Mapping Market Structure Evolution

Maximilian Matthe (Goethe University Frankfurt, Germany)

matthe@wiwi.uni-frankfurt.de

Daniel M. Ringel (University of North Carolina at Chapel Hill, USA)

dmr@kenan-flagler.unc.edu

Bernd Skiera (Goethe University Frankfurt, Germany)

skiera@wiwi.uni-frankfurt.de

Available Online Companions

Interactive Browser Tool: www.evomap.io

Python Package: https://github.com/mpmatthe/evomap

Mapping Market Structure Evolution

Abstract

A common element of market structure analysis is the spatial representation of firms' competitive positions on maps. Such maps typically capture static snapshots in time. Yet, competitive positions tend to change. Embedded in such changes are firms' trajectories, that is, the series of changes in firms' positions over time relative to all other firms in a market. Identifying these trajectories contributes to market structure analysis by providing a forward-looking perspective on competition, revealing firms' (re)positioning strategies and indicating strategy effectiveness. To unlock these insights, we propose EvoMap, a novel dynamic mapping framework that identifies firms' trajectories from high-frequency and potentially noisy data. We validate EvoMap via extensive simulations and apply it empirically to study the trajectories of more than 1,000 publicly listed firms over 20 years. We find substantial changes in several firms' positioning strategies, including Apple, Walmart, and Capital One. Because EvoMap accommodates a wide range of mapping methods, analysts can easily apply it in other empirical settings and to data from various sources.

1 INTRODUCTION

Firms need to understand the competitive structure of their market to develop effective strategies and create sustainable competitive advantages (Hunt 1983; Rao and Sabavala 1986). Such knowledge allows managers to assess how well their offerings are positioned, develop clear positioning strategies, and identify critical elements of their strategies (Lilien and Rangaswamy 2004). Beyond positioning, market structure analysis informs new product development, product policy, competitive advertising, and pricing strategies (Urban et al. 1984). Moreover, broader strategic considerations such as defining a firm's business, assessing threats and opportunities, and allocating resources also require a profound understanding of competition in a market (Day et al. 1979). Hence, market structure analysis is an essential ingredient in the strategic market-planning process (Day 1984).

A common element of market structure analysis is the derivation of a market structure map, that is, a spatial representation of firms' competitive positions relative to one another based on some measure of their competitive relationships (DeSarbo et al. 1993). Market structure maps visually summarize large amounts of information, which facilitates decision-making, helps managers think more strategically, and provides them with explorative tools to identify competitors, discover submarkets, and guide positioning and differentiation decisions (Smelcer and Carmel 1997; Lilien and Rangaswamy 2004).

With the increasing availability of extensive data on firm positioning and consumer perception, market structure maps' informative potential continues to grow. Recent studies in marketing leverage various data sources such as consumer search, shopping baskets, product reviews, and social media engagement to create market structure maps (Kim et al. 2011; Lee and Bradlow 2011; Netzer et al. 2012; Tirunillai and Tellis 2014; France and Ghose 2016; Ringel and Skiera 2016; Gabel et al. 2019; Liu et al. 2020; Yang et al. 2021). These maps are typically static; that is, they typically represent an isolated snapshot of a market's structure in a single period. Yet, changes in consumers' needs and preferences, new competitors' entries, or incumbents' repositioning lead to evolving rather than static market structures (Lambkin and Day 1989; D'Aveni 1994; Ning and Villas-Boas 2021).

Recognizing that market structures tend to evolve rather than remain static (Elrod et al. 2002), earlier marketing research proposed approaches to examine market structures across more than one period (DeSarbo and Carroll 1985; Moore and Winer 1987; Cooper 1988; Mela et al. 1998). However, the application of these

earlier approaches in the more recent marketing literature is rare — even though many studies emphasize the potential value of monitoring the evolution of market structure over time (e.g., Kim et al. 2011; Lee and Bradlow 2011; Netzer et al. 2012; Ringel and Skiera 2016; Wedel and Kannan 2016; Gabel et al. 2019). A possible reason might be that, as we later show, mapping the evolution of market structure in today's large and fragmented markets is not easily possible with existing approaches.

We suggest that mapping the evolution of market structure across multiple periods provides additional insights beyond those of single snapshots in time. These additional insights, elaborated in what follows, emerge from what we call firms' trajectories—the evolutionary paths of firms' positions over time relative to all other firms in a market (hereafter referred to as "firms' trajectories" for brevity). In contrast to merely describing the change in firms' positions from one snapshot of market structure to another, trajectories capture underlying trends in competitors' positions over multiple periods. The joint examination of firms' trajectories in a market reveals the evolution of its structure and offers a unique forward-looking perspective on how market structure might look in the future.

Specifically, we suggest that firms' trajectories can contribute the following insights to market structure analysis:

(1) **A forward-looking perspective on competition:** By comparing other firms' trajectories with their own firm's trajectory, managers can better anticipate their competitive situation in the future. While converging trajectories indicate competitive threats, diverging trajectories point to an increase in differentiation. Likewise, new entrants' trajectories reveal where they seek to position themselves in the market, thereby enabling incumbents to (a) better qualify the competitive threat that these new entrants might pose and (b) do so at an early stage. Finally, a more gradual convergence of multiple firms' trajectories away from their previous competitors can point to a new submarket's emergence.

(2) **A lens on competitors' strategic intentions:** Sudden changes in other firms' trajectories can point to repositioning efforts of those firms. By monitoring other firms' trajectories, managers can detect (or see the impact of) such repositioning, which they may need to account for in their own strategy formulation.

(3) **An indicator of strategy effectiveness:** Any change in a firm's trajectory should ideally align with the firm's strategic objectives. Therefore, monitoring their firm's trajectory enables managers to better

evaluate the effectiveness of their marketing strategy. Moreover, doing so helps to assess the impact of marketing actions, such as (re-) positioning efforts or marketing mix decisions (i.e., whether or not they "moved the needle").

Consider, for example, a scenario where a change in strategy prompts a firm's management to reposition the firm (e.g., change the firm's name, introduce a new logo, or adjust the marketing mix). Such changes are common, as recent examples from the tech sector illustrate (e.g., Odeo / Twitter, Netflix, or Facebook / Meta). Because repositioning usually requires substantial investments, managers must monitor the impact and effectiveness of their efforts. Likewise, investors likely want to evaluate whether the current management achieves the desired transition to a new position. Firms' trajectories can provide the desired information about the strategy change's effectiveness, as they reveal how the firm's position changes relative to its former competitors (i.e., its previous position) and its new competitors (i.e., its targeted position).

Moreover, the firm needs to redefine its competitive set. As repositioning efforts of multiple firms can coincide, the firm may face both new and former competitors after successfully repositioning itself. Firms' trajectories can help managers anticipate their future competitive situation by providing a forward-looking perspective on competition. Thereby, they can help inform corrective actions where necessary. For example, some of Meta's (previously Facebook) strongest competitors (e.g., Google) are also positioning themselves in the virtual reality space (e.g., the metaverse). At the same time, firms from other parts of the market, such as Gaming developers (e.g., Epic Games) and hardware manufacturers (e.g., Sony), are pushing into the realm of virtual reality. While there is little that firms can do about other firms' trajectories, these trajectories can alert them to emerging threats (i.e., unanticipated competition) and opportunities (e.g., potential collaborators) at an early stage.

This article aims to (1) provide a novel dynamic mapping framework that reveals firms' trajectories and (2) illustrate the additional market structure insights that these trajectories create. Identifying firms' trajectories is, however, not trivial. While the marketing literature previously proposed several approaches for mapping changes in market structure, we find that these approaches are unfit to identify firms' trajectories. Common to extant approaches is the idea of independently generating a sequence of static maps and connecting them so that changes in a market's structure become apparent. By inspecting the change of firms' positions from one static map to the next, marketers sought to learn how the market under analysis evolved. However, such a

simple approach is limited in its capacity to accurately reveal market structure evolution because it (1) tends to generate a sequence of maps that are misaligned, (2) fails to uncover trends that persist over multiple periods, and (3) is sensitive to even low levels of noise in the data.

Specifically, a sequence of independently generated maps tends to be misaligned for two reasons. First, the coordinate systems of subsequent maps can change arbitrarily (Schönemann and Carroll 1970). Second, modern mapping methods typically do not have a unique solution but can approximate any given data in multiple ways. Thus, they can produce a sequence of substantially different maps when applied sequentially. Consequently, changes in the absolute positions of firms, or even in the relative distances between them, can be arbitrary. So, they do not necessarily correspond to actual changes in market structure. To mitigate the misalignment problem, marketers traditionally resort to Procrustes Analysis (Gower 1975). Procrustes Analysis tries to align multiple maps after their estimation by transforming their coordinate systems (e.g., rotations and reflections). Procrustes Analysis can be an effective approach when the sequence of maps covers only a few periods and when one uses linear mapping methods such as metric ratio Multidimensional Scaling (MDS) or centroid scaling (see, for instance, Moore and Winer 1987; Borg and Groenen 2005). However, as we later show, Procrustes Analysis works less well with modern (nonlinear) mapping methods commonly used in contemporary market structure research (e.g., Ringel and Skiera 2016; Gabel et al. 2019; Yang et al. 2021) or when the analysis covers many periods.

Furthermore, changes in firms' positions across multiple (more than two) successive maps can be erratic when one generates each map independently; that is, they can move back and forth instead of evolving continuously. Such erratic changes in map positions thwart any attempt to faithfully reveal persistent trends in firms' competitive positions, that is, their trajectories.

Finally, because even tiny changes in the underlying relationship structure can affect each individual map, a sequence of maps is very sensitive to noise in the data. This sensitivity to noise exacerbates successive maps' misalignment, curtails the revelation of trends, and may compromise the accurate representation of firms' actual relationships on each map.

The limitations of previous approaches for mapping market structure evolution result from the fact that a map inherently simplifies a set of competitive relationships based on a particular set of criteria—none of which covers the alignment of firms' positions across multiple maps or the identification of persistent trends over

multiple periods.

