Financial Econometrics Research Centre

WORKING PAPERS SERIES

WP05-03

Emergence of large cliques in random scale-free networks

Ginwestra Bianconi and Matteo Marsili

Emergence of large cliques in random scale-free networks

Ginestra Bianconi[1] and Matteo Marsili[1]
[1] *The Abdus Salam International Center for Theoretical Physics, Strada Costiera 11, 34014 Trieste, Italy*

In a network cliques are fully connected subgraphs that reveal which are the tight communities present in it. Cliques of size $c > 3$ are present in random Erdös and Renyi graphs only in the limit of diverging average connectivity. Starting from the finding that real scale free graphs have large cliques, we study the clique number in uncorrelated scale-free networks finding both upper and lower bounds. Interesting we find that in scale-free networks large cliques appear also when the average degree is finite, i.e. even for networks with power-law degree distribution exponents $\gamma \in (2, 3)$. Moreover as long as $\gamma < 3$ scale-free networks have a maximal clique which diverges with the system size.

PACS numbers: : 89.75.Hc, 89.75.Da, 89.75.Fb

Scale-free graphs have been recently found to encode the complex structure of many different systems ranging from the Internet to the protein interaction networks of various organisms [1, 2, 3]. This topology is clearly well distinguished from the Erdös and Renyi (ER) [4] random graphs in which every couple of nodes have the same probability p to be linked. In fact while scale-free graphs have a power-law degree distribution $P(k) \sim k^{-\gamma}$ and a diverging second moment $\langle k^2 \rangle$ when $\gamma < 3$, ER graphs have a Poisson degree distribution and consequently finite fluctuations of the nodes degrees. The degree distribution strongly affects the statistical properties of processes defined on the graph. For example, percolation and epidemic spreading which have very different phenomenology when defined on a ER graph or on a scale-free graphs [5, 6].

The occurrence of a skewed degree distribution has also striking consequences regarding the frequency of particular subgraphs present in the network. For example, ER graphs with finite average connectivity have a finite number of finite loops [4, 7]. On the contrary scale-free graphs have a number of finite loops which increases with the number N of vertices, provided that $\gamma \leq 3$ [8, 9]. The abundance of some subgraphs of small size – the so-called *motifs* – in biological networks has been shown to be related to important functional properties selected by evolution [10, 11, 12]. Among subgraphs, cliques play an important role. A clique of size c is a complete subgraph of c nodes, i.e. a subset of c nodes each of which is linked to any other. The maximal size $c_{\max}$ of a clique in a graph is called the clique number. Finding the clique number of a generic network is an NP-complete problem [13], even though it is relatively easy to find upper (c_+) and lower (c_-) bounds [14]. The clique number also provides a lower bound for the chromatic number of a graph, i.e. the minimal number of colors needed to color the graph [15]. Finally, cliques and overlapping succession of cliques have been recently used to characterize the community structure of networks [16, 17].

In ER graphs it is very easy to show that cliques of size $3 < c \ll N$ appear in the graph only when the average degree diverges as $\langle k \rangle \sim N^{\frac{c-3}{c-1}}$ with N [4]. On the other hand, real scale free networks, such as the Internet at the autonomous system level, contain cliques of size much larger than $c = 3$. For example, Fig. 1 reports upper and lower bounds c_+, c_- [14] for the size of the maximal clique of the Internet and protein interaction networks of c.elegans and yeast [18]. This shows that scale-free networks can have large cliques and that the clique number of the Internet graphs increase with the network size N.

Is the presence of such large cliques a peculiar property of how these networks are wired or is this a typical property of networks with such a broad distribution of degrees? This letter addresses this question and shows that scale free random networks do indeed contain cliques of size much larger than $c = 3$. We shall do this by computing the first two moments of the number $\mathcal{N}_c$ of cliques of size c in a network of N nodes. These provide upper and lower bounds for the probability $P(\mathcal{N}_c > 0)$ of finding cliques of size c in a network through the inequalities [4]

$$\frac{\langle \mathcal{N}_c \rangle^2}{\langle \mathcal{N}_c^2 \rangle} \leq P(\mathcal{N}_c > 0) \leq \langle \mathcal{N}_c \rangle. \tag{1}$$

Here and in the following the notation $\langle \ldots \rangle$ will be used for statistical averages. Eq. (1) in turn provide upper and lower bounds for the clique number $\underline{c} \leq c_{\max} \leq \overline{c}$: Indeed if $\langle \mathcal{N}_c \rangle \to 0$ for $c > \overline{c}$ as $N \to \infty$, we can conclude that no clique of size larger than $\overline{c}$ can be found. Likewise if for $c = \underline{c}$ the ratio $\langle \mathcal{N}_c \rangle^2 / \langle \mathcal{N}_c^2 \rangle$ stays finite, then cliques of size $c \leq \underline{c}$ can be found in the network with at least a finite probability. The results indicate that the finding in Fig. 1 are expected, given the scale free nature of these graphs. Our predictions are summarized in Table I. We find that the ER result $c_{\max} = 3$ extends to random scale free networks with $\gamma > 3$ whereas for $\gamma < 3$ the clique number $c_{\max}$ diverges with the network size N in a way which is extremely sensitive of the degree distribution of mostly connected nodes, i.e. to the precise definition of the cutoff.

