\tilde{x}_i) + \frac{1}{2}\sum_{i=1}^{N-1}\sum_{j=1}^{N-1} \alpha_{ij} \ln(\tilde{x}_i) \ln\left(\tilde{x}_j\right) \tag{19}$$

$$+ \sum_{i=1}^{M-1} \beta_i \varphi_i(y) + \beta_M \ln\left\|y\right\| + \frac{1}{2}\sum_{i=1}^{M-1}\sum_{j=1}^{M-1} \beta_{ij}\varphi_i(y)\varphi_j(y)$$

$$+ \sum_{i=1}^{M-1} \beta_{iM}\varphi_i(y) \ln\left\|y\right\| + \frac{1}{2}\beta_{MM}\left(\ln\left\|y\right\|\right)^2$$

$$+ \sum_{i=1}^{K} \delta_i z_i + \frac{1}{2}\sum_{i=1}^{K}\sum_{j=1}^{K} \delta_{ij} z_i z_j$$

$$+ \sum_{i=1}^{N-1}\sum_{j=1}^{M-1} \psi_{ij} \ln(\tilde{x}_i)\varphi_j(y) + \sum_{i=1}^{N-1} \psi_{iM} \ln(\tilde{x}_i) \ln\left\|y\right\|$$

$$+ \sum_{i=1}^{N-1}\sum_{j=1}^{K} \xi_{ij} \ln(\tilde{x}_i) z_j$$

$$+\sum_{i=1}^{M-1}\sum_{j=1}^{K}\zeta_{ij}\varphi_i(y)z_j+\sum_{i=1}^{K}\zeta_{Mi}\ln\|y\|z_i-u+v,$$

which can be easily estimated as a stochastic frontier model[18]. The Cobb-Douglas functional form is a special case with $\alpha_{ij}=0\,\forall\,i,j=1,\dots,N$, $\beta_{ij}=0\,\forall\,i,j=1,\dots M$, $\delta_{ij}=0\,\forall\,i,j=1,\dots,K$, $\psi_{ij}=0\,\forall\,i=1,\dots,N;j=1,\dots,M$, $\xi_{ij}=0\,\forall\,i=1,\dots,N;j=1,\dots,K$, and $\zeta_{ij}=0\,\forall\,i=1,\dots,M;j=1,\dots,K$.

B Monotonicity conditions in matrix form

We represent these monotonicity restrictions in matrix form as $R\theta\geq r$, where $\theta=(\alpha_0,\alpha_1,\dots,\alpha_{N-1},\alpha_{11},\dots,\alpha_{1,N-1},\alpha_{22},\dots,\alpha_{2,N-1},\dots,\alpha_{N-1,N-1},\beta_1,\dots,\beta_M,\beta_{11},\dots,\beta_{1M},$ $\beta_{22},\dots,\beta_{2M},\dots,\beta_{MM},\ \delta_1,\dots,\delta_K,\ \delta_{11},\dots,\delta_{1K},\ \delta_{22},\dots,\delta_{2K},\dots,\delta_{KK},\ \psi_{11},\dots,\psi_{1M},\ \psi_{21},\dots,$ $\psi_{2M},\dots,\psi_{N-1,M},\xi_{11},\dots,\xi_{1K},\xi_{21},\dots,\xi_{2K},\dots,\xi_{N-1,K},\zeta_{11},\dots,\zeta_{1K},\zeta_{21},\dots,\zeta_{2K},\dots,\zeta_{MK})$, is a vector of all estimated coefficients, $R=[R_{i,\theta}]$ is a matrix with one row i for each of the $N+M$ monotonicity restrictions and one column for each of the coefficients in θ, and r is a vector with one element for each of the $N+M$ monotonicity restrictions. The elements of $R=[R_{i,\theta}]$ are:

$$R_{i,\alpha_o}=0\,\forall\,i=1,\dots,N+M \tag{20}$$

$$R_{i,\alpha_j}=\begin{cases}1 & \text{if } i=j\\ 0 & \text{if } i\neq j\end{cases}\ \forall\,i=1,\dots,N-1;j=1,\dots,N-1 \tag{21}$$

$$R_{N,\alpha_j}=-1\,\forall\,j=1,\dots,N-1 \tag{22}$$

$$R_{N+i,\alpha_j}=0\,\forall\,i=1,\dots,M;j=1,\dots,N-1 \tag{23}$$

$$R_{i,\alpha_{jk}}=\begin{cases}0 & \text{if } i\neq j\wedge i\neq k\\ \ln(\tilde{x}_k) & \text{if } i=j\\ \ln\left(\tilde{x}_j\right) & \text{if } i=k\end{cases}\quad\forall\,i=1,\dots,N-1;j=1,\dots,N-1;k=j,\dots,N-1 \tag{24}$$

$$R_{N,\alpha_{jk}}=-\left(\ln\left(\tilde{x}_j\right)+I_{j\neq k}\ln(\tilde{x}_k)\right)\forall\,j=1,\dots,N-1;k=j,\dots,N-1 \tag{25}$$

$$R_{N+i,\alpha_{jk}}=0\,\forall\,i=1,\dots,M;j=1,\dots,N-1;k=1,\dots,N-1 \tag{26}$$

$$R_{i,\beta_j}=0\,\forall\,i=1,\dots,N;j=1,\dots,M \tag{27}$$

$$R_{N+i,\beta_j}=\begin{cases}-\Omega_{ji} & \text{if } j<M\\ -y_i/\|y\|^2 & \text{if } j=M\end{cases}\ \forall\,i=1,\dots,M;j=1,\dots,M-1 \tag{28}$$

[18]This equation is identical to equation (6).

