Zero Augmented Generalized Linear Models with Score Driven Parameters for Sparse Weighted Dynamical Networks (With Applications to Systemic Risk Forecasting ?)

Domenico Di Gangi

February 19, 2020

Domande aperte

- La motivazione del paper sarebbe :
  - Contribuire alla scarsa letteratura su modelli dinamici per reti sparse e pesate, unendo networks ZA and score driven models.
  - Affronta il problema della high dimensionality delle possibili strutture di dipendenza temporale con combinazione fitnesses (parametri node specific) + SD
  - Permette di disaccoppiare la propensione alle connessioni intrinseca di ogni nodo (fitnesses) dal contributo dovuto a un qualunque regressore. In un certo senso, é simile al paper sul DAR-Fitness per la persistenza. Quindi le fitnesses (binarie o pesate), quando usate congiuntamente con i regressori, potrebbe essere proposte come propensione ad avere connessioni (o connessioni di peso) che non sono spiegate dai regressori
  - Usando a validazione in sample e out of sample per model selection, consente di rispondere alla domanda: il regressore X é importante per la dinamica della rete? Quali sono le fitnesses al netto dell'effetto del regressore?
- Ha senso discutere anche il forecast della rete? i.e. combinazione di forecast binario e pesato? Forse si potrebbe fare per mostrare una possibile applicazione a financial stability.
- commenti sulla struttura?
  - non sono sicuro se aggiungere una sezione solo per i modelli score driven.
  - presento prima dinamica per parametri associati ai pesi e poi per la versione binaria. controintuitivo? Vorrei sottolineare che la dinamica SD per i pesi é la parte principale del lavoro. oltre a essere indipendente da quella dei links e quindi applicabile con altri modelli binari

Paper Structure

1. Introduction: we intend to model dynamical sparse weighted networks.

   The main challenges are: sparcity and high dimensionality.

   Review of existing literature.

   We use Zero Augmentation and Score Driven version of meaningful static models, adding the possibility of dependency on external variables.

   Paper outline.

2. Methods: Zero augmentation in general to address sparcity

   Score Driven version of static Zero Augmented model with one parameter per node to address high dimensionality. Our approach addresses the challenges mentioned and is flexible enough to allow for additional regressors (separately in links' presence and links' weights).

   Possibility to change the distribution of the weights

   Description of models and numerical tests independent of the sequence of binary probabilities. Combination with a similar approach for the binary network (SD-beta model with regressors) for a completely score driven dynamics

3. Applications: application to WTN and eMid that highlights the possibility of using standard model selection techniques (AIC, BIC, other ideas?) to compare different models. Both SD vs sequence of single snapshots, and different regressors in determining links' presence and weights for different weights distributions.

   Description of forecasting approach and out of sample validation and comparison of different models.

   Final application of forecasting debt rank in eMid?

1 Introduction

MOTIVATION FOR WEIGHTED NETWORKS

NEED FOR SPARCITY

We focus on models for **Dynamical Sparse Weighted Networks** described by sequences of matrices of positive real numbers $\left\{ Y_{ij}^{(t)} \right\}_{t=1}^{T}$.

MOTIVATION FOR DYNAMICS, DISTINCTION BETWEEN GROWTH AND "STATIONARY" (FIXED NUMBER OF NODES) MODELS

CHALLENGES AND LITERATURE REVIEW:
The goal is to model $P\left(Y_{ij}^{(t)} | \mathbf{Y}^{(t-1)}, \dots \mathbf{Y}^{(1)}\right)$. Even if we consider dependencies at one lag, the problem is extremely high dimensional. To cope with the issue of dimensionality, different approaches have been considered in the literature:

- Driven by system specific insights, the researcher selects a set of network statistics $G_i(\mathbf{Y})$ and estimates the dependency of links at time t on $G\left(\mathbf{Y}^{(t-1)}\right)$. Giraitis et al. (2016), for example, estimate a Tobit model with few regressors, for each link. Moreover they use a local-likelihood method to estimate time varying coefficients of the regression.
- Latent space models, where a set of parameters is associated to each node and an exogenous time evolution is assumed for those parameters, e.g. Sewell and Chen (2015) define a latent space Tobit
- Models that allow each one of the matrix elements $Y_{ij}^{(t)}$ to depend on each of the ${Y_{ij}}^{(t-1)}$ have also been considered in the literature:
  1. Billio et al. (2018) Estimate a tensor regression (very similar to a VAR on $vec(\mathbf{Y})$), with rank restrictions on the (huge) matrix of model's parameters . (Not clear how they take sparsity into account)
  2. Betancourt et al. (2018) Consider a penalized logistic auto-regression model for binary networks (basically a logistic regression for each link, using all lagged matrix elements, and also products, with a lasso penalization). The same approach can in principle be extended to sparse weighted networks, and in the ZA framework.

REVIEW OF SCORE DRIVEN MODELS

In order to review the score-driven models as introduced by Creal et al. (2013) and Harvey (2013), let us consider a sequence of observations $\left\{ y^{(t)} \right\}_{t=1}^{T}$, where each $y^{(t)} \in \mathbb{R}^M$, and a conditional probability density $P\left(y^{(t)} | f^{(t)}\right)$, that depends on a vector of time-varying parameters $f^{(t)} \in \mathbb{R}^K$. Defining the score as $\nabla^{(t)} = \frac{\partial \log P\left(y^{(t)} | f^{(t)}\right)}{\partial f^{(t)\prime}}$, a score-driven model assumes that the time evolution of $f^{(t)}$ is ruled by the recursive relation

$$f^{(t+1)} = w + \boldsymbol{\beta} f^{(t)} + \boldsymbol{\sigma} \boldsymbol{S}^{(t)} \nabla^{(t)}, \tag{1}$$

where w, $\boldsymbol{\sigma}$ and $\boldsymbol{\beta}$ are static parameters, w being a K dimensional vector and $\boldsymbol{\sigma}$ and $\boldsymbol{\beta}$ $K \times K$ matrices. $\boldsymbol{S}^{(t)}$ is a $K \times K$ scaling matrix, that is often chosen to be the inverse of the square root of the Fisher information matrix associated with $P\left(y^{(t)} | f^{(t)}\right)$, i.e. $\boldsymbol{S}^{(t)} = \mathbb{E}\left[\frac{\partial \log P\left(y^{(t)} | f^{(t)}\right)}{\partial f^{(t)\prime}} \frac{\partial \log P\left(y^{(t)} | f^{(t)}\right)}{\partial f^{(t)\prime}}' \right]^{-\frac{1}{2}}$.