The results of Table I are derived for the hidden variable ensemble proposed in Ref. [19, 20], where the link probability p between two nodes is replaced by a function $r(q_i, q_j)$ which depends on the fitness q_i and q_j of

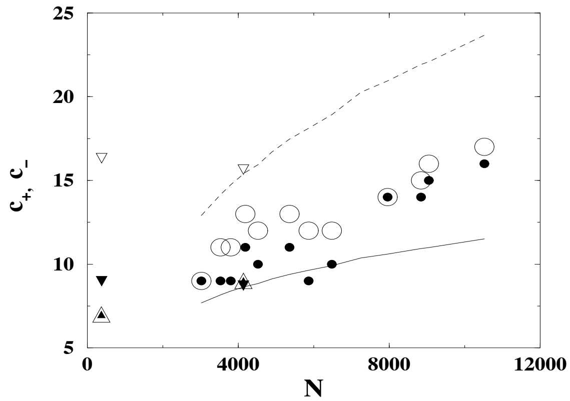

FIG. 1: The lower bound c_- (filled symbols) and the upper bound c_+ (empty symbols) of the clique number of the Internet graphs(circles) and the protein interaction networks of e.coli and yeast (triangles) [18] are shown as a function of the network size N. The lines (null hypothesis on Internet data) and the triangles pointing down (null hypothesis on protein interaction networks) indicates the upper bound (dashed line and empty symbols) and the lower bound (solid line and filled symbols) computed from Eq. (9) for random graphs constructed with the same properties of the considered real graphs.

	$\epsilon = 0$	$\epsilon \neq 0$
$\gamma > 3$	$c_{max} = 3$	
$2 < \gamma < 3$	$\underline{c} \leq c_{max} \leq \bar{c}$ $\bar{c} \simeq \sqrt{b} N^{\frac{3-\gamma}{4}}$ $\underline{c} \simeq \alpha \bar{c}^{2/3}$	$\underline{c} \leq c_{max} \leq \bar{c}$ $\bar{c} \simeq \frac{3-\gamma}{2} \frac{\log(N)}{\lvert\log(1-\epsilon)\rvert}$ $\underline{c} = (1-\alpha)\bar{c}$
$1 < \gamma \leq 2$	$\underline{c} \leq c_{max} \leq \bar{c}$ $\bar{c} \simeq \sqrt{b'} N^{\frac{1}{2\gamma}}$ $\underline{c} \simeq \alpha \bar{c}^{2/3}$	$\underline{c} \leq c_{max} \leq \bar{c}$ $\bar{c} \simeq \frac{1}{\gamma} \frac{\log(N)}{\lvert\log(1-\epsilon)\rvert}$ $\underline{c} = (1-\alpha)\bar{c}$

TABLE I: Scaling of the theoretically estimated upper and lower bound of the clique number of random scale-free networks with different exponents γ of the degree distribution. The precise definitions of $\bar{c}$ and $\underline{c}$ together with the expression for the constants b, b' are given in the text.

the end nodes i and j. Apart from its close relation with the ER ensemble, this choice is also convenient because it allows for a simple generalization of the results to networks with a correlated degree distribution [23]. Quite similar results can be derived for the Molloy-Reed ensemble [21] with the same approach (provided a cutoff is chosen appropriately to avoid double links among mostly connected nodes) . Other ensembles, such as that of Ref. [22] instead implicitly introduce a degree correlations for highly connected nodes and therefore require a different approach [23]. Given the extreme sensitivity of the clique number on details of the cutoff of the degree distribution, we also expect quite different results.

Hidden variable network ensemble As in Ref. [19] we generate a realization of a scale-free networks by the following procedure: *i)* assign to each node i of the graph a hidden continuous variable q_i distributed according a $\rho(q)$ distribution. Then *ii)* each pair of nodes with hidden variables q, q' are linked with probability $r(q, q')$. For random scale-free networks with uncorrelated degree distribution, we take $\rho(q) = \rho_0 q^{-\gamma}$ for $q \in [m, Q]$ and

$$r(q, q') = \frac{qq'}{\langle q \rangle N}. \tag{2}$$

The average degree $\langle k \rangle = \langle q \rangle$ is equal to the average fitness, and it diverges as $N \to \infty$ for $\gamma < 2$. Likewise, the degree k_i of node i follows a Poisson distribution with average q_i. Notice that a cutoff is needed in $\rho(q)$ to keep the linking probability $r(q, q')$ smaller than one. In particular, we will take require

$$Q = (1-\epsilon)\sqrt{\langle q \rangle N} \tag{3}$$

so that $r(Q, Q) = 1 - \epsilon$. For $\gamma > 3$, values of $q_i \approx Q$ will never occur, as the maximal $q_i \approx N^{1/(\gamma-1)} \ll Q$. We shall see that this is immaterial for the clique number, however. Instead, for $\gamma < 2$, $\langle q \rangle$ diverges with the cutoff, and hence $Q \sim N^{1/\gamma}$.

