How the 'fear factor' spreads across the global financial markets

Jonas Gottal

Preliminary results*of upcoming thesis at
Faculty of Mathematics, Computer Science and Statistics,
Ludwig Maximilian University of Munich

9th July 2022

Abstract

This excerpt serves as a summary of the application results of the underlying work as an attachment to the published code on GitHub. To apply our obtained knowledge described in the underlying thesis, we propose a use case in finance with real world data. Our interest targets Implied Volatility data and the relationships of this indicator between different countries. Since this measure is often used to capture the market's sentiment towards the volatility of a particular asset, it is also referred to as the "fear factor". [12] [10]

I. Introduction

The underlying thesis describes how agents can learn from data through probability theory by reformulating the task as exact inference and maximum likelihood. We focus on learning Bayesian networks and show why they provide optimal predictions, solutions to overfitting and how they even handle decision making with incomplete data. And since Bayesian networks provide the potential for causal interpretations, we conclude with causal discovery. This excerpt serves as a summary of the application results and the approach we propose is structured as follows: First, we discuss the terminology of financial markets and then analyse and process our raw data. After learning our Bayesian networks, we validate their predictive properties and compare them appropriately. Finally we apply methods of causal discovery to infer the actual impact of countries.

*Code available on GitHub: `https://github.com/trashpanda-ai/GitHub-Local/tree/master/Causal%20Bayesian%20Networks`

II. Terminology: Financial Markets and Derivatives

Besides dark pools there are two different ways to conclude a transaction: over-the-counter (OTC) between two firms, and exchange trades. At exchanges for every transaction the order book is filled with the relevant (and thus for us useful) information, while OTC and dark pool trades remain invisible to the public eyes. We further separate between the stock market and derivatives market, which is significantly bigger in terms of underlying assets. By exploiting minuscule price differences on two or more exchanges to their own benefit, arbitrageurs balance these markets. Thus we can utilise the information from the derivative exchanges to gain insights into the stock market and vice versa. But what are derivatives? As an umbrella term, it is vaguely defined as an agreement between two parties on a future transaction, where the value can be *derived* from a number of underlying variables. The two most popular derivatives are futures and options. While the future is a firm agreement to buy or sell an underlying asset *at* a specified time in the *future* at a specified

price, the option gives the holder the right – the *option* – to exercise the trade *up to* the specified time. These options can now be used to calculate the so called *implied volatility* – volatility *implied* by the option's prices. And while calculating historical volatility – via standard deviation – is straight forward, the calculation of implied volatility is more complicated and not part of this thesis. For example the famous Black-Scholes approach aims to approximate the price of an option by modelling the geometric brownian motion – taking into account for volatility, time and the current price of the underlying, among other variables. And if you apply this model in the opposite direction, you can calculate the implied volatility for the underlying using the current option prices. [12] [10]
Implied volatilities are calculated for countless financial products – but our focus is on national equity market indices, replicating a countries stock market performance in one instrument. We gathered a data set of implied volatility indices from the twelve most important countries for financial markets, summarised in Table 4. Intraday data would be optimal to analyse the simultaneous changes on each exchange, but unfortunately, we were only able to collect daily data (end of day). However, for a long time horizon of up to 10 years. [10]

III. Data Management: Preprocessing and Validation

Our data set (Table 4, Appendix) contains twelve different implied volatility indices, each replicating the implied volatility of national stock market indices. For reasons of readability, we refer to them in the following by the name of the respective country. The raw time series of daily implied volatility in percent points (Figure 1) demonstrates an important property for our further analyses: stationarity. Thus, our means, variances, and correlations do not vary excessively and we do not have to address the influence of a global trend within our time series. Also within non-dynamic Bayesian networks the underlying generating process can not change over time and the predictive accuracy of our models would differ at different points in time. Obviously our data shows some outliers (especially Russia in 2014 due to oil prices) and trends at more granular level, but in general our data provides sufficient quality. [8] [22]

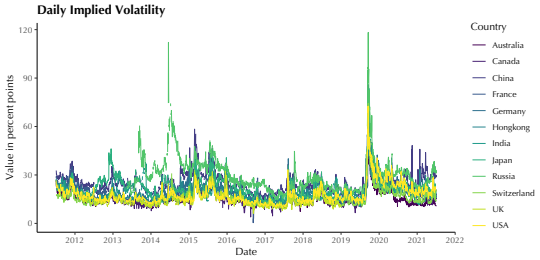

Figure 1: *Stationary time series: An index value of 30 means that the implied volatility of the underlying is considered to be at 30%.*

But we care mostly about the daily changes while the different markets interact and therefor calculated them as daily *log returns* for multiple reasons: The most important one is that a stock market can be approximated as a geometric Brownian motion and we can use *Itô's lemma* to derive that they are log-normally distributed. And since we are aiming for normally distributed variables, we chose the natural logarithm to achieve this. Furthermore this approach yields a normalised data set (Figure 2). [12] [10]

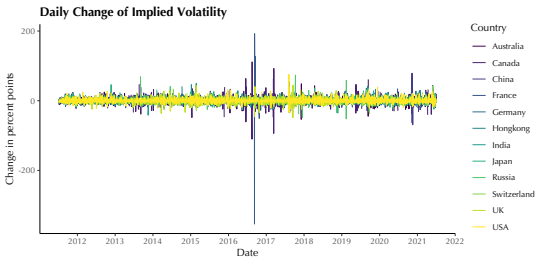

Figure 2: *Daily log returns.*

And since we want to analyse the impact of one country on another, but do not need to predict the exact value of the index, we are satisfied with the up- and downward movements. Thus, we discretise our data, with 1 represent-

ing positive and -1 negative changes. Due to the immense activity in these large financial instruments, there are no sideways movements and therefore no values of 0. As in Figure 3 illustrated, the values are balanced and not skewed in either direction.