However, this is not the only possible specification and different choices for the scaling are discussed in Creal et al. (2013).

OUR APPROACH

- **We propose** to start from a static model (in principle one that fits well the cross sectional data) and define its dynamical version using the SD approach. Applying this to the ZA-Gamma model results in the **Zero Augmented Score Driven (Gamma?) Network Model ZA-SD-Nets (or maybe weighted random graphs ?ZA-SD-WRG)**
- We consider Zero Augmented distributions to define a framework for a flexible description of **Sparse Weighted Networks**, already at the static level:

$$P\left(Y_{ij}=y\right)=\begin{cases}\left(1-p_{ij}\right) & for \quad y=0\\ \\ p_{ij}\, g\left(y|\mu_{ij},\sigma_{ij}\right) & for \quad y>0\end{cases}, \tag{2}$$

  where $0<p_{ij}<1$ and $g\left(y\right)$ is the density for a positive continuous random variable, with link specific parameters μ_{ij} and λ_{ij}. Motivation :

  - The main motivation for the ZA approach would be the possibility to decouple the probability of a link's presence from the distribution of it's weight. This seems to allow a greater flexibility compared to Tobit models. Flexibility that we intend to fully exploit in the **Dynamical** context.
  - The simplest choice for g would be the exponential. We consider the gamma because it allows to adjust the dispersion around the mean . Moreover, with the gamma, our approach includes, as a limiting case, the method of Cimini et al. (2014). This is interesting, since the latter has been found to be a good solution in reconstructing financial networks (see the wide comparative study of Anand et al., 2017). Additionally a recent maximum entropy model (cite the latest paper from Lucca that proposes once more an ensemble "physically" motivated) results in a ZA exponential. [1]

- Our approach combines the ZA distribution for static networks with the SD approach to the definition of model with dynamical parameters

  - It "decouples" the modeling of the weights, and their dynamics, from that of the links' presence. Basically we can use it in combination with any model for dynamical binary networks, e.g. constant (uniform or link specific) probabilities, ERGMs, SD-ERGM, neural networks based models[2]
  - It can be used as DGP to simulate realistic (in what sense?) dynamical sparse weighted networks

[1] More alternatives considered in the literature:

1. Rastelli (2018) proposes a Sparse Latent Position Model (SLPM), modifies the original LPM to account for nonnegative weighted edges using finite mixtures of exponential distributions. (Only Static)
2. A number of papers that model static weighted networks and do not take into account sparsity.
3. Models for dynamical weighted networks, cited in the following.

[2] Unfortunately this stream of literature has not been noticed in many paper on temporal networks so far (see, for example Perozzi et al., 2014; Trivedi et al., 2018; Goyal et al., 2018, 2019; Singer et al., 2019).

- It can be easily estimated via maximum likelihood (show it by simulating it as a DGP and estimating the static parameters) and can be used as a filter in case of misspecified dynamics (by filtering known DGPs for the time varying parameters or simulating a completely different network model and using it for forecast on synthetic data?)
- When estimated on data, copes with the challenge of dimensionality and allows to forecast links.
- It allows also to include and evaluate the contribution of exogenous regressors[3].
- As final application, and proof of flexibility, we can show how to use the framework for forecast of systemic risk from partial information.

2 Methods: Score Driven ZA Distributions for Sparse Weighted Dynamical Networks

We propose to describe the network with a set of independent random variables Y_{ij}, one for each link, and to model separately the probability of observing a link $\Theta(Y_{ij})$ and the probability to observe a specific weight $Y_{ij}^{(t)}$ conditional on the presence of that link. Although in principle there is nothing preventing us from considering a non trivial dependence among the random variables associated to each link, we decided for simplicity, to assume independence in the first version of this framework.

The idea of using Zero Augmentation, also indicated as Zero Inflation, to model sparse weighted networks has been advocated in many papers describing static networks .... but we are aware of no dynamical models for time varying networks using it.

In a general way a ZA description of a sparse network consists in assuming a distribution as follows:

$$P\left(Y_{ij}=y\right)=\begin{cases}\left(1-p_{ij}\right) & for \quad y=0\\ \\ p_{ij}g_{ij}\left(y\right) & for \quad y>0\end{cases} \tag{3}$$

While the conditional distribution can be arbitrary, in the following we will focus on two concrete examples: a gamma distribution, with density[4]

$$g_{ij}\left(y\right)=\frac{\left(\mu_{ij}\right)^{-\sigma_{ij}}\,y^{\left(\sigma_{ij}-1\right)}}{\Gamma\left(\sigma_{ij}\right)}e^{-\frac{y}{\mu_{ij}}}, \tag{4}$$

and a lognormal

$$g_{ij}\left(y\right)=\frac{1}{y\sigma_{ij}\sqrt{2\pi}}e^{\frac{-\left(\ln y-\mu_{ij}\right)^2}{2\sigma_{ij}^2}} \tag{5}$$

All the methods proposed in the following can be easily extended to consider different continuous distributions as well as discrete ones. We focus on these two examples for the sake of exposition. In Appendix 4 we describe our flexible python code[5] that can be used with different distributions, provided that the log-likelihood and a sampling method are available.

[3] or features that can be extracted using deep learning techniques for dimensionality reduction

[4] Note that, in the attempt of keeping the notation similar to the log-normal case, we use a somewhat unusual notation indicating the shape parameter for the Gamma distribution with the letter σ. This parameter is often indicated with an σ or a k in the literature.

