IDENTIFYING DYNAMICS IN STRATEGIC GROUPS

Gabriel I. Penagos-Londoño (penagosi@javeriana.edu.co)

Department of Economics – Pontificia Universidad Javeriana - Colombia

Felipe Ruiz-Moreno (felipe.ruiz@ua.es) – corresponding author

Department of Marketing - University of Alicante – Spain

ABSTRACT

We characterize the situation in the Spanish banking industry through the identification of strategic groups. We use a 13-year dataset for the period 1992–2004 and a sophisticated statistical model to identify these strategic groups. The primary contribution of this empirical study is to model the evolution of these strategic groups using a time inhomogeneous hidden Markov model (HMM) in which the time variable transition matrix captures institutions' group switching behavior. We consider a mixture model is the data generating process. Two strategic groups are identified. These groups are primarily characterized by size and other strategic variables. The probability of remaining in a group is generally high: 87.28% for SG1 and 61.84% for SG2. The probability of switching groups is low: 12.72% probability of switching from SG1 to SG2 and 38.16% probability of switching from SG2 to SG1. Banks in SG1 seem more stable over time; they have low levels of switching behavior and well-defined long-term behavior. Banks in SG2 seem to evolve in terms of group membership.

KEYWORDS: strategic groups; dynamics; hidden Markov models; banking

Citation:

Penagos-Londoño, G. and Ruiz-Moreno, F. (2019), "Identifying dynamics in strategic groups", *Journal of Modelling in Management*, Vol. 14 No. 2, pp. 408-429.

https://doi.org/10.1108/JM2-08-2018-0114

1. INTRODUCTION

Since Hunt's (1972) seminal work, the concept of strategic groups has generated a rich stream of literature.[1] A strategic group is a set of firms that use the same or similar strategies according to some strategic dimensions, resulting in homogeneous competitive actions within an industry (Caves & Porter 1977; Cool & Schendel 1987). Scholars continue to discuss strategic groups because of their impact on certain decisions that the firm must take. For example, strategic groups may affect the nature of competition within an industry.

In the strategic group literature, there is controversy regarding not only the usefulness of the concept for advancing strategic management research but also the question of whether strategic groups actually exist and, if so, what method is most suitable for identifying them (Murthi *et al.* 2013). Therefore, correctly identifying strategic groups and assessing how well they characterize a particular industry is primarily an empirical question, the answers to which provide valuable input to managerial decisions. In this stream of research, cluster analysis is probably the most popular method of identifying strategic groups (Mascarenhas & Aaker 1989; Fiegenbaum & Thomas 1990). Most of the cited studies focus on selecting variables that capture product market and resource commitments and applying cluster analysis across stable periods (González-Moreno & Sáez-Martínez 2008) based largely on heuristic procedures like Ward's method and k-means (Tuma *et al.* 2011). However, the generally weak statistical basis of such methods is a major drawback, and crucial segmentation questions such as determining the optimal number of segments cannot be answered satisfactorily by heuristic procedures. In addition, understanding the dynamics of strategic groups enriches traditional models of industrial economics and the strategic management literature.

Consequently, a primary goal of recent strategic group research has been to improve

[1] For detailed reviews of this literature, see McGee & Thomas (1986), Thomas & Venkatraman (1988), Ketchen *et al.* (2004), and DeSarbo *et al.* (2009).

theoretical and methodological techniques to identify groups over time. Accordingly, other methods have been proposed to overcome the limitations of cluster analysis and advance research on how best to identify strategic group dynamics. Some of the most recent advances in this field include a new cluster technique known as MCLUST (Zuñiga-Vicente *et al.* 2004), spatial clusterwise multidimensional scaling (DeSarbo *et al.* 2008, 2009), latent class regression analysis (Murthi *et al.* 2013), finite mixture-based models (DeSarbo *et al.* 2010), and hidden Markov models (Ebbes *et al.* 2010). All of these techniques have been developed to capture the varying, dynamic nature of strategic groups over time.

We present analysis of strategic groups in a specific industry using a sophisticated method that overcomes some of the aforementioned empirical criticisms. We enrich the literature by empirically addressing three considerations. First, we explore whether strategic groups existed and, if so, what form these groups took in the Spanish banking industry between 1992 and 2004. To do so, we use an enhanced version of the hidden Markov model (HMM) applied by Ebbes *et al.* (2010). Specifically, we propose an inhomogeneous HMM (instead of the homogeneous HMM proposed by the aforementioned authors). In this inhomogeneous HMM, the time-variable transition matrix captures institutions' group switching behavior between two years (t and $t+1$) in the study period. From a technical perspective, we extend the HMM by employing initial values for the priors adjusted to the data. Doing so adds flexibility because it allows us to control the amount of information included in the priors. Consequently, we allow for time-varying parameters in the distributions to analyze the evolution of strategic group membership in this industry. This enables us to detect changes in group strategy, changes in membership, and the stability of groups over time. This contribution expands the set of available tools to enable more accurate analysis of the dynamics of strategic groups. Second, we compare our proposed model with suitable benchmark models to examine its performance in explaining and diagnosing the evolution of strategic group membership in an industry. Finally, we delineate the strategic dimensions, or variables, to test for performance differences between groups.

The remainder of the paper is organized as follows: In Section 2, we provide a brief literature review, primarily focusing on strategic group research and the characteristics of the Spanish banking industry. The method is described in Section 3, and the variables are defined and operationalized. We present the results in Section 4. In Section 5, we discuss the key conclusions and implications of our study and consider possible avenues for future research.

2. THEORY: LITERATURE REVIEW

We begin this section with a brief review of the strategic group literature. We then justify the choice of the Spanish banking industry as the setting for our study.

2.1 Strategic group research

The strategic group literature includes the following major streams of research: the emergence of groups, within- and between-group rivalry, performance differences between groups, and the stability of group structures (Mas-Ruiz *et al.* 2014). We review each of these four streams separately.

The emergence of strategic groups. Several theories have attempted to explain the existence and creation of strategic groups. The earliest theories are based on differences in firms' assets and capabilities or attitudes toward risk (Caves & Porter 1977). These differences lead firms to invest in barriers to mobility and in forming groups. Tang & Thomas (1992) report that economic models of spatial competition and cognitive models of competitive structure can also explain the creation of strategic groups. Peteraf & Shanley (1997) define the identity of a strategic group as a set of mutual understandings that derive from interactions and ensure that members are aware of the logic that underpins their behavior and can therefore predict how other group members are likely to react. Dranove *et al.* (1998)

affirm that group-level effects on performance derive from strategic interactions among members. While strategic interactions and existing relations are critical to group-level effects on performance, mobility barriers help sustain these group-level effects by limiting entry to the group and enhancing strategic interactions among members (Más-Ruiz *et al.* 2014). Dranove *et al.*'s (1998) approach meets the scientific criterion of falsifiability in that a finding that showed the absence of systematic variation in profitability across groups would be equivalent to a finding that showed the non-existence of strategic groups (Murthi *et al.* 2013).

Rivalry within and between strategic groups. The literature that examines whether the degree of rivalry differs within and between groups is ambiguous (Porter 1976, 1979; Peteraf 1993; Cool & Dierickx 1993; Smith *et al.* 1997). On the one hand, several studies provide evidence that rivalry between strategic groups is greater than rivalry within groups (Caves & Porter 1977; Porter 1976, Peteraf 1993). Based on the resource-based view of the firm, these studies suggest that members of the same group have similar resources, so they will act and react to competitive disturbances similarly. On the other hand, Gimeno & Woo (1996) suggest that greater strategic distance makes tacit coordination easier by signaling whether a rival has overstepped its tacit boundary. According to the resource-based view of the firm, a greater within-group rivalry could result from homogeneity of resources among members (Barney 1991; Bogner & Thomas 1994) because each firm strives to achieve the same goals but does not have unique resources or isolation mechanisms that enable the firm to gain a competitive advantage (Smith *et al*. 1997).

Performance difference between groups. The earliest studies of this relationship indicate that owing to barriers to mobility, which produce rigidities, members of strategic groups have a relative cost advantage over other firms (Porter 1979). Thus, differences in performance between group members and outsiders tend to persist in the medium to long term (McGee & Thomas 1986). Thomas & Venkatraman (1988), Leask & Parker (2007), and Nair

& Kotha (2001), among others, report performance differences across strategic groups. However, empirical evidence remains inconclusive. Other studies have shown that mobility barriers provide insufficient theoretical support for the link between group membership and performance (Cool & Schendel 1987, 1988; Frazier & Howell 1983). Group members are keenly aware of their mutual dependence and often react in the same fashion to the same external stimuli (Caves & Porter 1977). This mutual dependence of firms within a strategic group makes the development of tacit agreements between member firms easier, which prevents other firms from becoming members.

