Quantitative Finance
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rquf20

Dynamical clustering of exchange rates

Daniel J. Fenn [a b], Mason A. Porter [c b], Peter J. Mucha [d e], Mark McDonald [f], Stacy Williams [f], Neil F. Johnson [g b] & Nick S. Jones [h i b j]

[a] Mathematical and Computational Finance Group, Mathematical Institute, University of Oxford, Oxford OX1 3LB, UK

[b] CABDyN Complexity Centre, University of Oxford, Oxford OX1 1HP, UK

[c] Oxford Centre for Industrial and Applied Mathematics, Mathematical Institute, University of Oxford, Oxford OX1 3LB, UK

[d] Carolina Center for Interdisciplinary Applied Mathematics, Department of Mathematics, University of North Carolina, Chapel Hill, NC 27599, USA

[e] Institute for Advanced Materials, Nanoscience and Technology, University of North Carolina, Chapel Hill, NC 27599, USA

[f] FX Research and Trading Group, HSBC Bank, 8 Canada Square, London E14 5HQ, UK

[g] Physics Department, University of Miami, Florida, Coral Gables 33146, USA

[h] Physics Department, Clarendon Laboratory, University of Oxford, Oxford OX1 3PU, UK

[i] Oxford Centre for Integrative Systems Biology, Oxford OX1 3QU, UK

[j] Department of Mathematics, Imperial College, London SW7 2AZ, UK

Version of record first published: 10 Aug 2012.

To cite this article: Daniel J. Fenn, Mason A. Porter, Peter J. Mucha, Mark McDonald, Stacy Williams, Neil F. Johnson & Nick S. Jones (2012): Dynamical clustering of exchange rates, Quantitative Finance, 12:10, 1493-1520

To link to this article: http://dx.doi.org/10.1080/14697688.2012.668288

Quantitative Finance, Vol. 12, No. 10, October 2012, 1493–1520

Dynamical clustering of exchange rates

DANIEL J. FENN*†‡, MASON A. PORTER§‡, PETER J. MUCHA¶‖, MARK MCDONALD⊥, STACY WILLIAMS⊥, NEIL F. JOHNSON††‡ and NICK S. JONES‡‡§§‡¶¶

†Mathematical and Computational Finance Group, Mathematical Institute, University of Oxford, Oxford OX1 3LB, UK
‡CABDyN Complexity Centre, University of Oxford, Oxford OX1 1HP, UK
§Oxford Centre for Industrial and Applied Mathematics, Mathematical Institute, University of Oxford, Oxford OX1 3LB, UK
¶Carolina Center for Interdisciplinary Applied Mathematics, Department of Mathematics, University of North Carolina, Chapel Hill, NC 27599, USA
‖Institute for Advanced Materials, Nanoscience and Technology, University of North Carolina, Chapel Hill, NC 27599, USA
⊥FX Research and Trading Group, HSBC Bank, 8 Canada Square, London E14 5HQ, UK
††Physics Department, University of Miami, Florida, Coral Gables 33146, USA
‡‡Physics Department, Clarendon Laboratory, University of Oxford, Oxford OX1 3PU, UK
§§Oxford Centre for Integrative Systems Biology, Oxford OX1 3QU, UK
¶¶Department of Mathematics, Imperial College, London SW7 2AZ, UK

(*Received 12 April 2010; in final form 15 February 2012*)

We use techniques from network science to study correlations in the foreign exchange (FX) market during the period 1991–2008. We consider an FX market network in which each node represents an exchange rate and each weighted edge represents a time-dependent correlation between the rates. To provide insights into the clustering of the exchange-rate time series, we investigate dynamic communities in the network. We show that there is a relationship between an exchange rate's functional role within the market and its position within its community and use a node-centric community analysis to track the temporal dynamics of such roles. This reveals which exchange rates dominate the market at particular times and also identifies exchange rates that experienced significant changes in market role. We also use the community dynamics to uncover major structural changes that occurred in the FX market. Our techniques are general and will be similarly useful for investigating correlations in other markets.

Keywords: Foreign exchange market; Networks; Community detection

JEL Classification: C02, F31

1. Introduction

Complex systems are composed of many interacting elements and can exhibit numerous forms of 'emergent' collective dynamics without the need for any external organizing principle (Boccara 2003). Such dynamics typically cannot be explained by studying the constituent parts in isolation, so a complex system must be analysed as a whole. Networks provide a tractable framework for the quantitative analysis of many complex systems by distilling them to their key elements (Albert and Barabási 2002, Newman 2003, Amaral and Ottino 2004, Caldarelli 2007). In such a representation, the elements of a system are represented as the network's nodes and the important interactions between them are represented as links that connect the nodes. (In this paper, we use the terms 'links' and 'edges' interchangeably.)

Financial markets exhibit many of the key properties that characterize complex systems: they are composed of many heterogeneous components that interact with each other and their environment nonlinearly in the presence of feedback (Mantegna and Stanley 2000, Amaral and Ottino 2004). An investigation of a financial market can

*Corresponding author. Email: dan.fenn@hsbcib.com

Quantitative Finance
ISSN 1469–7688 print/ISSN 1469–7696 online © 2012 Taylor & Francis
http://www.tandfonline.com
http://dx.doi.org/10.1080/14697688.2012.668288

be formulated as a network problem. Indeed, a wide range of financial assets have been investigated using network techniques, including equities (Mantegna 1999, Mantegna and Stanley 2000, Onnela *et al.* 2003), currencies (McDonald *et al.* 2005, 2008), commodities (Sieczka and Holyst 2009), bonds (Bernaschi *et al.* 2002), and interest rates (Matteo *et al.* 2004). Network analyses have the potential to provide new insights into financial data and the structure of markets, which may in turn lead to the development of better market models.

In the most common network description of a market, each node represents an asset, and each weighted link is a function (the same function for all links) of the pairwise temporal correlations between the two assets that it connects (Mantegna and Stanley 2000). In a typical financial network containing n assets, one can calculate a correlation coefficient between each pair of assets, so the network contains $n(n-1)/2$ links. Thorough, simultaneous investigation of the interactions is therefore difficult for even moderate values of n, so attaining an understanding of the market system necessitates some form of simplification.

The most prevalent method for reducing the complexity of a financial network is to construct a minimum spanning tree (MST) (Mantegna 1999, Mantegna and Stanley 2000, Bouchaud and Potters 2003, Onnela *et al.* 2003, 2004). An MST is generated using a hierarchical clustering algorithm (Duda *et al.* 2001), and it reduces a network to $n-1$ of its most important microscopic interactions. This approach has resulted in many useful financial applications, including the construction of a visualization tool for portfolio optimization (Onnela *et al.* 2003) and a means for identifying the effect of news and major events on market structure (McDonald *et al.* 2008). Nevertheless, an MST approach has several limitations, which we discuss in section 6.

An alternative simplification method is to coarse-grain a network and consider it at various mesoscopic scales. The properties of a market can then be understood by considering the dynamics of small groups of similar nodes. A widely-studied form of mesoscopic structure, known as a 'community' (Newman 2004a, Newman and Girvan 2004, Danon *et al.* 2005, Newman 2006a, Reichardt and Bornholdt 2006, Fortunato and Barthelemy 2007, Arenas *et al.* 2008, Porter *et al.* 2009, Fortunato 2010) is constructed from subsets of nodes that are more strongly connected to each other than they are to the rest of a network. Communities are of considerable interest to network scientists because they can correspond to behavioural or functional units (Guimerà and Amaral 2005, Porter *et al.* 2005, Adamcsek *et al.* 2006, Traud *et al.* 2011), so their identification can lead to a better understanding of dynamical processes (such as the spread of opinions and diseases) that operate on networks (Danon *et al.* 2005, Porter *et al.* 2009, Fortunato 2010). From a financial perspective, communities correspond to groups of closely-related assets, so community detection has the potential to suggest possible formulations for coarse-grained stochastic models of markets.

During the last decade, there has been an explosion of papers on networks with static connections between nodes, and research on dynamical systems on such networks has now also become ubiquitous (Newman 2003, Caldarelli 2007, Barrat *et al.* 2008). However, there has been much less research on networks that are themselves time-dependent (Onnela *et al.* 2007, Palla *et al.* 2007, Mucha *et al.* 2010), and a characterization of such networks is essential for a full understanding of dynamical processes on networks. One of the main reasons for the limited analysis of time-dependent networks is the difficulty of acquiring time-dependent data. Fortunately, financial markets are one of the most data-rich complex systems, providing a valuable source of accurate, high-frequency, time-series data. Financial data are therefore an important resource for developing tools and theories for describing time-dependent networks.

In the present work, we investigate community dynamics in a time-dependent foreign exchange (FX) market network. The FX network possesses a fixed number of nodes and evolving link weights that are determined by time-varying pairwise correlations between time series associated with each node. Therefore, in contrast to some other studies of financial networks, we analyse a fully-connected network and do not remove links below some threshold (Farkas *et al.* 2007). Community detection in networks of this kind is closely related to the problem of clustering multivariate time series (Liao 2005). We also track communities from the perspective of individual nodes, which removes the undesirable requirement of determining which community at each time step represents the descendant of a community at the previous time step. Previous dynamic community studies have attempted to track entire communities (Hopcroft *et al.* 2004, Palla *et al.* 2007) but (as discussed in section 7.2) some of these approaches can lead to equivocal mappings following community splits and mergers.

We demonstrate that exchange-rate community dynamics provides insight into correlation structures within the FX market and uncovers important exchange-rate interactions. We also show that large community reorganizations often accompany significant market events and that the details of such community adjustments can reveal trading behaviour that leads to these changes. We find that there is a relationship between an exchange rate's functional role within the market and its position within its community, and we identify exchange rates that experience significant changes in market role. Although we focus on the FX market, the techniques that we present are general and will be similarly insightful for other asset classes.

This paper builds on the results described in Fenn *et al.* (2009), which focused on the period 2005–2008 when the recent credit and liquidity crisis began. In the present work, we extend our earlier investigation to two additional time periods: 1991–1998 (before the introduction of the euro) and 1999–2003 (following the introduction of the euro). For each time interval, we identify communities of exchange rates, and we then compare the structure of the communities across the different periods. In addition,

we provide further examples of the effects of major market events on the dynamics of individual communities by considering community changes during the 1994 Mexican peso crisis and the 1997–1998 Asian currency crisis. The present work also extends Fenn *et al.* (2009) by comparing the results obtained using community detection with those produced by traditional clustering techniques. In particular, we demonstrate that the communities that we uncover are consistent with the clusters identified using linkage-clustering algorithms (Duda *et al.* 2001). Finally, in appendix A, we provide a detailed analysis of the sensitivity of our results to the choice of computational heuristic used to identify communities. We obtain the same aggregate conclusions, although there are differences in the communities identified using different heuristics.

The remainder of this paper is organized as follows. In section 2, we discuss the nature of the FX data and use it to derive a time-dependent network. We detect communities in the network in section 3 and discuss robust communities in section 4. We examine the properties of the communities in section 5, compare the detected communities to minimum spanning trees in section 6, and derive the roles of exchange rates within communities in section 7. We relate the major changes in community structure over time to significant changes in the FX market in section 8 and investigate the changes in the community roles of exchange rates in section 9. In section 10, we offer some conclusions. In appendix A, we discuss the effects of using different heuristics to identify optimal partitions of the FX networks into communities.

2. Data

The FX networks that we construct have $n = 110$ nodes, each of which represents an exchange rate of the form XXX/YYY (with XXX $\neq$ YYY), where XXX, YYY $\in$ {AUD, CAD, CHF, GBP, DEM, JPY, NOK, NZD, SEK, USD, XAU}† and we note that DEM $\rightarrow$ EUR after 1998. An exchange rate XXX/YYY indicates the amount of currency YYY that can be received in exchange for one unit of XXX.‡ Other authors have recently studied the FX market by constructing networks in which all nodes represent exchange rates with the same base currency, implying that each node can then be considered to represent a single currency (Górski *et al.* 2008). Exchange-rate networks formed with reference to a single base currency are somewhat akin to ego-centred networks studied in the social networks literature (Wasserman and Faust 1994). Ego-centred networks include links between nodes that all have ties to an *ego*, which is the focal node of the network. However, this approach has two major problems for FX networks. First, it neglects a large number of exchange rates that can be formed from the set of currencies studied and consequently also ignores the interactions between these rates. Second, the network properties depend strongly on the choice of base currency, and this currency is, in effect, excluded from the analysis. We therefore construct networks that include all exchange rates that can be formed from the studied set of currencies.

The return of an exchange rate with price $p_i(t)$ at discrete time t is defined by

$$R_i(t) = \ln \frac{p_i(t)}{p_i(t-1)}. \quad (1)$$

We take the price $p_i(t)$ as the mid-price of the bid and ask prices:

$$p_i(t) = \frac{p_i^{\mathrm{bid}}(t) + p_i^{\mathrm{ask}}(t)}{2}. \quad (2)$$

We use the last posted price within an hour to represent the price for the following hour. To calculate a return at time t, one needs to know the price at both t and $t-1$. To minimize the possibility of a price not being posted in a given hour, we focus on the FX market's most liquid period: 07:00–18:00 UK time. Nevertheless, there are still hours for which we do not have price data. (This usually occurs as a result of problems with the data feed.) One can calculate a return for hours with missing price data by assuming the last posted price or interpolating between prices at the previous and next time step (Dacorogna *et al.* 2001). However, to ensure that all time steps included in the study are ones at which a trade can actually be made, we take the stricter approach of omitting all returns for which one of the prices is not known. In order to ensure that the time series of exchange rates are directly comparable, we consequently remove a return from all exchange rates if it is missing from any rate.

For the period 1991–2003, we derive each exchange rate XXX/YYY with XXX, YYY $\neq$ USD from two USD rates. For example, we find the CAD/CHF price at each time step by dividing the USD/CHF price by the USD/CAD price. For the period 2005–2008, we derive each exchange rate not included in the set {AUD/USD, EUR/NOK, EUR/SEK, EUR/USD, GBP/USD, NZD/USD, USD/CAD, USD/CHF, USD/JPY, USD/XAU} from pairs of exchange rates in this set. For example, we find the USD/NOK price at each time step by dividing the EUR/NOK price by the EUR/USD price. Although this approach appears somewhat artificial, it matches the way in which many exchange rates are calculated in the actual FX market. For example, a bank customer wishing to

†These symbols represent: 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. We include gold in the study because it has many similarities with a currency (McDonald *et al.* 2005).

‡For each exchange rate, market participants can quote both bid and ask prices. Bid/ask prices give the different prices at which one can buy/sell currency, and the ask price tends to be larger than the bid price. For example, suppose that the exchange rate between EUR and USD is quoted as 1.4085/1.4086. A trader then looking to convert USD into EUR has to pay 1.4086 USD for each EUR, whereas a trader looking to convert EUR to USD receives only 1.4085 USD per EUR.

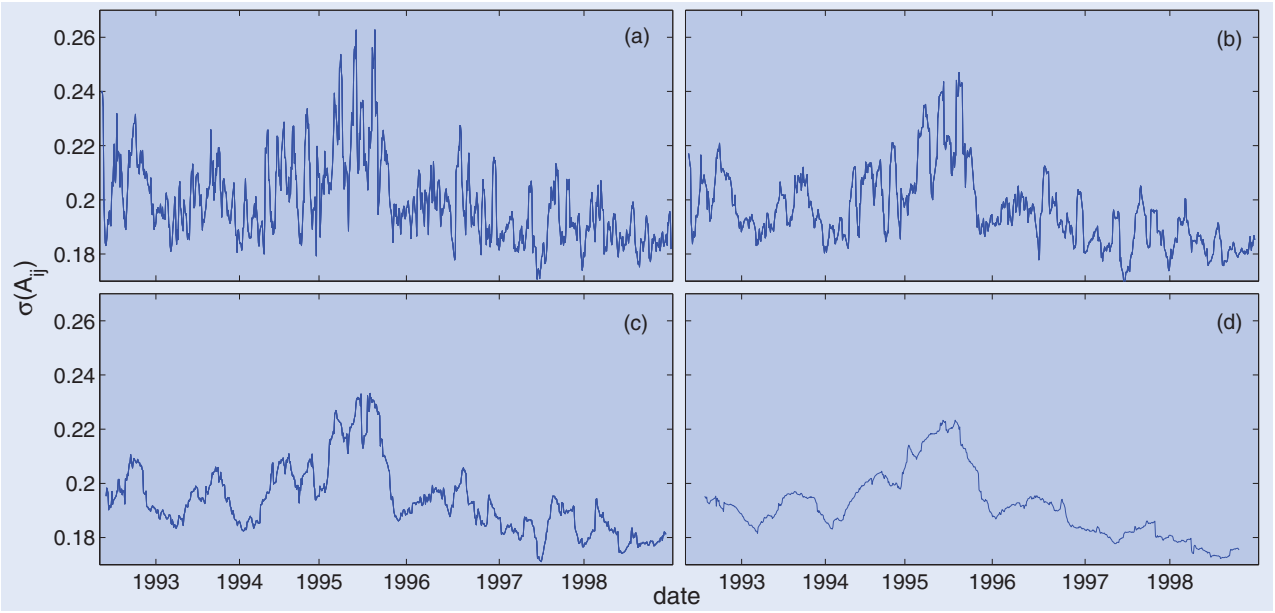

Figure 1. Standard deviation of the edge weights A_{ij} as a function of time for the period 1991–1998. For all panels, $\Delta t = 20$ (approximately two days). We show this calculation for (a) $T = 100$ hours, (b) $T = 200$ hours, (c) $T = 400$ hours, and (d) $T = 1200$ hours (corresponding to approximately 0.5, 1, 2, and 6 months, respectively).

convert CAD to NZD (or vice versa) will need to be quoted the CAD/NZD prices. Because this is not a standard conversion, the bank will not be able to quote a direct market price but will instead calculate a price using the more widely traded USD/NZD and USD/CAD exchange rates. Calculating the exchange rates in this way implies that there is some intrinsic structure inherent in the FX market. However, as shown in McDonald *et al.* (2005) and demonstrated further in sections 5.2 and 5.3 of this paper, this 'triangle effect' does not dominate the results.

