

# Dynamic communities in multichannel data: An application to the foreign exchange market during the 2007–2008 credit crisis

Daniel J. Fenn,<sup>1,2,a)</sup> Mason A. Porter,<sup>3,2</sup> Mark McDonald,<sup>4</sup> Stacy Williams,<sup>4</sup>  
Neil F. Johnson,<sup>5,2</sup> and Nick S. Jones<sup>6,7,2</sup>

<sup>1</sup>*Mathematical and Computational Finance Group, Mathematical Institute, University of Oxford, Oxford OX1 3LB, United Kingdom*

<sup>2</sup>*CABDyN Complexity Centre, University of Oxford, Oxford OX1 1HP, United Kingdom*

<sup>3</sup>*Oxford Centre for Industrial and Applied Mathematics, Mathematical Institute, University of Oxford, Oxford OX1 3LB, United Kingdom*

<sup>4</sup>*FX Research and Trading Group, HSBC Bank, 8 Canada Square, London E14 5HQ, United Kingdom*

<sup>5</sup>*Department of Physics, University of Miami, Coral Gables, Florida 33146, USA*

<sup>6</sup>*Department of Physics, Clarendon Laboratory, University of Oxford, Oxford OX1 3PU, United Kingdom*

<sup>7</sup>*Oxford Centre for Integrative Systems Biology, University of Oxford, Oxford OX1 3QU, United Kingdom*

(Received 27 March 2009; accepted 29 June 2009; published online 19 August 2009)

We study the cluster dynamics of multichannel (multivariate) time series by representing their correlations as time-dependent networks and investigating the evolution of network communities. We employ a node-centric approach that allows us to track the effects of the community evolution on the functional roles of individual nodes without having to track entire communities. As an example, we consider a foreign exchange market network in which each node represents an exchange rate and each edge represents a time-dependent correlation between the rates. We study the period 2005–2008, which includes the recent credit and liquidity crisis. Using community detection, we find that exchange rates that are strongly attached to their community are persistently grouped with the same set of rates, whereas exchange rates that are important for the transfer of information tend to be positioned on the edges of communities. Our analysis successfully uncovers major trading changes that occurred in the market during the credit crisis. © 2009 American Institute of Physics. [DOI: 10.1063/1.3184538]

**The past decade has seen an explosion of research on static networks and on detecting cohesive subnetworks known as communities. However, despite all of this attention, very little research has considered the problem of investigating communities in time-evolving networks; these occur across disciplines, from economics to biology. In this paper, we examine communities in an evolving, weighted, fully connected network that we obtained from high-frequency time series of currency exchange rates. Community detection in time-dependent weighted networks provides a new approach to the problem of clustering multivariate time series. In contrast to previous studies of dynamic communities in networks, we employ a node-centric approach that allows us to track the effects of the community evolution on the functional roles of nodes without having to track entire communities. To illustrate the power of our approach, we consider a foreign exchange (FX) market network during the recent credit crisis. Our analysis uncovers notable structural changes that occurred in the FX market during the crisis and identifies exchange rates that experienced significant, striking changes in market role.**

## I. INTRODUCTION

During the past ten years, there has been an explosion of research on networks (i.e., graphs) with static connections between nodes.<sup>1–3</sup> This has involved a remarkable interdisciplinary effort—with physicists, mathematicians, computer scientists, sociologists, and others all playing major roles. Among the most important ideas in network science is the investigation of mesoscopic network structures known as “communities,”<sup>4–13</sup> which are constructed from subsets of nodes that are more strongly connected to each other than they are to the rest of the network. The analysis of communities has provided striking insights into functional modules in several networks.<sup>6,11–18</sup>

However, despite this wealth of attention, there has been much less research that considers communities in the more general class of networks that have time-evolving weighted links.<sup>19</sup> Here, we examine the community dynamics in an evolving, fully connected FX market network over the period 2005–2008 (which includes the recent credit crisis). The network possesses a fixed number of nodes and evolving link weights that are determined using time-varying pairwise correlations between time series associated with each node. Community detection in networks of this kind is equivalent to the problem of clustering multivariate time series. Time series clustering has a range of applications, including finding groups of genes with similar expression patterns in mi-

<sup>a)</sup>Electronic mail: fenn@maths.ox.ac.uk.

croarray experiments and grouping functional MRI time series to identify regions of the brain with similar activation patterns.<sup>20</sup> Again, however, there has been little focus on the dynamics of clusters.

In this paper, we present a methodology for investigating cluster/community dynamics. Recently, scientists have begun to develop techniques to examine the dynamics of communities by tracking entire communities through time.<sup>19,21,22</sup> This requires a method for determining which community at each time step represents the descendant of a community at the previous time step, which can lead to equivocal mappings following splits and mergers. As an alternative, we investigate community dynamics from a *node-centric* perspective and demonstrate how the dynamics can affect the functional roles of individual nodes. Our analysis uncovers major changes that occurred in the FX market during the credit crisis. Among other things, we identify individual exchange rates that experienced significant changes in market role. Our approach can potentially provide similarly useful insights into other multivariate data sets from commerce to medicine.