Average number of cliques. A clique of size c is a set of c distinct nodes $\mathcal{C} = \{i_1, \dots, i_c\}$, each one connected with all the others. For each choice of the nodes, the probability that they are connected in a clique is

$$\prod_{i \neq j \in \mathcal{C}} r(q_i, q_j) = \prod_{i \in \mathcal{C}} \left(\frac{q_i}{\sqrt{\langle q \rangle N}} \right)^{c-1} \tag{4}$$

where we used Eq. (2). Fixing a small fitness interval Δq, let $n(q)$ be the number of nodes $i \in \mathcal{C}$ with fitness $q_i \in (q, q + \Delta q)$. The number of ways in which we can pick c nodes in the network with $n(q)$ nodes with fitness q can be expressed by combinatorial factors. Hence, with the shorthand $\mathcal{Q} = q/\sqrt{\langle q \rangle N}$,

$$\langle \mathcal{N}_c \rangle = \sum_{\{n(q)\}}' \prod_q \begin{pmatrix} N(q) \\ n(q) \end{pmatrix} \mathcal{Q}^{(c-1)n(q)} \tag{5}$$

where the sum is extended to all the sequences $\{n(q)\}$ satisfying $\sum_q n(q) = c$. Introducing such constraint by a delta function, we can perform the resulting integral by saddle point method, i.e.

$$\langle \mathcal{N}_c \rangle = \int_{-\pi}^{\pi} \frac{d\omega}{2\pi} e^{Nf(i\omega)} \simeq \frac{e^{Nf(y^*)}}{\sqrt{2\pi N |f''(y^*)|}} \tag{6}$$

where $f(y) = \frac{c}{N} y + \left\langle \log\left[1 + \mathcal{Q}^{c-1} e^{-y}\right] \right\rangle$, and we have taken the limit $\Delta q \to 0$. In Eq. (6) y^* is fixed by the saddle point condition

$$\frac{c}{N} = \left\langle \frac{\mathcal{Q}^{c-1}e^{-y^*}}{1+\mathcal{Q}^{c-1}e^{-y^*}} \right\rangle. \quad (7)$$

We present here an asymptotic estimate of $\langle \mathcal{N}_c \rangle$. Slightly more refined arguments, which do not add much to the understanding given here, can be used to derive an upper bound [23]. In the limit $N \to \infty$, the left hand side of Eq. (7) is small, hence to a good approximation $c \approx N\langle \mathcal{Q}^{c-1}\rangle e^{-y^*}$ [25]. Inserting this in Eq. (6) we find

$$\langle \mathcal{N}_c \rangle \approx \left(\frac{Ne\langle \mathcal{Q}^{c-1}\rangle}{c} \right)^c \sqrt{\frac{2\pi}{c}}. \quad (8)$$

Therefore, in order to have $\langle \mathcal{N}_c \rangle \to 0$ it is sufficient to take $c > \bar{c}$, where $\bar{c}$ is the solution of

$$Ne\langle \mathcal{Q}^{c-1}\rangle = c. \quad (9)$$

We consider now separately the case of scale-free networks with different exponents γ of the degree distribution.

- Networks with $\gamma > 3$
  Eq. (9) has no solution for $c > \gamma$. Indeed $N\langle \mathcal{Q}^{c-1}\rangle \sim N^{(3-\gamma)/2} \to 0$ in this range. For $c < \gamma$, the integral in $\langle \mathcal{Q}^{c-1}\rangle$ is no longer dominated by the upper cutoff, and it is hence finite. Therefore $N\langle \mathcal{Q}^{c-1}\rangle \sim N^{(3-c)/2}$ which implies that $\bar{c} = 3$. It is easy to see that this conclusion holds also if we take the natural cutoff $Q = aN^{\frac{1}{\gamma-1}}$.

- Network with $2 < \gamma < 3$
  Using Eq. (3), Eq. (9) becomes

  $$\frac{\bar{c}(\bar{c}-\gamma)}{(1-\epsilon)^{\bar{c}-\gamma}} \simeq bN^{(3-\gamma)/2} \quad (10)$$

  for $b = (\gamma-1)m^{(\gamma-1)}e\langle q\rangle^{(1-\gamma)/2}$. The solution depends crucially on whether $\epsilon = 0$ or not. In the former case $\bar{c} \sim N^{(3-\gamma)/4}$ increases as a power law of the system size, whereas for $\epsilon > 0$ it increases only as $\log N/\log(1-\epsilon)$, as detailed in Table I.