[5] Publicly available at

We associate four parameters $(\overleftarrow{\varphi}_i, \overrightarrow{\varphi}_i, \overleftarrow{\eta}_i, \overrightarrow{\eta}_i)$ to each node i, and relate them with the cross sectional structure of the distribution via the conditional expectation matrix as follows

$$E\left[Y_{ij}|y>0\right] = e^{\left(\overleftarrow{\varphi}_i + \overrightarrow{\varphi}_j\right)}, \tag{6}$$

hence

$$E\left[Y_{ij}\right] = p_{ij} e^{\left(\overleftarrow{\varphi}_i + \overrightarrow{\varphi}_j\right)}.$$

This choice fixes the following relations between $(\overleftarrow{\varphi}_i, \overrightarrow{\varphi}_i)$ and the parameters μ_{ij} of the gamma

$$\mu_{ij} = \frac{1}{\sigma_{ij}} e^{\left(\overleftarrow{\varphi}_i + \overrightarrow{\varphi}_j\right)}, \tag{7}$$

while for the lognormal we have

$$\mu_{ij} = \overleftarrow{\varphi}_i + \overrightarrow{\varphi}_j - \frac{\sigma_{ij}^2}{2}. \tag{8}$$

For both distributions we consider

$$\sigma_{ij} = e^{\left(\overleftarrow{\eta}_i + \overrightarrow{\eta}_j\right)}, \tag{9}$$

The choice of the exponential link between μ and φ is standard in the context of generalized linear models. This choice allows the parameters to be unbounded[6], and is the standard approach in generalized linear models to take into account the dependence on external variables X_{ij}. Indeed we will also consider the following specification[7]:

$$E\left[Y_{ij}|y>0\right] = e^{\left(\overleftarrow{\varphi}_i + \overrightarrow{\varphi}_j + \beta_{ij} X_{ij}\right)}, \tag{11}$$

where β_{ij} can be either composition of nodes' specific parameters $\beta_i + \beta_j$ (better $\beta_i \beta_j$?), or equal for all links $\beta_{ij} = \beta$.

A general dynamical version of the models in previous section is

$$P\left(Y_{ij}^{(t)} = y\right) = \begin{cases} \left(1 - p_{ij}^{(t)}\right) & for \quad y = 0 \\ p_{ij}^{(t)} g_{ij}^{(t)}(y) & for \quad y > 0 \end{cases} \tag{12}$$

where we allowed both the probability of observing a link and the conditional distribution of the weights to depend on time. We are going to discuss in detail a proposal for the temporal dependency of the previous equation. In this section we are going to focus on the distribution conditional on the link being present and in the next we will discuss the dynamics of $p^{(t)}$.

[6] That is very convenient since we are going to turn them into dynamical ones, and to keep the dynamics into bounded regions turns out to be complicated in this setting.

[7] In some applications to real data with large heterogeneity in the external variables, or the observed weights, we found helpful the following additional transformation

$$E\left[Y_{ij}|y>0\right] = e^{f_{LU}\left(\overleftarrow{\varphi}_i + \overrightarrow{\varphi}_j + \beta_{ij} X_{ij}\right)}, \tag{10}$$

where $f_{LU}(x) = \frac{L-U}{2} \tanh\left(\frac{2x-L-U}{L-U}\right) - \frac{L-U}{2} \tanh\left(-\frac{L-U}{L+U}\right)$ is a function that softly bounds its argument between L and U. This helps to avoid numerical overflow issues. In the following we omit this soft bound but, in the applications where it is used, we take it into account also in the computations of the scores, using the chain rule and the derivative

$$\frac{\partial f_{LU}}{\partial x} = 1 - \tanh^2\left(2\frac{x-L}{L-U} + 1\right)$$

2.1 Score-Driven Dynamics for the Weights

We propose to apply the score-driven methodology to the Zero Augmented Distributions for static networks, defined in the previous section. This approach allows us to promote any of the parameters $(\overleftarrow{\varphi}, \overrightarrow{\varphi}, \overleftarrow{\eta}, \overrightarrow{\eta})$, defined in the previous section, of the conditional density g in eq. (3) to dynamical ones. The resulting $\left(\overleftarrow{\varphi}^{(t)}, \overrightarrow{\varphi}^{(t)}, \overleftarrow{\eta}^{(t)}, \overrightarrow{\eta}^{(t)}\right)$ will have a stochastic evolution driven by the score of the static model, computed at different points in time according to eq. (1).

Our approach results is a framework for the description of sparse weighted dynamical networks, more than in a single model. In fact, for each choice of the conditional distribution we obtain a different model. We refer to this class as Score-Driven Zero Augmented Generalized Linear Models .

We point out that the model is well defined whatever the $p_{ij}^{(t)}$ are. Those can be obtained by any model for binary networks and then plugged into the previous definition. We describe the binary probabilities in terms of positive parameters $\pi_{ij}^{(t)}$ defined by

$$p_{ij}^{(t)} = \frac{1}{1+\pi_{ij}^{(t)}}$$

The loglikelihood of the static model in eq. (3) for a single observation $\mathbf{Y}$, omitting the temporal dependency, is

$$\log P\left(\mathbf{Y}|\mathbf{p},\mu\right) = \sum_{ij} -\log\left(1+\pi_{ij}\right) + \Theta\left(y_{ij}\right)\left[\log \pi_{ij} + \log g_{ij}\right] \tag{13}$$

and the score

$$\overleftarrow{s}_i\left(\varphi, \mathbf{Y}^{(t)}, \sigma\right) = \frac{\partial \log P}{\partial \overleftarrow{\varphi}_i} = \sum_j \Theta\left(y_{ij}^{(t)}\right)\frac{\partial \log g_{ij}}{\partial \overleftarrow{\varphi}_i}. \tag{14}$$