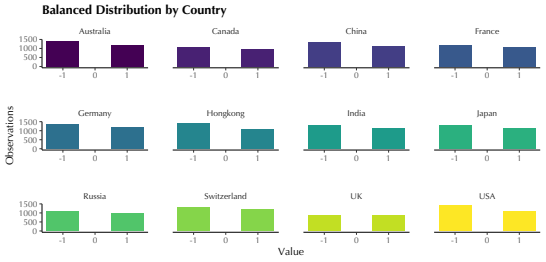

Figure 3: *A balanced distribution between -1 and 1.*

For validation purposes, we randomly split our dataset row-wise into a test and training set with the ratio ⅓ and ⅔, leaving still enough data to train reliable models. For the EM algorithm we use the incomplete data as is, but for the score based approaches we omit the rows of the train set with missing values – resulting in the EM algorithm seeing more data. The test set for both is without rows of missing values and therefore the same to be comparable for subsequent validation. [13] [9]

IV. Learning Networks from Implied Volatility Data

With our processed and discretised data, we can now learn our Bayesian network's structure and parameters. Due to the diversity of algorithms and score functions in `bnlearn`, we have to make some restrictions. As previously discussed hybrid and constraint-based approaches perform less accurately. [20] Therefore we only implemented three hybrid algorithms (*Restricted Maximization, Hybrid HPC, Max-Min Hill-Climbing*) to verify this thesis for our data, but mostly focus on the score functions in Table 1 optimised by Hill-Climbing and Tabu from prior sections. Furthermore we implement the EM algorithm from prior sections, with the maximisation step conducted by each Hill-Climbing and Tabu.

As the name suggests, information-theoretic scoring functions are based on information theory and often equivalent to a well established concept (e.g., BIC ⇔ MDL, Loglik ⇔ (cross) entropy), and Bayesian scoring functions utilise Bayes theorem and maximise the a-posteriori probability distribution of the networks. [2] [19] *Score-equivalence* means the scoring function assigns the same score to all networks in the same *Markov equivalence* class, denoted by the same conditional independencies. In general, these equivalence classes are Completed Partially Directed Acyclic Graphs (CPDAG) whose only directed arcs either belong to v-structures, as illustrated in Figure 4, or would lead to another v-structure or cycle if reversed. We will learn more about them in Section VI. [15] [19] [3]

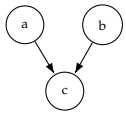

Figure 4: *v-structure*

To verify – or exclude – the presented scoring functions, as well as hybrid- and score-based algorithms, we run a *k*-fold cross validation to measure the loss of our data. This means we partition our data in k equal-sized parts, and learn the model with $k-1$ parts. To validate them, we calculate the negated expected log-likelihood – *negative entropy* – of the remaining partition k, where lower values are better. [15] [9] [19] The results in Figure 5 show, that we can not exclude certain combinations, as the overall loss is very similar.

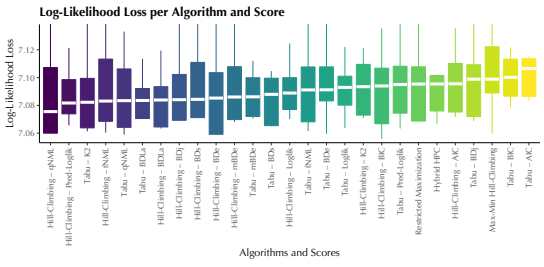

Figure 5: *k-fold cross validation with $k = 10$ folds to verify the fit of the algorithms and scoring functions to our data as implemented in R.*

We apply both the score-based and hybrid approaches by averaging $n = 100$ networks of a bootstrap approach: by estimating the confidence in individual arcs and setting their minimum threshold, we obtain a more robust structure overall as implemented in `R`. The following networks in Figure 6 are the results, but as they are differing in many arcs and orderings, we need to define a way to objectively validate and compare of our different models. [15] [19] [11] [7]

V. Evaluating Networks from Implied Volatility Data

To validate our learned Bayesian networks we exploit the binary representation of our data. We make predictions from our networks using `predict()` and compare them with our test data.

Sensitivity or true positive rate (TPR) is derived from the true positives TP, i.e., the correctly identified positives P from our test set: $\text{TPR} = \frac{\text{TP}}{\text{P}}$

Specificity or true negative rate (TNR) is derived from the true negatives TN, i.e., the correctly identified negatives N from our test set: $\text{TNR} = \frac{\text{TN}}{\text{N}}$

By plotting both the sensitivity and specificity in relation, we obtain the so called Receiver Operating Curve (ROC), where for both measures, 1 or 100% is the optimum, and if the curve is the diagonal, we observed a random process. In Figure 7, we can see the ROC for each country – or node in our network in this case – depicted for each algorithm for comparability. To further summarise these evaluations, we can calculate the Area Under the Curve (AUC) and populate Table 3 (Appendix), ordered by the sum of all values to rank our models. Again, a value of 0.5 indicates a random process. [9] [6] [13] [18]

As illustrated in Figure 7 and Table 3 (Appendix), most countries demonstrated significant predictive qualities. But *Canada*, *India*, and especially *Russia* are very close to a random process. We continue our analysis but consider the possibility of dropping them from our data set and therefore Bayesian networks.

We proceed by comparing the best performing networks (in terms of probability distributions) and illustrating their overlap, as shown in Figure 8.

After comparing the Bayesian networks in terms of structural similarities, we now examine the reliability of the predictions and the corresponding conditional probabilities. In Figure 9, we illustrate this confidence using arc strength and can show that we can expect reliable distributions for *France*, *Germany*, and *Switzerland* in particular. Table 2 provides an excerpt of some recognisable conditional probabilities.