- Network with $1 < \gamma < 2$
  Taking into account the divergence of $\langle q\rangle$ and $Q \sim N^{1/\gamma}$, Eq. (9) becomes

  $$\frac{\bar{c}(\bar{c}-\gamma)}{(1-\epsilon)^{\bar{c}-\gamma}} \simeq b'N^{1/\gamma} \quad (11)$$

  with $b' = \{(\gamma-1)[m(2-\gamma)]^{(\gamma-1)}\}^{1/\gamma}$. Again, for $\epsilon = 0$ and $\epsilon > 0$ we find different results, $\bar{c} \sim N^{1/(2\gamma)}$ and $\bar{c} \sim \log N/\log(1-\epsilon)$ respectively (see Table I).

Second moment of the average number of cliques. When computing the average number of some particular subgraphs in a random network ensemble the result might be dominated be extremely rare graphs with an anomalously large number of such subgraphs. In this cases, the average number of a subgraph does not provide a reliable indication of its value. In order to have more insight on the characteristics of typical networks we use the classical relation Eq. (1) of probability theory [4] which provides a lower bound for the probability that a typical graph contains at least one clique of size c. This requires us to compute the second moment $\langle \mathcal{N}_c^2 \rangle$ of the number of cliques of size c in the random graph ensemble. In order to do this calculation we are going to count the average number of pairs of cliques of size c present in the graph with an overlap of $o = 0, \ldots, c$ nodes. We use the notation $\{n(q)\}$ to indicate the number of the nodes with fitness q belonging to the first clique, $\{n_o(q)\}$ to indicate the number of nodes belonging to the overlap and $\{n'(q)\}$ to indicate the number of nodes belonging to the second clique but not to the overlap. We consider only sequences $\{n(q)\}, \{n'(q)\}, \{n_o(q)\}$ which satisfy $\sum_q n(q) = c$, $\sum_q n_o(q) = o$ and $\sum_q n'(q) = c - o$. With these conditions, following the same steps as for $\langle \mathcal{N}_c \rangle$ we get

$$\langle \mathcal{N}_c^2 \rangle = \sum_{o=0}^{c} \int dy \int dy^o \int dy' e^{N\langle f(y,y',y^o,\mathcal{Q})\rangle} \quad (12)$$

where

$$f(y,y',y^o,\mathcal{Q}) = \tfrac{1}{N}[yc + y'(c-o) + y^o o] + \\ + \log\left[1 + \left(e^{-y'} + e^{-y}\right)\mathcal{Q}^{c-1} + e^{-(y+y^o)}\mathcal{Q}^{2c-o-1}\right] \quad (13)$$

The evaluation of this integral by saddle point is straightforward. The key idea is that, in order to have $\langle \mathcal{N}_c^2 \rangle$ of the same order as $\langle \mathcal{N}_c \rangle^2$ one needs to require that the sum is dominated by configurations with non-overlapping cliques ($o \sim 0$). Using the estimate of $\langle \mathcal{N}_c \rangle$ derived above and the definition of $\bar{c}$, for $\gamma < 3$ we arrive at

$$P(\mathcal{N}_{\hat{c}} > 0) \geq \frac{\langle \mathcal{N}_c \rangle^2}{\langle \mathcal{N}_c^2 \rangle} \geq \left[1 + \tfrac{c(c-\gamma)(1-\epsilon)^{(\bar{c}-c)}e}{\bar{c}(\bar{c}-\gamma)}\right]^{-c}. \quad (14)$$

The lower bound for the clique number will depend on ϵ and $\bar{c}$.

In the case $\epsilon = 0$ lets define the clique size $\underline{c}$ satisfying

$$\frac{\underline{c}(\underline{c}-\gamma)e}{\bar{c}(\bar{c}-\gamma)} = \frac{1}{\underline{c}} \quad (15)$$

i.e. $\underline{c} \sim \bar{c}^{2/3}$. From Eq. (14) and the definition of $\underline{c}$ it follows that as $N, \bar{c} \to \infty$ the probability to have at least a clique of size $c = \underline{c}$ is finite, i.e.

$$P(\mathcal{N}_{\underline{c}} > 0) \geq \frac{1}{e}. \quad (16)$$

Instead in the case $\epsilon > 0$ for any $\alpha > 0$ the r.h.s. of Eq. (14) is very close to 1 for and clique sizes $\underline{c} = (1-\alpha)\bar{c}$ and $\bar{c} \gg 1/(\alpha\epsilon)$,i.e.

$$P(\mathcal{N}_{\underline{c}} > 0) \to 1. \quad (17)$$

This implies that for $\epsilon > 0$ the lower bound is very close to the upper bound $\underline{c} = (1-\alpha)\bar{c}$ for very large networks.