Given the observations $\left\{Y_{ij}^{(t)}\right\}_{t=1}^{T}$, we can apply the update rule in (1) to all or some elements of $\boldsymbol{\mu}$ or $\boldsymbol{\sigma}$. In this paper we will consider constant $\boldsymbol{\sigma}$ and

From the form of the loglikelihood for the Gamma[8] it follows that the score is

$$\overleftarrow{s}_i\left(\varphi, \mathbf{Y}^{(t)}, \sigma\right) = \sum_j \Theta\left(y_{ij}\right)\left(\frac{y_{ij}}{\mu_{ij}} - \sigma_{ij}\right) = \sum_j \Theta\left(y_{ij}\right)\frac{y_{ij}}{\sigma_{ij}E\left[y_{ij}|y_{ij}>0\right]} - \sigma_{ij}K_i, \tag{15}$$

[9] while for the log-Normal [10] we have

$$s_i\left(\varphi, \mathbf{Y}^{(t)}, \sigma\right) = \sum_j \frac{1}{\sigma_{ij}^2}\log\frac{E\left[y_{ij}|y_{ij}>0\right]}{y_{ij}} - K_i. \tag{16}$$

[8]The log likelihood for a gamma distribution is

$$\log g\left(y|\mu_{ij},\sigma_{ij}\right) = \left(\sigma_{ij}-1\right)\log y_{ij} - \log\Gamma\left(\sigma_{ij}\right) - \sigma_{ij}\log\mu_{ij} - \frac{y_{ij}}{\mu_{ij}}$$

[9]Setting the score equal to zero, we obtain the following

$$\varphi_i = \log\left(\frac{\sum_j \frac{y_{ij}\sigma_j}{e^{\varphi_j}}}{\sum_j \Theta\left(y_{ij}\right)\sigma_j}\right)$$

[10]The log-likelihood for the log-Normal is

$$\log g\left(y|\mu_{ij},\sigma_{ij}\right) = -\log y_{ij} - \log\sigma_{ij} - \frac{1}{2}\log 2\pi - \frac{\left(\log y_{ij}-\mu_{ij}\right)^2}{2\sigma_{ij}^2}.$$

In both cases the update rule is[11]

$$\overleftarrow{\varphi}_i^{(t+1)} = \overleftarrow{w}_i + \overleftarrow{B}_i \overleftarrow{\varphi}_i^{(t)} + \overleftarrow{A}_i \overleftarrow{s}_i \left(\varphi, \mathbf{Y}^{(t)}, \sigma\right) \tag{19}$$

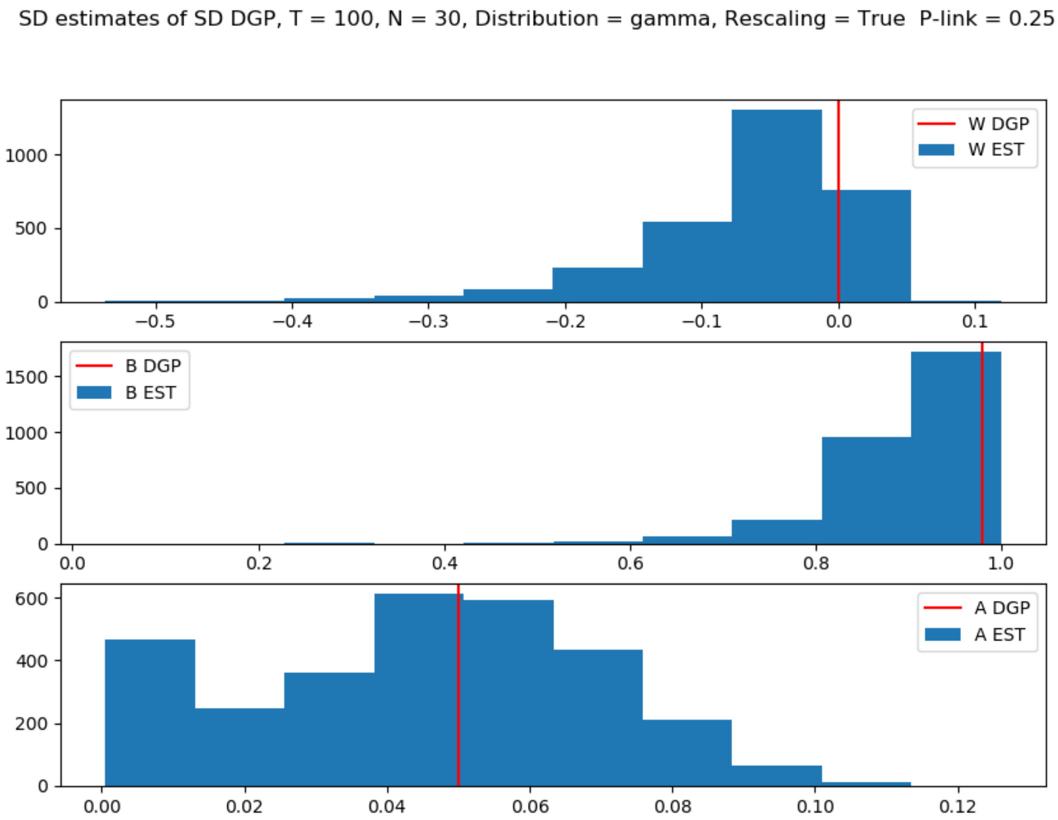


Figure 1: Some issues with the lognormal estimate.