Stability of a group structure. This stream of research focuses on understanding changes over time in group strategy, strategic group membership, and the number of strategic groups (Mascarenhas 1989). Despite certain theoretical developments in the domain of strategic groups that suggest that strategic group membership appears to be stable over time (Mascarenhas and Aaker 1989), most studies provide evidence that firms within a strategic group evolve (e.g., Ruiz 1999; Nair & Filer 2003; Rebière & Mavoori 2016). Changes over time in a given firm's strategy tend to be neither random nor uniform. Instead, these changes occur at specific moments when the industry suffers a shock of some type. This shock creates new opportunities or threats, to which firms react. This response to changes in the environment might cause an unusually high number of firms to change their strategies and, according to Dranove *et al*. (1998), could promote the adoption of collective strategies among group members. This source of change and discontinuity is an important part of the dynamic model of strategic groups. Other scholars (e.g., Cool & Schendel 1987, 1988; Cool & Dierickx 1993; Fiegenbaum & Thomas 1990 1993; Fiegenbaum *et al.* 1987, 1990) argue that firms modify their strategies in response not only to changes in the industry environment but also to the imitation activities of other firms that attempt to copy their behavior and to the market reception of their product positioning. From a strategic management/marketing and economics perspective, studying strategic group dynamics can help identify entry and exit barriers at the group level and explain how competitive dynamics change over time (Ebbes *et al.* 2010).

2.2. Spanish banking industry

The banking industry is a strategic sector in the global economy. Therefore, numerous studies of financial contexts have been carried out from a strategic group perspective. For some examples, see Amel & Rhoades (1989), Berg & Kim (1994), DeSarbo & Grewal (2008), Burke (1990), Ebbes *et al.* (2010), Epure *et al.* (2011), Más-Ruiz *et al.* (2005), Más-Ruiz & Ruiz-Moreno (2011), Más-Ruiz *et al.* (2014), Ray & Das (2010), Ruiz (1999), Spiller & Favaro (1984), and Zuñiga-Vicente *et al.* (2004).

According to Más-Ruiz *et al.* (2005), the Spanish banking market provides an interesting scenario for analyzing the domains of strategic groups. The Spanish banking industry's deregulation process ended in 1992. Two types of regulations have exerted a particularly strong influence on the Spanish banking industry: the regulation of firm behavior through price setting and the regulation of market structure through the control of savings banks' geographical expansion (Gual 1992; García-Cestona & Surroca 2008).

Despite the liberalization of prices and controls on fees in 1987, the regulation of interest rates since the 1960s removed price competition and forced the larger banks to compete by investing more in services and proximity to the customer by expanding their branch networks. The regulation of the Spanish banking industry's geographical expansion in the 1970s and 1980s created a market where financial firms operated at the national, regional, or local level. However, these geographical restrictions applied only to savings banks. Accordingly, savings banks were allowed to make strategic geographical choices, but these choices were limited to regional or local, but not national, markets. No such limits were imposed on commercial banks. Therefore, large financial firms that had branches in numerous regions faced a different (and more diversified) competitive structure and socioeconomic reality than small commercial banks and savings banks (Carbó *et al*. 2003).

However, factors such as the removal of branching restrictions on savings banks' geographical expansion in 1988 and Spain's accession to the European Union in 1986 led to the concentration of the banking industry. First, the removal of branching restrictions led to the nationwide expansion of large savings banks and the defensive formation of small savings banks into geographical groups. This defensive formation was achieved through an accelerated process of mergers and acquisitions, primarily involving savings banks that operated in the same markets. Second, Spain's accession to the harmonized European financial market also led to the consolidation of the banking sector through additional mergers and acquisitions, which drastically affected the domestic competitive environment. At the end of the 1980s, Spanish banks engaged in mergers and acquisitions to increase their size, compete in the broader European market, and preserve market power (Más-Ruiz *et al.* 2005).

3. METHOD

As stated in the introduction, the identification of strategic groups is riddled with controversy. Scholars such as Day *et al.* (1995) suggest that inconsistent results in prior studies may owe to variation in research designs, the lack of multiple criteria, and inappropriate selection methods in the identification of groups. After selecting a stable period, researchers face two major problems in identifying the strategic groups within an industry (Murthi *et al.* 2013). First, the composition of groups and the number of groups may vary according to the choice of strategic variables. Factor analysis is the most commonly used technique to identify relevant variables. The primary risk is that smaller factors derived from the factor analysis may be discarded even though these smaller factors may contain the most relevant clustering information (DeSarbo *et al*. 2009). The second problem relates to the choice of the most suitable method to identify the strategic groups. Although cluster analysis is the most common method for identifying groups, it has several limitations that are well documented in the literature (see Ketchen and Shook (1996), Murthi *et al*. (2013), and Ebbes *et al*. (2010) for details). DeSarbo *et al*. (2009) report that different methods provide different results for the same dataset.

Consequently, alternative methods have been proposed to overcome the limitations of cluster analysis and find a better way to identify strategic group dynamics. Thus, DeSarbo *et al*. (2009) propose a spatial clusterwise multidimensional scaling (MDS) technique that simultaneously identifies the number of strategic groups and strategic group membership, derives the underlying dimensions of the strategic groups, and models the evolution of the strategic groups over time. Murthi *et al*. (2013) advocate the use of latent class regression analysis, which is also known as a finite mixture model that is used to control for unobserved heterogeneity in the data, to create strategic groups based on similarities in the response coefficients of variables that are used to explain performance. Ebbes *et al*. (2010) propose the use of a homogeneous hidden Markov models (HMMs), which simultaneously identify strategic groups and explicitly model the transition probabilities across the identified strategic groups to account for potential time dependencies, enabling them to detect how strategic groups evolve over time. For each firm, the HMM approach identifies a strategic time path where the successive strategy states are linked through a first-order Markov process. Thus, the strategy that a firm adopts at time *t* depends on the strategy that the firm adopted in the previous measurement period. We use this HMM approach as a basis to identify the strategic groups and describe the dynamic path of strategic group membership over time.

3.1. Model

The assumption that underlies the concept of strategic groups is that groups are internally homogeneous but externally heterogeneous; strategic responses to certain market variables will be similar within groups but different among groups. Based on this assumption, we classify each financial institution into one of a given number of homogeneous classes. We model the evolution of strategic groups using a time inhomogeneous HMM in which the time-variable transition matrix captures institutions' group switching behavior between two years (t and $t+1$) in the study period. From a technical perspective, we extend the HMM by employing initial values for the priors adjusted to the data. Doing so adds flexibility because it allows us to control the amount of information included in the priors. Consequently, we allow for time-varying parameters in the distributions to analyze the evolution of strategic group membership in this industry. This enables us to detect changes in group strategy, changes in membership, and the stability of groups over time.

Strategic variables can be thought of as realizations of certain processes that characterize the groups. However, these hypothetical groups are not observed because the membership information has been lost. Therefore, it is straightforward to consider a mixture model is the data generating process. If we let K denote the number of such classes, such that $\mathrm{k} = 2, \dots, \mathrm{K}$, and T denote the number of periods indexed by the discrete time variable t ($\mathrm{t} = 1, \dots, \mathrm{T}$), we can model the distribution of the observed variables x_{ti} for bank i at time t as the following mixture density:

$$f(x_{ti}|x_{t-1}, \theta_{tk}) = \sum_{k=1}^{K} \omega_{tk} \varphi(x_{ti}|z_{ti} = k, \theta_{tk}) \qquad (1)$$

In Eq. 1, φ is the multivariate normal density distribution with parameters $\theta_{tk} = (\mu_{tk}, \Sigma_{tk})$, where $\mu = (\mu_{tk})$ are the mean vectors, $\Sigma = (\Sigma_{tk})$ are the covariance matrices, and $\mathcal{MN}(\mu_{tk}, \Sigma_{tk})$ or $\mathcal{MN}(\theta_{tk})$ is the multivariate normal probability distribution. Also in Eq. 1, $\omega_{tk} > 0$, $\sum_{k=1}^{K} \omega_{tk} = 1$, are the weights or mixing proportions. The densities are taken with respect to the Lebesgue measure. In each period, a bank can remain in the same class or dynamically change its membership to another class.

We assume that the number of classes remains stable over time and that the behavior is described by a set of strategic variables represented by an *R*-dimensional vector x_{ti}. The weights can be thought of as the relative class sizes for the period t or the probabilities of belonging to a specific class. For a specific period t, the vector $\omega_t = (\omega_{t1}, \dots, \omega_{tK})$ provides the weight distribution and belongs to the unit simplex $\mathcal{E}_K$:

$\mathcal{E}_K = \{\omega_t = (\omega_{t1}, \dots, \omega_{tK}) \in \mathbb{R}^K | \omega_t \cdot \mathbf{1} = 1, \omega_{tk} \in (0,1), \forall k, t\}$, where $\mathbf{1}$ is the $\mathbb{R}^K$ vector of ones.