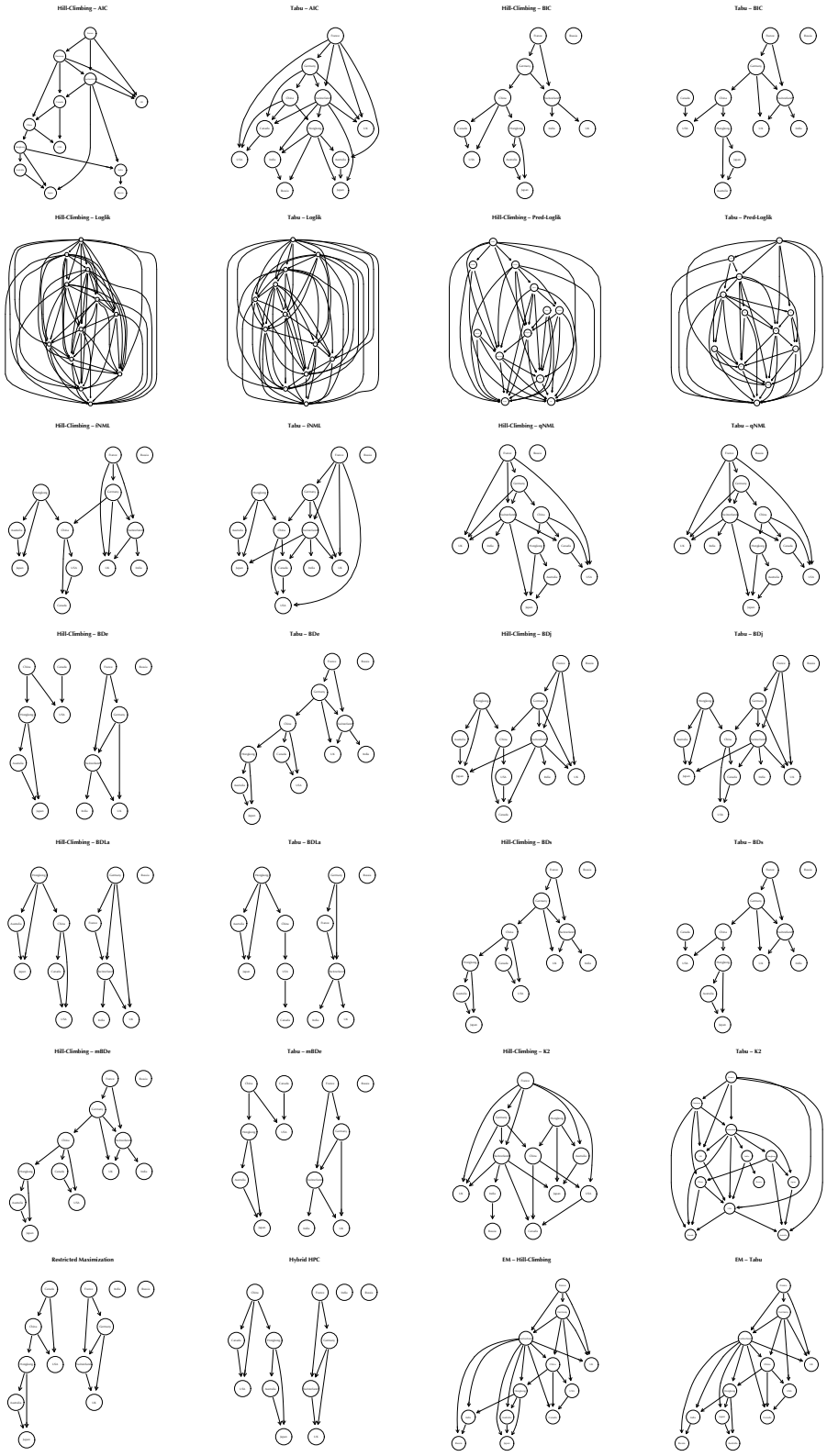

Figure 6: *Resulting learned Bayesian networks with dedicated algorithms and scoring functions*