Herein we propose a novel dynamic mapping framework called EvoMap, which overcomes previous approaches' limitations and is the first to accurately reveal firms' trajectories in evolving market structures. EvoMap can accommodate a wide range of static mapping methods, such that one can easily apply it to various empirical settings and various kinds of data. The basic idea of EvoMap is to combine the estimation of multiple successive maps into a joint optimization problem. Thereby, we introduce additional objectives to achieve alignment, persistence, and noise cancelation. We then merge the estimated sequence of maps into a single dynamic market structure map that reveals each firm's trajectory over time relative to all other firms' trajectories.

In this article, we implement EvoMap for a classic psychometric method (metric MDS), a nonlinear variant (Sammon Mapping; Sammon 1969), as well as a more recent innovation in computer science (t-SNE; Maaten and Hinton 2008). In an extensive simulation study, we validate that EvoMap accurately recovers market structure evolution across various possible market structures and different evolution patterns. Our simulation shows that EvoMap outperforms all previous approaches in its capacity to reveal accurate underlying trends in the changes in firms' positions. Our findings are consistent across all three static mapping methods (i.e., metric MDS, Sammon Mapping, and t-SNE); that is, the superior performance of EvoMap is not contingent on a particular mapping method.

We use EvoMap (paired with t-SNE) to empirically investigate the evolution of product market competition among publicly listed firms over two decades. Our analysis uses the Text-Industry Network Classification (TNIC) data that Hoberg and Phillips (2016) provide. These data capture how firms position themselves relative to other firms based on product descriptions in their annual report to the U.S. Securities and Exchange Commission (called Form 10-K). Using EvoMap, we identify the trajectories of 1,092 firms across 20 years and confirm the face validity of the revealed trajectories. Our analysis identifies individual firms' repositioning efforts and more global changes in market structure, such as the convergence of entire industries.

The remainder of this article is structured as follows: The following section embeds this article into the extant literature on dynamic market structure mapping and details the limitations of previous approaches. We then formally introduce our proposed framework. Subsequently, we validate and benchmark it against

alternative mapping approaches in an extensive simulation study. Finally, we use our framework to empirically investigate the evolution of product market competition based on the TNIC data. We close with a discussion on implications, limitations, and directions for future research.

2   EXTANT APPROACHES AND THEIR LIMITATIONS

This article contributes to the literature on market structure analysis, specifically to the literature stream that seeks to capture complex networks of competitive relationships among entities of interest (e.g., firms, brands, or products) in spatial representations (maps). Without loss of generality, we limit the subsequent text to "firms" as the object of analysis.

2.1   Background on Mapping Market Structure

Studies that analyze market structure in maps usually adopt one of two approaches to represent the firms in a market: (1) as vectors in higher-dimensional space or (2) as nodes in a network (Wei 2020). The vector space approach models firms as vectors of d dimensions, where each of the d dimensions corresponds to one of the firms' (potentially latent) attributes. One can then measure firms' competitive relationships as pairwise dissimilarities or similarities between these vectors. Alternatively, the network approach models firms as nodes connected by edges that represent competitive relationships (measured as pairwise similarities). In both approaches, the basis for deriving a market structure map is a square matrix of pairwise competitive relationships (hereafter referred to as a competitive relationship matrix).[1] Typically, the matrix is symmetric (or symmetrized as part of the mapping process), but methods exist for asymmetric relationships (DeSarbo et al. 2006).

Given a competitive relationship matrix $Y \in \mathbb{R}_0^{+ \, n \times n}$, static mapping methods fit a map—defined as a configuration of the set of firms I in lower-dimensional space $X = \left\{ x_i = \left(x_i^{(1)}, \dots, x_i^{(d_{map})} \right) \in \mathbb{R}^{d_{map}} \middle| i \in I \right\}, X \in \mathbb{R}^{n \times d_{map}}, \ n = |I|$—to the competitive relationship matrix Y as well as possible. The resultant configuration is usually two-dimensional to ease interpretation. Therefore, we also focus on two-dimensional maps (i.e., $d_{map} = 2$). Yet, one- or three-dimensional applications also

---

[1] Note that there also exist alternatives to such relationship-based methods, which use attribute ratings as input data (such as factor analysis). These input data typically consist of customer judgements on a set of known attributes. The data sources described in the introduction and used in contemporary studies, however, typically offer measures of relationships, rather than attributes. Therefore, we limit our discussion to relationship-based methods.

exist. Firms with strong competitive relationships appear close together on the map, whereas firms with weak competitive relationships appear more distant.

Commonly used mapping methods process either dissimilarity matrices (as in the case of MDS) or similarity matrices (as in the case of force-directed drawing, Fruchterman and Reingold 1991). While some mapping methods use higher-dimensional vector data as their input, they eventually transform it internally into a relationship structure (e.g., by taking distances). Both classes of methods share similar underlying concepts, and each representation translates easily into the other (e.g., by taking the inverse). Therefore, we do not differentiate between similarities or dissimilarities but refer to weak competitive relationships (low similarities / high dissimilarities) or strong competitive relationships (high similarities / low dissimilarities).

Earlier studies of market structure largely measured competitive relationships from surveys (Sabnis and Grewal 2012). However, more recent marketing studies draw on a range of alternative data sources to measure competitive relationships. These new types of data—which include data from online search (Kim et al. 2011), online forum discussions (Netzer et al. 2012), customer reviews (Tirunillai and Tellis 2014; France and Ghose 2016), clickstreams (Ringel and Skiera 2016), co-follower networks in social media (Culotta and Cutler 2016), social tags (Nam et al. 2017), firm-related images (Liu et al. 2020), or user engagement in social media (Yang et al. 2021)—are often regularly retrievable over time in high frequency. Taken together, the abundance, accessibility, and high frequency of data on competitive relationships provide new opportunities to study market structure evolution and to create insights beyond those of single snapshots in time.

2.2 Marketing Approaches for Mapping Market Structure Evolution

Initial approaches in marketing for analyzing market structure in maps over time typically followed one of two approaches: (I) fixing the coordinate system ex-ante, or (II) aligning mapping solutions ex-post.

2.2.1 Approach I: Fixing the Coordinate System Ex-Ante

The first approach fixes two dimensions as the map's coordinate system. For example, one might decide to use the attributes "performance" and "ease-of-use" for the two dimensions of the coordinate system. One can then plot each firm's position on these two attributes at each period to show how firms' positions evolved. In other words, one describes the evolution of a market's structure using a time-indexed scatter plot along two fixed dimensions. Examples from marketing research that apply this approach include Tirunillai and Tellis
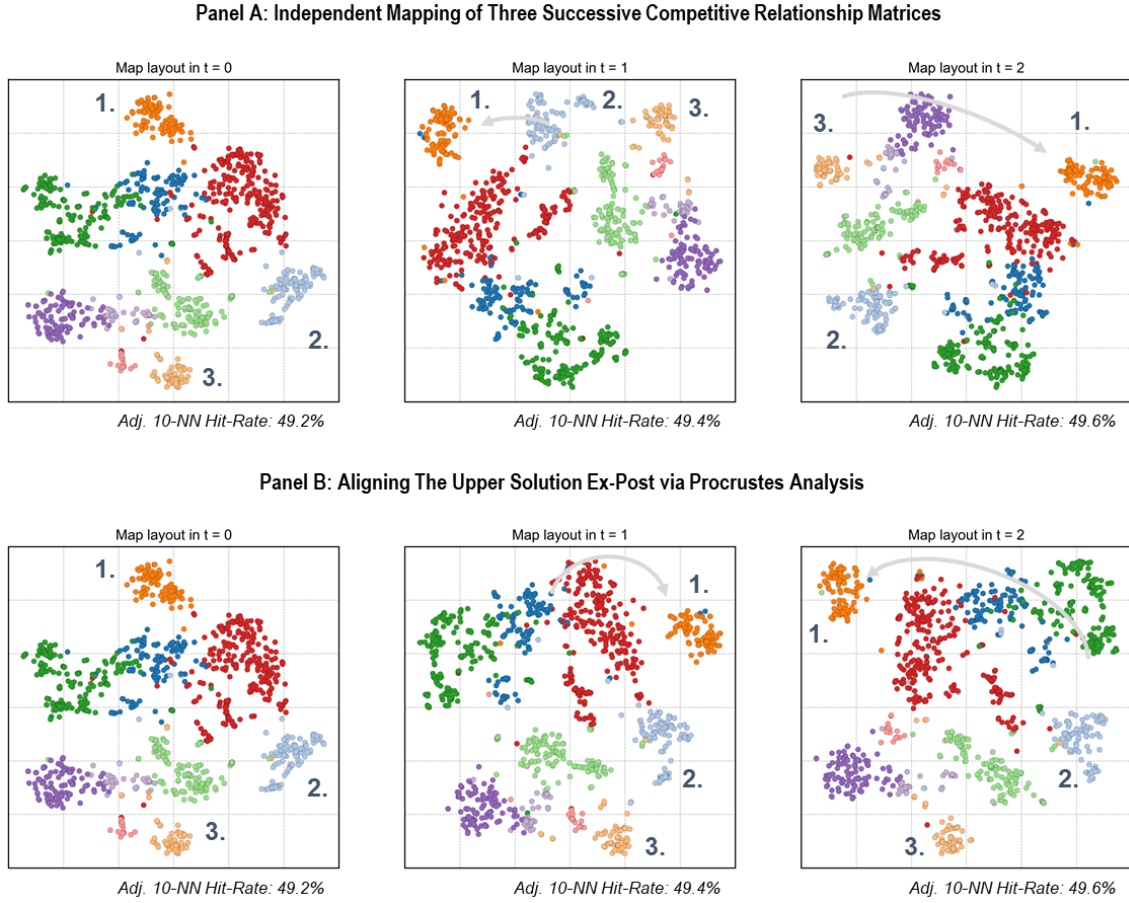
(2014) and van Heerde et al. (2004).