$$R_{i,\beta_{jk}} = 0\,\forall\, i = 1,\ldots,N; j = 1,\ldots,M; k = 1,\ldots,M \tag{29}$$

$$R_{N+i,\beta_{jk}} = \begin{cases} -\left(\varphi_k(y)\,\Omega_{ji} + I_{j\neq k}\varphi_j(y)\,\Omega_{ki}\right) & \text{if } j < M \wedge k < M \\ -\Omega_{ji}\ln\|y\| - \varphi_j(y)\,y_i/\|y\|^2 & \text{if } j < M \wedge k = M \\ -\ln\|y\|\,y_i/\|y\|^2 & \text{if } j = M \wedge k = M \end{cases} \forall\, i = 1,\ldots,M; j = 1,\ldots,M; k = j,\ldots,M \tag{30}$$

$$R_{i,\delta_j} = 0\,\forall\, i = 1,\ldots,N+M; j = 1,\ldots,K \tag{31}$$

$$R_{i,\delta_{jk}} = 0\forall\, i = 1,\ldots,N+M; j = 1,\ldots,K; k = 1,\ldots,K \tag{32}$$

$$R_{i,\psi_{jk}} = \begin{cases} 0 & \text{if } i \neq j \\ \varphi_k(y) & \text{if } i = j \wedge k < M \\ \ln\|y\| & \text{if } i = j \wedge k = M \end{cases} \forall\, i = 1,\ldots,N-1; j = 1,\ldots,N-1; k = 1,\ldots,M \tag{33}$$

$$R_{N,\psi_{jk}} = \begin{cases} -\varphi_k(y) & \text{if } k < M \\ -\ln\|y\| & \text{if } k = M \end{cases} \forall\, j = 1,\ldots,N-1; k = 1,\ldots,M \tag{34}$$

$$R_{N+i,\psi_{jk}} = \begin{cases} -\ln\left(\tilde{x}_j\right)\Omega_{ki} & \text{if } k < M \\ -\ln\left(\tilde{x}_j\right)y_i/\|y\|^2 & \text{if } k = M \end{cases} \forall\, i = 1,\ldots,M; j = 1,\ldots,N-1; k = 1,\ldots,M \tag{35}$$

$$R_{i,\xi_{jk}} = \begin{cases} 0 & \text{if } i \neq j \\ z_k & \text{if } i = j \end{cases} \forall\, i = 1,\ldots,N-1; j = 1,\ldots,N-1; k = 1,\ldots,K \tag{36}$$

$$R_{N,\xi_{jk}} = -z_k\,\forall\, j = 1,\ldots,N-1; k = 1,\ldots,K \tag{37}$$

$$R_{N+i,\xi_{jk}} = 0\,\forall\, i = 1,\ldots,M \,\wedge\, \forall\, j = 1,\ldots,M \,\wedge\, \forall\, k = 1,\ldots,K \tag{38}$$

$$R_{i,\zeta_{jk}} = 0\,\forall\, i = 1,\ldots N; j = 1,\ldots,N; k = 1,\ldots,K \tag{39}$$

$$R_{N+i,\zeta_{jk}} = \begin{cases} -z_k * \Omega_{ji} & \text{if } j < M \\ -z_k y_i/\|y\|^2 & \text{if } j = M \end{cases} \forall\, i = 1,\ldots,M; j = 1,\ldots,M; k = 1,\ldots,K \tag{40}$$

and the elements of r are:

$$r_i = \begin{cases} -1 & \text{if } i = N \\ 0 & \text{if } i \in \{1,\ldots,N-1,N+1,\ldots,N+M\} \end{cases} \tag{41}$$

The monotonicity restrictions $R\theta \geq r$ defined above impose monotonicity at one data point $(x_1,\ldots,x_N, y_1,\ldots,y_M, z_1,\ldots,z_K)$. If one wants to impose monotonicity at more than one data point, e.g., at each observation in a data set, one needs to create one R matrix and one r vector for each data point where monotonicity should be imposed and stack these matrices and vectors vertically:

$$\tilde{R} = \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_G \end{bmatrix} \tag{42}$$

$$\tilde{r} = \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_G \end{pmatrix}, \tag{43}$$

where G is the number of data points where monotonicity should be imposed, $R_g; g = 1, \ldots, G$ is the matrix R at the gth data point where monotonicity should be imposed, and $r_g; g = 1, \ldots, G$ is the vector r at the gth data point where monotonicity should be imposed. In this case, $\tilde{R}\theta \geq \tilde{r}$ defines the $G \cdot (N + M)$ monotonicity restrictions.

C Procedure to impose monotonicity conditions

Following the three-step procedure proposed by Henningsen and Henning (2009)[19], in the first step, we estimate the following (unrestricted) stochastic frontier model:

$$-\ln x_N = \ln D\left(\frac{x}{x_N}, y, z, \theta_u\right) - u + v, \tag{44}$$

where θ_u is the vector of the estimated unrestricted coefficients.

In the second step, we use a constrained optimization procedure to find the vector of restricted coefficients. This is a minimum distance estimation, i.e., we search for the vector of coefficients that is closest to (has the minimum distance to) the unrestricted vector of coefficients, conditional on fulfilling the monotonicity conditions and using the inverse of the variance-covariance matrix of the unrestricted coefficients as weighting matrix:

$$\hat{\theta}_r = \underset{\hat{\theta}_r}{\arg\min} \left(\hat{\theta}_r - \hat{\theta}_u\right)^\top \hat{\Sigma}^{-1} \left(\hat{\theta}_r - \hat{\theta}_u\right),$$
$$\textbf{s.t. } R\hat{\theta}_r \geq r,$$

[19]Other methods for imposing monotonicity conditions include O'Donnell and Coelli (2005).

where $\hat{\theta}_r$ is the vector of the estimated monotonicity-restricted coefficients and $\hat{\Sigma}$ is the estimated variance-covariance matrix of the unrestricted coefficients $\hat{\theta}_u$.

