



# The use of Bayesian networks for realist evaluation of complex interventions: evidence for prevention of human trafficking

Ligia Kiss<sup>1</sup> · David Fotheringham<sup>2</sup> · Joelle Mak<sup>3</sup> · Alys McAlpine<sup>3</sup> · Cathy Zimmerman<sup>3</sup>

Received: 14 June 2019 / Accepted: 8 February 2020  
© The Author(s) 2020

## Abstract

Complex systems and realist evaluation offer promising approaches for evaluating social interventions. These approaches take into account the complex interplay among factors to produce outcomes, instead of attempting to isolate single causes of observed effects. This paper explores the use of Bayesian networks (BNs) in realist evaluation of interventions to prevent complex social problems. It draws on the example of the theory-based evaluation of the Work in Freedom Programme (WIF), a large UK-funded anti-trafficking intervention by the International Labour Organisation in South Asia. We used BN to explore causal pathways to human trafficking using data from 519 Nepalese returnee migrants. The findings suggest that risks of trafficking are mostly determined by migrants' destination country, how they are recruited and in which sector they work. These findings challenge widely held assumptions about individual-level vulnerability and emphasize that future investments will benefit from approaches that recognise the complexity of an intervention's causal mechanisms in social contexts. BNs are a useful approach for the conceptualisation, design and evaluation of complex social interventions.

**Keywords** Complex systems · Realist evaluation · Bayesian network · Human trafficking · Forced labour · Nepal

---

Ligia Kiss and David Fotheringham have contributed equally to this work.

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s42001-020-00067-8>) contains supplementary material, which is available to authorized users.

---

✉ Ligia Kiss  
l.kiss@ucl.ac.uk

<sup>1</sup> Institute of Global Health, University College London, 30 Guilford Street, London, UK

<sup>2</sup> London, UK

<sup>3</sup> Gender Violence and Health Center, London School of Hygiene and Tropical Medicine, London, UK

## Introduction

### Realist evaluation of complex social systems

Failing to appreciate complex causal relationships in social interventions has often led to misinterpretations, over simplification or even harm [1–4]. New conceptual and methodological approaches offer promising alternatives for the evaluation of complex social interventions [5–7].

Over the past two decades, calls for research that respond to the challenges of explaining and evaluating complex systems have abounded both in Public Health and the Social Sciences [6–12]. Traditional approaches centered in a linear understanding of cause and effect have been deemed insufficient for the identification, implementation, and evaluation of effective responses to public health and social problems [2, 4]. Researchers called attention to the limitation of classical experimental designs, highlighting their inaccurate results and limited explanatory power [1, 2, 12].

At the same time, theories on complex adaptive systems have been increasingly used to understand social phenomena, and a trend in theory-based and realist evaluation has been gaining momentum [13, 14]. Realist evaluations focus on programme theories to examine the validity of assumptions and ideas underlying how, why and under which circumstances complex social interventions work [15].

The recent sharp increase in accessible computational power has amplified the range of methodological options for researchers to conduct analysis on complex systems [6, 7, 16]. These approaches are being used to examine how causes work together to produce outcomes instead of focusing solely on specific causal effects of single risk factors [2]. Agent-based models, microsimulation, dynamic systems, data mining and Bayesian Networks (BNs) are examples of these applications [6, 7, 11, 17].

BNs have been praised by the American Association for Evaluation for their ability to represent complex causal relationships, to use data from different sources, handle subjective and objective data, predict the effects of changes in intervention variables and incorporate programme variables across diverse contexts [18]. BNs hold the promise of disclosing potential causal links among nodes (variables), even in cross-sectional data sets [19], which make them a very useful approach in the conceptualisation, design and evaluation of complex interventions. They offer additional resources to structural equation models (SEMs)—which are a statistical technique that also address the issue of multiple influences on outcomes. However, SEMs do not have the predictive and diagnostic capabilities of BNs or the capability of accommodating smaller datasets and missing data [20, 21]. Because of these additional features, BNs have been identified as a powerful tool for reasoning under uncertainty arising from: gaps in knowledge about systems, imperfect understanding of systems, randomness in the mechanisms driving systems' behaviour or any combination of these factors [22].

This paper explores the use of BNs in realist evaluation of interventions to prevent complex social problems. It draws on the example of a theory-based

evaluation of the Work in Freedom Programme (WIF), a large UK-funded anti-trafficking intervention by the International Labour Organisation in South Asia. It aims to show how BNs can help focus attention on key aspects of a complex system thus allowing analysis to concentrate on the important relationships in a large multivariate data set. Moreover, the paper uses BNs to predict the effect of potential interventions—an overriding goal of this type of study.

### **Modern-slavery, trafficking and forced labour prevention**

Investments to prevent human trafficking, modern-slavery and forced labor have grown substantially over the past decade [23] (see SM1 for glossary of key terms). More than 4 billion USD have been spent in development aid to stop modern-slavery between 2000 and 2013, and 193 countries have pledged their commitment to end the problem by 2030 [24]. The range of organizations involved in responses to trafficking has also diversified considerably [25]. However, to date, these investments have been made around the world based on little or no evidence of their effectiveness [23, 26].

For instance, in 2017, the United Nations’ “Call to Action to End Forced Labour, Modern Slavery and Human Trafficking” lists awareness-raising to prevent exploitation as a key national-level strategy. Eighty-four states have endorsed this document [27]. Yet, to date, there remains virtually no evidence on the effectiveness of these interventions. The assumption underlying awareness-raising, knowledge-building or similar ‘safer-migration’ learning activities is that if individuals are informed of the risks associated with migration or are made aware of their rights, they will be less liable to fall prey to trafficking [28]. These initiatives are often targeted at women to promote gender empowerment and encourage female participants to assert their rights as women and as migrants [29, 30]. However, only recently has there been an interest in learning what makes migration safe in different contexts [26]. For example, recent research in Bangladesh suggests that empowerment interventions may be ineffective in preventing trafficking, and may even inadvertently cause harm to participants [31].

Gender has also been at the core of some state actions to prevent trafficking. Certain governments have imposed nationwide restrictions in the form of bans on female migration to prevent trafficking in women, which aimed to stop women from migrating to particular destinations (e.g., Gulf States) or work in specific sectors (e.g., domestic work) [32]. Governments were strongly criticized by advocates for allegedly increasing undocumented migration and thus placing women at higher risk of exploitation and abuse [32–35]. Yet, in reality, there is similarly limited evidence on how bans affect undocumented or illegal migration and the risk of trafficking [25].

In addition, international organizations and donor agencies are investing in livelihood strategies to promote migration by choice (versus by compulsion) as a means to prevent human trafficking [34]. These interventions seek to promote alternative livelihood options, provide micro-finance schemes, cash transfers, and career development opportunities to prevent distress migration, i.e., motivated by acute

economic necessity [29, 35]. Once again, these investments persist despite the scarcity of evidence showing the preventative capacity of livelihood schemes in relation to migration or human trafficking [23, 36].

Over the past 5 years, fair recruitment has been at the core of the International Labour Organization's (ILO) strategy to fight forced labour among migrants. In their 'tripartite' formula, the ILO has been working with employers, labour organisations and governments to promote fair recruitment laws and practices [37]. These mechanisms aim to tackle abuses by labour intermediaries through licensing schemes, regulation of recruitment networks and verification of the legitimacy of overseas recruitment agencies [38, 39]. These initiatives are in relatively early phases and thus the evidence is still scarce. However, in addressing recruitment, there has been inexcusably little examination of the large and important informal social and migrant networks that operate as local intermediaries and which are widely used by economic migrants around the world [31, 40].

Because concerns over 'human trafficking' emerged relatively rapidly over the past twenty years, many investments were made in the absence of intervention-focused evidence [41]. However, it is now clear that 'modern slavery', human trafficking and decent work are a long-term priority on the global agenda. For example, the Sustainable Development Goal 8.7 [42] highlights the international policy importance of anti-trafficking initiatives and the promotion of decent work. The time is now ripe for the second generation of studies aiming to develop evidence-informed intervention and evaluation designs [26, 43].

