

Score Driven Generalized Fitness Model for Sparse and Weighted Temporal Networks

Domenico Di Gangi^{a,*}, Giacomo Borgetti^b, Fabrizio Lillo^{c,b}

^a*ISTI-CNR, via G. Moruzzi 1, 56124 Pisa*

^b*Department of Mathematics, University of Bologna, Piazza di Porta San Donato 5, 40126 Bologna, Italy*

^c*Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126, Pisa, Italy*

Abstract

While the vast majority of the literature on models for temporal networks focuses on binary graphs, often one can associate a weight to each link. In such cases the data are better described by a weighted, or valued, network. An important well known fact is that real world weighted networks are typically sparse. We propose a novel time varying parameter model for sparse and weighted temporal networks as a combination of the fitness model, appropriately extended, and the score driven framework. We consider a zero augmented generalized linear model to handle the weights and an observation driven approach to describe time varying parameters. The result is a flexible approach where the probability of a link to exist is independent from its expected weight. This represents a crucial difference with alternative specifications proposed in the recent literature, with relevant implications for the flexibility of the model. Our approach also accommodates for the dependence of the network dynamics on external variables. We present a link forecasting analysis to data describing the overnight exposures in the Euro interbank market and investigate whether the influence of EONIA rates on the interbank network dynamics has changed over time.

Keywords: Temporal Networks, Weighted Networks, Score Driven Models, Interbank Market

1. Introduction

In the last two decades, networks, or graphs, have attracted an enormous amount of attention as an effective way of describing pairwise relations in complex systems ([1]). The ever increasing abundance, and variety of graph data has motivated a great deal of applications of statistical models to graphs [see,

*Corresponding author: digangidomenico@gmail.com

for example 2, for a review]. More recently, the availability of time varying networks' data has stimulated the development of models for temporal networks [3, 4, 5, 6, 7]. While the vast majority of this literature focuses on *binary graphs*, i.e. graphs that are defined solely by a set of nodes and a set of links between pairs of nodes, often one can associate a weight to each link. Links' weights are typically positive, discrete or continuous, numbers and can be associated, for instance, to the strength of the relation described by each link. In standard binary descriptions, such relevant information is completely lost. For example, in a network of exposures among financial institutions the weight could be the value of the credit. In this case, a binary network would describe in the same way a link associated to an exposure of 1 million to that associated to an exposure of 1 billion. Indeed it is very common for network data to have also informative weights associated with their links. Some additional examples of the importance of weights in networks can be found, for example in: the International Trade Network [8, 9], migration flows [10], scientific collaborations [11], transportation networks [12], just to mention a few. In these cases the data are better described by a weighted, or valued, network, that can be associated with a positive, real valued matrix $Y_{ij} \in \mathcal{R}^+$. Y_{ij} is the weight of the directed link from node i to node j , and $Y_{ij} = 0$ if the link is not present. Moreover, it is well known that real world networks, both binary and weighted, are very often found to be sparse, i.e. their adjacency matrices have an abundance of zero entries. That is the case, for example, of interbank networks ([13]), a class of weighted temporal networks of paramount importance, that are known to be extremely relevant to financial stability ([14, 15]), and have motivated the application and development of a number of statistical models for networks (see, for example, [16, 6] and references therein).

In spite of their relevance, networks' weights have received less attention in the literature on models for temporal networks. Indeed there are only a few models for temporal networks that take them into account. In this paper we propose a novel model for sparse and weighted temporal networks, that also accommodates for the dependency of the network dynamics on external covariates. Our efforts are originally motivated by the need to properly model weighted temporal network data, describing overnight exposures in the Euro interbank market (eMID). In a previous work [7], we disregarded the information associated with the links' weight and we only focused on the temporal evolution of binary relations. We achieved that by extending the Exponential Random Graph models (ERGM), a class of statistical models for random networks, whose first and probably most famous example is the Erdős-Rényi model [17]. In the novel framework, the ERGM parameters change over time following an observation-driven dynamics [18]. The relevant information for the time evolution is encoded in the filtration \mathcal{F}_t . It

determines the update of the time varying parameters through an autoregressive process whose innovation term corresponds to the scaled score of the observation probability mass function. Such specification, known as Dynamic Conditional Score-driven (DCS) or Generalized Autoregressive Score (GAS), has been pioneered and recently introduced in the econometric literature by [19, 20]. In this paper, we extend the well known fitness model [21, 22, 23, 24, 25, 26] for static binary networks, combining it with a simple generalized linear model to handle the weights and the DCS approach to time varying parameters. The generalization is non trivial. We need to explicitly account for the abundance of zeros that follow from the sparse nature of real world networks. We solve the issue by resorting to zero augmentation (ZA). The resulting modeling framework is very general and extremely flexible. It allows to decouple the probability of a link to exist from its expected weight – a fact that will prove to be crucial in the forecasting exercise – leaving us full flexibility concerning the specification of the weight distribution and the possibility to explore the influence of external covariates on the network’s dynamics. We provide convincing Monte Carlo evidence that the score driven model is an effective filter of the latent fitness dynamics in miss-specified settings. We document a clear computational advantage, in terms of mean squared error and mean absolute difference, with respect to competitor models also in presence of external covariates and omitted variables. Interestingly, by we exploitIn the empirical analysis, we explore from different perspectives the role played by the Euro Overnight Index Average (EONIA) rates on the dynamics of the lending relations. Consistently with similar findings discussed in the literature, we observe that lower interest rates are related with a reduction of network interconnectdness but an increase of the average liquidity flow for the loans that are present. However, the novelty of our results rely on two important facts: i) the time varying fitness model accounts explicitly for bank specific effects and thus provides a measure of the impact of EONIA rates decoupled from node specific effects; ii) our results leverage the full information available in the description of the eMID network without the need to collapse the network matrices into a single statistic, as done in [27] and [28]. Finally, concerning the link and weight persistence analysis, we complement the work of [29] highlighting the tendency of banks to form links whose weight is positively related with the weights at previous time steps. We call weight persistence this aspect of the dynamics of interbank relations.

The rest of the paper is organized as follows. The next section provides an overview of the main contributions proposed in the literature for the modeling of temporal weighted networks. Section 2 defines a novel observation driven model for sparse and weighted temporal network, generalizing the well-known fitness model for binary networks to the sparse weighted case and leveraging the score driven

approach to time varying parameters. Section 3 presents an extensive Monte Carlo analysis of the score driven generalized fitness model. Section 4 details the application of the new model to the eMID market. Section 5 draws the main conclusions and provides some ideas for future research.

1.1. Literature Review

An econometric model for weighted temporal networks can be described in terms of a probability distribution $P\left(Y_{ij}^{(t)}|\mathbf{Y}^{(t-1)}, \dots, \mathbf{Y}^{(1)}, \mathbf{X}_1^{(t)}, \dots, \mathbf{X}_K^{(t)}\right)$ that describes the probability of the weight $Y_{ij}^{(t)}$ of the link between node i and j at time t as potentially depending on previous realizations of the network and a set of contemporaneous, matrix valued, external variables $\mathbf{X}_1 \dots \mathbf{X}_K$. The external variables can be different for each link. Even disregarding the dependency on external variables, it is immediately evident that the matrix valued nature of networks' data implies that the problem is typically very high dimensional.

Formally, observations of a weighted temporal network $Y_{ij}^{(t)}$ are no different from balanced panel data, where the cross sectional index ij runs over all possible N^2 links. Thus, one could directly apply the non-linear panel regression methods [described, for example, in 30] to the sequence of positive valued matrix observations and estimate, for example, the effect of external variables. Such a direct application would nevertheless disregard the network structure and the fact that links observations are very much influenced by their association with specific nodes. Moreover, the high dimensional nature of the problem complicates the estimation and interpretation of link specific fixed effects. For this reasons, the direct approach of treating sequences of networks as standard panel data, possibly with lagged dependency, is not widespread in the literature, and restrictions to reduce the number of parameters to be estimated are used in most cases. For example, [5], driven by system specific insights, select a set of network statistics $G_i(\mathbf{Y})$ and estimate the dependency of each link $Y_{ij}^{(t)}$ on $G(\mathbf{Y}^{(t-1)})$, by means of a Tobit model with few covariates, for each link. Moreover they use a local-likelihood method to estimate time varying coefficients of the regression. The censored regression used in their paper has the downside of requiring a joint modeling of the presence of a link and of its weight. While this choice allows for straightforward estimates and builds on the well known Tobit regression, it models jointly the effect that a covariate has on the probability of observing a link and the one that it has on its expected weight. In this work we propose a more flexible method that, among other things, allows to disentangle the probability of a link being present from its expected weight.

Another widespread approach is that of dynamical latent space models ([31]) where a set of time

varying parameters is associated to each node, and, at each time, the probability of observing a link between two nodes depends on a measure of distance between two nodes in such a latent space. Latent space models have also been considered for sparse weighted networks [for example in 4], with the aim of inferring from the network’s dynamics the positions of nodes in a latent space. The resulting embedding of the nodes is typically analyzed and compared with available metadata on the nodes. While very informative, the analysis of networks’ embeddings is fundamentally different from the purpose of our work, as one of main aims is to leverage the availability of data on external variables that are expected to be related with network dynamics and estimate how much the latter depends on them.

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. For example, [32] estimate a tensor regression (very similar to a VAR on $vec(\mathbf{Y})$), with rank restrictions on the (huge) matrix of model’s parameters. Differently from our work, they do not take the sparse nature of networks explicitly into account. [33] 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.

We conclude this literature review citing some contributions that are relevant to the present work, even if they do not address directly the temporal evolution of weighted networks. Among the many papers that address the issue of modeling static sparse weighted networks, [34] consider a combination of the fitness and the gravity model of [35] to reconstruct sparse weighted financial networks. Even if they do not investigate temporal networks, their approach has proved to be very effective in modeling sparse weighted networks describing financial systems [as discussed in 36]. Another interesting recent contribution is [37], where the authors propose a modelling framework for weighted network data based on the compound Poisson model. They incorporate the binary fitness model as a special case but use an additional set of fitness to describe the distribution of the weights, in addition to the probability of a link being present.

Finally, we mention three recent contributions related to our work that model binary temporal networks. The first one is that of [6], where the authors consider a model with time varying fitness and combine it with Discrete Auto-regressive Models [38] to investigate link persistence in financial networks. Their approach is very much related to ours as it extends the binary fitness model to the temporal domain, by allowing the fitness to evolve in time. In their case the fitness follows a parameter driven dynamics. Moreover they consider a mechanism to copy links from the past, and explore the possibility

of decoupling the link persistence implied by this mechanism from the probability to form new connections captured by the fitness. The second one is the recent [39] which introduces a non-Markovian model of binary temporal networks based on an extension of Discrete Autoregressive Models. Interestingly, authors considers also a local likelihood estimation approach to take into account non-stationary and infer time varying parameters. The third paper, anticipated in the Introduction, is [7] where we extend the ERGMs for binary networks – a class which includes the fitness model – to the temporal context by allowing its parameters to evolve in time. The updating equation of the time varying parameters follows an autoregressive process whose innovation corresponds to the scaled score of the observation distribution. We show how to exploit the model as an effective filter of miss-specified dynamics and how to deal with cases where one does not have full knowledge of the likelihood function. The present paper owes a lot to the score-driven extension of the ERGM framework (SD-ERGM). Nonetheless, as it will be clear in the next section, the generalization of the static fitness model to the temporal sparse weighted case following the steps of the SD-ERGM is far from being trivial.

2. Score Driven Generalized Fitness Model

In this work, we propose a model for sparse and weighted temporal networks that we refer to as *score driven generalized fitness model*. Our model combines the well known *fitness model* - that was first discussed in [40], appropriately extended to a weighted zero augmented version to handle the weights, and the DCS approach to time varying parameter models [19, 20]. Before presenting the details of our model, we deem appropriate to review some preliminary concepts on the binary fitness model and the score driven framework.

2.1. Preliminary Concepts

The first concept that we use in our work is the fitness model. Multiple variations of the fitness model have been considered in the literature [21, 22, 23, 24, 25, 26] and the same specification of the probability for a binary, directed or undirected, networks is known also as *beta model* or *configuration model*. Hereafter we will define the fitness model as a model for binary networks where, in the case of directed networks, two parameters are assigned to each node i , the in-fitness $\overleftarrow{\theta}_i$ and the out-fitness $\overrightarrow{\theta}_i$, that describe the tendency of each node to form incoming and outgoing connections, respectively. Because the occurrence of a link is very likely influenced by both intrinsic and external factors, we focus

in the following on a specification that allows for the links' probability to depend also on the external, possibly link specific, covariates [26]. Hence the probability of a link to exist is described by

$$p_{ij} = \frac{1}{1 + e^{-(\overleftarrow{\theta}_i + \overrightarrow{\theta}_j + \mathbf{X}_{ij}\beta)}},$$

where β is the coefficient tied to the information of the, link specific, covariate \mathbf{X} . This model is very convenient because it allows for a parsimonious modeling of the matrix associated with a network, via node specific fitness.