We determine the weights of the edges connecting pairs of nodes in the FX networks using a function of the linear correlation coefficient ρ between the return time series for the corresponding exchange rates. The correlation between the exchange-rate returns R_i and R_j in a time window of T returns is given by

$$\rho(i,j) = \frac{\langle R_i R_j \rangle - \langle R_i \rangle \langle R_j \rangle}{\sigma_i \sigma_j}, \tag{3}$$

where $\langle \cdot \rangle$ indicates a time average over T returns and σ_i is the standard deviation of R_i over T. We use the linear coefficient $\rho(i, j)$ to measure the correlation between pairs of exchange rates because of its simplicity, but one could use alternative measures that are capable of detecting more general dependencies (Schelter *et al.* 2006). Our methods can be applied using any choice for $\rho(i, j)$. The weighted adjacency matrix $\mathbf{A}$ representing the network has components

$$A_{ij} = \frac{1}{2}[\rho(i,j) + 1] - \delta_{ij}, \tag{4}$$

where the Kronecker delta δ_{ij} removes self-edges. The matrix elements $A_{ij} \in [0, 1]$ quantify the similarity of each pair of exchange rates i and j. For example, two exchange rates i and j whose return time series are perfectly correlated will be connected by a link of weight 1.

We exclude self-edges in order to deal with simple graphs. This approach was also taken in a previous study of an equities network derived from a correlation matrix (Heimo *et al.* 2008). If we include self-edges, the node compositions of the identified communities are identical if one makes a small parameter change in the community-detection algorithm. We discuss the community-detection algorithm and the effect of including self-edges in sections 3 and 5.

We create a longitudinal sequence of networks by displacing time windows by $\Delta t = 20$ hours (approximately two trading days) and fix $T = 200$ hours (approximately one month of data). This choice of T, motivated in part by the example data in figure 1, represents a trade-off between over-smoothing for long time windows and overly-noisy correlation coefficients for small T (Onnela *et al.* 2003). Figure 2 demonstrates that the choice of Δt has a similar, but less pronounced, effect on the standard deviation of the edge weights, and we again select a compromise value. The time windows we use to construct the networks overlap, so the single-time networks are not independent but rather form an evolving sequence through time.

3. Community detection

Communities consist of cohesive groups of nodes that are more strongly connected to each other than they are to the rest of a network. They can represent functionally-important subnetworks (Girvan and Newman 2002, Danon *et al.* 2005, Guimerà and

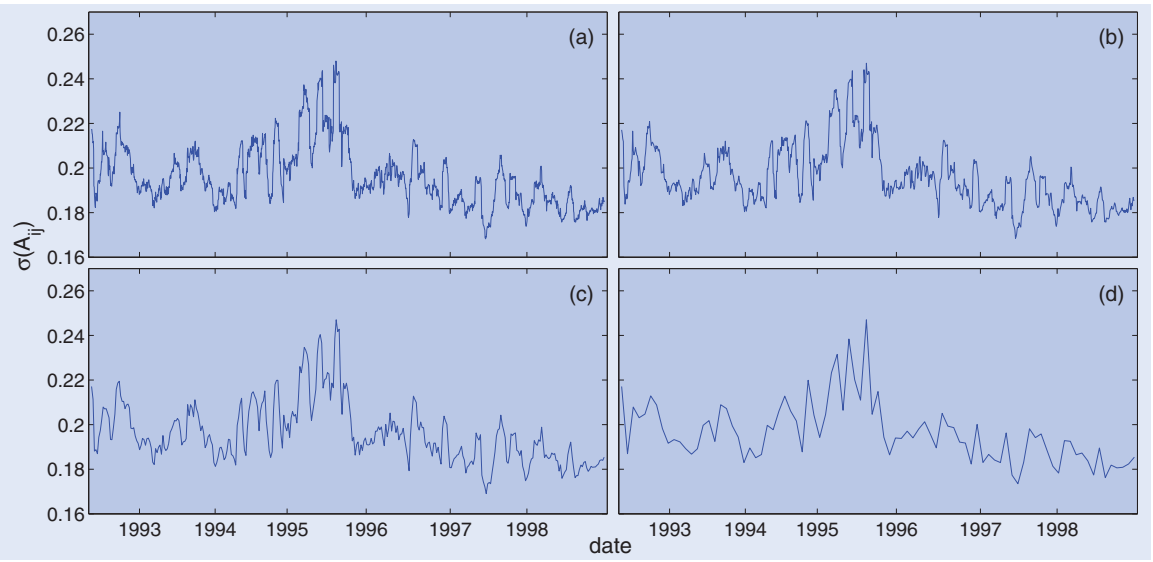

Figure 2. Standard deviation of the edge weights A_{ij} as a function of time for the period 1991–1998. For all panels, $T = 200$ hours. We show this calculation for (a) $\Delta t = 10$, (b) $\Delta t = 20$, (c) $\Delta t = 50$, and (d) $\Delta t = 200$ (corresponding to approximately 1 day, 2 days, 5 days, and 2 weeks, respectively).

Amaral 2005, Porter *et al.* 2005, Adamcsek *et al.* 2006, Porter *et al.* 2009, Fortunato 2010, Traud *et al.* 2011). Most prior studies of financial networks have found groups of closely-related assets using traditional hierarchical clustering techniques (Mantegna and Stanley 2000, Onnela *et al.* 2003, McDonald *et al.* 2005) or by thresholding to create a binary network (Farkas *et al.* 2007). In this paper, we identify communities in high-frequency, time-evolving, weighted networks using a technique based on the maximization of a quality function known as *modularity* (Newman and Girvan 2004). To our knowledge, other papers with similar approaches have not examined longitudinal networks or have considered networks of equities rather than exchange rates (Heimo *et al.* 2008). Dynamic communities have been investigated in biophysical data using methods based on modularity maximization (Shalizi *et al.* 2007). However, Shalizi *et al.* 2007 were concerned with the dynamics of functional communities that arise from coordinated behaviours taking place on a network rather than the community dynamics of the underlying network. (In section 7.2, we briefly discuss additional investigations of community dynamics in non-financial data using other community-detection techniques.)

The identification of communities using graph modularity is based on the idea that random networks are not expected to demonstrate community structure beyond fluctuations. Modularity therefore identifies communities by finding subsets of nodes that are more strongly connected to each other than one would expect for some null model. Let C be a partition of the n nodes in $\mathbf{A}$ into disjoint communities. The modularity Q of the partition C is given by

$$Q(C) = \frac{1}{2m}\sum_{ij}(A_{ij} - P_{ij})\delta(c_i, c_j), \tag{5}$$

where c_i is the community containing node i and P_{ij} denotes the expected weight of the link with which nodes i and j are connected in a null model. The quantity m represents the total edge weight in the network and is given by $m = \frac{1}{2}\sum_i k_i$, where $k_i = \sum_j A_{ij}$ is the strength (weighted degree) of node i. We identify communities by finding the partition C that maximizes Q. The most popular choice of null model is the Newman–Girvan (NG) model (Newman and Girvan 2004)

$$P_{ij} = \frac{k_i k_j}{2m}, \tag{6}$$

which preserves the expected strength distribution of the network and is closely related to the configuration model (Bollobas 2001). An alternative null model is a uniform model in which a fixed mean edge weight occurs between each pair of nodes (Porter *et al.* 2009).

We construct FX networks by calculating a correlation coefficient between every pair of exchange rates. This results in a weighted, fully-connected network. We include each exchange rate XXX/YYY and its inverse rate YYY/XXX in the network, because one cannot infer a priori whether a rate XXX/YYY will form a community with a rate WWW/ZZZ or its inverse ZZZ/WWW. However, the return of an exchange rate XXX/YYY is related to the return of its inverse YYY/XXX by

$$R_{\frac{XXX}{YYY}} = -R_{\frac{YYY}{XXX}}.$$

This implies that the correlation coefficients between these rates and a rate WWW/ZZZ are related by

$$\rho\left(\frac{XXX}{YYY}, \frac{WWW}{ZZZ}\right) = -\rho\left(\frac{YYY}{XXX}, \frac{WWW}{ZZZ}\right).$$

Consequently, every node has the same strength

$$k_i = \sum_j A_{ij} = \frac{1}{2}(n-2), \tag{7}$$

so the probability of connection in the NG null model $P_{ij} = k_i k_j / 2m$ is also constant and is given by

$$P_{ij} = \frac{n-2}{2n}. \tag{8}$$

In the case of our FX network, the NG model and the uniform null model are thus equivalent. However, the methods we present are general and can be applied to networks with non-uniform strength distributions. Additionally, every community has an equivalent inverse community. For example, if there is a community consisting of the three exchange rates XXX/YYY, XXX/WWW, and ZZZ/WWW in one half of the network, there must be an equivalent community formed of YYY/XXX, WWW/XXX, and WWW/ZZZ in the other half. The existence of an equivalent inverse community for each community implies that the FX network is composed of two equivalent halves at each time step. However, the exchange rates residing in each half change in time as the correlations evolve.

An important issue with using modularity as a quality function to identify communities is that modularity optimization can fail to find communities that are smaller than a scale that depends on the total size of a network and on the extent of interconnectedness between network communities (Fortunato and Barthelemy 2007). However, many modularity-optimization techniques can easily be adapted to other quality functions, and several alternatives have been proposed that avoid this resolution limit by uncovering communities at multiple scales (Reichardt and Bornholdt 2006, Arenas *et al.* 2008, Lancichinetti *et al.* 2009).

Reichardt and Bornholdt (2006) proposed a multi-resolution method in which a network **A** is represented as an infinite-range, *n*-state Potts spin glass in which each node is a spin, each edge is a pairwise interaction between spins, and each community is a spin state. The Hamiltonian of this system is given by

$$\mathcal{H}(\gamma) = -\sum_{ij} J_{ij} \delta(c_i, c_j), \tag{9}$$

where c_i is the state of spin i and J_{ij} is the interaction energy between spins i and j. The coupling strength J_{ij} is given by $J_{ij} = A_{ij} - \gamma P_{ij}$, where P_{ij} again denotes the expected weight of the link with which nodes i and j are connected in a null model and γ is a resolution parameter. One can find communities by assigning each spin to a state and minimizing the interaction energy in equation (9). Within this framework, community identification is equivalent to finding the ground-state configuration of a spin glass.

Tuning γ allows one to find communities at different resolutions. As γ becomes larger, there is a greater incentive for nodes to belong to smaller communities. The Potts method therefore allows the investigation of communities below the resolution limit of modularity. One can write a scaled energy Q_s in terms of the Hamiltonian in equation (9) as

$$Q_s = \frac{-\mathcal{H}(\gamma)}{2m}. \tag{10}$$

The modularity is then the scaled energy with $\gamma = 1$. Community detection using modularity optimization is therefore a special case of the Potts method.

Recently, an alternative version of the Potts method has been proposed that is able to deal with both positive and negative links (Traag and Bruggeman 2009). One can apply this technique to FX data using the correlation matrix ρ as the network adjacency matrix. Using this approach and a uniform null model, we found the same robust communities (see section 4 for a discussion of robust communities) as we identified using the Potts method and the adjacency matrix in equation (4). However, the Potts method for signed adjacency matrices did not identify the same robust communities when we employed the signed null model of Traag and Bruggeman (2009).

In this paper, we use the Potts method to detect communities of exchange rates in FX networks with adjacency matrices given by equation (4), and we employ the NG model of random link assignment $P_{ij} = k_i k_j / (2m)$ as a null model.† The number of possible community partitions grows at least exponentially with the number of nodes (Newman 2004b), so it is typically impossible computationally to sample the energy space by exhaustively enumerating all partitions (Brandes *et al.* 2008). Several different heuristic procedures have been proposed to balance the quality of the identified optimal partition with computational cost (Danon *et al.* 2005, Porter *et al.* 2009, Fortunato 2010). We minimize equation (9) at each resolution using the locally greedy Louvain algorithm (Blondel *et al.* 2008). We discuss the effect on our results of using different optimization heuristics in appendix A. We find the same aggregate conclusions, although there are some differences in the communities identified using different heuristics. We identify the same changes taking place in the FX market for each of the different algorithms that we use to minimize energy.

†If we include self-edges in the network, the strength of each node increases by 1. This, in turn, leads to a constant increase in the expected edge weights in the null model. For a network with self-edges, the expected edge weight between nodes i and j is given by $P^s_{ij} = n/[2(n+2)]$. This constitutes a shift by a constant value of $P^s_{ij} - P_{ij} = 2/[n(n+2)] \approx 1.62 \times 10^{-4}$ relative to a network in which self-edges are excluded. Self-edges always occur within a community, so they will always contribute to the summation in equation (9) irrespective of exactly how the nodes are partitioned into communities. This implies that self-edges play no role when determining an FX-network partition that minimizes the interaction energy at a particular resolution.

4. Robust community partitions

In some networks, similar community structure persists across a range of resolutions (Reichardt and Bornholdt 2006, Arenas *et al.* 2008, Fortunato 2010). As one increases the resolution parameter in the Potts method, there is an increased energy incentive for nodes to belong to smaller clusters. Network partitions that are robust across a range of resolutions are therefore significant because the communities do not break up despite an increasing incentive to do so. Communities in robust partitions have been found to correspond to communities imposed by construction in simulated networks and to known groupings in real-world networks (Arenas *et al.* 2008, Fortunato 2010). This suggests that communities that persist over a large range of resolutions potentially represent important structures.

We compare network partitions using the normalized variation of information $\hat{V}$ (Meilă 2007, see also Traud *et al.* 2011). The entropy of a partition C of the n nodes in $\mathbf{A}$ into K communities c^k ($k \in \{1, \ldots, K\}$) is

$$S(C) = -\sum_{k=1}^{K} p(k) \log p(k), \tag{11}$$

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 the kth community.† For a partition C, the entropy therefore indicates the uncertainty in the community membership of a randomly-chosen node. Given a second partition C' of the n nodes into K' communities, the mutual information $I(C, C')$ is given by

$$I(C, C') = \sum_{k=1}^{K} \sum_{k'=1}^{K'} p(k, k') \log \frac{p(k, k')}{p(k)\, p(k')}, \tag{12}$$

where $p(k, k') = |c^k \cap c^{k'}|/n$. The mutual information is the amount (averaged over all nodes) by which knowledge of a node's community in C reduces the uncertainty about its community membership in C'. The normalized variation of information $\hat{V}$ between C and C' is then given by

$$\hat{V}(C, C') = \frac{S(C) + S(C') - 2I(C, C')}{\log n}. \tag{13}$$

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. We will use equation (13) to compare partitions in networks with the same number of nodes and remark that one should not normalize by $\log n$ when comparing the variation of information in data sets with different sizes (Meilă 2007).

Variation of information is a desirable measure for quantifying the difference between partitions of a network because it is a metric on the space of community assignments and it thus satisfies the triangle inequality. Therefore, if two partitions are close to a third partition, they cannot differ too much from each other. It is also a local measure, so the contribution to $\hat{V}(C, C')$ from changes in a single community does not depend on how the rest of the nodes are clustered (Meilă 2007, Karrer *et al.* 2008).

One can identify robust communities by detecting communities at multiple resolutions and calculating $\hat{V}(C, C')$ between the network partitions for consecutive resolutions. Robust communities are revealed by intervals in which there are few spikes in $\hat{V}(C, C')$. In figure 3(a), we show $\hat{V}(C, C')$ between network partitions computed at 100 resolutions in the interval $\gamma \in [0.6, 2.1]$ separated by $\Delta\gamma = 0.015$. We focus on this interval in this example because all of the nodes are assigned to the same community at $\gamma = 0.6$ and all of the nodes are assigned to singleton communities at $\gamma = 2.1$. One can also identify robust communities by examining summary statistics that describe community structure as a function of the resolution parameter. We consider the number of communities N_c, the optimized modularity Q_s (see equation 10), the entropy S (see equation 11), and the rate of change of the energy with resolution $\mathrm{d}\mathcal{H}/\mathrm{d}\gamma$ (see equation 9 for the definition of $\mathcal{H}$). Robust communities correspond to plateaus (constant values) in curves of any of these quantities as a function of the resolution parameter. In figure 3(a), we plot curves for each of the summary statistics as a function of γ.

Figure 3(a) contains four principle plateaus, corresponding to partitions of the FX network into $N_c = 1$, 2, 20, and 110 communities. The first and last plateaus, respectively, represent all nodes in one community and all nodes in singleton communities. The second plateau represents one community of exchange rates and a corresponding community of inverse rates. The $N_c = 20$ plateau occurs over the interval $\gamma \in [1.34, 1.57]$, in which there is a single plateau in the N_c plot and a few smaller plateaus in each of the other plots. In contrast to the other plateaus, this one was not expected, so the robust communities over this interval can potentially provide new insights into the correlation structure of the FX market. Although the community configuration over this interval does not have maximal Q (i.e., it is not the community configuration corresponding to the maximum value of the traditional modularity, which is the scaled energy with $\gamma = 1$), it provides an appropriate resolution

†Note that the quantity c^k represents the kth community but that c_i is the set of nodes in the same community as node i.

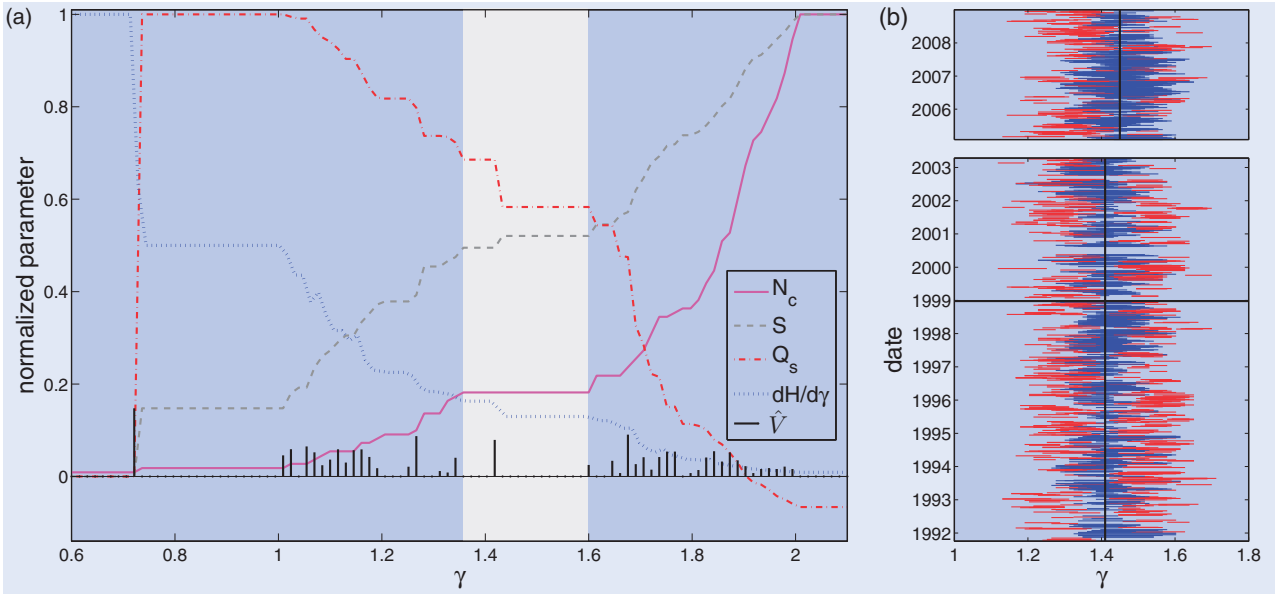

Figure 3. (a) The quantities N_c, S, Q_s, and $d\mathcal{H}/d\gamma$ (defined in the text), normalized by their maximum values, versus the resolution parameter γ for a single time window beginning on 17 March 1992. The lightly shaded rectangle highlights the main plateau. The bottom curve gives the normalized variation of information $\hat{V}$ between partitions at resolutions separated by $\Delta\gamma = 0.015$. (b) Position of the main plateau at each time step for the full period 1991–2008. Main plateaus (blue) containing the fixed resolution (set to $\gamma = 1.41$ for 1991–2003 and to $\gamma = 1.45$ for 2005–2008) and (red) not containing the fixed resolution. The solid black line separates the pre- and post-euro periods. Panel (b) is separated into two sections because we do not possess data for 2004.