The limitation of this approach is that one needs to assume that the chosen two attributes sufficiently explain all relevant differences between firms over time. To relax this limitation, market researchers can resort to three-way MDS. Three-way MDS is an extension of MDS applicable whenever multiple relationship matrices for the same objects (e.g., firms) are available (for instance, one for each of several respondents in a survey). In the context of market structure analysis, three-way MDS can be used to estimate two latent dimensions and firms' positions on them from multiple relationship matrices (Carroll and Chang 1970; DeSarbo and Carroll 1985; Cooper 1988; Mela et al. 1998). However, suppose these matrices represent successive periods. In such a case, three-way MDS assumes that the rank order of firms' positions remains the same on each dimension (and merely the importance of each dimension changes over time). Moreover, the two dimensions themselves need to remain constant over time. Yet, these assumptions might be challenging to justify in more complex and dynamic markets, where maps can be particularly insightful.

2.2.2   Approach II: Aligning Mapping Solutions Ex-Post

For the second approach, one independently applies a static mapping method (e.g., MDS or t-SNE) at each period. The result is a sequence of individual maps that are not necessarily aligned (see Panel A of Figure 1 for an example). One reason for such misalignment is that the maps are oriented differently (i.e., their coordinate system differs), as any mapping solution is unique only up to a linear transformation of its coordinate system (such as rotation or reflection).

To resolve the misalignment of independently generated maps, one can employ Procrustes Analysis (e.g., Moore and Winer 1987), which identifies linear transformations (i.e., translation, rotation, reflection, or scaling) to align successive maps as well as possible (Schönemann 1966; Schönemann and Carroll 1970). Procrustes Analysis tends to work well for methods that provide a linear mapping between the given relationship data and the proximities on the estimated map, such as metric ratio MDS (Borg and Groenen 2005). However, as shown in Panel B of Figure 1, Procrustes Analysis works less well for modern nonlinear mapping methods (e.g., t-SNE). While the successive maps in Panel B are more similar than in Panel A, substantial differences remain. In particular, the erratic shifts in cluster 1's position remain even after applying Procrustes Analysis.

Figure 1: Illustration of Ex-Post Alignment via Procrustes Analysis

Panel A: Independent Mapping of Three Successive Competitive Relationship Matrices

Map layout in t = 0	Map layout in t = 1	Map layout in t = 2
Adj. 10-NN Hit-Rate: 49.2%	*Adj. 10-NN Hit-Rate: 49.4%*	*Adj. 10-NN Hit-Rate: 49.6%*

Panel B: Aligning The Upper Solution Ex-Post via Procrustes Analysis

Map layout in t = 0	Map layout in t = 1	Map layout in t = 2
Adj. 10-NN Hit-Rate: 49.2%	*Adj. 10-NN Hit-Rate: 49.4%*	*Adj. 10-NN Hit-Rate: 49.6%*

Notes: Three maps fitted independently to three competitive relationship matrices using t-SNE (Panel A), aligned ex-post via Procrustes Analysis (Panel B). Coloring based on cluster assignments identified within the competitive relationship matrices using Louvain community detection by Blondel et al. (2008). Static mapping quality measured by Adj. 10-NN Hit-Rate (10 Nearest Neighbor Hit-Rate adjusted for random agreement (Chen and Buja 2009)). The arrow belonging to cluster 1 highlights the maps' misalignment.

One reason for such erratic shifts is that it is usually impossible to perfectly preserve an entire matrix of pairwise competitive relationships on a two-dimensional plane. In other words, a map is always an approximation of the underlying relationship matrix (Tversky and Hutchinson 1986). When the utilized mapping method's objective function has multiple local optima, different maps approximate the same competitive relationship matrix equally well. Thus, repeatedly applying static mapping methods tends to result in a sequence of misaligned maps.

As Figure 1 shows, such misalignment is particularly prevalent for modern nonlinear mapping methods (e.g., t-SNE) typically used when the number of objects (i.e., firms) or clusters (i.e., submarkets) in a map is large—as is the case in many of today's large markets (e.g., Ringel and Skiera 2016; Gabel et al. 2019). These nonlinear mapping methods are more flexible in how objects can be arranged on a map (for instance, by focusing more on placing objects with strong relationships nearby rather than objects with weak relationships

9

far apart). However, this flexibility comes at the cost of allowing the method to re-arrange weakly related objects in various ways without compromising the mapping quality (i.e., many local optima exist). Because these local optima are no longer unique up to linear transformations, the entailing sequence of maps cannot be aligned ex-post via Procrustes Analysis.

Some studies suggest resolving the local optima problem by various initialization strategies, such as using the same (i.e., fixed) random initialization in each period or initializing periods sequentially with the estimated map of the previous period (e.g., Gabel et al. 2019). However, since the estimated positions remain sensitive to even minor differences across successive competitive relationship matrices, these initialization strategies can easily fall short of producing well-aligned maps.

Moreover, even when using linear mapping methods (such as metric MDS), ex-post alignment may not always lead to desirable outcomes. Specifically, by generating maps independently, one risks overfitting each map to even minor variations in the relationship data that may constitute noise. As Procrustes Analysis only changes each map's orientation, it cannot remedy such overfitting. The resultant overfitting to each period makes it hard to reveal persistent trends and can impair the accurate representation of the market's structure. Moreover, Procrustes Analysis does not tie the alignment over time to changes in the input data. It merely attempts to make maps similar over time without considering which firms changed their relationships with other firms and, thus, their positions (and to what extent). In sum, these extant approaches are subject to limitations that make them less suitable for many of today's large, fragmented, and rapidly evolving markets.

2.3 Other Approaches for Mapping Evolving Relationship Data

The problem of mapping evolving relationships among objects is not unique to market structure analysis. Outside of the marketing literature, researchers proposed alternative approaches for visualizing changes in relationship data over time. For instance, in network visualization, Xu et al. (2013) propose increasing the stability of successive maps by preserving objects' positions as much as possible over time through a regularization scheme. Rauber et al. (2016) later apply the idea of Xu et al. (2013) to t-SNE. They show that their "Dynamic t-SNE" increases objects' stability in pairwise comparisons of subsequent snapshots in time relative to independently generated maps.

However, map stability alone is not sufficient when the objective is to study market structure evolution. Specifically, a very stable solution is undesirable when parts (or individual firms) of the market undergo

substantial changes while other parts (or individual firms) do not. Ideally, any firm's map position should be as stable as possible across a sequence of maps as long as there are no substantial changes in that firm's actual position (as indicated by its competitive relationships with other firms in the market). Moreover, Dynamic t-SNE merely links objects' positions across two successive periods. Yet, as we will show in an extensive simulation study, doing so still falls short of revealing persistent trends.

2.4 Implications for Market Research

The limitations of extant approaches for mapping market structure evolution (i.e., misalignment, lack of persistence, sensitivity to noise) curtail market researchers' abilities to study market structure evolution on multiple fronts:

First, when successive maps are misaligned, firms' positions change erratically (e.g., a group of firms 'jumping' to another place). As a result, it becomes impossible to identify firms' trajectories and characterize their directions and lengths (e.g., to differentiate firms whose competitive relationships changed substantially versus those whose competitive relationships did not change).

Second, faithfully identifying underlying trends in firms' positions (for instance, to extrapolate them into the future) is not easily possible when changes are not persistent but revert. Both sequences of maps in Figure 1 exhibit highly non-persistent changes: The average correlation of successive changes in firms' positions on these maps is -0.46 and -0.54, respectively, which indicates that any move of a firm into one direction of the map tends to be followed by a move into the opposite direction.

Third, the high sensitivity of individual maps to changes in the input data makes inference of actual market structure evolution harder. When, for instance, two firms move closer together (further apart) on a map, it is unclear whether this convergence (divergence) is due to (1) actual changes in the firms' positions, (2) noise in the data, or (3) different local optima (or a combination thereof). This problem remains for approaches that constrain the movement of all firms uniformly, as these approaches risk constraining the movement of some firms more (or less) than what is justified by the underlying data.

We overcome these limitations of past approaches with a novel dynamic mapping framework called EvoMap. In contrast to past approaches in the marketing literature, EvoMap jointly estimates a sequence of maps from a sequence of competitive relationship matrices instead of attempting to align maps ex-post. EvoMap thereby not only ensures stability in successive maps as proposed by Xu et al. (2013), but explicitly

11

accounts for heterogeneity in how strongly firms change their positions.

To this end, EvoMap uses a novel adaptive regularization scheme, which adjusts constraints for each firm based on the data at hand. To reveal persistent trends in the evolution of firms' positions, EvoMap additionally smooths the estimated trajectories across multiple periods by imposing further constraints on them. As a result, the estimated trajectories do not overfit individual periods (which would result in erratic oscillation), but reveal smooth underlying trends that ease map exploration and allow to extrapolate into the future. This "noise cancellation" property of EvoMap makes it particularly suitable for high-frequency data sources that are likely subject to a substantial amount of noise, as in the case of, for example, relationships derived from user-generated content or social media. Finally, in contrast to recent approaches (i.e., Rauber et al. 2016), we design EvoMap as a flexible framework that can easily accommodate various static mapping methods (e.g., MDS, Sammon Mapping, t-SNE, or methods not yet developed).

Table 1 summarizes the key differences between EvoMap and extant approaches for mapping market structure evolution.

Table 1: Alternative Approaches for Mapping Market Structure Evolution

Approach	Description	Various Mapping Methods	Alignment	Joint Estimation	Adaptive Regularization	Smoothing
Independent Mapping	Repeated independent application of a static mapping method	Yes	-	-	-	-
Fixed Initialization	Independent mapping using the same (i.e., a fixed) initialization	Yes	-	-	-	-
Sequential Initialization	Independent mapping, using the map in $t-1$ as initialization in t	Yes	-	-	-	-
Ex-post Alignment *(Gower 1975)*	Alignment of independent maps via Procrustes Analysis	Yes	Yes	-	-	-
Dynamic t-SNE *(Rauber et al. 2016)*	t-SNE aligned during optimization	-	Yes	Yes	-	-
EvoMap *(this article)*	Multiobjective optimization framework, implemented for Metric MDS, Sammon Mapping, and t-SNE	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**

3 FORMAL DESCRIPTION OF EVOMAP

Let $Y \in \mathbb{R}_0^{+\,n \times n}$ denote a competitive relationship matrix representing the market structure for the set of firms I with $n = |I|$. The entries of Y consist of symmetric measures of pairwise competitive relationships. We do not impose any restrictions on how these measures are derived. That is, they could represent measures of substitutability, brand switching, co-occurrence, or distances in a high-dimensional embedding space (among many others). We also do not restrict the type of relationships (i.e., distances/dissimilarities vs. similarities), which depends upon the chosen static mapping method.