In the third-step, we estimate the stochastic frontier model:

$$-\ln x_N - \ln D\left(\frac{x}{x_N}, y, z, \hat{\theta}_r\right) = a_0 - u + v \tag{45}$$

$$-\ln D\left(x, y, z, \hat{\theta}_r\right) = a_0 - u + v. \tag{46}$$

In contrast to equation (12) in Henningsen and Henning (2009), we only adjust the intercept in $\hat{\theta}_r$, i.e., α_0, but we do not adjust the other coefficients in $\hat{\theta}_r$ as done by Henningsen and Henning (2009). This is because the input distance function must be linearly homogeneous in input quantities, and an adjustment of the slope coefficient in $\hat{\theta}_r$ would either abandon the homogeneity restriction or—if one adjusts the non-estimated coefficients to maintain homogeneity—potentially abandon the monotonicity restrictions (that were just imposed). In order to avoid these problems, we do not adjust the slope coefficients in $\hat{\theta}_r$ by having the predicted distance values $D\left(x/x_N, y, z, \hat{\theta}_r\right)$ on the left-hand side of equation (45) rather than as explanatory variable on the right-hand side (as already suggested by Henningsen and Henning, 2009, in their footnote 7). However, we suggest adjusting the intercept in $\hat{\theta}_r$, i.e., α_0, by estimating coefficient a_0 (rather than restricting it to zero) because this allows the "height" of the restricted frontier and, thus, the general (in)efficiency level to change after restricting the shape of the frontier (by imposing monotonicity conditions) as the latter may change the general (in)efficiency level.

D Results based on the ordering of outputs that gives the best fit

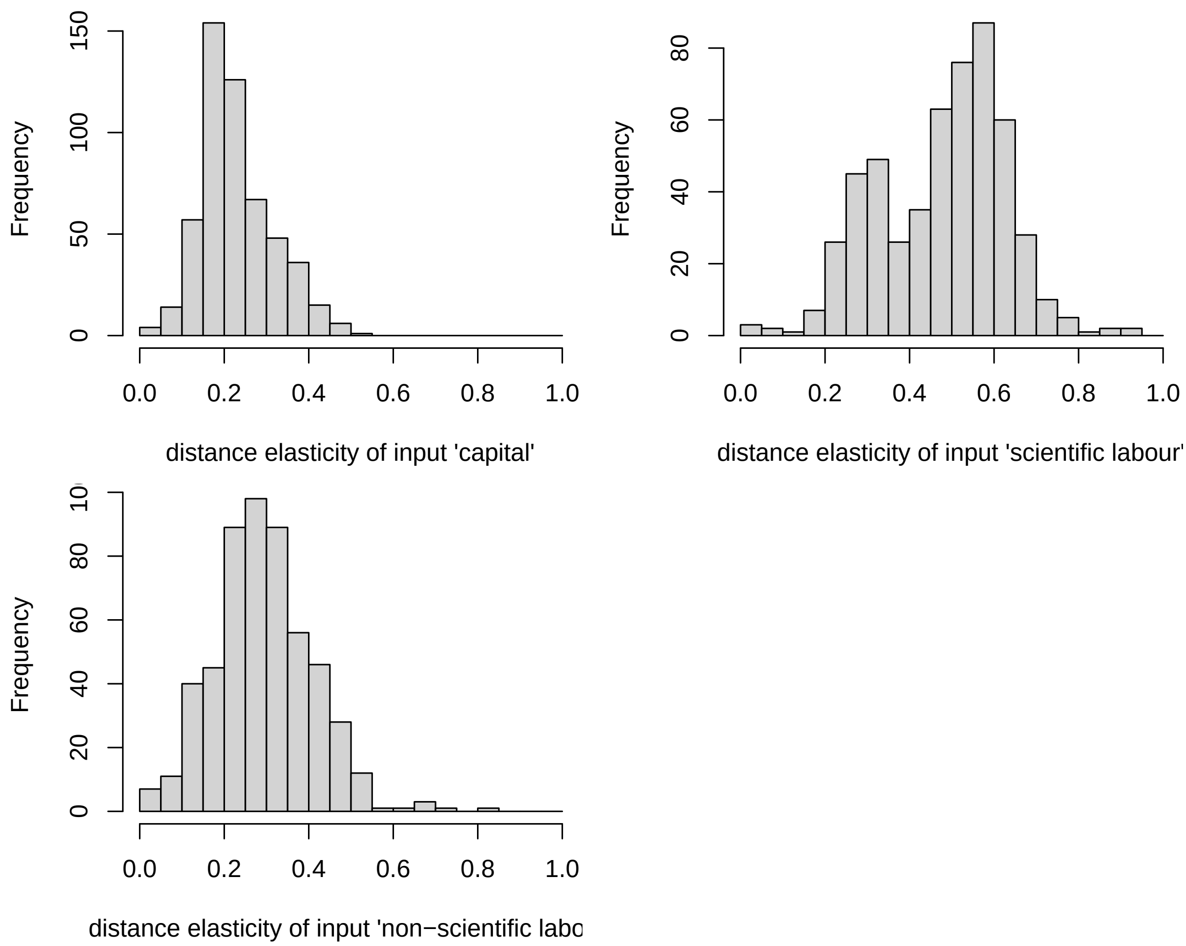

Figure 5: Distance elasticities of the inputs based on the ordering of outputs that gives the best fit