The second key ingredient of our approach is the DCS framework for time varying parameter models, introduced by [19] and [20]. In order to review it, let us consider a sequence of observations $\{z^{(t)}\}_{t=1}^T$, where each $z^{(t)} \in \mathbb{R}^M$, and a conditional probability density $P(z^{(t)}|f^{(t)})$, that depends on a vector of time varying parameters $f^{(t)} \in \mathbb{R}^K$. Defining the score as $\nabla^{(t)} = \frac{\partial \log P(z^{(t)}|f^{(t)})}{\partial f^{(t)'}}$, a score-driven model assumes that the time evolution of $f^{(t)}$ is ruled by the recursive relation

$$f^{(t+1)} = w + \mathbf{b}f^{(t)} + \mathbf{a}\mathcal{I}^{(t)}\nabla^{(t)}, \quad (1)$$

where w , \mathbf{a} and \mathbf{b} are static parameters, w being a K dimensional vector and \mathbf{a} and \mathbf{b} $K \times K$ matrices. $\mathcal{I}^{(t)}$ 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 $P(y^{(t)}|f^{(t)})$, i.e. $\mathcal{I}^{(t)} = \mathbb{E} \left[\frac{\partial \log P(y^{(t)}|f^{(t)})}{\partial f^{(t)'}} \frac{\partial \log P(y^{(t)}|f^{(t)})'}{\partial f^{(t)'}} \right]^{-\frac{1}{2}}$. However, this is not the only possible specification and different choices for the scaling are discussed in [19]. A score driven model can be regarded both as *data generating process* (DGP) or as a filter of an unknown dynamics. In both cases, the most important feature of (1) is the role of the score as the driver of the dynamics of $f^{(t)}$. The structure of the conditional observation density determines the score, from which the dependence of $f^{(t+1)}$ on the vector of observations $y^{(t)}$ follows. When the model is viewed as a DGP, the update results in a stochastic dynamics thanks to the random occurrence of $z^{(t)}$. When the score-driven recursion is regarded as a filter, the update rule in (1) is used to obtain a sequence of filtered $\{\hat{f}^{(t)}\}_{t=1}^T$. In this setting, the static parameters are estimated maximizing the log-likelihood of the whole sequence of observations [for a detailed discussion, see 20, 41],

$$(\hat{w}, \hat{\mathbf{b}}, \hat{\mathbf{a}}) = \arg \max_{(w, \mathbf{b}, \mathbf{a})} \sum_{t=1}^T \log P \left(z^{(t)} | f^{(t)} \left(w, \mathbf{b}, \mathbf{a}, \left\{ z^{(t')} \right\}_{t'=1}^{t-1} \right) \right).$$

Score driven models have seen an explosion of interest in recent years due to their flexibility and ease of estimation. Indeed, many state of the art wildly popular econometric models turn out to belong to the family of score driven models. Examples are the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model of [42], the Exponential GARCH model of [43], the Autoregressive Conditional Duration model of [44], and the Multiplicative Error Model of [45]. Moreover, there are motivations, originating in information theory, for the optimality of the score-driven updating rule [46]. The introduction of this framework in its full generality has been investigated from a theoretical point of view [41, 46, 47, 48] and opened the way to applications in various contexts ¹.

2.2. Definition of the Model

In order to introduce the *score driven weighted fitness model*, let us describe a sequence of networks with a set of random variables $Y_{ij}^{(t)}$, one for each link. We propose to use Zero Augmentation to model separately the probability of observing a link, $Y_{ij}^{(t)} > 0$, and the probability to observe a specific weight $Y_{ij}^{(t)}$. We prefer Zero Augmentation over censoring, as we believe the former to be much more flexible in the context of network data. With this choice the probability distribution for the link ij is

$$P\left(Y_{ij}^{(t)} = y\right) = \begin{cases} 1 - p_{ij}^{(t)} & \text{for } y = 0 \\ p_{ij}^{(t)} g_{ij}^{(t)}(y) & \text{for } y > 0. \end{cases} \quad (2)$$

where $g_{ij}^{(t)}(y)$ is the distribution for the positive continuous weight for link ij , conditional on the presence of the link.

We then model the binary temporal network and its weights by means of time varying fitness, allowing also for the dependency on external covariates X_{ij} . In our model the probability of a link to exist is described by

$$p_{ij}^{(t)} = \frac{1}{1 + e^{-\left(\overleftarrow{\theta}_i^{(t)} + \overrightarrow{\theta}_j^{(t)} + X_{ij}^{(t)} \beta_{bin}\right)}}, \quad (3)$$

where, for simplicity we consider only one external covariate $X_{ij}^{(t)}$ with its own associated parameter β_{bin} , while in general nothing prevents us to have multiple covariates. With this choice, the log-likelihood for

¹Please refer to <http://www.gasmodel.com/index.htm> for the updated collection of papers dealing with GAS models.

the single observation $(\mathbf{Y}^{(t)})$ in (2) is

$$\begin{aligned}
\log P\left(\mathbf{Y}^{(t)}|\overleftarrow{\theta}^{(t)}, \overrightarrow{\theta}^{(t)}, \beta_{bin}, \beta_w, \mathbf{X}^{(t)}\right) &= \sum_{ij} \left(\mathbb{I}\left(Y_{ij}^{(t)}\right) - 1\right) \left(\overleftarrow{\theta}_i^{(t)} + \overrightarrow{\theta}_j^{(t)} + \beta_{bin} X_{ij}^{(t)}\right) \\
&\quad - \log\left(1 + e^{-\left(\overleftarrow{\theta}_i^{(t)} + \overrightarrow{\theta}_j^{(t)} + \beta_{bin} X_{ij}^{(t)}\right)}\right) + \mathbb{I}\left(Y_{ij}^{(t)}\right) \log g_{ij}^{(t)} \\
&= \sum_i \overleftarrow{D}_i \overleftarrow{\theta}_i^{(t)} + \overrightarrow{D}_i \overrightarrow{\theta}_i^{(t)} \\
&\quad + \sum_{ij} \mathbb{I}\left(Y_{ij}^{(t)}\right) \beta_{bin} X_{ij}^{(t)} + \overleftarrow{\theta}_i^{(t)} + \overrightarrow{\theta}_j^{(t)} + \beta_{bin} X_{ij}^{(t)} \\
&\quad - \log\left(1 + e^{-\left(\overleftarrow{\theta}_i^{(t)} + \overrightarrow{\theta}_j^{(t)} + \beta_{bin} X_{ij}^{(t)}\right)}\right) + \mathbb{I}\left(Y_{ij}^{(t)}\right) \log g_{ij}^{(t)},
\end{aligned}$$

where the indicator function \mathbb{I} is zero if its argument is less or equal than zero and one otherwise. We denote the in and out degrees of vertex i at time t respectively as $\overleftarrow{D}_i = \sum_j \mathbb{I}\left(Y_{ij}^{(t)}\right)$ and $\overrightarrow{D}_i = \sum_j \mathbb{I}\left(Y_{ji}^{(t)}\right)$. As in the standard fitness model for binary networks, the time varying *binary fitness* parameters $\left(\overleftarrow{\theta}^{(t)}, \overrightarrow{\theta}^{(t)}\right)$ describe the tendency of nodes at time t to form links, that are not explained by the external covariate $\mathbf{X}^{(t)}$. In order to model the weights of the observed links, we consider a generalized version of the fitness model where we associate to each node i , at time t , also the parameters $\overleftarrow{\eta}_i^{(t)}, \overrightarrow{\eta}_i^{(t)}$, that we call *weighted fitness*. They describe the propensity of a node to have more or less heavy weights in incoming and outgoing links respectively, and are related to the distribution of the weights of present links $g_{ij}^{(t)}$ by

$$\mathbb{E}\left[Y_{ij}^{(t)}|Y_{ij}^{(t)} > 0\right] = e^{\left(\overleftarrow{\eta}_i^{(t)} + \overrightarrow{\eta}_j^{(t)} + X_{ij}^{(t)} \beta_w\right)}, \quad (4)$$

where we considered the dependency on a single external covariate $X_{ij}^{(t)}$ and indicated the associated regression coefficient with β_w to distinguish it from the coefficient for the binary part in (3). This choice of linking the weighted fitness to the mean of the distribution g provides dynamics and heterogeneity only to one parameter of the conditional distribution, as shown in the following with a concrete example. The score driven weighted fitness model can be defined for a generic distribution g , for both continuous and discrete data, as we discuss in Appendix Appendix A. Nevertheless in the following, for concreteness,

we will focus on the gamma distribution to model links' weights

$$g_{ij}(y) = \frac{(\mu_{ij})^{-\sigma} y^{(\sigma-1)}}{\Gamma(\sigma)} e^{-\frac{y}{\mu_{ij}}},$$

where σ is a positive constant. Given a sequence of observed weighted adjacency matrices $\{\mathbf{Y}^{(t)}\}_{t=1}^T$, we denote by $f^{(t)}$ a K dimensional vector, where $K = 4 \times N$, containing all the time varying fitness parameters $\overleftarrow{\theta}^{(t)}$, $\overrightarrow{\theta}^{(t)}$, $\overleftarrow{\eta}^{(t)}$, $\overrightarrow{\eta}^{(t)}$. With this notation, the model's distribution takes the following form

$$P\left(Y_{ij}^{(t)} = y | f^{(t)}, \beta_{bin}, \beta_w, \sigma\right) = \begin{cases} \frac{e^{-(\overleftarrow{\theta}_i^{(t)} + \overrightarrow{\theta}_j^{(t)} + X_{ij}^{(t)} \beta_{bin})}}{1 + e^{-(\overleftarrow{\theta}_i^{(t)} + \overrightarrow{\theta}_j^{(t)} + X_{ij}^{(t)} \beta_{bin})}} & \text{for } y = 0 \\ \frac{(\mu_{ij}^{(t)})^{-\sigma} \Gamma(\sigma)^{-1}}{1 + e^{-(\overleftarrow{\theta}_i^{(t)} + \overrightarrow{\theta}_j^{(t)} + X_{ij}^{(t)} \beta_{bin})}} y^{(\sigma-1)} e^{-\frac{y}{\mu_{ij}^{(t)}}} & \text{for } y > 0. \end{cases} \quad (5)$$

with

$$\mu_{ij}^{(t)} = \sigma^{-1} e^{(\overleftarrow{\eta}_i^{(t)} + \overrightarrow{\eta}_j^{(t)} + X_{ij}^{(t)} \beta_w)}.$$

We let the fitness, both binary and weighted, evolve in time, following the score-driven recursive update rule in (1), that in this case takes the form

$$f^{(t+1)} = w + b f^{(t)} + a \mathcal{I}^{(t)} \frac{\partial \log P(\mathbf{Y}^{(t)} | f^{(t)})}{\partial f^{(t)'}}, \quad (6)$$

where w , a and b are three K dimensional vectors of static parameters ². $\mathcal{I}^{(t)}$ is a $K \times K$ scaling matrix. Hence, conditionally on the value of the parameters $f^{(t)}$ at time t and the observed adjacency matrix $\mathbf{Y}^{(t)}$, the parameters at time $t+1$ are deterministic. The element k of the score for the in and out binary

²Hence, in our definition we have three static parameters for each time varying fitness. While in the general formulation of score driven models [as given, for example, in 19], the parameters b and a are defined as matrices, it is very common to impose some restriction on them in order to limit the number of parameters to be estimated, as we do here.

fitness takes the following form

$$\begin{aligned}\frac{\partial \log P(\mathbf{Y}|f^{(t)}, \beta_{bin}, \beta_w, \mathbf{X}^{(t)})}{\partial \overleftarrow{\theta}_k^{(t)'}} &= \overleftarrow{D}_k - \sum_j \frac{1}{1 + e^{-(\overleftarrow{\theta}_k^{(t)} + \overrightarrow{\theta}_j^{(t)} + X_{kj}^{(t)} \beta_{bin})}} \\ \frac{\partial \log P(\mathbf{Y}|f^{(t)}, \beta_{bin}, \beta_w, \mathbf{X}^{(t)})}{\partial \overrightarrow{\theta}_k^{(t)'}} &= \overrightarrow{D}_k - \sum_i \frac{1}{1 + e^{-(\overleftarrow{\theta}_i^{(t)} + \overrightarrow{\theta}_k^{(t)} + X_{ik}^{(t)} \beta_{bin})}}\end{aligned}\quad (7)$$

and does not depend on the choice of g . The element k of the score for the weighted in and out fitness are

$$\begin{aligned}\frac{\partial \log P(\mathbf{Y}|f^{(t)}, \beta_{bin}, \beta_w, \mathbf{X}^{(t)})}{\partial \overleftarrow{\eta}_k^{(t)'}} &= \sum_j \mathbb{I}(Y_{kj}^{(t)}) \frac{\partial \log g_{kj}}{\partial \overleftarrow{\eta}_k} = \sum_j \mathbb{I}(Y_{kj}^{(t)}) \left(\frac{Y_{kj}^{(t)}}{\mu_{kj}^{(t)}} - \sigma \right) \\ \frac{\partial \log P(\mathbf{Y}|f^{(t)}, \beta_{bin}, \beta_w, \mathbf{X}^{(t)})}{\partial \overrightarrow{\eta}_k^{(t)'}} &= \sum_i \mathbb{I}(Y_{ik}^{(t)}) \frac{\partial \log g_{ik}}{\partial \overrightarrow{\eta}_k} = \sum_i \mathbb{I}(Y_{ik}^{(t)}) \left(\frac{Y_{ik}^{(t)}}{\mu_{ik}^{(t)}} - \sigma \right).\end{aligned}\quad (8)$$

In (7) and (8), as scaling matrix we use the Hessian of the log-likelihood.