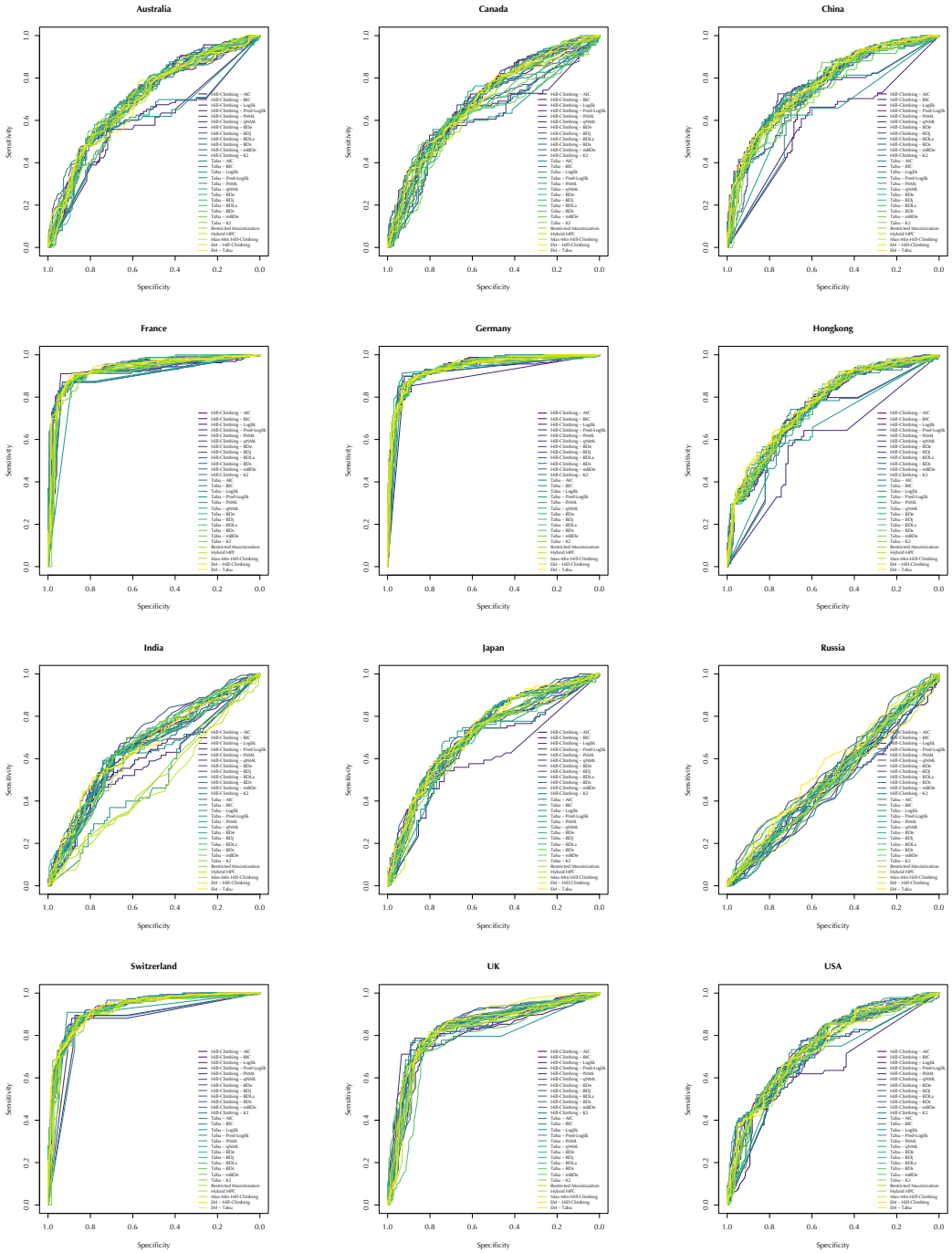

Figure 7: *ROC curves of all implemented algorithms per country (node)*

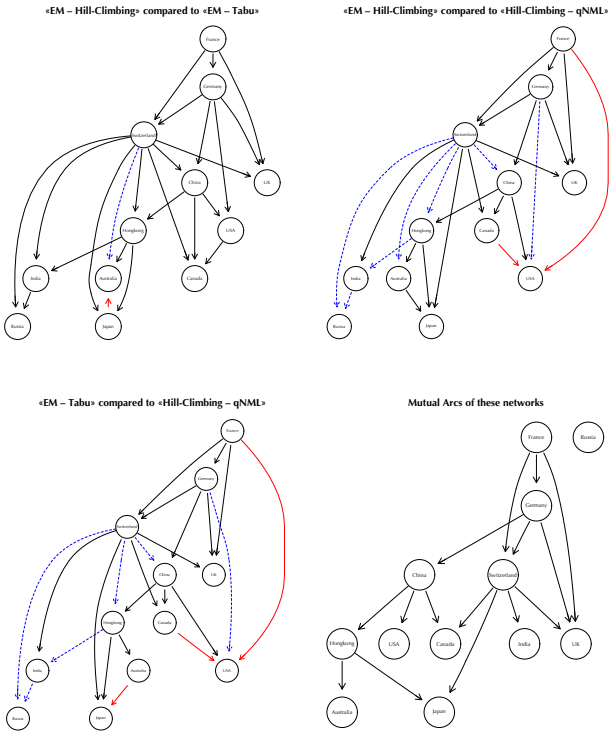

Figure 8: *Comparison of the three networks with the highest combined AUC scores*

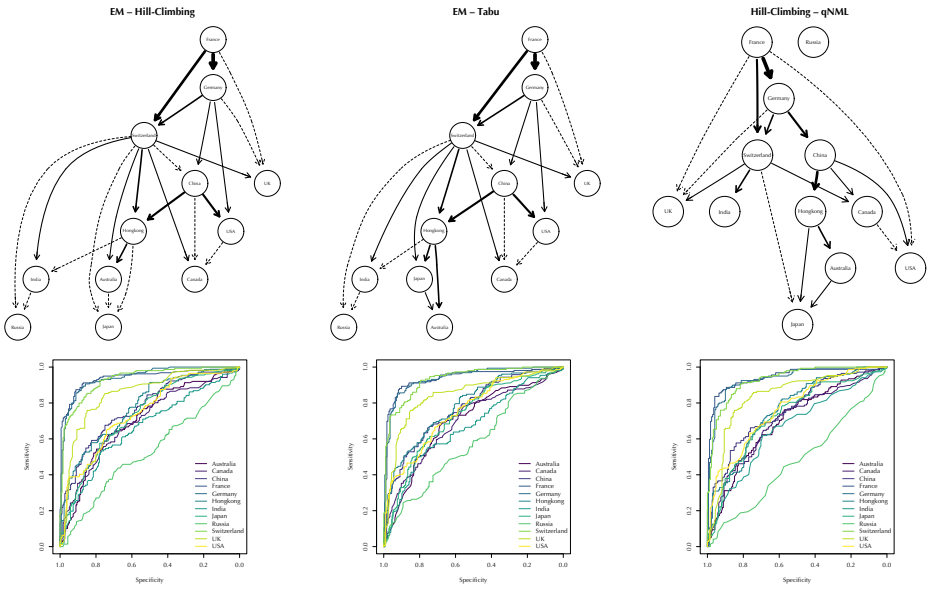

Figure 9: *Comparison of the three networks with the highest total AUC scores – confidence depicted as arc strength by Δ BIC scores.*

VI. Causal networks and final result

"I would rather discover one causal law than be King of Persia."

– Dēmókritos, *(460 – 370 B.C.) [16, p. 41]*

When we introduced this use case, we talked about the impact of the countries on each other. But up to this point, we have merely identified significant conditional probabilities – in other words, correlation, not causation. That is, why we now introduce and apply methods of causal discovery to infer the actual cause – the desired impact. [16] [4]

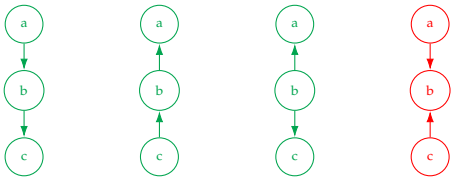

Figure 10: *Markov Equivalence*

We do not attempt to identify the best causal Bayesian network, but rather a small set of plausible causal Bayesian networks that fit our data. And there are several methods to learn a causal model directly, but we aim for a more comparable approach (here ROC in Section V) and thus derive causality from our networks by causal structure search. [8] For our further approach, it is important to be aware of the Markov equivalence classes: identical adjacencies – derived from the Markov condition – and the same *unshielded colliders*. Exemplified for three nodes in Figure 10 in green. The two cascading arrangements and the pattern of a common parent are all equivalent – only the v-structure, the unshielded collider, in red does not belong to the equivalence class. Most causal discovery approaches learn exclusively these equivalence classes. [16] We broaden our assumptions and in addition to acyclicity we define the following: [16] [17] [23] [5]

I *Causal Markov* assumption – an extension to our previous Markov assumption: instead of conditional probabilities, a deterministic relationship is implied and nodes are conditionally independent of its non-descendants, given its direct causes.

II *Causal faithfulness* or stability assumption – as it assumes all independencies to be stable and thus for any change in the parameters they will remain.