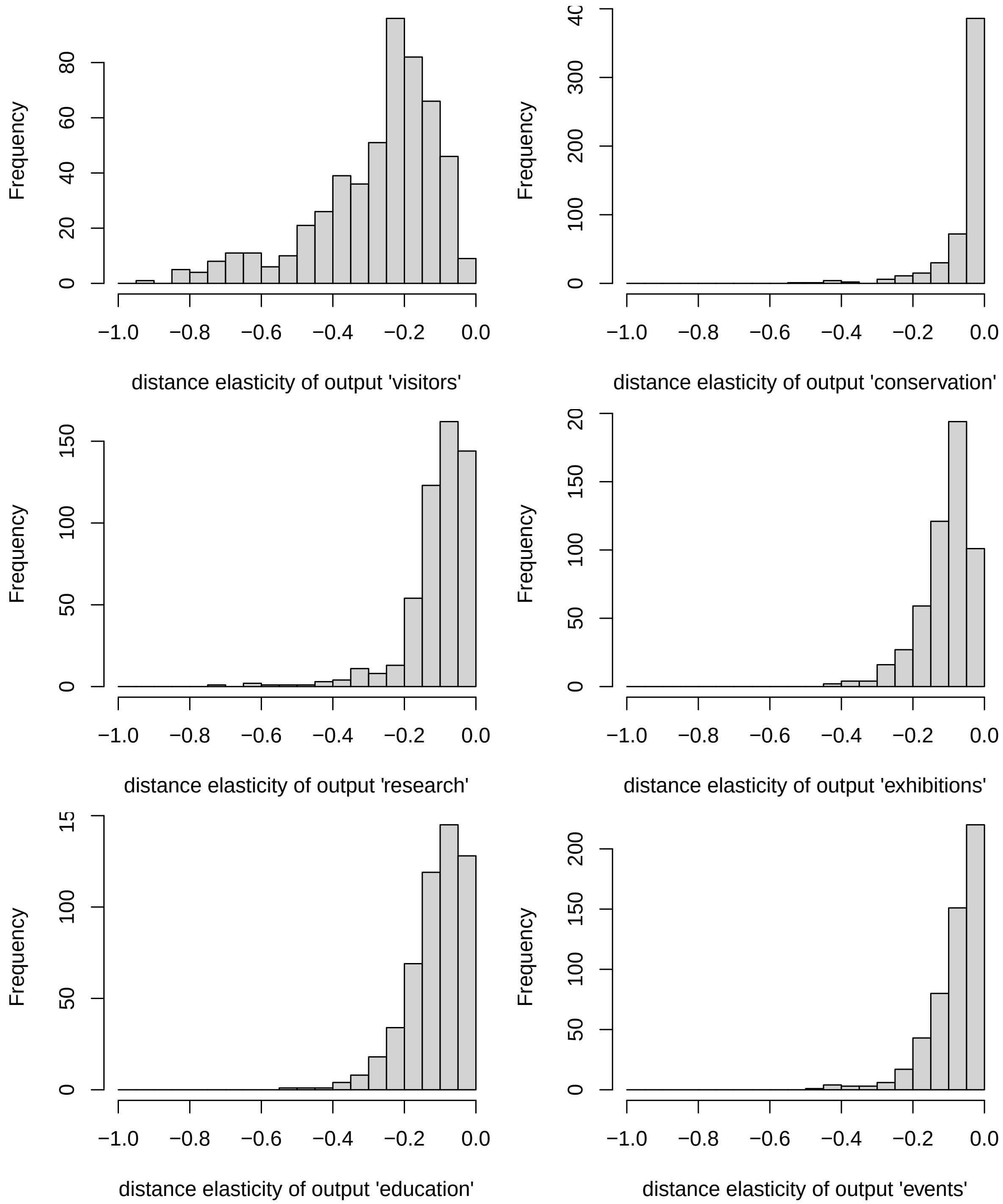

Figure 6: Distance elasticities of the outputs based on the ordering of outputs that gives the best fit

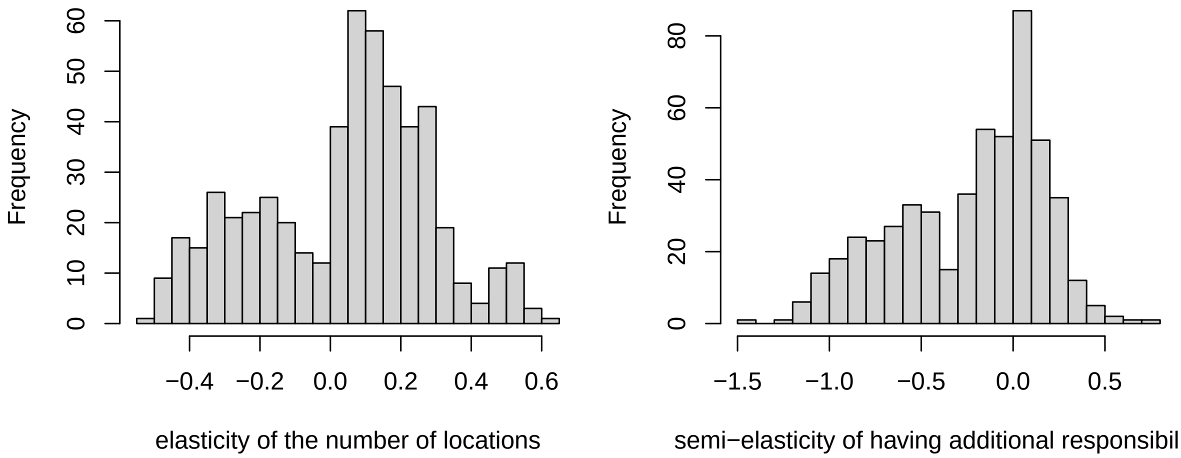

Figure 7: Elasticities and semi-elasticities of the environmental variables based on the ordering of outputs that gives the best fit