Conclusions In conclusion we have calculated upper and lower bounds for the maximal clique size c_{max} in uncorrelated scale-free network, showing that $c_{\max}$ diverges with the network size N as long as $\gamma < 3$. In particular large cliques are present in scale-free networks with $\gamma \in (2,3)$ and finite average degree. It is suggestive to put the emergence of large cliques for $\gamma < 3$ in relation with the persistence up to zero temperature of long range order in spin models defined on these graphs [24]. These results were derived within the hidden variable ensemble [19, 20], but the same method can be extended to other ensembles [21, 22] including those with a correlated degrees.

In Fig. 1 we compare the upper and lower bounds derived here for random scale-free graphs with the estimated clique number of real networks. These networks have many nodes with degree larger than that of the structural cutoff. Networks with such highly connected nodes cannot be considered as uncorrelated. The best approximation, within the class of uncorrelated networks discussed here, is provided by those with maximal cutoff ($\epsilon = 0$). The bounds of Fig. 1 have been derived from Eq. (9) and (15), assuming a random network with *i)* an exponent γ as measured from real data *ii)* the same number of nodes and links (i.e. the same average degree) and *iii)* a structural cutoff given by Eq. (3) with $\epsilon = 0$. Also notice that $\epsilon = 0$ yields the least stringent bounds.

Fig. 1 shows that generally the largest clique size $c_{\max}$ of real networks falls well within our bounds. Of course, accounting for the presence of correlations in the degree of highly connected nodes in these networks may provide more precise estimates. We saw that our estimates are very sensitive to the tails of the degree distribution and we expect it to depend also strongly on the nature of degree correlations. Preliminary results, extending the present calculation to correlated networks [22] where $r(q,q') = 1 - e^{-\alpha q q'}$ with the natural cutoff $Q \simeq N^{1/(\gamma-1)}$, indicates that the clique number can take values a factor two bigger than in real data [23]. These preliminary results underline the importance of extending this approach to correlated networks.

G. B. was partially supported by EVERGROW and by EU grant HPRN-CT-2002-00319, STIPCO.

[1] R. Albert and A.-L. Barabeási, *Rev. Mod. Phys.* **74**, 47 (2002).
[2] S. N. Dorogovtsev and J. F. F. Mendes, *Evolution of Networks* (Oxford University Press,Oxford,2003).
[3] R. Pastor-Satorras and A. Vespignani, *Evolution and Structure of the Internet* (Cambridge University Press, Cambridge, 2004).
[4] S. Janson, T. Luczak, A. Rucinski, *Random graphs* (John Wiley & Sons,2000).
[5] R. Cohen, K. Erez D. ben-Avraham and S. Havlin, *Phys. Rev. Lett.* **85** 4626 (2000).
[6] R. Pastor-Satorras and A. Vespignani, *Phys. Rev. Lett.* **86**, 3200 (2001).
[7] E. Marinari and R. Monasson, *J. Stat. Mech.* P09004 (2004).
[8] G. Bianconi and A. Capocci, *Phys. Rev. Lett.* **90** , 078701 (2003).
[9] G. Bianconi and M. Marsili, JSTAT P06005 (2005).
[10] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii and U. Alon, *Science* **298**, 824 (2002).
[11] R. Dobrin, Q. K. Beg, A.-L. Barabási and Z. N. Oltvai, *BMC Bioinformatics* **5** 10 (2004).
[12] A. Vazquez, R. Dobrin, D. Sergi, J.-P. Eckmann, Z. N. Oltvai, A.-L. Barabási, *PNAS* **101**, 17940 (2004).
[13] S. S. Skiena, in *The Algorithm Design Manual*, pp. 144 and 312-314 (New York: Springer-Verlag, 1997).
[14] Let K be maximal integer such that removing iteratively all nodes with degree less than K leaves a non-empty sub-graph, the so-called K-core, and let c_K be the clique number of the K-core. Then if $c_K > K$ then it is easy to show that $c_{\max} = c_K$. Else if $c_K \le K$ then we have $c_K \le c_{\max} \le K$. Hence we derive the bounds $c_- = \min(c_K, K) \le c_{\max} \le \max(c_K, K)$.
[15] T. R. Jensen, B. Toft, *Graph Coloring Problems* (New York: Wiley, 1994).
[16] I. Derenyi, G. Palla and T. Vicsek, *Phys. Rev. Lett.* **94** 160202 (2005).
[17] G. Palla, I. Derenyi, I. Farkas and T. Vicsek, *Nature* **435** 815 (2005).
[18] The Internet datasets we used are the ones collected by University of Oregon Route Views project, NLANR and the protein-protein interaction datasets are one listed in DIP database.
[19] G Caldarelli, A. Capocci, P. De Los Rios and M. A. Muñoz, *Phys. Rev. Lett.* **89**, 258702 (2002).
[20] M. Boguña and R. Pastor-Satorras, *Phys. Rev. E* **68**, 036112 (2003).
[21] M. Molloy and B. Reed, *Random Structures and Algorithms* **6**, 161 (1995).
[22] K.-L. Goh, B. Kahng and D. Kim, *Phys. Rev. Lett.* **87** 278701 (2001).
[23] G. Bianconi and M. Marsili (in preparation).
[24] S.Dorogotsev, A.V. Goltsev and J. F. F. Mendes, *Phys. Rev. E* **66** 016104 (2002);M. Leone, A. Vázquez, A. Vespignani and R. Zecchina,*Eur. Phys. J. B* **28** 191 (2002).
[25] Indeed Eq. (7) can be expanded in moments of $\mathcal{Q}^{c-1}$ and a direct calculation shows that the ratio of the $n^{\rm th}$ term to the first is $\langle \mathcal{Q}^{(c-1)n}\rangle e^{-ny^*}/\langle \mathcal{Q}^{c-1}\rangle e^{-y^*} \approx \frac{c-\gamma}{(c-1)n-\gamma+1}\left(\frac{c(c-1)}{bN}\right)^{n-1}$, which vanishes as $N \to \infty$ when $c \approx \bar{c}$.