### **Data on causal pathways to human trafficking, modern-slavery and forced labour can help close the evidence gap**

Conducting rigorous research on human trafficking, modern-slavery and forced labour is challenging because of the illegal, hidden and stigmatised nature of this extreme form of exploitation. Sampling frames for these populations are virtually nonexistent and previous data have come mostly from surveying survivors in post-trafficking assistance services. However, the ratio of survivors who are assisted versus the overall trafficked population is unknown [44]. In addition, these data do not permit a comparative analysis of the causes of forced labour, because all participants in the study will have necessarily experienced these abuses. To our knowledge, this paper provides some of the first evidence on forced labour risk factors drawn from a general returnee migrant population. The key advantage of this sampling approach is that it contains a mixture of migration outcomes—that is, not all migrants were identified as forced labour survivors. Thus, we were able to examine the risk versus protective factors for exploitation among a general returnee migrant population.

### **Research context: the Work in Freedom Programme**

The Work in Freedom Programme (WIF) is a project implemented by the International Labour Organisation (ILO) and funded by the United Kingdom's Department for International Development (DFID). Currently, in its second edition,

the programme aims to reduce the incidence of human trafficking among female migrants in South Asia [30]. It is a multi-country intervention developed around four central strategies: (i) pre-departure female empowerment and rights-based training; (ii) fair recruitment; (iii) improved policy and legal frameworks in origin and destination countries; and (iv) workers' collective representation. WIF's activities were designed to promote mobility by choice (versus by compulsion), fair recruitment to decent work, and safety and dignity for migrant workers [45]. The London School of Hygiene and Tropical Medicine (LSHTM) was contracted by DFID as an independent evaluator to help develop, adapt and evaluate WIF's community-based pre-departure component, that is the aforementioned component (i) pre-departure female empowerment and rights-based training [46]. It included door to door visits and two-days training of prospective migrant women on "safe and rights-based migration, financial literacy, rights at work, and how to recognize and protect themselves from the risks of trafficking". By 2018, More than 170,000 women in high-emigration communities in Nepal, Bangladesh and India had participated in WIF pre-departure activities. These activities were designed to increase prospective migrants' knowledge, awareness, information and skills. Specifically, the trainings promoted empowerment strategies that aimed to strengthen women's capacity to assert their rights in situations of disempowerment such as access to entitlements, relations with relatives, labour recruiters, agents of different types, border officials, employers and other stakeholders [45]. The data analysed in this paper were collected at the preliminary stages of the WIF implementation in Nepal.

## Human trafficking in Nepal

Among rural Nepalese populations, international labour migration is an important strategy for improving livelihoods. Migration in Nepal has increased sharply, since the 2000s and remittances represent up to a quarter of the country's GDP. The main destinations for Nepalese female international migrants are Malaysia, Qatar, Saudi Arabia and the UAE. Most migrant women are employed as domestic workers and caregivers, in hotels, catering, manufacturing and health and medical services. More than half (62%) use recruitment agencies to migrate [47]. As a consequence, Nepal has a thriving labour recruitment industry, with almost 800 firms licensed by the Government, and 50,000 brokers engaging in recruitment activities, mostly unlicensed and operating "illegally" [48]. Research on low-skill labour sectors has indicated that exploitation akin to modern slavery (such as forced or bonded labour) and human trafficking are pervasive [49, 50].

Over the past decade, the Nepalese government has acted repeatedly to meet international standards and comply with targets to fight modern slavery and human trafficking [50]. Among these actions, the Government of Nepal has put in place widespread restrictions on female migration. The policy response was intended as a protective measure against exploitation of female migrants abroad following a high number of trafficking cases and slave-like conditions reported by female Nepalese domestic workers, especially in the Gulf States. The most recent migration bans in Nepal include the 2008 ban on women's migration to Gulf countries, the 2009

ban on women's migration to Lebanon for domestic work, the 2012 ban on women below 30 years-old from migrating to Gulf countries for domestic work, and the 2014 ban on women migrating as domestic workers. Reports indicated that these gender discriminatory bans pushed many female migrants into risky migration through informal channels [32] and decreased women's control over their migration process and their workers' rights abroad [35]. In light of this high migration, high-risk context, increasing efforts have been made to protect female migrant workers from harm through pre-departure, community-based activities, such as the WIF interventions [45]. The theory-based evaluation conducted by the LSHTM focused on WIF's community-based pre-departure component, including door-to-door outreach and training in high-emigration communities in Nepal, Bangladesh and India. This paper describes results from the evaluation in Nepal which aimed to examine if WIF's intervention targets and rationale were supported by evidence.

## Methods

### Study design

The data analysed in this paper were collected in a cross-sectional survey with female returnee migrants in Nepal, as part of the theory-based evaluation of the aforementioned Work in Freedom Programme (WIF). The main objective of the study was to verify whether the programmatic assumptions of WIF were supported by relevant evidence.

### Data collection

ILO-Nepal selected five districts with high levels of female labour migration to implement WIF, based on field visits and consultations with local stakeholders. Three of the five selected districts for the intervention were included in our study: Morang, Chitwan and Rupandehi. The other two districts were excluded for logistical reasons.

In each district, implementing partners (organisations delivering the activities on behalf of the ILO) selected 6 Village Department Committees (VDCs), the smaller administrative unit in Nepal, where intervention activities could be delayed until the fieldwork was completed. In these three districts, the partners selected the VDCs. By the time LSHTM received ethical clearance to conduct the survey in Nepal, the implementing partners had already started roll out in VDCs considered by the local stakeholders to have a higher female migration prevalence. This meant that the VDCs selected for a delayed roll-out (and thus included for our study) were not those with potentially the highest migration prevalence. These VDCs were, however, more likely to have a higher prevalence rate of female migration than the country's average, because they were located in high migration districts.

As part of the WIF intervention, one peer educator and one social mobiliser were identified per VDC and trained on issues surrounding labour, female migration and the delivery of the intervention activities. Female labour migration can be a sensitive topic in some communities due to the perception that female migration is linked to sex work [28], and also because of the Government's restrictive policies on female migration [32]. For this reason, we opted to construct our sampling frame based on the WIF mapping instead of conventional household enumeration with the use of a roster to identify returnee migrants. Peer educators and social mobilisers were selected among residents of the VDCs and received intensive training to work with the community. They were responsible for identifying and compiling a list of prospective and returnee migrants through door-to-door visits to households. The list of returnee migrants that resulted from this exercise was our sampling frame. Women aged between 15 and 49 were eligible for inclusion in the returnee migrant surveys if they had worked abroad and returned. We invited all eligible women to take part in our surveys. We did not include data on prospective female migrants who did not report international migration in the present analysis. Interviews were conducted face-to-face by trained female Nepalese interviewers, using electronic tablets for data collection.

## Variables

The questionnaire was developed by the LSHTM team in collaboration with the local partner—Social Sciences Baha. It included questions on: demographic and socioeconomic characteristics; dates, destinations and sectors of previous migrations; decisions, plans, recruitment process and outcomes of the last migration; outcomes of previous migrations (Box 1).

## Statistical analysis

We applied Bayesian (Belief) Networks (BNs) to the modelling of interactions between variables related to trafficking. BNs are a well-established mathematical approach to modelling the relationships in multivariate systems. BNs make inference manageable in complex systems by discovering conditional independencies between variables allowing the, usually large, joint probability distribution to be factorized [51].

