

Hoffmann, Rasmus; Kröger, Hannes; Geyer, Siegfried

Article — Published Version

Social Causation Versus Health Selection in the Life Course: Does Their Relative Importance Differ by Dimension of SES?

Social Indicators Research

Provided in Cooperation with:

German Institute for Economic Research (DIW Berlin)

Suggested Citation: Hoffmann, Rasmus; Kröger, Hannes; Geyer, Siegfried (2019) : Social Causation Versus Health Selection in the Life Course: Does Their Relative Importance Differ by Dimension of SES?, Social Indicators Research, ISSN 1573-0921, Springer, Heidelberg, Vol. 141, Iss. 3, pp. 1341-1367,
<https://doi.org/10.1007/s11205-018-1871-x>

This Version is available at:

<https://hdl.handle.net/10419/200211>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by/4.0/>

Appendix - Model definition

We describe the system of equations in two steps for the sake of simplicity. However, all equations are estimated simultaneously. First, we describe the measurement models for the latent variables. While the specific approach is of course different, the general modelling approach is very similar to Warren (2009) and follows the idea of structural equation modelling as used by Bollen (1989). Second, we show the structural model with the interrelations among the latent variables. Latent variables are noted in upper case letters, observed indicators in lower case letters. We only present the equations for a generic version of the SES indicators, as they change for the three dimensions of SES. We use p as the index for the phase (childhood (C), adulthood (A), old age (O)) and k to indicate the k th indicator per period. The link function between the latent variables and the observed indicators in the measurement models is a linear model for linear observed indicators and an (ordered) probit link for those observed indicators which are categorical in nature (see summary statistics table). In case of categorical dependent observed indicators, the error terms are of course not part of the equation. Further we define the vector of control variables including all country dummies and age at interview as \mathbf{x} .

The measurement models are defined by:

$$ses_{p_k} = \alpha_{SES_{p_k}} + \beta_{SES_{p_k}} SES_p + e_{ses_{p_k}} \quad (1)$$

$$h_{p_k} = \alpha_{H_{p_k}} + \beta_{H_{p_k}} h_p + e_{h_{p_k}} \quad (2)$$

for continuous indicators in a linear model and for categorical indicators as:

$$P(ses_{p_k} = B) = g(\alpha_{SES_{p_k}, B} + \beta_{SES_{p_k}} SES_p) \quad (3)$$

$$P(h_{p_k} = B) = g(\alpha_{H_{p_k}, B} + \beta_{H_{p_k}} h_p) \quad (4)$$

where $g()$ is the probit function.

$$ses_{p_k} = g(\alpha_{SES_{p_k}} + \beta_{SES_{p_k}} SES_p) + e_{ses_{p_k}} \quad (5)$$

$$ses_{p_k} = g(\alpha_{SES_{p_k}} + \beta_{SES_{p_k}} SES_p) + e_{ses_{p_k}} \quad (6)$$

For identification purposes, in each of the measurement models we assume that $\beta_{H_{p_k}} = 1$ or $\beta_{SES_{p_k}} = 1$ which sets the scaling of the latent variable.

The structural model is defined the following set of equations. We use AR to denote autoregressive coefficients (i.e. health and SES on health and SES one phase earlier). We use HS as the index for the coefficients representing health selection (health predicting SES one phase later) and SC to denote social causation coefficients (SES predicting health one phase later):

$$SES_O = \gamma_{SES_{AR_O}} SES_A + \gamma_{HS_O} H_A + \mathbf{x} \gamma_{x_{SES_O}} + \epsilon_{SES_O} \quad (7)$$

$$SES_A = \gamma_{SES_{AR_A}} SES_C + \gamma_{HS_{A_{direct}}} H_C + \gamma_{edu_{SES}} education + \mathbf{x} \gamma_{x_{SES_A}} + \epsilon_{SES_A} \quad (8)$$

$$H_O = \gamma_{SC_O} SES_A + \gamma_{H_{AR_O}} H_A + \mathbf{x} \gamma_{x_{H_O}} + \epsilon_{H_O} \quad (9)$$

$$H_A = \gamma_{SC_{A_{direct}}} SES_C + \gamma_{H_{AR_A}} H_C + \gamma_{edu_H} education + \mathbf{x} \gamma_{x_{H_A}} + \epsilon_{H_A} \quad (10)$$

$$education = \alpha_{edu} + \gamma_{edu_{SES_C}} SES_C + \gamma_{edu_{H_C}} H_C + \mathbf{x} \gamma_{x_{edu}} + \epsilon_{edu} \quad (11)$$

The indirect effect of childhood SES and health on adulthood health and SES respectively as well as their total effect are calculated as:

$$\gamma_{SC_{A_{indirect}}} = \gamma_{edu_H} * \gamma_{edu_{SES_C}} \quad (12)$$

$$\gamma_{HS_{A_{indirect}}} = \gamma_{edu_{SES}} * \gamma_{edu_{H_C}} \quad (13)$$

$$\gamma_{SC_{A_{total}}} = \gamma_{SC_{A_{direct}}} + \gamma_{SC_{A_{indirect}}} \quad (14)$$

$$\gamma_{HS_{A_{total}}} = \gamma_{HS_{A_{direct}}} + \gamma_{HS_{A_{indirect}}} \quad (15)$$

All error terms in the model are assumed to be uncorrelated and normally distributed:

$$\mathbf{e}, \boldsymbol{\epsilon} \sim N(0, \boldsymbol{\sigma}) \quad (16)$$

$$\boldsymbol{\sigma} = \begin{bmatrix} \sigma_{e_{ses_{C_1}}} & 0 & \cdots & 0 \\ 0 & \sigma_{e_{ses_{C_2}}} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \sigma_{\epsilon_{H_O}} \end{bmatrix} \quad (17)$$

$\boldsymbol{\sigma}$ is a block diagonal matrix.

The exceptions are:

$$cov(\epsilon_{SES_O}, \epsilon_{H_O}) \neq 0 \quad (18)$$

$$cov(\epsilon_{SES_A}, \epsilon_{H_A}) \neq 0 \quad (19)$$

$$cov(SES_C, H_C) \neq 0 \quad (20)$$

The model therefore relies on the assumption of normality and independence of the error terms, especially in the structural model. This is a common assumption in most regression applications.

References

Bollen, K.A. (1989). Structural Equations with Latent Variables. John Wiley & Sons.

Warren, J.R. (2009). Socioeconomic Status and Health across the Life Course: A Test of the Social Causation and Health Selection Hypotheses. *Social Forces* 87(4):2125–2153.