Health & Research Journal

Vol 12, No 1 (2026)

Volume 12 Issue 1 January - March 2026


Health and Research Journal (peer review) ISSN:2459-3192

International Journal of Research and Review in Health Sciences

Volume 12 Issue 1 January – March 2026

Published in cooperation with the Postgraduate Program "Intensive Care Units", the Hellenic Society of Nursing Research and Education and the Helerga

How Social Care Services can be designed using cause-effect models and Bayesian analysis. A study in Scotland

Sotirios Raptis

doi: 10.12681/healthresj.38338

HEALTH AND RESEARCH JOURNAL
E-ISSN:2459-3192

(2026), Volume 12, Issue 1

RESEARCH ARTICLE

HOW SOCIAL CARE SERVICES CAN BE DESIGNED USING CAUSE-EFFECT MODELS AND BAYESIAN ANALYSIS. A STUDY IN SCOTLAND

Sotirios Raptis

Scottish Government, National Health Services Scotland (NHSS), UK/Scotland, Abertay University Dundee

Abstract

Objective: The paper aims to model Health-and-Social-Care (H&Sc) services as Cause–Effect (CE) groups within a Bayesian framework, using cross- and self-causation dynamics. The contribution is that the data used are open and never studied before and are posted online by National Health Services Scotland (NHSS).

Method and Material: The paper contributes to the ongoing discourse on how Machine Learning(ML), and Bayesian inference, can inform the purposes of policymaking. Cause–effect relationships and Bayesian methods can associate public services as causes-effects, through suitable likelihood functions and priors that binomial and normal distributions can imlplement. The study identified the optimal predictive distribution using the Maximum-a-Posteriori (MAP) estimation method. Moreover, the CE-matrix approach, enables the rep-resentation of multiple causes linked to a single effect-target in a tabular format, that facilitates interpretability and prediction.

Results: The findings indicate that services related to 'Alcohol' can be predictors of other effect-services, while home-based services were identified as causes of subsequent hospital admissions. Moreover, low-demand services were observed in earlier years, particularly those with no records after 1997, whereas higher-demand services were newly introduced in later years. These findings may offer in-sights into latent inter-service relationships, and inform policy development. The cross- and self-causation in a Bayesian framework, determined that the posterior can be predicted by 5 to 10 previous observations and this is significantly affected by the level of zero-padding (percentage of past no-records). In later years, the CE models yield more probable demand patterns. Cause–effect relation-ships were identified between smoking-related services, mental-health support, and the epidemiological index of Primary-1-Education children's Body-Mass-Index (BMI).

Conclusions: The conclusions drawn from this analysis may be particularly relevant for insurance providers and public policymakers, who can leverage Bayesian CE-linked service models for long-term care planning, especially for elderly and low-income populations. The validation of ser-vice interlinkages further enhances the potential for precise and efficient resource allocation.

Keywords: Bayesian probability, cause-effect cohorts, self- and cross-causality, data frames, services, prediction.

Corresponding Author: Sotirios Raptis, Scottish Government, National Health Services Scotland (NHSS), UK/Scotland, Abertay University Dundee. Email: sotiris.raptis.n@gmail.com

Cite as: Raptis, S. (2026). How social care services can be designed using cause-effect models and Bayesian analysis. A study in Scotland. Health and Research Journal,12(1),19-35. https://ejournals.epublishing.ekt.gr/index.php/HealthResJ

INTRODUCTION

The digitization of public services—particularly those related to social and health care—and the ability to apply for them online (e.g., subsidized housing or rent support) have facilitated the systematic recording of data pertaining to their delivery, as well as the associated benefits and impacts on service users. This also allows policymakers to capture operational problems such as unmet social needs or unnecessary pubic spending when the cost is high and the benefit is low. The allocation of public funds to ensure better care was observed to increasingly make use of modern analysis (econometrics) methods that depend on artificial intelligence (AI) whose part is machine learning (ML). ML is very roughly a process where raw data and known explanations for them (for example statistical distributions) are presented to a computer as pairs of known data (also called training data) so that the computer can learn from the pairs and then match new unknown data to possible explanations (for example, probabilities) for them that one seeks to find. ML can use concepts from statistics (statistical learning) when the learning process entails computing statistical parameters (for example sample mean or variance) from the known data so that the new explanations for the new data can be deducted using these parameters.

The current paper aims to provide a context for using ML (CE analysis) to well organise public resources. This is proposed to be achieved by grouping services according to the interrelationships of their demand patterns, as modeled through Cause–Effect (CE) frameworks. For example, resource planning for remote healthcare services—such as online consultations, community-based care, or services utilizing Internet of Things (IoT) technologies—can be improved when it is known that the demand for such services can be predicted or associated, as an effect, with underlying causes such as the number of individuals of advanced age (who may require community-based care due to limited mobility) or those residing in rural areas. Importantly, within the Cause–Effect (CE) modeling framework adopted in this study, a causal relationship does not necessarily imply that the 'cause' service chronologically precedes the 'effect' service in terms of need or occurrence.[1] Rather, it indicates that the criteria for establishing a CE relationship are met based on the temporal patterns of demand recorded in year-by-year digital datasets. A more detailed discussion of this modeling approach will be presented later in the paper. There have been concerns about the system of publicly funded social care in England and Scotland for more than 20 years.[2] Added value can be created to face the growing demand by connecting services using CE service models in the Bayesian setting and, in this study, two distinct types of models are employed. The first is based on the Bayesian analysis of self or cross-causation.[3] In these a service can be the Bayesian prediction of itself or of another one. The second is using the linear cross-regression through the CE matrix. It is based on linearly predicting a demand from the same and other demands. The second uses parameters such as other services to use as predictors or as their effects. Also, there can be time (years) over which it is best to link them. The first one finds the number of nonzero demand years that can be caused by a still adjusted number of past demands for the same datastream. The CE matrix model assigns a confidence level to any linkage that determines how well two or more services can be linked over a defined period. A motivation for this work was to assist policymaking and determine the effects of the policies applied on public services. An example can be how living quality factors can be affected after a policy is applied. The process of policymaking can be seen as a data process considering how many people took one or more services before and after some policy was applied. The paper is organized as follows: In the 1st section the nature of the data processed is better explained and the main analysis is given including the CE models and the rationale behind them as well the relevant literature is given for CE. Parts of this are the Bayesian causality models. Next, the data landscape is outlined and their format is explained. Then, the application on the PHS data (Public Health Scotland)[4] is explained. Next, in the same section the data are better explained and the notion of the service pack, (H&Sc), is contrasted with the concept of the single data stream tracked, time-series (TS). It is further explained how this facilitates detailed data linking using CE. The section on ―Research Method‖ briefly presents the CE matrices and the two Bayesian predictions of self and cross causation. It is also discussed how these can work along with CE pairing while in the same section, the CE map and the specifics of pairing a cause with an effect through an intervention is given. Indicative comparisons and results are presented in the same section and are accompanied by comparative co-plots or tabular forms. Emphasis is placed on the integration of Cause–Effect (CE) modeling with binomial and normal distributions, employed respectively as prior and likelihood functions. This framework enables the grouping of service data into cohorts, thereby supporting more effective and systematic organization of public services in a broader policy context.