## II. DATA

We construct networks with  $n=110$  nodes, where each node represents an exchange rate of the form XXX/YYY (with  $XXX \neq YYY$ ), and  $XXX, YYY \in \{\text{AUD, CAD, CHF, GBP, DEM, JPY, NOK, NZD, SEK, USD, XAU}\}$ . The currency symbols are defined as follows: AUD, Australian dollar; CAD, Canadian dollar; CHF, Swiss franc; EUR, euro; GBP, pounds sterling; JPY, Japanese yen; NOK, Norwegian krone; NZD, New Zealand dollar; SEK, Swedish krona; USD, US dollar; XAU, gold. An exchange rate XXX/YYY indicates the amount of currency YYY one can receive in exchange for one unit of XXX. We include gold (XAU) in the study because it behaves like a volatile currency.<sup>23</sup>

We define the weight of the link connecting nodes  $i$  and  $j$  using the time series of hourly exchange rate returns  $R_i(t)$  ( $i=1, 2, \dots, n$ ) from the hours 07:00–18:00 UK time over the period 2005–2008. The return of an exchange rate with price  $p_i(t)$  at discrete time  $t$  is defined by  $R_i(t) = \ln(p_i(t)/p_i(t-1))$ . We represent the resulting fully connected, weighted network by an adjacency matrix  $\mathbf{A}$  with components

$$A_{ij} = \frac{1}{2}(\rho_{ij} + 1) - \delta_{ij}, \quad (1)$$

where  $\rho_{ij} = (\langle R_i R_j \rangle - \langle R_i \rangle \langle R_j \rangle) / (\sigma_i \sigma_j)$  is the correlation coefficient between exchange rates  $i$  and  $j$  over a window of  $T$  returns, the Kronecker delta  $\delta_{ij}$  removes self-edges,  $\langle \cdot \rangle$  indicates a time-average over  $T$ , and  $\sigma_i$  is the standard deviation of  $R_i$  over  $T$ . The matrix elements  $A_{ij} \in [0, 1]$  quantify the similarity of two exchange rates. We use the linear correlation coefficient  $\rho_{ij}$  to measure the correlation between pairs of exchange rates because of its simplicity, but our methods are independent of this choice and can be employed using alternative measures capable of detecting more general dependencies.

We exclude self-edges in order to deal with simple graphs. This approach was also taken in a previous study of a financial network derived from a correlation matrix.<sup>24</sup>

However, we note that if we include self-edges, the node compositions of the identified communities are identical if we make a small parameter change in the community detection algorithm. We discuss the community detection algorithm and the effect of including self-edges in Sec. III and in endnotes 25 and 26.

We create a longitudinal sequence of networks by consecutively displacing the time windows by  $\Delta t=20$  h (approximately two trading days) and fix  $T=200$  h (approximately 1 month of data) as a compromise between oversmoothing and overly noisy correlation coefficients.<sup>27</sup>

In contrast to most prior studies of financial networks, we use recent community detection techniques<sup>11–13</sup> rather than traditional hierarchical clustering<sup>23,27,28</sup> and consider weighted links rather than thresholding them to create a binary network.<sup>29</sup> In Ref. 24, the authors also departed from such typical assumptions in a study of communities in a network of equities. However, they did not examine longitudinal networks. Dynamic communities have been investigated in biophysical data using recent methods,<sup>30</sup> but our work differs by focusing on quantifying and tracking the changes in community composition.

## III. COMMUNITY DETECTION

We represent each network  $\mathbf{A}$  as an infinite-range,  $n$ -state Potts spin glass so that each node is a spin, each edge is a pairwise interaction between spins, and each community is a spin state. We then find communities by assigning each spin to one state and minimizing the interaction energy of these states, which is given by the Hamiltonian<sup>8</sup>

$$\mathcal{H} = - \sum_{ij} J_{ij} \delta(c_i, c_j), \quad (2)$$

where  $c_i$  is the state of spin  $i$ . The interaction energy between spins  $i$  and  $j$  is  $J_{ij} = A_{ij} - \gamma p_{ij}$ , where  $\gamma$  is a resolution parameter and  $p_{ij}$  denotes the expected edge weight with which nodes  $i$  and  $j$  are connected in a null model of random link assignment. The Kronecker delta  $\delta(c_i, c_j)$  ensures that the sum is only taken over nodes belonging to the same community. We employ the standard Newman–Girvan null model (or prior) so that  $p_{ij} = k_i k_j / (2m)$ , where  $k_i = \sum_j A_{ij}$  is the strength (weighted degree) of node  $i$  and  $m = \frac{1}{2} \sum_{i,j} A_{ij}$  is the total edge weight in the network.<sup>7</sup> For the particular example network considered here, each of the nodes has the same strength  $k_i = (n-2)/2$  and the expected edge weight in the null model is given by

$$p_{ij} = \frac{n-2}{2n}, \quad (3)$$

so the null model is equivalent to the uniform case (with constant  $p_{ij}$ ). However, the methods we present are general and can also be applied to networks with nonuniform strength distributions.<sup>25</sup>

By tuning  $\gamma$ , we probe the community structure at different resolutions, with larger  $\gamma$  values corresponding to more fragmented communities.<sup>8,24</sup> At each resolution, we use a greedy algorithm<sup>31</sup> to minimize Eq. (2).<sup>26</sup>

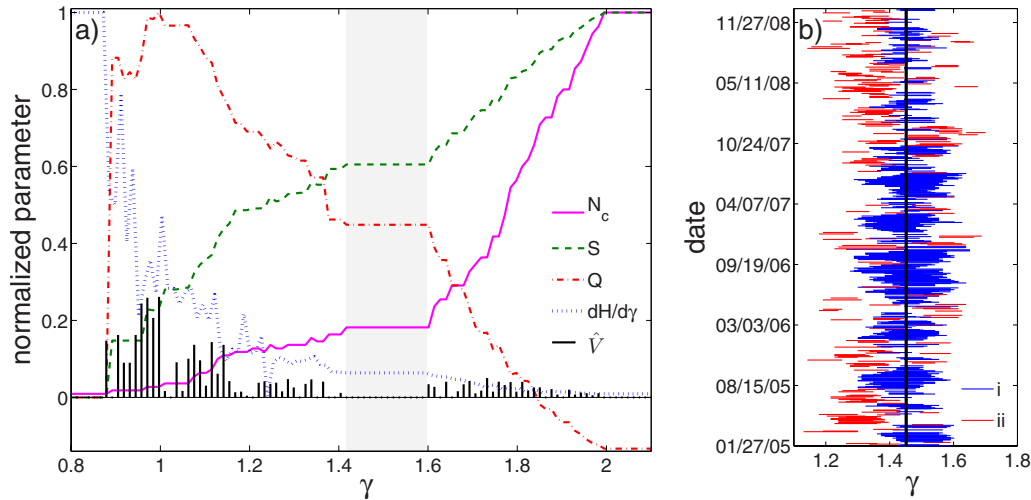


FIG. 1. (Color online) (a) The quantities  $N_c$ ,  $S$ ,  $Q$ , and  $d\mathcal{H}/d\gamma$  (normalized by their maximum values) as a function of the resolution parameter  $\gamma$  for a single time window beginning on 10 March 2005. The shaded gray area highlights the main plateau. The bottom curve gives the normalized variation of information between partitions at resolutions separated by  $\Delta\gamma \approx 0.013$ . (b) The position of the main plateau at each time step, showing the main plateaus (i) containing  $\gamma=1.45$  and (ii) not containing  $\gamma=1.45$ .

