 doi:10.20944/preprints201607.0007.v1

Article

Do the Flexible Employment Arrangements Increase Job Satisfaction and Employee Loyalty? Evidence from Bayesian Networks and Instrumental Variables

Eleftherios Giovanis

University of Verona, 37129 Verona, Italy; giovanis95@gmail.com; Tel.: +39-327-891-1154

Abstract: This study explores the relationship between job satisfaction, employee loyalty and two types of flexible employment arrangements; teleworking and flexi-time. The analysis relies on data derived by the Workplace Employee Relations Survey (WERS) in 2004 and 2011. A propensity score matching and least squares regressions are applied. Furthermore, Bayesian Networks (BN) and Directed Acyclic Graphs (DAGs) are employed in order to confirm the causality between employment types explored and the outcomes of interest. Finally, an instrumental variables (IV) approach based on the BN framework is proposed and applied in this study. The results support that there is a positive causal effect from these employment arrangements on job satisfaction and employee loyalty.

Keywords: Bayesian networks; directed acyclic graphs; employee loyalty; employment arrangements; flexi-time; job satisfaction; teleworking; workplace employment relations survey

JEL classifications: C11; C26; J28; J53

1. Introduction

Work especially the last years with the fast enhancement of technology and networks has been disconnected from a particular place and time and information technologies have made it possible for organizational workers to become untethered from their traditional office setting [1]. While the traditional place of work used to be the employer's premises, nowadays it is carried out in other locations, such as the employee's home or while traveling. Advances in technology reshape the relationship between work and home, where in some cases the traditional flow of employees from home to office is reversed [2]. The virtualization and this shock of the contemporary organisation has evolved as a vital necessity for the firms to be able to compete for workers globally and advances in information technology provide the means [3]. By 2016 it is estimated that around 90 million of self-employed and employed U.S. workers will work from home or from a remote location at least 2 to 3 days a week [4]. This study examines the relationship of teleworking or home-based working and flexi-time with job satisfaction and the employee loyalty in a sample of firms in Great Britain using the Workplace Employee Relations Survey (WERS) that took place waves in 2004 and 2011.

Teleworking is a term used to describe an alternative work arrangement that enables employees to work from anywhere other than the traditional work setting or employer's premises. Teleworking or telecommuting as it's sometimes called has gained increasing popularity and acceptance throughout the United States and the world [5]. According to the research study by Crandall and Gao [6] telework has become an international phenomenon. Thus, teleworkers spend some portion of their time away from the conventional workplace, working from home, and communicate by way of computer-based technology [7]. Previous studies have outlined the reasons for the growth of teleworking or other kinds of flexible type of employment, which are owned mainly to their perceived benefits. In particular, these benefits refer on both telework and employer including job satisfaction, increasing productivity, organizational loyalty, improved employee morale and

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2016 doi:10.20944/preprints201607.0007.v1

employer retention and saving in space office among others [8]-[9]. Secondly the relationship between flexible employment types, job satisfaction and employee loyalty is examined. This study adds to the previous literature by examining the above-mentioned linkages using Ordinary Least Squares (OLS) and Ordered Probit models based on propensity score matching, accounting for selection and heterogeneous bias. In addition, a Bayesian Network framework and Directed Acyclic Graphs (DAGs) representation are applied in order to examine and confirm the causal effect of flexible working employment arrangements on job satisfaction and employee's loyalty. The findings support a positive effect from the employment arrangements examined on both job satisfaction and employee loyalty. This can have possible policy implications not only to employees and firms, but to society overall, which are discussed later. It should be noticed that there are other flexible employment types, such as changing shifts, compressed hours meaning that the employee has the option to work the same weekly hours in 4 days instead of 5, the option of switching from part-time to full time and vice versa. However, the purpose of the paper is not to make an extensive analysis of all the possible flexible employment types, but to confirm whether there is a causal effect from the employment types to job satisfaction and employee's loyalty. Furthermore, as it has been mentioned before, the purpose of this study is to present an alternative way to investigate whether and which instrument can be used based on DAGs and BN framework, which can be otherwise difficult to be found.

The paper is organised as follows: Section 2 is devoted to the literature review on teleworking, job satisfaction and performance. Section 3 presents the methodology and the data used in the study. Section 4 considers the results and section 5 discusses the main findings of the study. Section 6 presents the concluding remarks and areas for future research.

2. Literature Review

In this section previous research studies on the association between teleworking and job satisfaction are briefly discussed. Organisation theorists have long recognized that any kind of interaction on the working environment can be an important determinant of job satisfaction. Sims et al. [10] suggest that jobs, offering opportunities feedback, friendship and interacting with other people can improve employee's job satisfaction. Previous studies note that face-to-face interaction is associated with positive outcomes [11]. Social interaction at work can facilitate social presence, foster mutuality and common ground and improve communication quality [12]-[13]. On the other hand, employees who face a small social presence at work and increased reliance on technology based job activities may experience lower levels of proper communication and less communication richness and quality [14]. Therefore, based on the previous researches teleworkers may report lower levels of job satisfaction owned to reduced frequency and quality of interaction with other people. Since this relationship is still unknown on a large scale study, this paper aims to examine the relationship between teleworking, job satisfaction and turnover intentions or employee loyalty.

On the other hand, recent research studies confirm a positive relationship between teleworking and job satisfaction [5], while other studies have found a curvilinear association, where increases of the teleworking hours increase the employee's job satisfaction up to a point, after which the effects slightly fade out [15]. Thus, the traditionally belief that the face-to-face interactions at work have positive effects on job satisfaction may be overestimated and over-generalised. Fonner and Roloff [16] using a sample of 89 teleworkers and 103 office-based employees applied a path analysis in order to test the adequacy of their mediation model and to examine the relationship between teleworking and job satisfaction. Additionally, they examined the indirect paths from telework to job satisfaction through work-life conflict, information stress exchange frequency and quality, stress from meetings and interruptions, general politics, and get ahead politics. Their results support that teleworking directly affects job satisfaction positively. This study adds to the previous research by examining the relationship between teleworking, flexi-time, job satisfaction and employee loyalty and it is compared with other flexible working arrangements. Previous studies explored the effects of precarious-compressed employment schemes for causal workers [17]-[18]. The studies found that causal workers report less flexible experience when it comes to daily working routines and are

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2016 doi:10.20944/preprints201607.0007.v1

more likely to report more instances of distress and social instability that can be traced back to job insecurity. However, this study explores two flexible employment schemes for permanent staff and not for causal workers. The difference is that amongst the firms explored, the employees have the option to choose one or more of the flexible working schedules examined.

3. Materials and Methods

3.1. The Estimated Model

The following job satisfaction function for individual *i*, in firm *k*, area-region *j* at time *t.*

$$JS_{i,j,k,t} = a_0 + a_1 WA_{i,j,k.t} + \alpha' z_{i,j,k,t} + \theta_t + l_j + l_j T + \varepsilon_{i,j,k,t} \tag{1}$$

$JS_{i,j,k,t}$ denotes the job satisfaction and the vector $WA_{i,j,k,t}$ is a dummy indicating whether or not the respondent is involved in the current type of working arrangement or not in firm *k,* in region *j* and in time *t*. Vector *z* includes individual and firm characteristics, such as age, education level, marital status, skills matching the job, the quality of relations between the managers and employees, whether the employees receive profit-related payments, whether the employees' payment is linked to the outcome of the performance evaluation, whether there are more than one establishments of this workplace in the UK, the firm type, such as public, private, charity and local government among others, whether the respondent-employee supervise other employees, whether the place-location of the product or service of the workplace is local, regional, national or international, and whether this workplace faces competition from other over-seas companies. Finally, the regression controls for standard travel to work areas (TTWA). Set l_j is the area fixed effects, which is expressed by the TTWA, θ_t is a time-specific vector, while l_jT is a set of the area-specific linear time trends which controls for unobservable, time-varying characteristics in the TTWA. Finally, $\varepsilon_{i,j,k,,t,}$ express the error terms which it is assumed to be *iid*. Standard errors are clustered at the TTWA level and job satisfaction function (1) is estimated using WERS which was conducted in 2004 and 2011 and has information about employee, employer and firm characteristics.

The working arrangements explored in this study are the following: flexi-time which means that there is no fixed or specific start or end of the job and the second employment type refers to employees who are home-based workers or teleworkers and work some days of the week at home.

Then relation (1) is estimated by replacing the dependent variable with the ordered variable "employee" loyalty and it is estimated with OLS and ordered Probit. It should be noticed that the propensity score matching [19] applied has been based on various algorithms, such as kernel and Mahalanobis, however the results remain the same. The final estimates are based on Mahalanobis algorithm.

3.2. Bayesian Networks and Instrumental Variables

Probabilistic models based on directed acyclic graphs (DAGs) have a long and rich tradition, which began with the geneticist Sewall Wright (1921). Variants have appeared in many fields, including artificial intelligence (AI) and cognitive sciences. The capability of inferring bidirectional relationships within a rigorous probabilistic foundation led to the rapid emergence and development of Bayesian networks as the choice method for uncertain reasoning in AI. The Bayesian Networks have been motivated and developed based on the conditional probability. Bayesian Networks rely on Bayes' theorem of probability theory to propagate information between nodes. As it is well known Bayes' theorem describes how prior knowledge about hypothesis A is updated by observed evidence B. The theorem relates the conditional and marginal probabilities of A and B as follows

$$P(A \mid B) = \frac{P(A) \cdot P(B \mid A)}{\int P(A) \cdot P(B \mid A) \cdot dB} \tag{2}$$

doi:10.20944/preprints201607.0007.v1

P(A) is the prior probability of the hypothesis or the likelihood that A will be in a particular state, prior to consideration of any evidence, P(B|A) is the conditional probability or the likelihood of the evidence, given the hypothesis to be tested and P(A|B) is the posterior probability of the hypothesis or the likelihood that A is in a particular state, conditional on the evidence provided. The integral in (2) represents the likelihood that the evidence will be observed, given a probability distribution. The presentation in the form of probabilities gives an explicit representation of uncertainty [20]. So far only simple problems with one or few variables have been considered. However, in real application learning problems, the main interest is the exploration of relationships among a large number of variables. Bayesian network is representation tool suited for this task, which encodes the joint probability distribution, physical or Bayesian, for a large set of variables.