As any score driven model, our model can be regarded both as a DGP or as a filter of a misspecified dynamics. When used as a DGP, it describes a stochastic dynamics because, at each time t , the adjacency matrix is not known in advance. It is randomly sampled from $P(\mathbf{Y}^{(t)}|f^{(t)})$ and then used to compute the score that, as a consequence, becomes itself stochastic. When the sequence of networks $\{\mathbf{Y}^{(t)}\}_{t=1}^T$ is observed, and the model is applied as a filter of the time varying parameters, the static parameters (w, b, a) , that best fit the data, can be estimated maximizing the log-likelihood of the whole time series.

Taking into account that each network $\mathbf{Y}^{(t)}$ is independent from all the others *conditionally* on the value of $f^{(t)}$, the log-likelihood can be written as

$$\log P\left(\left\{\mathbf{Y}^{(t)}\right\}_{t=1}^T | w, b, a\right) = \sum_{t=1}^T \log P\left(\mathbf{Y}^{(t)} | f^{(t)}\left(w, b, a, \left\{\mathbf{Y}^{(t')}\right\}_{t'=1}^{t-1}\right)\right).\quad (9)$$

Then the filtered time varying fitness $\hat{f}^{(t)}$ are obtained by an iterative application of (6) using as parameter values the maximum likelihood estimates (MLEs).

For ease of exposition, so far we introduced a baseline version of our model where the probabilities

depend on a single external covariate through the scalar β_{bin} and β_w . This implies uniform, i.e. equal across all links, dependency on the covariates. In order to explore node specific dependency on the external variables, we could easily consider the specification $X_{ij}(\beta_{bin_i} + \beta_{bin_j})$, where we associate two covariate coefficients to each node, one for outgoing links and one for incoming links.

2.2.1. Poisson Distribution for Discrete Weights and Inference from Partial Information

As mentioned before, our method is not restricted to a specific distribution for the weights and we can easily consider a different specification for g_{ij} in (2), instead of the gamma distribution. For example, in some settings it might be more appropriate to model weights as discrete variables. In this section we show how we can describe discrete weights by means of the Poisson distribution, defined as follows

$$g_{ij}(y) = \frac{\mu_{ij}^y e^{-\mu_{ij}}}{y!} \quad y \in \mathbb{N}.$$

If we consider a specification without external covariates, hence where $\mu_{ij}^{(t)} = \sigma^{-1} e^{(\overleftarrow{\eta}_i^{(t)} + \overrightarrow{\eta}_j^{(t)})}$, the log-likelihood takes the following form

$$\begin{aligned} \log P(\mathbf{Y}^{(t)} | f^{(t)}) &= \sum_i \overleftarrow{D}_i^{(t)} \overleftarrow{\theta}_i^{(t)} + \overrightarrow{D}_i^{(t)} \overrightarrow{\theta}_i^{(t)} + \sum_{ij} \overleftarrow{\theta}_i^{(t)} + \overrightarrow{\theta}_j^{(t)} - \log \left(1 + e^{-(\overleftarrow{\theta}_i^{(t)} + \overrightarrow{\theta}_j^{(t)})} \right) \\ &\quad + \sum_{ij} \mathbb{I}(Y_{ij}^{(t)}) \log g_{ij}^{(t)} \\ &= \sum_i \overleftarrow{D}_i^{(t)} \overleftarrow{\theta}_i^{(t)} + \overrightarrow{D}_i^{(t)} \overrightarrow{\theta}_i^{(t)} + \sum_{ij} \overleftarrow{\theta}_i^{(t)} + \overrightarrow{\theta}_j^{(t)} - \log \left(1 + e^{-(\overleftarrow{\theta}_i^{(t)} + \overrightarrow{\theta}_j^{(t)})} \right) \\ &\quad + \sum_i \left(\overleftarrow{S}_i^{(t)} \overleftarrow{\eta}_i^{(t)} + \overrightarrow{S}_i^{(t)} \overrightarrow{\eta}_i^{(t)} - \overleftarrow{D}_i^{(t)} e^{\overleftarrow{\eta}_i^{(t)}} - \overrightarrow{D}_i^{(t)} e^{\overrightarrow{\eta}_i^{(t)}} \right) - \log(Y_{ij}^{(t)}!), \end{aligned}$$

where we defined the in and out strengths as $\overleftarrow{S}_i^{(t)} = \sum_j Y_{ij}^{(t)}$ and $\overrightarrow{S}_i^{(t)} = \sum_j Y_{ji}^{(t)}$, respectively. Interestingly the log-likelihood of the model based on the Poisson distribution depends only on the degrees $(\overleftarrow{D}^{(t)}, \overrightarrow{D}^{(t)})$ and strengths $(\overleftarrow{S}^{(t)}, \overrightarrow{S}^{(t)})$ and, apart from the last not essential additive term, does not depend on the full matrix $\mathbf{Y}^{(t)}$. The score of $(\overleftarrow{\theta}, \overrightarrow{\theta})$ does not depend on the distribution of the weights

and reads as in (7), while the element k of the score for the weighted in and out fitness is

$$\frac{\partial \log P(\mathbf{Y}|f^{(t)})}{\partial \overleftarrow{\eta}_k^{(t)'}} = \overleftarrow{S}_k^{(t)} - \overleftarrow{D}_k^{(t)} e^{\overleftarrow{\eta}_k^{(t)}}$$

$$\frac{\partial \log P(\mathbf{Y}|f^{(t)})}{\partial \overrightarrow{\eta}_k^{(t)'}} = \overrightarrow{S}_k^{(t)} - \overrightarrow{D}_k^{(t)} e^{\overrightarrow{\eta}_k^{(t)}}.$$

Hence we could use the score driven update and estimate the model only relying on the sequences of strengths and degrees, ignoring the full matrix $\mathbf{Y}^{(t)}$. This has interesting implications as it allows for a parsimonious implementation of the inference procedure. Indeed, the observations can be compressed from a sequence of T matrices each with $N(N-1)$ elements³ to only 4 sequences of vectors each of length N , for a total of $4N$ elements. More importantly, it is not uncommon in practice not to have access to the full adjacency matrix, but only to the degree and strength sequences. This circumstance is so relevant that has stimulated an important stream of literature on the reconstruction of networks from partial information [49, 36, 13, 50]. We do not further investigate this aspect in this paper but leave it for future research.

3. Numerical Simulations in Misspecified Settings

In this section we discuss the results of extensive numerical simulations⁴ that we run to evaluate the performance of the score driven weighted fitness model as a misspecified filter, i.e. when the true DGP of the simulated data is not the same as the score driven update rule. This is the typical situation in practical applications, where the true DGP is unknown.

When modeling a sequence of observed weighted networks with time varying fitness as in Eq. (5), we compare the score driven update rule to two simple alternative filters. The first one is to specify the fitness, both binary and weighted, as static. This amounts to consider a zero augmented generalized linear model, accounting for node specific effects by means of the fitness. The probability of observing a link is that of the standard fitness model (3), where the parameters $(\overleftarrow{\theta}, \overrightarrow{\theta})$ do not change for all time steps $T = 1, \dots, T$, while the distribution of the weights depends on the constant weighted fitness

³Not considering the diagonal elements that would describe self-loops.

⁴The python code used for the simulations is available at <https://github.com/domenicodigangi/DynWGraphsPaper>

$(\overleftarrow{\eta}, \overrightarrow{\eta})$ such that

$$\mathbb{E} \left[Y_{ij}^{(t)} | Y_{ij}^{(t)} > 0 \right] = e^{(\overleftarrow{\eta}_i + \overrightarrow{\eta}_j + X_{ij}^{(t)} \beta_w)}.$$

The second alternative consists in estimating a static fitness model for each observed snapshot, i.e. at each time t . This procedure results in a sequence of single snapshot estimates for the fitness, where clearly the number of parameters to be estimated grows with the number of networks observed. This sequence of single snapshot estimates provides a filter for the time varying fitness, and has already been discussed, for the binary case, in [7].

In the rest of this section we compare the score driven weighted fitness model, when appropriate, with these two alternatives in three numerical experiments.

3.1. Filtering Time Varying Fitness

In our first numerical experiment we focus on filtering the fitness that evolves in time following a DGP different from the score driven dynamics. Specifically, we assess how accurately the score driven methodology allows us to filter the fitness when no external covariates are present, and compare it with the sequence of single snapshot MLEs. Since the focus here is on filtering time varying fitness, we do not consider the alternative version with constant fitness. In [7] we already carried out a similar comparison for the binary part of the model and we showed that, for the binary part, the sequence of cross sectional estimates clearly under-performs the score driven fitness model, both in numerical simulations and empirical applications. Here we repeat a similar exercise considering also the weighted fitness.

Specifically, we analyze sequences of networks sampled from Eq. (5) where each fitness, both binary and weighted, evolves according to an auto-regressive process of order one, $f_i^{(t+1)} = b_0 + b_1 f_i^{(t)} + \epsilon^{(t)}$ where $\epsilon \sim N(0, 0.1)$, $b_1 = 0.98$ and b_0 is chosen for each parameter in order to sample networks that resemble a real world realization. We also consider one example of a deterministic DGP where the fitness evolve in time following a sin function, as shown in Figure 1. In practice, in order to obtain realistic parameters values for the DGPs, we first estimate the fitness models, both binary and weighted on the first observation available for the temporal interbank network data that we describe in Section 4.2⁵. Then, we sample time series of 150 observations from the AR(1) processes for each fitness independently.

⁵The numerical values for the fitness used are available at <https://github.com/domenicodigangi/dynwgraphs>. We have checked that similar results hold true using different numerical values for the fitness parameters.

We use the simulated fitness paths as parameters to sample from the model specification in Eq. (5). Using only the sequence of sampled matrices as input, with the two methods we filter the time varying evolution of the fitness. For each fitness we compute the mean squared error (MSE) over the time series and then average it across nodes, keeping separate the MSE for binary and weighted fitness. Finally, we repeat the sampling and filtering procedure 50 times and average the obtained MSE across the repetitions. The results, showed in Table 1, confirm that, also for the weighted fitness, the score driven update rule

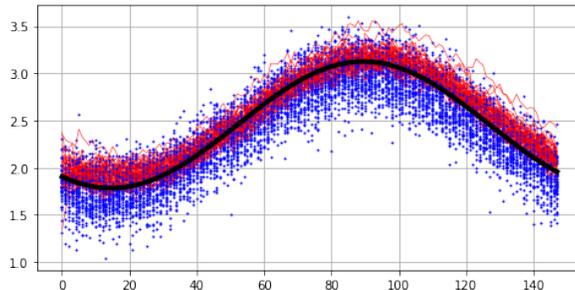


Figure 1: Filtered paths for one of the time varying weighted fitness parameters. The path from the true DGP is in black. The blue dots are the single snapshot estimates, and the red lines the paths filtered with the score driven generalized fitness model.

DGP	AR(1)		SIN	
Filter	SS	SD	SS	SD
Avg. MSE $(\overleftarrow{\theta}, \overrightarrow{\theta})$	0.36	0.11	0.66	0.04
Avg. MSE $(\overleftarrow{\eta}, \overrightarrow{\eta})$	0.54	0.25	0.6	0.18

Table 1: Results from the first experiment: MSE of the filtered fitness averaged across all nodes, for both the single snapshot (SS) and score driven (SD) filters.

is clearly a better choice in filtering misspecified paths for the time varying fitness.