The community structure of many networks is robust across a range of resolutions.<sup>8,9,11,12</sup> Robust communities obtained using the Potts method are significant because they persist even though there is a larger incentive for nodes to belong to smaller clusters as  $\gamma$  increases. We detect communities at 100 resolutions in the interval  $\gamma \in [0.8, 2.1]$ . At  $\gamma=0.8$  all of the nodes are assigned to the same community, and at  $\gamma=2.1$  they are all assigned to singleton communities. We compare the community compositions at consecutive values of  $\gamma$  using the normalized variation of information  $\hat{V}$ .<sup>15,32</sup> The entropy of a partition  $C$  of the  $n$  nodes in  $\mathbf{A}$  into  $K$  communities  $c^k$  ( $k \in \{1, \dots, K\}$ ) is

$$S(C) = - \sum_{k=1}^K P(k) \log P(k), \quad (4)$$

where  $P(k) = |c^k|/n$  is the probability that a randomly selected node belongs to community  $k$  and  $|c^k|$  is the size (set cardinality) of communities. (The quantity  $c^k$  is, therefore, the set of nodes labeled by  $k$ , while  $c_i$  is the set of nodes in the same community as node  $i$ .) Given a second partition  $C'$  of the  $n$  nodes into  $K'$  communities, we define

$$\hat{V}(C, C') = \frac{S(C) + S(C') - 2I(C, C')}{\log n}. \quad (5)$$

The quantity  $I(C, C')$ , the mutual information between  $C$  and  $C'$ , is given by

$$I(C, C') = \sum_{k=1}^K \sum_{k'=1}^{K'} P(k, k') \log \frac{P(k, k')}{P(k)P(k')}, \quad (6)$$

where  $P(k, k') = |c^k \cap c^{k'}|/n$ . The factor  $\log n$  normalizes  $\hat{V}(C, C')$  to the interval  $[0, 1]$ , with 0 indicating identical partitions and 1 indicating that all nodes are in individual communities in one partition and in a single community in the other.

One can find robust communities by examining summary statistics that describe community configurations as a function of the resolution parameter. Figure 1(a) shows, as a function of  $\gamma$ , the number of communities  $N_c$ , the rate of change in the energy with resolution  $d\mathcal{H}/d\gamma$ , the modularity  $Q$ ,<sup>5</sup> and the entropy  $S$  of each partition. One can write a scaled energy  $Q_s$  in terms of the Hamiltonian in Eq. (2) as  $Q_s = -\mathcal{H}/2m$ . The modularity  $Q$  is the scaled energy with  $\gamma=1$ .

There are four principle plateaus in Fig. 1(a), corresponding to partitions of the network into  $N_c=1, 2, 20$ , and 110 communities. The first and last plateaus, respectively, represent all nodes in a single community and all nodes in individual communities. The second plateau represents one community of exchange rates and a corresponding community of inverse rates. We include inverse exchange rates, because one cannot infer *a priori* whether a rate XXX/YYYY will form a community with a rate WWW/ZZZ or its inverse ZZZ/WWW. The existence of an equivalent inverse community for each community means that at each time step the network is composed of two identical halves. However, the exchange rates residing in each half change with time as the correlations evolve. For the example in Fig. 1(a), the other plateau occurs in the interval  $\gamma \in [1.41, 1.60]$ . Although the community configuration over this interval does not have maximal  $Q$ , it provides an appropriate resolution at which to investigate community dynamics due to its resolution robustness with  $\gamma$  and the financially interesting features of the detected communities. For the remainder of this paper, we will refer to this plateau as the “main” plateau.

#### IV. COMMUNITY PROPERTIES

One way to investigate the community dynamics is to select a resolution  $\gamma$  in the main plateau at each time step. As shown in Fig. 1(b), this plateau occurs over different  $\gamma$  intervals at different time steps and has different widths. These intervals need not share common resolution values, so this