The hidden (or latent) variable that determines membership is embedded in a discrete time process with finite space $z_{ti} \in \{1, \cdots, K\}$, which is assumed to be an irreducible aperiodic Markov chain with respect to *t*. This variable associates each observation x_{ti} with the normal distribution from which it is sampled. Consequently, the stochastic pair (x_{ti}, z_{ti}) is an HMM. Here, (x, z), which is taken as a data collection, is referred to as the *complete data*, where x is the *incomplete data* and z is the vector of hidden variables. The weights can be written in terms of the hidden variables as $\omega_{tk} = p(z_{ti} = k|\vartheta)$. The initial distribution is thus $\omega_k \doteq \omega_{0k} = p(z_{0i} = k|x_{t-1}, \xi)$, which is assumed to be ergodic and where $\omega = (\omega_1, \dots, \omega_K) \doteq p(z_{0i}|\xi)$.

Furthermore, conditional on knowing this hidden variable, the random variables x_{ti} are stochastically independent (Frühwirth-Schnatter 2006). The model becomes $x_{ti}|z_{ti} \sim \mathcal{MN}(\theta_{tk})$, with $p(z_{ti} = k) = \omega_{tk}$. Accordingly, the latent variable is distributed as a multinomial random variable $z_{ti} \sim \mathcal{M}(1; \omega_{t1}, \cdots, \omega_{tk})$.

We assume a transition matrix (ξ_t), which characterizes the process z_{ti} for each bank i. This transition matrix is allowed to change over time because the Markov chain is defined as time inhomogeneous. Thus, we can write the transition matrix as $\xi_{tjk} = p\left(z_{t,i} = k \middle| z_{t-1,i} = j\right)$

Because it is a row-stochastic matrix, $\xi_{tjk} \geq 0;\ \forall j, k \in \{1, \cdots, K\}$, and $\sum_{k=1}^{K} \xi_{tjk} = 1\ \ \forall t, \forall j \in \{1, \cdots, K\}$, where ξ_t takes values in the parameter space $\mathcal{E}_K^K$.

The parameters to be estimated are ω, $\xi = (\xi_t)$, and $\theta = (\theta_{tk})$, which takes values in the parameter space $\Theta \subseteq \mathbb{R}^Q$ for all k and t. The total parameter space to be estimated is $\Psi = \Theta^{TK} \times \mathcal{E}_{K-1}^{TK+1}$, where $\Psi \subseteq \mathbb{R}^S$ and $S = KTQ + T(K+1)(K-1)$. We say that the mixture density f is parametrized by $\vartheta \in \Psi$.

Posterior and prior distributions and full conditionals

Assume that, for each time point $t \leq T$, we obtain from the mixture f a sample of N independent and identically distributed draws $x = (x_{ti})_{i=1,\cdots,N;t=1,\cdots,T}, x \in \mathbb{R}^R$. In addition, the density of $x_t = (x_{si})_{i=1,\cdots,N;s=t}$ is allowed to depend on x_{t-1}. To construct the mixture likelihood function, we assume that the random variables (x_{ti}) for $t = 1, \ldots, T$ and $i = 1, \ldots, N$ are stochastically independent, conditional on knowing z_i (Ebbes *et al*. 2010). Therefore, the likelihood function reads:

$$l(\theta|x_i) = \prod_{t=1}^{T}\left(\sum_{k=1}^{K} p(z_{ti} = k|\theta, x_{t-1})\varphi(x_{ti}|z_{ti} = k, \theta)\right)$$

The expanded expression has K^T terms, leading to a problem that is computationally too expensive even for a modest set of observations (Casella *et al.* 2004). Moreover, the form of the likelihood function gives rise to multiple local maxima or even a function that is unbounded (Young 2008). This complexity also precludes the use of Bayes' estimators, inasmuch as the resulting expression for the posterior is $p(\vartheta|x_i) = l(\theta|x_i)p(\vartheta)$, where $p(\vartheta)$ is the prior for the parameter set Θ. The likelihood has to be understood as comprising a collection of paths followed by each bank i. The standard procedure to deal with the dimensionality problem is to exploit the HMM structure of the model. Thus, *the complete data likelihood*, $L(x_i, \mathrm{z}_i|\vartheta) = p(x_i|z_i, \vartheta)p(z_i|\vartheta)$, is used. We assume that the prior distribution is of the form $p(\vartheta) = \prod_k p(\theta_k) p(\xi)$. Given this expression for the likelihood, the *complete data posterior distribution* is therefore:

$$p(\vartheta|x_i, \mathrm{z}_i) = \left(\prod_{t=1}^{T} \xi_{t,z_{t-1,i},z_{ti}} \varphi\left(x_{ti}\middle|\theta_{tz_{ti}}\right) p(\xi_t) \prod_{k=1}^{K} p(\theta_{tk})\right) p(z_{0i}|\xi) \tag{2}$$

The model includes the constants $\delta = (\gamma, \lambda, \tau, \upsilon, \beta)$ for $k = 1, \cdots, K$ and $t = 1, \cdots, T$, and the parameters $(\omega, \xi, \mu, \Sigma)$. The priors, which depend on the vector of constants δ, are chosen in such a way that they are conjugated to the following complete data likelihood:

$$\omega \sim \mathcal{D}(\gamma);\ \xi_{tj} \sim \mathcal{D}(\gamma)$$

$$\mu_{tk}|\Sigma_{tk}, x, z \sim \mathcal{N}_r\left(\lambda_{tk}, \frac{\Sigma_{tk}}{\tau_{tk}}\right)$$

$$\Sigma_{tk}^{-1}|\mu_{tk}, x, z \sim \mathcal{W}_r(\upsilon, \beta_k) \tag{3}$$

where $\mathcal{D}$ is the Dirichlet distribution of order K with parameters $\gamma = (\gamma_k), \gamma_k > 0$; $\mathcal{N}_r$ is the multivariate normal distribution with *r*-dimensional vector mean λ_{tk}, $\lambda = (\lambda_{tk})$, and Σ_{tk}/τ_{tk} is the $r \times r$ −dimensional covariance matrix where $\tau = (\tau_{tk})$ are the adjusting parameters; and $\mathcal{W}_r$ is the Wishart distribution of an $r \times r$ −dimensional positive definite and symmetric matrix with υ degrees of freedom and fixed positive definite scale matrices $\beta = (\beta_k)$.

3.2. Estimation

To estimate the model parameters, we use a Markov chain Monte Carlo (MCMC) algorithm by constructing a Markov chain in Ψ, whose stationary (equilibrium) distribution is the complete data posterior distribution. To create the Markov chain, we use the Gibbs sampler (via the Clifford–Hammersley theorem), in which a sample of each parameter is successively drawn from the posterior conditional on the other parameters and the data. The simulations are used to estimate expectations of functions $f: \Psi \to \mathbb{R}$ with respect to the posterior:

$$E[f(\vartheta, z)|x] = \iint f(\vartheta, z) p(\vartheta, z|x) d\vartheta dz \tag{4}$$

These functions can be, for example, posterior moments of parameters or state variables such as $E[\theta|x]$, $E[z|x]$, or $Var(\theta|x)$. Given the simulations $(\vartheta_i, z_i) \sim p(\vartheta, z|x)$ for $i = 1, \dots, N$, the Monte Carlo estimates are given by:

$$\hat{E}[f(\vartheta, z)|x] = \frac{1}{N}\sum_{i=1}^{N} f(\vartheta_i, z_i)$$

This estimate converges a.s. to Eq. (4) as $N \rightarrow \infty$ under certain regularity conditions (Roberts & Smith, 1994). It is therefore customary to determine the full conditionals from the complete data posterior to be able to apply Gibbs sampling in order to extract random samples from the posterior distribution. After some manipulation of the equations in Eq. (3) we arrive at the following:

$$\omega_k \sim \mathcal{D}(\gamma_k + N_k);\ \xi_{tk} \sim \mathcal{D}(\gamma_k + N_k)$$

$$\mu_{tk}|\Sigma_{tk}, x, z \sim \mathcal{N}_r\left(\frac{\tau_{tk}\lambda_{tk} + S_{tk}^x}{\tau_{tk} + N_k}, \frac{\Sigma_{tk}}{\tau_{tk} + N_k}\right)$$

$$\Sigma_{tk}^{-1}|\mu_{tk}, x, z \sim \mathcal{W}_r(\upsilon + N_k, \beta_k + M_{tk}^V + S_{tk}^V) \qquad (5),\ \text{where}$$

$$S_{tk}^x = \sum_{i=1}^{N} x_{ti} I_{\{z_i = k\}}$$

$$S_{tk}^V = \sum_{i=1}^{N} (x_{ti} - \mu_{tk})(x_{ti} - \mu_{tk})' I_{\{z_{ti} = k\}}$$

$$M_{tk}^V = (\mu_{tk} - \lambda_{tk})(\mu_{tk} - \lambda_{tk})' \qquad (6)$$

where $I_{\{z_{ti} = k\}}$ is the indicator function that takes the value 1 when the condition in curly brackets is met, and 0 otherwise. The latent variable z_{ti} is sampled from the conditional posterior distribution $P(z_i|x_i, \vartheta)$. The multimove method (Frühwirth-Schnatter 2006) relies on the forward filtering backward sampling algorithm implemented for each path z_i as a mathematical object.