We used the data set described above to learn a BN from scratch with minimal a priori assumptions about the relationships between the variables. This predominantly unconstrained approach allows assessment of various major features of the data set, free of domain knowledge assumptions. The Greedy Thick Thinning algorithm [52] was used for the structural learning phase of the model construction. This algorithm is based on the Bayesian Search approach [53]. In the thickening phase, it begins with an empty graph and iteratively adds the next arc that maximally increases the marginal likelihood of the data given the model. This is repeated until no new arcs can be added that will increase the likelihood. Next, in

**Box 1** Main study variables

Variable	Description
Forced labour	Forced labour as defined by the International Labour Organisation [45], includes three dimensions: unfree recruitment, work and life under duress and impossibility of leaving employer. Within each dimension there are indicators of involuntariness and penalty, classified as strong and medium indicators. Each dimension was coded as positive if the woman experienced both involuntariness and penalty, and had at least one strong indicator in any of these two categories. For more detailed explanations of the calculation, please see Mak et al. [61]
Age of migration	Age at date of interview minus date of the most recent departure for migration
Caste	Coded according to the DHS classification Bennett et al. [69] and other key sources of information (UNDP 70; Bhattachan et al. [71])
Education	Highest level of education attended
Literacy	Able to read or write a short note or message
Debt	Coded in three categories: no migration expenses, incurred debt and did not incur debt associated with the migration
Migration for economic reasons	women who responded that the following factors 'very much' influenced their decision to migrate: financial responsibility towards ill family members, pay debt of a family member, pay other type of debt, make additional income for the family or unemployment near their home
Migration for interest or curiosity	Women who reported that the following factors influenced their decision very much: everyone around is migrating, they heard positive experiences from other migrants, they wanted to see a foreign country or they had family at the destination
Awareness of risks	Women who answered that, before leaving their home, they were aware that people who migrate for work could be deceived about their work details and conditions
Network of influences	People who may have influenced women's migration plans, including husbands or in-laws; friends, family or neighbours who were returned migrants; friends, family or neighbours who were not returnee migrants; and agent or broker
Autonomy in migration	Whether the woman made the migration plans by herself, jointly with others or if someone else made the plans for her without her input
Recruitment situation	Coded into four categories: did not use recruiter, used licensed recruiter, used non-licensed recruiter; and used a recruiter but does not know if they were licensed
Pre-migration training	Positive for women who said they attended any training course or programme before their most recent migration
Migration documentation and processes	Women who obtained the following previously to their most recent migration: <i>obtained medical clearance, obtained labour permit, obtained a visa, obtained foreign employment life term insurance and paid into the Welfare fund</i>
Contract	Coded positive for women who reported they had a work contract before migration



**Box 1** (continued)

Variable	Description
Exploitation in prior migrations	Women who reported being exposed to any of the following, in migration experiences prior to her most recent migration: physically abused such as hit with fist or object or worse, by employer, their family, or their friends; financially cheated of what was owed/promised to them; prevented from leaving their employer; seriously injured doing their job
Migration bans	Each variable <i>ban1</i> , <i>ban2</i> , <i>ban3</i> and <i>ban4</i> was coded positive if women who migrated during the time when these bans were in place. Variable <i>ban1</i> refers to the 2008 ban on embarking to Gulf countries and Malaysia; <i>ban2</i> is the 2009 ban on migration to Lebanon; <i>ban3</i> 2012 is the ban on women aged less than 30 migrating to Arab States; and <i>ban4</i> is the 2014 ban on all women migrating as domestic workers

the thinning phase, it repeatedly removes arcs until no arc deletion will increase the likelihood.

Although the BN structure was primarily derived from the data, a limited set of constraints were imposed based on our prior knowledge. We divided the variables into 4 tiers such that variables in higher tiers could not influence variables in lower tiers. This enables us to forbid certain types of influence, which we know are logically impossible. The variable *caste* was assigned to Tier 1, *literacy* and *education* to Tier 2, and all other predictors were placed in Tier 3. The outcome variables (*work and life under duress*, *unfree recruitment*, *impossibility of leaving the employer* and *forced labour*) were assigned to Tier 4. In the BN, the forced labour variable is specified as a deterministic node instead of a probabilistic (chance) node as it is related deterministically to the three outcome variables by definition.

The choice of structure learning algorithm can affect the final BN. There are a large number of different algorithms in general use which can be assigned to one of the 3 main categories: score-based, constraint-based and hybrid. The Greedy Thick Thinning algorithm we used belongs to the score-based family. Recent work suggests that constraint-based and hybrid algorithms are often less accurate than score-based algorithms [54]. It is well understood that the networks produced using different algorithms can differ. We intend to explore this aspect of the methodology in future work. However, our approach incorporates independent validation of the model—prediction accuracy was calculated using the technique of leave-one-out cross-validation. This is the gold standard technique for preventing the influence of overfitting on model performance metrics. The results of the model evaluation provide strong independent support for the main conclusions we draw from the model.

The strength of influence between variables is calculated as the average effect of changing the state of a parent node on the probability distribution of states in the child node. For example, if we changed the state of the *destination* variable from Kuwait to India, how much would that change the distribution of the *worklife duress* variable.

Any number of simulations (what-if scenarios) can be run once a BN has been built. The state of any variable(s) can be set to a specific value(s). Alternatively, the probability distribution across states for a variable can be specified. The effects of these changes on other variables of interest can then be measured precisely. We have tested the effects of changes to variables that the model suggests strongly influence the outcomes of interest (*destination, recruiter*).

A sensitivity analysis [59] was performed using the method proposed by Kjaerulff and van der Gaag [60].

The implementation of BNs used in this paper was from the GeNIe Modeler software package (BayesFusion LLC, <https://www.bayesfusion.com/>).

## Why Bayesian networks?

Diverse research fields have recognized the power of BNs for discovering complex causal relationships in multivariate and even cross-sectional data sets. BNs can combine a priori knowledge with empirical data, cope well with missing values and predict the effects of different interventions on the outcome variables of interest. BNs can model non-linear interactions, multiple causal pathways, population heterogeneity and changes over time and across contexts [17, 22]. They can be a powerful tool to inform intervention development and evaluation of complex interventions [54–58]. We use BNs to analyze the interactions between risk and protective factors among Nepalese female migrants to: (a) test widely held assumptions in the anti-trafficking field; (b) identify promising intervention opportunities to prevent human trafficking; (c) simulate the effects of potential interventions on key variables along the causal pathway to forced labour.

## Results

### Characteristics of survey participants and their migration trajectory

The prevalence of forced labour during women's most recent migration was 90.4% (95%CI: 87.4, 92.8). Almost all migrant women experienced work-life duress (89.8%). More than half (52.0%) experienced unfree recruitment, and 19.1% reported being unable to leave their employer. Almost one in five women (19.9%) experienced all three of the above dimensions of forced labour.

The mean age at the most recent migration in our sample was 28 years-old (SD 7.7). The median time elapsed, since their return from migration was 4.8 years (mean 5, SD 2.5).

One in four women (24.9%) were illiterate and almost one in three migrants were from the lowest Dalit castes (31.4%). The majority of women (83.2%) in our sample reported a pressing economic need to migrate. Of those with a pressing economic reason, the majority (74.2%) also reported other motivations, such as interest or

curiosity. More than one-quarter of women (26.1%) said that being approached by a recruiter or agent influenced their decision to migrate.

The majority of the women (64.7%) had to pay for their migration expenses and more than half (55.7%) incurred debt to cover these expenses. More than one in ten women (11.2%) had an advance from the recruiter to pay for their migration and nearly one-quarter (22.4%) borrowed money from relatives.

The majority of women used a recruiter (65.5%) to help find or arrange a job and almost a third (29.3%) had help from family or friends who had migrated before.

Most women (60%) had migrated only once. Of those who migrated more than once (40%), 43.7% reported some type of exploitation or abuse during previous migrations. When asked about their knowledge of migration-related risks, over two-thirds of the women (67.8%) stated they were aware that migrants can be deceived about their work details and conditions. A detailed description of the sample can be found in SM 2.