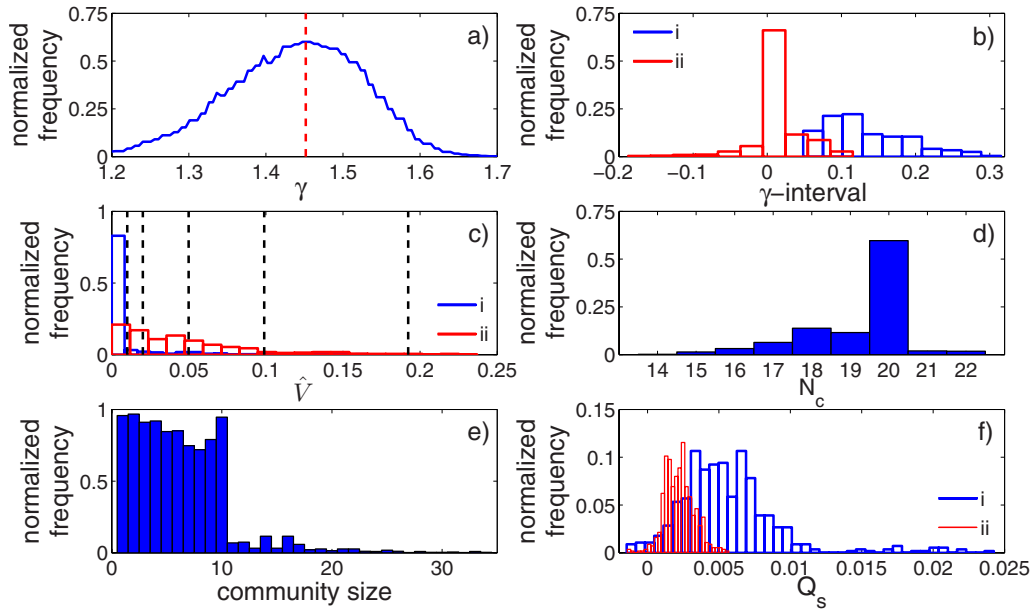


FIG. 2. (Color online) (a) Observed fraction of time steps that the resolution  $\gamma$  lies in the main plateau. The vertical line indicates the value  $\gamma=1.45$ , which lies in the largest number of main plateaus and is the resolution at which we investigate the community dynamics. (b) (i) Normalized sampled distribution of the main plateau width and (ii) normalized sampled distribution of the  $\gamma$ -distance between the main plateau and  $\gamma=1.45$ . (c) (i) Distribution of normalized variation of information between the community configuration at  $\gamma=1.45$  and the configuration corresponding to the main plateau and (ii) distribution of normalized variation of information between consecutive time steps. The vertical lines give the mean  $\hat{V}$  when (left to right) 1, 2, 5, 10, and 20 nodes are randomly reassigned to different communities (averaged over 100 reassignments for each time step). (d) The fraction of time steps at which  $N_c$  communities are observed. (e) The fraction of time steps at which a community of a given size is observed. (f) Comparison of the distribution of the scaled energy for (i) market data and (ii) 100 realizations of shuffled data.

method seems inappropriate because one would then be comparing communities obtained at different resolutions. We therefore fix  $\gamma=1.45$ , which, as shown in Fig. 2(a), is the value that occurs in the largest number of main plateaus. In order to demonstrate the validity of this choice, we show in Fig. 2(b) the distribution of the  $\gamma$ -distance between the fixed resolution  $\gamma=1.45$  and the main plateau, and in Fig. 2(c) the distribution of  $\hat{V}$  between the community configuration obtained using  $\gamma=1.45$  and that corresponding to the main plateau. The fixed resolution is in a  $\gamma$ -interval of less than 0.05 from the main plateau 88% of the time, and the community configurations of the main plateau and  $\gamma=1.45$  differ in the community assignments by fewer than two nodes 88% of the time. These results support our proposed method of investigating the community dynamics at a fixed  $\gamma$ .

Figure 2(d) shows that there are only small fluctuations in the number of communities into which the network is partitioned. Nevertheless, as shown in Fig. 2(c), there is a considerable variation in the extent of community reorganization between consecutive time steps. There are no changes between some steps, but more than 20 nodes change communities between others. Figure 2(e) shows that the observed community size distribution is bimodal and has a long tail extending to large community sizes. There is a large variation in the sizes of the communities observed at each time step.

To ensure that the identified communities are meaningful, we compare the scaled energies of the observed community partitions with those from shuffled data. We generate shuffled data by randomly reordering the elements of each of the real time series. By inspection, Fig. 2(f) shows that the

communities identified for the actual data are significantly stronger than those generated using the shuffled data. The sample mean scaled energy for the actual data is 0.0056 (with a standard deviation of 0.0036) and for shuffled data the sample mean is 0.0022 (with a standard deviation of 0.0010).

## V. NODE ROLES IN COMMUNITIES

Having considered the properties of entire communities, we now investigate the roles of nodes within communities. A node's identity is known at all times and its community is known at any one time. We can thus track community evolution from the perspective of individual nodes. We describe the relationship between a node and its community using various centrality measures. The betweenness centrality of a node  $i$  is computed by calculating, for every pair of vertices  $s$  and  $t$ , the fraction of geodesic paths between  $s$  and  $t$  that pass through  $i$  and then summing this fraction over all  $s$  and  $t$ .<sup>1,33</sup> We take the distance between nodes  $i$  and  $j$  as  $d_{ij} = 1/A_{ij}$  for  $i \neq j$  and 0 for  $i=j$ . In some sense, betweenness centrality measures the importance of each node for the spread of information around the network.

We also consider the community centrality of each node.<sup>7</sup> We define a scaled energy matrix  $\mathbf{J}$  by  $J_{ij}=A_{ij}-\gamma p_{ij}$ , where we again set  $p_{ij}=k_i k_j / (2m)$ . Following the notation in Ref. 7, this matrix can be expressed as  $\mathbf{J}=\mathbf{U}\mathbf{D}\mathbf{U}^T$ , where  $\mathbf{U}=(\mathbf{u}_1|\mathbf{u}_2|\dots)$  is the matrix of eigenvectors of  $\mathbf{J}$  and  $\mathbf{D}$  is the diagonal matrix of eigenvalues  $\beta_i$ . If  $\mathbf{D}$  has  $q$  positive eigenvalues, one can define a set of node vectors  $\mathbf{x}_i$  by

$$[\mathbf{x}_i]_j = \sqrt{\beta_j} U_{ij}, \quad j \in \{1, 2, \dots, q\}. \quad (7)$$



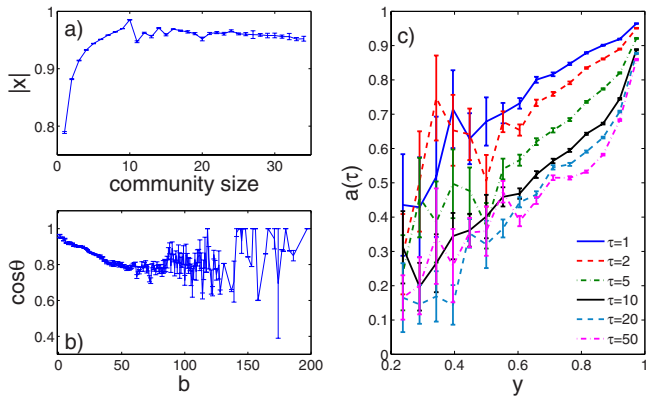


FIG. 3. (Color online) (a) Mean community centrality vs the size of the community to which a node belongs. (b) Mean community alignment vs the betweenness centrality of nodes. (c) Mean community autocorrelation vs the projected community centrality. [All error bars indicate the standard error (Ref. 34).]