Although convergence diagnosis has been developed, see Cowles & Carlin (1996), the accuracy of the results is unclear (Andrieu *et al*. 2003; Flegal & Gong, 2013), and theoretical results are not always directly applicable. We approach the issue heuristically and search for stationarity in the parameters. From initial experiments, we determine a fixed length of 10,000 steps. We also discard the initial transient, specifying a burn-in period of 10% of the total iterations.

It is important to consider label switching. Each component of the complete data likelihood function of a mixture model—Eq. (2)—cannot be differentiated to each other because they are exchangeable. In other words, the parameters θ_k are not identifiable. Therefore, the likelihood is invariant under the permutation of components, i.e., relabeling. To resolve this problem in other studies (Young 2008), scholars commonly impose an identifiability constraint on the parameters. However, this solution has drawbacks on the inference results. To deal with this problem, we use Stephens's (1997) approach and implement post processing identifiability. Although we report label switching in the simulations, the groups were so well separated that the identification was straightforward.

The choice of the constants $\delta = (\gamma, \lambda, \tau, \upsilon, \beta)$ is important because it affects the posterior distribution. We therefore follow Bensmail *et al*. (1997) and Frühwirth-Schnatter (2006) to choose these constants. Because we are considering a potentially rare segment with a few important institutions, we choose $\gamma = 4$. The allocations are initialized with a clustering algorithm based on the squared Euclidean distance (K-means), but an alternative approach would be to use a Kohonen neural network. Consequently, we have the initial guess for z, and the centroids are the initial values of λ. The priors are proper for $\upsilon > (p-1)/2$. Hence, we choose $\upsilon = 2.5 + (R-1)$, while $\tau = 1$. Eq. (5) shows that τ is interpreted as the relative weight assigned to the initial guess of the cluster mean. β is the product of the covariance sampling estimate by a factor that controls the precision. The mean and covariance matrices are initialized with the sampling estimates, while ω and ξ are randomly selected.

3.3. Model comparison

As stated before, alternative methods proposed in the literature have made significant contributions to our understanding of the evolution of strategic group membership within an industry. Therefore, it is of interest to examine the performance of some benchmark models in identifying strategic groups. The performance of these models can then be compared with that of our proposed model in terms of estimation robustness and the ability to accurately classify banks into strategic groups. This model comparison is also important because conclusions about strategic group dynamics based on these methods might have strategic as well as managerial implications for firms.

We perform this comparison by applying two prominent methods described in the literature to data on the Spanish banking industry for the period 1992 to 2004. We then compare these results with those of our proposed model using the same performance criteria. Because numerous studies have used cluster analysis to identify strategic groups based on heuristic procedures such as k-means, we perform this model clustering for each period. The determination of the optimal number of groups cannot be inferred by this model, so we state the same number of groups (k) to enable comparison with other models. This basic model has weaknesses, so more sophisticated models have been proposed to overcome the limitations of cluster analysis. One example is the HMM proposed by Ebbes *et al.* (2010). The membership of a bank is determined using latent states that follow a first-order Markov process. This represents a major contribution by simultaneously identifying strategic groups and explicitly modeling the long-run transition probabilities for the identified strategic groups described by a unique transition matrix. This model is also our main source of inspiration and provides the basis for the model presented in this paper.

We apply the three aforementioned methods: cluster analysis, HMM, and the model extension proposed in this paper. The goal is to infer their performance in identifying strategic groups and defining a more accurate picture of the evolution of strategic group membership within an industry. To compare the models, we compute the log-likelihood function of the nested mixture models. The likelihood function is as follows:

$$p(x_i, \vartheta) = \prod_{t=1}^{T} \left(\sum_{k=1}^{K} \varphi(x_{ti} | z_{ti} = k, \vartheta) p(z_t = k | x_{t-1}, \vartheta) \right)$$

This expression is the product of the one-step-ahead predictive densities (Frühwirth-Schnatter 2006).

$$p(x_i, \vartheta) = \prod_{t=1}^{T} \varphi(x_{ti} | x_{t-1}, \vartheta)$$

where

$$\varphi(x_{ti} | x_{t-1}, \vartheta) = \sum_{k=1}^{K} \varphi(x_{ti} | z_{ti} = k, x_{t-1}, \vartheta) p(z_t = k | x_{t-1}, \vartheta)$$

3.4. Sample and variables

Our sample comprises 25 savings banks and 13 commercial banks. Data correspond to a 13-year period (1992–2004). Boeker (1991) and Burke (1990) report a relationship between a bank’s size and geographical spread: National banks are generally larger than regional banks are. Furthermore, competition in banking is largely driven by geographical constraints, so customers are unwilling to travel for long distances to meet their banking needs (Ebbes *et al.* 2010). For these two reasons, the sample comprises the largest financial entities in Spain in terms of assets. We therefore study private banks and savings banks with a national and regional scope and extensive branch networks. Consequently, this sample excludes smaller financial entities that focus on local markets.

We analyze empirical data on the Spanish banking industry for the period after the end of the deregulation process in 1992. The subsequent increase in competition was a prominent characteristic of this period, resulting in critical changes in bank customers' behavior. The study period is characterized by an economic crisis that took place between 1992 and 1996, followed by drastic change in the national economic cycle starting with strong recovery in 1997.

Numerous variables can be used to identify strategic groups in the banking industry. We use a series of key strategic dimensions to identify the strategic groups (Ebbes *et al*. 2010; Frazier & Howell 1983; Lewis & Thomas 1990). $SD1_{it}$ is the total monetary value of loans issued in year *t.* We define loans to include the sum of loans to entities and loans to customers reported on the balance sheet. $SD2_{it}$ is the interest rate on loans issued by bank *i* in year *t*. Because banks' interest rates are not reported, we estimate the annual average for each bank based on ratios of loan revenues, including the ratio of fee income to the value of outstanding loans. The evolution of this series over time closely resembles that of the market interest rate cited by the Bank of Spain (Carbó *et al.* 2009). $SD3_{it}$ is the total value of deposits for bank *i* in year *t*. We define deposits as the sum of deposits from customers, deposits represented by negotiable shares, and other deposits reported on the balance sheet. $SD4_{it}$ is the interest rate for deposits issued by bank *i* in year *t.* Because banks' interest rates are not reported, for each bank, we estimate the annual average interest rate on deposits based on ratios of deposit expenses, including the ratio of fee expenses to the value of outstanding deposits. $SD5_{it}$ is the price of labor (i.e., personnel costs divided by number of employees) for bank *i* in year *t*. $SD6_{it}$ is the price of capital (i.e., operating costs other than personnel costs divided by fixed assets) for bank *i* in year *t*. $SD7_{it}$ is the operating costs for bank *i* in year *t*. This value was derived by multiplying the price of inputs by the quantity of inputs. $SD8_{it}$ is the number of ATMs for bank *i* in year *t.* This variable is an indirect indicator of the technological expansion of bank *i* (Martín & Sáez 1997), which directly affects its operating costs. $SD9_{it}$ is the ratio of assets to number of employees of bank *i* in year *t*. This variable reflects the volume of assets managed by each employee at firm *i* (García *et al*. 1998). $SD10_{it}$ is the regional GDP, which is an indicator of the economic activity in the bank's market (Lago & Salas 2005). This variable is computed for each firm as the branch-weighted average of GDP for regions where bank *i* operates (Maudos 2001). $SD11_{it}$ is the Herfindahl index, which indicates the concentration and power of bank *i* in year *t*. $SD12_{it}$ is the number of branches of bank *i* in year *t*. Finally, $SD13_{it}$ is an advertising index defined as the share of non-wage expenditure on total administrative costs for bank *i* in year *t*. This index is supposed to reflect the amount that bank *i* spends on advertising and other marketing activities to attract new customers (Barros & Modesto 1999).

We use a set of variables that capture information on rivals to test competition within the different strategic groups. For bank *i* in year *t*, $SD14_{it}$ is the weighted average interest rate on loans issued by rival banks. Following Carbó *et al*. (2009), we calculate this variable as a weighted average of the market size of the provinces where the bank has branches. The relative importance of each province to that bank's branch network is used to weight this average. For bank *i* in year *t*, $SD15_{it}$ is the weighted average interest rate on deposits made in rival banks. $SD16_{it}$ captures the branch networks of the rivals of bank *i* in year *t*. Variables $SD15_{it}$ and $SD16_{it}$ are measured for each bank using a similar weighting system as the one used for $SD14_{it}$ to characterize the market where bank *i* operates.