Related works

The work of Shpitser[5] discusses the role of hidden variables in making sense of observed clinical effects when the results are not clear. An interesting point is made that the CE models can promote evidence-based decision-making.[6,7] That is, link the evidence as measured (encoded) by cause variables to the effects that are the decision on actions to take as interventions or as options. Decision- making in CE studies may concern the best action to take considering causes (observed evidence) as well as ways to gauge the effects if the effects cannot be made more precise. An example is the quality of care or the qualities of life that are multi-parametric effects.[8] In the last referred-to paper, the quality-of-life factors taken can be psychological well-being, behavioral patterns, and social parameters. This is central in estimating causal impacts, and in comprehending unintended consequences. One can expand the CE analysis (either as causes or as effects) in a wider spectrum of co-variants where many causes can have many effects. Then, one can consider as many as possible co- founding factors and then the risk of having adverse effects is limited (as predicted). Cause–Effect (CE) models can be applied in drug safety research to estimate the risks associated with drug intake and to identify potential adverse reactions.[9] More broadly, CE methods are commonly employed in clinical decision-making.[10] In such a clinical setting the challenge of linking many clinical causes (for example, the administration of drugs to many effects like the clinical outcomes, is discussed. The CE models can be applied to policymaking.[11] In that context, the focus lies on the Cause–Effect (CE) relationships between policy actions and their outcomes. An intervention can be conceptualized as a data-driven process, wherein publicly provided services within the Health and Social Care (H&Sc) domain are viewed as forms of social or public interventions. These services have the potential to improve beneficiaries' quality of life, thereby effecting measurable changes in their living conditions. This can be measured, in turn, in a plethora of ways. A case in the EU (European Union) is analyzed in the European Labour Authority[12] with the use of suitable Key Performance Indicators (KPIs). The referred-to work is based on pre-set (H&Sc) policy goals that were to be attained. The CE pairing in CE models can also mean (not always) that a change or not of the demand of the effect (target) services can be sensitive to (influenced by) changes that occurred in the cause-services that preceded them. An example of service pairing can be how the housing market can change when the rents are partially or fully subsidized by public funds. The tenants can be charged less for rent due to an allowance in the rent offered as a social policy. Hence, this social policy can have an impact on the housing market. A discussion regarding the UK housing market is presented in the recent literature.[13] The referenced study presents the full spectrum of effects, encompassing both benefits and disadvantages—each of which is still classified as an effect within the Cause–Effect (CE) framework. An additional example of a CE relationship is the dependence of hospital admissions on alcohol misuse; alcoholism often leads individuals to seek medical assistance, resulting in increased rates of hospital admission. Then, by using CE models, one can predict this number and by that one can also foresee seasonal admission peaks for some reason. The time relationship between the causes and the effects is discussed in recent literature where the volatile nature of CE relations is advocated. Its dependence on the dynamics of the system under investigation—specifically, a chemical system in the referenced study—is thoroughly analyzed. The physical or chemical processes are typical examples of CE models but in this work, these can be inferred using ML methods such as prediction and Bayesian association. The model used in this work is introduced in Kay et al. work.[15] Indeed, we had a year series of demand data from 1981 to 1999. We can refer to different areas where the CE modeling can be applied as a context in making plans. The work of Oki et al.,[16] makes a clear definition of risks conceived in the context of CE analysis that are applied in classical construction engineering.

The effects may be not only beneficial but also adverse. Such negative outcomes can include delays that hinder project advancement or, in some cases, negate previously achieved progress. The CE model is determined by a set of parameters that usually depend on learning data from which the model emerged (derived). Consequently, sensitivity—defined as the degree to which the outcome varies in response to changes in a process parameter—also depends on those parameters. In cases where the Cause–Effect (CE) model is Bayesian in nature, represented generically as P(effect | cause), the parameters influencing sensitivity are inherently part of the Bayesian model itself. For example, the parameters of the Cause–Effect (CE) model can be incorporated into the formulation of the Bayesian prior or the likelihood function.[9] In the referred-to work, the so-called (—Causal Bayesian Network||)s (CBN)s are presented along with the innovation of encoding expert feedback as an added knowledge (Authoritative Medical Ontologies (AMO)s). In the present work, the equivalent of this and the dependence on different causes, as priors, is to consider as many services as possible causes. That is, the Health and Social Care (H&Sc) delivery parameters—namely, the service attributes listed in the table in Figure 1—can be incorporated into CE models. These models can link public resources utilized by the H&Sc system, such as information technology, personnel time, and financial allowances, to improved service delivery and potentially reduced operational costs. As the work in Yang et al.,[17] discusses there is a cost risk in considering more causes to an observed effect as in. The CE analysis can solve, thus, a dual problem. The first objective is to gain a clear understanding of the factors that contribute to high outcome costs. For example, unanticipated or unjustified expenditures—representing adverse effects—could have been predicted and potentially avoided. The second objective is to identify the most effective and cost- efficient solution for achieving a desired positive outcome, which may represent the intended goal of a given policy intervention.

The Bayesian and the CE matrix approach: How do they relate?

One H&Sc service can be linked, using either a Bayesian form or another CE model to one or more other H&Sc services as its causes that can also be its predictors. The approaches adopted here are (1) use the Bayesian models as a paradigm of self-causation, and (2) use the CE matrix as a paradigm of cross-causation. The self-causation refers to predicting the best prior, using MAP, from past samples of the same data stream. The cross-causation refers to using the common years where two or more data streams have all nonzero demands. Then, one can consider one data stream as the posterior that is 'caused' by at least two other data streams that can represent the prior data and the likelihood data. Here, the terms _likelihood_ and _prior_ are only used for facilitation since these relate to two different data streams. The Bayesian setting allows for such a generic approach since the model, $P(A/B) = P(A \wedge B)$, can link at least 3 distributions for the events (A^B is a joint event), (A/B is A conditioned on B). $P(B)$

These can be taken as virtually different services. The Auto-regressive–Moving-Average (ARMA) models are models where one can predict one data stream from itself using past records of it. The number of past records is the —order|| of the model. These models have a similarity (that is, self- prediction) to the Bayesian selfcausation models used here. The AMRA models consist of a self- generating service (self-prediction) and one can define the effect directly from the past samples (for example before 2011) of the same service and not seek the causes in other data streams. The typical formula for the ARMA model is given in equation (Eq.1) and is also the basic one for the CE matrix form where the elements of the matrix are computed using it.