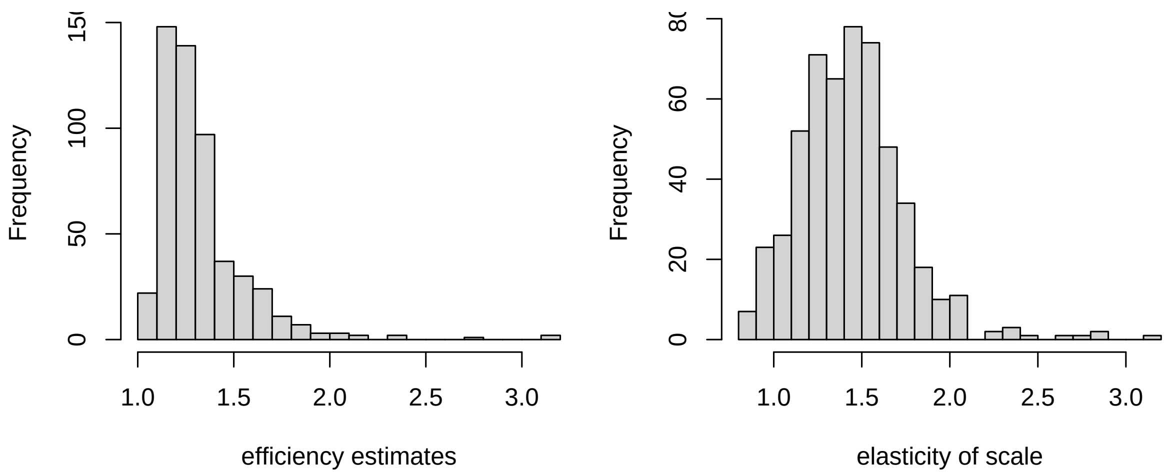

Figure 8: Efficiency estimates and elasticities of scale based on the ordering of outputs that gives the best fit

E Comparison of model-averaged estimates and estimates of the selected model

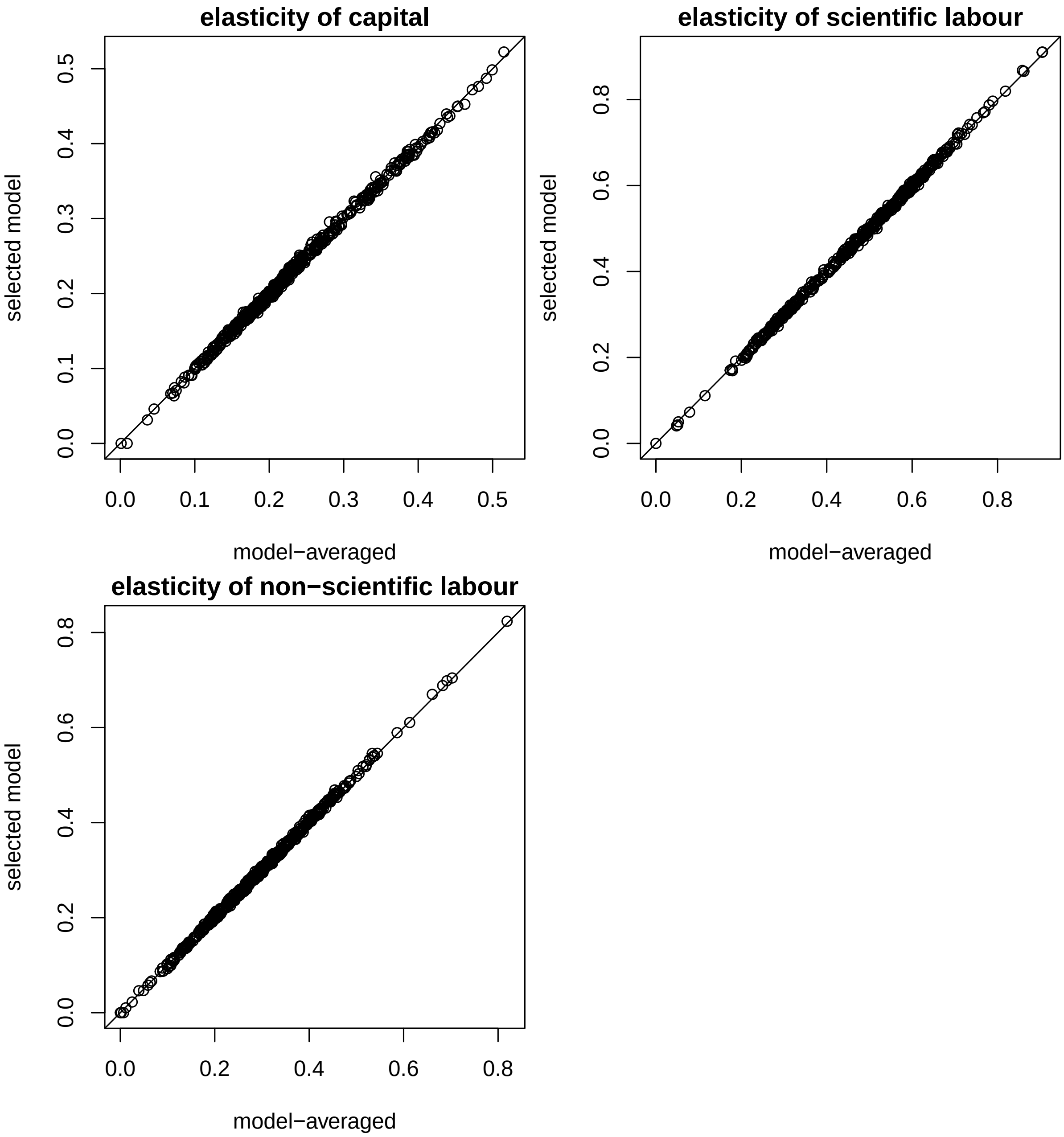

Figure 9: Comparison of the distance elasticities of the inputs based on model averaging and the ordering of outputs that gives the best fit