We also adopt the theory-based empirical approach proposed by Dranove *et al*. (1998). Under this approach, a strategic group exists if the performance of a firm is a function of group characteristics, after controlling for firm and industry characteristics. Therefore, after identifying the strategic groups, we measure differences across strategic groups using performance variables (consistent with Ebbes *et al*. (2010), and DeSarbo & Grewal (2008), among others). We use these differences to externally validate the identification of strategic groups because groups should also differ in terms of other variables besides those used in the estimation method (Dant & Gundlach 1999). To assess bank performance and the effectiveness of bank strategies, we use the following variables: $SD17_{it}$: loans to total assets; $SD18_{it}$: deposits to total assets; $SD19_{it}$: loans per employee; $SD20_{it}$: deposits per employee; $SD21_{it}$: Lerner index or net profit margin on loans; $SD22_{it}$: Lerner index or net profit margin on deposits; $SD23_{it}$: marginal operating costs of loans; and $SD24_{it}$: marginal operating costs of deposits. Because a bank's marginal cost is also part of the Lerner indices, we consider a translog function (Mas-Ruiz & Ruiz-Moreno 2017). This functional form is common in the analysis of banking markets because it deals with economies of scale and scope in multiproduct firms. This function has two inputs (labor and physical capital) and two outputs (loans and deposits). Its estimations allow us to calculate the marginal operating costs for every bank and the Lerner indices for loans and deposits. The AEB (*Asociación Española de Banca*), and the CECA (*Confederación Española de Cajas de Ahorros*) publish the necessary information to compute these variables.

4. RESULTS AND DISCUSSION

The study covers 38 banks, 13 strategic dimensions, 3 variables for measuring competition, and 8 performance variables based on annual data for the period 1992 to 2004. The data are arranged in a 494 by 24 matrix. We standardize the strategic dimensions used in the estimation procedure to ensure that the same scale is used. Data processing is performed using Matlab. We estimate the proposed model for K = 2, …,10 strategic groups in the banking industry. The model selection criterion is the Bayesian information criterion (BIC) (Schwarz 1978). The BIC suggests that K = 2 is the number of groups that best fits the data. As described earlier, to estimate the parameters of the probability distribution, we use Monte Carlo simulation with 10,000 steps, discarding the first 10% (burn-in period) to ensure the stability of the results. Table 1 shows the classification of banks for 2004.

Table 1. Classification of banks into two strategic groups for 2004

Strategic Group #1	Strategic Group #2	
	Banco Pastor	
	Bancaja	Caja de Burgos
	Banco de Andalucía	Caja Duero
	Banco de Valencia	Caja España
	Banco Guipuzcoano	Caja General de Ahorros de Granada
	Banco Popular Español	Caja Murcia
Banco Bilbao Vizcaya BBVA	Banco Sabadell	Caja San Fernando
Banco Santander	Banco Urquijo	Cajastur
Caja Madrid	Barclays Bank	CajaSur
La Caixa	Bilbao Bizkaia Kutxa	Caja Castilla La Mancha
Banesto	Caja Ahorros Inmaculada	Deutsche Bank SAE
Citibank España	Caixa Catalunya	El Monte
	Caixa Galicia	Ibercaja
	Caixanova	Kutxa
	Caixapenedes	Sa Nostra
	Caja de Ahorros de Navarra	Unicaja
	Caja de Ahorros del Mediterraneo	

Because we analyze the composition of both strategic groups for the period 1992 to 2004, we are able to quantitatively identify lines of continuity for banks. Thus, we can identify the group that the bank belonged to for the majority of the study period. Banco Bilbao Vizcaya BBVA, Banco Santander, Caja Madrid, and La Caixa were included in strategic group 1 (SG1) for much of the study period. Banesto started in strategic group 2 (SG2), but switched to SG1 in 1998 and remained in SG1, except in 2001. The aforementioned five entities in SG1 are also the largest in the banking industry. Finally, Citibank belonged primarily to SG2 although it sporadically moved to SG1 (in 2001 and 2004). Between 2002 and 2004, some banks (e.g., Banco Sabadell, Banco Urquijo, Barclays Bank, Caixapenedes, and Cajastur) temporarily belonged to SG1. Switching between clusters does not seem random. In fact, patterns emerge. To measure this behavior we use transition matrices shown in Table 2. Therefore, Table 2 shows the probability that a bank switches from strategic group 1 or 2, in rows, to strategic group 1 or 2, in columns. Each sub-table is identified by the initial year and the final year. The percentages on the diagonal therefore reflect the probability that a bank remains in the same group. Row probabilities add up to one, meaning that a bank that belongs to group i $(i = 1,2)$ can only stay in the same group or switch to the other group. Notably, according to Eq. (5), the results are an average of the last samples draws from the posterior distribution, Eq. (2). Thus, the entries in the transition matrices are the average of the last 9,000 sample points for each yearly transition.

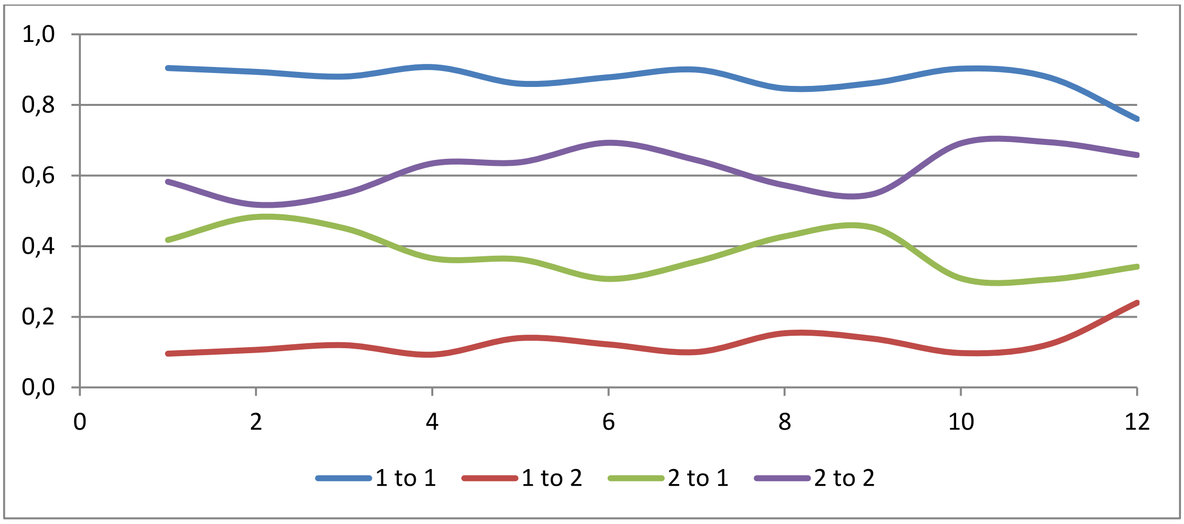
Table 2. Transition matrices between groups

Year	Strategic group	
1992–1993	**1**	**2**
1	90.44%	9.56%
2	41.76%	58.24%
1993–1994	**1**	**2**
1	89.36%	10.64%
2	48.27%	51.73%
1994–1995	**1**	**2**
1	88.04%	11.96%
2	45.08%	54.92%
1995–1996	**1**	**2**
1	90.71%	9.29%
2	36.57%	63.43%
1996–1997	**1**	**2**
1	86.02%	13.98%
2	36.22%	63.78%
1997–1998	**1**	**2**
1	87.83%	12.17%
2	30.72%	69.28%
Average	**1**	**2**
1	**87.28%**	**12.72%**
2	**38.16%**	**61.84%**

Year	Strategic group	
1998–1999	**1**	**2**
1	89.98%	10.02%
2	35.65%	64.35%
1999–2000	**1**	**2**
1	84.66%	15.34%
2	42.77%	57.23%
2000–2001	**1**	**2**
1	86.21%	13.79%
2	45.26%	54.74%
2001–2002	**1**	**2**
1	90.26%	9.74%
2	30.88%	69.12%
2002–2003	**1**	**2**
1	87.79%	12.21%
2	30.56%	69.44%
2003–2004	**1**	**2**
1	76.03%	23.97%
2	34.19%	65.81%

The last matrix, named *Average*, summarizes the information for all periods. It can be considered the transition matrix for the whole data period. It reflects the high probability of remaining in the same group. The probability of remaining in SG1 is 87.28%, and the probability of remaining in SG2 is 61.84%. The *Average* matrix also reflects the low probability of switching groups. The probability of switching from SG1 to SG2 is 12.72%, and the probability of switching from SG2 to SG1 is 38.16%. Additionally, Figure 1 shows the transition probabilities over time. The blue and red lines denote the probability of starting in SG1 and remaining in SG1 or switching to SG2, respectively. At any given time, the sum of the corresponding points on both lines is equal to 1. The behavior of both probability lines reflects the stability over time. Only for the last period does the probability of switching from SG1 to SG2 increase to almost 24%.