2.2 Filter Misspecified Temporal Evolution

2.3 Dynamics of the Binary Network: Score Driven Beta Model with Explanatory Variables

The weights dynamics in our approach is independent from the binary dynamics, and we can combine any sequence of $\left\{p^{(t)}\right\}_{t=1}^{T}$ with our proposed model. Nevertheless in the following we consider a specific example of binary dynamics, the Score Driven beta model, that is very much related with the one introduced in.. for the weights. We complement the model presented in

[11]We could also consider

$$\overleftarrow{\varphi}_i^{(t+1)} = \overleftarrow{w^0}_i + \overleftarrow{w^1}_i \overleftarrow{Z}_i^{(t)} + \overleftarrow{B}_i \overleftarrow{\varphi}_i^{(t)} + \overleftarrow{A}_i \overleftarrow{s}_i \left(\varphi, \mathbf{Y}^{(t)}, \sigma\right) \tag{17}$$

where $\overleftarrow{Z}_i^{(t)}$ is an additional external variable. We could use this extra flexibility to model a trend in the dynamics of the time varying parameters, hence we will consider:

$$\overleftarrow{\varphi}_i^{(t+1)} = \overleftarrow{w^0}_i + \overleftarrow{w^1}_i t + \overleftarrow{B}_i \overleftarrow{\varphi}_i^{(t)} + \overleftarrow{A}_i \overleftarrow{s}_i \left(\varphi, \mathbf{Y}^{(t)}, \sigma\right) \tag{18}$$

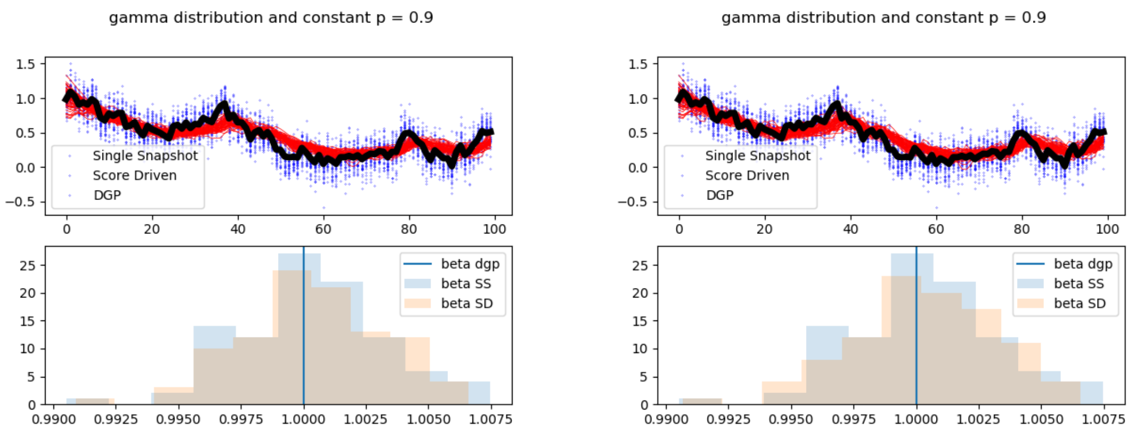


Figure 2: No rescaling left, with rescaling right. The average RMSE over 50 Simulations is 0.15, 0.133, 0.124, respectively for SS, SD-NO-RESC, SD-RESC

(cite our paper) allowing for binary probabilities to depend on additional regressors, alongside two parameters for each node.

The definition of the static beta model reduces to assuming that the probability for the presence of each link can be written as

$$p_{ij} = \frac{1}{1 + e^{-\overleftarrow{\theta}_i - \overrightarrow{\theta}_j}}$$

hence, with the notation used so far $\pi_{ij} = e^{-\overleftarrow{\theta}_i - \overrightarrow{\theta}_j}$. With the aim of increasing the flexibility of the model we consider (propose?) a simple extension that allows the probability of observing a link to additionally depend on link specific regressors:

$$p_{ij} = \frac{1}{1 + e^{-(\overleftarrow{\theta}_i + \overrightarrow{\theta}_j + X_{ij}\beta_{ij})}}$$

In this model, there are no impediments in writing the explicit dependence of the likelihood on the parameters $\overleftarrow{\theta}$ and $\overrightarrow{\theta}$, that can be found in (**??**). we can consider the SD-beta model that allows parameters $\overleftarrow{\theta}$ and $\overrightarrow{\theta}$ to follow a SD dynamics.

3 Applications

In this section we show that our approach can be easily applied to real world networks and discuss how the flexibility that it allows can be exploited to gain new insight into their dynamics. We consider the publicly available data on international trade hence focusing on the dynamical network of bilateral trades between countries. As a second example we consider data from eMid: a section of the European (mostly Italian) interbank market.

The first aim of this applications is to use our model to "test" for the relevance of different effects in determining the dynamics of links' weights. In doing so we refer to standard approaches to model selection (refs to general model selection AIC and BIC) applied to choose among different versions of the model proposed in the previous sections.

We conclude this section with a discussion of how our zero augmented dynamical model can be applied to the forecast of future links' presence and weights. We then use these methods to compare the out of sample performances of different models.

3.1 Model Selection For Real Networks

STATIC VS DYNAMICAL?
SINGLE SNAPSHOTS VS SCORE DRIVEN ?
REGRESSORS?
DISTRIBUTION CHOICE? ONLY FOR WEIGHTED

3.1.1 Out of Sample Evaluation: Forecasting Sparse weighted Networks

Once again we stress that the dynamics of links' weights is independent from that of the binary network, hence the model proposed so far can be combined with any other method for the forecast of binary networks. Ideally we can evaluate the best performing model for link prediction in binary networks (as discussed for example in Haghani and Keyvanpour, 2017), and combine it with our score driven approach for the dynamics of the weights.