The magnitude  $|\mathbf{x}_i|$  is the *community centrality*. A node with high community centrality plays an important role in its local neighborhood irrespective of the community boundaries.

One can also define a community vector

$$\mathbf{w}_k = \sum_{i \in c^k} \mathbf{x}_i \tag{8}$$

for each community  $k$  with members  $c^k$ . Nodes with high community centrality are strongly attached to their community if their node vector is also aligned with their community vector. We then define a *projected community centrality*  $y_i$  by

$$y_i = \mathbf{x}_i \cdot \hat{\mathbf{w}}_k = |\mathbf{x}_i| \cos \theta_{ik}, \tag{9}$$

where  $\hat{\mathbf{w}}_k$  is the unit vector in the direction of  $\mathbf{w}_k$ . We refer to the quantity  $\cos \theta_{ik}$  as the *community alignment*. The community alignment is near 1 when a node is at the center of its community and near 0 when it is on the periphery. A node with a high community centrality that is located in the center of its community has a high projected community centrality and hence is attached strongly to its community (and can be considered to be very influential within it). We normalize  $|\mathbf{x}_i|$  and  $y_i$  by the maximum value at each time step.

We investigate the persistence through time of nodes' communities by defining a community autocorrelation. For a node  $i$  with community  $c_i(t)$  at time  $t$ , the autocorrelation  $a_i^t(\tau)$  of its community after  $\tau$  time steps is defined by

$$a_i^t(\tau) = \frac{|c_i(t) \cap c_i(t + \tau)|}{|c_i(t) \cup c_i(t + \tau)|}. \tag{10}$$

This is a node-centric version of a quantity considered in Ref. 19. Importantly, this measure does not require one to determine which community at each time step represents the descendant of a community at the previous time step. Instead, the communities are identified from the perspective of individual nodes.

In Fig. 3(a), we show the mean normalized community centrality as a function of community size (averaging over all nodes belonging to communities of the same size). The community centrality increases with community size up to sizes of about ten members. For larger communities,  $|\mathbf{x}_i|$  re-

mains approximately constant. Nodes with high  $|\mathbf{x}_i|$  therefore tend to belong to large communities, so exchange rates with high community centrality tend to be closely linked with many other rates. Figure 3(b) shows the mean betweenness centrality versus the community alignment (averaging the community alignment over all nodes with the same betweenness). As the betweenness centrality of a node increases, its community alignment tends to decrease, implying that nodes that are important for the transfer of information are usually located away from the centers of communities. (The standard errors for nodes with very high betweenness are large because there are only a few nodes with such values.)

In Fig. 3(c), we show the mean community autocorrelation versus the projected community centrality. We calculate the mean autocorrelation by splitting the range of  $y$  into 15 equally-spaced bins and then averaging over all autocorrelations falling within these bins. (The observed relationships are robust for reasonable variations in the number of bins.) As one would expect, the community autocorrelation for the same projected community centrality is smaller for larger  $\tau$ . For all values of  $\tau$ , the mean community autocorrelation increases with  $y$ . This suggests that nodes that are strongly connected to their community are likely to persistently share that community membership with the same subset of nodes. In contrast, exchange rates with a low  $y$  experience regular changes in the set of rates with which they are clustered.

## VI. EFFECTS OF THE CREDIT CRISIS ON NETWORK STRUCTURE

Thus far, we have considered the community properties aggregated over all time steps. We now investigate the community dynamics in specific time intervals. In particular, we focus on the insights that such shorter-term dynamics can provide into the changes that occurred in the FX market during the recent credit crisis. Figure 4(a) shows a contour plot of the normalized distribution of link weights at each time step. The mean link strength remains constant through time because of the inclusion in the network of each exchange rate and its inverse, but [as one can see in Figs. 4(a) and 4(b)] there is a large variation in the standard deviation of the link strengths. Figure 4(b) shows that the scaled energy is closely related to the standard deviation of the link weights. This is expected because the standard deviation increases as a result of the strengthening of strong ties and the weakening of weak ties.

In Fig. 4(c), we show the normalized variation of information  $\hat{V}$  between the community configurations at consecutive time steps. Significant changes in configuration are indicated by large spikes in  $\hat{V}$ . The correlation coefficient between  $\hat{V}$  and the absolute change in  $Q_s$  between consecutive time steps is 0.47, and that between  $\hat{V}$  and the absolute change in  $\sigma(A_{ij})$  is 0.27. Changes in  $\sigma(A_{ij})$  are therefore not always a good indication that there has been a significant change in the community configuration of the network.

The seriousness of the credit crisis first became widely recognized when on 19 July 2007, Federal Reserve chairman Ben Bernanke warned in testimony to the U.S. Congress that the crisis in the U.S. subprime lending market could cost up