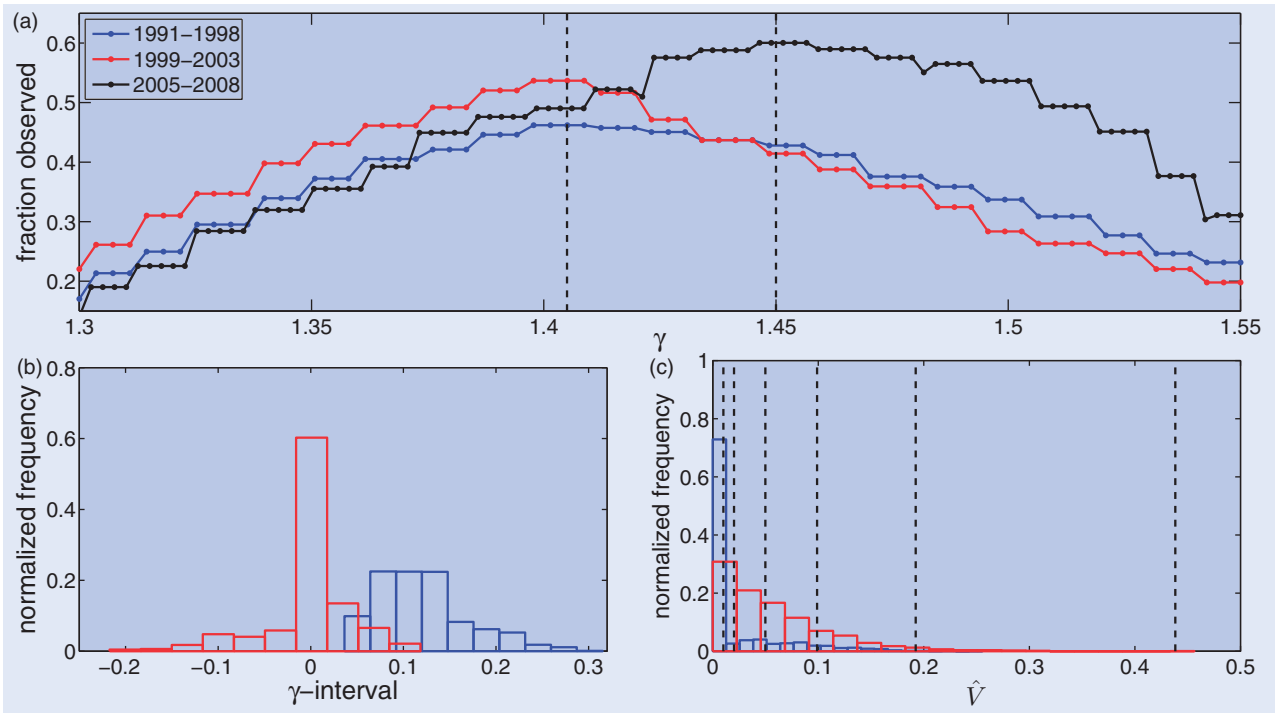

Figure 4. (a) Observed fraction of time steps that the resolution γ lies on the main plateau. The vertical lines indicate $\gamma = 1.41$, which lies in the highest number of main plateaus for the period 1991–2003, and $\gamma = 1.45$, which lies in the highest number of main plateaus for 2005–2008. These are the resolutions at which we investigate community dynamics over the two periods. For the full period 1991–2008, we show in panel (b) the normalized sampled distribution of the main plateau width (blue) and the normalized sampled distribution of the γ-distance between the main plateau and the fixed resolution (red). We label the x-axis in panel (b) as 'γ-interval' for both the main plateau width and the γ-distance between the main plateau and the fixed resolution value. The distance is exactly 0 for 53% of the time steps. Again for 1991–2008, we show in panel (c) the distribution of the normalized variation of information between the community structure detected at the fixed resolution and the community structure corresponding to the main plateau (blue) and the distribution of the normalized variation of information between consecutive time steps (red). The value of $\hat{V}$ is exactly 0 for 64% of the time steps. The vertical lines give the mean $\hat{V}$ when (left to right) 1, 2, 5, 10, 20, and 50 nodes are assigned uniformly at random to different communities (averaged over 100 reassignments for each time step).

at which to investigate community dynamics due to its robustness 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.

5. Dynamic community detection

5.1. Choosing a resolution

To investigate community dynamics, we first choose a resolution parameter at which to detect communities at each time step. One approach is always selecting a resolution γ in the main plateau. As shown in figures 3(b) and 4(a), this plateau occurs over different γ-intervals at different time steps and has different widths. These intervals need not share common resolution values, so this method seems inappropriate because one would then be comparing communities obtained from many different resolutions. Therefore, we fix the resolution at the value that occurs within the largest number of main plateaus. As shown in figure 4(a), this corresponds to $\gamma = 1.41$ for the period 1991–2003 and to $\gamma = 1.45$ for the period 2005–2008.†

In order to demonstrate the validity of this technique, we show in figure 4(b) the distribution of the γ-distance from the fixed resolution to the main plateau, and we show in figure 4(c) the distribution of the normalized variation of information between the community configuration obtained at the fixed resolution and that corresponding to the main plateau. Both distributions are strongly peaked at 0. The fixed resolution is a γ-distance of less than 0.05 from the main plateau 91% of the time for the period 1991–1998, 93% of the time for 1999–2003, and 88% of the time for 2005–2008. The community configurations of the main plateau and the fixed resolution differ in the community assignments of fewer than 5 nodes in 78% of time steps for the period 1991–1998, in 83% of time steps for 1999–2003, and in 88% of time steps for 2005–2008. For the majority of time steps, the community configuration at the fixed resolution is hence identical or very similar to the configuration corresponding to the main plateau. This supports our proposed method of investigating community dynamics at a fixed γ for each period.

5.2. Testing community significance

The scaled energy (see equation 10) measures the strength of communities compared with some null model, so large scaled energies indicate more significant communities. To ensure that the identified communities are meaningful, we perform a permutation test (Good 2005) and compare the scaled energies of the observed network partitions with the scaled energies for network partitions obtained using shuffled data. For the period 1991–2003, we generate shuffled data for each of the USD exchange rates by randomly reordering the returns of the corresponding time series. We create shuffled data for each of the non-USD exchange rates using the shuffled USD time series and the triangle relations described in section 2. We then calculate new correlation matrices for these shuffled time series, form new adjacency matrices, and find the communities and scaled energies for each of the new networks. Similarly, for the period 2005–2008, we shuffle the returns for each of the exchange rates in the set {AUD/USD, EUR/NOK, EUR/SEK, EUR/USD, GBP/USD, NZD/USD, USD/CAD, USD/CHF, USD/JPY, USD/XAU} and calculate the return time series for each of the rates not in this set by applying the triangle relations to these shuffled time series. This procedure conserves the return distribution for each of the original USD exchange rates for the period 1991–2003 and for each of the rates in the above set for 2005–2008. The shuffling, however, destroys the temporal correlations. Any structure in the shuffled data therefore emerges as a result of the triangle relationships. The shuffled data therefore provides some insights into the effects of the triangle relations on the properties of the actual data.

Figure 5(a) shows that the communities identified for the actual data are significantly stronger than those generated using shuffled data. The sample mean scaled energy for the actual data is 0.011 (with a standard deviation of 0.0061); for the shuffled data, the sample mean is 0.0039 (with a standard deviation of 0.0013). The communities that we observe for the actual data are therefore significantly stronger than the communities for randomized data in which the structure results from the triangle effect. This provides strong evidence that the communities represent meaningful structures within the FX market, so these communities can provide insights into the correlation structure of the market. We now consider properties of these communities in detail.

5.3. Community properties

Figure 5(b) shows that the number of communities into which the FX network is partitioned exhibits only small fluctuations during the period 1991–2008. Nevertheless, as shown in figure 4(c), there is a considerable variation in the extent of community reorganization between consecutive time steps. No nodes change community assignment between some steps, whereas more than twenty nodes change community assignments between others. Figure 5(c) shows that the community size distribution is bimodal for all three periods, and its tail extends to large community sizes. There is therefore a large variation in the

†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 γ, then we find identical communities in the corresponding network with self-edges using a resolution parameter $\gamma^s = \gamma p_{ij}/p^s_{ij} = \gamma(n+2)(n-2)/n^2$, where p^s_{ij} denotes the null model when self-edges are included. 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 $\gamma_s = 1.4495$.

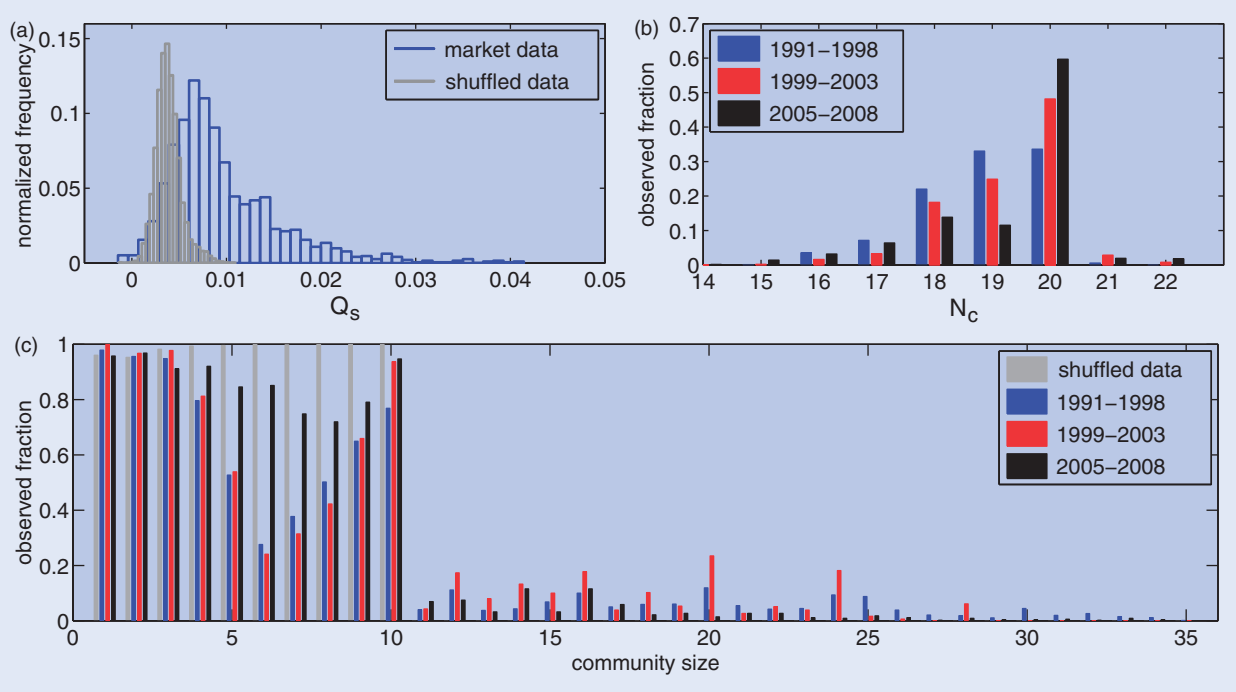

Figure 5. (a) Comparison of the distribution of the scaled energy for 1991–2003 for market data (blue) and 100 realizations of shuffled data (grey). (b) Fraction of time steps at which N_c communities are observed for 1991–1998 (blue), 1999–2003 (red), and 2005–2008 (black). (c) Fraction of time steps at which a community of a given size is observed for 1991–1998 (blue), 1999–2003 (red), 2005–2008 (black), and shuffled data (grey). The shuffled data distribution combines the results for the periods 1991–2003 and 2005–2008. The distributions are almost identical for the two periods.

sizes of the communities observed at each time step for all three periods. However, the minimum between the two peaks is not as deep for the period 2005–2008, and it has shifted from a community size of six nodes to a size of eight.

The peak in the size distribution for communities with 10 members occurs as a result of the number of currencies in the network. For each of the eleven currencies XXX $\in$ {AUD, CAD, CHF, GBP, DEM, JPY, NOK, NZD, SEK, USD, XAU}, there are 10 exchange rates XXX/YYY with XXX as the base currency (and 10 equivalent inverse rates YYY/XXX). We derive most of the exchange rates in a set of rates with the same base currency by applying the triangle relation (see section 2) to pairs of exchange-rate time series; one of the rates is common across all of the exchange rates in the base-currency set, and the other rate is different for each rate in the set. For example, for the period 1991–2003, we derive the CAD/DEM exchange rate from the USD/CAD and USD/DEM rates, whereas we derive the CAD/GBP rate from the USD/CAD and USD/GBP rates. Exchange rates with the same base currency are, therefore, often correlated, and they consequently have a tendency to form communities with 10 members. However, it is not possible for all currencies to form a 10-member base-currency community at each time step. If there is no additional structure beyond these base-currency correlations that emerge as a result of the triangle relation, then one would expect to observe communities with $1, 2, \ldots, 10$ members at each time step (and equivalent communities of inverse rates). Figure 5(c) shows that this size distribution is indeed observed for shuffled data. However, figure 5(c) also shows that the community-size distribution for market data is significantly different, so the community-detection techniques are uncovering additional FX market correlations. This again demonstrates that the triangle effect is not dominating the results.

The frequently-observed communities shown in table 1 demonstrate the variation in community size. Some of the most common communities are single exchange rates, such as USD/CAD, which are formed of two closely-related currencies. Table 1 also highlights that communities often consist of exchange rates with the same base currency. McDonald *et al.* (2008) used the relative clustering strengths of groups of exchange rates with the same base currency to provide insights into the effects of important news and events on individual currencies. The relative sizes of different base-currency communities can provide similar information. For example, if we observe a community of ten CHF/YYY exchange rates and a community of three DEM/YYY rates, then the larger size of the CHF/YYY community suggests that CHF is more important than DEM in the market at this time.

It is also worth noting that the most frequently observed community of 10 exchange rates with the same base currency is the gold (XAU) community. We include gold in our study because there are many similarities between it and a currency. However, gold also tends to be more volatile than most currencies, so it is unsurprising that the gold rates often form their own community. Consequently, the absence of a large gold community at a given time is an indication that another currency is particularly influential.

Importantly, the identified communities do not always contain exchange rates with the same base currency,

Table 1. Examples of frequently-observed communities for the pre-euro period 1991–1998 and for the two post-euro periods (1999–2003 and 2005–2008). The quantity Fr denotes the fraction of time steps at which each community is observed.

Period	Community	Fr
1991–1998	USD/CAD	0.62
	DEM/CHF	0.45
	NZD/{CAD, USD}	0.33
	AUD/{CAD, NZD, USD}	0.32
	XAU/{AUD, CAD, CHF, DEM, GBP, JPY, NOK, NZD, SEK, USD}	0.28
	SEK/{AUD, CAD, CHF, DEM, GBP, JPY, NOK, NZD, USD, XAU}	0.17
	DEM/NOK	0.16
	AUD/{CAD, NZD, USD, XAU}	0.14
	GBP/{CHF, DEM, NOK}	0.12
1999–2003	EUR/CHF	0.88
	USD/CAD	0.67
	XAU/{AUD, CAD, CHF, EUR, GBP, JPY, NOK, NZD, SEK, USD}	0.64
	NOK/{CHF, EUR}	0.59
	SEK/{CHF, EUR, NOK}	0.51
	GBP/{CAD, USD}	0.24
	NZD/{AUD, CAD, CHF, EUR, GBP, JPY, NOK, SEK, USD}	0.21
	JPY/{CAD, GBP, USD}	0.17
	AUD/{CAD, CHF, EUR, GBP, JPY,NOK, SEK, USD}	0.14
2005–2008	XAU/{AUD, CAD, CHF, EUR, GBP, JPY, NOK, NZD, SEK, USD}	0.91
	EUR/CHF	0.65
	AUD/NZD	0.39
	CAD/{AUD, CHF, EUR, GBP, JPY, NOK, NZD, SEK, USD}	0.39
	GBP/{CHF, EUR}	0.35
	SEK/{CHF, EUR}	0.33
	NZD/{AUD, CAD, CHF, EUR, GBP, JPY, NOK, SEK, USD}	0.26
	NOK/{CHF, EUR, SEK}	0.21
	GBP/{CHF, EUR, NOK, SEK}	0.20

which provides insights into changes in the inherent values of different currencies. For example, consider a community containing several exchange rates with CHF as the base currency and several rates with DEM as the base currency. The fact that the exchange rates are in the same community suggests that they are correlated. The structure of this community also provides information about the inherent values of CHF and DEM. Exchange rates of the form XXX/YYY quote the value of one currency in terms of another currency, so if the price of XXX/YYY increases, it is not clear whether this is because XXX has become more valuable or because YYY has become less valuable. However, if one observes that the price of XXX increases with respect to several different YYY over the same period, then one expects that the value of XXX has increased. Therefore, returning to our example, if one observes a community of several CHF/YYY and DEM/YYY exchange rates for many different YYY, then one expects that these rates are positively correlated. Because the values of CHF and DEM have increased versus several other currencies, we expect that the inherent values of both CHF and DEM have increased.