An adverse effect may be part of an unknown relation between a cause-variable and an effect one. In [18] the role of volunteers is discussed in determining the gains (effect variables) from a range of interventions (cause-alternatives) so that policymakers can make informed decisions. These decisions can concern how the healthcare services (and therapies) are delivered or chosen out of large lists of candidates ones. The social care plans examined in this study are analogous to the patterns of service uptake by beneficiaries, reflecting how social care services are accessed and utilized. Maybe one can measure how one service can cause different services (as its effects) so long as the other services can be defined using temporal CE models. In [19] the averaged treatment effect (ATE) is discussed as a way to assess the change of a cause to an effect service or an intervention. ATE implies a changing CE model in time or suggests that the CE model may not apply for the entire treatment period. In this work, we would need to take the average response of the likely effects from year #1 (1981) to the final year #39 (2019). Due to the heavy zero padding (zero padding was for no service record found), a 53.8 % of the 110 services had at least three records from 1981 to 2019 (that is the most frequent year span met in the data). That means that the CE models rather linked services of quite varying duration. The average year span for those services that lasted more than 1 year was 10 years. That also means that the CE models (here Bayesian as well) likely linked 10-year services as effects to 10-year services as their causes.

MATERIALS

The software

The data were loaded for processing using SPARQL (data loading and data filtering). SPARQL supported the implementation of datacubes (data segmentation per attribute/dimension).

The data landscape

The basic data structure can be observed in Table I where the level breakdown of the services is given for indicative cases (3 out of 55). The data utilized in this study comprised publicly available Health and Social Care (H&Sc) datasets published on the Public Health Scotland (PHS) website[20] as of June 2019. These datasets included counts of patients who accessed publicly funded services. 110 different H&Sc service demand data were examined as raw data. Then, they were aligned in time, with zero padding, so that one could have demands for all services and for all years from 1981 to 2019 as on NHSS website.[20] The time-alignment process was achieved with zero padding to recover missing records as records that did not exist in one or more years. A plot of the 11 main service demand sums (sums per service pack) where the summed demands across their attributes (conditions they were offered in) is shown in Fig. 2. The acronyms used per attribute and level is in parentheses and follows the pattern (*X, Y, Z+with X: pack ID, Y: property ID examined, Z: the level/value of the property). The attributes are boldfaced and their levels are separated by commas inside the parentheses. A service pack can have as many TS as each of its components. That is, the separate year series/TS per attribute/level. Fig. 2 shows the plots of the sums of the counts of the demands per H&Sc (pack). It was deemed necessary to consider a large year span of 39 years and not only take non-zero TS. In fact, there are, even in 1981, services of very low demand that cannot be shown in the collective diagrams here. The later recordings of the demand (after 2000) were of 3 or more scales higher. For example, the TS|| S6. Client.Group.of.in.Care.Home (Home sector/voluntary sector)_|| or (HSC with pack ID=6, service ID:79 [out of the 110 in total]) had (176944) hits whereas the TS _S12_ (HSC with pack ID:12, service ID:110) had only (5). It is worth studying CE relationships in a longer span since CE pairs do not have to be adjacent in time. That means that the result of an applied H&Sc policy can manifest itself in the short or in the longer term. It was found that CE can link with confidence most of the health and social care services as effects or as causes and only very few among them were not the effects of other services. It was also true that the health and social care data emerged from a few basic packs (10), that is, from services that were tailored to specific audiences such as (1) low-income people, (2) adults, (3) young people smoking, (4) young people suffering from alcoholism, (5) the elderly, (6) very young mothers, (7) people with mental health issues, (8) people who receive care from distance or at home, (9) newborns with very low weight, (10) people in need of community housing, etc. This naturally groups the social care beneficiaries. In addition, the Cause–Effect (CE) relationships highlighted which specific services or population groups could be associated with others, either as causes or effects. For instance, CE analysis revealed a linkage between female adults of a certain age group and male adults of a similar age residing in subsidized housing. Hence, adults of the same age regardless of gender needed free or subsidized housing and their numbers were very comparable (CE linked). That is, they were captured as population segments needing special health and social care plans for them. One can then think that one such service is the cause and its counterpart (as for age band or gender) is the effect services. Then, one can predict the demand for the effect service when knowing the demand for the cause service. Moreover, one can link services of quite different natures such as services offered to low-income people as causes of services related to subsidized housing as an effect-social care service.

Research Method

The CE matrix model

The CE matrix model entails both causes and effects and is defined for a cause service's time series (TS),$\hat{x}(t)$ = ,$x1(t)$,..., $xN(t)$-T, and for its observed effect's TS$\hat{y}(t)$ = ,$y1(t)$,..., $yN(t)$-T at the year point (t) with the form in equation (Eq.1):

$$\hat{y_i} = \hat{x_i}^T * \hat{\beta} + \epsilon_i, \ni\ i \in \quad [1, N] \text{Eq. (1)}$$

, for $\hat{b} \in R1\times k$and $\epsilon i \approx N(0, \sigma 2)$that are IID (Independent and identically distributed) variables. A CE model typically links an effect service to one or more cause services and assigns a confidence to any such pair. This confidence determines to what extent is the relationship valid.

The CE matrix normally links 110 services (as possible causes) to 110 other services (as possible targets). It is given in equation (Eq.2):

$$M_{CE} = \{P_{i,j}\}_{110,110}, \ni\ P_{i,j} \in \quad [0,1] \text{ Eq. (2)}$$

The exact model for the binary (CE)s is given in equation (Eq.3):

$$CE_{threshold} = CE_{bn} = \begin{matrix} 1, & P_{i,j} \geq thr \\ 0, & P_{i,j} < thr \end{matrix} \quad \text{Eq. (3)}$$

We can use cut values on the output of the CE model (the service CE connections, that is) and remove links with low CE linkage probabilities. This was carried out using various probabilities. Indicative ones are $P1$ = 0.5or $P2$ = 0.9. The remaining links have one ―target‖ (the effect service on the left) and one or more causing services (on the right). For example, it was found that the service with (ID=40, 'Smoking prevalence and deprivation. Self-assessed general health.Good', as an effect- service, links to five cause-services with confidence above (P=0.1). Then, the service (ID=8, 'Smoking prevalence in young people - Self-assessed. Smoking Behaviour .Occasional.smoker') with a threshold (0.1) links to three, and, the service (ID=19, 'Smoking prevalence in. young people.S Self-assessed .Gender.Female') with a threshold (0.1) links to five. Also, the service (ID=34, 'Smoking prevalence in. young people.S Self-assessed .Gender. Male') with a threshold (0.1) links to two others, and, the service (ID=10, 'Smoking prevalence and deprivation. Self-assessed general health. Bad) with a threshold (0.5) links to one. Moreover, the service (ID=21, 'Smoking prevalence in young people - Self-assessed. Smoking Behaviour . Non smoker') with a threshold (0.5) links to two. The names of these services can be found in Table. 1 if we order all attributes and all levels as the table unfolds downwards. There is no pre-determined way to apply a threshold on specific CE connections. The suitable cut values can be a challenging research question also tackled in other areas as in [21]. The last work states that to best link two data sets we need to check the statistical significance of their connection.

That means a high CC still bears a level of significance. This significance can be based on the MCMC method (a repeated random sampling from a wider distribution). Another way is to check if alternative methods can provide more insights can such be the KL test or other tests. These may suggest the connection or not of two data streams.