Next the directed acyclic graphs (DAGs) and the Bayesian Network (BN) used in this study for causal inference are described. The graphical structure $G= (\mathbf{V}, \mathbf{E})$ of a BN is a directed acyclic graph DAG where $\boldsymbol{V}$ denotes the vertex or node set and $\mathbf{E}$ represents the edge set as $\boldsymbol{V_i} \rightarrow \boldsymbol{V_j}$ The notation $\mathbf{P}_i^G$ is used to denote the parent set of V_i in G. $\mathbf{p}_i^j$ is used to denote the *j-th* configuration of the parents of $\boldsymbol{V_i}$: $\mathbf{P}_i \in \{\mathbf{p}_i^1 \ \ldots, \mathbf{p}_i^{q_i}\}$. Based on that the definition of BN is:

Definition 1 (Bayesian network) (Pearl, 2000; Neapolitan, 2003): A Bayesian network model *M* over a set of variables $\mathbf{V} = \{X_1, \ldots, X_N\}$ is a pair $(\Gamma, \boldsymbol{\theta})$, where $G(\mathbf{V})$ is a DAG over $\mathbf{V}$ and $\boldsymbol{\theta}$ is a set of conditional probabilities: $\theta = \{\theta_{ijk} \colon \forall\ (ijk)\}$ such that $(\theta_{ijk} = X_i = x_i^k \mid P_i = p_i^j)$.

A BN is a graphical structural model that encodes probabilistic relationships among the variables of interest [21][1]. A graph *G*(*V*,*E*) can be referred to as a directed acyclic graph (DAG), when the edges *E* linking nodes-set of variables *V* are directed and acyclic. *Directed* means that edges *E* represent direct causal effects, while *acyclic* means that the directed edges do not form circles [23]-[25]. Following Heckerman's [21] notation, a generic graph is presented in Figure 1. The arrow between *T* and *F* in figure 1 means that *T* may have a direct causal effect on *F*. Similarly, for the arrow between *B* and *T* or *A* and *C* or *B* and *C*. In the case where there are *missing arrows*, it is implied that the strong assumption of no direct causal effect between two variables is rejected, which is so-called "strong null" hypothesis of no effect. All variables directly or indirectly caused by a given variable are called its *descendants*. The descendants of *T* are *F* and *Y*, while the *descendants* of *B* are *C, D, T* (B's children), *E* (D's and T's child), *F* (T's child) and *Y* (child of *A, C, D, E, F*) and similarly for the remained nodes-variables. On the other hand, *parents* are the variables that direct cause another variable. In figure 1 the only parent of *F* is *T*, while the only parent of *T* is *B*. A similar definition to *descendants*, working on the opposite way, is the variables that directly and indirectly cause of another variable and are called *ancestors.* For example the *ancestors* of *F* are *T* and *B*, while the *ancestors* of *E* are *B*, *D*, and *T*. *Paths* are sequences of adjacent arrows that traverse any given variable at most once. The arrows along a path may point in any direction. For example if *B* is the treatment and *F* is the outcome *then B→T→F* is the only causal path.

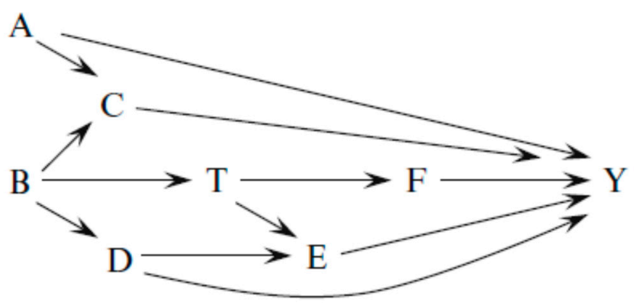


Figure 1. An example of a Directed Acyclic Graph (DAG).

[1] Major advances have been made in inferring causal relationships from observational data [22]-[23]

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2016 doi:10.20944/preprints201607.0007.v1

The DAG defines a factorization of the joint probability distribution of V = {X_1, ... ,X_N }, often called the global probability distribution, into a set of local probability distributions, one for each variable. The form of the factorization is given by the Markov property of Bayesian networks which states that every random variable X_i directly depends only on its parents:

$$p(x) = \prod_{i=1}^{m} p(x_i \mid par_i) \tag{3}$$

Applying the chain rule of probability, we have:

$$p(x) = \prod_{i=1}^{m} p(x_i \mid x_1, ..., x_{i-1}) \tag{4}$$

The causal Markov assumption is that each node is independent of its non-descendants in the graph conditional on its parents in the graph.

Definition 2 (One-step ahead conditional independence non-causality). *X* does not strongly cause *Y* one-step ahead given a set of covariates *Z* and *Y* does not cause *X* given a set of covariates *K* if (5)-(6) hold.

$$Y_{i,t} \perp X_{i,t-1} \mid Z_{i,t} \tag{5}$$

$$X_{i,t} \perp Y_{i,t-1} \mid K_{i,t} \tag{6}$$

$Y_{i,t} \subseteq Z_{i,t} \subseteq \Omega_{i,t}$ and $X_{i,t} \subseteq K_{i,t} \subseteq \Omega_{i,t}$. and Ω is the set of all covariates included in sets *K* and *Z*, for individual *i* and time *t*. The symbol $\perp$ is used to express independence.

Definition 3 (Conditional independence non-causality). The conditional independence *X* is conditional independent from *Y* in the edge set *E* iff $Y_{i,t} \perp X_{i,t} \mid \Omega_{i,t}$

The independence assumptions discussed above and are represented by the graph imply that parameters need to be estimated because the probability distribution for each variable depends only on the node's parents as it is shown in relations (3)-(4). Using the factorisation equation (4) it allows the network factorisation in such a way that it considers each node and its parents in isolation from the rest of the model variables. Otherwise, without employing this factorisation, far more parameters would be required to be estimated and therefore to specify the causal-effect relationships by a fully connected network and "unfactorable" model. Thus, employing factorisation model (4) the very complex models can be estimated avoiding the combinatorial explosion problem. In Figure 1 the Markov condition for *F* to *B* entails the following conditional independence relation:

$$F \perp B \mid T \tag{7}$$

More specifically, (7) implies that nodes *F* and *B* are independent given *T* as there is no direct edge connecting them. A similar interpretation can be derived for the remained nodes. In appendix some examples of DAGs, as well as the importance of the factorisation relation (4) are presented. More specifically, (6) implies that nodes *F* and *B* are independent as there is no direct edge connecting them and given *T*. A similar interpretation can be derived for the remained nodes.

The causal Markov assumption is the central assumption that defines BN. According to this assumption, each node is independent of its non-descendants in the graph, conditional on its parents in the graph. In other words, given a node's immediate cause, we can disregard the causes of its ancestors.

Lemma 1: Suppose Γ is a Bayesian network, and *Y* is a leaf node, where a leaf node is defined as the node that has no children. Let Γ' be the Bayesian network obtained from Γ by removing *Y*. Let $\mathbf{\Omega}$ be the set of all nodes in Γ. Then it will be:

 doi:10.20944/preprints201607.0007.v1

$$P_{\Gamma}(\mathbf{\Omega}) = P_{\Gamma'}(\mathbf{\Omega}) \tag{8}$$

Proof: (see appendix)

So far the discussion introduced one relationship between probability distributions and DAGs, namely the Markov condition. Causal Markov only produces a set of independence relations from a causal graph, but says nothing about whether there are additional independence relations. The faithfulness assumption states that a BN graph *G* and a probability distribution *P* are faithful to one another if every one and all independence relations valid in *P* are those entailed by the Markov assumption on *G* and the factorisation in (4). Faithfulness assumption additionally indicates that these independence relations derived from Causal Markov is the exact set of independence relations.

Definition 4 (Faithfulness). ([23], [26]): Suppose we have a joint probability distribution P of the random variables in some set *V* and a DAG *G*=(*V*,*E*). We say that (*G*,*P*) satisfies the faithfulness condition if based on the Markov condition, *G* entails all and only conditional independencies in *P*. That is, the following two conditions hold:

- (*G*,*P*) satisfies the Markov condition.
- All conditional independencies in *P* are entailed by *G*, based on the Markov condition.

When (*G*,*P*) satisfies the faithfulness condition, we say *P* and *G* are faithful to each other, and that *G* is a perfect map of *P*.

Next one very important definition for the DAG and BN, which is the *d-separation* is discussed. In addition, in appendix some examples are presented showing the importance of this condition.

Definition 5 (d-separation). ([23], [26]):: Let $G = (V,E)$ be a DAG, $A \subseteq V$, X and Y be distinct nodes in $V—A$, and h be a chain between X and Y. Then h is blocked if one of the following cases holds:
There is a node $S \in A$ on the chain h and the edges incident to S on h meet head-to-tail at S.
There is a node $S \in A$ on the chain h and the edges incident to S on h meet tail-to-tail at S.
There is a node S such that S and all of S's descendants are not in A on the chain h and the edges incident to A on h met head-to-head at S.

The *d*-separation condition is especially important and useful in constructing a BN because it controls for possible confounds as in the form of *S* described here. In other words, a set of variables *S*, *d*-separate variable *X* from *Y*, if and only if *S* blocks every path from *X* to *Y*. Graphically, *d*-separation usually exhibits two main cases: firstly $X \rightarrow S \rightarrow Y$ and secondly $X \leftarrow S \rightarrow Y$. The intuition behind this graphical representation is that *X* and *Y* are independent from each other conditioned on *S*. In the first case *X* causes *Y* through *S*, while in the second case *X* and *Y* have a common cause *S*. To ascertain whether a particular conditional independence statement $X \perp Y | S$ is implied the possible paths from any node in *X* to any node in *Y* are considered. Any such path is blocked if it includes a node such that either the arrows on the path meet either head-to-tail or tail-to-tail at the node, and the node is in *S*, such as the relations $X \rightarrow S \rightarrow Y$ and $X \leftarrow S \rightarrow Y$ or the arrows meet head-to-head at the node, and neither the node, nor any of its descendants, is in *S*. If all paths are blocked, *X* is d-separated from *Y* given *S*, and the joint distribution over all of the variables in the graph will then satisfy $X \perp Y | S$.