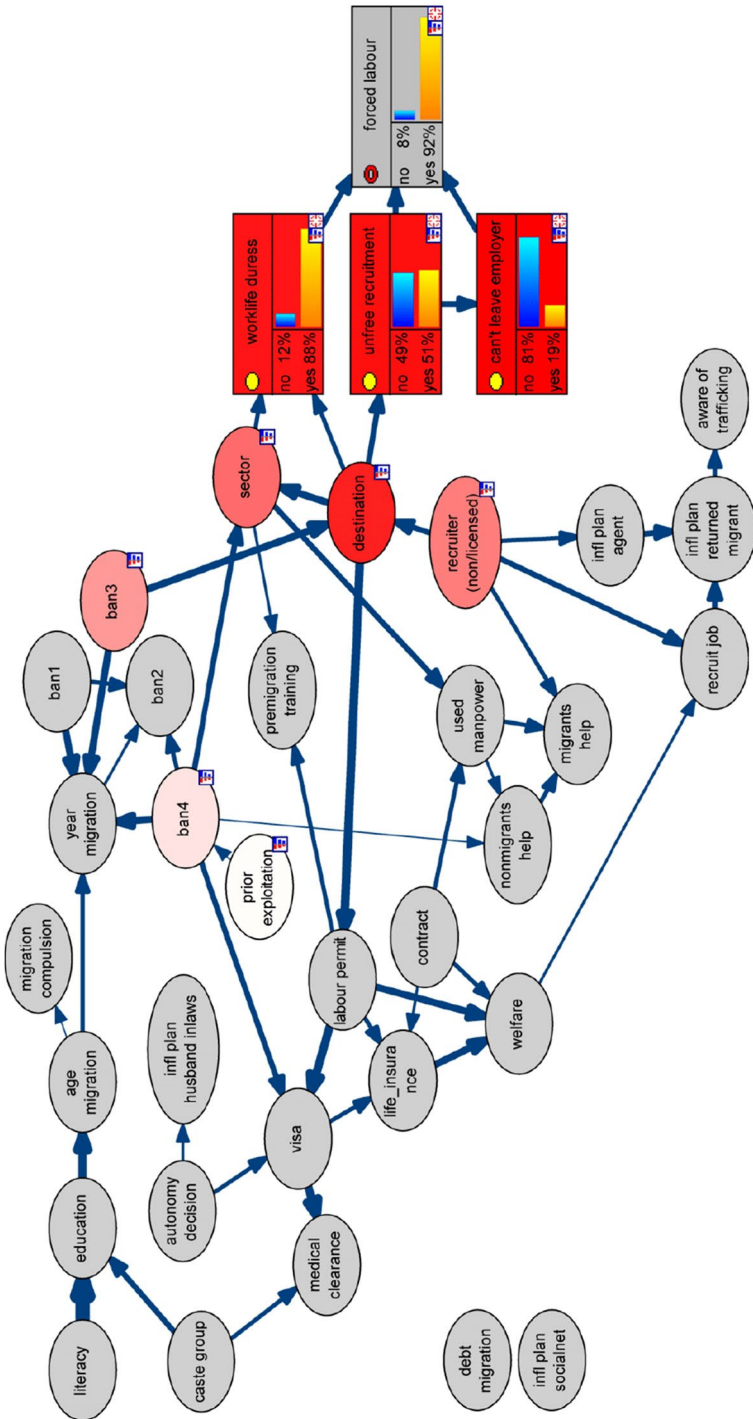
### Causal pathways to forced labour among female migrants from Nepal

Figure 1 shows the BN derived from the WIF data set. The nodes are colour-coded according to how sensitive the outcome variables are to small changes in the probabilities in each ancestor node. The nodes to which the outcomes were most sensitive are the proximal nodes which is generally the case in BNs. The first order variables (*destination* and *sector*) are those with a direct causal influence on any of the three outcome variables. The only other variables with any influence on the outcomes are *recruiter (non/licensed)*, *ban3*, *ban4* and *prior exploitation*, and their effects are mediated via the *sector* and *destination*, respectively. Any variable that had no direct or indirect influence on an outcome variable is shown in grey. These nodes are not ancestor nodes of the outcome variables, which means they did not have any influence on the outcomes in our model.

*Destination* influenced *work and life under duress* both directly, and indirectly via the work *sector*. *Destination* also directly influenced *unfree recruitment*.

### Prediction accuracy

Table 1 shows the prediction accuracy of the BN model. For each migrant in the data set, we use the states of each of the independent variables in the model to predict the probabilities of the three outcome variables. If the actual value (true or false) of the outcome variable for that migrant is predicted by the model, i.e., has a greater than 50% probability, this counts as a correct prediction. The results in Table 1 are split according to whether the reported outcome is true or false. The model is best at predicting work-life duress and worst at predicting unfree recruitment. As is often the case with binary model outputs, the more frequent output state in the data is the one that is predicted most accurately for all three outcomes. This is extreme in the case of the *can't leave employer* outcome, where the model has over-generalised and predicted all migrants are false for this variable. There is nonetheless clear predictive



**Fig. 1** Data derived BN model with sensitivity analysis for the WIF data set (N = 519). Stronger colours indicate nodes whose parameters (probabilities) have the greatest effect on the target outcomes. Strength of influence is represented by the thickness of the arrows

**Table 1** Prediction accuracy for the model in Fig. 1. The results were generated using leave-one-out cross-validation

	Unfree recruitment	Impossibility leaving employer	Work and life under duress	Forced labour
False	44.1% (112/254)	100% (420/420)	45.3% (24/53)	63.2% (24/38)
True	82.3% (218/265)	0% (0/99)	95.7% (446/466)	94.6% (455/481)
Average	63.6% (330/519)	80.9% (420/519)	90.6% (470/519)	92.3% (479/519)

power in the model. This is confirmed by the ROC Area Under Curve metric, which is above chance (0.5) for all three outcomes.

### What-if scenarios

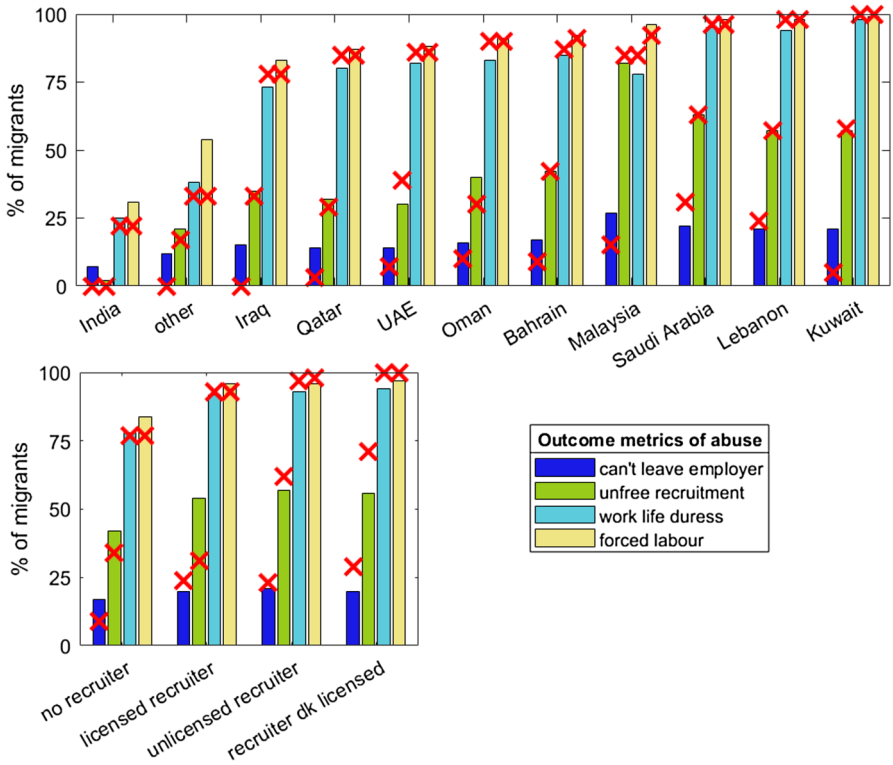
The sensitivity analysis, shown by the colors of the nodes in Fig. 1, indicates that changes to the variable *destination* have the largest influence on the outcomes. Hence, the first application of our model was to observe the effects of setting the *destination* variable (called ‘setting evidence’ in the language of BNs). This is equivalent to saying: if we knew the *destination* was country X, what would we expect the probabilities of the outcomes to be? The results are shown in Fig. 2.

The results in Fig. 2 indicate that the outcomes are highly sensitive to the *destination* variable. The disparity in the probability of experiencing abuse across the different countries means that small changes in the probabilities of migrating to each country can lead to large changes in the probability of forced labour. In particular, migrants going to India have a much reduced incidence of adverse outcomes compared to all other countries.