In this case, we only need to know to which extent this is true. This work is investigative and aims to suggest ways to design or improve offered H&Sc policies using demand data Finally, one could improve (refine) the cut values of the CE connections after tracking (updating), in the long run, the found of the CE links

The rationale is to check, in the longer term, that the validity of such a connection is preserved over time and also face the problem of spurious connections. This assumes that spurious connections cannot hold for long. It also has to be noted that the temporal overlap of two (otherwise stated as) is not necessarily a condition of CE linkage. The linkage of many effects to many causes can be explained by the equations introduced in [22]. Between any two variables (X, Z) one can find other factors (Y)s (one or more) that can link them under the role of intervention variables. This link limits the probability of spurious connections since these cannot always satisfy these equations. The intervention variables can be co-variants that link the CE pairs but may be often unobserved or hidden and may come into play in the background. These can be denoted as (Y)s and can link the causing service X with one of the observed outcomes Z. Cause–Effect (CE) theory, in this context, should be understood as a modeling framework for representing real-world phenomena—such as the outcomes of policies, medical interventions, or other events. These connections are often characterized by a degree of uncertainty or ambiguity regarding the direction and strength of causality. As a result, the interpretation of what constitutes a likely cause for a given effect may not always be straightforward and can vary depending on context or perspective. (2) presents the CE which is a probabilities matrix where each row represents one target service and all its columns contain the attached confidence of each of the 110 services to the the target. The CE confidences found were in the interval, $P_{CE} \in \quad [0.001, 0.9939516]$ and with frequencies that fell exponentially as one moved to higher confidences. There were 1046 different confidences as computed using CE models. The maximum count is 600 which is the most popular count (or the equivalent) among the 1046 found. The maximum number of causes per different effect is 109 (=110-1). This is when the effect is not the same as its cause. However, the way that the 109 services can be combined, in theory, to cause the 110 − th (effect) service $\sum_{i=1}^{i=109} \binom{109}{i}$ where, $\binom{109}{i} = \frac{109!}{(109-i)! \times i!}$. This is so because we can have combinations of any number (n) of cause-services, n ∈ [1, 109] to have an effect-service (the 110-th). Hence, 600 seems a reasonable number of paired combinations for CE confidences above (0.1) which is what the maps in Fig. 3.a illustrate. The maps in Fig. 3.a, and Fig. 3.c relate, using singles paired. That is, a single cause on the Y-axis to a single effect point on the X-axis.s The map follows a color mapping of the confidence levels (as per the color scale shown). The three plots in Fig. 3.a, show the map produced for a minimum cut level, (P > 0.1). Then, in Fig. 3.b the map is for (P > 0.5), and finally in Fig. 3.c the CE map is for (P > 0.9). The exponentially falling frequencies expand on what is shown from Fig. 3.a to Fig. 3.c for the case of 100 cut values, CutV alue ∈ [0, step = 0.01, 0.9]. The final number of valid CE pairs was 4. These are not tautologies (map one to itself) and were produced for high cut values ($P \in$,0.5,0.9)) . This count was first reached at a probability cut value in (0.8272727(0.827) < cutV alue <= 0.9). That means that for those values in (cutValue > 0.5) one can find only 4 CE pairs of public services. These services are: {"S6.Client.Group.of.in.CareHome (older people aged 65 and older)", "*S6.BMIDistribution.in.Primary1.Children.Client.Group.of.in.CareHo (All Adults)*"}, with a (cut Value = 0.8272727(= 0.827)).

The time dependence in the CE matrix model

Vidushi[22] very broadly represents the CE models as observed triplets of services of the form in equation (Eq. 4):

$$\begin{gathered} CE_i(t)=\{X_i(t), Y_i(t), Z_i(t)\} \\ i \in \quad [1,110],\, t \in \quad [t_1, t_2] \\ horizonStart \in [30, t_2] \\ \{X, Y, Z\} \in D \end{gathered} \qquad \text{Eq. (4)}$$

Then, the CE model links the three basic services in the generic relationship: Z ≈ X + Y. The CE model was trained using data from year ($t1$) through year ($t2$) and was tested to link services from year ($t2$) through the end. The duration, (di) is taken here the same for all services and is:di = 39 − $t2$. As discussed, both times can apply within the duration (di) that is the horizon of the specific CE model. Such variables can be responsible for the risk observed as possible adverse effects when the causes occur. That is, one can represent the adverse effects that entail a risk as new effect variables. Also, the variables denoted as (Y)s can be the co-founders. These can be other causes that come into play and can be clinical or social factors. These can come into account directly, that is, by affecting the outcome of the administered intervention (that is, the change of the vaccination parameters). They can also affect that outcome (effect) indirectly, by affecting (or by being affected by) other co-founding factors. CE can infer/investigate that as part of the CE triplet (cause, cofounder(if any), effect). This can help policymakers to make informed decisions and allows effectiveness in estimating the impacts of applying new measures.

The Bayesian method discussed above, $P\ P(effect/cause) = P\,\frac{(effect)^{*}P(cause/effect)}{\sum_{effect}(P(cause/effect)^{*}P(cause))}$, was implemented as well. One can consider as an effect the TS of a target service and as a potential cause all suitable TS from the rest of the services. With self-causation, the case of causing the current demand in year, $t$from past demands of the same service was implemented. The Bayesian probability was computed that some service $i$is consumed by a number of people in year, t ∈ [2, 39] provided that it has previously been taken by other numbers of people as the cause of the current demand. This can be formulated by the generic equation (Eq. 5):

$$\begin{gathered} P(HSC_{i,t})= \\ P(HSC_{i,t}/HSC_{i,t-lag},\ldots) \\ i \in \quad [1,110],\, t \in \quad [2,39],\, lag \in \quad [1,38] \end{gathered} \qquad \text{Eq. (5)}$$

Above, $P(HSCi,t)$is the fact that the service, i, is taken by $(HSCi,t)$

people in year, t. This was modeled as a binomial distribution (BD). This was decided given that the shape of the demands as seen in Fig. 2 suggests an arrival model (nonzero record arrival). The quantity $P(effect)$is the prior and it was approached by a generic Gaussian distribution centered on the average number of people who took the service by the year t, $P(effect) \approx N(\bar{HSC}_{i},t, STDHSCi,t)$and a standard deviation, $STDHSCi,t$that is the deviation of the counts of those people in the same period. This is also the most expected count for the year we are looking for, t. The quantity, $P(cause/effect)$is the likelihood and in this case, it is a BD since we count the hits or no hits per year. The BD is given by the well-known formula in equation (Eq. 6):

$$P(HSC_t \geq 0) = \binom{t}{t_1}, p_1^t \times q^{t-t_1} \qquad \text{Eq. (6)}$$