III *Causal sufficiency* assumption – of no common confounders: all the common causes of our variables are already present in our network.

Our causal structure search is based on the *Inductive Causation* algorithm from Pearl and Meek. The gist is the inductive reasoning of Ockham's razor to discard any model for which we find a more compact *minimal* model that represents our data equally. This *inferred causation* means in essence, that a is supposed to to have a causal influence on b if there is $a \rightarrow b$ 'in every minimal structure consistent with the data' [16, p. 45]. Thus, the topology of our Bayesian network – the underlying DAG – is sufficient for our causal structure search. [14] [17] [16] [5]
As we learn a causal Markov equivalence class, only unshielded colliders present in our underlying Bayesian network are to be adapted. Any further v-structures must be avoided, as illustrated in Figure 11. These rules are sound, as any other orientation in these patterns lead to further unshielded colliders or acyclicity. This approach is illustrated as pseudo code in Figure 14 and implemented in R (Appendix). [14] [17]

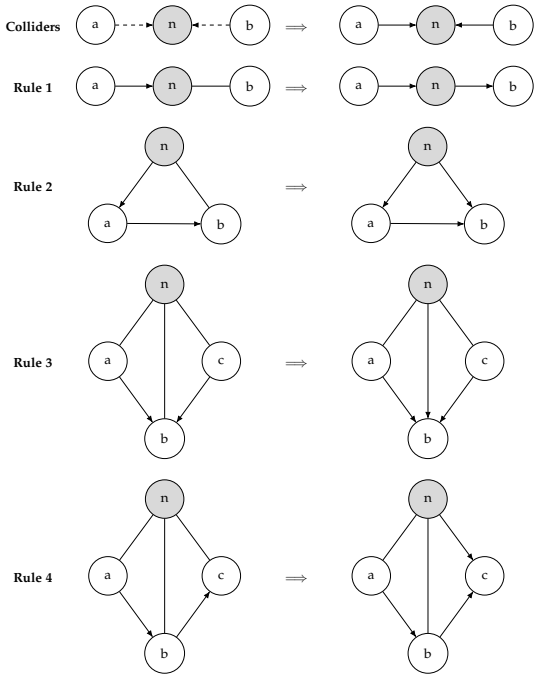

Figure 11: *After orienting the unshielded colliders present in our original Bayesian network once, we follow these four rules of the Inductive Causation algorithm repeatedly until convergence to obtain a maximally oriented Partially Directed Acyclical Graph (PDAG).*

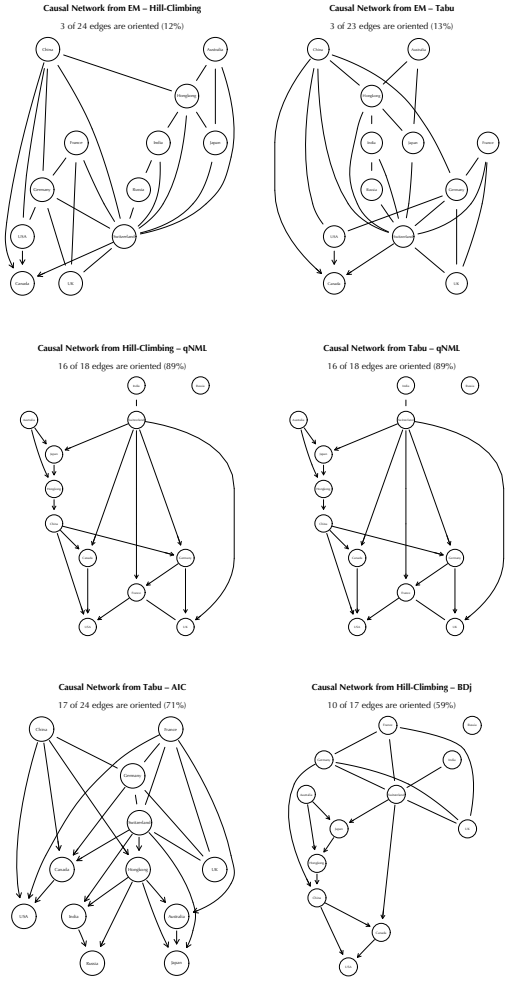

Figure 12: *Comparison of the networks from the highest quantile in terms of AUC scores.*

Now we can apply our algorithm from Figure 14 and discuss the results. In Figure 12 our two best performing networks could barely be oriented, only the Hill-Climbing – qNML network has significant orientations. But as we are interested in causal networks with both many orientations and good predictive qualities, we expand our focus to the best quantile in terms of AUC. Thus the additional causal networks in Figure 12. Also for Tabu – qNML we can orient most of the edges and only *India – Switzerland* and *UK – France* is missing a directed arc. Furthermore, as *Russia* is no part of the network and thus no cause for any effects, we can now remove it from our data set due to its low significance in our ROC analysis. *Canada* and *India* will remain in our data set, as their AUC scores are not necessarily random.

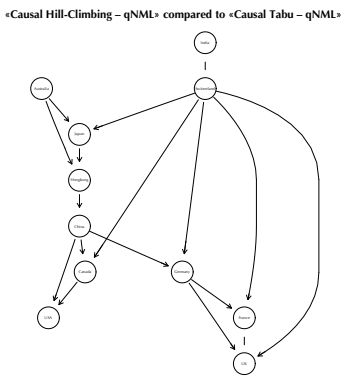

Figure 13: *Final causal structure: overlap of our networks with the most orientations in the highest quantile of total AUC.*

For the two causal networks with the most orientations, it can be argued that *Switzerland* is the financial hub. As described in the Appendix, *China*'s underlying is traded in the USA, but tracks Chinese equities traded on the Hongkong stock exchange. This explains the connection between *China*, *USA*, and *Hongkong*. So without *China* – or with more reliable data on China's implied volatility – *Switzerland* divides the network into a western and an eastern hemisphere with dedicated sub-clusters. Thus, the time zones and consequently the tradable hours could be an important generating process for the data. Although we cannot draw many conclusions, *Switzerland*'s role as a global financial epicenter seems very intuitive.