6. Comparison with linkage clustering

Perhaps the best-known approach for studying a network of financial assets is to consider the minimum spanning tree (MST) of the network. This is closely related to a dendrogram (i.e., a hierarchical tree). MSTs have been used regularly in studies of equity markets to identify clusters of stocks that belong to the same market sector (Mantegna 1999, Bouchaud and Potters 2003, Onnela *et al*. 2003, 2004). In this section, we briefly consider the limitations of this approach for community detection and describe the additional information that the Potts method can provide.

MSTs are constructed using the agglomerative hierarchical clustering technique known as single-linkage clustering (Duda *et al*. 2001, Porter *et al*. 2009). Agglomerative methods start with n singleton clusters and create a hierarchy by sequentially linking clusters based on their similarity. The similarity of clusters c and c' is usually expressed as a distance $\mathcal{D}(c, c')$, which is determined by considering the distance d_{ij} between each node $i \in c$ and each node $j \in c'$. In single-linkage clustering, the distance between clusters is given by

$$\mathcal{D}(c, c') = \min_{\substack{i \in c \\ j \in c'}} d_{ij}. \tag{14}$$

It thus represents an extreme because it joins clusters based on the minimum distance between nodes in each cluster. An alternative is average-linkage clustering, for which

$$\mathcal{D}(c, c') = \frac{1}{|c||c'|} \sum_{i \in c} \sum_{j \in c'} d_{ij}. \tag{15}$$

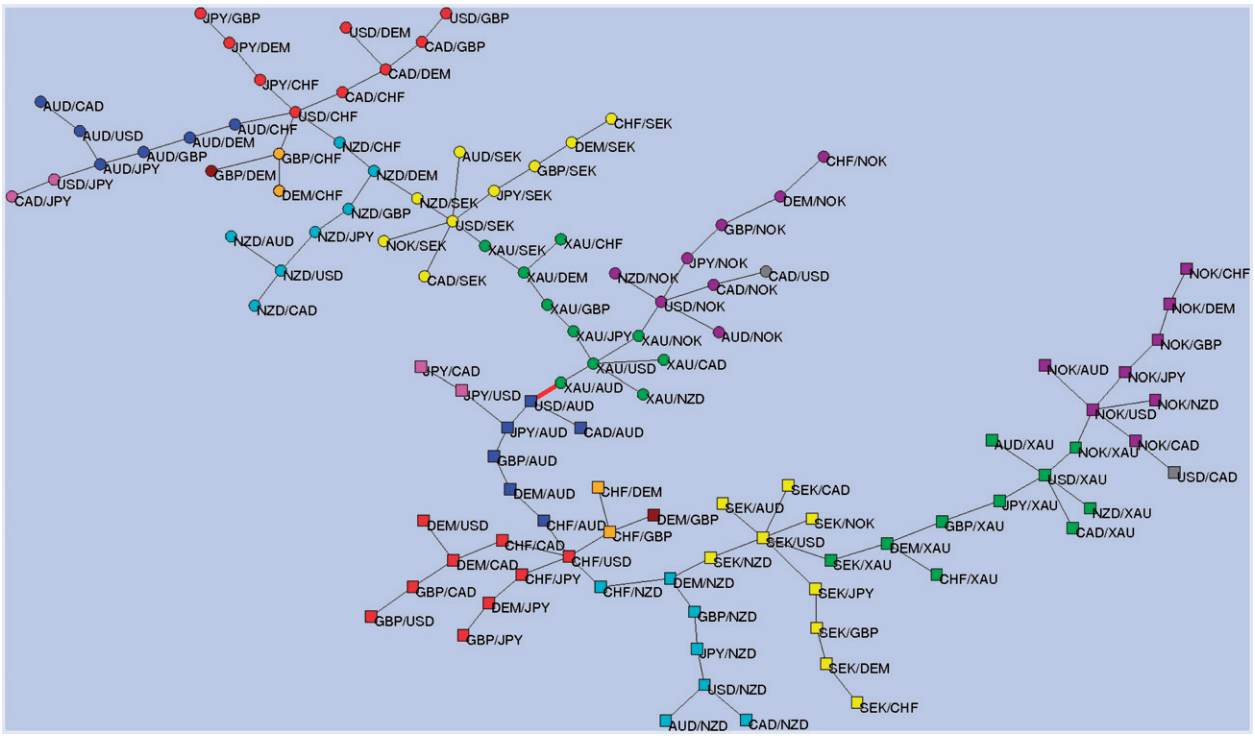

Figure 6. A minimum spanning tree for the FX network formed from a time window of returns beginning on 18 September 1991. The tree is split into two identical halves (indicated by ○ and □), which are connected via the edge (shown in red) between the XAU/USD and USD/AUD exchange rates. For each community of exchange rates, there is an equivalent community of inverse rates in the other half of the tree. We colour each node according to its community membership determined using the Potts method with $\gamma = 1.41$, and we show each community of exchange rates in the same colour as the corresponding community of inverse rates.

For financial networks, the standard measure used for d_{ij} is a particular nonlinear transformation of the correlation coefficient $\rho(i, j)$. The distance elements are given by (Mantegna 1999, Mantegna and Stanley 2000) the formula

$$d_{ij} = \sqrt{2[1 - \rho(i,j)]}. \tag{16}$$

The distance d_{ij} takes values in the interval [0,2], and similar nodes have small values of d_{ij}. An MST possesses $n-1$ links and is appealing because it is much simpler to analyse than the full network with $n(n-1)/2$ links. A dendrogram provides an alternative representation of the output of a linkage-clustering algorithm that shows the full hierarchical structure (Duda *et al.* 2001, Porter *et al.* 2009). At the first level of a dendrogram, there are n singleton clusters. As one climbs the vertical distance scale of a dendrogram, clusters are combined until all nodes are contained in a single community at the top of the dendrogram.

In earlier studies of equity markets, clusters of closely-related assets were identified based on their proximity on the branches of an MST (Mantegna 1999, Onnela *et al.* 2003, 2004) and by finding the disconnected groups of assets that remained when all tree links weaker than some threshold were removed (Bouchaud and Potters 2003). Similar computations have found clusters of assets by considering an entire network and removing edges below some threshold or alternatively by starting with a network with no links and iteratively adding links above an increasing threshold (Onnela *et al.* 2004, Garas *et al.* 2008). In figure 6, we show an example of an MST of exchange rates. We colour the nodes in this tree according to their community membership as determined using the Potts method. The MST is partitioned into two halves, with communities of exchange rates in one half and equivalent communities of inverse exchange rates in the other. In this example, nodes belonging to the same community are always grouped contiguously in the MST, but this is not always the case.

The main problem with single-linkage clustering (and, as a consequence, with MSTs) is that clusters can be joined as a result of single pairs of elements being close to each other even though many of the elements in the two clusters are rather dissimilar. An MST then contains weak links that might be misinterpreted as being more financially meaningful than they actually are (Onnela *et al.* 2004). It is also difficult to determine where the community boundaries lie on an MST. For example, a branch of an MST might include nodes belonging to a single community or the nodes might belong to several communities. As an example of this phenomenon, and of the additional clustering information provided by the Potts method, consider the branch at the far right of the tree shown in figure 6. By simply considering this MST, one might have inferred the existence of a cluster that includes all of the NOK/YYY rates and USD/CAD. However, the Potts method highlights the fact that USD/CAD forms a singleton community and that NOK/XAU belongs to a community with the XXX/XAU rates. This observation might provide information as to the relative importances of NOK and XAU in the market over this period.

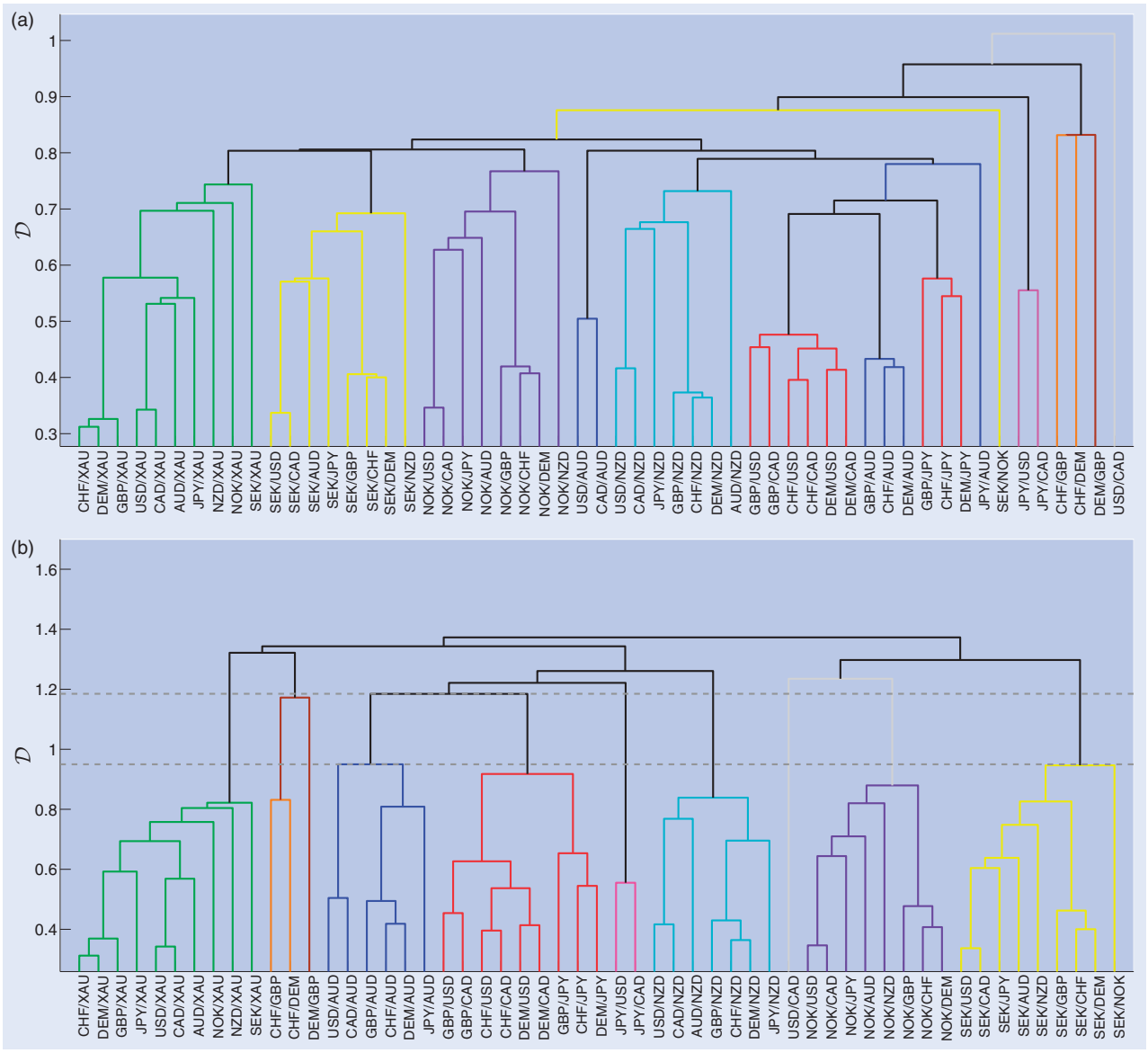

Figure 7. Dendrograms showing hierarchical clustering of exchange rates for one half of the FX network for a time window of returns beginning on 18 September 1991. We colour each exchange rate according to its community membership determined using the Potts method with $\gamma = 1.41$. We generated the dendrograms using (a) single-linkage clustering and (b) average-linkage clustering. The dashed grey lines in panel (b) highlight the range over which the communities correspond to the communities of the main plateau identified using the Potts method.

In figure 7(a), we show a dendrogram generated using the same single-linkage clustering algorithm used to produce the MST in figure 6. If the distances between different dendrogram levels are reasonably uniform, then no clustering appears more 'natural' than any other (Duda *et al.* 2001). However, large distances between levels (i.e., the same clusters persist over a large range of distances) might indicate the most appropriate level at which to view the clusters. This is analogous to investigating communities that are robust over a range of resolutions. The clusterings observed at some levels of figure 7(a) correspond closely with the communities identified using the Potts method, but there is no level at which they correspond exactly. The levels are also reasonably evenly distributed along the distance axis. In the dendrogram in figure 7(b), which we generated using average-linkage clustering, there is a range of distances over which the clustering does not change. The clustering observed over this interval is identical to the community configuration corresponding to the main plateau found using the Potts method. Therefore, in this case, average-linkage clustering and the Potts method identify the same robust communities.

7. Exchange-rate centralities and community persistence

Thus far, we have considered the properties of entire communities. We now investigate the roles of nodes within communities.

7.1. Centrality measures

We describe relationships between a node and its community using various centrality measures. In the social networks literature, such measures are used to measure the roles of nodes within networks and to identify which nodes are the most important or most prominent (Wasserman and Faust 1994). Because there are multiple notions of importance, many different centrality measures have been proposed (Valente *et al.* 2008). In the present context, we use centrality measures to identify exchange rates that occupy important positions in the FX market.

The *betweenness centrality* of nodes is defined using the number of geodesic paths between pairs of vertices in a network (Freeman 1977, Newman 2003). We calculate node betweenness by taking the distance between nodes i and j as

$$d_{ij} = \begin{cases} 0 & \text{if } i = j \text{ or } A_{ij} = 1, \\ 1/A_{ij} & \text{otherwise.} \end{cases} \tag{17}$$

The betweenness centrality b_i of node i is then given by

$$b_i = \sum_s \sum_t \frac{g^i_{st}}{G_{st}}, \quad \text{for } s, t \neq i \text{ and } s \neq t, \tag{18}$$

where G_{st} is the total number of shortest paths from node s to node t and g^i_{st} is the number of shortest paths from s to t passing through i. Betweenness centrality is used widely in social network analysis to quantify the extent to which people lie on paths that connect others. Nodes with high betweenness can be construed to be important for facilitating communication between others in a network, so betweenness is used to help measure the importance of nodes for the spread of information around a network (Valente *et al.* 2008).

We also consider the community centrality of each node (Newman 2006a). We employ the scaled energy matrix $\mathbf{J}$, with components $J_{ij} = A_{ij} - \gamma P_{ij}$, where we again set $P_{ij} = k_i k_j/(2m) = (n-2)/(2n)$. Following the notation in Newman (2006a), the energy matrix can be expressed as $\mathbf{J} = \mathbf{U}\mathbf{D}\mathbf{U}^{\mathrm{T}}$, where $\mathbf{U} = (\mathbf{u}_1|\mathbf{u}_2|\cdots)$ is the matrix of eigenvectors of $\mathbf{J}$, and $\mathbf{D}$ is the diagonal matrix of eigenvalues β_i. If $\mathbf{D}$ has q positive eigenvalues, then one can define a set of node vectors $\{\mathbf{x}_i\}$ of dimension q by

$$[\mathbf{x}_i]_j = \sqrt{\beta_i} U_{ij}, \quad j \in \{1, 2, \ldots, q\}, \tag{19}$$

where $[\mathbf{x}_i]_j$ indicates the jth element of the node vector of node i. The magnitude $|\mathbf{x}_i|$ is the *community centrality*. Nodes with high community centrality play an important role in their local neighbourhood, irrespective of community boundaries.

One can also define a community vector

$$\mathbf{w}_k = \sum_{i \in c^k} \mathbf{x}_i \tag{20}$$

for each community c^k. Nodes with high community centrality are strongly attached to their community if their node vector is also aligned with their community vector. One defines *projected community centrality* y_i by (Newman 2006a)

$$y_i = \mathbf{x}_i \cdot \hat{\mathbf{w}}_k = |\mathbf{x}_i| \cos\theta_{ik}, \tag{21}$$

and we refer to the quantity $\cos\theta_{ik}$ as the *community alignment*. The community alignment is near 1 when a node is at the centre of its community, and it is near 0 when it is on the periphery. Nodes with high community alignment are located near the centre of their community and have a high projected community centrality, so they are strongly attached to their community and can be construed to be highly influential within it. The number of positive eigenvalues of $\mathbf{J}$ can vary between time steps, so we normalize $|\mathbf{x}_i|$ and y_i by their maximum value at each time step.

7.2. Community tracking

A node's identity is known at all times and its community is known at any given time. We can thus track community evolution from the perspective of individual nodes. 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 $a^t_i(\tau)$ of its community after τ time steps is defined by

$$a^t_i(\tau) = \frac{|\, c_i(t) \cap c_i(t+\tau) \,|}{|\, c_i(t) \cup c_i(t+\tau) \,|}. \tag{22}$$

This is a node-centric version of a quantity considered in Palla *et al.* (2007) that, importantly, does not require one to determine which community at each time step represents the descendant of a community at the previous time step. Palla *et al.* (2007) detected communities using a method known as k-clique percolation. They tracked communities by defining the descendant of community a at time step t as the community at $t+1$ that had the maximum edge overlap with a. Several other approaches have been proposed for identifying community descendants using different measures to quantify the node overlap rather than the edge overlap between communities at different time steps. See, for example, Toyoda and Kitsuregawa (2003), Berger-Wolf and Saia (2006), Falkowski *et al.* (2006), and Asur *et al.* (2007). Methods that identify descendent communities based on maximum node or edge overlap can, however, lead to equivocal mappings following splits and mergers. For example, consider a community $c^f(t)$ that splits into two communities, $c^g(t+1)$ and $c^h(t+1)$, at the following time step. If the extent of overlap between $c^f(t)$ and $c^g(t+1)$ is identical to that between $c^f(t)$ and $c^h(t+1)$, then one will need to make an arbitrary choice as to which community represents the descendant of $c^f(t)$. See section 4.4.6 of Fenn (2011) for a more detailed discussion of this point and a review of the community dynamics literature. In order to avoid this ambiguity, we identify communities from the perspective of individual nodes instead of tracking whole communities.