Having established that as a misspecified filter of the fitness, the sequence of single snapshot estimates under-performs the score driven alternative, also for the weighted fitness, let us mention a second limitation of the single snapshot approach. In the usual empirical setting, only one network realization is available at each time step. In such a situation, we might not be able to jointly estimate the sequence of single snapshots estimates and the coefficients β_{bin} , or β_w , due to the low number of observations per parameter. This is the case, for instance, if we consider as covariate a variable that is uniform across all links i.e. $X_{ij}^{(t)} = x^{(t)}$ for all i, j , a case that we consider in the empirical application of Section 4.2. An

identification issue arises for the single snapshot estimates for both the binary and weighted parameters. In fact, the probability of the sequence of observations, given the sequence of weighted fitness, remains unchanged under the following transformation

$$\begin{aligned}\overleftarrow{\eta}_i^{(t)} &\rightarrow \overleftarrow{\eta}_i^{(t)} + c_1 x^{(t)}, \quad \forall i \\ \overrightarrow{\eta}_i^{(t)} &\rightarrow \overrightarrow{\eta}_i^{(t)} + c_2 x^{(t)}, \quad \forall i \\ \beta_w &\rightarrow \beta_w + c_3,\end{aligned}$$

for any choice of (c_1, c_2, c_3) such that $c_1 + c_2 + c_3 = 0$, since for each t such a transformation does not change the sum $\overleftarrow{\eta}_i^{(t)} + \overrightarrow{\eta}_j^{(t)} + \beta_w x^{(t)}$ ⁶. Hence, in the model specification with a uniform external covariate, we cannot identify the parameter of interest β_w and this prevents us to use sequences of single snapshot estimators. With a simple change of notation we can see that the same issue arises for the binary parameters. We point out that the model with static fitness, that we use for comparison, and our score driven version, do not suffer from this identification issue, as in both cases the number of static parameters to be estimated does not increase with the number of time steps observed. For instance, in the score driven model, the sequence of time varying parameters $(\overleftarrow{\eta}^{(t)}, \overrightarrow{\eta}^{(t)})$ for $t = 1, \dots, T$, is not estimated directly but follows from the score driven update rule (6) that is uniquely identified given the sequence of observations and the static parameters (w, b, a) . For this reason, in the rest of this section, we compare the score driven weighted fitness model only with the alternative having constant fitness.

3.2. External Covariates and Omitted Variables Misspecification

We present here two numerical experiments to assess how effective the score driven weighted fitness model can be in estimating the dependency of the network dynamics on external covariates. We show that, in synthetically generated datasets, introducing the binary and weighted time varying fitness reduces errors due to unobserved variables.

Our second experiment is designed to show that the score driven model properly describes the effect of external covariates, even when the time varying fitness is generated by a misspecified DGP. Moreover it highlights the importance of taking into account time varying node specific effects. In fact, assuming the fitness to be constant, when they are actually time varying, results in poor estimates of the dependency

⁶The need to fix an identification condition in fitness models, even without external covariates, is well known as we discuss in Appendix B.

on external covariates. In order to show that, we consider samples from the model in Eq. (5), where we let the fitness evolve with the AR(1) DGP of the previous experiment. We also assume that the network dynamics depends on the realization of two independent, predetermined, external variables, \mathbf{X}_{bin} and \mathbf{X}_w . The first covariate enters the binary part of the DGP

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

and the second one influences the expected weights in 5 as follows

$$\mu_{ij}^{(t)} = \sigma^{-1} e^{(\overleftarrow{\eta}_i + \overrightarrow{\eta}_j + X_{w_{ij}}^{(t)} \beta_w)}.$$

We fix $\beta_{bin} = 1$, $\beta_w = 1$ and consider two possible specifications for the DGP of the synthetic external covariates. In the first one, both external variables are scalar, $X_{bin_{ij}}^{(t)} = x_{bin}^{(t)}$ and $X_{w_{ij}}^{(t)} = x_w^{(t)}$ for all i, j , and follow an AR(1) process equal to the one followed by the fitness in the previous example ⁷. In the second specification, we set $X_{bin_{ij}}^{(t)} = \mathbb{I}(Y_{ij}^{(t-1)})$ and $X_{w_{ij}}^{(t)} = \mathbb{I}(Y_{ij}^{(t-1)}) \log(Y_{ij}^{(t-1)})$ for all i, j . The latter DGP, due to the explicit dependency of the network at time t from its realization at previous time $t - 1$, simulates a temporal network with link persistence, i.e. a higher probability of observing at time t a link if this was present at time $t - 1$ and weight persistence, i.e. the weight at time t , if the link is present, is affected by its weight at time $t - 1$.

We sample 50 sequences of networks, and external covariates, each of 150 time steps. For each sampled sequence, we estimate the score driven weighted fitness model maximizing the likelihood of the static parameters. We then compute the MSE between simulated and estimated values of β_{bin} and β_w across the sample. In order to test the effect of neglecting time varying effects, we repeat the same procedure for a version of the model with static fitness and report the results in Table (2). It emerges clearly that not taking into account the dynamics of the fitness can severely deteriorate the estimation of the effect of external covariates.

In order to motivate our third and last numerical experiment, let us mention that the topic of estimation errors due to omitted variables has been discussed widely in the econometric literature [51, 52, 30]. The standard approach to mitigate it consists in using control variables. This approach has some

⁷We set the parameter b_0 so that the unconditional mean of the AR(1) process is equal to 1.

DGP	AR(1) Scalar External Regressors		Persistent links and weights	
Filter	Constant Fitness	SD Fitness	Constant Fitness	SD Fitness
MSE β_{bin}	0.14	0.06	0.23	0.02
MSE β_w	0.34	0.02	0.18	0.015

Table 2: Results for the second experiment: MSE of the estimated regression coefficients (average over 50 samples). The fitness dynamics follow an AR(1) process. Two columns on the left: external regressors are scalar and follow two AR(1) processes. Two columns on the right: regressors induce a persistent dynamics in both links and weights.

known downsides [refer for instance to 53, 54, 55] and, most importantly, it is not always viable since the data on appropriate controls might not be available. Indeed, this is very common for financial networks, as the one that we consider in Section 4.2, where the variables that one would like to use as controls are likely to be privacy protected and often unavailable to researchers. Considering, for example the case of interbank networks, we could very well expect the current leverage of a bank to have a significant influence on the decision to borrow or lend, i.e. create interbank links. Nevertheless it is very unlikely for this information to be available, at the frequency required. This issue is even more frequent when the datasets are anonymous, and the identity of the nodes is not known. For these reasons, in the third experiment we show that allowing the fitness to vary in time is extremely beneficial to mitigate the errors due to omitted variables, at least in the context of controlled numerical simulations. We assume the network dynamics, both of the links and of the weights, to be determined by two external covariates. We then assume that, for whatever reason, only one of them is observed. Specifically the considered DGP is similar to the one in Eq. (5) but with parameters defined as

$$p_{ij}^{(t)} = \frac{1}{1 + e^{-(\beta_{1,bin}x_1^{(t)} + \beta_{2,bin}x_2^{(t)})}},$$

and

$$\mu_{ij}^{(t)} = \sigma^{-1} e^{(\beta_{1,w}x_1^{(t)} + \beta_{2,w}x_2^{(t)})}.$$

We assume that $x_2^{(t)}$ is not observed and assess the effect of omitting it when estimating the coefficients β_{bin_1} and β_{w_1} . We show that the score driven fitness compensates the impact of the neglected variable on the estimates. The external variables $x_1^{(t)}$ and $x_2^{(t)}$ are independent and follow an AR(1) model with high persistence, $b_1 = 0.98$. We sample the DGP and estimate the coefficients 50 times for a sequence of networks 150 time step long. We then compare the MSE for the estimates of parameters $\beta_{1,bin}$ and

DGP	No Fitness and two regressors		
Filter	No Fitness	Constant Fitness	SD Fitness
MSE β_{bin_1}	5.66	0.53	0.01
MSE β_{w_1}	133	0.39	0.08

Table 3: Misspecified filtering of a DGP with two covariates and no fitness. First column: MSE for the estimates of β_{bin_1} and β_{w_1} , when no fitness is used (average over 50 samples). Second column: MSE for estimates with constant fitness. Third column: MSE using a model with score driven fitness.

$\beta_{1,w}$, obtained using three different specifications of the model in (5): one without node heterogeneity, hence no fitness, where the probability of observing a link depends only on the observed covariate. The second is a model with constant fitness a the model with score driven fitness and the third uses the observed external covariate. From the results that we report in Table (3), it follows that considering a model with time varying fitness is extremely beneficial when the DGP is misspecified and does not feature node specific effects. We believe that the results presented in this section strongly support the choice to describe temporal weighted networks by means of time varying fitness. Moreover, they give us clear insights to interpret the results of Section 4, where we find a clear advantage, in terms of goodness of fit, in using the score driven weighted fitness model to describe the empirical data.

4. Link and Weight Dynamics in the Italian e-MID

We apply our model to the interbank overnight loans market described as a temporal weighted network. Interbank markets are an important point of encounter for banks' supply and demand of extra liquidity, and have received much attention in the literature [see 56, for a review]. In particular, e-MID has been investigated in many papers [see, for example 57, 58, 6, 59, and references therein]. We use data from the e-MID, a market where banks can extend loans to one another for a specified term and/or collateral. Our dataset contains the list of all credit transactions in each day from June 6, 2009 to February 27, 2015. In our analysis, we investigate the interbank network of overnight loans, aggregated weekly. The standard approach in the literature to model temporal interbank networks is to disregard the size of the exposures and consider only the presence or absence of links, i.e. to consider only the binary network. Thanks to the flexibility of the score driven weighted fitness model, we are able to take into account and explicitly model the strength of the links. We consider a link from bank j to bank i as present at week t if bank j lent money overnight to bank i , and the associated weight as the total

amount lent over that period. This results in a set of $T = 298$ weekly aggregated weighted networks. For a detailed description of the dataset, we refer the reader to [59].

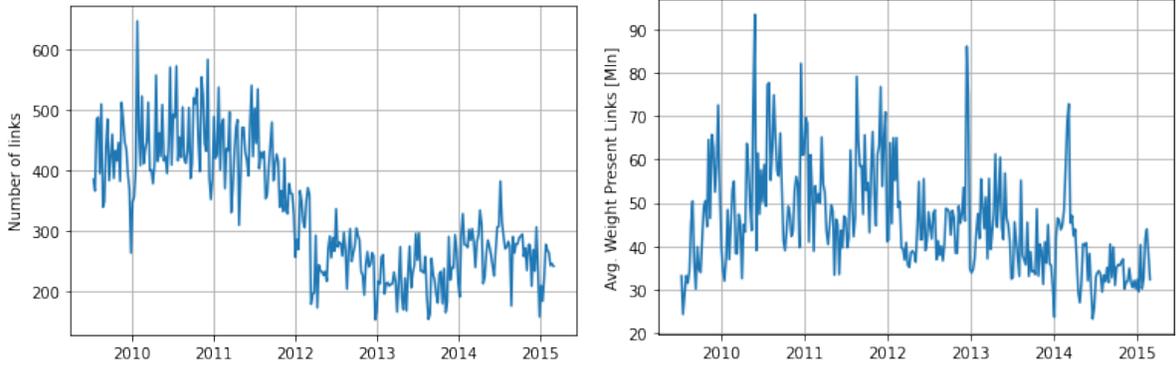


Figure 2: Left panel: number of links present in the data at each time step. Right panel: average weight in Millions of Euro of the present links.

As it is evident from the left panel of Figure 2, the number of links in e-MID is significantly lower in the second half of the dataset. In particular it started declining in 2011, most likely as a consequence of the European sovereign debt crisis and of the ECB unconventional measures, namely the long term refinancing operations (LTROs) that took place on the 22nd of December 2011 and on the 29th of February 2012 (see [59] for an in-depth discussion). The number of links then fluctuated around a new lower level since the beginning of 2012. As discussed in [59], the decreased number of links corresponds to a lower number of banks being active in the market. Both the network density⁸ and the average weight of present links has not followed a similar clear transition to a different level.

4.1. Link and Weight Prediction

As a first empirical application, we explore the possibility of using our approach to predict the presence of future links and the value of their weights. The problem of link prediction in temporal networks is extremely relevant in practical applications and has been discussed widely in the literature on binary networks [60, 61, 62, 63]. It can be defined in multiple ways depending on the context, the type of data at hand, and whether we want to predict the existence of a link in a partially observed network or the presence of a link in a future network. We focus on the latter case that is often referred to as temporal

⁸Number of links present divided by the number of possible links, given the number of active banks.

link prediction. On the other hand, the prediction of weights in weighted temporal networks has received much less attention so far mainly due to the lack of models suited to describe both links and weights. For this reason, here we run an exercise using the full dataset described in the previous section and focus on forecasting the weights of the network at time $t + 1$, using only a subset of the information available at time t . This means that, when forecasting observations at time $t + 1$, we use only observations from $t - T_{train}$ to t for training, extremes included. This is very much in line with what we did in [7], with the important difference that there we only considered binary link prediction, i.e. the sole prediction of links' presence. For a detailed discussion of how to employ the binary component of our model for the link prediction, we refer the reader to our previous work.