Definition 6. (Partial Correlation): For $i \neq j \in 1, \ldots, p$, $k \in X_r$, let $\rho_{i,j|k}$ be the partial correlation between X_i and X_j given X_r and X_r denotes the rest of the variables.

Based on this definition we have that $X_i \perp\!\!\!\perp X_j | X_r \Leftrightarrow \rho_{i,j|k}$. Next the Fisher's Z test for the conditional independence is presented [23]-[27]:

$$\rho_{XY|C} = 0 \tag{9}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2016 doi:10.20944/preprints201607.0007.v1

$$z(\rho_{XY|C^{,n}}) = \frac{1}{2}\sqrt{n-|C|-3}\log\frac{(|1+\rho_{XY|C}|)}{(|1-\rho_{XY|C}|)} \tag{10}$$

|C| is the number of variables in C and n is the length of the sample. If X,Y,C~N under the null hypothesis of zero partial correlation:

$$z(\hat{\rho}_{XY|C^{,n}}) \sim N(0,1) \tag{11}$$

The test for independence is based on the PC algorithm [23] at significance level α. Kalisch and Buhlmann [27] show that the choice of α is not too important. However, a significance level α=0.05 is used. The pseudo-code of the PC algorithm is presented in figure 2.

Step 1:

Start with the complete undirected graph, $C^{\sim}$ with vertices $V = X_1, \ldots, X_p$. Then:

Step 2:

Set $l = -1$ and $C = C^{\sim}$

Step 3:

Increase l by one. For all pairs of adjacent nodes:

- Check for conditional independence
- Remove edge (X_i, X_j) if $X_i \perp\!\!\!\perp X_j|rest$

Step 4:

Repeat step 2 until $l = m$ or until each node has fewer than $l - 1$ neighbours

And let mr each $\in$ max l, m denote the stopping level of the algorithm and q be the maximum number of neighbours

In plain words the above pseuso-code of the PC algorithm works on the following simple steps.

- For each X and Y, see if X ⊥Y; if so, remove their edge.
- For each X and Y which are still connected, and each third variable Z1, see if X ⊥ Y|Z; if so, remove the edge between X and Y.
- For each X and Y which are still connected, and each third and fourth variables
- Z1 and Z2, see if X ⊥ Y|Z1,Z2; if so, remove their edge.

For each X and Y which are still connected, see if X ⊥ Y| all the p − 2 other variables; if so, remove, their edge

Figure 2. PC algorithm pseudo-code for the estimated DAG.

Next this section discusses the possibilities of using the DAG and BN as a tool for discovering candidate instrumental variables. A DAG representation where an instrumental variable *I* can meet the conditions and which can be used into the analysis is presented in figure 3, where the instrumental variable *I* is related to the cause of interest *X* and influences *Y* only through its impact on *X* and at least one control variable blocks the other path, such the variable-node *S*, which can be also a set of variables **S**. In the case of figure 3 in order to use variable *I* as an instrument, we should condition on *S* but not on *W* because it is descendant of *Y* and the common cause of *W* coming from *I* and *Y* will lead to selection bias as it has been described in the methodology part and the *d-separation* condition. More specifically, a variable *I* qualifies as an IV for *X* (factor of interest) and *Y* (outcome of interest if the following three conditions are met: a) *I* is statistically independent of all joint common causes of *X* and *Y*; b) *I* is not independent of *X*; and c) the effect of *S* on *Y* is mediated solely by *X*.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2016 doi:10.20944/preprints201607.0007.v1

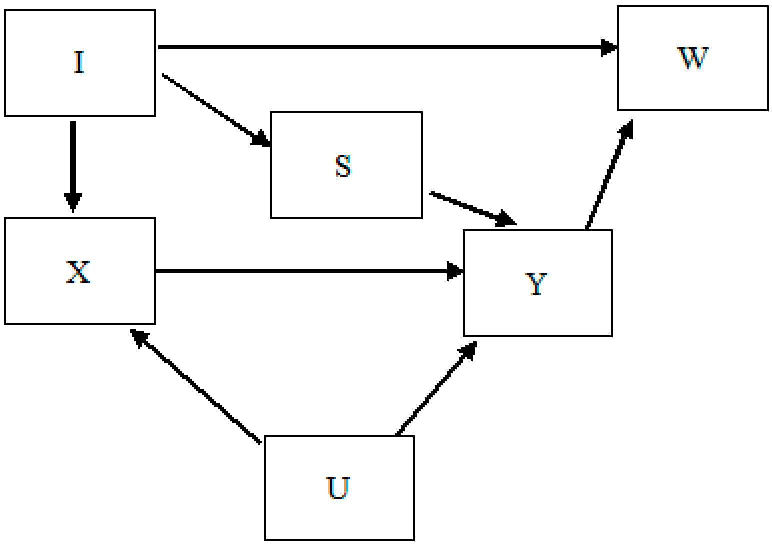


Figure 3. Illustration of IV conditions

Next the second condition of choosing an instrumental variable is presented in figure 4. Based on that we have the following Lemma.

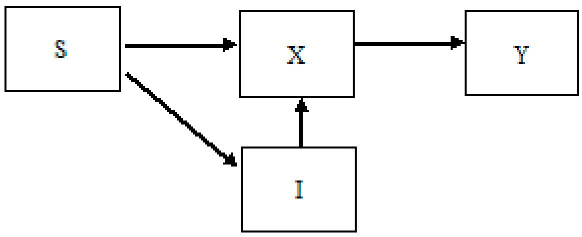


Figure 4. IV conditions given a Set of **S**.

Lemma 2. *Given a path diagram G and which contains the direct edge X→Y, then a variable I can be an instrumental variable for X given S, which is a set of variables that does not contain any variable from I, X, Y, or Desc(Y) if the following conditions are met:*

- In the path diagram G, X and I are connected given S
- In the path diagram $G_{\backslash X \to Y}$ which is formed by removing X→Y from G, I and Y are d-separated given S.

Then the direct causal effect τ_{yx} is given by:

$$\tau_{yx} = \frac{\sigma_{IY \cdot S}}{\sigma_{IX \cdot S}} \tag{12}$$

Lemma 2 can be extended into a set of instrumental variables I_i for i=1,….,k. In that case the direct causal effects $\tau_{yx_1}, \ldots\ldots, \tau_{yx_k}$ can be solved by a system of k equations as:

$$\sigma_{I_i Y \cdot S} = \sum_{j=1}^{k} \sigma_{I_i X_j \cdot S} \tau_{YX_j} \tag{13}$$

for i=1,…k

 doi:10.20944/preprints201607.0007.v1

For instance when $X \perp Y$ then it is Cov(X,Y)=0. In relation (11), $\sigma_{IY.Z}$ is the conditional covariance between I and Y given S=s, while similarly $\sigma_{IX.Z}$ is defined as the conditional covariance between I and X given S=s. Considering the Cov(I,Y|S) it will be:

$$\begin{aligned} &Cov(I,Y \mid S) = E[I,Y \mid S] - E[I \mid S]E[Y \mid S] = \\ &E[(I,Y - E[I \mid S]Y - IE[Y \mid S] + E[I \mid S]E[I \mid S]) \mid S] = \\ &E[I,Y \mid S] - E[I \mid S]E[Y \mid S] - E[I \mid S][Y \mid S] + E[I \mid S]E[Y \mid S] = \\ &E[I,Y \mid S] - E[I \mid S]E[Y \mid S] \end{aligned} \tag{14}$$

In a similar fashion the Cov(I,X|Z) we will have:

$$E[I,X \mid S] - E[I \mid S]E[X \mid S] \tag{15}$$

Thus, this test can be applied in other studies using IV approach and to define whether the IV is proper or not, especially regarding the selection bias. Plugging (14)-(15) in (12) it will be:

$$\tau_{yx} = \frac{Cov[I,Y \mid S]}{Cov[I,X \mid S]} \tag{16}$$

And equivalently becomes:

$$\tau_{yx} = \frac{E[I,Y \mid S] - E[I \mid S]E[Y \mid S]}{E[I,X \mid S] - E[I \mid S]E[X \mid S]} \tag{17}$$

3.3. Data

The Workplace Employment Relations Study (WERS) series commenced in 1980 and took place six times until 2011. The 2004 and 2011 Panel Survey was conducted in a random sub-sample of workplaces and the surveys are conducted to managers and employees. This is useful for the analysis, since the regressions control not only for employee characteristics, but also for firm characteristics, such as competition and market area.

The first outcome of interest is the job satisfaction, which is an ordered variable measured in a Likert scale from 1 (very dissatisfied) to 5 (very satisfied). Similarly, employee loyalty is an ordered variable answering to the question "To what extent do you agree or disagree with the statement that you are loyal to your firm-organisation" measured in a scale from 1 (strongly disagree) to 5 (strongly agree). In table 1 the summary statistics for the outcomes of interest-job satisfaction and employee retention- and the factors of interest which are the employment arrangements explored in this study are presented. The average job satisfaction and employee loyalty levels are relatively high to our sample with average very close to 4. Regarding the employment arrangements explored in this study, the lowest percentage of participation belongs to teleworking.

Additional, variables are not presented, as the descriptive statistics do not give any additional insights, for this reason a correlation matrix among the outcomes of interest and the employment arrangements is presented in table 2. In all cases there is a positive relationship among the two types of employment, as well as, a positive association between employee loyalty, job satisfaction and employment types. Additional factors are not explored as they are used as controls in to the regressions; however, some correlation statistics show for instance that wage and higher education degree are positive associated with flexi-time and teleworking.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2016 doi:10.20944/preprints201607.0007.v1

Table 1. Summary Statistics.