VII. Critique

Even if impacts can be discovered without chronological temporal information, causal discovery can only benefit from it. Thus, with intraday data this approach would be even more reliable. Also, with respect to *measurement errors*, with more granular data our models would be more resilient, e.g., to the influence of high-frequency traders and other speculators. Currently, we have to assume that our multivariate time series of ten years is invariant. However, since the underlying generating process changes frequently, this likely results in an under-fit for our averaged static model. Thus, *dynamic* Bayesian networks could be implemented with a shorter (but denser) time frame and validated by rolling windows and tested against recent unseen data. [17] [8] [15]

Overall, our approach of inferred causation is based on assumptions that are unlikely to apply to our use case and need to be carefully evaluated. We included only the largest global financial players, no other countries, and most importantly, no other variables. Thus, on the one hand, the *selection bias* and, on the other hand, the assumption that there are no common confounding factors are very fragile: why should implied volatility not also be influenced by completely different processes? Thus, even if the IC algorithm works with latent variables, we need to include more detailed and comprehensive data from more potential influencers and then implement a dynamic approach. [1] [8] [17]

Appendix

Abbreviation	Name	Category	Score equivalent
AIC	Akaike Information Criterion	Information-theoretic	✓
BIC	Bayesian Information Criterion	Information-theoretic	✓
fNML	factorised Normalized Maximum Likelihood	Information-theoretic	✗
qNML	quotient Normalized Maximum Likelihood	Information-theoretic	✓
Loglik	Log-likelihood	Information-theoretic	✓
Pred-Loglik	Predictive Log-likelihood	Information-theoretic	✓
BDe	Bayesian Dirichlet equivalent	Bayesian	✓
BDj	Bayesian Dirichlet equivalent (Jeffrey's prior)	Bayesian	✗
BDLa	locally averaged Bayesian Dirichlet	Bayesian	✗
BDs	Bayesian Dirichlet sparse	Bayesian	✗
mBDe	modified Bayesian Dirichlet equivalent	Bayesian	✗
K2	K2	Bayesian	✗

Table 1: *Overview of all scoring functions available using* bnlearn *and implemented with Hill-Climbing and Tabu. [2] [21] [19]*

Conditional Probability	EM – Hill-Climbing	EM – Tabu	Hill-Climbing – qNML
$P(Germany \mid France)$	0.83	0.83	0.86
$P(Switzerland \mid France, Germany)$	0.85	0.85	0.83
$P(\neg UK \mid \neg France, \neg Germany, \neg Switzerland)$	0.85	0.85	0.85
$P(Hongkong \mid China, Switzerland)$	0.69	0.69	0.59
$P(China \mid Switzerland, USA)$	0.74	0.74	0.68
$P(USA \mid China, Germany)$	0.69	0.69	0.66

Table 2: *Excerpt of some recognisable conditional probabilities conducted with* `cpquery`*: since the structures are not identical, these values cannot be retrieved directly from the conditional probability tables.*

[0]Tradable Hours pulled from dedicated website of exchange.
[1]Sometimes referred to as AXVI.
[2]Sometimes referred to as Nikkei 225 VI.

Algorithm 1 Inductive Causation based on Pearl and Meek

Input: Bayesian network (BN) from `bnlearn`
Output: Partially Directed Acyclic Graph (PDAG) in `bnlearn`

procedure Inductive Causation(BN)
  $PDAG \leftarrow$ skeleton(BN)
  for nodes n in $PDAG$ **do**
    if neighbours$(n) > 1$ **then**
      for permutations (a, b) of neighbours(n) **do**
        if arcs $a \rightarrow n$ and $b \rightarrow n$ exist in BN **then**
          if no edge between a, b **then**
            add arcs $a \rightarrow n$ and $b \rightarrow n$ ▷ C
          end if
        end if
      end for
    end if
  end for
  new.arcs = TRUE
  while new.arcs = TRUE **do**
    new.arcs = FALSE
    for nodes n in $PDAG$ **do**
      if neighbours$(n) > 1$ **then**
        for permutations (a, b) of neighbours(n) **do**
          if no edge between a, b **then**
            if $(a \rightarrow n)$ and **not** $(b \rightarrow n)$ **then**
              add arc $n \rightarrow b$ and new.arcs = TRUE ▷ R1
            end if
          end if
          if b in descendants(n) and **not** $(b \rightarrow n)$ **then**
            add arc $n \rightarrow b$ and new.arcs = TRUE ▷ R2
          end if
        end for
      end if
      if neighbours$(n) > 2$ **then**
        for permutations (a, b, c) of neighbours(n) **do**
          if $(a \rightarrow b)$ and $(c \rightarrow b)$ **then**
            add arc $n \rightarrow b$ and new.arcs = TRUE ▷ R3
          end if
          if $(a \rightarrow b)$ and $(b \rightarrow c)$ **then**
            add arc $n \rightarrow c$ and new.arcs = TRUE ▷ R4
          end if
        end for
      end if
    end for
  end while
end procedure

Figure 14: *The algorithm based on* Pearl *and* Meek *uses* `bnlearn` *methods "*`skeleton`*", "*`neighbours`*" and "*`descendants`*" and only adds arcs if the PDAG remains acyclical.*