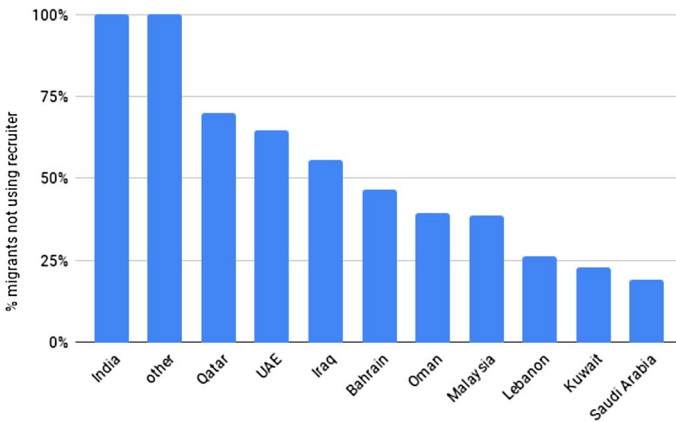
In addition, Fig. 2 shows a good correspondence between the model predictions and the raw data. This confirms the results of the prediction errors shown in Table 1.

### Recruiter—destination influence

Our model suggests that the effect of *recruiters* is mediated via *destination*. An analysis of the relationship between the two variables suggests an underlying mechanism. The raw data in Fig. 3 show that no migrants to India used a recruiter. This contrasts with the proportions using recruiters when going to other countries. This implies an explanation for the effect that recruiters have on the forced labour outcomes—namely that not using a recruiter will mean that a migrant is more likely to travel to India, where they will be less likely to experience abuse. It could be argued that the direction of causality could run in either direction: (a) if a migrant uses a recruiter, the recruiter will be most likely to direct her towards migration to countries other than India (*recruiter* causes *destination*) or (b) if a migrant decides she wants to migrate to a country other than India, she is more likely to use a recruiter (*destination* causes *recruiter*). Although domain knowledge suggests that (b) may be more likely, the structure learning algorithm used to build the model concluded (a) is the dominant effect.



**Fig. 2** Relationship between destination country (top chart) and forced labour outcomes and recruiters (bottom chart) and forced labour outcomes. The bars show the model predictions and the red crosses show the corresponding proportions taken from the raw data set



**Fig. 3** Variation in the use of recruiters by destination country. These are the proportions taken from the raw data set

## Migration bans

The model identifies an influence of the migration bans on migration outcomes. In particular *ban3* (ban on domestic workers younger than 30 from migrating to Arab States) influenced the *destination* variable and *ban4* (ban on women migrating as domestic workers) influenced the *sector* variable. These results are intuitively reasonable given the nature of the bans—*ban3* has a specified *destination* target and *ban4* has a specified *sector* target. However, in our sample, there were 22 women who defied *ban3*, i.e., were under 30 and migrated to an Arab state to do domestic work. Of these women, 16 (73%) used a recruiter whose license status was unknown, and the remainder did not use a recruiter. Thus the ban appears to have been very successful in shutting down the licensed recruitment channel to the Arab states.

## Discussion

The Work in Freedom Program was among the largest multi-country initiatives to try to prevent human trafficking in a region blighted by these abuses. The British Government's investment in this research offered an unprecedented learning opportunity for the field. This paper presents a valuable new approach that demonstrates the importance of evidence to intervention development and the promise of a second generation of more sophisticated anti-trafficking research and evaluation methods. Importantly, this paper demonstrates how BNs can be used for the conceptualisation, design and evaluation of complex social interventions. The results of our model suggest that the underlying theory for the WIF community-based intervention rested on misguided assumptions. Migrant women's empowerment is unlikely to prevent human trafficking or exploitation at a population level.

Almost all of the Nepali returnee women in our sample (90%) experienced forced labour. The vast majority experienced labour exploitation with differing degrees of restrictions on their freedom, deception, coercion, violence and abuse. Moreover, our previous study in the Dolakha district of Nepal found that 73% of returnee men experienced forced labour [61]. While we cannot conclude that these enormously high figures for labour exploitation are nationally representative, they nonetheless suggest the substantial risk of exploitation among international labour migrants. It is likely that official data from records on identified trafficked people derived from anti-trafficking rescue efforts are only revealing the tip of the iceberg. Most cases that fulfil the ILO criteria for forced labour will be invisible to surveys conducted at post-trafficking services. To fully understand how to address human trafficking will require evidence from population-based surveys similar to that reported here, which are not restricted to identified cases, but instead have the power to more accurately describe the extent and nature of the problem in a community. While data of the type reported here are difficult to obtain, their advantages are clear.

## Implications for interventions

The usefulness of modelling the causal relations between variables using BNs emerges when considering interventions to reduce forced labour. The model directly suggests which factors may merit intervention efforts (ancestor/coloured nodes) and which do not (non-ancestor/grey nodes) in this or similar contexts. For example, the model suggests that it would not be an efficient use of resources to try to affect the use of manpower agencies (*used manpower*). Whereas trying to improve employment and working conditions in specific *destination* countries would be more likely to affect the outcomes.

As is always the case in such studies, we cannot rule out the possibility of hidden variable(s) which could account for the relationship between destination country and forced labour. Any factor which caused migrants to choose particular destination countries *and* caused them to experience forced labour could produce the relationships represented in our BN. The clearest way to test whether destination country has a *direct causal* influence is to investigate the effect of intervention efforts applied directly to the *destination* factor. We believe that further analysis of the bans imposed by the Nepalese government can help clarify the effect of *destination* on the likelihood of forced labour.

Contrary to recent findings on the higher incidence of trafficking resulting from the migration bans by the Nepalese government [32], our study suggests that women migrating during the most recent bans (ban on women less than 30 years-old migrating to Arab States and ban on all women migrating for domestic work) were less likely to experience forced labour. On the other hand, the 2008 ban on migration of low skilled migrants to the Gulf countries and the 2009 ban on female migration to Lebanon—both imposing restrictions to risky destinations—had no significant effect on forced labour. However, even if migration bans can have a protective effect against women's entry into situations of forced labour, they can also further limit livelihood options and agency, reinforcing gender asymmetries [25].

Women's personal demographic and socioeconomic characteristics had very limited influence on their probability of experiencing exploitation or any dimension of forced labour. This is a finding that screams for attention from the anti-trafficking donors and policy-makers, who remain heavily invested in interventions to address generic individual vulnerabilities such as poverty and education levels. What affects exploitative outcomes is: where migrants go, how they are recruited and in which sector they work. These factors are much more influential than an individual's age or personal financial situation—which may not differentiate their risk status at all. These findings suggest that for our study population at least, prevention should focus on the recruitment process, migration pathways (*destination, sector*) and working conditions at the destination—that is, the employment conditions and employer. Interventions should refrain from targeting girls and women solely based on caste, education, family situation or pre-departure poverty levels—because in a world with limited resources, these types of investments will likely have negligible effects.

Our findings confirm that raising an individual's awareness of general migration-related risks is not an effective strategy for reducing their likelihood of forced labour.



In fact, women's own prior experiences of labour exploitation in migration had only a weak effect on their probability of experiencing forced labour in their next migration experience. These findings provide further evidence that the current approach to awareness-raising is an ineffective anti-trafficking strategy. An individual's direct or indirect awareness of risks as assumed or explained in present pre-migration training rarely seems to affect their probability of entering into high risk migration.

The finding that a migrant's destination location is a primary risk factor for forced labour corresponds with the literature that describes the increased risks among workers who migrate to certain countries—especially countries where employment contracts bind employees to employers [25]. Reports repeatedly indicate that migration sponsorship programmes, such as the *kafala* system in the Gulf Cooperation Council (GCC) countries, increase the vulnerability of foreign construction and domestic workers [62]. Indeed, abusive practices such as non-payment of wages, long working hours, unsafe accommodation, limited access to justice and violence have been widely reported by domestic workers in the GCC [63]. Women migrating to these countries are more likely to use a recruiter, which may also increase the costs of migration and risk of deception [62]. Conversely, women travelling to India are less likely to experience the extreme aspects of forced labour compared to the other destination locations. This difference may be due to the open border between Nepal and India, which means that migrants are less tied to employers and recruiters for their initial migration or for their return. In addition, the costs and logistics of return might be easier for migrants living in India.