	Mean	Standard deviation	Minimum	Maximum
		Panel A: Dependent Variables		
Job Satisfaction	3.8134	0.8939	1	5
		Panel B: Employment Arrangements		
Flexi-time (Yes)	39.08	Teleworking (Yes)	15.87	
Flexi-time (No)	60.92	Teleworking (No)	84.13	

Table 2. Correlation Matrix.

	Job satisfaction	Loyalty	Flexi-time
Loyalty	0.4637*** (0.0000)		
Flexi-time	0.1322*** (0.0000)	0.1473*** (0.0000)	
Teleworking	0.1187*** (0.0000)	0.1368*** (0.0000)	0.2763*** (0.0000)

P-values in parentheses, *** p<0.01

4. Results

The OLS estimates after the propensity score matching are presented in table 3. The results show that the relationship between job satisfaction and the employment arrangements explored is positive and significant. The highest magnitudes are presented in the case of the flexi-time, followed by the teleworkers. The remained coefficients show that elder workers are more satisfied with their job, while the married or couples and divorced are more likely to report higher levels of job satisfaction than the singles. Regarding education and the level of skills to job matching, those with first degree in some cases report lower levels of job satisfaction, which is consistent with other studies, arguing that more educated people have higher expectations about their pecuniary and non-pecuniary returns form their job and thus are more easily disappointed and dissatisfied [28]-[30]. The same applies with the matching skills to job. More precisely, those who stated that their skills match the job almost the same or bit lower report higher levels of job satisfaction than those who stated that are over-qualified. The competition, and the status of company are insignificant factors, while on the other hand, performance related payments schemes and supervising other employees increase the job satisfaction levels, as well as those who state that the quality of relations between employees and managers is either good or very good, relatively to those who reported low levels of manager-employee relations. Wage presents mixed results among the employment arrangements examined. More specifically, the wage is insignificant in the low scales, while it becomes significant in the high salary scales. Similarly in table 4 the relationship between employee loyalty and the employment arrangements is significant. The remained controls are not reported as the conclusions are similar.

Furthermore, the variables of whether the respondent has at least one dependent child aged 0-2 years old or older and the percentage of the employees using computer in the firm are included. In all cases the coefficients are insignificant, suggesting that these variables can be used as instruments since are not correlated with the outcome. However, other variables are not also significant, but they might be necessary to be employed as controls-or parents-using the terminology of BN. More specifically, relation (13) implies that the IV approach should consider certain variables that can be used as controls, as it will be shown later in this part.

In table 5 the Ordered Probit estimates for the two employment arrangements are reported. The positive and significant coefficient of flexi-time and teleworking is presented; however the magnitude is higher since the Ordered Probit models follows a different empirical estimation procedure than OLS.

Table 3. Propensity Score Matching and OLS for Job Satisfaction

VARIABLES	(1)	(2)	VARIABLES	(1)	(2)
Flexi-time	0.1564***		Depend. children 0-2 (Yes)	0.0067	0.0002
	(0.0221)			(0.0257)	(0.0259)
Home-Teleworking		0.1437***	Skills matching job (much higher)		
		(0.0276)			
Performance appraisal (Yes)	0.0226	0.0420*	Skills matching with job-bit higher	0.2210	0.1465
	(0.0168)	(0.0220)		(0.1716)	(0.1608)
Wage (reference £141-£180 per week			Skills matching with job-the same	0.4142**	0.4040***
Wage - £181-£220 per week	0.0372	-0.0293		(0.1634)	(0.1531)
	(0.0839)	(0.0771)	Skills matching with job-bit lower	0.3702**	0.3551**
Wage - £221-£260 per week	0.1575*	0.0510		(0.1637)	(0.1534)
	(0.0816)	(0.0763)	Skills matching with job-much lower	0.1597	0.1167
Wage - £261-£310 per week	0.1253*	0.0444		(0.1645)	(0.1544)
	(0.0689)	(0.0742)	Quality of relations (very poor)		
Wage - £431-£540 per week	0.1349*	0.0288			
	(0.0782)	(0.0745)	Quality of relations-Poor	0.3935***	0.4278***
Wage - £681-£870 per week	0.1491*	0.1155		(0.0801)	(0.0801)
	(0.0803)	(0.0764)	Quality of relations-Neither good nor bad	0.6272***	0.7114***
				(0.0756)	(0.0754)
Wage - £871 or more per week	0.2329**	0.2021**			
	(0.0932)	(0.0882)	Quality of relations-Good	0.9503***	1.0009***
Gender (Female)	0.0989***	0.0978***		(0.0744)	(0.0742)
	(0.0278)	(0.0258)	Quality of relations-Very Good	1.2973***	1.3624***
Age (reference category 16-17)			Number of establishments (reference many)	(0.0771)	(0.0765)
Age (18-19)	-0.2096*	-0.2621**	single	0.0194	0.0253
	(0.1188)	(0.1032)		(0.0305)	(0.0281)
Age (20-21)	-0.1433	-0.2173**	Sole in UK-foreign	-0.0710	-0.1276*
	(0.1180)	(0.1007)		(0.0618)	(0.0640)
Age (22-29)	-0.0114	-0.1221	Supervise other employees (No)	-0.1157***	-0.0932***
	(0.1044)	(0.0897)		(0.0258)	(0.0243)
Age (30-39)	0.0508	-0.0653	Market Area (reference-Local)		
	(0.1051)	(0.0905)	Market Area-Regional	0.0816**	0.0561*
Age (40-49)	0.0542	-0.0606		(0.0368)	(0.0339)
	(0.1069)	(0.0923)	Market Area-National	0.0398	-0.0105
Age (50-59)	0.1086	-0.0036		(0.0311)	(0.0287)
	(0.1083)	(0.0937)	Market Area-International	0.0653*	0.0698**
Age (60-64)	0.3996***	0.1965*		(0.0386)	(0.0351)
	(0.1143)	(0.1010)			
Age (65 and above)	0.3262**	0.1597	Performance related payments (Yes)	0.0331**	0.0535**
	(0.1343)	(0.1155)		(0.0140)	(0.0221)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2016 doi:10.20944/preprints201607.0007.v1

Table 3 (cont.) Propensity Score Matching and OLS for Job Satisfaction.

VARIABLES	(1)	(2)	VARIABLES	(1)	(2)
Marital status (reference category single)			Body established by Royal Charter	0.0892	0.1114
				(0.0724)	(0.0685)
Marital status-Married or couple	0.0432	0.1232*	Competition from abroad	-0.0172	-0.0485
	(0.0819)	(0.0731)		(0.0783)	(0.0758)
Marital status-Divorced	0.1395***	0.0939**	Government-owned limited	-0.1238	-0.1120*
	(0.0451)	(0.0432)		(0.0815)	(0.0672)
Marital status-Widowed	0.0359	0.0270	Computer Use		
	(0.0299)	(0.0277)			
Education level (reference primary school)			Member of Union Trade (reference Yes)		
Education level-GCSE D-E levels	-0.0211	-0.0417*	Member of union (No, but in the past)	0.0527	0.0262
	(0.0242)	(0.0224)		(0.0331)	(0.0231)
Education level-GCSE B-S levels	-0.0456	-0.0314	Member of union (No never)	0.0835**	0.0934***
	(0.0320)	(0.0299)		(0.0302)	(0.0217)
Education level-GCSE A-AS levels	-0.0452	-0.0670***	Completion from abroad (A lot)		
	(0.0287)	(0.0259)	Completion from abroad-Little	-0.0214	0.0122
Education level-First degree	-0.0600**	-0.0762***		(0.0512)	(0.0376)
	(0.0305)	(0.0280)	Completion from abroad-No	-0.0338	0.0102
Education level-Higher degree	-0.0055	0.0117		(0.0458)	(0.0299)
	(0.0510)	(0.0459)	Computer Use	-0.00085	-0.00088
				(0.0036)	(0.0037)
Type of Firm (reference Public)			Dependent children 0-2 years old	-0.0105	0.0052
				(0.0391)	(0.0448)
Private limited company	0.0114	-0.0376	Dependent children >2	0.0025	0.0016
	(0.0281)	(0.0258)		(0.0260)	(0.0237)
Company limited by guarantee	0.0748	0.0068	No. Observations	7,691	7,503
	(0.0633)	(0.0561)	R-Square	0.2777	0.2659
Partnership	0.0475	0.0107			
	(0.0515)	(0.0486)			
Trust/Charity	0.0456	-0.0317			
	(0.0512)	(0.0490)			

The estimates for flexi-time and teleworking are respectively reported in columns (1) and (2). Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 4. Propensity Score Matching and OLS for Employee Loyalty.

Variables	**(1)**	**(2)**
Flexi-time	0.1428***	
	(0.0111)	
Teleworking		0.1158***
		(0.0136)
Observations	7,140	7,368
R-squared	0.3232	0.3087

Robust standard errors in parentheses, *** p<0.01

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2016 doi:10.20944/preprints201607.0007.v1

Table 5. Propensity Score Matching and Ordered Probit.

Variables	**(1) DV: Job Satisfaction**	**(2) DV: Job Satisfaction**	**(3) DV: Employee Loyalty**	**(4) DV: Employee Loyalty**
Flexi-time	0.2085***		0.1913***	
	(0.0318)		(0.0162)	
Teleworking		0.2008***		0.1830***
		(0.0406)		(0.0403)
Observations	7,291	7,503	7,140	7,368
Pseudo R-Square	0.1323	0.1243	0.1493	0.1439
Wald chi square	4,578.65	4,725.33	4,326.97	4561.24
	[0.000]	[0.000]	[0.000]	[0.000]

Robust standard errors in parentheses, *P*-values within brackets *** p<0.01; Columns (1)-(2) are estimated respectively for flexi-time and teleworking and the dependent variable is the job satisfaction. Similarly, in columns (3)-(4) the respective estimates using employee loyalty as the dependent variable are presented.