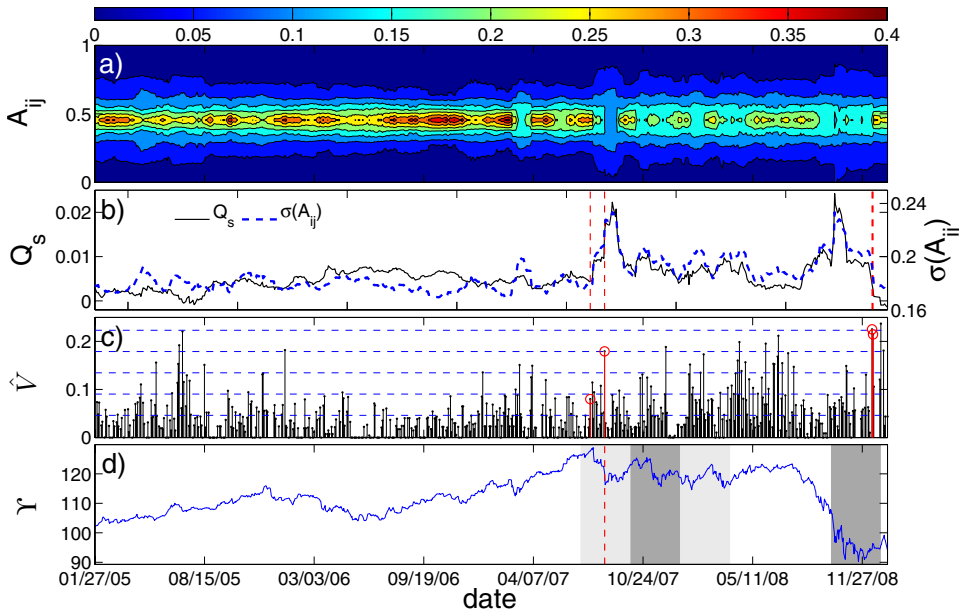


FIG. 4. (Color online) (a) Normalized distribution of link weights at each time step. (b) Scaled energy  $Q_s$  and standard deviation of the link weights. (c) Normalized variation of information between the community configurations at consecutive time steps. The horizontal lines show (from bottom to top) the mean of  $\hat{V}$  and 1, 2, 3, and 4 standard deviations above the mean. The vertical lines in (b) and (c) show the time steps when the following dates enter the rolling time window: 20 July 2007, 15 August 2007, and 16 December 2008 (two large reorganizations at consecutive time steps are marked by circles following this date). (d) Carry trade index  $Y$ . The vertical line shows 15 August 2007 and the shaded blocks (from left to right) indicate Q3 2007, Q4 2007, Q1 2008, and Q4 2008.

to \$100 billion. This announcement marks the start of a prolonged increase in  $Q_s$ , beginning as 20 July 2007 enters the rolling time window. There is no significant community reorganization at this time step.

The most important effect of the credit crisis on the FX market was its impact on the carry trade. The first significant community reorganization of the crisis occurred on 15 August 2007 as a result of massive “unwinding” of the carry trade. This change accompanied a large increase in  $Q_s$ . The carry trade consists of selling low interest rate “funding currencies” such as the JPY and CHF and investing in high interest rate “investment currencies” such as the AUD and NZD. It yields a profit if the interest rate differential between the funding and investment currencies is not offset by a commensurate depreciation of the investment currency.<sup>35</sup> The carry trade requires a strong appetite for risk, so unwinding of the trade tends to occur during periods in which there is a decrease in the available credit. A trader unwinds a carry trade position by selling his/her holdings in investment currencies and buying funding currencies.

One approach to quantifying carry trade activity is to consider the returns that could have been achieved using a carry trading strategy. The total return for a currency trading strategy can be split into two parts: a return resulting from the price changes and a return resulting from the interest rate differentials between the currencies. We can estimate the return that could have been achieved for a particular carry trade strategy using historical price and interest rate time series. We consider a common strategy in which one buys equal proportions of the three major currencies with the highest interest rates and sells equal proportions of the three currencies with the lowest interest rates. In Fig. 4(d), we show the cumulative return index  $Y$  for this trading strategy. Large negative returns result in large decreases in  $Y$ , which are therefore likely to indicate significant unwinding of the carry trade.

In Fig. 5(a), we show the observed communities before and after 15 August 2007. Figure 4(d) shows that leading up

to 15 August 2007, there has been some unwinding of the carry trade, so the initial configuration includes a community containing exchange rates of the forms AUD/YYY, NZD/YYY, and XXX/JPY (which all involve one of the key carry trade currencies). During 15 August 2007–17 August 2007, there is a sharp increase in carry trade unwinding, which results in this community increasing in size as it incorporates other XXX/JPY rates as well as some XXX/CHF and XXX/USD rates. The presence of many exchange rates involving one of the key carry trade currencies in a single community clearly demonstrates the significance of the trade over this period. Some of the exchange rates included in the commu-

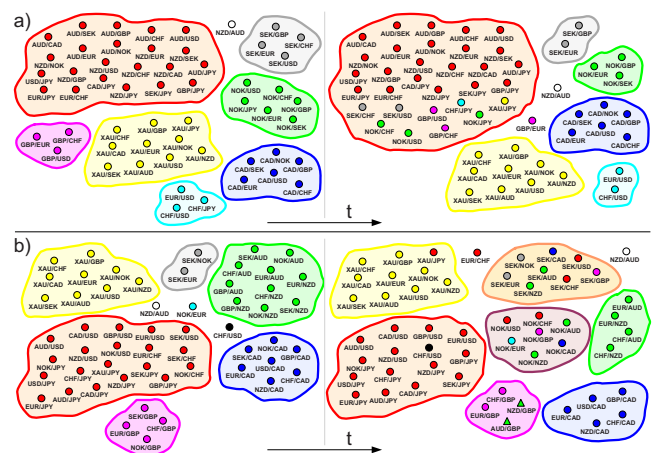


FIG. 5. (Color online) Schematic representation of the change in the community configuration in one half of the FX market network following: (a) 15 August 2007, when there was significant unwinding of the carry trade and (b) 16 December 2008, when the U.S. Federal Reserve cut the funding interest rate from 1% to 0%–0.25%. The node colors after the community reorganization correspond to the community assignments before the change. If the parent community of a post-reorganization community is obvious, it is drawn in the same color as its parent. The nodes represented as triangles resided in the opposite half of the network before the community reorganization.

nity are also somewhat surprising and provide insights into the range of currencies used in the carry trade.