Static mapping methods estimate the map $\hat{X} \in \mathbb{R}^{n \times 2}$ for any given competitive relationship matrix Y by minimizing the cost function $C(X): \mathbb{R}^{n \times 2} \to \mathbb{R}_0^+$, which measures the discrepancy between firms' competitive relationships in the matrix Y and the relative proximities of their positions on a map X:

$$\hat{X} = \underset{X \in \mathbb{R}^{n \times 2}}{\arg \min}\, C(X) \tag{1}$$

Note that different cost functions C give rise to different mapping methods. We assume that C is a non-negative loss function such that smaller values correspond to better solutions. For instance, C could represent Kruskal's

13

Stress function (for MDS) or Kullback-Leibler divergence (for t-SNE). If solution quality *increases* with increasing values of C or if C takes negative values (e.g., for a log-likelihood function), one first needs to transform C to use it within our framework (e.g., invert it).

Applied to a sequence of competitive relationships matrices $(Y_t)_{t=1,...,T}$, static mapping methods derive a sequence of maps $(\hat{X}_t)_{t=1,...,T}$ such that each map $\hat{X}_t$ at time t preserves the competitive relationships in Y_t as well as possible. In addition, we pursue the following three objectives:

1) Successive maps should be aligned. That is, the total changes in positions on successive maps $\|\hat{X}_t - \hat{X}_{t-1}\|, t = 2, ..., T$ should be small.

2) Changes in firms' positions should persist across successive periods; that is, their trajectories should reveal gradual underlying trends rather than oscillate back and forth. Specifically, successive movements $(\nabla \hat{X}_t)_{t=2,...,T}$, where $\nabla \hat{X}_t = \hat{X}_t - \hat{X}_{t-1}$, should exhibit non-negative serial correlation.

3) Changes in firms' positions on the map should reflect actual changes in market structure. That is, any firm's trajectory length on the map should be related to the degree of change in its actual position (as indicated by its competitive relationships).

Meeting one objective might come at the cost of another. Consider a sequence of relationship matrices subject to some degree of change. An utterly stable sequence of maps $\hat{X}_1 = \hat{X}_2 = \cdots = \hat{X}_T$, for instance, aligns successive maps perfectly but reduces the goodness-of-fit of each map. Moreover, the absence of changes on the map does not reflect the existing changes in the competitive relationship matrices. Thus, we cannot expect to meet all objectives perfectly. We, therefore, balance these objectives in the following joint optimization problem.

3.1 Optimization Problem for EvoMap

The basis for EvoMap is a static mapping method such as MDS, Sammon Mapping, or t-SNE. We design EvoMap as a flexible framework to accommodate various static mapping methods. This flexibility is important for marketing research because it allows researchers to select the mapping method that is most suitable for their specific purpose, empirical setting, and data source. Notably, EvoMap's flexibility opens it to future advances in static mapping methods.

To meet the three additional mapping objectives outlined in the previous subsection, we proceed as follows:

Rather than fitting each map independently to the respective competitive relationship matrix for all $t \in \{1, \dots T\}$, we fit the sequence of maps jointly to the sequence of competitive relationship matrices. Doing so allows us to (1) incorporate information about maps in successive periods into the estimation of each map and to (2) add additional objectives.

Formally, we derive the sequence of maps $\left(\hat{X}_t\right)_{t=1,\dots,T}$ as

$$\left(\hat{X}_t\right)_{t=1,\dots,T} = \underset{X_1,\dots,X_T \in \mathbb{R}^{n \times 2}}{\arg\min} C_{total}(X_1,\dots,X_T), \tag{2}$$

where

$$C_{total}(X_1,\dots,X_T) = \sum_{t=1}^{T} C_{static}(X_t) + \alpha \cdot C_{temporal}(X_1,\dots,X_T) \tag{3}$$

In equation (3), the first term represents the static component of EvoMap's cost function, which equals the sum of the cost function C_{static} of the selected static mapping method evaluated at each period $t \in \{1,\dots,T\}$. This static component of EvoMap's cost function seeks to preserve the competitive relationships in Y_t on the map X_t as closely as possible for every $t \in \{1,\dots,T\}$. The second term represents the temporal component of the cost function, which seeks to meet the three additional objectives outlined before. The hyperparameter $\alpha \in \mathbb{R}_0^+$ balances the relative importance of the cost functions' static vs. temporal component. We specify $C_{temporal}$ as follows:

$$C_{temporal}(X_1,\dots,X_T) = \sum_{i=1}^{n} f_w(i) \sum_{k=1}^{p} \sum_{t=k+1}^{T} 1_{[i \in I_{t,k}]} \left\| \nabla^k x_{i,t} \right\|^2, \tag{4}$$

where

$1_{[i \in I_{t,k}]} : I \to \{0,1\}$ denotes an indicator function equal to one if firm i is present in time t and the k preceding periods: $1_{[i \in I_{t,k}]} := \prod_{l=0}^{k} 1_{[i \in I_{t-l}]}$, where $1_{[i \in I_t]} := \begin{cases} 1 & if & i \in I_t \\ 0 & & i \notin I_t \end{cases}$,

$f_w : I \to \mathbb{R}^+$ is a positive weight function defined on the set of firms I,

$p \in \mathbb{N}$ is a second hyperparameter of positive integer values that controls the degree of smoothing, and

$\nabla^k x_{i,t} \in \mathbb{R}^2$ is the k-th (backward) difference of firm i's map position at time t:

$$\nabla^k x_{i,t} := \nabla^{k-1} x_{i,t} - \nabla^{k-1} x_{i,t-1} \qquad (5)$$

For $k = 0$, $\nabla^k x_{i,t}$ simply is firm i's position at time t:

$$\nabla^0 x_{i,t} := x_{i,t} \qquad (6)$$

For $k = 1$, $\nabla^k x_{i,t}$ equals the difference vector between firm i's positions in t and $t - 1$:

$$\nabla^1 x_{i,t} = x_{i,t} - x_{i,t-1} \qquad (7)$$

For $k > 1$, $\nabla^k x_{i,t}$ represents a higher-order difference vector (e.g., for $k = 2$, it corresponds to the difference in differences of successive positions, analogous to the acceleration). We outline the underlying intuition of the terms in equation (4) in what follows. We also provide an illustrative example of its calculation together with exemplary trajectories in Online Appendix B.

First, to align successive maps, we link firms' positions in successive periods in our cost function. Specifically, we penalize all changes between successive maps, measured as the squared Euclidean norm $\left\|\nabla^1 x_{i,t}\right\|^2$ of the first differences between each firm's positions in t and $t - 1$, summed over all firms[2]. Doing so penalizes large changes in firms' positions over time and thereby aligns successive maps.

To achieve our second aim (identifying trajectories that capture underlying trends), the penalty term $\left\|\nabla^1 x_{i,t}\right\|^2$ alone is insufficient as merely two successive periods impact each firm's position in a given period. As a result, $\left\|\nabla^1 x_{i,t}\right\|^2$ would, for instance, penalize oscillation of firms' positions equally strong as a smooth trajectory into one direction, as long as the total length of the movement path is the same (for a concrete example, see Online Appendix B). Therefore, we extend the penalty term such that more than two successive periods impact each firm's position in a given period. To do so, we penalize higher-order differences between any firm's successive positions $\left\|\nabla^k x_{i,t}\right\|^2$. For $k = 2$, for instance, this term penalizes large differences in differences (analogous to the acceleration of a firm's moves on the map). Effectively, these penalty terms

---

[2] The squared Euclidean norm is a natural choice because it is differentiable everywhere. In principle, non-Euclidean metrics, such as the ℓ_1-norm, could be chosen as well. Doing so would slightly alter the gradient calculations, see Online Appendix A.

impose stronger constraints on the resultant trajectories by penalizing more complex trajectories (for instance, erratic back-and-forth movement). As a result, the estimated trajectories become less sensitive to changes in individual periods and better recover underlying trends across multiple periods. Hyperparameter p controls the degree of smoothing.

To achieve our third aim (i.e., to link changes on the map to actual changes in market structure), we further augment the cost function with the weighting function f_w. Until now, the outlined specification imposes constraints on all firms' trajectories uniformly—regardless of how much their competitive relationships with others changed. Doing so, however, harms the ability to differentiate between firms with dynamic and static positions. Specifically, uniform penalties across all firms are undesirable because firms may change their competitive relationships with other firms at different rates and to different extents. Consider, for instance, a new entrant heading towards the market's incumbents (i.e., its competitive relationships with all incumbents are strengthening). Overall, the entrant's competitive relationships are subject to substantial changes, whereas any incumbent's competitive relationships may change only marginally (only the competitive relationship with the new entrant changes substantially). Ideally, the resultant sequence of maps should capture such heterogeneity by retaining incumbents at their current positions while allowing the new entrant to change its position towards the incumbents (such that its trajectory points toward them).

The weighting function $f_w: I \rightarrow \mathbb{R}^+$ accounts for such heterogeneity across firms by determining how each firm i is affected by the previously outlined constraints. Specifically, firms whose competitive relationships changed very little should be constrained from undergoing large changes in their positions on the map (i.e., they should receive high weights). In contrast, firms with substantial changes in their competitive relationships should be free to change their positions on the map more (i.e., they should receive low weights). Thus, we specify f_w as a monotonically decreasing function of the total change in firm i's competitive relationships. Specifically:

$$f_w(i) := exp(-bz_i) \tag{8}$$

where

$$b := \frac{1}{\max(\{z_i | i \in I\})}, \text{ and}$$

$$z_i := \sum_{t=2}^{T} \|Y_{i,t} - Y_{i,t-1}\|^2.$$

Here, $Y_{i,t}$ denotes the i-th row of the relationship matrix Y_t. Thus, $\|Y_{i,t} - Y_{i,t-1}\|^2$ captures the total change in firm i's relationships with all other firms between periods $t-1$ and t. z_i captures the total change in firm i's relationships to all other firms across all periods. Note that the scale of z_i depends on the scale of the input data, that is, the entries in Y (for instance, pairwise similarities are typically bound by [0, 1], whereas pairwise distances can be much larger than 1). To account for such potential differences in the input data's scale, we use the normalizing constant b in equation (8).