where(t) ∈ [2, 39] is the time one considers the CE association,$(t1)$ is the number of nonzero years up to time t <= 39 or the time of equivalent 'arrivals'. The BD model can compute the number, $(n)$of nonzero years by year (t). That means that the demand for some service of up to (n)years can be the effect of various numbers of people who took the service before. The nonzero demands can cause not only a current nonzero demand but a range of current demands with variable confidences. Therefore, one can seek the specific current demand that is most likely caused by a specific string of known demands of the past. This is reminiscent of the MAP (Maximum A Posteriori) methods that is actually computed in equation (5). Fig. 3.d shows a self-causation example for service, ―S3. Age.13‖, and for the year (with ID=30) or (1981+30=2011) where the posterior (POST) of the candidate counts of patients, as defined before, drops as the counts increase. In a cross-causing paradigm the posterior, the likelihood (LIK) and the prior (PRIOR) are different H&Sc services. In the cross-causation, we would need to approach these 3 distributions from the observed data that are the co-occurrences. That means that the LIK, the POST, and the PRIOR would be known models that are better precised using the frequency of how often specific demand levels of the considered services occur in the same year. The BD model has two parameters which are the rate of the arrivals, $(p)$and the rate of no arrivals, $(1-p)$. The first can be computed per each of the 110 services from theratio $p = \frac{n_{NZ}}{39}, q = 1-p$, The p_{NZ} is the number of years for which we have a record for some service. The connection of the CE models to the Bayesian setting is more evident in the figures from Fig. 3.e to Fig. 3.g. The best posteriors seem relatively high for a large number of services and a specific expected level of the demand (here is 2) as seen in Fig. 3.g. This points to the agreement across the services that one service can cause a similar demand on one or more other services. This can be cross-causality. When the level of the demand is higher (for example 15 nonzero years or more) then the number of services for which this is expected as a best posterior falls as we can see in Fig. 3.f or in Fig. 3.e. One can also see that public services that launched later (for example after 2000) likely have longer periods of nonzero records and the relevant posterior is significant only for them. This can be observed in Fig. 3f.

The cross-causation used in a Bayesian setting

The cross-causation method uses demand patterns across the 110 services. The idea is that services that have common demand levels over a specific year span can likely be CE pairs. The services that have a longer span can be the causes of the services that start later. The equations (5) and $P(effect/cause)$can apply here as well when the cause service is a longer service than the effect one is. The challenge is to capture common patterns of demand of at least two years from the 39, overall, that are shared by at least two services. Let us assume we seek such a demand pattern over the span of any year $t_1 \in \quad [1,39]$to any year $t_2 > t_1 \in \quad [2,39]$such that$t_2 > t_1$. Let us assume that the demand pattern is $\widehat{D} = \{d_{t_1}, d_{t_2},..., d_{t_{n-1}}, t_2 \leq 39\}$. The table in equation (Eq. 1) summarizes some CE pairs (that is, cause-service to the rest (as effect-services) after the first). One can see that this was possible for several year-spans as for $t_1 = 20(1981+20=2001)$, $t_1 = 30(1981+30=2011)$and for services that belonged to the same H&Sc pack. The CE linkage (and CE confidence) varied across ages or genders or degrees of health. The cause services can be more and begin from the second one in the 1st column in Table 2 and are contained between the brackets. Although the ones shown are not totally different services (belong t to the same H&Sc pack) the table shows that one gender can cause the other gender to take the service or that ages interact in some way when taking the same service.

RESULTS AND DISCUSSION

Major findings

The frequency of the demand can be caused by past samples when the year is earlier than 2019. Using the models in equations (Eq. 5) and (Eq. 6) it was proved that the nonzero demands (hits per year) fall exponentially as the level of the demand increases. Lower demand services and lasting for up to 10 nonzero years were marked as more likely to occur thus more likely caused. For these levels, the MAP was higher than for longer ones (services of a longer span). Also, this trend proved not dependent on the service. For each year we can compute for each expected (caused) level of demand the candidate posteriors for it. Then, MAP chooses the best probability for that value (that maximizes the 'POST'). An expected quadruplet for (n)= 3 nonzero years for the service (ID=1) and considering the years up to ($t = 30$) is given in Fig. 3.d.

One can further adjust these models and set a cut value for the demand only for nonzeros years that cannot be less than the cut value. Then, the models can predict the years that can cause a level of demand above some level. The condition one can use in the binomial model can become more complex and seek more precise patterns of the demand.

One finding was that a cause service can be variably linked to more effect-services. For example, one found the service (or social cohort) "*S1.Age.13*" was dominant and was linked to (1) *S9* (the sum of the demands for that), (2) *S4* (the sum of the demands for that), (3) *"S6.Client.Group.or.in.CareHome"*, and so on. This can also imply the existence of co-founders, that is, other cause-services that cause these services (as their effects). Another finding was that most of the CE relationships link different (H&Sc)s, that is, services that do not belong to the same H&Sc pack. This helps in planning and in designing policies since one can associate, using CE pairs, different parts of the population (that is, as cohorts) and plan the resources needed. Also, applcations of CE analysis can be found in pattern matching and information discovery (data mining) [23].

Another finding was that the CE matrix linked the social cohort "*Intensive Home Care. Gender. Male*" to (1) "*S1.Age.13*" (P = 0.168), to (2) "*S1.SmokingBehavior.OccasionalSmoker*" (P = 0.157), to (3) to "*S1.Age.13-15*" (P = 0.157), to "*S1.OccasionalSmoker*" (P = 0.157), to (4) "*S1.Age.13*" (P = 0.157). Again, all these (P)s are above $P = 0.1$. One can see that smoking habits can also relate to males who are in need of intensive home care. This cause can also be related to the elderly.

More CE pairs were found such as "*S6. Client Group Of In Care Home*" (that is, Primary 1 children living in care homes) that as an effect can be linked to (1) "*S1.Age.13*" (P = 0.541), to (2) "*S1.Age.All*" (P = 0.166), to (3) "*S1.Gender.Female*" (P = 0.164), to (4) "*S1 . Smoking Behaviour . NonSmoker*" (P = 0.161), to (5) "*S1 . Smoking Behaviour. Regular Smoker*" (P = 0.102). As expected, children at the age of 13 or 15 as well as young people of any age or gender who smoke are related, through CE pairs. Also, it was found that children raised in care homes by young people with smoking habits are involved with gender as "causes" for them to live in care homes. The age seems more relevant (has higher confidence) to that than smoking habits.

Other CE pairs suggest that "*S8 . RegisteredPatients*" (patients registered with GPs) as an effect can be linked to (1) "*S1.Age.13*" (P = 0.108), to (2) "*S1.Age.15*" (P = 0.109). This suggests that young people at the age of 13 or 15 tend to register with GPs with an average confidence.

Furthermore, an interesting connection is that people who declared as being in the pack "*S9*" ("*Smoking prevalence in young people SALSUS . Smoking Behaviour . Regular smoker*") can be the "result" of "*S1.Age.13*" with (P = 0.5). This suggests that young people aged 13 tend not to have mental problems that is expected since mental problems begin usually after some age.