In figure 5 the estimated DAG for teleworking is presented while the BN estimates considering both employment arrangements explored in the study, are reported in table 6. A similar representation is observed for the flexi-time, but its associated DAGs is not presented here. Also separate estimates for each employment arrangements is taking place since it is difficult to disentangle their effects, when they are included into the same regression. The reason is that regression presents over-control bias where some variables block the causal effect from the variable of interest to the outcome. For instance coming back to figure 1, F blocks-off the causal effect from T to Y since there is no direct effect (arrow) from the former to the latter. Similarly, in this case teleworking may block –off the causal effect from flexi-time to job satisfaction. Thus, one solution is to not include them in to the same DAG, while the second solution is to incorporate them into the same DAG and BN and applying the factorization relation (4) wherever necessarily.

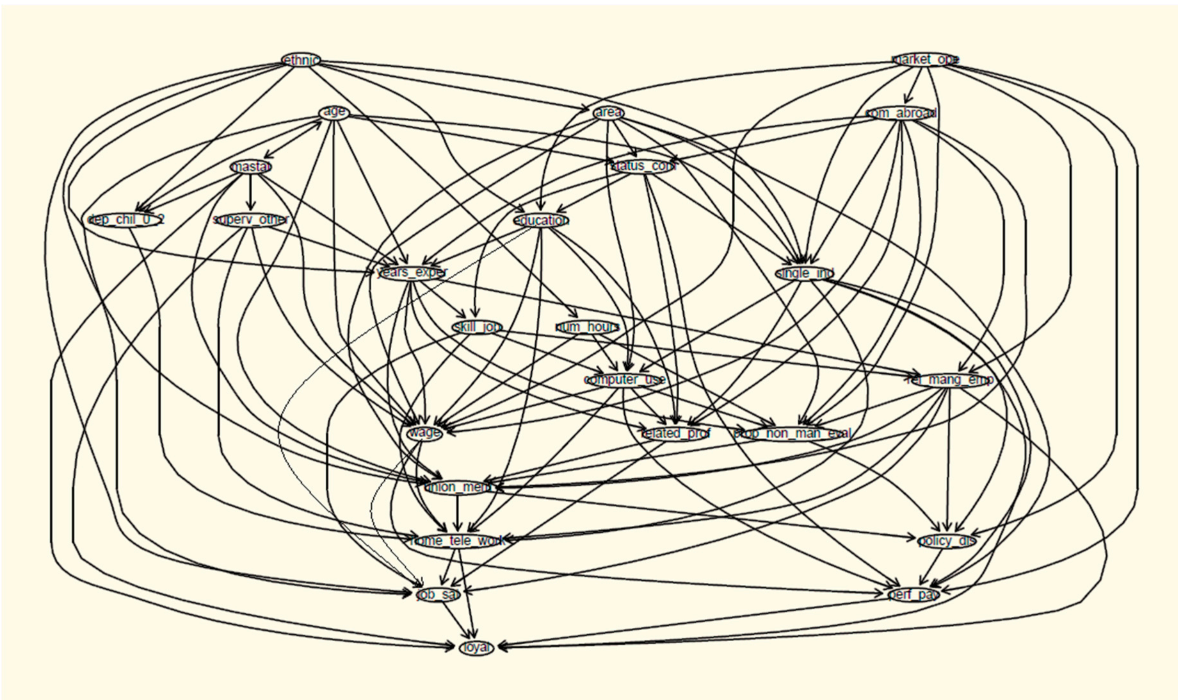


Figure 5. DAG for Teleworking.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2016 doi:10.20944/preprints201607.0007.v1

Table 6. BN and DAG Estimates for Job Satisfaction, Employee Loyalty and Employee Arrangements.

VARIABLES	(1) OLS	(2) OLS	(3) OP	(4) OP
Panel A: Dependent Variable Job Satisfaction				
Flexi-time	0.1643***		0.2939***	
	(0.0193)		(0.0458)	
Teleworking		0.1671***		0.2725***
		(0.0335)		(0.0718)
Observations	7,291	7,503	7,291	7,503
Panel B: Dependent Variable Employee Loyalty				
Flexi-time	0.1523***		0.2744***	
	(0.0182)		(0.0468)	
Teleworking		0.1503***		0.2632***
		(0.0285)		(0.0718)
Job satisfaction (reference category-very dissatisfied)				
Job satisfaction-dissatisfied	0.4224***	0.3727***	0.8140***	0.7089***
	(0.0594)	(0.0629)	(0.1478)	(0.1556)
Job satisfaction-neither dissatisfied nor satisfied	0.6246***	0.5652***	1.1978***	1.0815***
	(0.0545)	(0.0580)	(0.1370)	(0.1447)
Job satisfaction- satisfied	1.0116***	0.9881***	2.1370***	2.0731***
	(0.0529)	(0.0565)	(0.1352)	(0.1432)
Job satisfaction-very satisfied	1.3999***	1.3654**	3.4082***	3.2731***
	(0.0571)	(0.0624)	(0.1483)	(0.1562)
Observations	7,140	7,368	7,140	7,368

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.01; Columns (1)-(2) are estimated with OLS respectively for flexi-time and teleworking. Similarly, in columns (3)-(4) the respective estimates using ordered Probit are reported. In Panel A the dependent variable is job satisfaction and in panel B the dependent variable is employee loyalty.

Before we proceed to the findings conclusion, the abbreviations of the variables in the DAG figure 5 are described. Variables *ethnic*, *mastat*, *superv_other*, *num_hours* and *age* indicate respectively ethnicity, marital status, supervising other employees, number of weekly hours worked and age. The other variables have as following: dependent children 0-2 years old (dep_chil_0_2), years worked (years_exp), education level (education), status of company (status_com), number of firm establishments (single_ind), whether the skills match to employee's work (skill_job), area market of the firm (market_ope), TTWA (area), quality of relations between managers and employees (rel_mang_emp), wage (wage), percentage of employees in the firm using computer (computer_use), whether the employee is member of union trade or staff association (union_memb), formal written policies for equal opportunities in the institution (policy_dis), performance pay schemes (perf_pay), related profit schemes (related_prof), proportion of non-managerial staff under performance evaluation (prop_non_man_eval) home-teleworkers (home_tele_work), job satisfaction (job_sat) employee loyalty (loyal).

Applying the factorization model (4) and the d-separation the causal effect of the teleworking on job satisfaction is a regression of itself and its parents-*computer_use*, *wage*, *education*, *union_memb*, *rel_mang_emp*, *skill_job*, *years_exp*, *single_ind* and *superv_other*. More specifically, in columns (1)-(2) of table 6 the OLS estimates for flexi-time and teleworking are respectively presented. Similarly, in columns (3)-(4) the respective ordered Probit (OP) results are reported. In Panel A the dependent variable is job satisfaction, while in Panel B the dependent variable is employee loyalty.

The causal effect of teleworking on job satisfaction and employee loyalty is higher than those found in tables 3-5. More specifically, according to BNs the coefficients for job satisfaction and employee loyalty are 0.1671 and 0.1503, while the respective coefficients with OLS are 0.1437 and 0.1158 lower by 15-23 per cent. Similarly, the effects derived by OP and OP based on DAGs are underestimated by 35-45 per cent. A similar DAG is estimated for the flexi-time; however is not

 doi:10.20944/preprints201607.0007.v1

presented here, but its causal effects are reported in table 6. More precisely, the effects on job satisfaction and employee loyalty are underestimated by 15-35 per cent.

Various other conclusions can be derived from DAG in figure 5. For instance the regression should not condition on employee loyalty since it is caused by both job satisfaction and teleworking, leading to selection bias. Similarly, if we would like to derive the causal effect of wage on job satisfaction, a regression including the wage and its parents will take place. Another example is computer use, where its causal effect is blocked-off by teleworking. In this case also the regression should include the *computer use* and its parents, in order to estimate the causal effect of computer percentage use on job satisfaction. However, two things are concluded. Firstly, conditioning on teleworking, the causal effect of co*mputer use* is blocked-off from teleworking leading to over-control bias, as it has been discussed in the methodology section and since there is no direct effect-arrow to job satisfaction. Thus, in this case the front-door and back-door criteria are applied.

Secondly, coming back to the figure 5 computer use can be used as an instrumental variable because is directly related to teleworking, and is conditioned on *related_prof* which the latter affects the job satisfaction. Thus, the IV should be conditioned on at least one other variable which causes job satisfaction. On the other hand, the *computer use* is conditioned on the performance payment schemes (perf_pay) and it can be used as an instrument variable in the case of the employee loyalty.

Third, following Lemma 2 *computer use* and *dependent child 0-2 years old* can be used as instrumental variables given a set S which is related to both the instrumental variable and the factor of interest which is the teleworking expressed as X in figure 4. The variables included in the set S are *skill_job, single_ind* and *education* when the outcome of interest is employee loyalty, while only the *single_ind* is included in the set S, when the dependent variable is the job satisfaction, since the remained variables are correlated with the outcome. Regarding, the *dependent child 0-2 years old,* the variables *age, marital status* and *ethnicity* are included in the set S when the outcome of interest is employee loyalty, while in the case of job satisfaction only the marital status is considered.

In this case, according to table 7 and the 2SLS estimates in column (2), the causal effect of teleworking on job satisfaction employing as instrument to teleworking the percentage of employees using computer in the institution-firm is 0.1751. For the flexi-time working arrangement, based on the BN the same instrumental variable is employed, as well as, whether there is dependent child 0-2 years old in the employee’s family. In all cases the estimates confirm the impact of these working schedules on job satisfaction and employee loyalty and it is 0.1627, which is higher than those found in the previous estimates. Similarly the effects of teleworking on job satisfaction and employee loyalty is found equal at 0.1988 and 0.1779 respectively which are 38 and 53 percent higher than the respective estimates found by OLS. In columns (3)-(4) the IV-DAG estimates using the Lemma 2 are reported. As it has been discussed in the methodology section the IV-DAG regressions are similar with the 2SLS with the difference that the factors of interest, which is flexi-time and teleworking, are conditioned on a specific set of variables **S**. This is more proper since the DAG can account for the three types of biases discussed previously. The estimates are close with the respective ones found by the 2SLS. The results suggest that a causal effect of the flexible employment arrangements explored in this study have a causal effect on employee’s job satisfaction and loyalty to the workplace with positive impact.