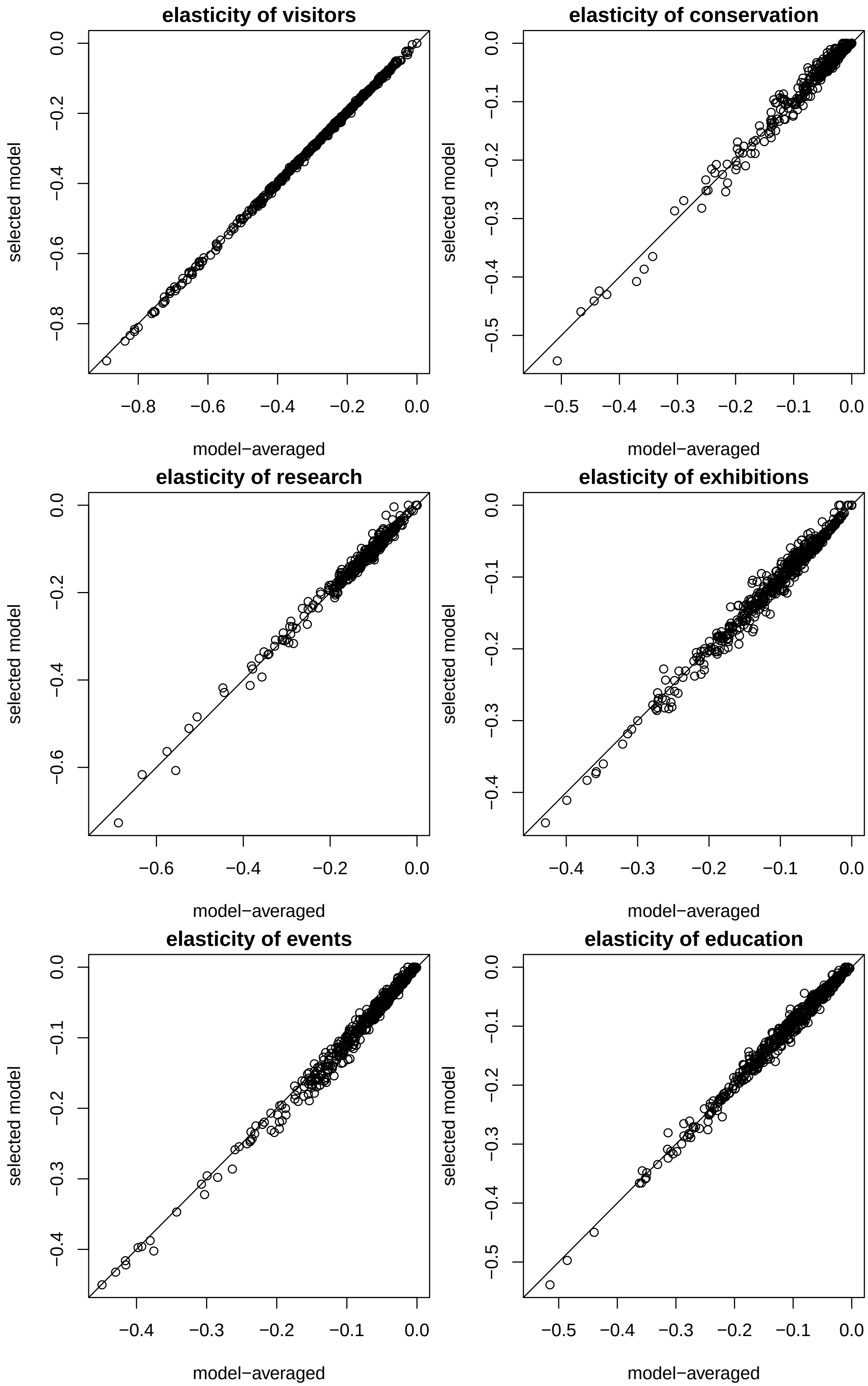

Figure 10: Comparison of the distance elasticities of the outputs based on model averaging and the ordering of outputs that gives the best fit

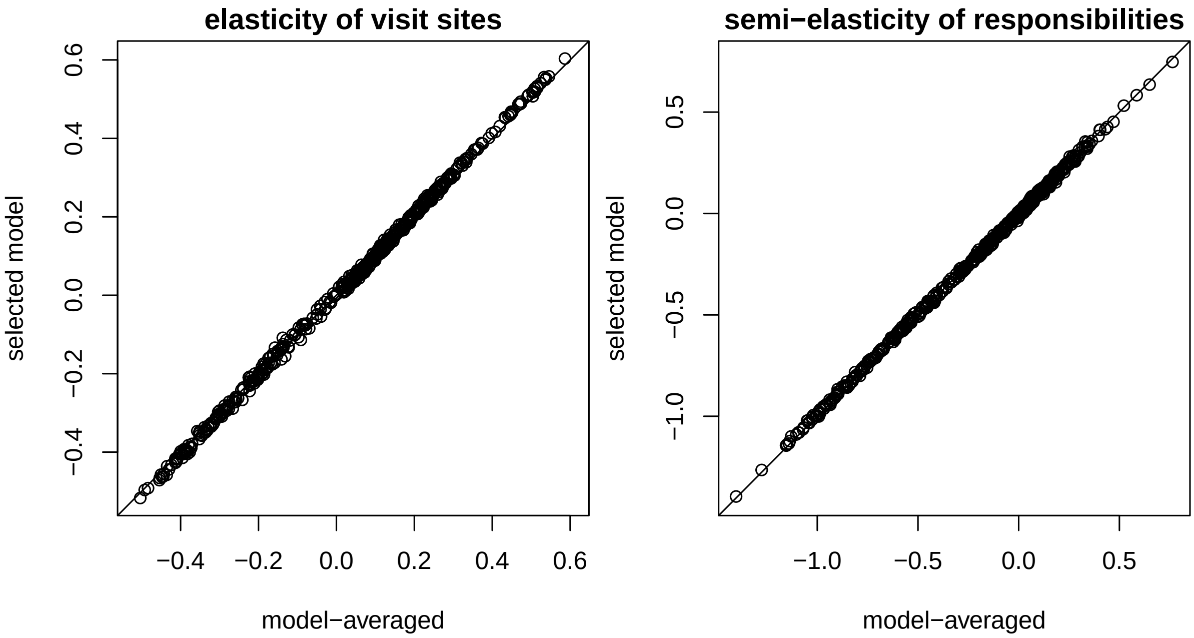

Figure 11: Comparison of the distance (semi-)elasticities of the 'environmental' variable based on model averaging and the ordering of outputs that gives the best fit