Another strong CE connection found was that "S4" ("*Smoking bahvior in young people.SALSUS*) is linked as an effect to: (1) "*S1.Age.13*" (P=0.574) as its cause. This link is stronger and as such it only pairs two services. The justification is roughly the same as the previous ones that link BMI as an index to young people (young people are linked to their BMI index). Moreover, "*S6. Client Group of in Care Home*" is also linked to the same category of services (cohorts) as to "*S1.Age.15*" (P = 0.560) and to "*S1.Age.13*" (P = 0.531) with high chances. The services that were reported as more likely effects of other ones considering likely CE pairings (one effect, more causes) were those for (P > 0.1) and are: (1) "*S6. Care Home Sector . Private Sector*"(11 causes), (2) "*S5.BMI Epidemiological . Weight Category . Clinical Obese...Severely Obese*" (10 causes), (3) "*S11.Age.15*" (8 causes), (4) "*S11.Age.All*" (8 causes), (5) "*S11. Gender . All*" (8 causes). One can see that services related to young people (children) directly (newborns weight), or, indirectly, such as those who live in care homes where there are children and they are parents of them have been noted to them as causes. The criterion was not the level of the CE confidence attached to them (still P > 0.1) as CE pairs but the Bayesian popularity (count) of them of being a Bayesian cause to them in the data. The results from the Bayesian analysis are presented in figures from Fig. 3.e to Fig. 3.g. These figures convey the interplay of the best posteriors for various no-zero demand levels and various year spans. The posterior is the best confidence we have in having future demands caused by past demands.

CONCLUSIONS, CONTRIBUTIONS, AND FUTURE DIREETIONS

This paper explores how Health and Social Care (H&Sc) service demands can be analyzed using the Cause–Effect (CE) matrix model in conjunction with Bayesian inference. These methods enable the identification of services that function as causal or effect variables within the broader system. A key contribution of this study lies in its application to openly available data published by National Health Services Scotland (NHSS)—a dataset that, to the best of the authors' knowledge, has not been previously examined in this contextAnother contribution of this study is its relevance to the broader discourse on the application of Machine Learning (ML) and statistical inference in policymaking. Specifically, it addresses the underexplored area of linking public services—viewed as effects—to other services, whether within the same or different population groups, as their potential causes. This dimension of causal inference has received limited attention in the context of policy-driven service planning. Moreover, the paper presents and discusses a conceptual rationale for linking the risks associated with new policy implementation to the notion of spurious connections. It explores how ARMA-based prediction methods can be integrated with the Cause–Effect (CE) matrix framework, and how, within a Bayesian context, future service demands can be inferred from past demands or correlated with other concurrent demands. As emphasized, such inferred connections may be spurious, introducing an element of risk. This risk stems from the inherent uncertainty in determining which relationships are valid and reliable enough to justify policy decisions and the allocation of public resources. The paper highlighted and discussed this of risk using a probabilistic context. CE does not have time limits (co-occurrences that need time overlapping) for binding causes to effects and data do not need to overlap in time. The causes can have time-lagged effects or can they be linked to latent, or non-observable data that may link to both the causes and their effects. The CE model was trained using a 6- to 10-year horizon (from roughly 1981 to 2010) that was a common span to most services from 2011 to 2015. The Bayesian model considered the fully spanned services of 39 years and accepted zeros as real values when no service records were found. The CE models are not widely used in public services planning and it is an under-represented area of research. The paper attempted to show the wide variety of CE models. Another innovation introduced is that public services were analyzed as medical intervention data (causes, interventions, effects). Also the time-lag was long (39 years).

The paper suggested the use of CE models to link service demand data over the span of 39 years. As a methodology, it resembles linear regression (LR) methods where the goal is to predict the demand of a single target service from the demands of other services (called its *predictors*) and then group them all (target and its predictors) into a single class of services so that again social care resources can be collectively allocated. CE and LR methods are both ML methods. A further proposal can be to only match cause-services to effect-service services that present a strong year overlap. That is, pick-up as a candidate cause services the services that are offered in the same year span as the effect services, and not relax this condition as was the case in this work. Another proposal would be to pre-classify demand data, that is, create prior data groupings before CE models are applied and apply the CE models to each of these initial groups. This offers the advantage that the data per such group will be more homogenous (better cross-correlated). On the other hand, it may result in very little data (service demands) per such initial group. Then, almost all services within a group will be the cause of the rest of the services in the same group. This would offer no practical insights to policymakers. Another more practical approach would be to "force" groupings, that is, use no algorithms but use only clinical (or social care experience from practice) to create "prior" groupings in the data and then find either causes or effects for these groups algorithmically. In general, the validation of a method relies primarily on practical application, as a definitive 'ground truth' is typically absent—except in cases involving training datasets where true data pairings are known. Results may vary depending on the methodological approach employed, and even within the same model, outcomes can differ based on parameter selection. Machine Learning (ML) techniques are particularly sensitive to the specifics of each case. This context-dependency represents both a limitation—due to the difficulty of formal verification—and a significant strength, as ML methods can yield highly accurate and nuanced insights that may surpass the capabilities of conventional analytical techniques.

Acknowledgements

The author gratefully acknowledges Public Health Scotland (PHS) and National Health Services Scotland (NHSS) for providing access to the datasets used in this study. Financial support for part of this work was provided by Abertay University's 'O924: Integrated Modelling of Health and Social Care Pathways' fund, the 'R861: Systems Approaches to Cancer (2016–2019)' initiative, and, in part, by the Scottish Government.

Funding

This research did not receive specific external funding. However, during part of the work, the author was supported by a research stipend from Abertay University, Dundee, and received additional funding from the Scottish Government.

Notes on contributor(s)

The author is the only contributor of the paper. Data providers are referred-to in the reported literature. The author would like to thank PHS and NHSS for suggesting these open data are suitable for health and social care analysis and Abertay University, Dundee for hosting early parts of this work

REFERENCES

1. Castro, M., Mendes Junior, P.R., Soriano-Vargas, A. et al. Time series causal relationships discovery through feature importance and ensemble models. Sci Rep 13, 11402 (2023). DOI : Available at: https : //doi.org/10.1038/s41598 − 023 − 37929 − w
2. Simon Bottery, et al. A fork in the road: next steps for social care funding reform . The King's Fund . Available at: https : //www.kingsfund.org.uk/publications/fork − road − social − care − funding − reform
3. Koopmans E, Schiller DC. Understanding Causation in Healthcare: An Introduction to Critical Realism. Qual Health Res. 2022 Jul;32(8-9):1207-1214. DOI: 10.1177/10497323221105737. Epub 2022 Jun 1. PMID: 35649292; PMCID: PMC9350449.Available at: https : //www.ncbi.nlm.nih.gov/pmc/articles/PMC9350449/
4. Public health Scotland. Data and intelligence (2020). A − Z Subject Index; 2020. Available at: https : //www.isdscotland.org/A − to − Z − index/index.asp.
5. Shpitser I. Identification in Causal Models With Hidden Variables. J Soc Fr Statistique (2009). 2020 Jul;161(1):91-119. Epub 2020 Jun30. PMID: 33240555; PMCID: PMC7685307. https : //www.ncbi.nlm.nih.gov/pmc/articles/PMC7685307/
6. Olier, I., Zhan, Y., Liang, X. et al., Causal inference and observational data, BMC Med Res Methodol 23, 227 (2023) . Available at: https : //bmcmedresmethodol.biomedcentral.com/ articles/10.1186/s12874 − 023 − 02058 − 5 DOI: https://doi.org/10.1186/s12874-023-02058-5
7. Albert Maydeu-Olivares, Dexin Shi, Amanda J. Fairchild, Estimating causal effects in linear regression models with observational data: The instrumental variables regression model . Psychological methods 1 April 2020 Psychology, Economics . Available at: https : //www.semanticscholar.org/paper/Estimating − causal − effects − in − linear − regression − with − Maydeu − Olivares − Shi/b91a8c4401b93f6db1dd52a1e33ad88522fc5747 . DOI:10.1037/met0000226Corpus ID: 195879758