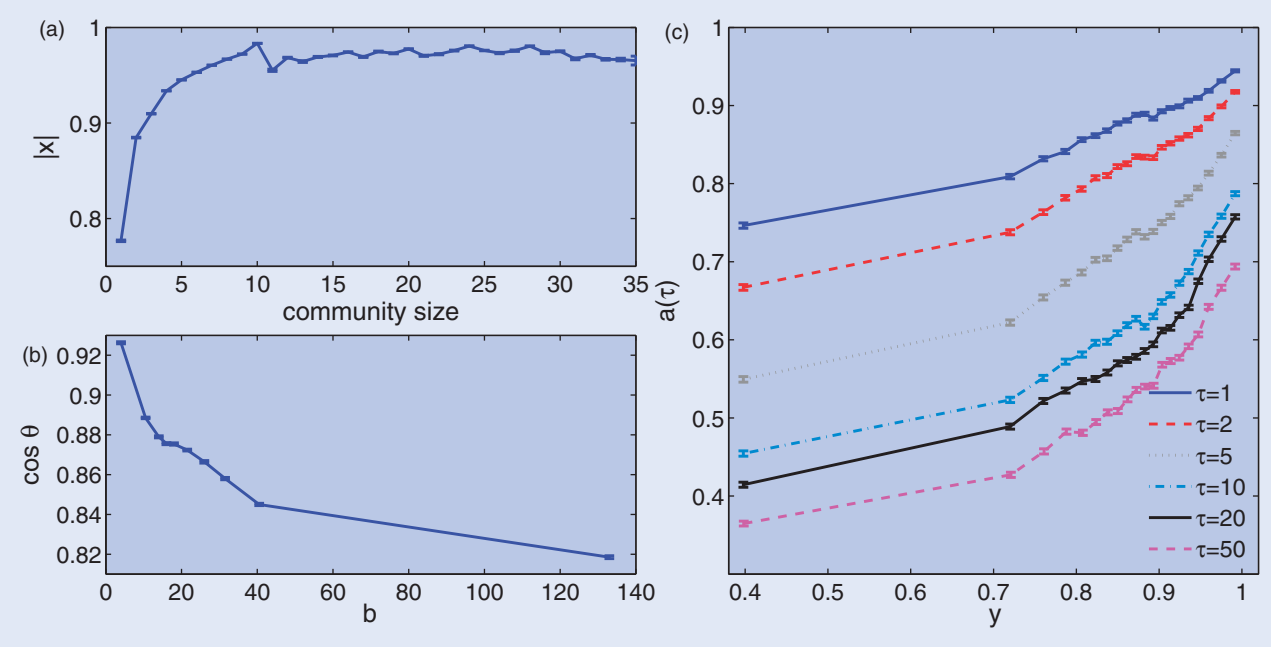

Figure 8. (a) Mean community centrality versus community size. (b) Mean community alignment versus node betweenness centrality. (c) Mean community autocorrelation versus projected community centrality. All error bars indicate the standard error (Berry and Lindgren 1990).

Table 2. The 10 exchange rates with the highest values of betweenness centrality, community centrality, and projected community centrality for each of the three periods. We rank the exchange rates for each centrality according to their mean rank over all time steps. For each exchange rate XXX/YYY, the corresponding inverse rate YYY/XXX has the same betweenness centrality, community centrality, and projected community centrality.

	1991–1998			1999–2003			2005–2008		
Rank	*b*	|x|	*y*	*b*	|x|	*y*	*b*	|x|	*y*
1	NOK/SEK	CHF/AUD	USD/DEM	AUD/NZD	SEK/XAU	USD/XAU	USD/CAD	JPY/XAU	EUR/XAU
2	AUD/XAU	CHF/NZD	USD/CHF	NZD/CAD	CHF/CAD	EUR/USD	AUD/NZD	USD/XAU	GBP/XAU
3	AUD/NZD	CHF/XAU	USD/XAU	AUD/CAD	EUR/XAU	EUR/XAU	AUD/CAD	NZD/XAU	CHF/XAU
4	AUD/CAD	CHF/CAD	CHF/CAD	JPY/CAD	NOK/XAU	GBP/XAU	NOK/SEK	CAD/XAU	EUR/CAD
5	CHF/SEK	DEM/AUD	CHF/AUD	NOK/SEK	CHF/NZD	EUR/CAD	USD/GBP	GBP/XAU	SEK/XAU
6	NZD/XAU	SEK/AUD	CHF/NZD	USD/AUD	CHF/XAU	USD/CHF	NZD/CAD	SEK/XAU	USD/XAU
7	CAD/XAU	DEM/XAU	DEM/CAD	USD/NZD	EUR/CAD	CHF/XAU	USD/JPY	CHF/XAU	EUR/NZD
8	DEM/SEK	SEK/XAU	DEM/AUD	USD/JPY	EUR/NZD	NOK/XAU	USD/AUD	NOK/XAU	JPY/XAU
9	NZD/CAD	NOK/AUD	USD/AUD	GBP/JPY	SEK/NZD	EUR/NZD	CHF/NOK	CHF/NZD	AUD/XAU
10	DEM/NOK	DEM/NZD	DEM/NZD	CHF/SEK	NOK/NZD	CHF/NZD	GBP/AUD	AUD/XAU	NOK/XAU

7.3. *Exchange-rate roles*

In figure 8(a), we show the mean normalized community centrality of exchange rates as a function of community size. (For each community size, we calculated the mean value of $|\mathbf{x}_i|$ over all nodes that were members of a community of that size.) The community centrality increases with community size up to sizes of about 10 members. For larger communities, $|\mathbf{x}_i|$ remains 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. Table 2 shows the 10 exchange rates that tend to have the highest values of betweenness centrality, community centrality, and projected community centrality. For all three periods, CHF/NZD, CHF/XAU, and SEK/XAU have one of the 10 highest community centralities, so they are closely tied to many other rates. For 1991–2003, exchange rates formed from one of the major European currencies—DEM (and then EUR, after its introduction) or CHF—and one of the commodity currencies† also tend to have high community centrality. For 2005–2008, however, XAU rates encompass nearly all of the exchange rates with the highest values of $|\mathbf{x}_i|$.

†A country is said to have a 'commodity currency' if its export income depends heavily on a commodity. For example, AUD, NZD, and CAD are all considered to be commodity currencies.

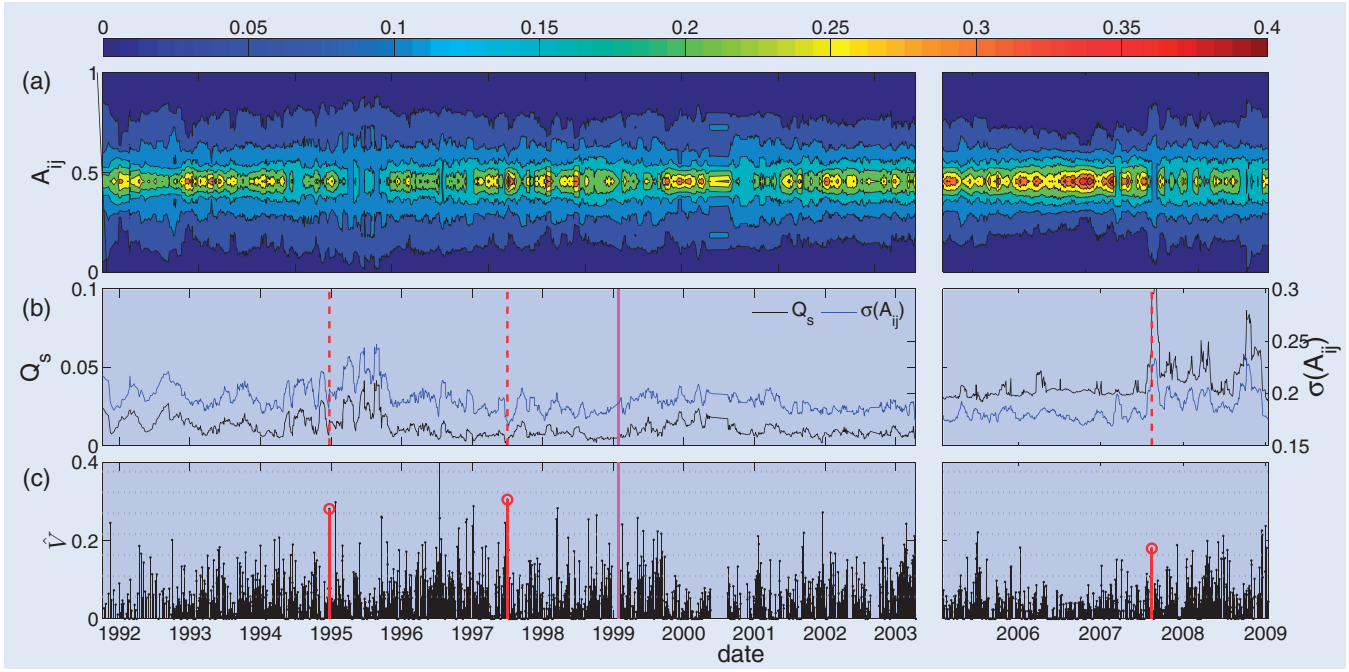

Figure 9. (a) Normalized distribution of the link weights at each time step. (b) Scaled energy Q_s and standard deviation of the link weights A_{ij}. (c) Normalized variation of information $\hat{V}$ between the community configurations at consecutive time steps. The dashed horizontal lines show (from bottom to top) the mean of $\hat{V}$ and 1, 2, 3, 4, 5, and 6 standard deviations above the mean $\hat{V}$. The solid (magenta) vertical lines in panels (b) and (c) separate the pre- and post-euro periods. The red vertical lines show the time steps when 22 December 1994, 7 February 1997, and 15 August 2007 enter the rolling time window. These dates correspond, respectively, to the devaluation of the Thai baht during the Asian currency crisis, the flotation of the Mexican peso following its sudden devaluation during the tequila crisis, and significant unwinding of the carry trade during the 2007–2008 credit crisis. Each panel is separated into two sections because we do not possess data for 2004.

Figure 8(b) shows the mean betweenness centrality versus the community alignment. We calculate the mean community position by splitting the range of b into 10 bins containing equal numbers of data points and then averaging over all community positions falling within these bins. (The observed relationships are robust for reasonable variations in the number of bins.) Nodes with high betweenness centralities tend to have small values for their community position, implying that nodes that are important for information transfer are usually located on the edges of communities. Table 2 shows that for all three periods, NOK/SEK, AUD/NZD, and AUD/CAD all tend to have high values of betweenness centrality on average. They are therefore located on the edges of communities and are important for information transfer. Interestingly, for the post-euro period (1999–2008), several USD exchange rates also seem to be important for information transfer, but no USD rates regularly have high betweenness values for the pre-euro period. In contrast, XAU exchange rates are important for information transfer for the pre-euro period but not after the euro was introduced.

In figure 8(c), we show the mean community autocorrelation versus the projected community centrality. We calculate the mean autocorrelation by splitting the range of the projected community centrality into 20 bins containing equal numbers of data points and then averaging over all autocorrelations falling within each bin. (Again, the observed relationships are robust for reasonable variations in the number of bins.) As one would expect, the community autocorrelation for the projected community centrality of a given node is smaller for larger τ. More interesting, we find for all values of τ that the mean community autocorrelation increases with projected community centrality. This suggests that nodes that are strongly connected to their community are persistently likely to share that community membership with the same subset of nodes. In contrast, exchange rates with low values of projected community centrality experience regular changes in the set of rates with which they are clustered.

Table 2 shows the exchange rates with the highest projected community centralities, which in turn reveals the most persistent communities. For 1991–2003, approximately half of the 10 exchange rates with the highest projected community centralities also appear in the list of the 10 rates with the highest community centralities. For 2005–2008, however, the lists of exchange rates with the highest community centralities and projected community centralities are dominated by the same set of XAU exchange rates (though the rankings differ). For 1991–2003, the exchange rates with the highest projected community centralities again include rates formed of DEM (and EUR) or CHF and one of the commodity currencies. However, there are also several USD exchange rates with high values of projected community centrality that don't have high values of community centrality. This suggests that these USD rates do not have strong links with a large number of other exchange rates, but that they strongly influence the rates within their own communities.

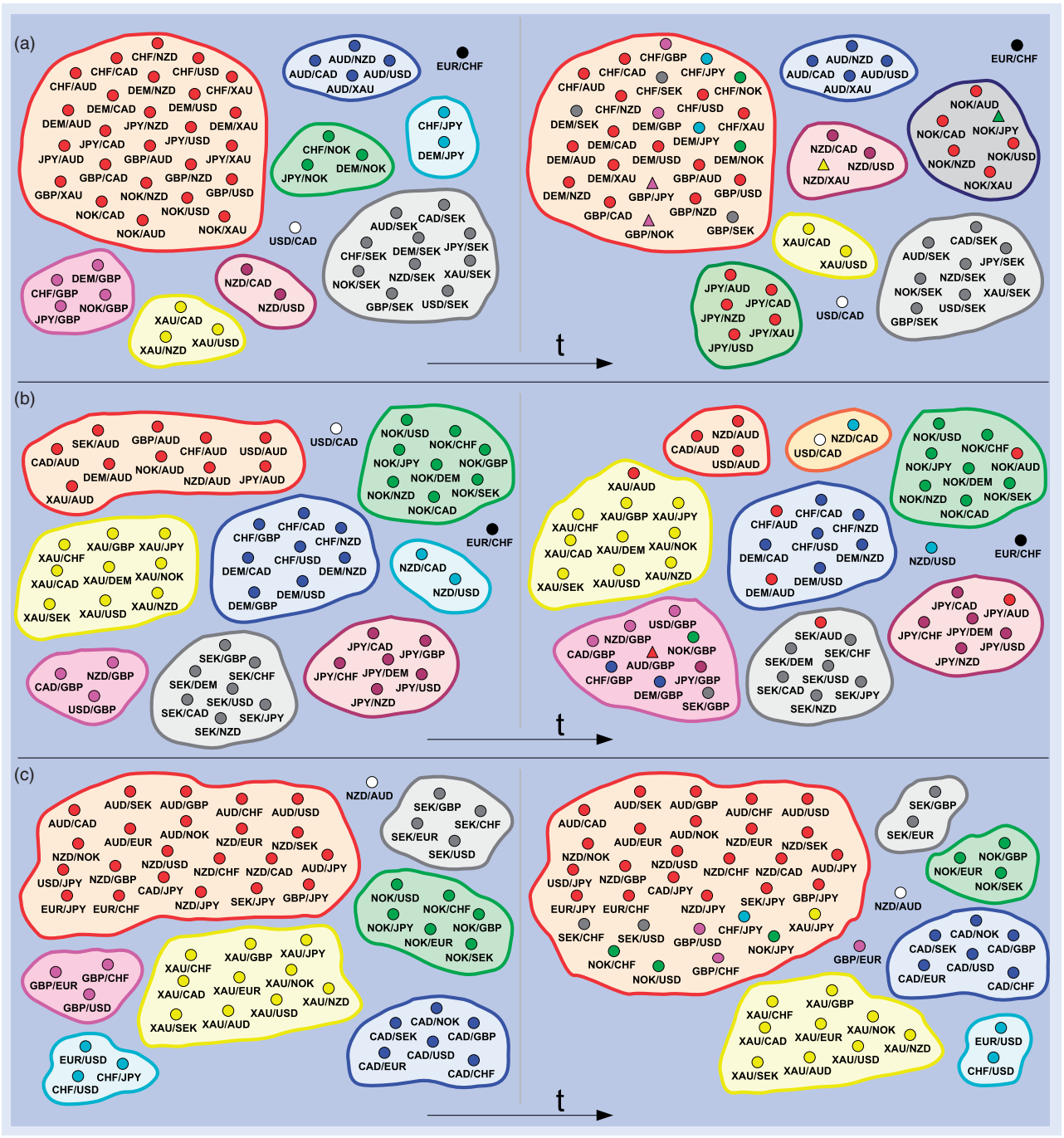

Figure 10. Schematic representation of changes in community structure in one half of the FX market network for several events. (a) The Mexican tequila crisis: the depicted reorganization followed 22 December 1994, when the Mexican peso was allowed to float after a sudden devaluation. (b) The Asian currency crisis: the depicted reorganization followed 2 July 1997, when Thailand devalued the baht. (c) Carry-trade unwinding: the depicted reorganization followed 15 August 2007, when there was significant unwinding of the carry trade during the 2007–2008 credit and liquidity crisis. The node colours after the community reorganization correspond to the communities before the changes. If the parent community of a community after the reorganization is obvious, we draw it using the same colour as its parent. The nodes drawn as triangles resided in the opposite half of the network before the reorganization of community structure.

8. Major community changes

We now investigate the insights that short-term community dynamics can provide into changes in the FX market. Figure 9(a) shows a contour plot of the normalized distribution of the 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 figures 9(a,b)] there is a large variation in the standard deviation of the link strengths. The scaled energy and standard deviation of link weights are closely related. This is expected because the standard deviation increases as a result of the strengthening of strong ties and the weakening of weak ones.

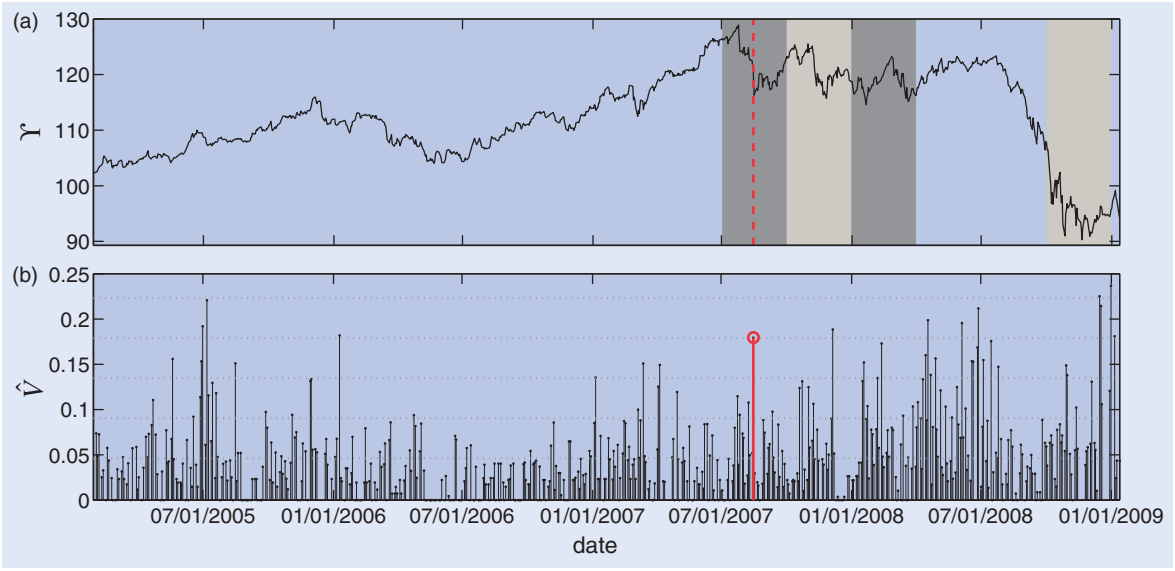

Figure 11. (a) Carry-trade index Υ. The vertical line again shows 15 August 2007 and the shaded blocks (from left to right) show Q3 2007, Q4 2007, Q1 2008, and Q4 2008. (b) Normalized variation of information between community configurations at consecutive time steps for 2005–2008. The horizontal lines show (from bottom to top) the mean of $\hat{V}$ and 1, 2, 3, and 4 standard deviations above the mean. The red vertical line in (b) shows 15 August 2007, when there was a marked increase in unwinding of the carry trade.