For the purposes of links' forecasting and /or out of sample evaluation, we have different alternatives :

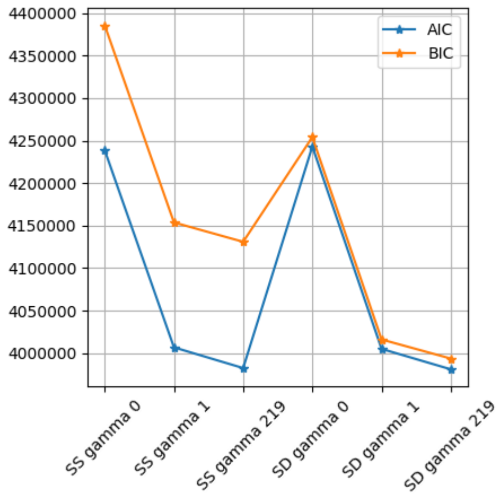
1. forecast $Y_{ij}^{(t)}$ with $E\left[Y_{ij}^{(t)}|Y_{ij}^{(t-1)}\right]$ (forecast a dense, probably always fully connected, network)
2. given the $p_{ij}^{(t)}$ obtain a single network $A_{ij}^{(t)}\left(p_{ij}^{(t)}\right)$ and only for the non zero elements compute $E\left[Y_{ij}^{(t)}|y>0\right]$ [12]
3. We can repeat the previous point for each point on the ROC curve and maybe average? what about the error for the false positive links?
4. repeat sample of the ZA distribution, compute one measure of forecast error each time and then average. Useful for model selection not for actual forecast.

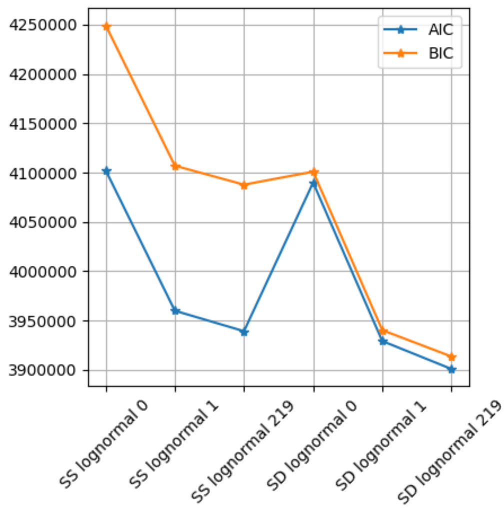
3.2 World Trade Network

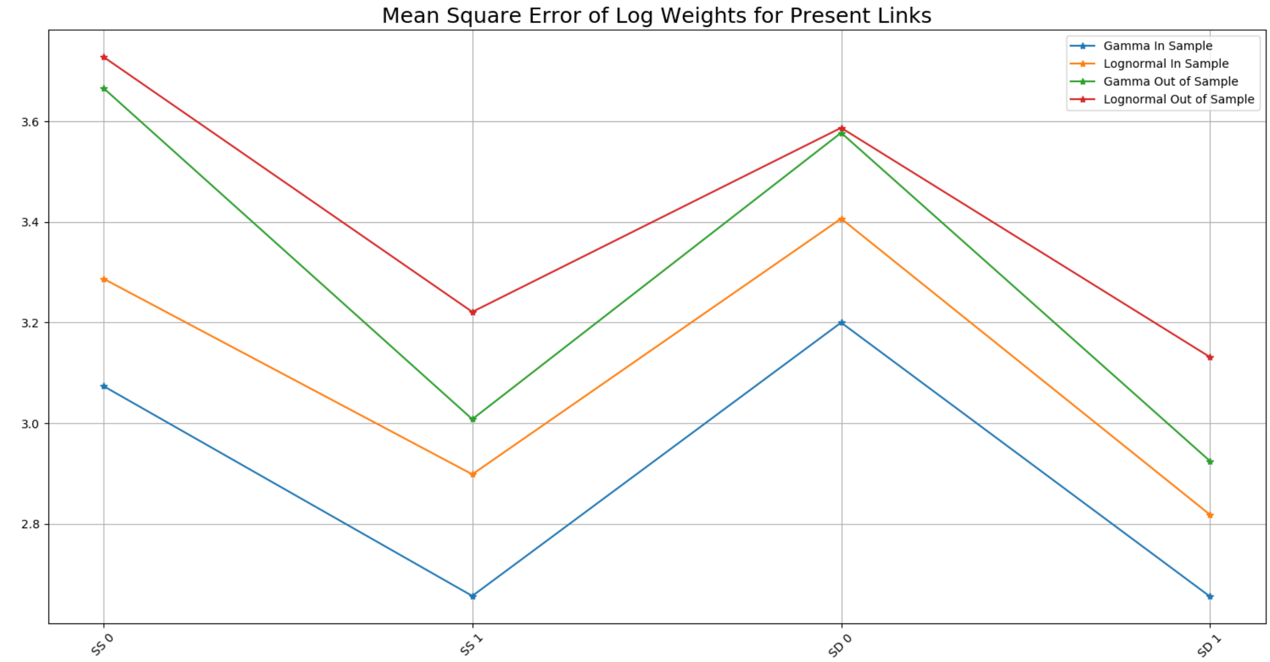
DATASET DESCRIPTION, LITERATURE REVIEW

3.2.1 Weighted

- $T = 56$, $T_{train} = 46$, out of sample tests on the remaining 10 time points
- In smaple:
  - likelihood ratios, for train sample strongly suggests models with more parameters: Single Snapshot over Score driven versions, and inclusion of regressors
  - AIC and BIC both suggest Score Driven models over Single Snapshot, log-normal over gamma and regressors (with more parameters) over non regressors







3.2.2 Binary

3.3 Italian Interbank Market

DATASET DESCRIPTION, LITERATURE REVIEW
TEST A LINK PERSISTENCY TERM in the logit
TEST LIBOR EFFECT IN THE LOGIT
Consider also the possibility of adding node specific regressors in the score driven update equation, e.g. the libor interest rate for eMid data

[12]Another possibility would be to use the SD update to compute $\varphi^{(t)}$ (using only observations at $t-1$), compute the expected strengths sequence (depends on $A_{ij}^{(t)}\left(p_{ij}^{(t)}\right)$) and then distribute the weights using RAS. A criterium would be needed for the extra weights. This method might allow us to take into account the fact that often banks split their weights in simple fractions (e.g. one third and two thirds): potremmo verificare quali banche nel passato mostrano propensione a divisione in frazioni ben definite e dare prioritá a queste nell'allocare i pesi. alloco prima i le piú grandi con maggiore propensione a integer division e poi il resto.