Lastly, we add an indicator function $1_{[i \in I_{t,k}]}$ to include cases where not all firms are observable for the entire observation period (e.g., due to market entry or exit).

Combining equations (3) and (4), we fit the sequence of maps according to (2), where

$$C_{total}(X_1, \ldots, X_T) =$$

$$\sum_{t=1}^{T} C_{static}(X_t) + \alpha \sum_{i=1}^{n} f_w(i) \sum_{k=1}^{p} \sum_{t=k+1}^{T} 1_{[i \in I_{t,k}]} \|\nabla^k x_{i,t}\|^2 \tag{9}$$

3.2 Python Implementation of EvoMap

Different mapping methods can be superior in different empirical contexts and for different tasks (France and Akkucuk 2021). Therefore, we designed EvoMap independent of a particular static mapping method (the choice of which only determines the specification of C_{static}, its corresponding component in the gradient, and potentially its optimization technique). As outlined before, this flexibility allows market researchers to decide which static mapping method is most suitable for their specific empirical setting.

In this article, we implement EvoMap for three popular mapping methods to demonstrate that EvoMap is compatible with (I) traditional psychometric methods, (II) their nonlinear advancements, as well as (III) recent innovations in computer science:

(I)   metric MDS,

(II)   Sammon Mapping (Sammon 1969), and

(III)   t-distributed Stochastic Neighborhood Embedding, t-SNE (Maaten and Hinton 2008).

We implement EvoMap in Python, adopt each mapping method's respective static cost function, and derive the cost-minimizing solutions via iterative optimization with adaptive step sizes. Specifically, we use gradient

descent with backtracking (via step halving) for metric MDS and Sammon Mapping. For t-SNE, we use momentum-based gradient descent with early exaggeration. Further, we automatically adjust initial step sizes via exponential decay to ensure convergence. We provide the gradient derivations in Online Appendix A.

Our Python implementation draws on the NumPy library (https://numpy.org) for numerical computing and the Numba compiler (https://numba.pydata.org) for performance optimization. Our implementation is available as a Python package with this journal article; see https://github.com/mpmatthe/evomap. We provide further details on EvoMap's implementation (including the specific package versions and runtime estimates) in Online Appendix E.

3.3 Hyperparameter Selection

The hyperparameters α and p give market researchers additional control over their mapping solutions. To illustrate and quantify the hyperparameters' effects on the resulting sequence of maps, we use the following three metrics:

(I)   Static goodness-of-fit, measured as the average Hit-Rate of k-nearest neighbor recovery across all T periods and n firms, adjusted for random agreement: $\frac{1}{T}\sum_{t=1}^{T} AR_k^*$, where $AR_k^* = \frac{1}{nk}\sum_{i=1}^{n} a_{ik} - \frac{k}{n-1}$ and a_{ik} denotes the number of firm i's shared k-nearest neighbors in the data and on the map (France and Carroll 2007; Chen and Buja 2009). The choice of k depends upon the application and can be adjusted to obtain a measure of local recovery (when k is small) or global recovery (when k is large). Given that market structure analysis typically focuses on identifying close competitors (rather than remote ones), we follow Ringel and Skiera (2016) and set k to 10.

(II)   Misalignment, measured as the average movement path length across successive positions: $\frac{1}{(T-1)}\sum_{t=2}^{T}\frac{1}{n}\sum_{i=1}^{n}\lVert x_{i,t} - x_{i,t-1}\rVert$, using the Euclidean norm.

(III)   Persistence, measured as the Pearson correlation coefficient between successive changes in all firms' positions (averaged across both dimensions of the map).

Figure 2 (panel A) shows that when α is set very low, the static goodness-of-fit of each map (i.e., the average Hit-Rate) is highest. Yet, maps tend to be misaligned (panel B), and the resultant trajectories on successive maps tend to evolve erratically (panel C). As α increases, misalignment decreases (panel B), and the trajectories' persistence tends to increase (panel C), while—at some point—the static goodness-of-fit of

each map decreases considerably (panel A).

The second hyperparameter p thereby defines the highest order of differences considered in the penalty term. As such, p effectively sets the degree of smoothing by controlling how many successive periods impact each firm's map position at a given period. For $p = 1$ the penalty only considers first-order differences (corresponding to two successive periods). For $p = 2$ the penalty also considers second-order differences (corresponding to three interdependent periods) and so on (i.e., $p + 1$ periods). A higher value of p sets higher constraints, producing smoother trajectories that are less sensitive to changes in individual periods.

As evident from Figure 2, higher values of p make it cheaper (in terms of static goodness-of-fit) to achieve more persistent trajectories (panels A and C). Conditional on a specific value of α, persistence is substantially higher for increasing values of p (see panel C of Figure 2), while the corresponding decrease in static goodness-of-fit is small (see panel A of Figure 2). Hence, one can choose to trade off small decreases in static goodness-of-fit against much smoother trajectories by imposing higher constraints (i.e., setting p to a higher value).

Figure 2: Impact of Hyperparameters (α and p) on Quality of Mapping Solutions
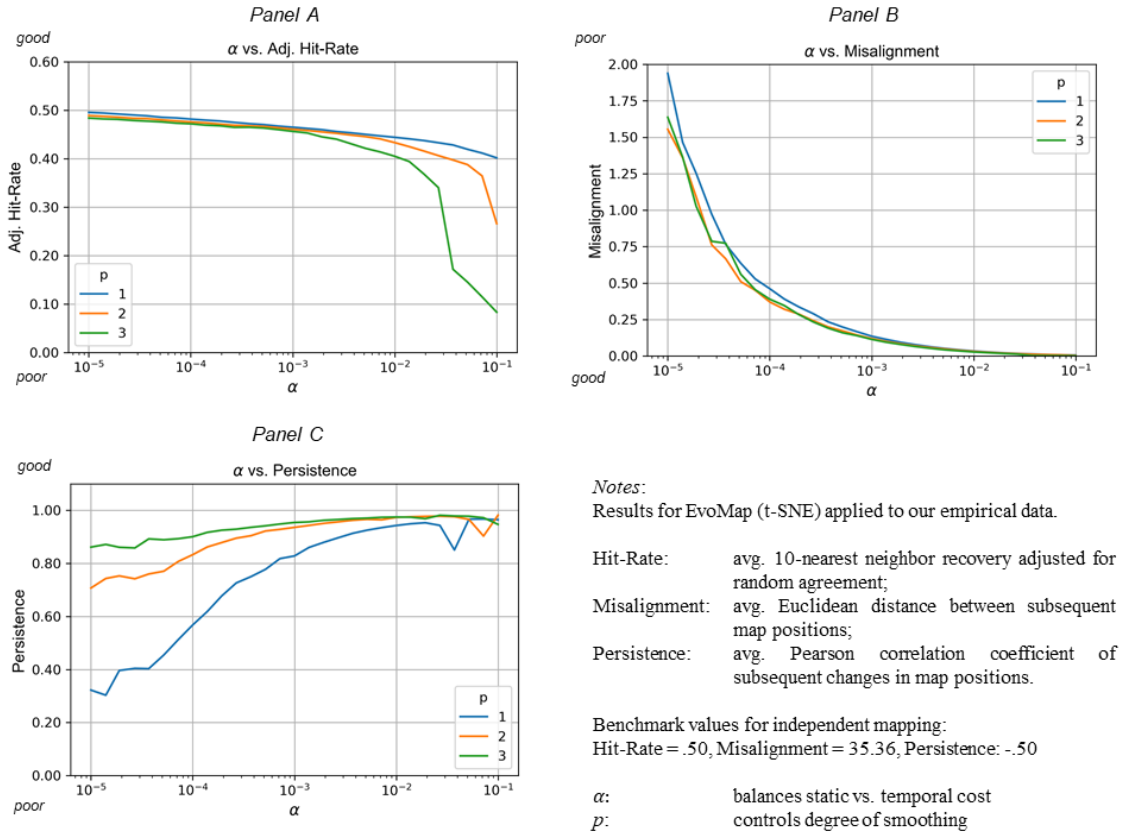
Based on the observed trade-offs, we propose the following strategy for tuning the hyperparameters α and

p. First, we select a threshold for how much static goodness-of-fit we are willing to "trade" for better dynamic mapping quality. For instance, we might set this threshold to 5% in terms of the average Hit-Rate. Then, starting with low α values (e.g., α near zero), we increase α until the static mapping quality drops below this threshold. In consequence, the entailing sequence of maps will align as strongly as possible without impairing the static mapping quality beyond the threshold. We repeat the procedure for multiple values of p (starting with $p = 1$) to obtain a set of candidate value pairs (α, p) of which we pick the one that produces the visually most appealing solution. Importantly, EvoMap is an unsupervised learning framework that is explorative. Therefore, there is no straightforward procedure for finding an optimal hyperparameter combination for each specific data set (such as cross-validation in supervised learning). We, therefore, recommend that analysts always test multiple hyperparameter combinations and fine-tune them via visual inspection.