Our data also indicate that female economic migration from the districts in our study is not necessarily motivated solely by economic distress. This is important, because there has been an emphasis in the literature on migrants as “abject victims of structural compulsion” [25]. Instead, our data reveal that migrants often value their choices, in spite of structural constraints. Many migrants are attracted by the financial opportunity that migration can bring. Moreover, with growing international attention to migration, such as the UN Global Compact on Migration, investments in local livelihoods remain a focus of substantial attention. However, interventions that focus on alternative local livelihoods as either anti-trafficking or migration safety strategies will wish to take account of: the diverse aspirations among prospective migrants; local consumption expectations (by age and gender) that have resulted from a globalized economy; and government policies in origin and destination countries. Local livelihood options that are not competitive with job opportunities abroad in terms of prospective earnings are unlikely to succeed.

Employment contracts have been seen by some as the “harbinger” of freedom and equality for modern-slaves [25]. Our findings, however, suggest that work contracts do not have an intrinsic protective role for female labour migrants from Nepal. Practices such as contract substitutions and hidden clauses may be partially to blame for this [39]. Therefore, it may be worth questioning whether it is useful for pre-migration training and information sessions to focus on helping individuals negotiate contracts or review contract content, when enforcement of clauses is not possible.

Our data also suggest that the recruiters used by the women in Nepal influenced their risk of forced labour, via the *destination*. Recent qualitative research findings have highlighted the diversity of recruitment practices and the role of recruiters in

facilitating safer migration [31, 64]. Our results indicate that, on average, using a recruiter *decreased* women's likelihood of safer migration. These results should not be used to vilify recruiters in general, because we still have very little evidence on the role of recruiters and the spectrum of recruitment services in most contexts. This finding indicates that some current recruitment practices may increase risks of forced labour or that migrants choosing alternatives to current recruitment practices (e.g., social network facilitated migration) might be protected by these other pathways. The complex bureaucracy, multi-locality of risks and procedural uncertainty of international migration processes make labour recruiters indispensable for certain destinations [31, 48, 64]. This was evidenced in our model by the link between *recruiter* and the many documents and processes related to migration. Previous research has shown that information asymmetries between migrants and recruiters/employers, tend to shift costs to migrants and impact on their employment security and stability [62]. Initiatives such as the ITUC Migrant Advisor platform have offered migration-related information, but these types of information formats have yet to be evaluated for their effects on work outcomes [65].

## Limitations

We opted to base our sampling frame on the WIF mapping of returned migrants because of probable underreporting of migration status by women in our formative research in Nepal. Although the ILO has devised a strategy with the engagement of local experts and the implementing agencies have visited a high proportion of households in the selected villages, the sample for this study cannot be considered representative of the overall population of female migrants in the selected clusters. Women who previously migrated to India may be particularly under-represented in our sample despite our explicit attempts to encourage WIF peer-educators to include India. India is not considered as 'international' migration in the local context because of its proximity and open border. Women who were still in work in foreign countries at the time of the survey were obviously excluded from our sample. This could create 2 opposing biases—are they more likely to be experiencing good conditions and hence choosing to stay or conversely are unable to leave. Our study cannot shed light on the relative weight of these factors but some sampling bias is likely. At the same time, despite a lack of systematic data, reports indicate that temporary migration based on short-term contracts is the main form of migration among Nepalese low-wage labourers. Most international migrants from Nepal return home [66, 67].

Another limitation of our study was the impossibility of exploring some intrinsic gendered aspects of migration. We were not able to use household information to better understand their position and financial role in the household, as data on these variables were only available for the time of the interview, after the women returned from their previous migration. Their circumstances and their household socio-economic situation could have changed considerably, since their most recent migration and, therefore, these variables were not included in this study.

For most relationships, the directions of arcs in the BN were automatically inferred from the data. However, for some of the relationships, the algorithm was not able to infer the direction. In these cases, we used domain knowledge to manually select the most likely direction. In particular, this was required for subnetworks of interrelated variables, where the direction of influence was unclear. This could suggest a possible hidden variable influencing all of the variables in the sub-network. Examples of these sub-networks were ‘migration documentation and processes’ (i.e. *visa, labour permit, medical clearance, welfare payment*), and ‘timing of migration’ (i.e., *bans and year of migration*).

The potential for recall bias is probably the main limitation of our study. Women were asked to recall experiences that happened several years ago. The passage of time will almost certainly have affected the completeness and accuracy of their recollection of events and perceptions at the time of migration.

Despite these limitations, this is one of the first papers to examine the multicausality of human trafficking based on empirical data.

### **Empirical and methodological contributions**

This paper recognizes the pressing need for evidence to inform investments in human trafficking, modern-slavery and forced labour prevention [23, 26]. The importance of these findings cannot be overstated. For over two decades now, the international community has been programming based on the idea that there are patterns of individual-level vulnerability to human trafficking. Yet, our findings suggest that this assumption is unsafe. In fact, trying to target individual characteristics may be misdirecting the investments which could be better spent on, for instance, changing harmful recruitment practices, making positive social migrant networks stronger or discouraging migration to countries where women and men are most likely to be treated as ‘slaves’.

Compared to other complex social problems such as obesity, smoking or intimate partner violence, research on human trafficking grapples with the additional challenge of longitudinal data collection in transnational contexts among hard-to-reach populations. Moreover, trafficking often involves multiple agents (e.g., migrants, recruiters, officials, employers) acting in informal and hidden labour sectors with conflicting interests and constrained by Estate-level laws, regulations and policies. This complexity suggests that intervention research needs to focus on the systemic and dynamic causal pathways that influence human trafficking. Current research needs to go beyond analytical approaches that assume linear, average effects of exposures on single outcomes affecting homogenous populations.

From a realist evaluation perspective, the use of BNs allowed us to identify important gaps in the WIF’s theory. The findings from our analysis indicate that WIF’s reliance on individual knowledge and empowerment as a mechanism for prevention of human trafficking is unlikely to produce safe and profitable

migration for Nepalese female migrants. This finding is in line with results from LSHTM's qualitative evaluation of WIF in Bangladesh [31].

While BNs have been used to model social-behavioural mechanisms in other contexts [17, 19, 57], this is the first attempt to use BNs to analyse the processes underlying labour migration. The results of our study confirm that BNs are a useful tool for modelling complex systems and assisting realist evaluations. One of the key advantages of the approach is that it provides a visual representation of the variables and their causal inter-relationships, which facilitates understanding compared to a model formulated purely of mathematical symbols. The additional features such as sensitivity analyses, strength of influences, what-if scenarios and the ability to combine a priori domain knowledge with empirical data, provide a rich set of tools to aid the understanding of mechanisms in complex multivariate systems. In this paper, we have shown how BNs can help focus attention on key aspects of a mechanism thus allowing further detailed analysis to be concentrated on the important relationships in a relatively large multivariate data set. Moreover, we used BNs to simulate the effect of interventions—an overarching goal of this type of study.

BNs should not be regarded as a substitute for other quantitative impact evaluation methods [58]. Nonetheless, they can offer important contributions to the cumulative evidence-base on the mechanisms and effects of complex social interventions. Indeed, the advantage of using BNs in Realist Evaluation is in the possibility of combining robust inference within a complex system (from the network of probabilistic dependencies) with strong theory (based on the analysis of context-mechanism-outcome configurations proposed by realist evaluation). Realist evaluation needs both causal reasoning and empirical evidence to address the question of “what works, for whom and in what circumstances” [68]. BNs can be a valuable instrument in empirically elucidating the complex interrelationships investigated in Realist Evaluations, both from an inferential and explanatory perspective.

**Acknowledgements** This research was conducted by the London School of Hygiene and Tropical Medicine, with financial support from the United Kingdom Department for International Development (contract number: PO 5732). The International Labour Organisation also supported initial stages of data collection.

**Author contributions** Dr. Kiss co-designed the original study with Ms Mak, conceived the article, co-designed the analysis strategy with Dr. Fotheringham and Ms McAlpine. Dr. Kiss and Dr. Fotheringham led on interpretation of the results and were responsible for the first draft of the paper. Dr. Fotheringham conceived the analytical strategy, lead on data analysis and writing of the results. Dr. Mak co-designed the study with Dr. Kiss, developed the instruments with Dr. Zimmerman and Dr. Kiss, coordinated data collection and cleaning, and critically revised the manuscript, offering intellectual input. Ms McAlpine contributed to formulating the analytical strategy and revised the article for important intellectual content. Dr. Zimmerman contributed to data interpretation and revised the paper critically for important intellectual input.