Algorithm	*Australia*	*Canada*	*China*	*France*	*Germany*	*Hongkong*	*India*	*Japan*	*Russia*	*Switzerland*	*UK*	*USA*	∑ **AUC**
EM – Tabu	0.714457	0.710526	0.801356	0.935114	0.935456	0.804403	0.667599	0.748692	0.603444	0.911532	0.838552	0.763433	**9.434565**
EM – HC	0.715970	0.699063	0.788222	0.932255	0.934769	0.806075	0.659308	0.734171	0.589861	0.912596	0.836478	0.763318	**9.372085**
Tabu – BIC	0.716891	0.709326	0.780017	0.927320	0.924025	0.785356	0.632751	0.731162	0.571701	0.910025	0.816523	0.737721	**9.242818**
Tabu – FNML	0.727862	0.710461	0.779239	0.925556	0.935750	0.778078	0.634877	0.737049	0.516135	0.903620	0.839178	0.744050	**9.231854**
Tabu – K2	0.732895	0.715789	0.789562	0.924690	0.932283	0.777406	0.613455	0.707663	0.549853	0.905504	0.813478	0.741924	**9.204501**
Tabu – QNML	0.725049	0.692155	0.782548	0.924020	0.930256	0.785356	0.641712	0.716706	0.522342	0.902441	0.826567	0.750082	**9.199233**
Hill-Climbing – FNML	0.732648	0.706234	0.770505	0.925245	0.915162	0.782209	0.638295	0.723639	0.537456	0.902277	0.825826	0.736996	**9.196491**
Hill-Climbing – AIC	0.735428	0.718026	0.781621	0.922157	0.928522	0.776505	0.613798	0.719126	0.534986	0.903063	0.830765	0.723925	**9.187923**
Hill-Climbing – QNML	0.726464	0.682451	0.782597	0.922925	0.931172	0.783996	0.615221	0.717229	0.536978	0.906208	0.822829	0.741891	**9.169960**
Hill-Climbing – BDe	0.741118	0.713257	0.767841	0.927451	0.913690	0.788274	0.633798	0.737359	0.474184	0.908649	0.817215	0.744380	**9.167216**
Hill-Climbing – K2	0.734688	0.699918	0.764367	0.924755	0.929307	0.772391	0.627535	0.707319	0.535678	0.903931	0.820096	0.740523	**9.160507**
Tabu – AIC	0.732599	0.712303	0.779156	0.922092	0.933543	0.771489	0.608140	0.716281	0.528170	0.902572	0.828921	0.724618	**9.159882**
Hill-Climbing – BIC	0.720148	0.675559	0.780463	0.930686	0.923159	0.791175	0.592981	0.739976	0.520580	0.909304	0.824904	0.746572	**9.155507**
Max-Min Hill-Climbing	0.730592	0.690000	0.745029	0.926160	0.909046	0.774898	0.618344	0.736149	0.506520	0.902441	0.824640	0.758851	**9.122671**
Hill-Climbing – mBDE	0.708684	0.691003	0.758941	0.928497	0.905089	0.786848	0.624755	0.732045	0.469377	0.908747	0.818269	0.748368	**9.080622**
Tabu – BDe	0.717780	0.690872	0.767080	0.930850	0.904288	0.786126	0.626815	0.729674	0.463779	0.906863	0.809855	0.746242	**9.080224**
Tabu – mBDE	0.736003	0.681694	0.768337	0.929739	0.893936	0.787044	0.611689	0.731096	0.481725	0.907387	0.806958	0.742764	**9.078374**
Restricted Maximization	0.720477	0.662122	0.758346	0.926471	0.908163	0.772046	0.656086	0.721154	0.473690	0.906618	0.813922	0.731227	**9.050322**
Hybrid HPC	0.719243	0.689079	0.744334	0.928529	0.908605	0.767899	0.629955	0.726828	0.455267	0.908256	0.811206	0.747429	**9.036630**
Hill-Climbing – Loglik	0.641453	0.690983	0.707243	0.918612	0.889331	0.760078	0.605792	0.662410	0.532438	0.864247	0.797363	0.707817	**8.777766**
Tabu – Loglik	0.640951	0.690655	0.700178	0.915822	0.891117	0.758780	0.602654	0.665205	0.527449	0.865212	0.801603	0.695978	**8.755604**
Average	**0.717686**	**0.696737**	**0.766523**	**0.926140**	**0.917937**	**0.780782**	**0.626455**	**0.720997**	**0.520553**	**0.902452**	**0.820245**	**0.739910**	9.136417

Table 3: *Area under the ROC curves – ranked highest to lowest*

Index ticker	**Associated country**	**Underlying assets**	**Traded exchange**	**Tradable hours** (UTC)[1]	**Available timeframe**	**Data source** *[accessed 27-April-2022]*
A-VIX[a]	Australia	S&P/ASX 200	ASX (Sydney)	20:00 – 02:00	Jan 2012 – Dec 2021	https://www.investing.com/indices/s-p-asx-200-vix-historical-data
VIXI	Canada	S&P/TSX 60	TSX (Toronto)	04:30 – 11:00	Jan 2012 – Dec 2021	https://www.investing.com/indices/s-p-tsx-60-vix-historical-data
VXFXI	China	FXI	CBOE (Chicago)	04:30 – 11:00	Jan 2012 – Dec 2021	https://www.investing.com/indices/cboe-china-etf-volatility-historical-data
VCAC	France	CAC 40	Euronext Paris	11:01 – 19:30	Jan 2012 – Dec 2021	https://www.investing.com/indices/cac-40-vix-historical-data
VDAX-NEW	Germany	DAX	Deutsche Börse (Frankfurt)	10:00 – 18:30	Jan 2012 – Dec 2021	https://www.investing.com/indices/vdax-historical-data
VHSI	Hongkong	HSI	HKEX (Hongkong)	17:30 – 00:00	Jan 2012 – Dec 2021	https://www.investing.com/indices/hsi-volatility-historical-data
NIFVIX	India	NIFTY	NSE (Mumbai)	14:45 – 21:00	Jan 2012 – Dec 2021	https://www.investing.com/indices/india-vix-historical-data
JNIV[b]	Japan	Nikkei	Osaka Exchange	18:00 – 00:15	Jan 2012 – Dec 2021	https://www.investing.com/indices/nikkei-volatility-historical-data
RVI	Russia	RTSI	MOEX (Moscow)	13:00 – 21:45	Nov 2013 – Dec 2021	https://www.investing.com/indices/russian-vix-historical-data
VSMI	Switzerland	VMI	SIX Switzerland Exchange (Zurich)	11:00 – 19:30	Jan 2012 – Dec 2021	https://www.onvista.de/index/VSMI-VOLATILITAETS-Index-7911812
VFTSE	UK	FTSE 100	Euronext Amsterdam	11:01 – 19:30	Aug 2012 – Jun 2019	https://www.investing.com/indices/ftse-100-vix-historical-data
VIX	USA	S&P500	CBOE (Chicago)	04:30 – 11:00	Jan 2012 – Dec 2021	https://www.investing.com/indices/volatility-s-p-500-historical-data

Table 4: *Mapping of implied volatility data and their underlying asset to dedicated countries, including stock exchanges, tradable hours and available time frame with sources.*