Figure 1. Evolution of transition probabilities



The purple and green lines show the probability of starting in SG2 and remaining in SG2 or switching to SG1, respectively. Although banks are more likely to stay in SG2 once they are there, the way that these curves evolve is complex. Indeed, there is a high probability of switching to SG1 in periods 2 and 9 and of staying in SG2 in other periods. As before, at any given time, the sum of the corresponding points on both lines is equal to 1. This behavior suggests that banks in SG2 might evolve for long periods in terms of group membership. For example, banks in SG2 were more likely to switch to SG1 between 1997 and 2001, which corresponds to an economic recovery period in which mergers were taking place in the Spanish banking market. This transformation in the banking industry offered a new opportunity for banks in SG2 to modify their strategies in response to the new competitive environment. In contrast, banks in SG1 appear to be more stable over time, with low levels of switching behavior and well-defined long-term behavior.

After aggregating over the entire period, we present the descriptive statistics for the strategic dimensions (unstandardized form) used in the identification of the strategic groups shown in Table 3. The information is presented in general terms and by strategic group.

Table 3. Summary statistics for the two strategic groups in the Spanish banking sector

Variable	Overall mean	Overall SD	SG	Mean	SD
$SD1_{it}$: quantity of loans	12,544,999	21,869,379	SG1	57,678,799	39,274,379
			SG2	7,005,850	8,339,919
$SD2_{it}$: interest rate for loans	0.101	0.040	SG1	0.092	0.035
			SG2	0.103	0.041
$SD3_{it}$: quantity of deposits	11,582,853	19,719,999	SG1	51,898,118	36,036,869
			SG2	6,635,071	7,447,541
$SD4_{it}$: interest rate for deposits	0.044	0.029	SG1	0.050	0.039
			SG2	0.043	0.027
$SD5_{it}$: price of labor	45.051	8.883	SG1	52.449	8.00
			SG2	44.143	8.563
$SD6_{it}$: price of capital	0.494	0.972	SG1	0.879	2.516
			SG2	0.447	0.526
$SD7_{it}$: operating costs	566,633	1,033,656	SG1	2,719,365	1,695,634
			SG2	302,434	461,247
$SD8_{it}$: number of ATMs	759.50	1,157.49	SG1	3,171.96	2,018.54
			SG2	463.43	457.08
$SD9_{it}$: assets/employees	3,488.97	1,594.26	SG1	4,978.65	1,926.45
			SG2	3,306.14	1,449.07
$SD10_{it}$: regional GDP	27,297,294	21,369,153	SG1	42,885,908	19,122,516
			SG2	25,384,145	20,859,144
$SD11_{it}$: Herfindahl index	0.170	0.034	SG1	0.157	0.012
			SG2	0.172	0.036
$SD12_{it}$: branch networks	661.45	802.36	SG1	2,397.41	1,153.30
			SG2	448.40	382.38
$SD13_{it}$: advertising index	0.321	0.059	SG1	0.314	0.075
			SG2	0.322	0.057

Table 3 provides statistics for the variables for each group. The results indicate that the means and standard deviations are well separated because of a clear classification of the banks. The results also indicate major differences between the strategic groups for many of the variables that were used for their identification. The values of strategic dimensions such as quantity of loans and deposits, operating costs, number of ATMs, assets per employee, and branch networks are larger for SG1 than for SG2. These findings support the assertion at the beginning of this section about the size of the banks in SG1. In fact, deregulation and the institutional structure in the Spanish banking industry mean that size is a defining characteristic of this market. Hence, certain scholars (e.g., Freixas 1996; Más-Ruiz *et al*. 2005) have used size to identify strategic groups. The strategic dimension of regional GDP indicates that banks in SG1 have a national scope and a strong presence in the major cities. Many banks that belong to SG2 have a stronger presence in a few regional markets, but only a few symbolic branches in the rest of the national market. Some other strategic dimensions such as the Herfindahl index and the prices of labor and capital also reveal major differences between SG1 and SG2. Finally, minor differences emerge for some strategic dimensions such as advertising ratio and interest rates on loans and deposits.

We also examine some variables (SD14–SD16) that reflect competition within each group and some bank performance variables (SD17–SD24) that measure differences between groups to provide external validity and support this classification of strategic groups. These variables appear in Table 4. We use ANOVA to detect significant differences. The results in Table 4 do not show significant differences in rivals' interest rates on loans and deposits. These variables measure the prices for loans and deposits over a 13-years period. Differences might be diluted over such a long period because price wars in the loans and deposits markets are more common over short periods. However, the variable that indicates the competition using the strategic variable of the branch network reflects a significant difference between groups. Thus, the branch structure of the Spanish banking industry is the result of rivalry in a context of regulated prices; the broadest branch networks belong to the largest firms, which can provide their clients with the most convenient service. In fact, large Spanish banks achieve economies of scale by increasing their number of accounts per branch. Product differentiation is inherent to market leadership because most consumers apply the "what the majority buys must be the best" selection criterion to reduce the costs of searching for the best deal, which constitutes a major advantage (Rhoades 1985).

Table 4. Competition and performance variables

Variable	Total mean	Total SD	Sig.	By SG	Mean	SD
Competition variables						
$SD14_{it}$: rivals' interest rates for loans	0.096	0.039	0.939	SG1	0.095	0.039
				SG2	0.096	0.039
$SD15_{it}$: rivals' interest rates for deposits	0.047	0.030	0.352	SG1	0.044	0.029
				SG2	0.048	0.030
$SD16_{it}$: rivals' branch network	1,119.54	786.99	**<.0001**	SG1	1,555.60	423.32
				SG2	1,066.02	804.76
Performance variables						
$SD17_{it}$: loans to total assets	0.751	0.233	**0.035**	SG1	0.688	0.078
				SG2	0.759	0.245
$SD18_{it}$: deposits to total assets	0.722	0.269	**0.0002**	SG1	0.592	0.145
				SG2	0.738	0.276
$SD19_{it}$: loans per employee	2,564.11	1,156.02	**<.0001**	SG1	3,394.65	1,262.34
				SG2	2,462.18	1,101.42
$SD20_{it}$: deposits per employee	2,464.94	1,179.92	**0.0005**	SG1	2,988.69	1,333.94
				SG2	2,400.66	1,144.88
$SD21_{it}$: loans Lerner index	0.209	0.188	0.543	SG1	0.194	0.201
				SG2	0.211	0.187
$SD22_{it}$: deposits Lerner index	-0.192	0.434	**<.0001**	SG1	-0.802	0.460
				SG2	-0.117	0.368
$SD23_{it}$: loans marginal costs	0.023	0.014	**0.0642**	SG1	0.027	0.018
				SG2	0.022	0.014
$SD24_{it}$: deposits marginal costs	0.031	0.025	**<.0001**	SG1	0.046	0.039
				SG2	0.029	0.023

The results highlight significant differences between groups for most performance variables used to externally validate the classification. Almost all performance variables are statistically different between the groups. The largest banks in SG1 perform better than banks in SG2 with respect to variables that relate to the quantity of loans and deposits per employee. These variables are indicators of efficiency and economies of scale in the largest banks. However, banks in SG1 have lower ratios of loans and deposits to total assets than do banks in SG2. Despite significant differences in the estimated marginal operating costs of loans and deposits, banks in SG1 seemingly have higher marginal operating costs, especially in deposits. It is also notable that the Lerner index in deposits is negative but significantly different between banks in SG1 and SG2. Although deposits are not a profitable product in themselves, they allow banks to capture and retain customers. Through this tying arrangement, deposits allow banks to exercise market power in the loan market (Carbó *et al*. 2009). Finally, these performance differences among groups support our approach and the approach by Dranove *et al*. (1998), according to which strategic groups exist if there are performance differences between these groups.

4.1. Results of model comparison

The optimal number of strategic groups, k = 2, for the HMM implemented by Ebbes *et al.* (2010) matches the number for our extended model (hereinafter HMM2). For the cluster analysis, the optimal number of groups cannot be inferred by this method. Therefore, we estimate this method for k = 2 groups to enable direct comparison between the three methods.

The resulting grouping is similar but not identical for the three approaches. Both dynamic models, HMM and HMM2, group banks on a stable basis throughout the time period and yield the same classification of banks for the two groups in terms of the median. However, the classification using the k-means algorithm is more inconsistent along the analyzed period. In this classification, Caja Madrid is not part of SG1. Table 5 summarizes the results of the classification of banks into strategic groups based on the estimation of the three models: cluster analysis, HMM, and the HMM2 model proposed in this paper.