**Data availability** The dataset is not publicly available, but the authors can analyse individual requests for access. Researchers who are interested in obtaining the dataset should contact the first author.

## Compliance with ethical standards

**Conflicts of interest** The authors have no conflicts of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Fink, D., & Keys, K. M. (2017). Wrong answers: when simple interpretations create complex problem. In A. M. El-Sayed & S. Galea (Eds.), *Systems science and population health* (pp. 25–35). New York: Oxford University Press.
2. Keyes, K. M., & Galea, S. (2017). Commentary: The limits of risk factors revisited: is it time for a causal architecture approach? *Epidemiology*, *28*(1), 1–5.
3. Galea, S., & Vaughan, R. D. (2017). Can population health science counter in-kind dangerous oversimplifications? A public health of consequence. *American Journal of Public Health*, *107*(10), 1538–1540.
4. Rutter, H., Savona, N., Glonti, K., Bibby, J., Cummins, S., Finegood, D. T., et al. (2017). The need for a complex systems model of evidence for public health. *The Lancet*, *390*(10112), 2602–2604.
5. Shiell, A., Hawe, P., & Gold, L. (2008). Complex interventions or complex systems? Implications for health economic evaluation. *BMJ*, *336*(7656), 1281–1283.
6. El-Sayed, A. M., & Galea, S. (2017). *Systems science and population health* (pp. 25–35). New York: Oxford University Press.
7. Kaplan, G. A., Diez-Roux, A. V., Simon, C. P., & Galea, S. (2017). *Growing inequality: Bridging complex systems, population health, and health disparities*. Washington DC: Westphalia Press.
8. Rockhill, B. (2005). Theorizing about causes at the individual level while estimating effects at the population level: Implications for prevention. *Epidemiology*, *16*(1), 124–129.
9. Shy, C. M. (1997). The failure of academic epidemiology: Witness for the prosecution. *American Journal of Epidemiology*, *145*(6), 479–484.
10. Byrne, D., & Callaghan, G. (2013). *Complexity theory and the social sciences: The state of the art*. London: Routledge.
11. Bousquet, A., & Curtis, S. (2011). Beyond models and metaphors: Complexity theory, systems thinking and international relations. *Cambridge Review of International Affairs*, *24*(01), 43–62.
12. Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. London: Sage.
13. Fletcher, A., Jamal, F., Moore, G., Evans, R. E., Murphy, S., & Bonell, C. (2016). Realist complex intervention science: Applying realist principles across all phases of the Medical Research Council framework for developing and evaluating complex interventions. *Evaluation*, *22*(3), 286–303.
14. Bonell, C., Fletcher, A., Morton, M., Lorenc, T., & Moore, L. (2012). Realist randomised controlled trials: a new approach to evaluating complex public health interventions. *Social Science & Medicine*, *75*(12), 2299–2306.
15. Dalkin, S. M., Greenhalgh, J., Jones, D., Cunningham, B., & Lhussier, M. (2015). What's in a mechanism? Development of a key concept in realist evaluation. *Implementation Science*, *10*(1), 49.
16. Epstein, J. M. (1999). Agent-based computational models and generative social science. *Complexity*, *4*(5), 41–60.
17. Whitney, P., White, A., Walsh, S., Dalton, A., & Brothers, A. (2011). Bayesian Networks for Social Modeling. In J. Salerno, S. JayYang, D. Nau, & S. K. Chai (Eds.), *Social computing, behavioral-cultural modeling and prediction*. New York: Springer.
18. Sticha, P., Axelrad, E., & Schisler-Do, T. P. Program Evaluation using Bayesian Networks: An Introduction. American Evaluation Association. Slide Presentation October 16, 2013. Consulted in 5th November 2018 at: Program Evaluation Using Bayesian Networks: An introduction.

19. McNally, R. J., Heeren, A., & Robinaugh, D. J. (2017). A Bayesian network analysis of posttraumatic stress disorder symptoms in adults reporting childhood sexual abuse. *European Journal of Psychotraumatology*, 8(sup3), 1341276.
20. Gupta, S., & Kim, H. W. (2008). Linking structural equation modeling to Bayesian networks: Decision support for customer retention in virtual communities. *European Journal of Operational Research*, 190(3), 818–833.
21. Uusitalo, L. (2007). Advantages and challenges of Bayesian networks in environmental modelling. *Ecological modelling*, 203(3–4), 312–318.
22. Stewart, G. B., Mengersen, K., & Meader, N. (2014). Potential uses of Bayesian networks as tools for synthesis of systematic reviews of complex interventions. *Research synthesis methods*, 5(1), 1–2.
23. Davy, D. (2016). Anti-human trafficking interventions: How do we know if they are working? *American Journal of Evaluation*, 37(4), 486–504.
24. Gleason, K. A., & Cockayne, J. (2018). *Official development assistance and SDG Target 8.7: Measuring aid to address forced labour, modern slavery, human trafficking and child labour. Research brief*. New York: United Nations University.
25. Davidson, J. O. C. (2015). *Modern slavery: The margins of freedom*. London: Palgrave MacMillan Springer.
26. Kiss, L., & Zimmerman, C. (2019). Human trafficking and labor exploitation: Toward identifying, implementing, and evaluating effective responses. *PLOS Medicine*, 16(1), e1002740.
27. United Nations. (2017). Call to action to end forced labour, modern slavery and human trafficking. In *72nd UN General Assembly*. New York: United Nations. Retrieved from [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/759332/End-Force-d-Labour-Modern-Slavery1.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/759332/End-Force-d-Labour-Modern-Slavery1.pdf).
28. Van der Laan, P., Smit, M., Busschers, I., & Aarten, P. (2011). Cross-border trafficking in human beings: prevention and intervention. *Campbell Systematic Reviews*. Retrieved from <https://www.publicsafety.gc.ca/lbrr/archives/cnmcs-plcng/cn34643-eng.pdf>
29. Tzvetkova, M. (2002). NGO responses to trafficking in women. *Gender & Development*, 10(1), 60–68.
30. International Labour Organization. (2017). *DFID announces £13 million for ILO Work in Freedom programme: ILO* (cited 2019 12 March 2019). Retrieved from [https://www.ilo.org/global/about-the-ilo/newsroom/news/WCMS\\_608112/lang--en/index.htm](https://www.ilo.org/global/about-the-ilo/newsroom/news/WCMS_608112/lang--en/index.htm).
31. Blanchet, T., Biswas, H., Zaman, A., & Lucky, M.A. (2018). From risks to rights: evaluation of a training programme for women aspiring to migrate for work (report). Dhaka: 2018. Retrieved from <https://www.lshtm.ac.uk/files/swift-bangladesh-evaluation-report-2018.pdf>.
32. International Labor Office (ILO). (2015). *No easy exit: Migration bans affecting women from Nepal*. Geneva: ILO. Retrieved from [https://www.ilo.org/global/topics/forced-labour/publications/WCMS\\_428686/lang--en/index.htm](https://www.ilo.org/global/topics/forced-labour/publications/WCMS_428686/lang--en/index.htm).
33. Kaufman, M. R., & Crawford, M. (2011). Research and activism review: Sex trafficking in Nepal: A review of intervention and prevention programs. *Violence Against Women*, 17(5), 651–665.
34. International Organisation for Migration (IOM). (2017). *Migration: The move from rural to urban by choice 2017* (cited 2019 12 March). Retrieved from <https://www.iom.int/news/migration-making-move-rural-urban-choice>.
35. Idris, I. (2017). *Interventions to combat modern slavery. K4D Helpdesk Report*. Brighton, UK: Institute of Development Studies. Retrieved from <https://assets.publishing.service.gov.uk/media/5a5f23f240f0b652634c6f4d/Interventions-to-combat-modern-slavery.pdf>.
36. Fratzke, S., & Salant, B. (2017) *Understanding the impact of livelihood opportunities and interventions on migration patterns: Rapid evidence assessment*. London: UK Department for International Development. Retrieved from [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/685188/Migration-LivelihoodsI.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/685188/Migration-LivelihoodsI.pdf).
37. International Labor Organization (ILO). (2017) *Fair Recruitment Initiative: Fostering fair recruitment practices, preventing human trafficking and reducing the costs of labour migration*. Geneva: ILO. Retrieved from [https://www.ilo.org/wcmsp5/groups/public/---ed\\_norm/---declaration/documents/publication/wcms\\_320405.pdf](https://www.ilo.org/wcmsp5/groups/public/---ed_norm/---declaration/documents/publication/wcms_320405.pdf).
38. Andrees, B., Nasri, A., & Swiniarski, P. (2015). *Fair recruitment initiative: Regulating labour recruitment to prevent human trafficking and to foster fair migration: models, challenges and opportunities*. Geneva: ILO. Retrieved from [https://www.ilo.org/wcmsp5/groups/public/---ed\\_norm/---declaration/documents/publication/wcms\\_377813.pdf](https://www.ilo.org/wcmsp5/groups/public/---ed_norm/---declaration/documents/publication/wcms_377813.pdf).