Overall, BN can be a very useful tool or empirical research allowing us to find proper instrumental variable wherever possible. However, the results confirm that IVs are not always necessarily into BN framework for causality. A similar application for the remained employment types can be considered. Another point that it should be noted is that in many studies the principal component analysis (PCA) is used in order to reduce the number of variables into the analysis. However, PCA is based on assumptions which are not always met. More specifically, the first assumption is that the dimensionality of data can be efficiently reduced by linear transformation, but this is not always met, since points of an input set are positioned on the surface of a hypersphere, no linear transformation can reduce dimension. The second drawback of PCA is the fact that directions maximizing variance do not always maximize information. BN is not limited from these assumption and it allows us to decide which variables should be included into the regressions analysis without

 doi:10.20944/preprints201607.0007.v1

losing information [31]-[32]. Moreover, the interpretation can be more difficult since we are not working with the original variables and the principal components are heavily affected by the scaling of the variables.

Table 7. 2SLS and IV-DAG Estimates for Job Satisfaction and Employee Loyalty.

VARIABLES	(1) 2SLS	(2) 2SLS	(3) IV-DAG	(4) IV-DAG
Panel A: Dependent Variable Job Satisfaction				
Flexi-time	0.1751***		0.1710***	
	(0.0323)		(0.0292)	
Teleworking		0.1988***		0.1853***
		(0.0213)		(0.0230)
Observations	7,032	7,352	7,084	7,395
R-squared	0.2776	0.3050	0.2662	0.2748
Weak instrument test	67.835	68.967	74.214	70.467
	[0.000]	[0.000]	[0.000]	[0.000]
Exogeneity test	4.643	0.5191	0.046	0.2446
	[0.2213]	[0.7715]	[0.9299]	[0.5743]
Panel B: Dependent Variable Employee Loyalty				
Flexi-time	0.1627***		0.1561**	
	(0.0345)		(0.0674)	
Teleworking		0.1779**		0.1720***
		(0.0862)		(0.0292)
Observations	6,940	7,178	7,011	7,226
R-squared	0.1818	0.2523	0.1737	0.2176
Weak instrument test	62.754	63.766	72.834	74.214
	[0.005]	[0.000]	[0.005]	[0.000]
Exogeneity test	4.515	3.219	1.946	2.616
	[0.2333]	[0.3536]	[0.1630]	[0.2754]

Robust standard errors in parentheses, p-values within brackets, *** p<0.01, ** p<0.05, * p<0.10; Columns (1)-(2) are estimated with OLS respectively for flexi-time and teleworking using 2SLS Similarly, in columns (3)-(4) the respective estimates using IV-DAG are reported. In Panel A the dependent variable is job satisfaction and in panel B the dependent variable is employee loyalty

5. Discussion

Since the job satisfaction and employee retention have a central role to firm's organization and policy, but also are topics of the policy makers' agenda for the improvement of the society's well-being, Bayesian networks can have important policy implications, as causal inference has a central role in well-being and policy making. These implications can be extended and applied in many other domains of well-being and public policy, including life satisfaction, leisure and public health and policy generally. Since the natural experiments are very difficult to be found and many times may not be under the researcher's control, while the instrumental variables are very difficult to be found and be convincible, Bayesian Networks is an alternative tool which can be useful, when the former cases are absent. BN can be applied not only to observation data, which data are very useful for controlling for various characteristics, which in the majority of the natural experiments are missed, but they can be applied to randomized experiments as well (Pearl, 2000, 2009; Spirtes et al., 2000). The methodology framework followed in this study suggests that BN can be a valuable instrument for deriving plausible causal effects using observational data, as well as, it provides a very useful graphical representation which allows us to consider the three types of bias discussed before; the over-control, confounding and selection bias. Moreover, BN and DAGs can be a valuable

 doi:10.20944/preprints201607.0007.v1

tool for testing and obtaining candidate variables as instruments to the factor of interest, which were the employment arrangements in this study. Furthermore, DAGs allows us to account for selection bias when a variable is chosen as instrument. For instance, it has been discussed that a variable which is a descendant of both factor of interest (employment arrangement in this case) and the outcome of interest (job satisfaction or employee loyalty) cannot be considered as instrument, since it is correlated or affected by both employment arrangements and the outcome.

6. Conclusions

The findings of this study suggest that there is a positive effect from the two employment arrangements examined, on job satisfaction and employee loyalty. This may indicate that these types of employment can allow the employees to use them as means of relief from stressful conditions, coming mainly from commuting at work and the traffic congestion. Moreover, these types of working arrangements, may give to employees more autonomy and control of the working schedule and to allow them to adjust it on their needs, including family demands and obligations and leisure activities. Furthermore, future research might take place on how these employment arrangements improve the labour productivity and the firm performance, through job satisfaction, as well as, how much costs are saved in terms of office, equipment and other labour related costs. For example, employers can afford to lease or purchase smaller, less expensive facilities, pay less for energy and electricity and purchase fewer supplies. In addition, this study showed that in the case of the implementation of this type of employment arrangements and especially the employees who are involved in teleworking, are more likely to report higher levels of loyalty than those who do not implement them. Overall, managing turnover intention is a challenge for many organizations that incur very high costs as a result of voluntary turnover and retaining good workers is critical to any organization, public or private. This is important especially for valuable and high skilled employees, while the costs associated with the new employees recruitment is also high and this usually takes time. Thus, flexible employment arrangements, including teleworking can be a solution to turnover intention reduction, increasing the job satisfaction and improving the well-being of employees and saving costs for organizations. However, most of these issues have not been examined here but are proposed for future research and application. Another concluding remark of this study is that BN and DAG, as it has been discussed, offer an alternative way of deriving causality using observational data and surveys, with various policy implications and implementations to workplaces, employees, employers and to the society overall. Finally, the study presented and discussed an alternative way of identifying possible candidate instrumental variables using the BN and DAG framework, which otherwise can be difficult if not impossible.

Acknowledgments: This work was supported by the *Marie* Skłodowska-Curie Individual Fellowship (IF) grant [652938-TELE]. The author gratefully acknowledges the funding provided by European Commission to carry out this research. This study used panel data from the 2004 and 2011 Workplace Employment Relations Surveys (WERS). The WERS was conducted by NatCen Social Research on behalf of the Department for Business, Innovation and Skills, the Economic and Social Research Council, the UK Commission for Employment and Skills, the Advisory, Conciliation and Arbitration Service and the National Institute of Economic and Social Research. The data was distributed by the UK Data Archive at the University of Essex.

Conflicts of Interest: The authors declare no conflict of interest.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2016 doi:10.20944/preprints201607.0007.v1

Appendix

Proof Lemma 1.

$$P_\Gamma(\mathbf{\Omega}) = \sum_Y P_\Gamma(\mathbf{\Omega}, Y) = \sum_Y [\prod_{W\in\Omega} P(W \mid par(W))] P(Y \mid par(Y)) =$$

$$\prod_{W\in\Omega} P(W \mid par(W)) \sum_Y P(Y \mid par(Y)) = \prod_{W\in\Omega} P(W \mid par(W)) = P_{\Gamma'}(\mathbf{\Omega})$$

The third equality holds because is a leaf node; thus Y is not in X and cannot be in any $par(W)$ for any $W \in \Omega$. Also the fourth equality holds because probability sums to one.

Proposition 1. Let **X** be a set of nodes in a Bayesian network Γ and suppose **X** is ancestral. Let Γ' be the Bayesian network obtained from Γ' by removing all nodes outside **X**. Then

$$P_\Gamma(\mathbf{X}) = P_{\Gamma'}(\mathbf{X}) \tag{A1}$$

Proof:

First step is to find a leaf node and then remove it. Next we get Γ'. According to *Lemma* 1 the probability distribution of **X** remains unchanged throughout the procedure.

Proposition 2. Let **X**, **Y**, and **S** be three disjoint sets of nodes in a Bayesian network such that their union is the set of all nodes. If **S** d-separates **X** and **Y** as above then $\mathbf{X} \perp \mathbf{Y} \mid \mathbf{S}$.

Proof:

Let $\mathbf{S}_1$ be the set of nodes in **S** that have parent in **X** and let assume that $\mathbf{S}_2 = \mathbf{S} \setminus \mathbf{S}_1$. The latter shows that $\mathbf{S}_2$ is member of **S**, but not member of $\mathbf{S}_1$ defined by \. Because **S** d-separates **X** and **Y** then:

For any $W \in \mathbf{X} \cup \mathbf{S}_1$, $par(W) \subseteq \mathbf{X} \cup \mathbf{S}$ and
For any $W \in \mathbf{Y} \cup \mathbf{S}_2$, $par(W) \subseteq \mathbf{Y} \cup \mathbf{S}$
Then let us consider:

$$P(\mathbf{X},\mathbf{S},\mathbf{Y}) = \prod_{W\in \mathbf{X}\cup\mathbf{S}\cup\mathbf{Y}} P(W \mid par(W)) = [\prod_{W\in \mathbf{X}\cup\mathbf{S}_1} P(W \mid par(W))][\prod_{W\in \mathbf{S}_2\cup\mathbf{Y}} P(W \mid par(W))]$$

And

$\prod_{W\in \mathbf{X}\cup\mathbf{S}_1} P(W \mid par(W))]$ is a function of **X** and **S**, while $\prod_{W\in \mathbf{Y}\cup\mathbf{S}_2} P(W \mid par(W))$ is a function of **Y** and **S**.

Theorem 1. (Global Markov property) (Pearl, 2000; Neapolitan, 2003): Given a Bayesian network, let X and Y be two variables and **S** be a set of variables that does not contain X or Y. If **S** d-separates X and Y, then

$$X \perp Y \mid \mathbf{S} \tag{A2}$$

Proof:

Based on the proposition 1 it can be assumed that $an(\{X,Y\} \cup \mathbf{S})$ equals the set of nodes. Thus, $X \perp Y \mid \mathbf{S}$ in original network iff it is true in the restriction onto the ancestral set and **S** d-separates X and Y in original network iff it is true in the restriction onto the ancestral set. Next let **X** be the set of all nodes that are NOT d-separated from X by **S** and let Y be the set of all nodes that are neither in **X** or **Z**. Because of proposition 2 it is $\mathbf{X} \perp \mathbf{Y} \mid \mathbf{S}$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2016 doi:10.20944/preprints201607.0007.v1

There must exist functions $f(\mathbf{X},\mathbf{S})$ and $g(\mathbf{S},\mathbf{Y})$, such that $P(\mathbf{X},\mathbf{S},\mathbf{Y})=f(\mathbf{X},\mathbf{S})g(\mathbf{S},\mathbf{Y})$. Also it should be noticed that $X \in \mathbf{X}$ and $Y \in \mathbf{Y}$.