While Figure 2 provides empirical evidence for EvoMap's ability to produce sequences of well-aligned maps with persistent firm trajectories, it is confined to a single application in which the ground truth is unknown. To obtain a more comprehensive and rigorous understanding of EvoMap's capabilities, we proceed with an extensive simulation study in the next section.
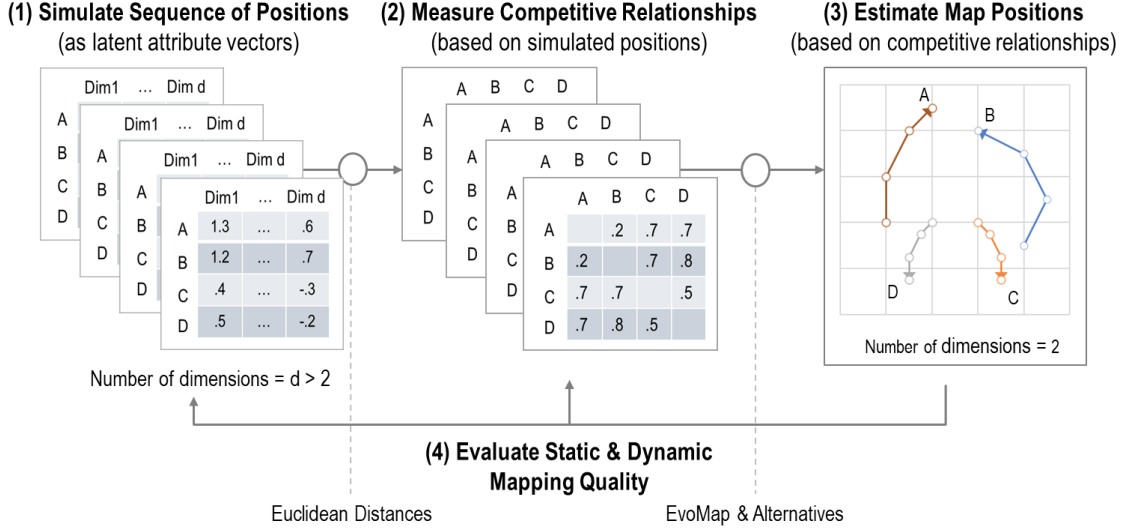
4 SIMULATION STUDY

The objectives of our simulation study are fourfold: (1) to validate EvoMap by testing its ability to accurately recover various simulated market structures' evolution, (2) to compare EvoMap with alternative dynamic mapping approaches, (3) to study EvoMap's ability to distinguish underlying trends from noise in the data (i.e., avoid overfitting), and (4) to test EvoMap's suitability as a flexible framework that can accommodate various static mapping methods.

Our simulation study comprises four steps (see Figure 3 for a schematic overview). In the first step, we simulate firms' positions as vectors in a higher-dimensional attribute space (number of dimensions $d > 2$) with pre-imposed cluster structures representing different submarkets. Next, we gradually adjust firms' simulated positions according to three evolution scenarios: (1) emergence of a new submarket, (2) shifts in individual firm's positions, and (3) market entry of new competitors. We design these three evolution scenarios to capture fundamental patterns that we expect to observe in the evolution of market structures. Our three proposed evolution scenarios' primary drivers are new consumer needs, firms' repositioning and

differentiation efforts, innovations, and new market entries. Finally, we add a varying degree of Gaussian white noise to all simulated positions to make our simulations more realistic.

Figure 3: Schematic Overview of the Four Steps of our Simulation Study

In the second step, we derive a competitive relationship matrix for each period by measuring the pairwise squared Euclidean distances among all (noisy) simulated positions in the higher-dimensional attribute space. As a result, we obtain a time-indexed sequence of competitive relationship matrices representing the simulated market structure's evolution and some degree of noise (where each sequence corresponds to one of the three evolution scenarios). In the third step, we apply EvoMap and various alternatives to these sequences of competitive relationship matrices to estimate two-dimensional representations, i.e., sequences of maps. Notably, we use noisy measurements for estimation (i.e., the relationships measured after adding noise to the simulated positions). In the fourth and final step, we evaluate each sequence of maps with regard to the desirable properties of dynamic market structure maps, namely, (1) successive maps should be aligned, (2) changes in firms' positions should persist over successive periods, and (3) changes in firms' positions should reflect actual changes in market structure. Note that we evaluate the resultant maps (estimated based on noisy measurements) against the actual market structure (that is, the simulated positions before adding any noise).
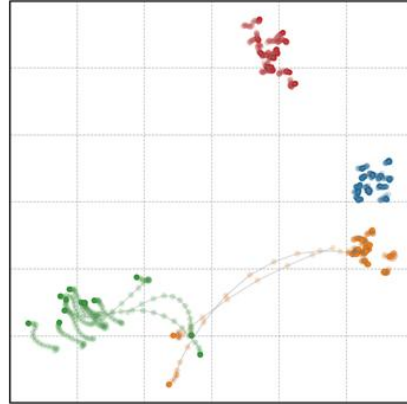
Appendix A provides a detailed description of our data-generating process. To illustrate the evolution scenarios, we generate six evolving market structures and map them with EvoMap in Figure 4.

22

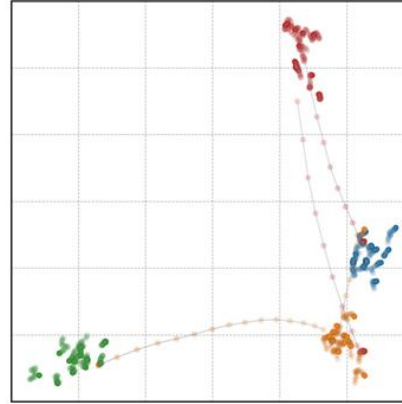**Figure 4: Examples of Simulated Evolving Market Structures**
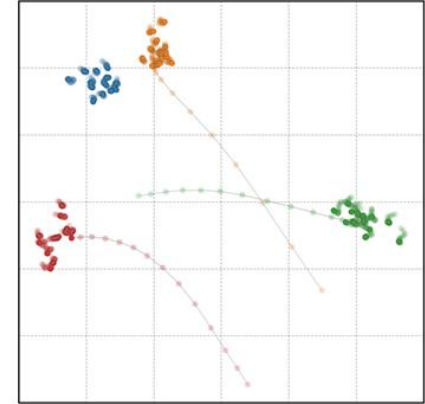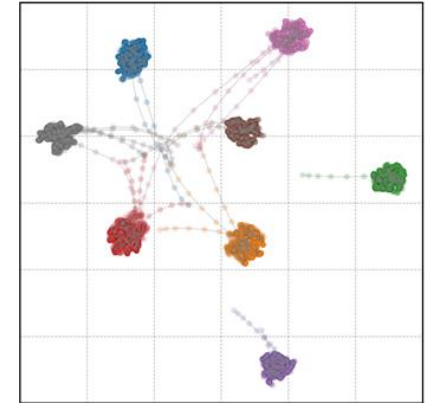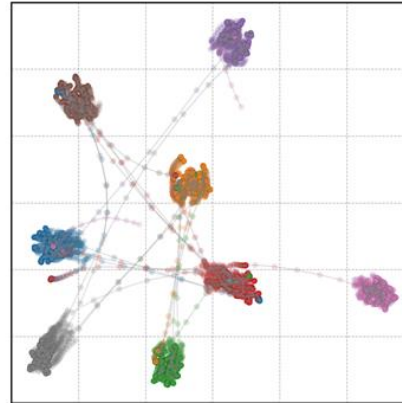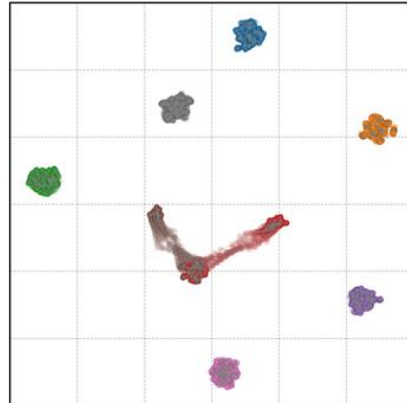
Notes: Six dynamic market structure maps, estimated by EvoMap (implemented for MDS and t-SNE) for two combinations from the simulation parameter space in Appendix Table A-1 (each simulated once for each of the three evolution scenarios). Each bubble represents one firm's position. Shaded paths indicate the estimated trajectories. Colors indicate submarket membership. We set the remaining simulation parameters to their medium values.

4.1 Evaluation Criteria and Comparison

To evaluate and compare dynamic mapping approaches, we use the three metrics introduced in Section 3.3: Hit-Rate (HIT-RATE), Misalignment (MIS-ALIGN), and Persistence (PERS). To further evaluate how well the estimated trajectories correspond to actual changes in market structure, we introduce a fourth metric: Change-Correlation (C-CORR), the Pearson correlation coefficient between the movement path lengths of firms' simulated and estimated positions, respectively. The intuition behind C-CORR is that firms should change their positions on a map in relation to how strongly their actual positions have changed.

HIT-RATE and C-CORR represent goodness-of-fit measures, which capture how well the estimated positions and trajectories fit the simulated data. MIS-ALIGN and PERS represent descriptive measures, which qualify different aspects of the estimated maps and trajectories. Naturally, alternative specifications for these metrics are possible, which we test in Online Appendix C.

Our simulation study comprises the following six dynamic mapping approaches: (1) EvoMap, (2) Dynamic t-SNE[3], (3) ex-post alignment via Procrustes Analysis, (4) sequential initialization of t-SNE, (5) fixed initialization of t-SNE, and (6) independent mapping (the benchmark for static mapping quality). Here, we focus on t-SNE as the static mapping method as (I) we also use it later in our empirical application, and (II) it is frequently used in recent marketing studies (e.g., Gabel et al. 2019; Yang et al. 2021). Appendix B also evaluates EvoMap paired with additional mapping methods (i.e., Stress-based metric MDS and Sammon Mapping).

4.2 Simulation Results

Visual inspection (as shown in Figure 4) confirms that EvoMap generates well-aligned maps (i.e., the coarse-grained configurations remain stable) in which individual positions still change over time. Firms do not erratically jump around in the estimated sequence of maps. Instead, firms tend to follow relatively gradual movement paths. Firms unaffected by the induced evolution scenarios are visually separable; they retain relatively stable positions. As a result, we can easily discern firms' trajectories (marked as shaded movement paths in Figure 4).

---

[3] Downloaded from https://github.com/paulorauber/thesne, latest version as of 03/22/2016.