In figure 9(c), we also show the normalized variation of information $\hat{V}$ between the community configurations at consecutive time steps. Large spikes in $\hat{V}$ indicate significant changes in community structure during a single time step and potentially also indicate important market changes. The correlation coefficient between $\hat{V}$ and the absolute change in Q_s between consecutive time steps is 0.39 over the period 1991–2003 and 0.47 over the period 2005–2008. (The absolute change in a quantity η from time t to time $t + \Delta t$ is defined as $|\eta(t + \Delta t) - \eta(t)|$.) The correlation between $\hat{V}$ and the absolute change in $\sigma(A_{ij})$ is 0.28 over the period 1991–2003 and 0.27 for 2005–2008. Changes in Q_s are thus a better indicator than changes in $\sigma(A_{ij})$ that there has been a change in the community configuration of the FX network.

In figure 10, we show three example community-structure reorganizations—two in which $\hat{V}$ is more than four standard deviations larger than its mean and a third in which it is more than two standard deviations above its mean.

8.1. Mexican peso crisis

Figure 10(a) shows the reorganization on 22 December 1994, when the Mexican peso was floated following its sudden devaluation.† This change is accompanied by an increase in the scaled energy Q_s. Although we do not include the Mexican peso in the set of investigated exchange rates, it appears that its devaluation was a sufficiently serious event that it led to major changes in the community relationships of the studied rates. Before 22 December 1994, the largest community consisted of a group of exchange rates of the form {AUD, CAD, NZD, USD, XAU}/{CHF, DEM, GBP, JPY, NOK}. After the flotation, the largest community consisted of a set of exchange rates formed from the major European currencies (CHF, DEM, and GBP). It is also noteworthy that there is only a small gold (XAU) community during this period, which—as noted previously—often indicates that another currency is particularly important in the market.

8.2. Asian currency crisis

Figure 10(b) shows the community changes following 2 July 1997, when the Thai baht was devalued during the Asian currency crisis. As with the peso, although we did not include the baht in the set of studied rates, its devaluation appears to have had a significant effect on the FX market. There is a large stable gold cluster during the whole period. Before 2 July 1997, there is also a large AUD cluster. After the devaluation, however, this cluster breaks up and the previously-small GBP cluster increases in size. This suggests that GBP is playing a more prominent market role after the devaluation. Although the reasons for the changes in the sizes of the AUD and GBP communities are not obvious, both adjustments suggest a sharp and significant change in the correlation structure of the market.

8.3. Credit crisis

The final example, which we show in figure 10(c), reveals significant community reorganization following 15 August 2007, and it illustrates one of the major effects on the FX network of the recent credit and liquidity crisis.

†For a floating exchange rate, the value of the currency is allowed to fluctuate according to the FX market. Prior to its flotation, the peso had been pegged to the US dollar.

This example also demonstrates changes in community structure that occurred as a result of a trading change that directly affected the studied rates.

The most important effect of the credit crisis on the FX market during the period 2005–2008 was its impact on the carry trade. The carry trade consists of selling low interest rate 'funding currencies' such as JPY and CHF and investing in high interest rate 'investment currencies' such as AUD and NZD. It yields a profit if interest-rate differentials between funding and investment currencies are not offset by commensurate depreciation of investment currencies (Brunnermeier *et al*. 2008). The carry trade is one of the most commonly used FX trading strategies and requires a strong appetite for risk, so the trade tends to 'unwind' during periods in which there is a decrease in 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 can be achieved using a carry-trade strategy. In figure 11(a), we show the cumulative return index Υ from trading using a common carry-trade strategy. We consider a strategy in which one buys equal weights of the three major currencies with the highest interest rates and sells equal weights of the three major currencies with the lowest interest rates. This is a dynamic trading strategy because the relative interest rates of currencies change over time. For example, consider the situation in which the interest rate of currency A (which initially has the third highest interest rate) decreases below the rate of currency B (which initially has the fourth highest interest rate). In order to maintain the strategy of only holding the three currencies with the highest interest rates at any time, one would re-balance the carry portfolio by selling the holding of currency A and buying currency B. The frequency at which such re-balances occur depends on the frequency at which the relative interest rates change. The returns from a carry strategy like this are widely construed by market participants to provide a good gauge of carry-trade activity. Large negative returns result in large decreases in Υ, which are therefore likely to indicate significant unwinding of the carry trade.

In figure 11(b), we focus on the period 2005–2008 from figure 9(c). Again, large spikes indicate significant changes in the community configuration over a single time step. Figure 11(b) shows that a significant reorganization of community structure occurred on 15 August 2007. (In figure 10(c), we showed the observed communities before and after this date.) This community reorganization is a result of massive unwinding of the carry trade. Figure 11(a) shows that, leading up to 15 August 2007, there was some unwinding of the carry trade, so the initial configuration includes a community containing exchange rates of the form AUD/YYY, NZD/YYY, and XXX/JPY (which all involve one of the key carry-trade currencies). In figure 11(a), it is also clear that there is a sharp increase in carry-trade unwinding following this date. The right network partition in figure 10(c) highlights this increase as the carry-trade community increases in size by incorporating other XXX/JPY rates as well as some XXX/CHF and XXX/USD rates. The presence of a large number of exchange rates involving one of the key carry-trade currencies in a single community demonstrates the significance of the trade over this period. Importantly, some of the exchange rates included in the carry-trade community are also somewhat surprising and provide insights into the range of currencies used in the carry trade over this period.

The above discussion illustrates that one can identify major changes in the correlation structure of the FX market by finding large values of $\hat{V}$ between time steps. Having identified significant changes, one can then gain a better understanding of the nature of such changes and potentially also gain insights into trading changes taking place in the market by investigating the adjustments in specific communities. We have discussed three examples in which the observed changes are obviously attributable to a major FX market event. However, there are also several time steps in which significant community reorganizations occur for which the cause is much less obvious, and the investigation of dynamic communities might help shed light on concomitant market changes.

9. Visualizing changes in exchange-rate roles

In this final section, we investigate changes in the relationships between specific exchange rates and their communities. We begin by defining within-community z-scores, which directly compare the relative importances of different nodes to their community (Guimerà and Amaral 2005). We describe the roles of individual nodes at each time step using the within-community projected community centrality z-score z^y and the 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

$$z_i^y = \frac{y_i - \bar{y}_{c_i}}{\sigma_{c_i}^y}, \tag{23}$$

where $\bar{y}_{c_i}$ is the mean of y_i over all nodes in c_i and $\sigma_{c_i}^y$ is the standard deviation of y_i in c_i. The quantity z_i^y measures how strongly connected node i is to its community compared with other nodes in the same community. Similarly, if node i has betweenness centrality b_i, then

$$z_i^b = \frac{b_i - \bar{b}_{c_i}}{\sigma_{c_i}^b}, \tag{24}$$

where $\bar{b}_{c_i}$ is the average of b_i over all nodes in c_i and $\sigma_{c_i}^b$ is the standard deviation of b_i in c_i. The quantity z_i^b indicates the importance of node i to the spread of information compared with other nodes in its 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

†For a within-community z-score to be well defined, a node must belong to a community containing two or more nodes.

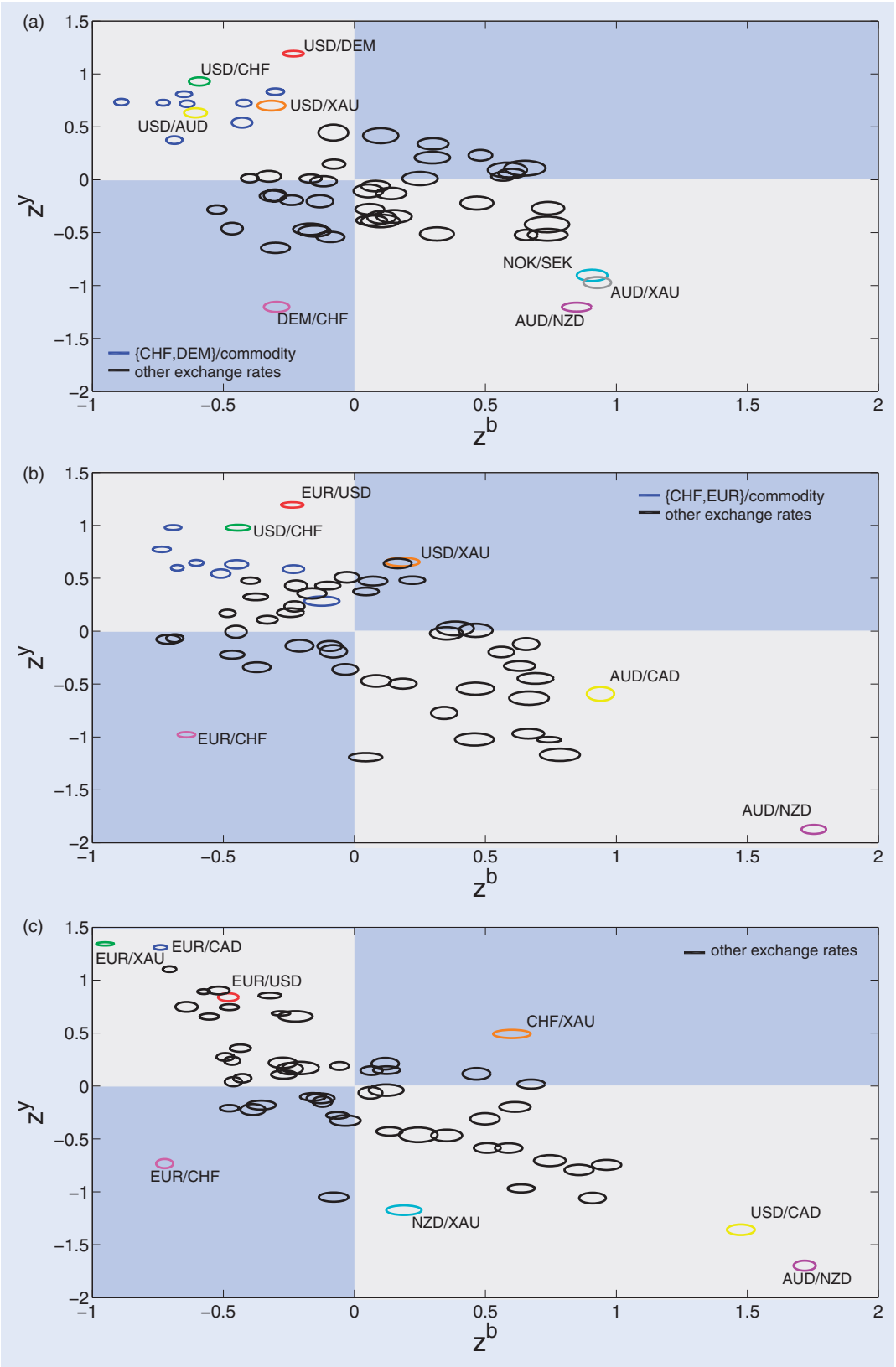

Figure 12. Node positions in the (z^b, z^y) plane averaged over all time steps for the periods (a) 1991–1998, (b) 1999–2003, and (c) 2005–2008. The radii of each elliptical marker equal the standard deviations in the z-scores for the corresponding node, and they are scaled by a factor of 1/15 for visual clarity.

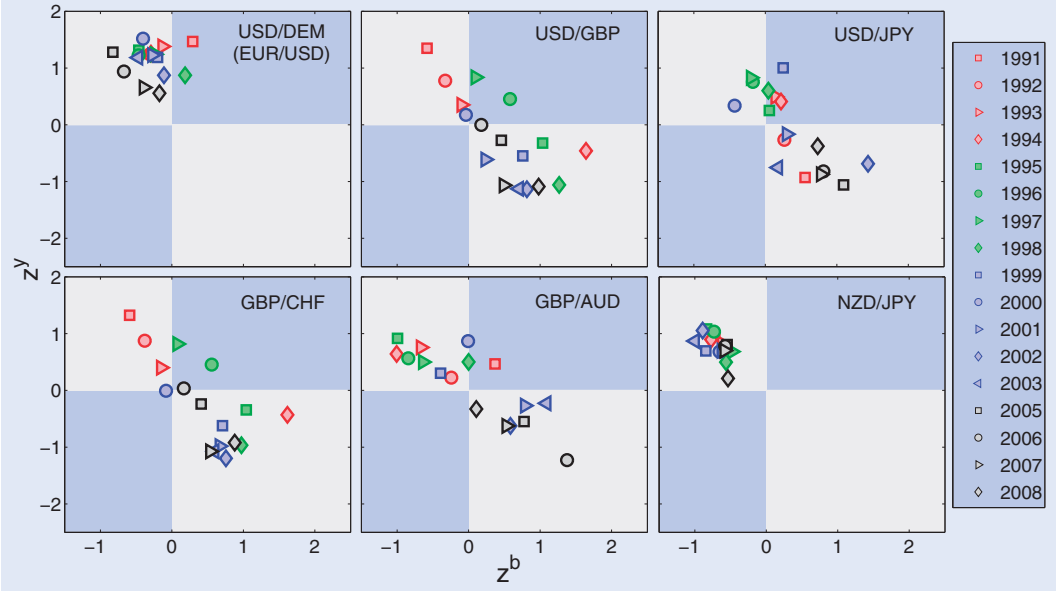

Figure 13. Annual node-role evolutions in the (z^b, z^y) plane for the full period (1991–2008).

and provide information that cannot be gained by simply considering individual exchange-rate time series.

We remark that our methods are robust with respect to the choice of measures used to construct the parameter plane: we obtain similar results using other notions, such as dynamical importance (Restrepo *et al.* 2006) instead of betweenness centrality and the within-community strength z-score (Guimerà and Amaral 2005) instead of projected community centrality. Obviously, one can also do analogous computations using other nodes properties.

9.1. Mean roles

In figure 12, we show the mean position of each exchange rate over the three periods and highlight some rates that play particularly prominent roles. For example, USD/DEM (and then EUR/USD after the introduction of the euro) regularly had the strongest connection to its community for 1991–2003, but EUR/XAU was more strongly connected to its community for 2005–2008. The importance of USD/DEM and EUR/USD is unsurprising, given that these rates had the highest daily trading volume (Galati *et al.* 2002). This provides a reality check that our methods uncover useful information about the roles of minor exchange rates. Other exchange rates, such as NOK/SEK and AUD/NZD, were less influential within their communities but were very important for the transfer of information around the FX network.

The (z^b, z^y) plots also highlight exchange rates that play similar roles in the FX market. For example, exchange rates formed from one of the major European currencies—DEM or CHF—and one of the commodity currencies—AUD, CAD, or NZD (or the commodity XAU)—are located close together in the upper left quadrant of the (z^b, z^y) plane for 1991–2003. This prominent similarity is not present for 2005–2008.

9.2. Annual roles

We can also gain insights into the temporal dynamics of exchange-rate roles by examining changes in the positions of the rates in the (z^b, z^y) plane over different time periods. Changes in a node's position in the (z^b, z^y) plane reflect changes in the membership of a node's community as well as changes in b and y. In figure 13, we show six example annual role evolutions. We determine the annual roles by averaging z^y and z^b over all time steps in each year. We see, for example, that the NZD/JPY exchange rate maintained a consistently influential role within its community over the full period. Similarly, the EUR/USD rate maintained the same influential role played by the USD/DEM rate before the introduction of the euro.

Other rates changed roles during the studied period. The GBP/USD and GBP/CHF exchange rates evolved in a similar manner, as they changed from being strongly influential within their communities before 1994 to being less influential within their communities but more important for information transfer after 1994. The role of both GBP/AUD and USD/JPY varied significantly over the period 1991–2008. From 2001 onwards, GBP/AUD became less influential within its community but more important for information transfer. Interestingly, the USD/JPY rate had its highest within-community influence in the late 1990s during a period of Japanese economic turmoil. One can construct similar plots to study role changes of other exchange rates. These role plots provide a useful tool for visualizing changes in exchange-rate correlations.

9.3. Quarterly roles

We also investigate higher-frequency changes in exchange-rate market roles using shorter time intervals. In figure 14, we show quarterly role changes over the period 1995–1998 for six exchange rates—including USD/DEM

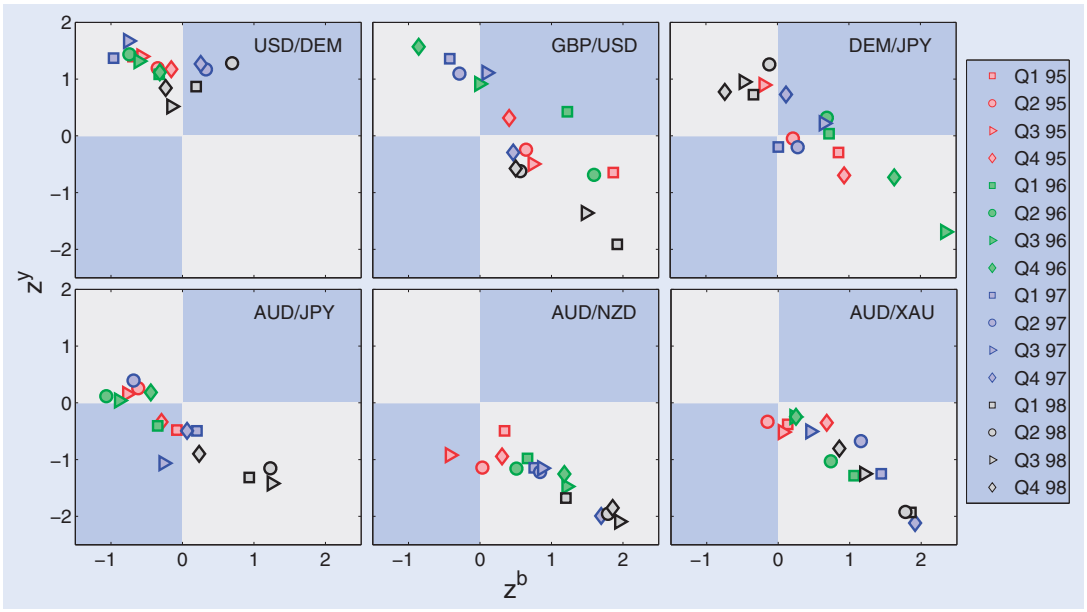

Figure 14. Quarterly node-role evolutions in the (z^b, z^y) plane for the period 1995–1998.