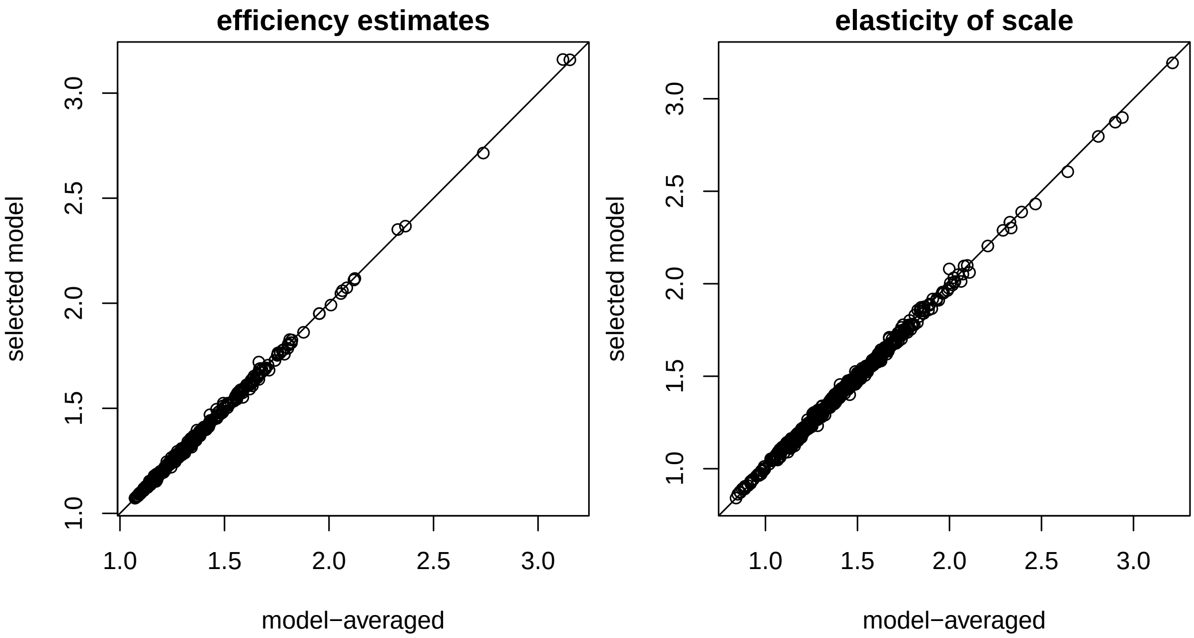

Figure 12: Comparison of the distance efficiency estimates and elasticities of scale based on model averaging and the ordering of outputs that gives the best fit