Table 2: Results of Simulation Study

| Dynamic Mapping Approach	Evolution Scenario I: New Submarket Emergence				Evolution Scenario II: Shifts in Positions				Evolution Scenario III: Market Entry			
	MIS-ALIGN	PERS	C-CORR	HIT-RATE	MIS-ALIGN	PERS	C-CORR	HIT-RATE	MIS-ALIGN	PERS	C-CORR	HIT-RATE
(1) EvoMap (t-SNE)	**0.129** (0.083)	**0.869** (0.085)	**0.773** (0.177)	**0.673** (0.098)	**0.179** (0.077)	**0.872** (0.086)	**0.881** (0.125)	**0.665** (0.104)	**0.156** (0.069)	**0.884** (0.062)	**0.852** (0.101)	**0.657** (0.113)
(2) Dynamic t-SNE (Rauber et al. 2016)	21.448 (35.178)	0.409 (0.343)	0.213 (0.285)	0.406 (0.156)	21.196 (34.196)	0.349 (0.294)	0.333 (0.315)	0.399 (0.156)	19.299 (31.389)	0.354 (0.289)	0.325 (0.283)	0.399 (0.143)
(3) Ex-post Alignment (t-SNE)	12.160 (14.923)	-0.404 (0.190)	0.107 (0.403)	0.650 (0.113)	12.148 (14.544)	-0.361 (0.191)	0.167 (0.242)	0.643 (0.120)	12.407 (14.027)	-0.372 (0.190)	0.113 (0.238)	0.634 (0.128)
(4) Sequential Initialization (t-SNE)	1.958 (1.455)	0.103 (0.325)	0.160 (0.332)	0.642 (0.117)	2.820 (2.085)	-0.018 (0.224)	0.386 (0.259)	0.635 (0.123)	2.646 (1.865)	0.048 (0.252)	0.316 (0.229)	0.626 (0.129)
(5) Fixed Initialization (t-SNE)	4.184 (3.567)	-0.341 (0.197)	0.004 (0.340)	0.650 (0.112)	5.108 (4.112)	-0.306 (0.176)	0.145 (0.183)	0.644 (0.119)	5.132 (4.198)	-0.303 (0.181)	0.084 (0.156)	0.634 (0.128)
(6) Independent Mapping (t-SNE)	21.594 (16.694)	-0.483 (0.097)	-0.165 (0.323)	0.650 (0.113)	21.592 (16.169)	-0.482 (0.096)	-0.039 (0.100)	0.643 (0.120)	22.574 (15.390)	-0.489 (0.099)	-0.091 (0.126)	0.634 (0.128)

Notes: Reported metrics are averages over 729 simulation iterations per evolution scenario corresponding to the parameter space shown in Appendix Table A-1. Standard deviation in parentheses. Best result per metric is marked in bold. MIS-ALGN (Misalignment): Misalignment of successive maps, measured as the average length of all movement in positions. PERS (Persistence): Persistence of trajectories, measured as the average Pearson correlation coefficient of successive changes in positions. C-CORR (Change-Correlation): Average Pearson correlation coefficient of trajectory length on the map vs. trajectory length in the simulated positions. HIT-RATE (Hit-Rate): Average 10-NN Hit-Rate of nearest neighbor recovery, adjusted for random agreement.

Table 2 reports the results of our comparison, organized by the three market evolution scenarios using the previously outlined metrics Hit-Rate (HIT-RATE), Misalignment (MIS-ALIGN), Persistence (PERS), and Change-Correlation (C-CORR). We report all metrics as means (standard deviation in parentheses) across 729 combinations for each of the three evolution scenarios. Thus, in total, we evaluate each approach on 2,187 simulated evolving market structures.

The results show that EvoMap consistently outperforms alternative approaches across all metrics and in all three evolution scenarios by a substantial margin. In contrast to all alternatives, EvoMap estimates well-aligned maps (MIS-ALIGN is lowest), featuring gradually evolving trajectories (PERS is highest), which correspond well to the actual (i.e., simulated) changes in market structure (C-CORR is highest). Notably, static mapping quality is also higher than for all remaining approaches. To understand why, recall that we evaluate static mapping quality against the actual (simulated) positions (i.e., without any noise) but estimate maps using the noisy relationship matrices (i.e., the relationships measured after adding noise to the simulated positions). Thus, any approach that merely considers data from a single period cannot separate the actual positions from the additional noise in each period. EvoMap, in contrast, considers information from multiple successive periods when estimating the positions in each given period. Thereby, we can (partially) cancel-out noise in each period. As a result, the static mapping quality increases (HIT-RATE is highest).

Considering the alternative approaches in more detail, we make the following observations: Our results confirm our earlier hypothesis that ex-post alignment via Procrustes Analysis performs poorly when applied to maps generated by nonlinear methods like t-SNE. Additional analyses (presented in Appendix B) show that Procrustes Analysis performs better when paired with MDS—yet, still falls short of EvoMap (with MDS). In regard to the different initialization strategies, we find that sequential initialization performs best (specifically, MIS-ALIGN is low, indicating rather well-aligned maps). However, sequential initialization still exhibits strong sensitivity to minor changes (as indicated by low Persistence). Moreover, it does not recover actual changes in market structure well (C-CORR is low).

Finally, we find that Dynamic t-SNE achieves moderate Persistence and Change Correlation values but cannot consistently create well-aligned maps with good static mapping quality (indicated by high Misalignment and low Hit-Rates). Moreover, Dynamic t-SNE's results are sensitive to the different types of market structure and often reach only poor local optima (indicated by low Hit-Rates and high standard

deviations across all metrics). Thus, in Online Appendix C, we test whether we can improve Dynamic t-SNE's performance using EvoMap's optimization procedure. We find that doing so improves its performance, but results are still below the performance of EvoMap by a substantial margin.

4.3 Additional Analyses

EvoMap introduces several novelties. We conduct additional analyses with our simulated data to better understand the benefits of these novelties.

First, we investigate the benefits of EvoMap over alternative approaches when paired with other static mapping methods (i.e., MDS and Sammon Mapping). We find that EvoMap also produces superior results (see Appendix B for details).

Second, we disentangle the impact of adaptive regularization and smoothing on map quality. Our analysis reveals that both contribute individually to map quality, but the combination of both yields the highest goodness-of-fit (see Online Appendix C for details).

Third, we investigate EvoMap's "noise cancellation" property more closely using hold-out validation. Specifically, we test the extent to which EvoMap avoids overfitting individual maps to noise in the data by holding out particular periods during the estimation and predicting their positions based on the estimated trajectories. In contrast to alternative approaches, EvoMap substantially mitigates the risk of overfitting (see Appendix C).

Finally, we test the sensitivity of our findings towards market structure characteristics in Online Appendix D. Regression analysis shows that the reported results do not suffer from pronounced sensitivity to specific market structure characteristics. Thus, EvoMap is well applicable for a wide range of market structures.

5 EMPIRICAL STUDY

In what follows, we aim to empirically show that EvoMap creates additional market structure insights beyond traditional (i.e., static) mapping methods. Specifically, we seek to (1) reveal firms' trajectories using EvoMap, (2) test the estimated trajectories for face validity, (3) and illustrate the additional insights that firms' trajectories provide. To do so, we use EvoMap to empirically investigate the evolution of publicly listed firms' positions over two decades.

Specifically, we apply EvoMap to the Text-based Network Industry Classification (TNIC) data by Hoberg and Phillips (2016). These data consist of a time-indexed sequence of pairwise similarity measures among all

publicly listed firms in the United States. Hoberg and Phillips (2016) derive these measures of pairwise firm similarity based on firms' product descriptions in their 10-k SEC filings.

We chose these data for four reasons: First, these data span more than 20 years. Such an extended period provides the ideal empirical setting to assess EvoMap's ability to identify trajectories that reveal persistent underlying changes in firms' positions. Second, TNIC data are well established and validated in the Finance and Economics literature. Previous research leveraged these data to, for instance, study mergers and acquisitions, firms' reactions to product market threats, the relationship between competition and the cost of capital, or innovation strategies in the IT industry (Hoberg and Phillips 2010; Valta 2012; Hoberg et al. 2014; Kim et al. 2016; Li and Zhan 2019). Third, TNIC data feature well-known publicly listed firms, which facilitates assessing the face validity of the identified market structure and its evolution. Lastly, TNIC data provide sufficient complexity in market structure (i.e., many competitors, many submarkets) and include many sources of market structure evolution (e.g., changes in product portfolios, product innovations, or technological change).

5.1 Sample Construction

We retrieved the TNIC data from the authors' website https://hobergphillips.tuck.dartmouth.edu/ (accessed on 02/17/2021). The data comprise a time-indexed list of pairwise similarity measures, which we transform into a sequence of similarity matrices. Our observation period ranges from 1998 to 2017, with one observation per year, totaling 20 observations for each firm pair. The set of firms changes every year due to entries (i.e., IPOs) and exits (such as defaults, delisting, or acquisitions). To study firms' trajectories across an extended period, we focus our analysis on firms present in each period. We prune the firm network to only those firms who (I) have at least three product market competitors in the sample in each year and (II) whose total similarity to other firms in the sample exceeds 1% in each year. Setting such a minimum threshold is required. Otherwise, our sample would include firms that barely relate to any other firm in our sample, such that we could not faithfully estimate their positions. Our final sample comprises 1,092 firms across 20 years. We match our sample to further data from Compustat (via the "gvkey" identifier) to obtain descriptive variables, such as the firm names and SIC codes.

5.2 Dynamic Mapping of TNIC Data

We apply EvoMap (with t-SNE) to our full sample; That is, to the pairwise relationships among all 1,092 firms

for all 20 years. We set hyperparameters α and p using our hyperparameter tuning procedure that we proposed in Section 3.3: We quantitatively evaluate a fine grid of 150 potential combinations (3 values for p and 50 values for α, as depicted in Figure 2), identify candidate values that trade-in approximately 5% of goodness-in-fit (in terms of the adjusted 10-NN Hit-Rate) relative to independent mapping, and select the final hyperparameters based on visual evaluation of the candidate values' results. Specifically, we set $\alpha = 6.92 \times 10^{-4}$ and $p = 3$.