List of other working papers:

2005

1. Shaun Bond and Soosung Hwang, Smoothing, Nonsynchronous Appraisal and Cross-Sectional Aggreagation in Real Estate Price Indices, WP05-17
2. Mark Salmon, Gordon Gemmill and Soosung Hwang, Performance Measurement with Loss Aversion, WP05-16
3. Philippe Curty and Matteo Marsili, Phase coexistence in a forecasting game, WP05-15
4. Matthew Hurd, Mark Salmon and Christoph Schleicher, Using Copulas to Construct Bivariate Foreign Exchange Distributions with an Application to the Sterling Exchange Rate Index (Revised), WP05-14
5. Lucio Sarno, Daniel Thornton and Giorgio Valente, The Empirical Failure of the Expectations Hypothesis of the Term Structure of Bond Yields, WP05-13
6. Lucio Sarno, Ashoka Mody and Mark Taylor, A Cross-Country Financial Accelorator: Evidence from North America and Europe, WP05-12
7. Lucio Sarno, Towards a Solution to the Puzzles in Exchange Rate Economics: Where Do We Stand?, WP05-11
8. James Hodder and Jens Carsten Jackwerth, Incentive Contracts and Hedge Fund Management, WP05-10
9. James Hodder and Jens Carsten Jackwerth, Employee Stock Options: Much More Valuable Than You Thought, WP05-09
10. Gordon Gemmill, Soosung Hwang and Mark Salmon, Performance Measurement with Loss Aversion, WP05-08
11. George Constantinides, Jens Carsten Jackwerth and Stylianos Perrakis, Mispricing of S&P 500 Index Options, WP05-07
12. Elisa Luciano and Wim Schoutens, A Multivariate Jump-Driven Financial Asset Model, WP05-06
13. Cees Diks and Florian Wagener, Equivalence and bifurcations of finite order stochastic processes, WP05-05
14. Devraj Basu and Alexander Stremme, CAY Revisited: Can Optimal Scaling Resurrect the (C)CAPM?, WP05-04
15. Ginwestra Bianconi and Matteo Marsili, Emergence of large cliques in random scale-free networks, WP05-03
16. Simone Alfarano, Thomas Lux and Friedrich Wagner, Time-Variation of Higher Moments in a Financial Market with Heterogeneous Agents: An Analytical Approach, WP05-02
17. Abhay Abhayankar, Devraj Basu and Alexander Stremme, Portfolio Efficiency and Discount Factor Bounds with Conditioning Information: A Unified Approach, WP05-01

2004

1. Xiaohong Chen, Yanqin Fan and Andrew Patton, Simple Tests for Models of Dependence Between Multiple Financial Time Series, with Applications to U.S. Equity Returns and Exchange Rates, WP04-19
2. Valentina Corradi and Walter Distaso, Testing for One-Factor Models versus Stochastic Volatility Models, WP04-18
3. Valentina Corradi and Walter Distaso, Estimating and Testing Sochastic Volatility Models using Realized Measures, WP04-17
4. Valentina Corradi and Norman Swanson, Predictive Density Accuracy Tests, WP04-16
5. Roel Oomen, Properties of Bias Corrected Realized Variance Under Alternative Sampling Schemes, WP04-15
6. Roel Oomen, Properties of Realized Variance for a Pure Jump Process: Calendar Time Sampling versus Business Time Sampling, WP04-14