F Model-averaged distance elasticities

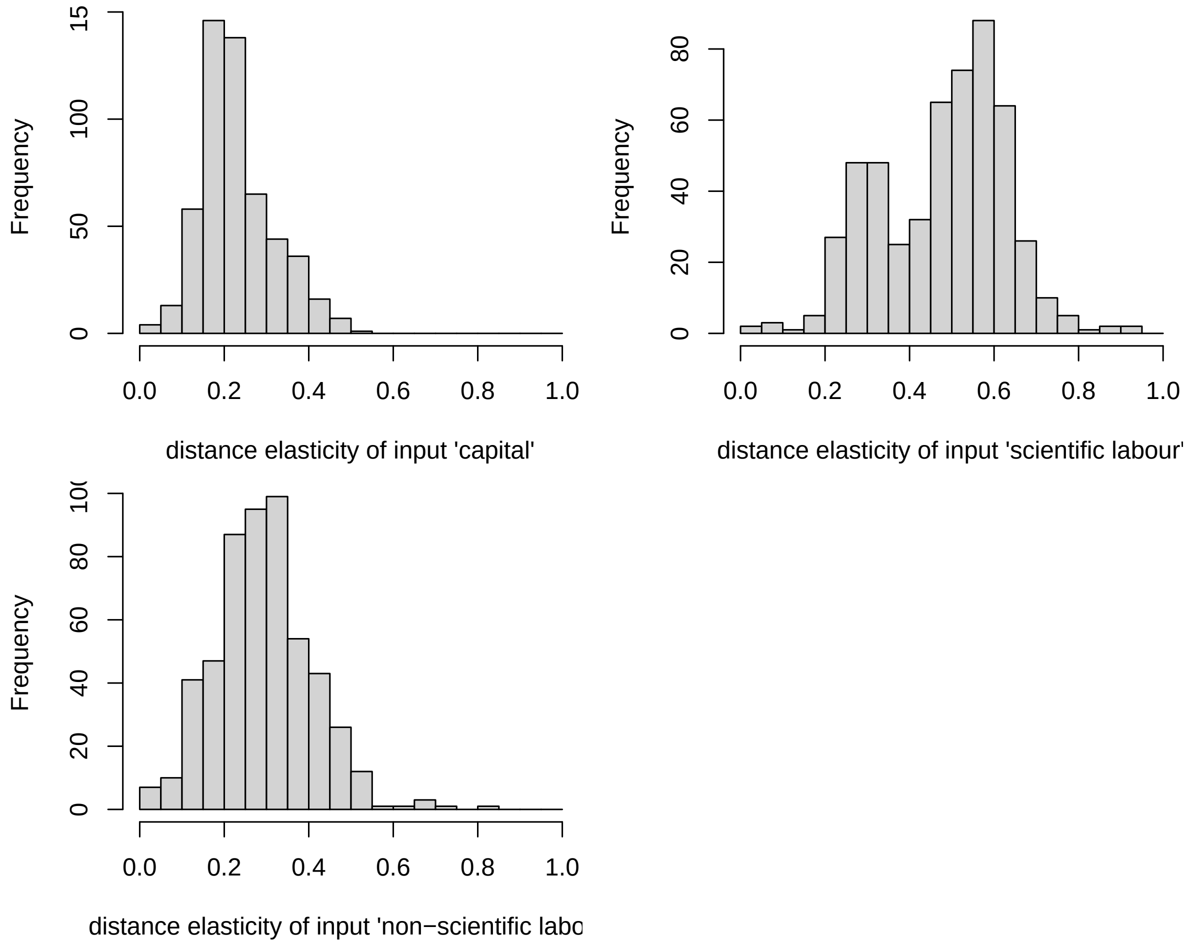

Figure 13: Model-averaged distance elasticities of the inputs

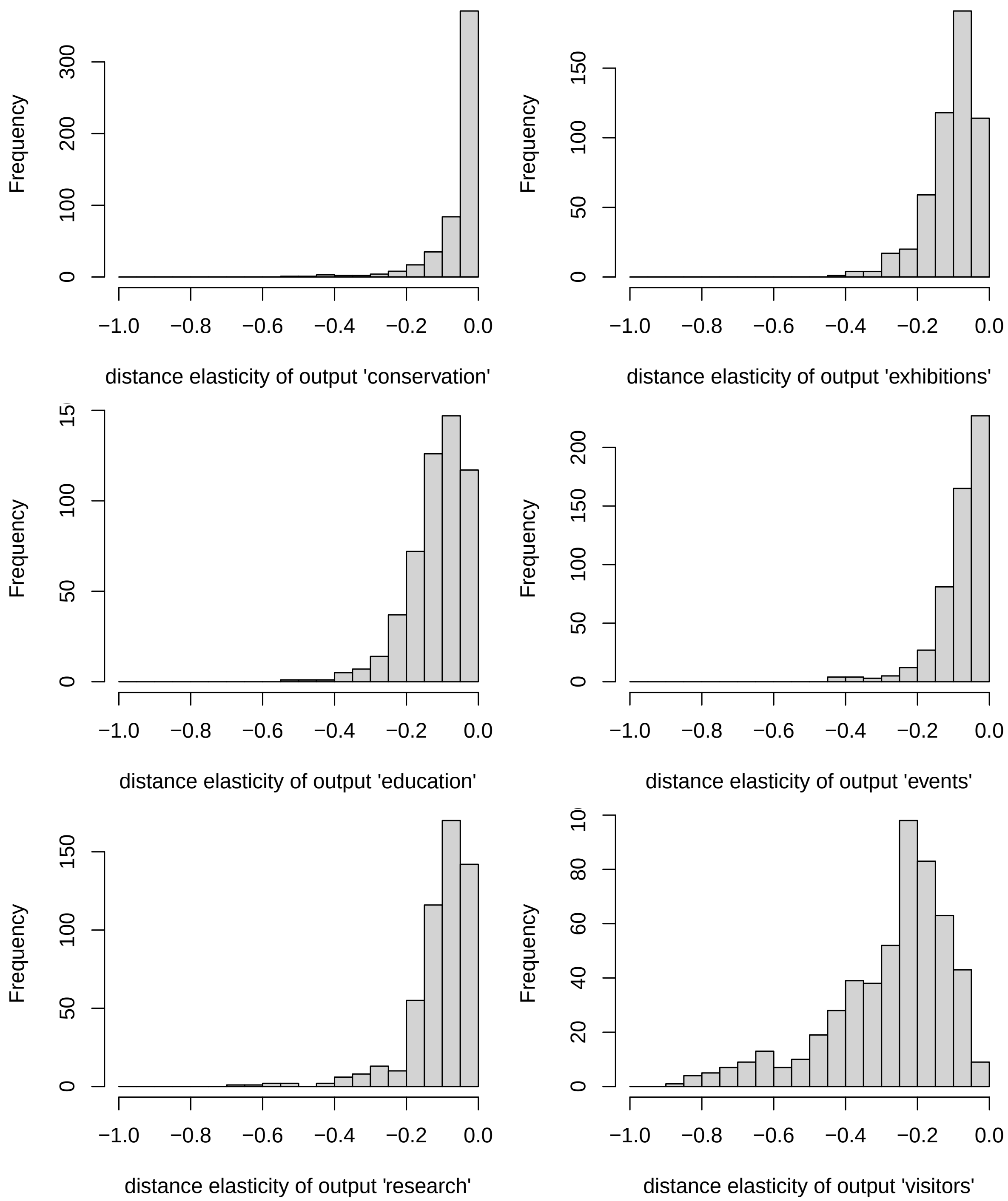

Figure 14: Model-averaged distance elasticities of the outputs

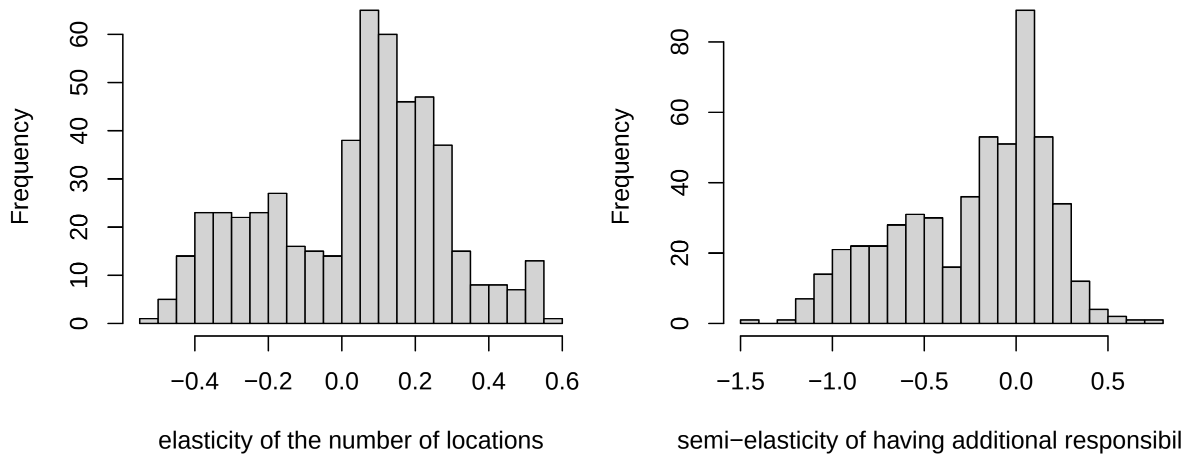

Figure 15: Model-averaged distance elasticities and semi-elasticities of the environmental variables