Our first goal is to assess how effective can a model with time varying fitness in Eq. (5) be in forecasting the weights and to benchmark the proposed score driven approach against an alternative model. To this end, we compare the forecasts obtained using two methods, both based on Eq. (5) and differing only in their dynamics for the fitness. The first approach uses the score driven model to forecast the fitness at time $t + 1$ with the update rule of Eq. (6). This is very easy to do in practice since, as mentioned in Section 2.2, the score driven fitness at time $t + 1$ are deterministic conditionally on the observations at time t . The second approach is a combination of the sequence of single snapshot estimates, described at the beginning of Section 3, and a set of AR(1) models, one for each fitness. In the latter approach, we first obtain the sequence of single snapshot estimates on the training set and then use them to estimate an AR(1) process for each fitness. Then, we use the estimated AR(1) models to forecast one value of the fitness at time $t + 1$. Practically, we repeatedly estimate the two models on rolling windows of length $T_{train} = 100$ time steps and, once for each estimate, we forecast the first out-of-sample observation for each train window. We then compare the weights of present link at time $t + 1$ with the expected values obtained from the two models and quantify the error by the mean squared error between the logarithms of observed and predicted weights

$$\text{MSE Log.} = \frac{\sum_{ij} \mathbb{I} \left(Y_{ij}^{(t+1)} \right) \left(\log \left(\mathbb{E} \left[Y_{ij} | \widehat{\eta}^{(t+1)}, \widehat{\eta}^{(t+1)} \right] \right) - \log \left(Y_{ij}^{(t+1)} \right) \right)^2}{\sum_{ij} \mathbb{I} \left(Y_{ij}^{(t+1)} \right)},$$

where $\widehat{\eta}^{(t+1)}$ and $\widehat{\eta}^{(t+1)}$ are the forecasts for the in and out weighted fitness, obtained using only

observations up to time t . Similarly we compute the mean absolute difference (MAD)

$$\text{MAD Log.} = \frac{\sum_{ij} \mathbb{I}(Y_{ij}^{(t+1)}) \left| \log \left(\mathbb{E} \left[Y_{ij} \mid \widehat{\eta}^{(t+1)}, \widehat{\eta}^{(t+1)} \right] \right) - \log(Y_{ij}^{(t+1)}) \right|}{\sum_{ij} \mathbb{I}(Y_{ij}^{(t+1)})}.$$

We compare the logarithms of predicted and observed weights because the distribution of observed weights is quite heterogeneous. It roughly spans five order of magnitudes, and directly comparing the weights would result in a measure of goodness of fit mainly describing the fit of the largest weights ⁹.

Method	SD	SS - AR(1)	Diebold-Mariano (p-value)
MSE Log.	0.859	0.882	1.73×10^{-7}
MAD Log.	0.726	0.737	1.21×10^{-7}

Table 4: Weight prediction exercise: MSE and MAD between the logarithms of observed and predicted weights. First and second columns: results from the score driven (SD) model and the approach based on single snapshot estimates and AR(1) processes ((SS) - AR(1)), respectively. Third column: p-values of a Diebold-Mariano test for the null hypothesis that the two forecasts are equivalent.

From the results reported in Table 4 we conclude that, similarly to the binary case, the score driven approach is a better choice to predict the weights with respect to a prediction based on sequence of single snapshot estimates, both in terms of MSE and MAD for the logarithms. The Diebold-Mariano [64, 65] test rejects the null hypothesis that the two forecasts are statistically equivalent.

4.1.1. Comparison With Link Specific Regressions

Among the models for weighted temporal networks reviewed in Section 1.1, [5] is the most relevant reference for the current work. Authors run a forecasting exercise on an interbank network of overnight loans. Differently from our case, they focus on the sterling interbank market that is very dense and somewhat small, when compared with the e-MID dataset. This is due to the structure of the UK interbank market that they consider. Indeed the fraction of links present at every time step for the full temporal network is always above 40%, while in our data it is typically equal to 6%. Furthermore, [5] run a weight forecasting exercise restricted to the sub-network composed by the 4 largest banks in the system, thus increasing the density of the actual considered network. They propose to model each link by means of a Tobit regression estimating the parameters with a local likelihood approach. Aside

⁹Similar results hold using measures of relative error for the comparison between predicted and observed weights.

Model	T_{train}	MSE Log.	MAD Log.	AUC
Localized Tobit	100	2.351	1.067	0.714
ZA Regression	100	2.785	1.240	0.830
Localized Tobit	200	2.267	1.043	0.795
ZA Regression	200	2.912	1.282	0.867
SD Generalized Fitness	100	0.859	0.726	0.896

Table 5: Results of the link and weights prediction exercise on e-MID data. We compare the Localized Tobit model of [5] with a simple Zero Augmented regression that uses the same regressors. We also report the results of running the same exercise with the score driven generalized fitness model.

their methodological contributions, a key element of their work is the accurate selection of regressors to capture relevant network properties that are expected to influence the future network structure. They design a set of simple functions of the network at the previous time step and use them as regressors. Specifically, for the link going from bank i to bank j , they consider as regressors the weight of that same link at the previous time step (*lagged*) plus the following quantities

- lagged total daily amount lent by i to all other banks except j ;
- lagged total daily amount borrowed by j from all other banks except i ;
- lagged total daily amount lent by j to all other banks except i ;
- lagged total daily amount borrowed by i from all other banks except j ;
- lagged total daily lending and borrowing not involving either i or j .

Here the goal is to compare their approach to link forecasting with ours, in an application to the e-MID data introduced in the previous sections, and investigate the reasons of different performances in links and weights prediction. In order to guide our intuition, in the following we also consider a version of their regression based on a simple Zero Augmented approach instead of the censoring that defines the Tobit approach. For the sake of simplicity, in this third approach, we separately consider a logistic regression on the binary part and a linear regression for the weights, both estimated with a standard, non local, maximum likelihood approach. As before we repeatedly estimate the models on rolling windows of length T_{train} and report the results in Table 5. It emerges clearly that the score driven generalized fitness model achieves better results in forecasting both links and weights, for the dataset that we considered, with respect to the localized Tobit regression of [5]. We believe this result to be mainly driven by two factors. The first one is that the Tobit model uses censoring to assign a finite probability to zero observations,

i.e. missing links. Thus, the probability of observing a link and the expected values of the weights are modelled by the same set of parameters. In contrast, the generalized score driven fitness model employs zero augmentation to separately model the probability of observing a link and the expected weight of present links. Indeed, this intuition is confirmed by looking at the performances of the simple Zero Augmented regression, that significantly improves the prediction of the existence of links – as assessed by the area under the curve (AUC) measure – but achieves slightly worst performances on predicting the weights of existing links with respect to the local likelihood Tobit. We empirically found the performances of both models to deteriorate significantly for the links that are rarely observed in the training set. In fact, as we show in Figure 3, both the out-of-sample log-MSE and AUC deteriorate when computed for subsets of links with a low fraction of non zero observations in the training set. Interestingly, the log MSE of the ZA regression seems to be monotonically decreasing when the density in the training set increases, while the Tobit regression does not. We interpret this as an additional indication of the advantages in separating the modelling of links’ presence from their expected weight.

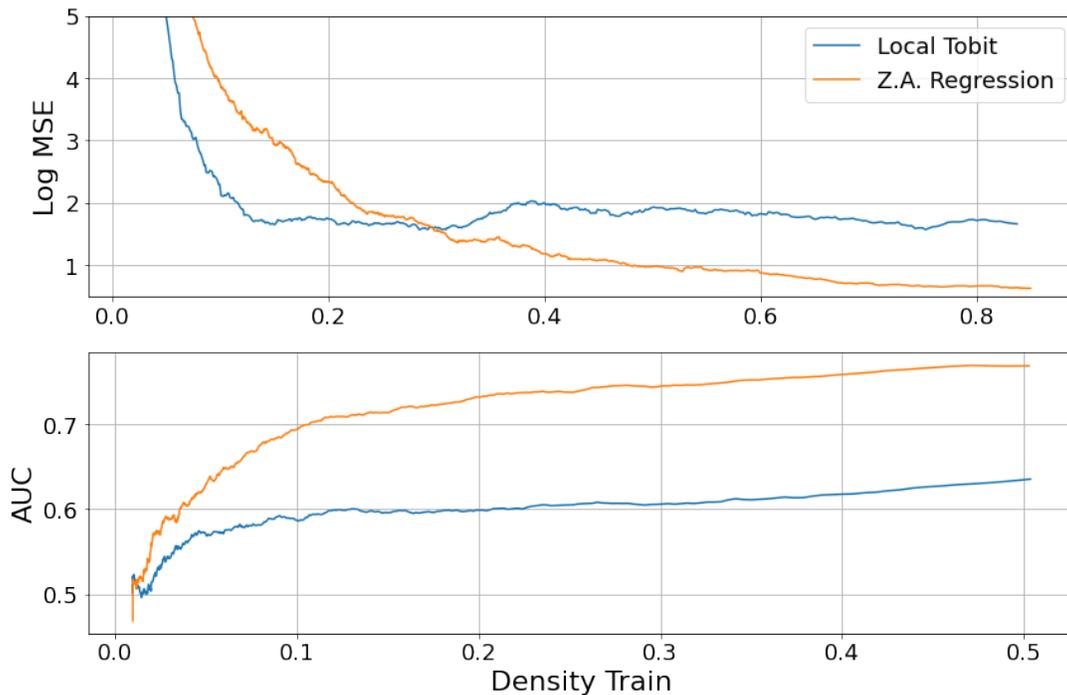


Figure 3: Out-of-sample Forecast accuracy as a function of the density in the training set for the models based on link specific regressions. Each line is obtained by averaging the log MSE (top panel) and the AUC (bottom panel) over rolling subsets of 1000 links sorted by the fraction of non zero observations in the training sample.

The second factor is that approaches based on running one regression for each link are extremely over-parametrized. They require the estimation of a large number of parameters that scales with the number of links ($N(N - 1)$), as opposed to our approach which requires the estimation of 6 parameters per each node, thus scaling as N . While the idea of running a separate model for each link has clear advantages for parallelizing the execution, such a high number of parameters can result in overfit and poor out-of-sample generalization, when the dataset at hand does not allow for a large number of time steps to be used in training. Indeed we corroborate this intuition by noting that the results of both the localized Tobit and the Zero Augmented regression improve if we consider a training set of 200 time steps, instead of one of 100 (see the third and fourth lines of Table 5).

4.2. *The Effect of Interest Rates on Interbank Lending*

In this section, by means of the score driven model, we investigate the effect of interest rates on the dynamics of the interbank network data introduced above. To track average interest rates, we use the EONIA benchmark. EONIA “*is a measure of the effective interest rate prevailing in the Euro interbank overnight market. It is computed as a weighted average of the interest rates on unsecured overnight contracts on deposits denominated in Euro, as reported by a panel of contributing banks*”¹⁰.

Intuitively we expect that banks’ funding rates and the topology of the interbank market are deeply related. This relation is of clear interest from the point of view of policymakers and has received much attention in the literature [see, for instance, 27, 66, 67, 68, 28]. Of particular relevance for the results discussed in this section is the work of [27] that investigated the effects of banks’ characteristics and the conditions of the network as a whole on the interest rates that each bank faces on the interbank market. They exploited a remarkable dataset, obtained from Norges Bank real time gross settlement system, that allowed them to model bank specific interest rates as dependent on a set of variables and controls, including overall market’s liquidity. From basic supply and demand reasoning one would expect that excess liquidity in the market would have a negative pressure on the interest rates on average, and indeed they found that interest rates tend to be lower when the overall liquidity available on the market is higher.

A second relevant work for the purposes of this section is [28], where authors considered data on the e-MID interbank market for a period ranging from the beginning of 2006 to the end of 2012, focusing solely

¹⁰Definition from <https://stats.oecd.org/>.

on the binary part of the daily temporal network of overnight loans. They computed various aggregated network statistics for each time step, thus obtaining one univariate time series for each statistic. Among other quantities, they computed the density of each network – defined as the number of connections as a proportion of all possible connections – and, using a standard linear regression, found it to be positively related with EONIA.

In the following, we explore the impact of interest rates on the probability of observing each link and on the expected weight of observed links by applying the score driven weighted fitness model to the e-MID dataset, using as external covariate the EONIA rate. For our estimates, we use a training set comprising the first 80% of time steps and left the last 20% to assess goodness of fit out-of-sample. Moreover, similarly to what done in the numerical simulations discussed in Section 3, we compare the score driven time varying fitness with two alternative specifications: a model without fitness and one with constant fitness, as defined in Section 3, both with EONIA as the only external variable. We then compare their goodness of fit both in-sample and out-of-sample. In Table 6 we show the results that clearly confirm the importance of including time varying fitness to improve goodness of fit. We report the Bayesian Information Criterion (BIC) for each model, computed separately for the likelihood of observing links and the likelihood of their weights, to compare goodness of fit in-sample. The binary and weighted parts of each model are evaluated separately out of sample. We quantify out of sample accuracy in predicting link’s presence by means of the AUC, while for the weights we compute the MSE of the logs of the weights, only for the present links. Since the model without fitness is clearly not a good

Model	No Fitness	Constant Fitness	SD Fitness
BIC Bin	1.75×10^6	0.53×10^6	0.45×10^6
BIC Weight	1.90×10^6	1.46×10^6	1.46×10^6
AUC - Test Set	0.48	0.82	0.92
MSE Log. - Test Set	58.15	1.01	0.78

Table 6: EONIA effect on e-MID: in sample BIC, for both the binary and weighted part of model (5), AUC for the out of sample evaluation of the binary part, and the MSE of the logarithms for the out of sample evaluation of the weighted part. First and second columns: results for a model without and with constant fitness, respectively. Third column: score driven fitness.

fit for the data we do not discuss it further, and in Table 7 we report the estimated regression coefficients using the models with constant and score driven fitness. The model with score driven fitness is clearly the best fit for the data, both in sample and out of sample, as measured by the metrics reported in

Model	Constant Fitness	SD Fitness
β_{bin}	0.69 ± 0.06	0.29 ± 0.05
β_w	0.022 ± 0.029	-0.13 ± 0.02

Table 7: EONIA effect on e-MID. Estimates of the regression coefficients using models with constant or score driven fitness.