In Fig. 5(b), we show an example of a significant community reorganization that accompanied a decrease in  $Q_s$ . This reorganization occurred following 16 December 2008 when the U.S. Federal Reserve cut the federal funds interest rate from 1% to 0%–0.25%—the lowest rate ever recorded.

We have considered three examples in which large community reorganizations are clearly attributable to particular market events, demonstrating the effectiveness of our methods. However, there are several other large community changes that are not easily associated with specific events. Investigating these community changes might provide insights into market changes that are not otherwise obvious.

Finally, we consider the changes in the relationships between specific nodes and their communities during the crisis period. We begin by defining within-community  $z$ -scores, which directly compare the relative importances of different nodes to their community.<sup>16</sup> We describe the roles of individual nodes at each time step using the within-community projected community centrality  $z$ -score  $z^y$  and within-community betweenness centrality  $z$ -score  $z^b$ . If a node  $i$  belongs to community  $c_i$  and has projected community centrality  $y_i$ , then  $z_i^y = (y_i - \bar{y}_{c_i}) / \sigma_{c_i}^y$ , where  $\bar{y}_{c_i}$  is the average of  $y$  over all nodes in  $c_i$  and  $\sigma_{c_i}^y$  is the standard deviation of  $y$  in  $c_i$ . The quantity  $z_i^y$  measures how strongly node  $i$  is connected to its community. Similarly, if the same node has betweenness centrality  $b_i$ , then  $z_i^b = (b_i - \bar{b}_{c_i}) / \sigma_{c_i}^b$ , where  $\bar{b}_{c_i}$  is the average of  $b$  over all nodes in  $c_i$  and  $\sigma_{c_i}^b$  is the standard deviation of  $b$  in  $c_i$ . The quantity  $z_i^b$  indicates the importance of node  $i$  to the spread of information compared to other nodes in the same community. The positions of nodes in the  $(z^b, z^y)$  plane thereby illuminate the roles of the associated exchange rates in the FX market.

We investigate the time dynamics of exchange rate roles by examining quarterly changes in their positions in the  $(z^b, z^y)$  plane. The quarterly roles are determined by averaging  $z^y$  and  $z^b$  over all time steps in each quarter. Changes in a node's position in the  $(z^b, z^y)$  plane reflect changes in its community assignment as well as changes in  $b$  and  $y$ .

We show four example role evolutions in Fig. 6. The USD/XAU rate provides an interesting example due to the persistence of its community. From 2005–2008, the USD/XAU node has shifted from being an important information carrier within the XAU community to being more central to this community. This period of higher influence coincides closely with the period of financial turmoil. The CHF is widely regarded as a “safe haven” currency,<sup>36</sup> so one might expect USD/CHF to behave in a manner similar to USD/XAU. However, the CHF is also a key carry trade currency. The use of the CHF as both a safe haven and a carry trade currency means that USD/CHF does not move in the same direction as USD/XAU in the  $(z^b, z^y)$  plane. Instead the USD/CHF exchange rate is an important information carrier during the credit crisis. The AUD/JPY and NZD/JPY exchange rates change from being important for information transfer to being influential within their communities. These exchange rates are the ones that are most widely used for the carry

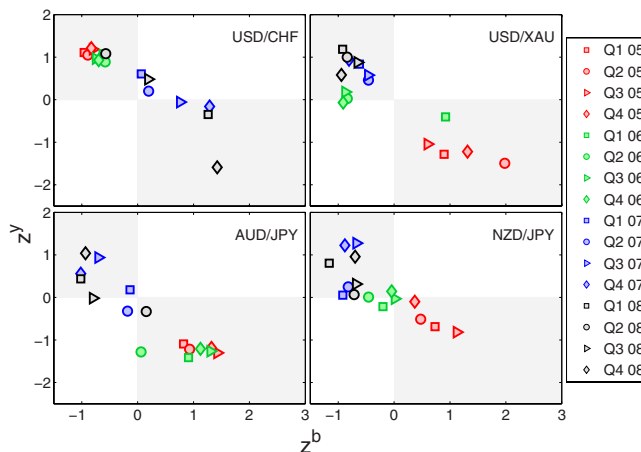


FIG. 6. (Color online) Quarterly node role evolutions in the  $(z^b, z^y)$  plane.

trade, so their increased importance is unsurprising. In fact, both rates were very influential during the third and fourth quarters in 2007, the first quarter in 2008, and the fourth quarter of 2008, when there was significant carry trade activity [see Fig. 4(d)].

### VII. CONCLUSIONS

In summary, we have analyzed evolving community structure in time-dependent networks to provide new insights into the clustering dynamics of multichannel/multivariate time series. We focused, in particular, on a node-centric community analysis that allows one to follow the time dynamics of the functional roles of individual nodes within networks. As an illustration of our approach, we examined the FX market network during a period that included the 2007–2008 credit crisis. We demonstrated that the FX market has a robust community structure over a range of resolutions and showed that there is a relationship between an exchange rate's functional role and its position within its community. Our analysis of the community dynamics successfully uncovered important structural changes that occurred in the FX market during the credit crisis and identified exchange rates that experienced significant changes in market role. Our methodology should be similarly insightful for other multivariate data sets.

### ACKNOWLEDGMENTS

We thank S. D. Howison, A. C. F. Lewis, M. A. Little, P. J. Mucha, J.-P. Onnela, S. Reid, and S. J. Roberts and his group for suggestions and code. We acknowledge HSBC bank for providing the data. N.S.J. acknowledges the BB-SRC and EPSRC-GB. M.A.P. acknowledges a research award (Grant No. 220020177) from the James S. McDonnell Foundation.

<sup>1</sup>M. E. J. Newman, *SIAM Rev.* **45**, 167 (2003).  
<sup>2</sup>R. Albert and A.-L. Barabási, *Rev. Mod. Phys.* **74**, 47 (2002).  
<sup>3</sup>G. Caldarelli, *Scale-Free Networks* (Oxford University Press, Oxford, 2007).  
<sup>4</sup>M. E. J. Newman, *Eur. Phys. J. B* **38**, 321 (2004).  
<sup>5</sup>M. E. J. Newman and M. Girvan, *Phys. Rev. E* **69**, 026113 (2004).