R code (RStudio)

Package	Version
Base R	4.2.0
BiocGenerics	0.41.2
bnlearn	4.7.1
dplyr	1.0.8
forcats	0.5.1
ggplot2	3.3.5
gRain	1.3.9
graph	1.73.0
gRbase	1.8.7
gtools	3.9.2
...	...
kableExtra	1.3.4
pROC	1.18.0
purrr	0.3.4
readr	2.1.2
reshape2	1.4.4
Rgraphviz	2.39.1
scales	1.2.0
stringr	1.2.0
tibble	3.1.6
tidyr	1.2.0
tidyverse	1.3.1
viridis	0.6.2

Table 5: *Utilised packages and their versions.*

Code 1: k-fold cross validation with $k = 10$ to verify the fit of the algorithms and scoring functions to our data. Due to the runtime a progress bar is incorporated.

```
validate_scores <-
  function(complete.data,
           k,
           score.algorithms,
           scores,
           hybrid.algorithms) {
```

Code 2: Bootstrap approach to implement pipeline for averaged models with $n = 100$ runs including a progress bar due to the long runtime.

```
build_models <-
  function(incomplete.training_data,
           complete.training_data,
           n,
           threshold,
           score.algorithms,
           scores,
           hybrid.algorithms) {
```

Code 3: Evaluation of our learned models by comparing predictions with test data and exporting the dedicated ROC plots. Due to the runtime a progress bar is incorporated.

```
roc_plots <- function(complete.testing_data) {
```

Code 4: Inductive Causation algorithm based on Pearl *and* Meek *from Figure 14 – implemented on top of* ***bnlearn*** *package with try blocks to avoid acyclicity. To get a better insight into the algorithm, we provide exports with information about the applied rule and the associated nodes after each change in the structure. In addition, the final export provides the ratio of oriented arcs in percent, as this is one of our main interests.*

```
inductive_causation <- function(my_bn) {
```

References

[1] Borgelt, C., and Kruse, R. A critique of inductive causation. In *Symbolic and Quantitative Approaches to Reasoning and Uncertainty* (Berlin, Heidelberg, 1999), A. Hunter and S. Parsons, Eds., Springer Berlin Heidelberg, pp. 68–79.

[2] Carvalho, A. M. Scoring functions for learning bayesian networks. Tech. rep., Instituto Superior Técnico, Technical University of Lisbon, 2009.

[3] Castelletti, F., Consonni, G., Vedova, M. L. D., and Peluso, S. Learning Markov Equivalence Classes of Directed Acyclic Graphs: An Objective Bayes Approach. *Bayesian Analysis 13*, 4 (2018), 1235 – 1260.

[4] Chicharro, D., and Panzeri, S. Algorithms of causal inference for the analysis of effective connectivity among brain regions. *Frontiers in Neuroinformatics 8* (2014).

[5] Druzdzel, M. J. The role of assumptions in causal discovery. In *8th Workshop on Uncertainty Processing (WUPES-09)* (September 2009), pp. 57 – 68.

[6] Ertel, W. *Grundkurs Künstliche Intelligenz - Eine praxisorientierte Einführung*, 5th ed. Springer-Verlag, 2020.

[7] Friedman, N., Goldszmidt, M., and Wyner, A. Data analysis with bayesian networks: A bootstrap approach. In *UAI '99: Proceedings of the 15th Annual Conference on Uncertainty in Artificial Intelligence* (1999), Morgan Kaufmann, pp. 196 – 205.

[8] Glymour, C., Zhang, K., and Spirtes, P. Review of causal discovery methods based on graphical models. *Frontiers in Genetics 10* (2019).

[9] Hastie, T., Tibshirani, R., and Friedman, J. *The Elements of Statistical Learning*, 2nd ed. Springer New York, 2009.

[10] Hull, J. C. *Options, Futures, and other Derivatives*, 11th ed. Pearson, Boston, 2021.

[11] Imoto, S., Kim, S., Shimodaira, H., Aburatani, S., Tashiro, K., Kuhara, S., and Miyano, S. Bootstrap analysis of gene networks based on bayesian networks and nonparametric regression. *Genome Informatics 13* (2002), 369 – 370.

[12] Joshi, M. S. *The Concepts and Practice of Mathematical Finance*, 2nd ed. Mathematics, Finance and Risk. Cambridge University Press, 2008.

[13] Kauermann, G., Küchenhoff, H., and Heumann, C. *Statistical Foundations, Reasoning and Inference*. Springer International Publishing, 2021.

[14] Meek, C. Causal inference and causal explanation with background knowledge. In *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence* (San Francisco, CA, USA, 1995), UAI'95, Morgan Kaufmann Publishers Inc., pp. 403–410.

[15] Nagarajan, R., Scutari, M., and Lèbre, S. *Bayesian Networks in R*. Springer New York, 2014.

[16] Pearl, J. *Causality*, 2nd ed. Cambridge University Press, Cambridge, UK, 2009.

[17] Pearl, J., and Verma, T. A theory of inferred causation. In *Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning* (San Francisco, CA, USA, 1991), KR'91, Morgan Kaufmann Publishers Inc., pp. 441–452.

[18] Russell, S. J., and Norvig, P. *Artificial Intelligence – A Modern Approach*, 4th ed. Pearson, 2021.

[19] Scutari, M. Package 'bnlearn'. Tech. rep., CRAN, 2022.

[20] Scutari, M., Graafland, C. E., and Gutiérrez, J. M. Who learns better bayesian network structures: Accur-

acy and speed of structure learning algorithms. *International Journal of Approximate Reasoning 115* (2019), 235–253.

[21] Scutari, M., Vitolo, C., and Tucker, A. Learning bayesian networks from big data with greedy search: computational complexity and efficient implementation. *Statistics and Computing 29*, 5 (2019), 1095 – 1108.

[22] Shumway, R. H., and Stoffer, D. S. *Time Series Analysis and Its Applications*, 4th ed. Springer International Publishing, 2017.

[23] Spirtes, P., and Glymour, C. An algorithm for fast recovery of sparse causal graphs. *Social Science Computer Review 9*, 1 (1991), 62 – 72.