Table 6. Moreover, the parameters estimated by the score driven fitness model are always statistically significant while the β_w estimated from the constant fitness model is not. We interpret this discrepancy as a sign that disregarding the fitness dynamics can lead us also to miss-guided qualitative interpretations, consistently with the numerical results discussed in Section 3.2.

From the estimate $\beta_{bin} = 0.29 \pm 0.05$, we can deduce that, in the considered period, the probability of observing a link in the network is positively related with the interest rates, hence the lowering of interest rates tends to reduce the overall market interconnectdness, even taking into account bank specific effects captured by the fitness. This result is coherent with the relation between network density and EONIA found in [28], although the approach based on standard regression on aggregated network statistics is different from ours. In fact, thanks to the time varying binary fitness in our model, the estimated effect of EONIA is decoupled from bank specific effects that are instead accounted for by the fitness. Such a separation of bank specific effects from the impact of a covariate is instead not possible when considering the density of the whole network as done in [28].

For what concerns the effect on the liquidity exchanged through the observed links, i.e. the links' weights, the estimated $\beta_w = -0.13 \pm 0.02$, indicates that the weight of the observed overnight loans is negatively related with the average interest rate in the market. Our result is coherent with the work [27] on the Norwegian interbank market, but our methodology allows us to explore a different aspect of the relation between liquidity and interest rates. In fact, that result regards the relation between bank specific rates and aggregated liquidity, while we explore the relation between average rates on the market and the weight of present links, controlling for time varying bank specific effects by means of the time varying fitness. Thanks to zero augmentation and separate modelling of links and weights, our finding is directly related with the average magnitude of the overnight loans that are actually present, more than with the total liquidity in the market. In summary, the data considered indicate that lower interest rates are related with a reduction of network interconnectdness but an increase of the average liquidity flow for the loans that are present.

Finally, we mention that, as we show in Appendix Appendix C, if we do not include EONIA as external covariate the dynamical fitness tend to correlate with it. This corroborates the fact that the estimated coefficients are found to be statistically significant.

We point out that our results on the relation between the dynamics of the e-MID interbank network and EONIA are obtained leveraging the full information available in the description of a temporal network as a temporal sequence of matrices, and considering the impact of the covariate both on the probability of each individual link and the expected weight of observed links. Differently from [27] and [28], we do not need to collapse the matrices into a single network statistic to estimate the effects of external variables. We directly use matrix valued network data and, thanks to the time varying latent fitness parameters, we can decouple the impact of EONIA from unobserved time varying node specific effects. The advantage of using matrix valued network data will become even more evident in the next section where we consider link specific covariates and carry out an analysis that would be impossible with standard regression methods on univariate network statistics.

4.3. Link and Weight Persistence

As a final application, we use our model to contribute to the literature on the persistence in interbank networks [69, 70, 29, 6] by exploring both the persistence of links and that of the weights. The existence of privileged lending relations between pairs of banks is a well known phenomenon and it is often referred to as *preferential trading* [69]. The motivations behind it can be explained by the relevance of strong lending relationships between banks as a way to overcome monitoring of creditworthiness and limit the risk of counter-party default [70]. The existence of preferential trading behaviours has been assessed quantitatively by means of statistical methods specifically developed for the purpose [29], in the case of binary networks. Additionally, models for binary temporal networks have been proposed that explicitly take it into account [6].

In this section we exploit the flexibility of our model and estimate the effect of two predetermined covariates that are meant to capture the persistence of links and weights. For what concerns link persistence of the binary network, we explore how the presence of a link at time $t - 1$ influences the probability of observing a link at time t . That amounts to use $\mathbb{I}\left(Y_{ij}^{(t-1)}\right)$ as covariate in Eq. (3). To assess persistence in the links' weights, we estimate the effect of the weight of a link at $t - 1$ in determining its weight at time t by using $\log\left(Y_{ij}^{(t-1)}\right)$ as covariate in Eq. (4). Let us recall that we have tested numerically the possibility to estimate such effects in synthetically generated data in Section 3.

As in the previous section, we compare three models, a model without fitness, one with constant fitness,

Model	No Fitness	Constant Fitness	SD Fitness
β_{bin}	0.064 ± 0.038	2.875 ± 0.045	2.048 ± 0.045
β_w	1.050 ± 0.003	0.073 ± 0.002	0.064 ± 0.002
BIC Bin	5.66×10^6	4.46×10^6	4.16×10^6
BIC Weight	2.61×10^6	1.45×10^6	1.45×10^6
AUC - Test Set	0.748	0.882	0.932
MSE Log. - Test Set	21.22	0.90	0.76

Table 8: Results on estimates of link persistence in e-MID. One column for each one of the three alternative model specifications. In the third and fourth rows, we show the in sample BIC for the binary and weighted parts of the model in (5), respectively. The last two rows are out of sample measures of goodness of fit. The fifth one is the out of sample AUC for the binary part. The last row is the out of sample MSE of the logarithms of the weights.

and the score driven fitness model, all using the same external covariates and two scalar coefficients, β_{bin} and β_w that quantify the persistence of links and weights respectively. The results in Table 8 confirm that neglecting node specific time varying effects results in worst fitting of the data. This is evident by looking at the superior performances, both in-sample and out-of-sample of the models with score driven fitness, with respect to those without or with constant fitness. The three model specifications all result in positive coefficients both for the binary and the weighted covariates. With the best performing model among those three, the model with score driven fitness, we estimate $\beta_{bin} = 2.048 \pm 0.045$. This indicates that globally the presence of a link at time $t - 1$ positively impacts the probability of observing that same link at time t . This is in agreement with the current consensus in the literature, supporting preferential trading behaviours in e-MID, that has been validated empirically only on the binary part of temporal interbank networks, for example by [29] and [6]. The novel aspect of our analysis lays in the estimated $\beta_w = 0.064 \pm 0.002$, that highlights a weight persistence effect. This result complements the analysis of [29], as they considered the weighted networks of the number of loans between each pair of banks, neglecting altogether the amount lent for each loan. By design, our model allows us to highlight the tendency of banks to form links whose weight is positively related with the weight at previous steps, a tendency that we might refer to as weight persistence.

5. Conclusions

In this work, we proposed a model for the description of sparse and weighted temporal networks that extends the well known fitness model for static binary networks. In the new score driven weighted fitness model, we also model links' weights with an additional set of fitnesses. Both binary and weighted fitness follow a stochastic dynamics driven by the score of the conditional likelihood. Additionally, we also considered the possibility for the network dynamics to depend on a set of external covariates.

Our numerical simulations proved the advantages of the score driven fitness over static fitness and over a sequence of standard cross sectional estimates. As an empirical application, we investigated the determinants of the dynamics of links and weights in the e-MID interbank network. We proved that there is a significant advantage in using score driven time varying parameters to forecast weights, with respect to single snapshot estimates. We exploited the flexibility of the new model to estimate the impact of the EONIA rate in determining the links and weights dynamics. We used it to inform the discussion on persistence in interbank networks providing empirical evidence of weights' persistence.

We run an empirical analysis on the prediction of links and weights that highlighted the superior performances of our approach with respect to alternatives based on single snapshot estimation and link specific regressions. Most notably, for the dataset that we considered, we found our model to attain clearly superior performances with respect to the local likelihood Tobit model by [5]. In short, we believe this to be a direct consequence of our modelling choices that allow for a flexible description with a moderate number of parameters. Specifically, using zero augmentation we describe separately the probability of a link to exist and its expected weight, and leveraging the score driven approach we only need to estimate order N parameters instead of N^2 .

Our work provides several perspectives for future research. First, the possibility to jointly model and predict the presence of a link and the associated weight could find relevant applications in the financial stability literature. Weighted financial networks are known to be among the determinants of systemic risk and their dynamical description has so far neglected the role of the weights. Second, score driven weighted fitness model could be applied on multiple instances of real world sparse and weighted temporal networks, where the standard approach of ignoring the weights might result in significant information loss. Finally, we plan to investigate the temporal evolution of the community structure of real world networks. Community detection has attracted an enormous amount of attention in various streams of literature [71], in particular in the context of temporal networks [72]. We believe that the score driven

fitness approach could provide a flexible modeling framework and offer valuable support to assess the degree of persistence of a given partition of nodes into groups.

References

- [1] A.-L. Barabási, *Linked: The New Science Of Networks*, Basic Books, 2002 (2002).
- [2] M. Newman, *Networks: an introduction*, Oxford University Press, 2010 (2010).
- [3] S. Hanneke, W. Fu, E. P. Xing, et al., Discrete temporal models of social networks, *Electronic Journal of Statistics* 4 (2010) 585–605 (2010).
- [4] D. K. Sewell, Y. Chen, Latent space models for dynamic networks, *Journal of the American Statistical Association* 110 (512) (2015) 1646–1657 (2015).
- [5] L. Giraitis, G. Kapetanios, A. Wetherilt, F. Žikeš, Estimating the dynamics and persistence of financial networks, with an application to the sterling money market, *Journal of Applied Econometrics* 31 (1) (2016) 58–84 (2016).
- [6] P. Mazzarisi, P. Barucca, F. Lillo, D. Tantari, A dynamic network model with persistent links and node-specific latent variables, with an application to the interbank market, *European Journal of Operational Research* 281 (1) (2020) 50–65 (2020).
- [7] D. Di Gangi, G. Bormetti, F. Lillo, Score-driven exponential random graphs: A new class of time-varying parameter models for dynamical networks, *arXiv:1905.10806* (2019).
- [8] E. E. Leamer, J. Levinsohn, International trade theory: the evidence, *Handbook of International Economics* 3 (1995) 1339–1394 (1995).
- [9] G. Fagiolo, J. Reyes, S. Schiavo, The evolution of the world trade web: a weighted-network analysis, *Journal of Evolutionary Economics* 20 (4) (2010) 479–514 (2010).
- [10] G. Fagiolo, G. Santoni, Revisiting the role of migrant social networks as determinants of international migration flows, *Applied Economics Letters* 23 (3) (2016) 188–193 (2016).
- [11] M. E. Newman, Scientific collaboration networks. ii. shortest paths, weighted networks, and centrality, *Physical Review E* 64 (1) (2001) 016132 (2001).
- [12] A. Barrat, M. Barthelemy, R. Pastor-Satorras, A. Vespignani, The architecture of complex weighted networks, *Proceedings of the National Academy of Sciences* 101 (11) (2004) 3747–3752 (2004).

- [13] K. Anand, I. van Lelyveld, Á. Banai, S. Friedrich, R. Garratt, G. Halaj, J. Figue, I. Hansen, S. M. Jaramillo, H. Lee, et al., The missing links: A global study on uncovering financial network structures from partial data, *Journal of Financial Stability* 35 (2018) 107–119 (2018).
- [14] F. Allen, A. Babus, Networks in finance, in: P. R. JKleindorfer, Y. J. R. Wind, R. E. Gunther (Eds.), *The network challenge: strategy, profit, and risk in an interlinked world*, Social Science Research Network, 2011, Ch. 21, pp. 367–379 (2011).
- [15] A. G. Haldane, R. M. May, Systemic risk in banking ecosystems, *Nature* 469 (7330) (2011) 351–355 (2011).
- [16] L. Bargigli, G. Di Iasio, L. Infante, F. Lillo, F. Pierobon, The multiplex structure of interbank networks, *Quantitative Finance* 15 (4) (2015) 673–691 (2015).
- [17] P. Erdős, A. Rényi, On random graphs i., *Publ. Math. Debrecen* 6 (1959) 290–297 (1959).
- [18] D. R. Cox, G. Gudmundsson, G. Lindgren, L. Bondesson, E. Harsaae, P. Laake, K. Juselius, S. L. Lauritzen, Statistical analysis of time series: Some recent developments, *Scandinavian Journal of Statistics* (1981) 93–115 (1981).
- [19] D. Creal, S. J. Koopman, A. Lucas, Generalized autoregressive score models with applications, *Journal of Applied Econometrics* 28 (5) (2013) 777–795 (2013).
- [20] A. C. Harvey, *Dynamic models for volatility and heavy tails: with applications to financial and economic time series*, Vol. 52, Cambridge University Press, 2013 (2013).
- [21] P. W. Holland, S. Leinhardt, An exponential family of probability distributions for directed graphs, *Journal of the American Statistical Association* 76 (373) (1981) 33–50 (1981).
- [22] G. Caldarelli, A. Capocci, P. De Los Rios, M. A. Muñoz, Scale-free networks from varying vertex intrinsic fitness, *Physical Review Letters* 89 (2002) 258702 (2002).
- [23] D. Garlaschelli, M. I. Loffredo, Maximum likelihood: Extracting unbiased information from complex networks, *Physical Review E* 78 (2008) 015101 (2008).
- [24] S. Chatterjee, P. Diaconis, A. Sly, et al., Random graphs with a given degree sequence, *The Annals of Applied Probability* 21 (4) (2011) 1400–1435 (2011).