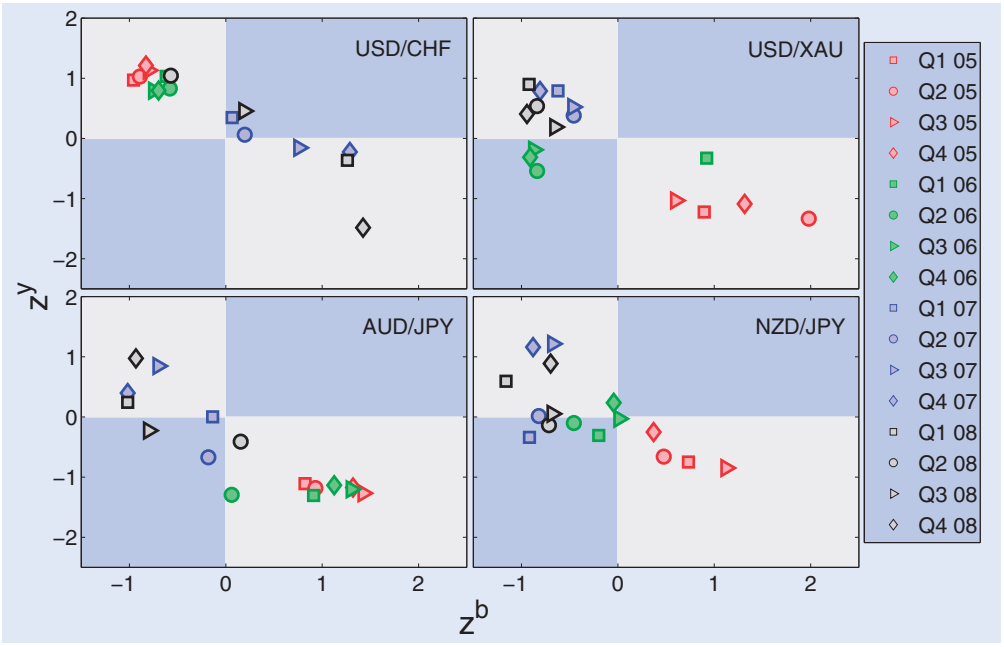

Figure 15. Quarterly node-role evolutions in the (z^b, z^y) plane for the period 2005–2008.

and GBP/USD, for which we also show the annual changes in figure 13. USD/DEM plays a relatively influential role within its community with both quarterly and annual time aggregations, whereas there is significant variation in the role of GBP/USD with both quarterly and annual time aggregations. We also show other examples for which we did not show annual changes. The role of DEM/JPY varied considerably over the period 1995–1998: in particular, it was an important information carrier for the last two quarters in 1996, whereas it was influential within its community throughout 1998. In contrast, AUD/JPY moves from being unimportant for information transfer to being an information carrier during 1998. Additionally, AUD/NZD and AUD/XAU were both always information carriers, and AUD/NZD was particularly important for information transfer during 1998.

Finally, we consider some examples of quarterly role evolutions for the period 2005–2008, which we discussed in our recent short paper (Fenn *et al.* 2009). Figure 15 shows quarterly role changes for four exchange rates during the period 2005–2008. The USD/XAU rate provides an interesting example due to the persistence of its community over this period. For 2005–2008, the USD/XAU node shifted from being an important information carrier within the XAU community to being more influential in this community. This period of higher influence coincides closely with the period of financial

turmoil during 2007–2008. CHF is widely regarded as a 'safe haven' currency (Ranaldo and Söderlind 2010), so one might expect USD/CHF to behave in a similar manner to USD/XAU. However, CHF is also a key carry-trade currency. Because CHF is used both as a safe haven and as a carry-trade currency, the USD/CHF node 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 2007–2008 credit crisis. Over the same period, the AUD/JPY and NZD/JPY exchange rates change from being important for information transfer to being influential within their communities. The AUD/JPY and NZD/JPY rates were most influential within their respective communities during Q3 and Q4 2007 and during Q1 and Q4 2008. Figure 11(b) shows that there was significant carry-trade activity in all of these periods, so it is unsurprising that two exchange rates that are widely used for this trade should increase in importance. This provides a further demonstration that the positions of exchange rates in the (z^b, z^y) parameter plane can provide important insights into the roles of exchange rates in the FX market.

10. Conclusions

To conclude, we have demonstrated that a network analysis of the FX market is useful for visualizing and providing insights into the correlation structure of the market. In particular, we investigated community structure at different times to provide insights into the clustering dynamics of exchange-rate time series. We focused on a node-centric community analysis that allows one to follow the temporal dynamics of functional roles of exchange rates within the market. We thereby demonstrate that there is a relationship between an exchange rate's functional role and its position within its community. We indicated that exchange rates that are located on the edges of communities are important for information transfer in the FX market, whereas exchange rates that are located in the centre of a community have a strong influence on other rates within that community. We also demonstrated that the community structure of the FX market can be used to determine which exchange rates dominate the market at each time and identified exchange rates that experienced significant changes in market role.

Our analysis successfully uncovered significant structural changes that occurred in the FX market, including ones that resulted from major market events that did not impact the studied exchange rates directly. We also demonstrated that community-structure reorganizations at specific times can provide insights into changes in trading behaviour and highlighted the prevalence of the carry trade during the 2007–2008 credit and liquidity crisis. Although we focused on networks of exchange rates, our methodology should be similarly insightful for multivariate time series of other asset classes.

Acknowledgements

We thank M.D. Gould, S.D. Howison, A.C.F. Lewis, M.A. Little, J.M. McPherson, J. Moody, J.-P. Onnela, S. Reid, and S.J. Roberts and his group for discussions and code. We acknowledge HSBC bank for providing the data. DJF acknowledges a CASE award from the EPSRC and HSBC bank. NSJ acknowledges support from the BBSRC and EPSRC and the grants EP/I005986/1, EP/H046917/1, and EP/I005765/1. MAP acknowledges a research award (#220020177) from the James S. McDonnell Foundation. PJM's contribution was funded by the NSF (DMS-0645369).

References

Adamcsek, B., Palla, G., Farkas, I.J., Derényi, I. and Vicsek, T., CFinder: locating cliques and overlapping modules in biological networks. *Bioinformatics*, 2006, **22**, 1021–1023.

Albert, R. and Barabási, A.L., Statistical mechanics of complex networks. *Rev. Mod. Phys*., 2002, **74**, 47–97.

Amaral, L.A.N. and Ottino, J.M., Complex networks. *Eur. Phys. J. B*, 2004, **38**, 147–162.

Arenas, A., Fernández, A. and Gómez, S., Analysis of the structure of complex networks at different resolution levels. *New J. Phys*., 2008, **10**, 053039, doi: 10.1088/1367-2630/10/5/053039.

Asur, S., Parthasarathy, S. and Ucar, D., An event-based framework for characterizing the evolutionary behavior of interaction graphs . In *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '07)*, 12–15 August 2007, San José, CA, pp. 913–921, 2007 (ACM: New York).

Barrat, A., Barthélemy, M. and Vespignani, A., *Dynamical Processes on Complex Networks*, 2008 (Cambridge University Press: Cambridge, UK).

Berger-Wolf, T.Y. and Saia, J., A framework for analysis of dynamic social networks. In *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD* '06), 20–23 August 2006, Philadelphia, PA. pp. 523–528, 2006 (ACM: New York).

Bernaschi, M., Grilli, L. and Vergni, D., Statistical analysis of fixed income market. *Physica A*, 2002, **308**, 381–390.

Berry, D.A. and Lindgren, B.W., *Statistics: Theory and Methods*, 1990 (Brooks/Cole: Pacific Grove, CA).

Blondel, V.D., Guillaume, J., Lambiotte, R. and Lefebvre, E., Fast unfolding of communities in large networks. *J. Stat. Mech*., 2008, **2008**, P10008, doi:10.1088/1742-5468/2008/10/P10008.

Boccara, N., *Modeling Complex Systems*, 2003 (Springer-Verlag: New York).

Bollobas, B., *Random Graphs*, 2nd ed., 2001 (Cambridge University Press: Cambridge, UK).

Bouchaud, J.P. and Potters, M., *Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management*, 2003 (Cambridge University Press: Cambridge, UK).

Brandes, U., Delling, D., Gaertler, M., Görke, R., Hoefer, M., Nikoloski, Z. and Wagner, D., On modularity clustering. *IEEE Trans. Knowl. Data Eng*., 2008, **20**, 172–188.

Brunnermeier, M.K., Nagel, S. and Pedersen, L.H., In *Carry Trades and Currency Crashes*., edited by D. Acemoglu, K. Rogoff, and M. Woodford, Vol. 23,

p. 14473, 2008 (NBER Working Paper Series). Available online at SSRN: http://ssrn.com/abstract=1297722 (accessed 10 April 2010).
Caldarelli, G., *Scale-Free Networks*, 2007 (Oxford University Press: Oxford, UK).
Dacorogna, M.M., Gençay, R., Müller, U., Olsen, R.B. and Pictet, O.V., *An Introduction to High-Frequency Finance*, 2001 (Academic Press: San Diego, CA).
Danon, L., Díaz-Guilera, A., Duch, J. and Arenas, A., Comparing community structure identification. *J. Stat. Mech.*, 2005, **2005**, P09008, doi:10.1088/1742-5468/2005/09/P09008.
Duda, R.O., Hart, P.E. and Stork, D.G., *Pattern Classification*, 2001 (Wiley: New York).
Falkowski, T., Bartelheimer, J. and Spiliopoulou, M., Mining and visualizing the evolution of subgroups in social networks. In *Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence (WI '06)*, 18–22 December 2006, Hong Kong, pp. 52–58, 2006 (IEEE Computer Society: Washington, DC).
Farkas, I., Ábel, D., Palla, G. and Vicsek, T., Weighted network modules. *New J. Phys.*, 2007, **9**, 180–198.
Fenn, D.J., Network communities and the foreign exchange market. PhD thesis, 2011 (University of Oxford: UK). Available at http://people.maths.ox.ac.uk/porterm/research/DJF_thesis_FINAL.pdf (accessed 18 April 2012).
Fenn, D.J., Porter, M.A., McDonald, M., Williams, S., Johnson, N.F. and Jones, N.S., Dynamic communities in multichannel data: an application to the foreign exchange market during the 2007–2008 credit crisis. *Chaos*, 2009, **19**, 033119, doi: 10.1063/1.3184538.
Fortunato, S., Community detection in graphs. *Phys. Rep.*, 2010, **486**, 75–174.
Fortunato, S. and Barthelemy, M., Resolution limit in community detection. *Proc. Natl. Acad. Sci. USA*, 2007, **104**, 36–41.
Freeman, L.C., A set of measures of centrality based on betweenness. *Sociometry*, 1977, **40**, 35–41.
Galati, G., Jeanneau, S. and Widera, R., Triennial Central Bank Survey: foreign exchange and derivatives market activity in 2001. Technical report, 2002 (Bank for International Settlements).
Garas, A., Argyrakis, P. and Havlin, S., The structural role of weak and strong links in a financial market network. *Eur. Phys. J. B*, 2008, **63**, 265–271.
Girvan, M. and Newman, M.E.J., Community structure in social and biological networks. *Proc. Natl. Acad. Sci. USA.*, 2002, **99**, 7821–7826.
Good, B.H., de Montjoye, Y.A. and Clauset, A., The performance of modularity maximization in practical contexts. *Phys. Rev. E*, 2010, **81**, 046106, doi: 10.1103/PhysRevE.81.046106.
Good, P., *Permutation, Parametric, and Bootstrap Tests of Hypotheses*, 2005 (Springer-Verlag: New York).
Górski, A.Z., Drożdż, S. and Kwapień, J., Scale free effects in world currency exchange network. *Eur. Phys. J. B*, 2008, **66**, 91–96.
Guimerà, R. and Amaral, L.A.N., Functional cartography of complex metabolic networks. *Nature*, 2005, **433**, 895–900.
Guimerà, R., Sales-Pardo, M. and Amaral, L.A.N., Modularity from fluctuations in random graphs and complex networks. *Phys. Rev. E*, 2004, **70**, 025101, doi: 10.1103/PhysRevE.70.025101.
Heimo, T., Kumpula, J., Kaski, K. and Saramäki, J., Detecting modules in dense weighted networks with the Potts method. *J. Stat. Mech.*, 2008, **2008**, P08007, doi: 10.1088/1742-5468/2008/08/P08007.
Hopcroft, J., Khan, O., Kullis, B. and Selman, B., Tracking evolving communities in large linked networks. *Proc. Natl. Acad. Sci. USA*, 2004, **101**, 5249–5253.
Karrer, B., Levina, E. and Newman, M.E.J., Robustness of community structure in networks. *Phys. Rev. E*, 2008, **77**, 046119, doi: 10.1103/PhysRevE.77.046119.
Lancichinetti, A., Fortunato, S. and Kertéssz, J., Detecting the overlapping and hierarchical community structure of complex networks. *New J. Phys.*, 2009, **11**, 033015, doi: 10.1088/1367-2630/11/3/033015.
Liao, T.W., Clustering of time series data—a survey. *Pattern Recognit.*, 2005, **38**, 1857–1874.
Mantegna, R.N., Hierarchical structure in financial markets. *Eur. Phys. J. B*, 1999, **11**, 193–197.
Mantegna, R.N. and Stanley, H.E., *An Introduction to Econophysics*, 2000 (Cambridge University Press: Cambridge, UK).
Matteo, T.D., Aste, T. and Mantegna, R.N., An interest rates cluster analysis. *Physica A*, 2004, **339**, 181–188.
McDonald, M., Suleman, O., Williams, S., Howison, S. and Johnson, N.F., Detecting a currency's dominance or dependence using foreign exchange network trees. *Phys. Rev. E*, 2005, **72**, 046106, doi: 10.1103/PhysRevE.72.046106.
McDonald, M., Suleman, O., Williams, S., Howison, S. and Johnson, N.F., Impact of unexpected events, shocking news, and rumours on foreign exchange market dynamics. *Phys. Rev. E*, 2008, **77**, 046110, doi: 10.1103/PhysRevE.77.046110.
Meilă, M., Comparing clusterings—an information based distance. *J. Multivariate Anal.*, 2007, **98**, 873–895.
Mucha, P.J., Richardson, T., Macon, K., Porter, M.A. and Onnela, J.-P., Community structure in time-dependent, multiscale, and multiplex networks. *Science*, 2010, **328**, 876–878.
Newman, M.E.J., The structure and function of complex networks. *SIAM Rev.*, 2003, **45**, 167–256.
Newman, M.E.J., Detecting community structure in networks. *Eur. Phys. J. B*, 2004a, **38**, 321–330.
Newman, M.E.J., Fast algorithm for detecting community structure in networks. *Phys. Rev. E*, 2004b, **69**, 066133, doi: 10.1103/PhysRevE.69.066133.
Newman, M.E.J., Finding community structure in networks using the eigenvectors of matrices. *Phys. Rev. E*, 2006a, **74**, 036104, doi: 10.1103/PhysRevE.74.036104.
Newman, M.E.J., Modularity and community structure in networks. *Proc. Natl. Acad. Sci. USA*, 2006b, **103**, 8577–8582.
Newman, M.E.J. and Girvan, M., Finding and evaluating community structure in networks. *Phys. Rev. E*, 2004, **69**, 026113, doi: 10.1103/PhysRevE.69.026113.
Onnela, J.-P., Chakraborti, A., Kaski, K., Kertész, J. and Kanto, A., Dynamics of market correlations: taxonomy and portfolio analysis. *Phys. Rev. E*, 2003, **68**, 056110, doi: 10.1103/PhysRevE.68.056110.
Onnela, J.-P., Kaski, K. and Kertész, J., Clustering and information in correlation based financial networks. *Eur. Phys. J. B*, 2004, **38**, 353–362.
Onnela, J.-P., Saramäki, J., Hyvönen, J., Szabó, G., Lazer, D., Kaski, K., Kertész, J. and Barabási, A.L., Structure and tie strengths in mobile communication networks. *Proc. Natl. Acad. Sci. USA*, 2007, **104**, 7332–7336.
Palla, G., Barabasi, A.L. and Vicsek, T., Quantifying social group evolution. *Nature*, 2007, **446**, 664–667.
Porter, M.A., Mucha, P.J., Newman, M.E.J. and Warmbrand, C.M., A network analysis of committees in the United States House of Representatives. *Proc. Natl. Acad. Sci. USA*, 2005, **102**, 7057–7062.
Porter, M.A., Onnela, J.-P. and Mucha, P.J., Communities in Networks. *Not. Am. Math. Soc.*, 2009, **56**, 1082–1097, 1164–1166.
Ranaldo, A. and Söderlind, P., Safe haven currencies. *Rev. Finance*, 2010, **14**, 385–407.
Reichardt, J. and Bornholdt, S., Statistical mechanics of community detection. *Phys. Rev. E*, 2006, **74**, 016110, doi: 10.1103/PhysRevE.74.016110.
Restrepo, J.G., Ott, E. and Hunt, B.R., Characterizing the dynamical importance of network nodes and links. *Phys. Rev. Lett.*, 2006, **97**, 094102, doi: 10.1103/PhysRevLett.97.094102.

Schelter, B., Winterhalder, M. and Timmer, J., *Handbook of Time Series Analysis*, 2006 (Wiley-VCH: Weinheim, Germany).

Shalizi, C.R., Camperi, M.F. and Klinkner, K.L., Discovering functional communities in dynamical networks. In *Statistical Network Analysis: Models, Issues, and New Directions*, edited by E.M. Airoldi, D.M. Blei, S.E. Feinberg, A. Goldenberg, E.P. Xing, and A.X. Zheng, pp. 140–157, 2007 (Springer-Verlag: New York).

Sieczka, P. and Holyst, J.A., Correlations in commodity markets. *Physica A*, 2009, **388**, 1621–1630.