7. Richard Clarida, Lucio Sarno, Mark Taylor and Giorgio Valente, The Role of Asymmetries and Regime Shifts in the Term Structure of Interest Rates, WP04-13
8. Lucio Sarno, Daniel Thornton and Giorgio Valente, Federal Funds Rate Prediction, WP04-12
9. Lucio Sarno and Giorgio Valente, Modeling and Forecasting Stock Returns: Exploiting the Futures Market, Regime Shifts and International Spillovers, WP04-11
10. Lucio Sarno and Giorgio Valente, Empirical Exchange Rate Models and Currency Risk: Some Evidence from Density Forecasts, WP04-10
11. Ilias Tsiakas, Periodic Stochastic Volatility and Fat Tails, WP04-09
12. Ilias Tsiakas, Is Seasonal Heteroscedasticity Real? An International Perspective, WP04-08
13. Damin Challet, Andrea De Martino, Matteo Marsili and Isaac Castillo, Minority games with finite score memory, WP04-07
14. Basel Awartani, Valentina Corradi and Walter Distaso, Testing and Modelling Market Microstructure Effects with an Application to the Dow Jones Industrial Average, WP04-06
15. Andrew Patton and Allan Timmermann, Properties of Optimal Forecasts under Asymmetric Loss and Nonlinearity, WP04-05
16. Andrew Patton, Modelling Asymmetric Exchange Rate Dependence, WP04-04
17. Alessio Sancetta, Decoupling and Convergence to Independence with Applications to Functional Limit Theorems, WP04-03
18. Alessio Sancetta, Copula Based Monte Carlo Integration in Financial Problems, WP04-02
19. Abhay Abhayankar, Lucio Sarno and Giorgio Valente, Exchange Rates and Fundamentals: Evidence on the Economic Value of Predictability, WP04-01

2002

1. Paolo Zaffaroni, Gaussian inference on Certain Long-Range Dependent Volatility Models, WP02-12
2. Paolo Zaffaroni, Aggregation and Memory of Models of Changing Volatility, WP02-11
3. Jerry Coakley, Ana-Maria Fuertes and Andrew Wood, Reinterpreting the Real Exchange Rate - Yield Diffential Nexus, WP02-10
4. Gordon Gemmill and Dylan Thomas , Noise Training, Costly Arbitrage and Asset Prices: evidence from closed-end funds, WP02-09
5. Gordon Gemmill, Testing Merton's Model for Credit Spreads on Zero-Coupon Bonds, WP02-08
6. George Christodoulakis and Steve Satchell, On th Evolution of Global Style Factors in the MSCI Universe of Assets, WP02-07
7. George Christodoulakis, Sharp Style Analysis in the MSCI Sector Portfolios: A Monte Caro Integration Approach, WP02-06
8. George Christodoulakis, Generating Composite Volatility Forecasts with Random Factor Betas, WP02-05
9. Claudia Riveiro and Nick Webber, Valuing Path Dependent Options in the Variance-Gamma Model by Monte Carlo with a Gamma Bridge, WP02-04
10. Christian Pedersen and Soosung Hwang, On Empirical Risk Measurement with Asymmetric Returns Data, WP02-03
11. Roy Batchelor and Ismail Orgakcioglu, Event-related GARCH: the impact of stock dividends in Turkey, WP02-02
12. George Albanis and Roy Batchelor, Combining Heterogeneous Classifiers for Stock Selection, WP02-01

2001

1. Soosung Hwang and Steve Satchell , GARCH Model with Cross-sectional Volatility; GARCHX Models, WP01-16
2. Soosung Hwang and Steve Satchell, Tracking Error: Ex-Ante versus Ex-Post Measures, WP01-15
3. Soosung Hwang and Steve Satchell, The Asset Allocation Decision in a Loss Aversion World, WP01-14
4. Soosung Hwang and Mark Salmon, An Analysis of Performance Measures Using Copulae, WP01-13
5. Soosung Hwang and Mark Salmon, A New Measure of Herding and Empirical Evidence, WP01-12

6. Richard Lewin and Steve Satchell, The Derivation of New Model of Equity Duration, WP01-11
7. Massimiliano Marcellino and Mark Salmon, Robust Decision Theory and the Lucas Critique, WP01-10
8. Jerry Coakley, Ana-Maria Fuertes and Maria-Teresa Perez, Numerical Issues in Threshold Autoregressive Modelling of Time Series, WP01-09
9. Jerry Coakley, Ana-Maria Fuertes and Ron Smith, Small Sample Properties of Panel Time-series Estimators with I(1) Errors, WP01-08
10. Jerry Coakley and Ana-Maria Fuertes, The Felsdtein-Horioka Puzzle is Not as Bad as You Think, WP01-07
11. Jerry Coakley and Ana-Maria Fuertes, Rethinking the Forward Premium Puzzle in a Non-linear Framework, WP01-06
12. George Christodoulakis, Co-Volatility and Correlation Clustering: A Multivariate Correlated ARCH Framework, WP01-05
13. Frank Critchley, Paul Marriott and Mark Salmon, On Preferred Point Geometry in Statistics, WP01-04
14. Eric Bouyé and Nicolas Gaussel and Mark Salmon, Investigating Dynamic Dependence Using Copulae, WP01-03
15. Eric Bouyé, Multivariate Extremes at Work for Portfolio Risk Measurement, WP01-02
16. Erick Bouyé, Vado Durrleman, Ashkan Nikeghbali, Gael Riboulet and Thierry Roncalli, Copulas: an Open Field for Risk Management, WP01-01