- [25] T. Yan, C. Leng, J. Zhu, et al., Asymptotics in directed exponential random graph models with an increasing bi-degree sequence, *The Annals of Statistics* 44 (1) (2016) 31–57 (2016).
- [26] T. Yan, B. Jiang, S. E. Fienberg, C. Leng, Statistical inference in a directed network model with covariates, *Journal of the American Statistical Association* 114 (526) (2019) 857–868 (2019).
- [27] Q. F. Akram, C. Christophersen, Interbank overnight interest rates-gains from systemic importance, Working Paper 11, Norges Bank (2010).
- [28] C. Brunetti, J. H. Harris, S. Mankad, G. Michailidis, Interconnectedness in the interbank market, *Journal of Financial Economics* 133 (2) (2019) 520–538 (2019).
- [29] V. Hatzopoulos, G. Iori, R. N. Mantegna, S. Micciché, M. Tumminello, Quantifying preferential trading in the e-mid interbank market, *Quantitative Finance* 15 (4) (2015) 693–710 (2015).
- [30] J. M. Wooldridge, *Econometric analysis of cross section and panel data*, MIT press, 2010 (2010).
- [31] B. Kim, K. H. Lee, L. Xue, X. Niu, et al., A review of dynamic network models with latent variables, *Statistics Surveys* 12 (2018) 105–135 (2018).
- [32] M. Billio, R. Casarin, M. Iacopini, Bayesian tensor regression models, in: *Mathematical and Statistical Methods for Actuarial Sciences and Finance*, Springer, 2018, pp. 149–153 (2018).
- [33] B. Betancourt, A. Rodríguez, N. Boyd, Investigating competition in financial markets: a sparse autologistic model for dynamic network data, *Journal of Applied Statistics* 45 (7) (2018) 1157–1172 (2018).
- [34] G. Cimini, T. Squartini, N. Musmeci, M. Puliga, A. Gabrielli, D. Garlaschelli, S. Battiston, G. Caldarelli, Reconstructing topological properties of complex networks using the fitness model, in: *Social Informatics*, Springer International Publishing, 2014, pp. 323–333 (2014).
- [35] J. Tinbergen, *Shaping the world economy; suggestions for an international economic policy* (1962).
- [36] K. Anand, B. Craig, G. Von Peter, Filling in the blanks: Network structure and interbank contagion, *Quantitative Finance* 15 (4) (2015) 625–636 (2015).
- [37] A. Gandy, L. A. Veraart, Compound poisson models for weighted networks with applications in finance, *Mathematics and Financial Economics* 15 (2021) 131–153 (2021).

- [38] P. A. Jacobs, P. A. Lewis, Discrete time series generated by mixtures. III. autoregressive processes (DAR (p))., Tech. rep., Naval Postgraduate School Monterey Calif (1978).
- [39] O. E. Williams, P. Mazzarisi, F. Lillo, V. Latora, Non-markovian temporal networks with auto- and cross-correlated link dynamics, *Physical Review E* (in press) (2022).
- [40] E. Zermelo, Die berechnung der turnier-ergebnisse als ein maximumproblem der wahrscheinlichkeit-rechnung, *Mathematische Zeitschrift* 29 (1) (1929) 436–460 (1929).
- [41] F. Blasques, S. J. Koopman, A. Lucas, Maximum likelihood estimation for generalized autoregressive score models, Tech. rep., Tinbergen Institute Discussion Paper (2014).
- [42] T. Bollerslev, Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* 31 (3) (1986) 307 – 327 (1986).
- [43] D. B. Nelson, Conditional heteroskedasticity in asset returns: A new approach, *Econometrica* (1991) 347–370 (1991).
- [44] R. F. Engle, J. R. Russell, Autoregressive conditional duration: a new model for irregularly spaced transaction data, *Econometrica* (1998) 1127–1162 (1998).
- [45] R. Engle, New frontiers for arch models, *Journal of Applied Econometrics* 17 (5) (2002) 425–446 (2002).
- [46] F. Blasques, S. J. Koopman, A. Lucas, Information-theoretic optimality of observation-driven time series models for continuous responses, *Biometrika* 102 (2) (2015) 325–343 (2015).
- [47] F. Blasques, A. Lucas, A. van Vlodrop, Finite sample optimality of score-driven volatility models, Tinbergen Institute Discussion Paper 17-111/III (2017).
- [48] A. Harvey, R. Ito, Modeling time series when some observations are zero, *Journal of Econometrics* 214 (1) (2020) 33–45 (2020).
- [49] P. E. Mistrulli, Assessing financial contagion in the interbank market: Maximum entropy versus observed interbank lending patterns, *Journal of Banking & Finance* 35 (5) (2011) 1114 – 1127 (2011).

- [50] A. Gandy, L. A. Veraart, Adjustable network reconstruction with applications to cds exposures, *Journal of Multivariate Analysis* 172 (2019) 193–209 (2019).
- [51] W. H. Greene, *Econometric analysis*, New Jersey: Prentice Hall, 2000 (2000).
- [52] H. Barreto, F. Howland, *Introductory econometrics: using Monte Carlo simulation with Microsoft excel*, Cambridge University Press, 2006 (2006).
- [53] Z. Griliches, Estimating the returns to schooling: Some econometric problems, *Econometrica* (1977) 1–22 (1977).
- [54] A. Yatchew, Z. Griliches, Specification error in probit models, *The Review of Economics and Statistics* (1985) 134–139 (1985).
- [55] K. A. Clarke, The phantom menace: Omitted variable bias in econometric research, *Conflict Management and Peace Science* 22 (4) (2005) 341–352 (2005).
- [56] C. Green, Y. Bai, V. Murinde, K. Ngoka, I. Maana, S. Tiriongo, Overnight interbank markets and the determination of the interbank rate: A selective survey, *International Review of Financial Analysis* 44 (2016) 149–161 (2016).
- [57] G. Iori, G. De Masi, O. V. Precup, G. Gabbi, G. Caldarelli, A network analysis of the italian overnight money market, *Journal of Economic Dynamics and Control* 32 (1) (2008) 259–278 (2008).
- [58] K. Finger, D. Fricke, T. Lux, Network analysis of the e-mid overnight money market: the informational value of different aggregation levels for intrinsic dynamic processes, *Computational Management Science* 10 (2-3) (2013) 187–211 (2013).
- [59] P. Barucca, F. Lillo, The organization of the interbank network and how ecb unconventional measures affected the e-mid overnight market, *Computational Management Science* 15 (1) (2018) 33–53 (2018).
- [60] L. Lü, T. Zhou, Link prediction in complex networks: A survey, *Physica A: statistical mechanics and its applications* 390 (6) (2011) 1150–1170 (2011).
- [61] P. Wang, B. Xu, Y. Wu, X. Zhou, Link prediction in social networks: the state-of-the-art, *Science China Information Sciences* 58 (1) (2015) 1–38 (2015).

- [62] V. Martínez, F. Berzal, J.-C. Cubero, A survey of link prediction in complex networks, *ACM computing surveys (CSUR)* 49 (4) (2016) 1–33 (2016).
- [63] S. Haghani, M. R. Keyvanpour, A systemic analysis of link prediction in social network, *Artificial Intelligence Review* (2017) 1–35 (2017).
- [64] F. X. Diebold, R. S. Mariano, Comparing predictive accuracy, *Journal of Business & Economic Statistics* 20 (1) (2002) 134–144 (2002).
- [65] D. Harvey, S. Leybourne, P. Newbold, Testing the equality of prediction mean squared errors, *International Journal of Forecasting* 13 (2) (1997) 281–291 (1997).
- [66] G. Iori, K. Burcu, J. Olmo, Bank characteristics and the interbank money market: a distributional approach, *Studies in Nonlinear Dynamics & Econometrics* 19 (3) (2015) 249–283 (2015).
- [67] L. Arciero, R. Heijmans, R. Heuver, M. Massarenti, C. Picillo, F. Vacirca, How to measure the unsecured money market: The eurosystem’s implementation and validation using target2 data, *International Journal of Central Banking* 12 (1) (2016) 247–280 (2016).
- [68] A. Temizsoy, G. Iori, G. Montes-Rojas, Network centrality and funding rates in the e-mid interbank market, *Journal of Financial Stability* 33 (2017) 346–365 (2017).
- [69] G. Weisbuch, A. Kirman, D. Herreiner, Market organisation and trading relationships, *The Economic Journal* 110 (463) (2000) 411–436 (2000).
- [70] J. F. Cocco, F. J. Gomes, N. C. Martins, Lending relationships in the interbank market, *Journal of Financial Intermediation* 18 (1) (2009) 24–48 (2009).
- [71] M. A. Javed, M. S. Younis, S. Latif, J. Qadir, A. Baig, Community detection in networks: A multidisciplinary review, *Journal of Network and Computer Applications* 108 (2018) 87–111 (2018).
- [72] G. Rossetti, R. Cazabet, Community discovery in dynamic networks: a survey, *ACM Computing Surveys (CSUR)* 51 (2) (2018) 35 (2018).
- [73] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, S. Chintala, Pytorch: An imperative style, high-performance deep

learning library, in: H. Wallach, H. Larochelle, A. Beygelzimer, F. d Alché-Buc, E. Fox, R. Garnett (Eds.), Advances in Neural Information Processing Systems 32, Curran Associates, Inc., 2019, pp. 8024–8035 (2019).

URL <http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>

[74] C. Spearman, The proof and measurement of association between two things. (1961).

Appendix A. Alternative Distributions for the Weights

In the main text, we employed the gamma distribution to model the links’ weights. It is extremely easy to move to different specifications. If, for example, we consider the log-normal distribution, we can substitute its parametric form

$$g_{ij}(y) = \frac{1}{y\sigma\sqrt{2\pi}} e^{-\frac{(\ln y - \mu_{ij})^2}{2\sigma^2}}$$

in place of the gamma distribution, and obtain a different update equation for the time varying parameters. The fitness would be driven by the score of the log-normal

$$\nabla_i \left(\mathbf{Y}^{(t)}, \overleftarrow{\eta}, \overrightarrow{\eta}, \sigma \right) = \sum_j \left(\frac{1}{\sigma^2} \log \frac{\mathbb{E} \left[Y_{ij}^{(t)} | Y_{ij}^{(t)} > 0 \right]}{Y_{ij}^{(t)}} - \mathbb{I}(Y_{ij}) \right).$$

Indeed, extending our approach to different specifications is practically straightforward. The interested reader would only need to compute the log-likelihood, sample from the distribution, and compute the score and scaling matrix. Conveniently the log-likelihood and sampling methods are already available for many distributions in PyTorch [73], upon which we built the core library for this work available at <https://github.com/domenicodigangi/dynwgraphs>.

Appendix B. Fitness Identification

It is easy to see that, as for the static version, the fitness parameters of the fitness model for a directed network are not identified. This fact is well known for the binary fitness model [25] and remains true for the weighted fitness that we consider here. Let us indicate here with $\overleftarrow{\varphi}$ and $\overrightarrow{\varphi}$ the vectors of in and out fitness respectively, without specifying whether they are binary or weighted. With this notation it

is immediately clear that the following transformation

$$\begin{aligned}\overleftarrow{\varphi}_i &\rightarrow \overleftarrow{\varphi}_i + c, \quad \forall i \\ \overrightarrow{\varphi}_i &\rightarrow \overrightarrow{\varphi}_i - c, \quad \forall i\end{aligned}$$

does not change the sum $\overleftarrow{\varphi}_i + \overrightarrow{\varphi}_j + \beta X_{ij}$ and leaves the probability distribution unchanged, for any $c \in \mathbb{R}$. The parameters are not uniquely identified and, in order to compare estimates across different time steps, we need to impose an identification restriction, as commonly done in the literature [25, 6]. In this work we require that

$$\sum_i \overleftarrow{\varphi}_i = \sum_j \overrightarrow{\varphi}_j,$$

at each time step.

Appendix C. Filtered Fitness Dynamics Wwith External Covariates

Here, we inspect the dynamical behaviour and the role of the fitness filtered with the score driven weighted fitness model, defined in Section 2.2, in presence of external covariates. In particular, we discuss how the dynamical evolution of the filtered fitness changes when we consider the dependency on EONIA, estimated as described in Section 4.2, with respect to the case of no covariate.