Toyoda, M. and Kitsuregawa, M., Extracting evolution of web communities from a series of web archives. In *Proceedings of the 14th ACM Conference on Hypertext and Hypermedia (HYPERTEXT '03)*, 26–30 August, Nottingham, UK, pp. 28–37, 2003 (ACM: New York).

Traag, V.A. and Bruggeman, J., Community detection in networks with positive and negative links. *Phys. Rev. E*, 2009, **80**, 036115, doi: 10.1103/PhysRevE.80.036115.

Traud, A.L., Kelsic, E.D., Mucha, P.J. and Porter, M.A., Comparing community structure to characteristics in online collegiate social networks. *SIAM Rev.*, 2011, **53**, 526–543.

Valente, T.W., Coronges, K., Lakon, C. and Costenbader, E., How correlated are network centrality measures? *Connections*, 2008, **28**, 16–26.

Wasserman, S. and Faust, K., *Social Network Analysis: Methods and Applications*, 1994 (Cambridge University Press: Cambridge, UK).

Appendix A: Robustness of results—alternative computational heuristics

In the main text, we detected all communities using the Louvain locally greedy algorithm (Blondel *et al.* 2008). However, as we noted in section 3, several alternative heuristics exist. We now investigate whether the choice of heuristic has any effect on the results described in this paper.

Good *et al.* (2010) demonstrated that there are extreme near-degeneracies in the energy function, as the number of low-energy solutions can scale exponentially or faster with the number of nodes. Given this, it is unsurprising that different energy-optimization heuristics can yield very different partitions for the same network. Good *et al.* suggested that the reason for such behaviour is that different heuristics sample different regions of the energy landscape. Because of the potential sensitivity of results to the choice of heuristic, one should treat individual partitions output by particular heuristics with caution. However, one can have more confidence in the validity of partitions if different heuristics produce similar results.

In this section, we compare the results for the Louvain algorithm (Blondel *et al.* 2008) with those for a spectral algorithm (Newman 2006b) and simulated annealing (Guimerà *et al.* 2004) for the 563 networks that we constructed for the period 2005–2008.

A.1. Comparison of partition energies

We begin by comparing the energy $\mathcal{H}$ of the optimal partitions at the resolution $\gamma = 1.45$. Figure A1 shows the distribution of energies for the different algorithms and demonstrates that the Louvain algorithm and simulated annealing find lower-energy partitions than the spectral algorithm. For the remainder of this section, we will only compare the Louvain and simulated-annealing algorithms because of the higher energy of the spectral partitions.

A.2. Temporal changes in communities

First, we compare the network partitions identified by the two heuristics for each network. In figure A2, we show the distribution of the normalized variation of information between the community partitions identified using the Louvain and simulated-annealing algorithms. The two methods identify identical partitions for 19% of the networks; for 83% of the networks, the partitions differ in

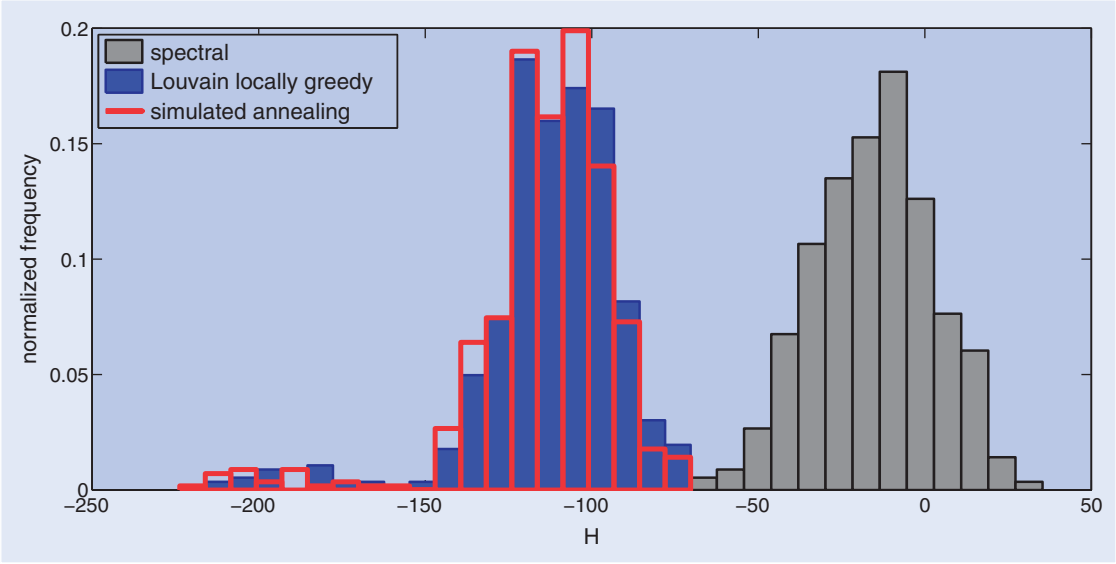

Figure A1. Distribution of the energy $\mathcal{H}$ of the optimal partition for networks over the period 2005–2008 computed using different optimization algorithms.

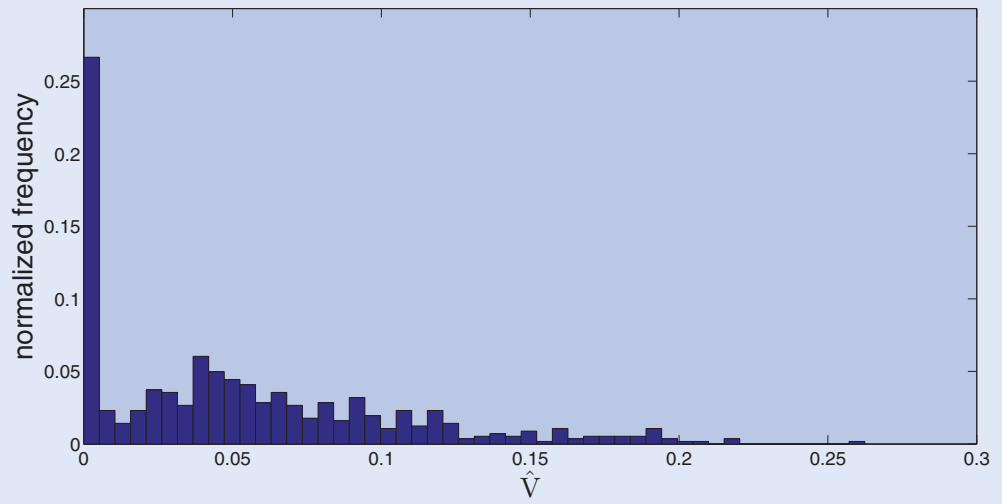

Figure A2. Distribution of the normalized variation of information between network partitions identified using the Louvain algorithm and simulated annealing for networks during the period 2005–2008.

their assignment of nodes to communities by fewer than 10 nodes. There is therefore strong agreement between the partitions obtained by the two heuristics, but there are also differences that warrant further investigation.

In section 8, we identified significant changes in community structure by comparing changes in the scaled energy Q_s (see equation 10) between consecutive time steps and by calculating the normalized variation of information between community partitions at consecutive time steps (see figure 9). The correlation between the scaled energy Q_s of the partitions obtained using the two optimization heuristics is 0.99, and the correlation between the changes in Q_s between consecutive time steps for the two heuristics is 0.93. The correlation between the normalized variation of information between partitions at consecutive time steps is 0.36. The scaled energy correlations are clearly extremely high. However, there are differences in the timings of some major reorganizations identified by the normalized variation of information. To compare the timings of major events, we identify time steps at which the normalized variation of information between consecutive partitions is more than a certain number of standard deviations larger than the mean normalized variation of information between consecutive partitions. We find that the algorithms identify 40% of 1-standard-deviation events at the same time steps and 33% of 2.5-standard-deviation events at the same time steps. The methods therefore agree reasonably well. However, the differences also suggest that one should be cautious when using normalized variation of information to identify major reorganizations in community structure.

A.3. Example community comparison

One time step at which both heuristics identify a large change in community structure is 15 August 2007 which, as described in section 8.3, was a day when there was a significant increase in carry-trade unwinding. It is worth considering the communities at this time step in detail to help assess the similarity of the results for the two heuristics. In figure A3(a), we show the communities that we identified using the Louvain algorithm (Blondel *et al.* 2008) immediately before and after 15 August 2007; in figure A3(b), we show communities that we identified using simulated annealing (Guimerà *et al.* 2004) for the same time steps. Figure A3(a) shows that, leading up to 15 August 2007, there was some unwinding of the carry trade, so the initial configuration includes a community containing exchange rates of the form AUD/YYY, NZD/YYY, and XXX/JPY (which all involve one of the key carry-trade currencies). After 15 August 2007, as the volume of carry-trade unwinding increases, this community incorporates other XXX/JPY rates as well as some XXX/CHF and XXX/USD rates. Although the communities in figure A3(b) for the simulated-annealing algorithm are not identical to those in figure A3(a), they are very similar. The main difference is that, for the simulated-annealing algorithm, there are two carry-trade communities before 15 August 2007: one community containing exchange rates of the form AUD/YYY and NZD/YYY (where we note that AUD and NZD are both carry-trade investment currencies) and another community containing exchange rates of the form XXX/CHF and XXX/JPY (where we note that CHF and JPY are both carry-trade funding currencies). After 15 August 2007, as carry-trade unwinding increases, these two communities combine and two other exchange rates also join the community. The resulting merged community is very similar to the largest community identified at the same time step using the Louvain algorithm.

Figure A3 therefore illustrates that there are only small differences in the community structures obtained from the two heuristics. In fact, as figure A2 demonstrates, the two algorithms agree in the assignment of all but about 10 nodes approximately 80% of the time. Importantly, figure A3 highlights that the two heuristics reveal similar changes in the FX market even when there are differences in the precise community configurations that they identify.

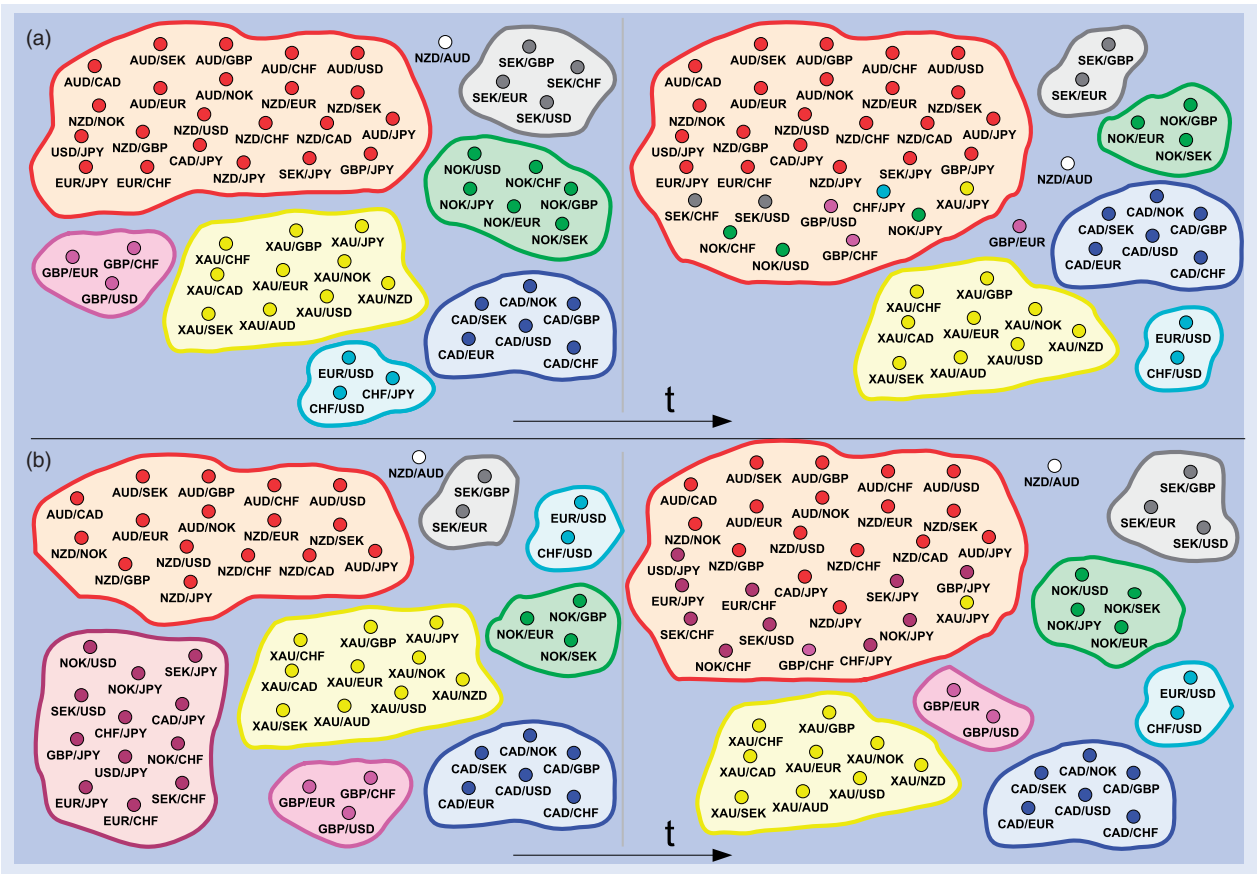

Figure A3. Comparison of changes in community structure in one half of the FX market network over the same period for different optimization heuristics. We show a schematic of the communities for the period following 15 August 2007, when there was significant unwinding of the carry trade during the 2007–2008 credit and liquidity crisis. We identified communities using (a) the Louvain locally greedy algorithm (Blondel *et al.* 2008) and (b) a simulated-annealing algorithm (Guimerà *et al.* 2004). The node colours after the community reorganization correspond to the communities before the change. If the parent community of a community after the reorganization is obvious, we draw it using the same colour as its parent. The nodes drawn as triangles resided in the opposite half of the network before the community-structure reorganization.

A.4. Node-role comparison

As an additional comparison, we investigate the effect of different computational heuristics on exchange-rate roles (see section 9). In figure A4, we compare quarterly role evolutions over the period 2005–2008 for the exchanges rates shown in figure 15. Although there are slight differences in the positions of the exchange rates in the (z^b, z^y) plane for some periods, we obtain the same aggregate conclusions. For example, for both heuristics, AUD/JPY is most influential within its community (high z^b) during Q3 and Q4 2007 and during Q1 and Q4 2008, and it is less influential (but more important for information transfer) during 2005 and 2006.

The positions in the (z^b, z^y) plane are similarly close for all of the other exchange rates. We quantify the differences in the positions for the two heuristics by calculating the mean and standard deviation of the change in position over all exchange rates and over all time periods. More precisely, we average the change in position of every node in the (z^b, z^y) plane over every quarter. The mean change in position in both the z^b and z^y directions is less than 10^{-4}; the standard deviations are 0.15 and 0.17, respectively. However, because the changes in position are likely to cancel out (i.e., an increase in z^b for one exchange rate is likely to be offset by a decrease in z^b for another exchange rate), it is more informative to calculate the mean and standard deviation of the absolute changes in position in the z^b and z^y directions. In the z^b direction, the mean absolute change in position is 0.08, and the standard deviation is 0.13; in the z^y direction, the mean absolute change is 0.09, and the standard deviation is 0.15. The mean differences in positions in the (z^b, z^y) plane are therefore very small for the two heuristics and, as figure A4 demonstrates, both algorithms uncover the same role changes in the FX market for the different exchange rates.

Finally, we also checked the relationships shown in figure 8 for community centrality versus community size, community alignment versus betweenness centrality, and community autocorrelation versus projected community centrality. Using simulated annealing, we find the same trends that we uncovered with the Louvain algorithm.

The results of this section demonstrate that, although there are differences in the communities identified using different optimization heuristics, the aggregate conclusions are the same. We identify the same changes taking

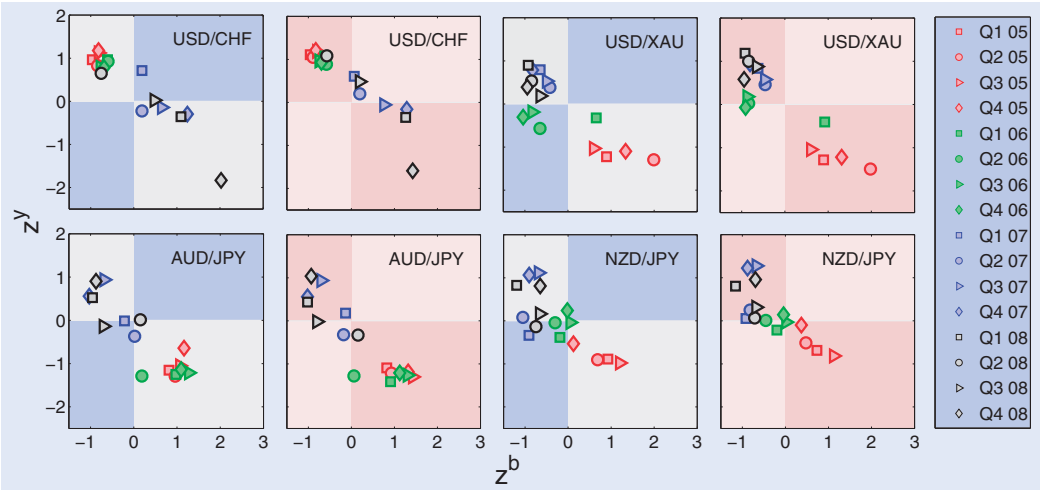

Figure A4. Comparison of the quarterly node-role evolutions in the (z^b, z^y) plane for the period 2005–2008 for communities identified using the locally greedy Louvain algorithm (Blondel *et al*. 2008) and simulated annealing (Guimerà *et al*. 2004). The plots with white and blue shading show results for the Louvain algorithm and the plots with pink shading show results for simulated annealing.

place in the FX market whether we use the Louvain algorithm or simulated annealing to minimize energy. The fact that we obtain very similar results using different optimization techniques, despite these techniques sampling different regions of the energy landscape, gives confidence that the effects that we uncover are genuine and that our results are robust. In practice, the Louvain algorithm is preferable to simulated annealing because of the computational cost of the latter. For example, on the machine that we used to perform the computations, the Louvain algorithm converged on an optimal community partition for all 563 networks over the period 2005–2008 in 5 minutes and 24 seconds. For the same networks, the simulated-annealing algorithm took about 36 hours.