4 Appendix A: DynWNets, An Open Source Repository

The python code is available at `https://github.com/domenicodigangi/DynWNets`
A VERY BRIEF DESCRIPTION OF THE REPOSITORY AND ONE EXAMPLE SCRIPT FOLLOWED BY A DISCUSSION OF HYPERPARAMETERS TUNING AND THE USE OF PYTORCH AND BACKPROP

References

Anand, K., I. van Lelyveld, dm Banai, S. Friedrich, R. Garratt, G. Haaj, J. Fique, I. Hansen, S. M. Jaramillo, H. Lee, J. L. Molina-Borboa, S. Nobili, S. Rajan, D. Salakhova, T. C. Silva, L. Silvestri, and S. R. S. de Souza (2017). The missing links: A global study on uncovering financial network structures from partial data. *Journal of Financial Stability*.

Betancourt, B., A. Rodrguez, and N. Boyd (2018). Investigating competition in financial markets: a sparse autologistic model for dynamic network data. *Journal of Applied Statistics 45*(7), 1157–1172.

Billio, M., R. Casarin, and M. Iacopini (2018). Bayesian tensor regression models. In *Mathematical and Statistical Methods for Actuarial Sciences and Finance*, pp. 149–153. Springer.

Cimini, G., T. Squartini, N. Musmeci, M. Puliga, A. Gabrielli, D. Garlaschelli, S. Battiston, and G. Caldarelli (2014). Reconstructing topological properties of complex networks using the fitness model. In *Social Informatics*, pp. 323–333. Springer International Publishing.

Creal, D., S. J. Koopman, and A. Lucas (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics 28*(5), 777–795.

Giraitis, L., G. Kapetanios, A. Wetherilt, and F. Žikeš (2016). Estimating the dynamics and persistence of financial networks, with an application to the sterling money market. *Journal of Applied Econometrics 31*(1), 58–84.

Goyal, P., S. R. Chhetri, and A. Canedo (2019). dyngraph2vec: Capturing network dynamics using dynamic graph representation learning. *Knowledge-Based Systems*.

Goyal, P., N. Kamra, X. He, and Y. Liu (2018). Dyngem: Deep embedding method for dynamic graphs. *arXiv preprint arXiv:1805.11273*.

Haghani, S. and M. R. Keyvanpour (2017). A systemic analysis of link prediction in social network. *Artificial Intelligence Review*, 1–35.

Harvey, A. C. (2013). *Dynamic Models for Volatility and Heavy Tails: With Applications to Financial and Economic Time Series*. Econometric Society Monographs. Cambridge University Press.

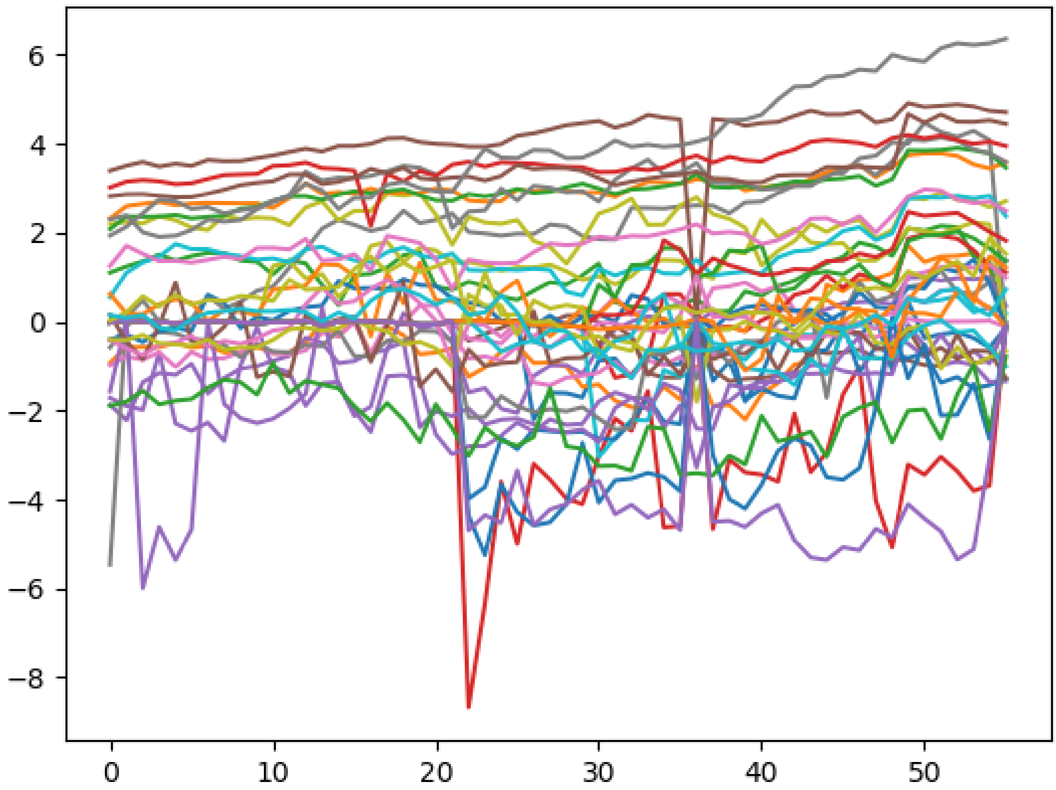
Perozzi, B., R. Al-Rfou, and S. Skiena (2014). Deepwalk: Online learning of social representations. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 701–710. ACM.