To this end, for a subset of the most active links ¹¹, we compute the sum of the in and out fitness corresponding to their probabilities, and expected weights. For example, for link (i, j) we consider $\widehat{\theta}_i + \widehat{\theta}_j$ and $\widehat{\eta}_i + \widehat{\eta}_j$. We compute these quantities using two sets of filtered fitness. The first set is obtained by estimating and running the score driven weighted fitness model as a filter without considering any external covariate, similarly to what we did in Section 4. The second set is obtained with a similar approach, now including in the model specification also the dependency on EONIA, exactly as in Section 4.2. Then, given these two sets of fitness sums we look at their Spearman rank correlation [74] over time with EONIA, thus obtaining four correlations for each considered link: for each of the two sets of filtered fitness (with and without EONIA as covariate) we have one correlation for the binary fitness and one for the weighted fitness. We consider the sum of fitness because, as it is evident from (3) and (4), they are related to the probability of observing a link and to the expected weight of that link. Moreover,

¹¹We consider the links that are present at least 5% of the weeks.

given that the EONIA rate had a substantial drop around the middle of 2012, as showed in Figure C.4, in computing the correlation with the sums of fitness we split the data set in two periods and consider them separately.

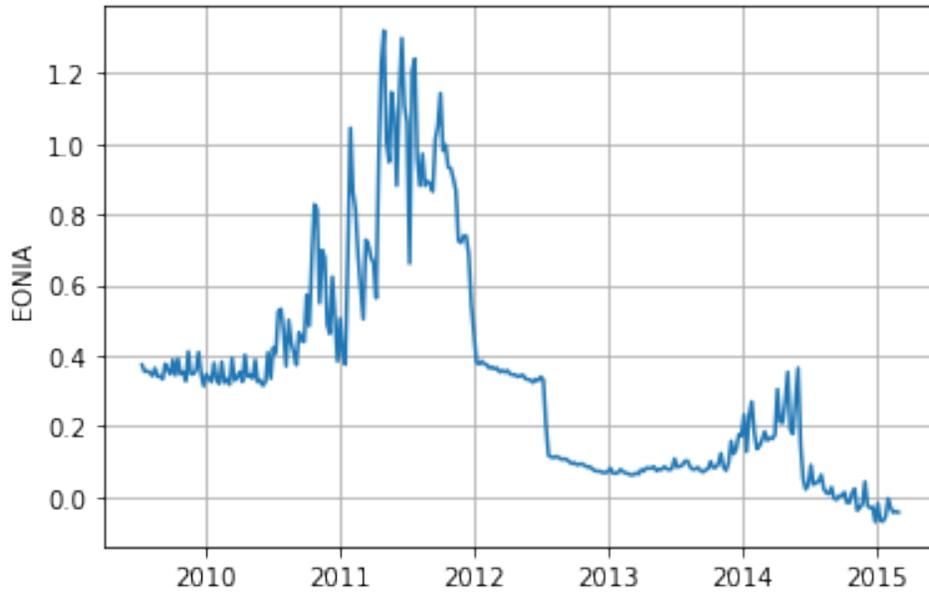


Figure C.4: Weekly average of EONIA rate over the considered time period.

The plots shown in Figure C.5 highlight the different behaviour of the fitness, with respect to EONIA, when the latter is explicitly considered as an external covariate. We notice that in the first period, i.e. when EONIA rates are higher, the correlation between the fitness sums and EONIA are significantly different when we include EONIA as covariate or we do not. To statistically confirm the difference we run a Kolmogorow-Smirnov test having as null hypothesis the fact that the two distributions are identical. In the first period, the null hypothesis is rejected at 5% confidence level for both the binary and weighted part. In the second period instead, we cannot reject the hypotheses that the two distributions are the same. Hence the relation of the fitness with EONIA in the second period does not seem to be affected by whether the latter is explicitly included or not as external covariate.

We believe that the tendency of the dynamical fitness to correlate with EONIA, when the latter is not considered explicitly as an external variable, corroborates the fact that the estimated coefficients are found to be significantly different from zero.

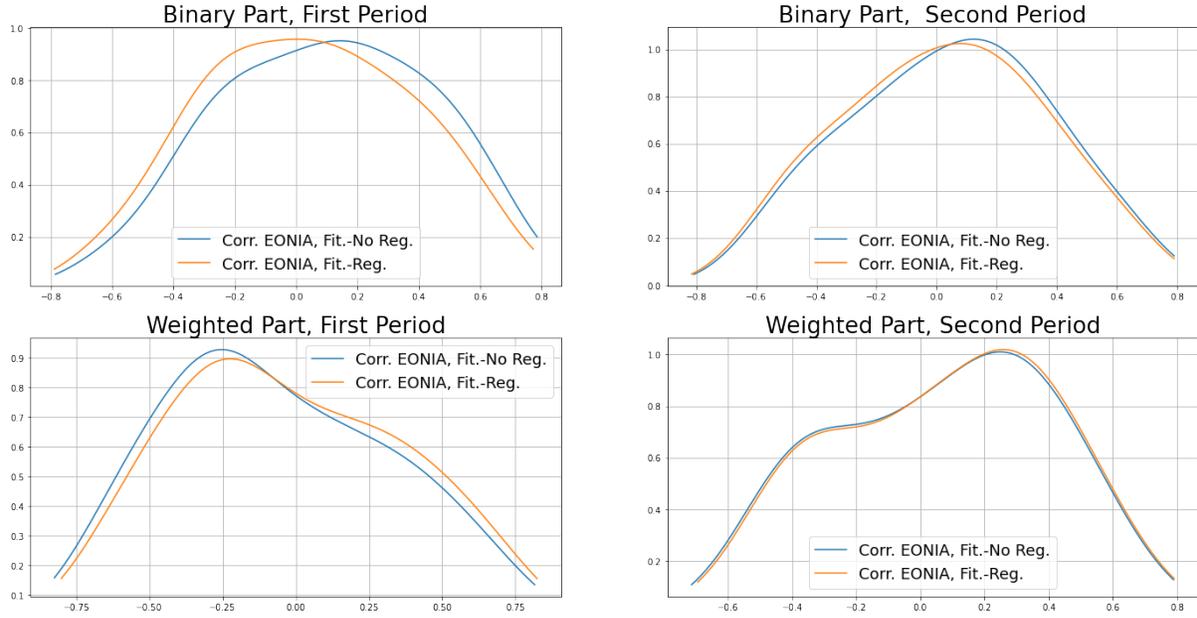


Figure C.5: Distribution of the Spearman correlation between the fitness sums and EONIA computed in two periods. The first period ranges from the beginning of the sample up to 15/07/2012, the second one from 22/07/2012 to the end of the sample. Left panels: correlations computed for the first period; right panels: correlations computed in the last period. Top panels: correlations between EONIA and the sums of binary fitness; bottom panels: correlations with the weighted fitness.

In order to better understand this behaviour we repeated the same analysis artificially modifying the regression coefficients before filtering the fitness. We keep everything equal except the values of β_{bin} and β_w . Instead of their MLE values, we now artificially set the values to $\beta_{bin} = 3$ and $\beta_w = 3$ in Figure C.7 and $\beta_{bin} = -3$ and $\beta_w = -3$ in Figure C.6. In practice, we change the regression coefficients and then filter the time varying fitness using the score driven update rule. We notice that, when the regression coefficient is artificially inflated, see Figure C.7, the fitness sums tend to negatively correlate with EONIA. While, when we force the coefficients to be negative, the fitness positively correlate with EONIA. In both cases the filtered fitness behaviour tends to mitigate the impact of the spurious regression coefficients. These effect is more evident in the first period than in the second, as expected due to the higher values of EONIA.

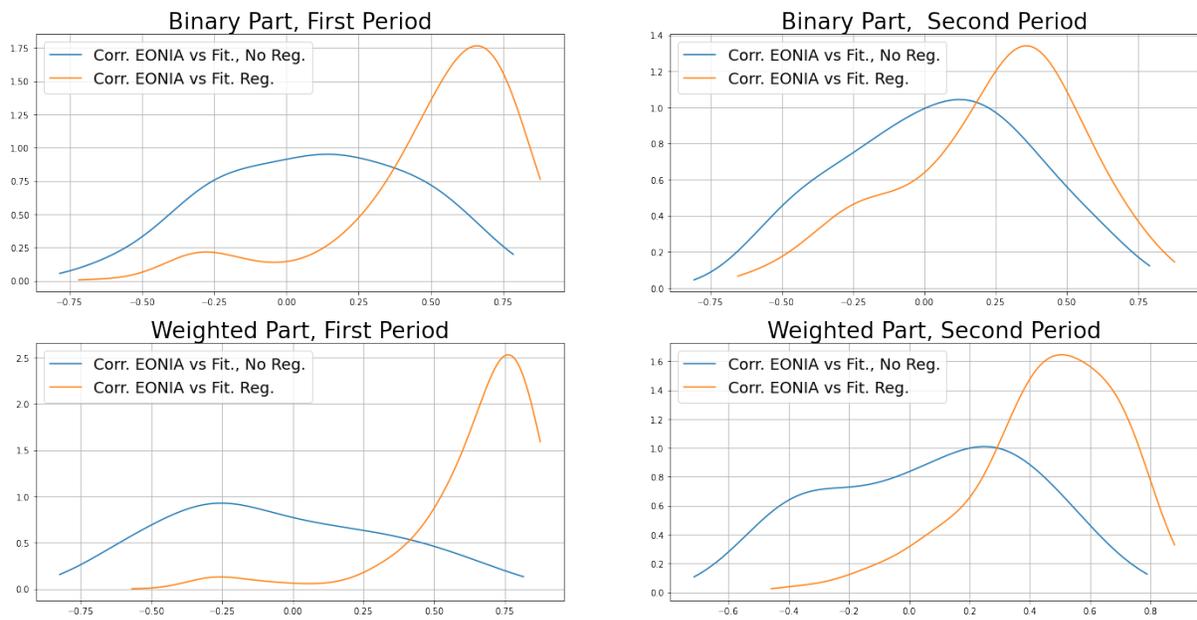


Figure C.6: Similar plots as in Figure C.5, now repeated after forcing the values of the regression coefficients to $\beta_{bin} = -3$ and $\beta_w = -3$.

for each one of 1860.0 links compute spearman of its fitness sum vs eonia

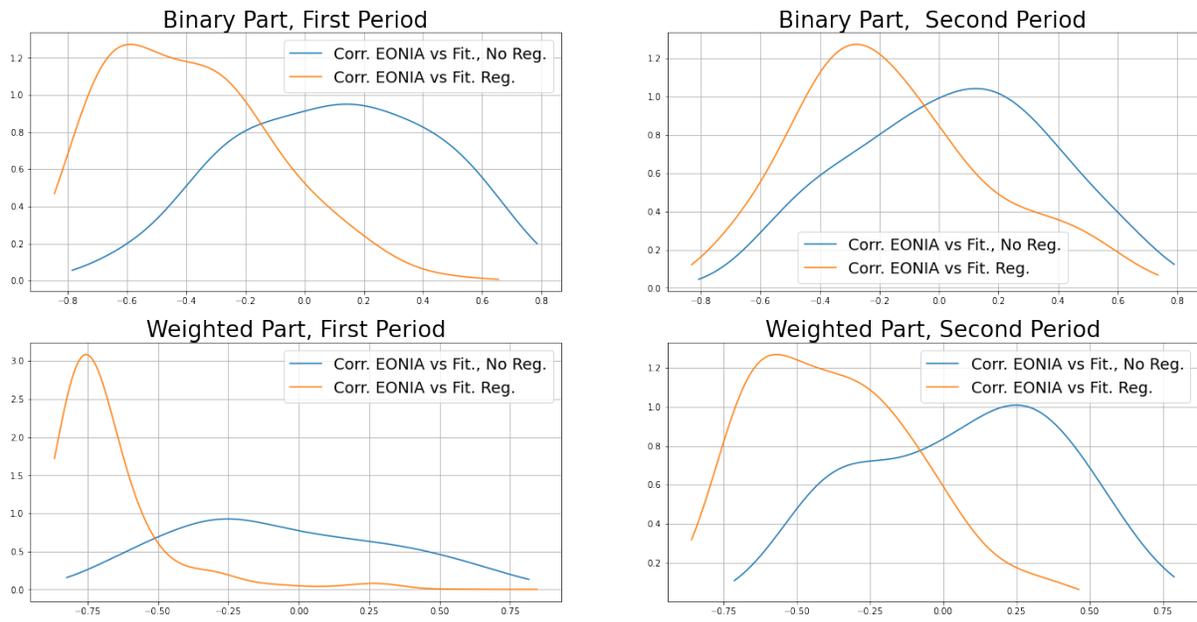


Figure C.7: Similar plots as in Figure C.5, now repeated after forcing the values of the regression coefficients to $\beta_{bin} = 3$ and $\beta_w = 3$.