2000

1. Soosung Hwang and Steve Satchell , Valuing Information Using Utility Functions, WP00-06
2. Soosung Hwang, Properties of Cross-sectional Volatility, WP00-05
3. Soosung Hwang and Steve Satchell, Calculating the Miss-specification in Beta from Using a Proxy for the Market Portfolio, WP00-04
4. Laun Middleton and Stephen Satchell, Deriving the APT when the Number of Factors is Unknown, WP00-03
5. George A. Christodoulakis and Steve Satchell, Evolving Systems of Financial Returns: Auto-Regressive Conditional Beta, WP00-02
6. Christian S. Pedersen and Stephen Satchell, Evaluating the Performance of Nearest Neighbour Algorithms when Forecasting US Industry Returns, WP00-01

1999

1. Yin-Wong Cheung, Menzie Chinn and Ian Marsh, How do UK-Based Foreign Exchange Dealers Think Their Market Operates?, WP99-21
2. Soosung Hwang, John Knight and Stephen Satchell, Forecasting Volatility using LINEX Loss Functions, WP99-20
3. Soosung Hwang and Steve Satchell, Improved Testing for the Efficiency of Asset Pricing Theories in Linear Factor Models, WP99-19
4. Soosung Hwang and Stephen Satchell, The Disappearance of Style in the US Equity Market, WP99-18
5. Soosung Hwang and Stephen Satchell, Modelling Emerging Market Risk Premia Using Higher Moments, WP99-17
6. Soosung Hwang and Stephen Satchell, Market Risk and the Concept of Fundamental Volatility: Measuring Volatility Across Asset and Derivative Markets and Testing for the Impact of Derivatives Markets on Financial Markets, WP99-16
7. Soosung Hwang, The Effects of Systematic Sampling and Temporal Aggregation on Discrete Time Long Memory Processes and their Finite Sample Properties, WP99-15
8. Ronald MacDonald and Ian Marsh, Currency Spillovers and Tri-Polarity: a Simultaneous Model of the US Dollar, German Mark and Japanese Yen, WP99-14
9. Robert Hillman, Forecasting Inflation with a Non-linear Output Gap Model, WP99-13
10. Robert Hillman and Mark Salmon , From Market Micro-structure to Macro Fundamentals: is there Predictability in the Dollar-Deutsche Mark Exchange Rate?, WP99-12
11. Renzo Avesani, Giampiero Gallo and Mark Salmon, On the Evolution of Credibility and Flexible Exchange Rate Target Zones, WP99-11
12. Paul Marriott and Mark Salmon, An Introduction to Differential Geometry in Econometrics, WP99-10

13. Mark Dixon, Anthony Ledford and Paul Marriott, Finite Sample Inference for Extreme Value Distributions, WP99-09
14. Ian Marsh and David Power, A Panel-Based Investigation into the Relationship Between Stock Prices and Dividends, WP99-08
15. Ian Marsh, An Analysis of the Performance of European Foreign Exchange Forecasters, WP99-07
16. Frank Critchley, Paul Marriott and Mark Salmon, An Elementary Account of Amari's Expected Geometry, WP99-06
17. Demos Tambakis and Anne-Sophie Van Royen, Bootstrap Predictability of Daily Exchange Rates in ARMA Models, WP99-05
18. Christopher Neely and Paul Weller, Technical Analysis and Central Bank Intervention, WP99-04
19. Christopher Neely and Paul Weller, Predictability in International Asset Returns: A Re-examination, WP99-03
20. Christopher Neely and Paul Weller, Intraday Technical Trading in the Foreign Exchange Market, WP99-02
21. Anthony Hall, Soosung Hwang and Stephen Satchell, Using Bayesian Variable Selection Methods to Choose Style Factors in Global Stock Return Models, WP99-01

1998

1. Soosung Hwang and Stephen Satchell, Implied Volatility Forecasting: A Compaison of Different Procedures Including Fractionally Integrated Models with Applications to UK Equity Options, WP98-05
2. Roy Batchelor and David Peel, Rationality Testing under Asymmetric Loss, WP98-04
3. Roy Batchelor, Forecasting T-Bill Yields: Accuracy versus Profitability, WP98-03
4. Adam Kurpiel and Thierry Roncalli , Option Hedging with Stochastic Volatility, WP98-02
5. Adam Kurpiel and Thierry Roncalli, Hopscotch Methods for Two State Financial Models, WP98-01