- <sup>6</sup>L. Danon, A. Díaz-Guilera, J. Duch, and A. Arenas, *J. Stat. Mech.* **2005**, P09008 (2005).
- <sup>7</sup>M. E. J. Newman, *Phys. Rev. E* **74**, 036104 (2006).
- <sup>8</sup>J. Reichardt and S. Bornholdt, *Phys. Rev. E* **74**, 016110 (2006).
- <sup>9</sup>A. Arenas, A. Fernández, and A. Gómez, *New J. Phys.* **10**, 053039 (2008).
- <sup>10</sup>S. Fortunato and M. Barthelemy, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 36 (2007).
- <sup>11</sup>S. Fortunato and C. Castellano, *Encyclopedia of Complexity and Systems Science* (Springer, Berlin, 2009).
- <sup>12</sup>S. Fortunato, "Community detection in graphs," e-print arXiv:0906.0612.
- <sup>13</sup>M. A. Porter, J.-P. Onnela, and P. J. Mucha, "Communities in networks," *Not. Am. Math. Soc.* (in press); e-print arXiv:0902.3788.
- <sup>14</sup>M. A. Porter, P. J. Mucha, M. E. J. Newman, and C. M. Warmbrand, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 7057 (2005).
- <sup>15</sup>A. L. Traud, E. D. Kelsic, P. J. Mucha, and M. A. Porter, "Community structure in online collegiate social networks," e-print arXiv:0809.0960.
- <sup>16</sup>R. Guimerà and L. A. N. Amaral, *Nature (London)* **433**, 895 (2005).
- <sup>17</sup>B. Adamcsek, G. Palla, I. J. Farkas, I. Derényi, and T. Vicsek, *Bioinformatics* **22**, 1021 (2006).
- <sup>18</sup>A. C. F. Lewis, N. S. Jones, M. A. Porter, and C. M. Deane, "The function of communities in protein interaction networks," e-print arXiv:0904.0989.
- <sup>19</sup>G. Palla, A.-L. Barabási, and T. Vicsek, *Nature (London)* **446**, 664 (2007).
- <sup>20</sup>T. W. Liao, *Pattern Recogn.* **38**, 1857 (2005).
- <sup>21</sup>J. Hopcroft, O. Khan, B. Kulli, and B. Selman, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 5249 (2004).
- <sup>22</sup>T. Falkowski, J. Bartelheimer, and M. Spiliopoulou, *Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence* (IEEE Computer Society, Washington, DC, 2006).
- <sup>23</sup>M. McDonald, O. Suleman, S. Williams, S. Howison, and N. F. Johnson, *Phys. Rev. E* **72**, 046106 (2005).
- <sup>24</sup>T. Heimo, J. S. Kumpula, K. Kaski, and J. Saramäki, *J. Stat. Mech.* **2008**, P08007 (2008).
- <sup>25</sup>If we include self-edges, the strength (weighted degree) of each node will increase by 1. This, in turn, leads to a constant increase in the expected edge weight in the null model. For a network with self-edges, the expected edge weight is given by  $p_{ij}^s = n/[2(n+2)]$ , a shift by a constant value of  $p_{ij}^s - p_{ij} = 2/[n(n+2)] \approx 1.62 \times 10^{-4}$  relative to the network in which self-edges are excluded. Self-edges always occur within a community, so they will always contribute to the summation in Eq. (2) irrespective of exactly how the nodes are partitioned into communities. This implies that self-edges play no role when selecting the community partition that minimizes the interaction energy at a particular resolution.
- <sup>26</sup>In order to find equivalent communities in the network in which self-edges are included, it is necessary to decrease the resolution parameter to compensate for the increase in the constant expected edge weight in the null model. If we identify communities in the network in which self-edges are excluded using the resolution parameter  $\gamma$ , then we find identical communities in the corresponding network with self-edges using a resolution parameter  $\gamma^s = \gamma p_{ij} / p_{ij}^s = \gamma(n+2)(n-2)/n^2$ . For example, if we identify communities in the network without self-edges using a resolution of  $\gamma = 1.4500$ , then we identify equivalent communities in the network with self-edges using a resolution parameter of  $\gamma_s \approx 1.4495$ .
- <sup>27</sup>J.-P. Onnela, A. Chakraborti, K. Kaski, J. Kertész, and A. Kanto, *Phys. Rev. E* **68**, 056110 (2003).
- <sup>28</sup>R. N. Mantegna and H. E. Stanley, *An Introduction to Econophysics* (Cambridge University Press, Cambridge, 2000).
- <sup>29</sup>I. Farkas, D. Ábel, G. Palla, and T. Vicsek, *New J. Phys.* **9**, 180 (2007).
- <sup>30</sup>C. R. Shalizi, M. F. Camperi, and K. L. Klinkner, in *Statistical Network Analysis: Models, Issues, and New Directions* (Springer-Verlag, New York, 2007).
- <sup>31</sup>V. D. Blondel, J. Guillaume, R. Lambiotte, and E. Lefebvre, *J. Stat. Mech.* **2008**, P10008 (2008).
- <sup>32</sup>M. Meilä, *J. Multivariate Anal.* **98**, 873 (2007).
- <sup>33</sup>L. C. Freeman, *Sociometry* **40**, 35 (1977).
- <sup>34</sup>D. A. Berry and B. W. Lindgren, *Statistics: Theory and Methods* (Brooks-Cole, Belmont, MA, 1990).
- <sup>35</sup>M. K. Brunnermeier, S. Nagel, and L. H. Pedersen, NBER Working Paper Series 14473, available at SSRN: <http://ssrn.com/abstract=1297722> (2008).
- <sup>36</sup>A. Rinaldo and P. Söderlind, available at SSRN: <http://ssrn.com/abstract=1097593> (2007).