Then let be $\mathbf{X}'=\mathbf{X} \setminus \{X\}$ and be $\mathbf{Y}'=\mathbf{Y} \setminus \{Y\}$. Then we have $P(X, \mathbf{X}', \mathbf{S}, Y, \mathbf{Y}) = f(X, \mathbf{X},\mathbf{S})g(\mathbf{S}, Y, \mathbf{Y})$.

Consequently it will be:

$$\mathbf{P}(X,Y,\mathbf{S}) = \sum_{\mathbf{X}'\mathbf{Y}'} P(X,\mathbf{X}',\mathbf{S},Y,\mathbf{Y}') = \sum_{\mathbf{X}'\mathbf{Y}'} f(X,\mathbf{X}',\mathbf{S},Y)g(\mathbf{S},Y,\mathbf{Y}') =$$
$$[\sum_{\mathbf{X}'} f(X,\mathbf{X}',\mathbf{S},Y)][\sum_{\mathbf{Y}'} g(\mathbf{S},Y,\mathbf{Y}')] = f'(X,\mathbf{S})g'(\mathbf{S},Y)$$

And that is (A2) or $X \perp Y | \mathbf{S}$

In figure A1 simple graphical proof of the factorization relation (4) is presented. In this case the aim is to explore the casual effect from X to Y and it will be as:

$$P(Y|Z,X) = \frac{P(Z,X,Y)}{P(Z,X)} = \frac{p(X)P(X|Z)P(Y|X)}{p(X)P(X|Z)} = P(Y|X) \qquad \text{(A3)}$$

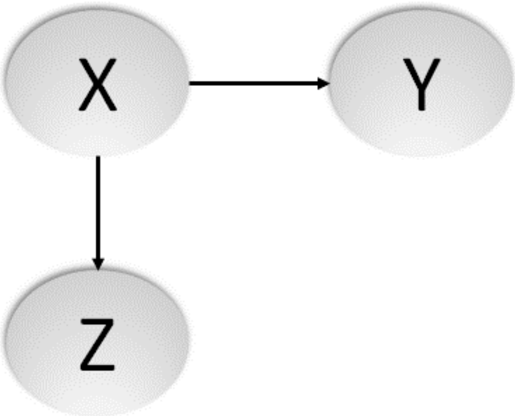


Figure A1. X causes Y.

It becomes obvious that it is not necessary to estimate the probabilities of $P(Z|X)$ or $P(Z)$ and $P(X)$ since the only parent of Y is X. Moreover, Z is a descendant of Z, but Z is not confounder either of X or Y, thus it should not be considered in the regression analysis. In other words DAGs show how the regression can be done correctly.

Next the typical example of the alarm problem (Pearl, 1988) is presented. In this example we estimate the probability of making the call to police because the alarm was deactivated. This might be a very simplified example; however it serves as a good exercise to show the importance of the factorisation equation (4). The figure shows that the reason someone will call is that the alarm has gone off. Thus in this case the only cause to call is the fact the alarm has been off. But this is not enough, as based on (4) also the parents of alarm should be included. Going one step back one cause of the alarm being off is the probability that a burglar broke into the house. Thus, the model should condition also on this probability. However, the additional information that an earthquake took place reduces the belief that it was merely a burglar but we cannot be totally sure for that. Nevertheless, the information given that also the radio has been also off gives an additional information and belief that the earthquake can be a very probable cause of the fact that the alarm gone off. In that case the model should condition on both parents of the alarm, which is the earthquake and the burglar. Without considering the equation (4) the probability of the alarm system-problem would be:

$$P(C,A,R,E,B) = P(B)P(E|B)P(R|E,B)P(A|R,B,E)P(C|A,R,B,E) \qquad \text{(A4)}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2016 doi:10.20944/preprints201607.0007.v1

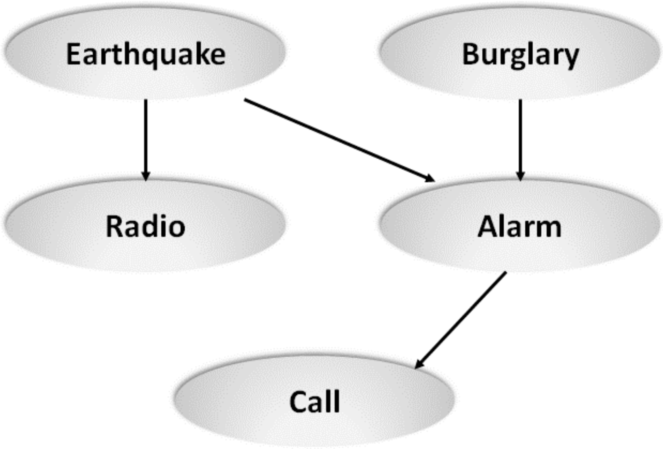


Figure A2. Factorisation Example for Alarm Problem

Using the factorization equation (4) the probability of the system is simplified to

$$P(C, A, R, E, B) = P(E)P(R \mid E)P(A \mid E, B)P(C \mid A) \tag{A5}$$

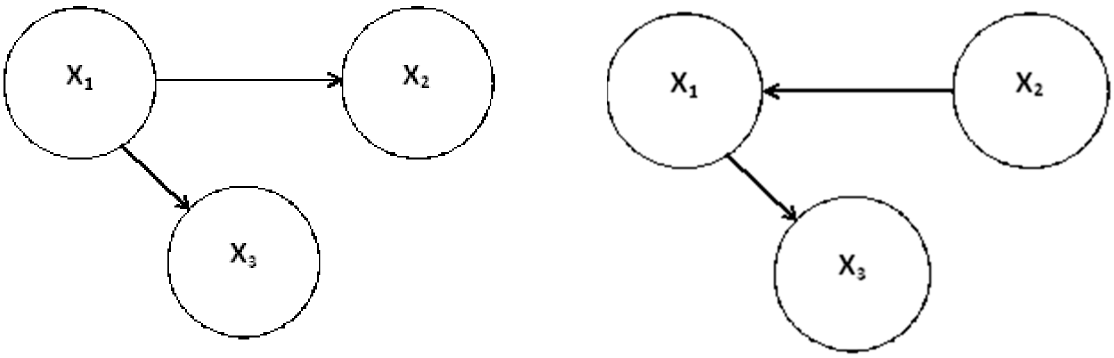
However, there might be issues regarding the direction of the association between the variables. For example in figure A3.a. is:

$$P(X_1, X_2, X_3) = P(X_3 \mid X_1)P(X_2 \mid X_1)P(X_1) \tag{A6}$$

which is equivalent with the probability in figure A3.b and it is:

$$P(X_1, X_2, X_3) = P(X_3 \mid X_1)P(X_2 \mid X_1) = P(X_3 \mid X_1)P(X_1 \mid X_2)P(X_2) \tag{A7}$$

which holds by symmetry property.



A3.a A3.b

Figure A3. An example of Arc reversal-Bayes rule

Similarly if we have the situation in figure A4.a.

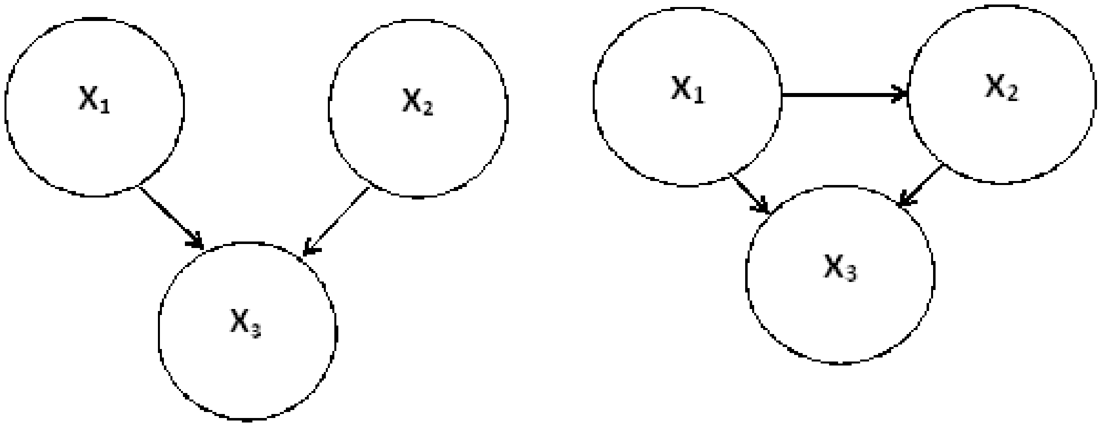
$$P(X_1, X_2, X_3) = P(X_3 \mid X_2, X_1)P(X_2)P(X_1) \tag{A8}$$

And this is equivalent with figure A4.b. which is:

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2016 doi:10.20944/preprints201607.0007.v1

$$P(X_1, X_2, X_3) = P(X_3, X_2 \mid X_1)P(X_1) = P(X_2 \mid X_3, X_1)P(X_3 \mid X_1)P(X_1) \quad \text{(A9)}$$

This is one of the issues and thus one very important definition for the DAG and BN, which is the *d-separation* is it has been discussed in the main text.