39. Sorrentino, L., & Jokinen, A. (2014). *Guidelines to prevent abusive recruitment, exploitative employment and trafficking of migrant workers in the Baltic Sea region. European Insituite for Crime Prevention and Control (HEUNI). Publication Series No.78.* Helsinki: Hakapaino Oy. Retrieved from <https://www.cbss.org/wp-content/uploads/2012/11/ADSTRINGO-Guidelines.pdf>.
40. Epstein, G. S. (2008). Herd and network effects in migration decision-making. *Journal of Ethnic and Migration Studies*, 34(4), 567–583.
41. Constraints to a robust evidence base for anti-trafficking interventions. (2017). Anti-trafficking review (8).
42. United Nations. (2017). *Sustainable Development Goal 8: Target and indicators 2017* (cited 2019 12 March). Retrieved from <https://sustainabledevelopment.un.org/sdg8>.
43. Cockayne, J., Gleason, K., Cholette, E., Savel, M., Enica, O., & Harris, B. (2018). *Using data and evidence to achieve target 8.7.* New York: United Nations University Centre for Policy Research.
44. Tyldum, G., & Brunovskis, A. (2005). Describing the unobserved: Methodological challenges in empirical studies on human trafficking. *International Migration*, 43(1–2), 17–34. <https://doi.org/10.1111/j.0020-7985.2005.00310.x>.
45. International Labour Organisation, UK Department for International Development. (2018). *Work in Freedom: reducing the vulnerability to trafficking of women and girls in South Asia and the Middle East: supporting informed migration, fair recruitment and decent work.* ILO, DFID. Retrieved from [https://www.ilo.org/wcmsp5/groups/public/---asia/---ro-bangkok/---sro-new\\_delhi/documents/publication/wcms\\_526175.pdf](https://www.ilo.org/wcmsp5/groups/public/---asia/---ro-bangkok/---sro-new_delhi/documents/publication/wcms_526175.pdf).
46. London School of Hygiene and Tropical Medicine. (2016). *South Asia Work in Freedom Transnational Evaluation (SWiFT).* London: LSHTM. Retrieved from <https://www.lshtm.ac.uk/files/swift-evaluation-overview.pdf>.
47. Nepal National Human Rights Commission. (2014). *Trafficking in persons especially on women and children in Nepal.* Kathmandu: NHRC.
48. Jones, K., Morin, S., Dobree, P., & Irimu, W. (2015). *The migrant recruitment industry: Profitability and unethical business practices in Nepal, Paraguay and Kenya.* Geneva: ILO.
49. Sijapati, B., Limbu, A., Khadka, M., & Aryal, S. (2011). *Trafficking and forced labour in Nepal: A review of the literature.* Kathmandu: Himal Books.
50. US Department of State. (2018). *Trafficking in Persons report 2018: Nepal.* US Department of State. Retrieved from <https://www.state.gov/documents/organization/282803.pdf>.
51. Barber, D. (2012). *Bayesian reasoning and machine learning.* Cambridge: Cambridge University Press.
52. Cheng, J., Bell, D. A., & Liu, W. (1997). An algorithm for Bayesian belief network construction from data. In *Proceedings of AI & STAT'97*, Citeseer.
53. Cooper, G. F., & Herskovits, E. (1992). A Bayesian method for the induction of probabilistic networks from data. *Machine learning*, 9(4), 309–347.
54. Scutari, M., Graafland, C. E., & Gutiérrez, J. M. (2019). Who learns better Bayesian network structures: Accuracy and speed of structure learning algorithms. *International Journal of Approximate Reasoning*, 115, 235–253.
55. Spiegelhalter, D. J. (1998). Bayesian graphical modelling: A case-study in monitoring health outcomes. *Applied Statistics*, 47(1), 115–133.
56. Le, Q. A. (2010). *New approaches using probabilistic graphical models in health economics and outcomes research.* Dissertation. Faculty of the USC Graduate School, University of Southern California.
57. Horny, M. (2014). *Bayesian networks. Technical report no. 5.* Boston University School of Public Health, Department of Health Policy and Management. Retrieved from <https://www.bu.edu/sph/files/2014/05/bayesian-networks-final.pdf>.
58. Giffoni, F., Salini, S., & Sirtori, E. (2018). Evaluating business support measures: The Bayesian Network approach. *Evaluation*, 24(2), 133–152. <https://doi.org/10.1177/1356389018767179>.
59. Castillo, E., Gutiérrez, J. M., & Hadi, A. S. (1997). Sensitivity analysis in discrete Bayesian networks. *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans*, 27(4), 412–423.
60. Kjørulff, U., & van der Gaag, L. (2000). Making sensitivity analysis computationally efficient. *Proceedings of the Sixteenth Annual Conference on Uncertainty in Artificial Intelligence (UAI 2000)* (pp. 317–325).
61. Mak, J., Abramsky, T., Sijapati, B., Kiss, L., & Zimmerman, C. (2017). What is the prevalence of and associations with forced labour experiences among male migrants from Dolakha, Nepal?

- Findings from a cross-sectional study of returnee migrants. *British Medical Journal Open*, 7(8), e015835.
62. Malit, F. T., Jr., & Naufal, G. (2016). Asymmetric information under the Kafala sponsorship system: Impacts on foreign domestic workers' income and employment status in the GCC countries. *International Migration*, 54(5), 76–90.
  63. International Labour Office. (2016). *Intertwined: An study of employers of migrant domestic workers in Lebanon*. Geneva: ILO.
  64. Kern, A., & Müller-Böker, U. (2015). The middle space of migration: A case study on brokerage and recruitment agencies in Nepal. *Geoforum*, 65, 158–169.
  65. International Trade Union Confederation. (2014). *ITUC Launches Migrant Worker "Recruitment Adviser" Platform*. Retrieved October 15, 2018, from <https://www.ituc-csi.org/ituc-launches-migrant-worker>.
  66. Labor Migration for Employment. *A Status Report For Nepal: 2015/2016 & 2016/2017*. Ministry of Labour and Employment, Government of Nepal, 2018. Retrieved from <https://asiafoundation.org/wp-content/uploads/2018/05/Nepal-Labor-Migration-status-report-2015-16-to-2016-17.pdf>.
  67. Bedford, R. (2008). United Nations Expert Group Meeting on International Migration and Development in Asia and the Pacific. United Nations Economic and Social Commission for Asia and the Pacific Population Division, Department of Economic and Social Affairs Bangkok, Thailand 20–21 September 2008.
  68. Emmel, N., Greenhalgh, J., Manzano, A., Monaghan, M., & Dalkin, S. (2018). *Doing realist research*. London: Sage.
  69. Bennett, L., Dahal, DR., Govindasamy, P. (2008). Caste, ethnic and regional identity in nepal: further analysis of the 2006 nepal demographic and health survey. Macro international inc.
  70. United Nations Development Fund (2008). The dalits of nepal and a new constitution: A resource on the situation of dalits in nepal, their demands and the implications for a new constitution. Kathmandu: United Nations Development Programme (UNDP).
  71. Bhattachan, KB., Sunar, TB., & Bhattachan, YK. (2009). Caste-based discrimination in nepal. New delhi: Indian institute of dalit studies.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.