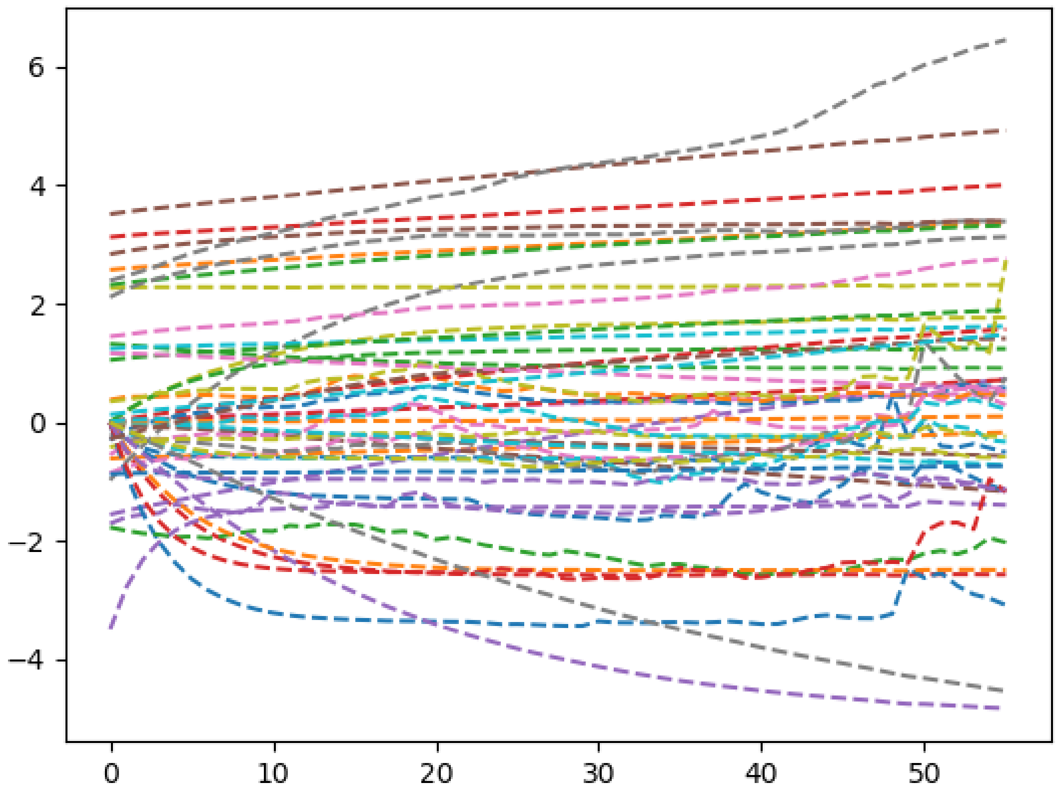
Rastelli, R. (2018). The sparse latent position model for nonnegative weighted networks. *arXiv preprint arXiv:1808.09262*.

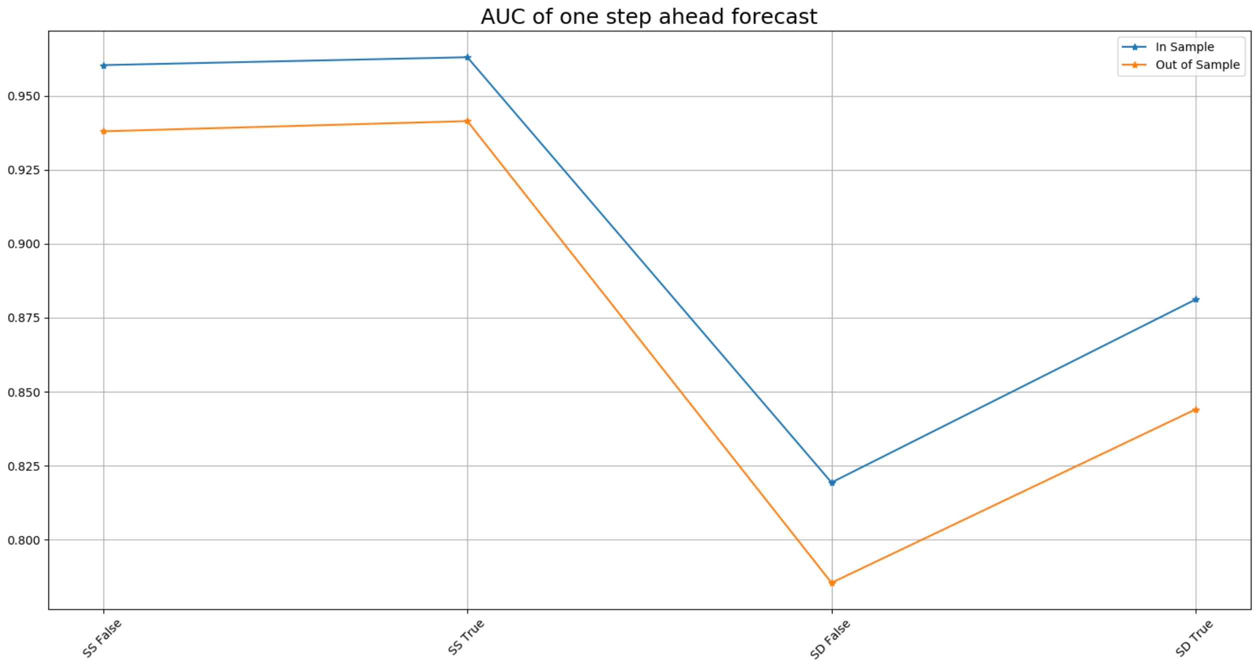
Sewell, D. K. and Y. Chen (2015). Latent space models for dynamic networks. *Journal of the American Statistical Association 110*(512), 1646–1657.

Singer, U., I. Guy, and K. Radinsky (2019). Node embedding over temporal graphs. *arXiv preprint arXiv:1903.08889*.

Trivedi, R., M. Farajtabar, P. Biswal, and H. Zha (2018). Representation learning over dynamic graphs. *arXiv preprint arXiv:1803.04051*.


6
4
2
0
−2
−4
−6
−8
0
10
20
30
40
50


6
4
2
0
−2
−4
0
10
20
30
40
50



AUC of one step ahead forecast
In Sample
Out of Sample
0.950
0.925
0.900
0.875
0.850
0.825
0.800
SS False
SS True
SD False
SD True