8. Ganopoulou (2023). Maria Ganopoulou, Dimitrios Koparanis , et al. Causal Structure assessment in Health-Related Quality of Life questionnaires||, Conference: 35th Panhellenic and 1st International Statistics Conference, Athens, Greece, Available at: https : //www.researchgate.net/publication/371169040 . DOI:10.13140/RG.2.2.32272.58881 Salman Afsar,
9. Hengyi Hu,Larry Kerschberg, Improving Causal Bayesian Networks Using Expertise in Authoritative Medical Ontologies, ACM Transactions on Computing for Healthcare, Volume 4, Issue 4, Article No.: 20pp 1–32, Available at: https : //dl.acm.org/doi/10.1145/3604561 , DOI:https://doi.org/10.1145/3604561
10. Joachim P. Sturmberg, James A. Marcum, From cause and effect to causes and effects, 3 March 2017 . Available at: https : //onlinelibrary.wiley.com/doi/full/10.1111/jep.13814, DOI:https://doi.org/10.1139/er-2016- 0109
11. Stefano Cucurachi and Sangwon Suh, Cause-effect analysis for sustainable development policy, Environmental Reviews . Available at: https : //cdnsciencepub.com/doi/abs/10.1139/er – 2016 – 0109
12. European Labour Authority Measuring the effectiveness of policy approaches and performance of enforcement authorities , November 2022, Available at: https : //www.ela.europa.eu/sites/def ault/f iles/2023 – 02/Output – paper – from – plenary – thematic – discussion – measuring – the – effectiveness – of – policy – approaches – and – performance – of – enforcement – authorities – %282022%29.pd
13. BPF. The Impact of Rent Control on the Private Rented Sector ||, Available at: https : //bpf.org.uk/media/6296/2023 - 03 – the – impact – of – rent – control – on – the – private – rented – sector – bpf – f inal.pdf4 . May 2023
14. Wenkai Hu, Jiandong Wang, Fan Yang, Banglei Han, Zhen Wang, Analysis of time-varying cause-effect relations based on qualitative trends and change amplitudes, Computers & Chemical Engineering, Volume 162,2022, 107813,ISSN 0098-1354, Available at: https : //www.sciencedirect.com/science/article/abs/pii/S00981354220015DOI: https://doi.org/10.1016/j.compchemeng.2022.107813
15. KAY H. BRODERSEN, FABIAN GALLUSSER, JIM KOEHLER, NICOLAS REMY AND STEVEN L. SCOTT, ||INFERRING CAUSAL IMPACT USING BAYESIAN STRUCTURAL TIME-SERIES MODELS|| ,The Annals of Applied Statistics 2015, Vol. 9, No. 1, 247–274. Available at: https : //static.googleusercontent.com/media/research.google.com/en//pubs/archive/41854.pdf . DOI: 10.1214/14- AOAS788.2015
16. Oki Oktaviani , Budi Susetyo, Bambang Purwoko Kusumo Bintoro, Risk Management Model using Cause and Effect Analysis in Industrial Building Project ||, International Journal of Research and Review, Vol.8; Issue: 8; August 2021 . Available at: https : //www.ijrrjournal.com/ IJRRV ol.8I ssue.8Aug2021/IJRR032.pdf, DOI: https://doi.org/10.52403/ijrr.20210832 . www.ijrrjournal.com . ISSN: 2349-9788; P-ISSN: 2454-2237
17. Yang C, Delcher C, Shenkman E et al. Expenditure variations analysis using residuals for identifying high health care utilizers in a state Medicaid program. BMC Med Inform Decis Mak. 2019. Available at: https :; 19(1) : 131
18. Atul Gupta, Joseph R. Martinez, Amol S. Navathe, SELECTION AND CAUSAL EFFECTS IN VOLUNTARY PROGRAMS: BUNDLED PAYMENTS IN MEDICARE, Working Paper 31256, NBER WORKING PAPER SERIES. Available at: nber.org/system/f iles/workingpapers/w31256/w31256.pdf
19. Chen Chen, Bin Huang, Michal Kouril, et al. An application programming interface implementing Bayesian approaches for evaluating effect of time-varying treatment with R and Python,Front. Comput. Sci., 16 August 2023, Volume 5 - 2023. Sec. Software, Available at: https : //www.frontiersin.org/articles/10.3389/ fcomp.2023.1183380/full . DOI: https://doi.org/10.3389/fcomp.2023.1183380
20. Scottish Governement, statistics.gov.scot . Available at: https : //statistics.gov.scot/datahome
21. Aristotelis Koskinas, Eleni Zaharopoulou, George Pouliasis , Ilias Deligiannis, 'Estimating the Statistical Significance of Cross–Correlations between Hydroclimatic Processes in the

Presence of Long–Range Dependence'. Available at : https : //www.itia.ntua.gr/el/getf ile/2234/1/documents/earth – 03 − 00059 − v3.pdf ,

22. Vidushi Adlakha and Eric Kuo Statistical causal inference methods for observational research in PER: a primer . Available at: https : //arxiv.org/pdf /2305.14558.pdf
23. Md Saiful Islam, Md Mahmudul Hasan (2018) A Systematic Review on Healthcare Analytics: Application and Theoretical Perspective of Data Mining, Healthcare', 2018, pp. 6-54. Available at: 10.3390/healthcare6020054

HEALTH AND RESEARCH JOURNAL
E-ISSN:2459-3192
(2026), Volume 12, Issue 1

ANNEX

TABLE 1. Services names, years, attributes and most levels break-down.