Table 5. Comparison of classification of banks into strategic groups (in terms of median in the time period)

Bank	**Cluster analysis**	**HMM**	**HMM2**
Banco Bilbao Vizcaya BBVA	SG1	SG1	SG1
Banco Santander	SG1	SG1	SG1
Caja Madrid	SG2	SG1	SG1
La Caixa	SG1	SG1	SG1
Other banks	SG2	SG2	SG2

Having obtained the optimized log-likelihoods, we use two criteria to compare the models. We employ Akaike's information criterion (AIC) (Akaike, 1974) and the Bayesian information criterion (BIC) (Schwarz, 1978). Table 6 shows the results of the pairwise comparison of the different models.

Table 6. Model comparison

Fitted models	**Cluster**	**HMM**	**Cluster**	**HMM2**	**HMM**	**HMM2**
Log-likelihood	-866.6	7698.7	-866.6	7314.8	7698.7	7314.8
AIC	6543.1	-10587.3	6543.1	-9819.5	-10587.3	-9819.5
BIC	31567.3	14436.8	31567.3	15204.7	14436.8	15204.7

Both dynamic models, HMM and HMM2, fit the data better than the cluster analysis. This observation is supported by the values for the AIC and BIC. Regarding the comparison of the two dynamic models, although the HMM has the best fit to the Spanish banking data, this difference is small, and the log-likelihood values are quite similar. As expected, both dynamic models fit the data significantly better than the cluster analysis and perform similarly (see AIC and BIC). However, the enhanced version presented in this paper explicitly models the transition matrix for every pair of years (*t* and *t+1*) within the study period instead of providing a unique general transition matrix for the whole period. This model enhancement improves the set of available tools to enable more accurate analysis of the dynamics of strategic groups.

5. CONCLUSIONS

One of the primary goals of the recent strategic group research is to develop theoretical and methodological techniques to improve the identification of group dynamics over time. We are aware of the importance of studying the dynamics of strategic groups to understand how competitive dynamics change over time (Ebbes *et al*. 2010). We therefore propose a hidden Markov model (HMM) to examine the dynamics of strategic groups for a given period. We estimate the switching behavior of banks that move from one group to another on a yearly basis to show the dynamic strategic group structure of the Spanish banking industry. The primary contribution of this empirical research is to model the development of the strategic groups using a time inhomogeneous HMM, for which the time-variable transition matrix captures institutions' group switching behavior for every pair of years (*t* and *t+1*) included in the study period. We allow for time-varying parameters in the distributions, thereby accounting for the dynamics of the strategic groups. From a technical perspective, our prior initial values are more general than those of Ebbes *et al*. (2010). This aspect adds flexibility when controlling the amount of information included in the priors.

Two strategic groups in the Spanish banking industry are detected. These groups are primarily characterized by size and other strategic dimensions. Results show that banks in SG1 are highly likely to remain in this group and that banks in SG2 are less likely to do so. We also detect a low probability of switching groups: approximately 12.72% probability of switching from SG1 to SG2 and 38.16% probability of switching from SG2 to SG1. Thus, banks in SG1 appear to be more stable over time, with low transients and well-defined long-term behavior. In contrast, the group membership of banks in SG2 might evolve over long periods.

One limitation of this research is the need for a longer study period to use more years of data to establish a converging path or trend. The importance and role of each strategic dimension in the identification of strategic groups is unknown. It would therefore be useful to determine the role of each strategic dimension to help the bank improve its general strategy. Another limitation of this study relates to the study setting. The study is based on Spanish data. It should be extended to global markets to enhance the robustness of the conclusions. On the model side, further studies should vary the number of clusters from year to year to allow for different numbers of strategic groups over time. Allowing for a different number of groups each year would help with group formation to determine the degree of fragmentation of the market at a given time. Also, the model would benefit from a balance between parsimony and a better fit to the data. The variables could be selected dynamically, in the same process of convergence, and the algorithm could detect the key variables to determine the strategic groups and eliminate the variables that are not used for such a purpose. Finally, although this model captures the way in which strategic groups change over time in response to changes in market conditions, greater consideration should be given to the interplay between firm, group, and industry characteristics in the analysis of strategic group dynamics (Schimmer & Brauer 2012). Therefore, the study of how firm characteristics can explain dynamic changes would be an interesting avenue for future research.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6), 716-723.

Amel, D. F., & Rhoades, S. A. (1989). Empirical evidence on the motives for bank mergers. *Eastern Economic Journal*, 15(1), 17–27.

Andrieu, C., De Freitas, N., Doucet, A., & Jordan, M. I. (2003). An introduction to MCMC for machine learning. *Machine learning*, 50(1–2), 5–43.

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, 17(1), 99–120.

Barros, F., & Modesto, L. (1999). Portuguese banking sector: a mixed oligopoly? *International Journal of Industrial Organization*, 17(6), 869–886.

Bensmail, H., Celeux, G., Raftery, A. E., & Robert, C. P. (1997). Inference in model-based cluster analysis. *Statistics and Computing*, 7(1), 1–10.

Berg, S. A., & Kim, M. (1994). Oligopolistic interdependence and the structure of production in banking: an empirical evaluation. *Journal of Money, Credit and Banking*, 26(2), 309–322.

Boeker, W. (1991). Organizational strategy: An ecological perspective. *Academy of Management Journal*, 34(3), 13–635.

Bogner, W., & Thomas, H. (1994). *Core competence and competitive advantage: a model and illustrative evidence from the pharmaceutical industry.* In G. Hamel, & A. Heene (Eds.), *Competence-Based Competition* (pp. 111–144). Chichester: John Wiley.

Burke, L. (1990). *Geographic scope strategy, market structure and interfirm rivalry: strategic groups in California banking, 1978-1985.* PhD Dissertation, University of California, Berkeley.

Carbó, S., Fernández, J., Humphrey, D., & Maudos, J. (2009). Estimating the intensity of price and non-price in banking. *Banks and Bank Systems*, 4(2).

Carbó, S., López, R., & Rodríguez, F. (2003). Medición de la competencia en los mercados bancarios regionales [Measuring competition in regional banking markets]. *Revista de Economía Aplicada* 32 (vol.XI).

Casella, G., Robert, C. P., & Wells, M. T. (2004). Mixture models, latent variables and partitioned importance sampling. *Statistical Methodology*, 1(1), 1-18.

Caves, R., & Porter, M. (1977). From entry barriers to mobility barriers. *The Quarterly Journal of Economics*, 91(2), 241–261.

Cool, K. O., & Schendel, D. (1987). Strategic group formation and performance: The case of the US pharmaceutical industry, 1963–1982. *Management Science*, 33(9), 1102–1124.

Cool, K. O., & Schendel, D. (1988). Performance differences among strategic group members. *Strategic Management Journal*, 9(3), 207–223.

Cool, K., & Dierickx, I. (1993). Rivalry, strategic groups and firm profitability. *Strategic Management Journal*, 14(1), 47–59.

Cowles, M. K., & Carlin, B. P. (1996). Markov chain Monte Carlo convergence diagnostics: a comparative review. *Journal of the American Statistical Association*, 91(434), 883–904.

Dant, R. P., & Gundlach, G. T. (1999). The challenge of autonomy and dependence in franchised channels of distribution. *Journal of Business Venturing*, 14(1), 35–67.

Day, D. L., Lewin, A. Y., & Li, H. (1995). Strategic leaders or strategic groups: A longitudinal data envelopment analysis of the U.S. brewing industry. *European Journal of Operational Research*, 80(3), 619–638.

Desarbo, W. S., & Grewal, R. (2008). Hybrid strategic groups. *Strategic Management Journal*, 29(3), 293–317.

DeSarbo, W. S., Grewal, R., & Wang R. (2009). Dynamic strategic groups: deriving spatial evolutionary paths. *Strategic Management Journal*, 30, 1420–1439.

DeSarbo, W. S., Grewal, R., Hwang, H., & Wang, Q. (2008). The simultaneous identification of strategic/performance groups and underlying dimensions for assessing an industry's competitive structure. *Journal of Modelling in Management*, *3*(3), 220-248.

DeSarbo, W. S., Wang, Q., & Blanchard, S. J. (2010). Exploring intra-industry competitive heterogeneity: The identification of latent competitive groups. *Journal of Modelling in Management*, *5*(2), 94-123.

Dranove, D., Peterat, M., & Shanley, M. (1998). Do strategic groups exist? An economic framework for analysis. *Strategic Management Journal*, 19, 1029–1044.

Ebbes, P., Grewal, R., & DeSarbo, W. S. (2010). Modeling strategic group dynamics: A hidden Markov approach. *Quantitative Marketing and Economics*, 8(2), 241–274.

Epure, M., Kerstens, K., & Prior, D. (2011). Bank productivity and performance groups: A decomposition approach based upon the Luenberger productivity indicator, *European Journal of Operational Research*, 211(3), 630–641.