A4.a A4.b

Figure A4. A second example of Arc reversal-Bayes rule

Proof Lemma 2:

$$\tau_{yx} = \frac{E[I,Y \mid Z] - E[I \mid Z]E[Y \mid Z]}{E[I,X \mid Z] - E[I \mid Z]E[X \mid Z]} = b_{iv} = \frac{E[Y \mid I]}{E[X \mid I]} \quad \text{(A9)}$$

Next we present some examples of the *d-separation* regarding the three sources of bias; the over-control, the confounding and the endogenous bias. In figure A5 a very simple example of the over-control bias is presented. Here A has an influence on B, which in turn has an influence on C. Obviously, evidence about A will influence the certainty of B, which then influences the certainty of C. Similarly, evidence about C will influence the certainty of A through B. On the other hand, if the state of B is known, then the channel is blocked, and A and C become independent; we say that A and C are d-separated given B. This is in line with the back-door criterion and the factorisation equation (4), where the model should control only for the parent of the factor or treatment of interest or the confounders of the factor and the outcome of interest and not on the descendants of the treatment.

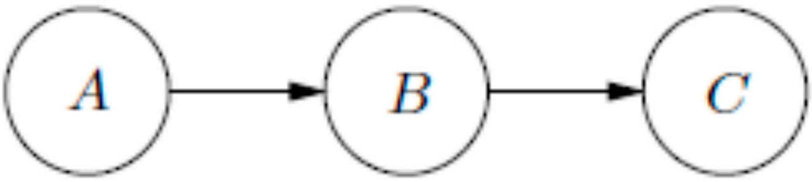


Figure A5. Over-control bias.

The second case is the confounding bias where BN and DAG are able to account for and one simple example is presented in figure A6. This is the well-known issue where the failure to condition on a common cause or the confounders into a regression will lead to confound bias.

 doi:10.20944/preprints201607.0007.v1

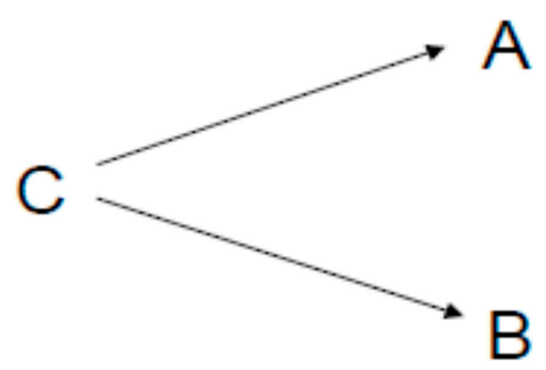
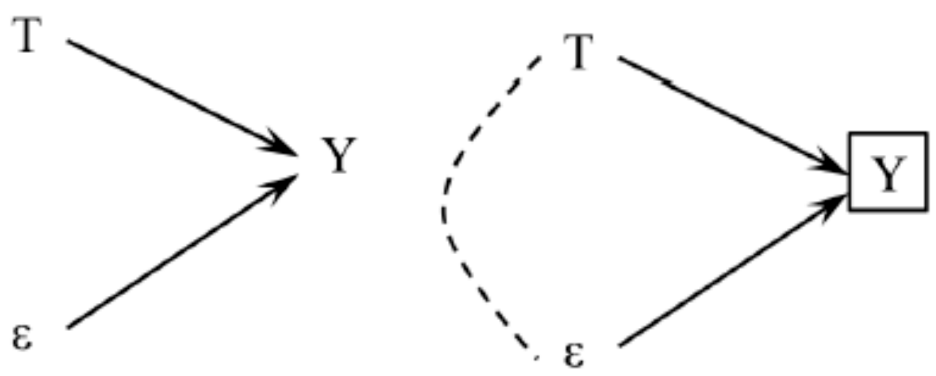


Figure A6. Confounding bias

The next example refers to the endogenous bias and the study by Hausman and Wise [33]. In this case the authors want to estimate the effect of education on income derived by the New Jersey Income Maintenance experiment. In figure A7.a the effect of T (education) on Y (Income) is identified. However, Hausman and Wise [33] restrict the sample to low earner with Y<$5000. Conditioning on Y in that case, a non-causal association between T and ε is induced, creating endogenous bias, as it can be seen in figure A7.b.



A7.a **A7.b**

Figure A7. The effect of education on income

References

1. Golden, T.D.; Veiga, J.F.; Simsek, Z. Telecommutings differential impact on work-family conflict: Is there no place like home?' *Journal of Applied Psychology* **2006**, *91*, 1340-1350.
2. Bailyn, L. Freeing Work from the Constraints of Location and Time. *New Technology, Work, and Employment*, **1988**, *3*, 143–152.
3. WorldatWork. Telework 2011'. A WorldatWork Special Report Based on Data Collected by The Dieringer Research Group Inc. and WorldatWork **2011**, http://www.worldatwork.org/waw/adimLink?id=53034.
4. Lister, K.; Harnish, T. The state of teleworking in the U.S.: how individuals, businesses, and government benefit. Teleworker Research Network. **2011** http://www.workshifting.com/downloads/downloads/Telework-Trends-US.pdf
5. Gajendran, R.S.; Harrison, D.A.. The good, the bad, and the unknown about telecommuting: Meta-analysis of psychological mediators and individual consequences. *Journal of Applied Psychology*, **2007**, *92*, 1524-1521.
6. Crandall, W.; Gao, L . An update on telecommuting: Review and prospects for emerging issues. S.A.M. *Advanced Management Journal*, **2005**, *70*(3), 30-37.
7. Nilles, J.M. *Making telecommuting happen: A guide for telemanagers and telecommuters*. New York: John Wiley & Sons, **1994**

 doi:10.20944/preprints201607.0007.v1

8. Bélanger, F. Workers' propensity to telecommute: An empirical study. *Information Management*, **1999**, *35*(3), 139-153.
9. Potter, E.E. Telecommuting: The future of work, corporate culture, and American society. *Journal of Labor Research*, **2003**, *23*(1), 73-83.
10. Sims, H.P.; Szilagyi, A.D.; Keller, R.T. The measurement of job characteristics. *Academy of Management Journal*, **1976**, *19*, 195-212.
11. Olson, J.S.; Teasley, S.; Covi, L.; Olson, G. The (currently) unique advantages of collocated work. In P. Hinds and S. Kiesler (Eds.: Distributed work (113-135) Cambridge, MA: MIT Press, **2002**.
12. Short, J;, Williams, E.; Christie, B. *The social psychology of telecommunications*. London, UK: Wiley, **1976**
13. Burgoon, J.K.; Bonito, J.A.; Ramirez, A.; Dunbar, N.E.; Kam, K.; Fischer, J. Testing the interactivity principle: Effects of mediation, propinquity, and verbal and nonverbal modalities in interpersonal interaction. *Journal of Communication* **2002**, *52*, 657-677.
14. Lowry, P. B.; Roberts, T.L.; Romano, N.C.; Cheney, P.D.; Hightower, R.T. The impact of group size and social presence on small-group communication: Does computer-mediated communication make a difference? *Small Group Research*, **2006**, *37*, 631-661.
15. Golden, T.D.; Veiga, J. F. The impact of extent of telecommuting on job satisfaction: Resolving inconsistent findings. *Journal of Management*, **2005**, *31*, 301- 318.
16. Fonner, K.L.; Roloff, M.E. Why teleworkers are more satisfied with their jobs than are office-based workers: When less contact is beneficial. *Journal of Applied Communication Research*, **2010**, *38*, 336-361.
17. Velayutham, S. Precarious experiences of Indians in Australia on 457 temporary work visas. *Economic and Labour Relations Review*, **2013**, *24*(3), 340–361.
18. Wilson, S.; Ebert, N. Precarious work: economic, sociological and political perspectives. *Economic and Labour Relations Review*, **2013** *24*(3), 263–279.
19. Rosenbaum, P.R.; Rubin, D.B. The central role of the propensity score in observational studies for causal effects. *Biometrika*, **1983**, *70*(1), 41–55.
20. Bromley, J.; KJackson, N.A.; Clymer O.J.; Giacomello A.M.; Jensen F.V. The use of Hugin to develop Bayesian networks as an aid to integrated water resource planning". Environmental Modelling & Software, **2005**, *20*(2), 231-242.
21. Heckerman, D.E. Bayesian networks for data mining. *Data Mining and Knowledge Discovery*, **1997**, 1, 79-119.
22. Pearl, J. *Probabilistic Reasoning in Intelligent Systems*. San Mateo, California, Morgan Kaufmann, **1988**.
23. Spirtes, P.; Glymour, C.; Scheines, R. *Causation, Prediction, and Search*. 2nd ed, MIT Press, Cambridge, MA., **2000**.
24. Pearl, J. *Causality. Models, Reasoning, and Inference*. Cambridge University Press, Cambridge, **2000**.
25. Pearl, J. Causal inference in statistics: An overview. *Statistical Surveys*, **2009**, *3*, 96–146.
26. Neapolitan, R.E. *Learning Bayesian Networks*. Prentice Hall Series in Artificial Intelligence, Pearson Education: New Jersey, **2003**.
27. Kalisch, M.; Buhlmann, P. Estimating High-Dimensional Directed Acyclic Graphs with ¨ the PC-Algorithm. *Machine Learning*, **2007**, *8*, 613–636.
28. Clark, A.E.; Oswald, A.J. Satisfaction and Comparison Income. *Journal of Public Economics*, **1996**, *61*, 350-381.
29. Hamermesh, D.S. The changing distribution of job satisfaction. *Journal of Human Resources*, **2000**, *36*, 1–30.
30. Giovanis, E.; Ozdamar, O. Structural Equation Modelling and the Causal Effect of Permanent Income on Life Satisfaction: The Case of Air Pollution Valuation in Switzerland. *Journal of Economics Surveys*, **2016**, *30*(3), 430-459.
31. Lee, Y.K.; Lee, E.R.; Park, B.U. Principal Component Analysis in very High-Dimensional Spaces. *Statistica Sinica*, **2012**, *22*, 933-956.
32. Karamizadeh, S.; Shahidan, M.A.; Azizah, A.M.; Mazdak, Z.; Hooman, A. *Journal of Signal and Information Processing*, 2013, 4, 173-175.
33. Hausman, J.A.; Wise, D.A. Social experimentation, truncated distributions and efficient estimation. Econometrica , **1977**, *45*, 919–38