(ID)s	**(HSc)s full names (packs names)**	**Attributes(As), and levels of them(Ls) (A(Value1:L1,…,ValueN :LN))**	**Years**
1	Self-assessed young people smoking (S1)	**Age**:A1(13:L1,15:L2,All:L3), **Gender**:A2(M:L1,F:L2) **SIMD (***Scottish Index of Multiple Deprivation***) quintiles**:A3 (Most deprived:L1,…,Least deprived:L5) **Smoking behaviour**:A4 (Non.smoker:L1, Occasional smoker:L2, Regular smoker:L3) **Self assessed general health**:A5 (Bad:L1,Fair:L2, Good:L3,Very bad:L4, Very good:L5) **SIMD quintiles**:A6 (Most deprived:L1 ,…, Least deprived:L5)	98-10
2	Smoking and deprivation (S2)	**Age**:A1(13:L1, 15:L2, All:L3) **Gender**:A2 (M:L1, F:L2) **Smoking Behaviour**:A3 (Non.smoker-L1, Occasional smoker-L2, Regular smoker-L3) **Self assessed general health**:A4 (Bad:L1, Fair:L2, Good:L3,Very bad-L4, Very good-L5)	98-10
3	Epidemiological BMI in Primary 1 Children (S3)	**Gender**:A1(M:L1,F:2), **Weight Category Epidemiological**:A2 (Healthy Weight:L1, Obese:L2, Overweight:L3, Overweight...Obese:L4, Underweight:L5) **Age**:A1(13:L1,15:L2,All:L3)	0 5 – 0 9
4	Smoking prevalence and deprivation (S4)	**Gender:A2(M:L1,F:L2)** **Weight Category – Epidemiological (Healthy Weight:L1,Obese:L2, Overweight:L3, Overweight-Obese:L4,Healthy.Weight:L5, Severely Obese:L6)**	98-10
5		**Gender:A2(M:L1,F:L2,ALL:L3)**	10-16
6	Primary 1 Children - Body Mass Index – Clinical		12-16
	Primary 1 BMI Distribution – Main Client Group in Care Home	**Adults:A1(**Adults **with learning disabilities:L1, Adults with Mental Health Problems:L2, Adults with Physical Disabilities:L3,All.Adults:L4, Older People Aged 65 and Older: L5, Other Groups:L6)** **Care Home Sector:A2**(All.Sectors:L1,Local Authority and NHS Sector:L2, Private Sector:L3,Voluntary.Sector:L4)	04-16
7			1981-2019
	Number OF General Practices Registered Patients	**Type Of Tenure:A3**(All:L1, Owned Mortgage/Loan:L2, Owned Outright:L3,Rented: L4) **Household Type:A4(Adults:L1,All:L2,Pensioners:L3, With Children:L4)** **Age:A5(16-34:L1,16-64:L2,35.64:L3,65.And Over:L4, All:L5)** **Gender:A6(M:L1,F:L2,ALL:L3)** **Limiting Long term Physical or Mental Health Condition:A7(All:L1,…)**	
8			04-16
9		**Birth Weight(…)**	10-16
10			10-16
		…	12-16

11	Mental Wellbeing SSCQ(*Scottish Surveys Core Questions*) Low Birthweight home Intensive Home Care Home.Care.Services	Age:A1(...), Gender:A2(...), Ever drank: A3(...), ...	

TABLE 2. Services sets where the first is the effect and the rest are the causes using the method of common demand patterns in years from t_1to t_2.

Services combinations (first is the effect one)	**Year span** t_1, t_2
Primary 1 BMI Distribution . Main Client Group in Care Home . Adults with Learning Disabilities . [All ages]	20, 30
Smoking prevalence in young people . SALSUS - Self.assessed.general health. Good[All genders]	30, 39
Smoking prevalence and deprivation. SALSUS - SIMD.quintiles [deprivations from level 1[most deprived] to 4[least deprived]]	25,30
Smoking prevalence and deprivation SALSUS Self.assessed. general.health. [Bad,...,Very good]	30,39

FIGURE 1. Logarithms of summed counts (all times series tracked per service) for the 11 packs of H&Sc services over the span of 39 years

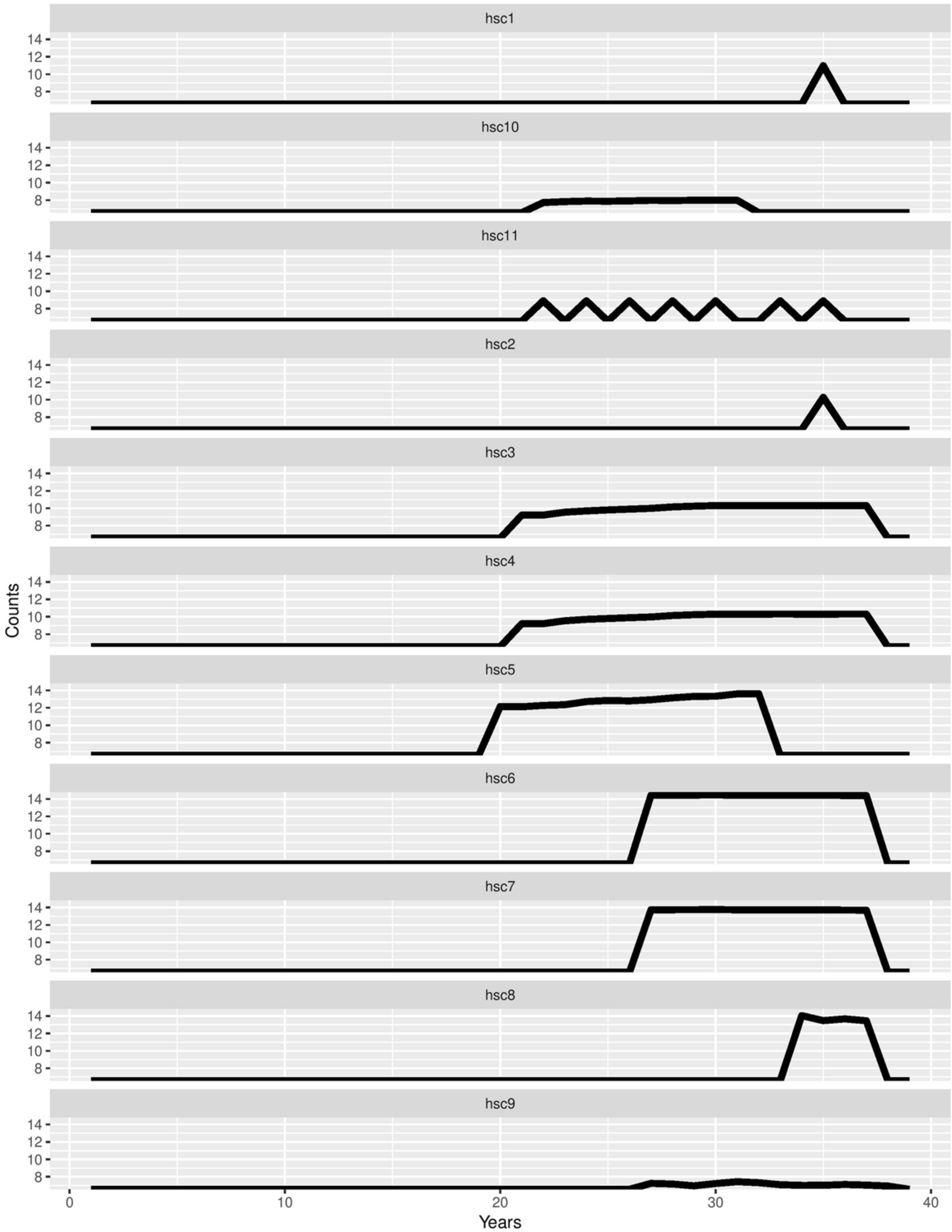

FIGURE 2. CE pairs with (3a) threshold=0.1, (3b) threshold=0.5, (3c) no threshold (> 0), (3e) How different services support the occurrence of 20 non zero demands('arrivals') in years from 20 (2001) to 30 (2011), (3f) same for 15 no zero years from year 15 (1996) to year 30 (2011), (3g) same for 2 arrivals in years from 5 (1986) to year 30 (2011). As we change the thresholds for the accepted confidence in CE relationships the number of valid CE pairs falls exponentially (points kept are the light-colored ones).