Fiegenbaum, A., & Thomas, H. (1990). Strategic groups and performance: the US insurance industry, 1970–84. *Strategic Management Journal*, 11(3), 197–215.

Fiegenbaum, A., & Thomas, H. (1993). Industry and strategic group dynamics: competitive strategy in the insurance industry, 1970–84. *Journal of Management Studies*, 30(1), 69–105.

Fiegenbaum, A., Sudharshan, D., & Thomas, H. (1987). The concept of stable strategic time periods in strategic group research. *Managerial and Decision Economics*, 8(2), 139-148.

Fiegenbaum, A., Sudharshan, D., & Thomas, H. (1990). Strategic time periods and strategic groups research: Concepts and an empirical example. *Journal of Management Studies*, 27(2), 133–148.

Flegal, J.M., & Gong, L. (2013). Relative fixed-width stopping rules for Markov chain Monte Carlo simulations. *arXiv preprint arXiv*:1303.0238.

Frazier, G. L., & Howell, R. D. (1983). Business definition and performance. *The Journal of Marketing*, 59-67.

Freixas X. (1996). *Los Límites de la Competencia en la Banca Española [Limits of rivalry in the Spanish banking]*. Fundación BBV, Madrid.

Frühwirth-Schnatter, S. (2006). *Finite mixture and Markov switching models*. New York: Springer-Verlag.

García, N., Polo, Y., & Urquizu, P. (1998), Relación Entre el Nivel de Servicios y Tipo de Interés [The relationship between the service standards and interest rates]. *Estudios Empresariales*, 96, 34–45.

García-Cestona, M., & Surroca, J. (2008). Multiple goals and ownership structure: Effects on the performance of Spanish savings banks, *European Journal of Operational Research*, 187(2), 582–599.

Gimeno, J., & Woo, C. Y. (1996). Hypercompetition in a multimarket environment. The role of strategic similarity and multimarket contact on competitive de-escalation. *Organization Science*, 7, 322–341.

González-Moreno, Á., & Sáez-Martínez, F. J. (2008). Rivalry and strategic groups: what makes a company a rival? *Journal of Management & Governance*, 12(3), 261–285.

Gual, J. (1992). La competencia en el sector bancario español [Competition in the Spanish banking sector]. Fundación BBVA, Bilbao.

Hunt, M. S. (1972). *Competition in the major home appliance industry, 1960-1970.* Unpublished PhD dissertation, Harvard University.

Ketchen, D. J., & Shook, C. L. (1996). The application of cluster analysis in strategic management research: an analysis and critique. *Strategic Management Journal*, 17(6), 441–458.

Ketchen, D. J., Snow, C. C., & Hoover, V. L. (2004). Research on competitive dynamics: Recent accomplishments and future challenges. *Journal of Management*, 30(6), 779-804.

Lago, R., & Salas, V. (2005). Market power and bank interest rate adjustments. *Documentos de Trabajo del Banco de España* (0539).

Leask, G., & Parker, D. (2007). Strategic groups, competitive groups and performance within the UK pharmaceutical industry: Improving our understanding of the competitive process. *Strategic Management Journal*, 28(7), 723-745.

Lewis, P., & Thomas, H. (1990). The linkage between strategy, strategic groups, and performance in the UK retail grocery industry. *Strategic Management Journal*, 11(5), 385-397.

Martín, M., & Sáez, F. J. (1997). Desregulación e intensificación de la competencia. El impacto del mercado único bancario sobre la competitividad del sistema bancario español [Deregulation and the intensification of competition. The impact of the single banking market on competition in the Spanish banking system]. *Perspectivas del Sistema Financiero,* 59, 3–20.

Mascarenhas, B. (1989). Strategic group dynamics. *Academy of Management Journal*, 32(2), 333–352.

Mascarenhas, B., & Aaker, D. A. (1989). Mobility barriers and strategic groups. *Strategic Management Journal*, 10(5), 475-485.

Mas-Ruiz, F. J., & Ruiz-Moreno, F. (2011). Rivalry within strategic groups and consequences for performance: The firm-size effects. *Strategic Management Journal*, 32(12): 1286–1308.

Mas-Ruiz, F. J., & Ruiz-Moreno, F. (2017). How Strategic Groups Act Competitively Within and Across Markets. *Managerial and Decision Economics*, 38(7), 1017-1032.

Más-Ruiz, F. J., Nicolau-Gonzálbez, J. L., & Ruiz-Moreno, F. (2005). Asymmetric rivalry between strategic groups: Response, speed of response and ex ante vs. ex post competitive interaction in the Spanish bank deposit market. *Strategic Management Journal*, 26(8), 713-745.

Mas-Ruiz, F. J., Ruiz-Moreno, F., & Ladrón de Guevara, A. (2014). Asymmetric rivalry within and between strategic groups. *Strategic Management Journal*, 35(3), 419–439.

Maudos, J. (2001). Rentabilidad, estructura de mercado y eficiencia en la banca [Profitability, market structure and banking efficiency]. *Revista de Economía Aplicada*, 9(25), 193–207.

McGee, J., & Thomas, H. (1986). Strategic groups: Theory, research and taxonomy. *Strategic Management Journal*, 7, 141–160.

Murthi, B. P. S., Rasheed, A. A., & Goll, I. (2013). An empirical analysis of strategic groups in the airline industry using latent class regressions. *Managerial and Decision Economics*, 34(2), 59–73.

Nair, A., & Filer, L. (2003). Cointegration of firm strategies within groups: a long-run analysis of firm behavior in the Japanese steel industry. *Strategic Management Journal*, 24(2), 145-159.

Nair, A., & Kotha, S. (2001). Does group membership matter? Evidence from the Japanese steel industry. *Strategic Management Journal*, 22(3), 221-235.

Peteraf, M. (1993). Intra-industry structure and the response toward rivals. *Managerial and Decision Economics*, 14, 519–528.

Peteraf, M., & Shanley, M. (1997). Getting to know you: a theory of strategic group identity. *Strategic Management Journal*, 18(S1), 165–186.

Porter M. (1976). *Interbrand choice, strategy, and bilateral market power.* Cambridge, MA: Harvard University Press.

Porter M. (1979). The structure within industries and companies' performance. *The Review of Economics and Statistics*, 61, 214–227.

Ray, S. C., & Das, A. (2010). Distribution of Cost and Profit Efficiency: Evidence from Indian Banking. *European Journal of Operational Research*, 201(1): 297–307.

Rebiere, P., & Mavoori, H. (2016). French Pharmaceutical strategic clusters: A hybrid inducto-deductive framework for visual analysis of competitive spaces. *Strategic* Change, 25(4), 343-360.

Rhoades, S. A. (1985). Market share as a source of market power: Implications and some evidence. *Journal of Economics Business*, 37, 343–363.

Roberts, G. O., & Smith, A. F. (1994). Simple conditions for the convergence of the Gibbs sampler and Metropolis-Hastings algorithms. *Stochastic Processes and Their Applications*, 49(2), 207–216.

Ruíz, F. J. M. (1999). Dynamic analysis of competition in marketing: strategic groups in Spanish banking. *International Journal of Bank Marketing*, 17(5), 233–250.

Schimmer, M., & Brauer, M. (2012). Firm performance and aspiration levels as determinants of a firm’s strategic repositioning within strategic group structures. *Strategic Organization*, 10(4), 406-435.

Schwarz, G. E. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6(2), 461–464. doi:10.1214/aos/1176344136. MR468014

Smith, K. G., Grimm, C. M., & Wally, S. (1997). Strategic groups and rivalrous firm behavior: towards a reconciliation. *Strategic Management Journal*, 18(2), 149–157.

Spiller, P. T, & Favaro, E. (1984). The effects of entry regulation on oligopolistic interaction: the Uruguayan banking sector. *Rand Journal of Economics*, 15(2), 244–254.

Stephens, M. (1997). *Bayesian Methods for Mixtures of Normal Distributions.* Unpublished PhD thesis, University of Oxford, Oxford.

Tang, M. J., & Thomas, H. (1992). The concept of strategic groups: Theoretical construct or analytical convenience. *Managerial and Decision Economics*, 13(4), 323–329.

Thomas, H., & Venkatraman, N. (1988). Research on strategic groups: Progress and prognosis. *Journal of Management Studies*, 25(6), 537-555.

Tuma, M. N., Decker, R. & Scholz, S. W. (2011). A survey of the challenges and pitfalls of cluster analysis application in market segmentation. *International Journal of Market Research*, 53(3), 391-414.

Young, D. S. (2008). An overview of mixture models. *Statistics Surveys*, 0, 1–24. *arXiv preprint arXiv*:0808.0383.

Zúñiga-Vicente, J. Á., de la Fuente-Sabaté, J. M., & Suarez Gonzalez, I. (2004). Dynamics of the strategic group membership–performance linkage in rapidly changing environments. *Journal of Business Research*, 57(12), 1378–1390.