**Figure 5: Static Market Structure Maps for the TNIC sample in 1998 and 2017**
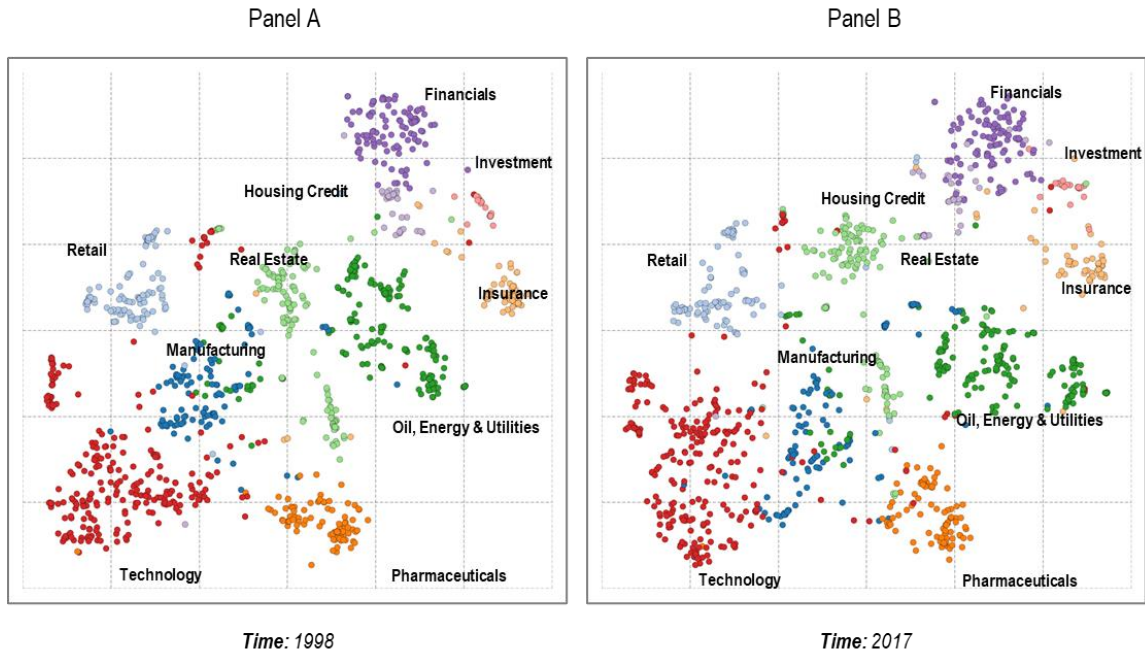
Notes: Each bubble represents one firm, colors indicate submarkets. Submarkets are labeled (manually) based on firms included in each submarket. Submarket assignments identified within the competitive relationship matrices using Louvain community detection prior to mapping. Adj. 10-NN Hit-Rates: 45.1% (left), 45.4% (right).

Figure 5 shows the first and the last map from EvoMap's estimated map sequence (i.e., 1998 and 2017). Both maps in Figure 5 reveal various submarkets (as indicated by the pronounced clustering of the bubbles in the maps) that are well aligned. To better understand the visualized submarket structures and to test for face validity, we manually label each submarket based on the firms included in each submarket. For instance, the technology submarket (bottom left) includes firms such as Apple, Microsoft, or Western Digital, whereas the financials submarket (top right) includes firms such as Capital One or Wells Fargo. All submarkets have an easily identifiable theme. Moreover, the relative positioning is intuitive (for instance, financials, investment, and insurance are positioned close together in the top right corner.

The comparison of two aligned static market structure maps, as presented in Figure 5, offers some insight

into the evolution of market structure over time (e.g., changes in the submarket structure). However, it does not readily reveal how individual firms changed their positions. A dynamic market structure map, on the other hand, can reveal such changes in market structure based on firms' trajectories. We present EvoMap's dynamic market structure map across 20 years in Figure 6.

Figure 6: EvoMap's Dynamic Market Structure Map (1998 - 2017)
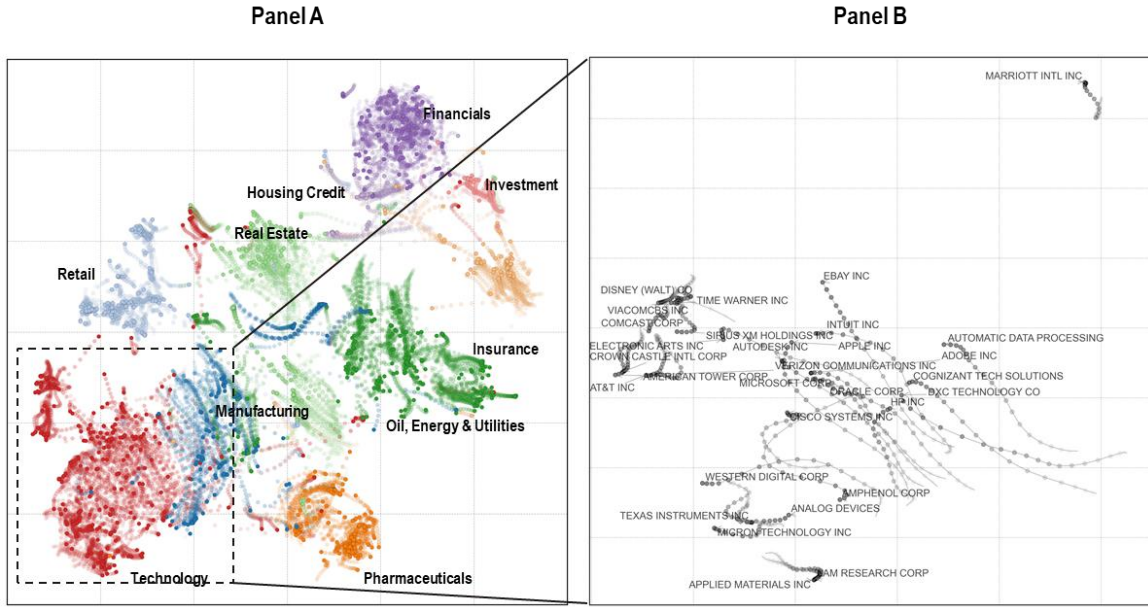
Notes: Overlay of 20 maps estimated by EvoMap. Each bubble represents one firm. More transparent bubbles indicate earlier positions. Shaded movement paths highlight firms' trajectories over time. Panel B zooms in on the top 30 firms by market value (as of 2017) in the bottom left area. Average adjusted 10-NN Hit-Rate: 44.6%.

Figure 6 (Panel A) reveals all firms' trajectories from 1998 until 2017 as smooth, shaded paths. We find that the global market structure (i.e., relative positions of submarkets) remains stable and individual maps are well aligned. Although most firms change their position only to a small extent, some firms substantially change their competitive positions (as indicated by several long trajectories). Moreover, while the global submarket structure remains relatively stable, there is considerable movement of firms within and across submarkets. For instance, within the bottom-left part of the map (highlighted in Panel B of Figure 6), technology firms (such as Apple) move closer to communication firms (such as AT&T), reflecting the increasingly important role of communication technologies for these firms' products.

5.3   Evaluation

In what follows, we evaluate EvoMap's trajectories with respect to (1) their face validity and (2) the substantive insights they can create.

5.3.1 Face Validity

Figure 6 suggests that most firms do not change their positions relative to other firms in the market much (i.e., most trajectories are relatively short). However, several firms substantially changed their positions (as indicated by longer trajectories). Our observation confirms the intuition that there is heterogeneity regarding the extent to which firms change their competitive position over time. Although greater changes are easier to see in Figure 6, most firms adjust their position to some extent (that is, we observe almost no zero-length trajectories).

To test the face validity of the estimated trajectories, we average their length on the SIC code level. Intuitively, if the identified trajectories accurately uncover changes in firms' positions, firms in less dynamic markets should (on average) have shorter trajectories than firms in more dynamic markets. To evaluate whether this relationship holds, we aggregate the average trajectory length by market membership (measured as the two-digit SIC code). Table 3 shows the 10 SIC codes with the longest and shortest average trajectory (measured as their movement path's total length in Figure 6). Our findings are intuitive because we observe the shortest average trajectories in markets with rather stable product portfolios (e.g., printing, food stores, apparel). That is, firms in these markets, including the New York Times or Abercrombie and Fitch, did not exhibit substantial changes in the products they sell. For instance, while fashion trends evolve, fashion products (i.e., apparel) change little. Likewise, while the content published by news publishers changes constantly, their products themselves remain rather similar. For instance, the New York Times still sells news. In contrast, we observe the longest average trajectories in markets with rather dynamic product portfolios (e.g., retail, business services, trade, or computer equipment). Firms in these markets, including Adobe or Intuit, change their products relative to their peers much more substantially. Many software firms, for instance, have been shifting their solutions to cloud-based subscription models, whereas wholesale trade firms, such as United Natural Food, constantly need to adjust their product portfolios in response to changing trends (e.g., the emergence of meat substitutes, vegan foods or sustainable products).

Table 3: Average Trajectory Length per SIC Code

Longest Trajectories			Shortest Trajectories		
SIC Description	SIC Code	Rank	SIC Description	SIC Code	Rank
Engineering, Accounting, Research, and Management Services	87	1	Security & Commodity Brokers, Dealers, Exchanges & Services	62	28
Miscellaneous Retail	59	2	Automotive Dealers and Gasoline Service Stations	55	29
Construction - General Contractors & Operative Builders	15	3	Home Furniture, Furnishings and Equipment Stores	57	30
Nondepository Credit Institutions	61	4	Apparel and Accessory Stores	56	31
Wholesale Trade - Nondurable Goods	51	5	General Merchandise Stores	53	32
Business Services	73	6	Amusement and Recreation Services	79	33
Real Estate	65	7	Food Stores	54	34
Transportation Equipment	37	8	Eating and Drinking Places	58	35
Primary Metal Industries	33	9	Printing, Publishing and Allied Industries	27	36
Fabricated Metal Products	34	10	Apparel, Finished Products from Fabrics & Similar Materials	23	37

Notes: Two-digit SIC codes ordered by their firms' average trajectory length. We excluded SIC codes with less than five firms in our sample. SIC descriptions obtained from https://mckimmoncenter.ncsu.edu/2digitsiccodes/, last accessed 02/21/2022

5.3.2 Substantive Insights

EvoMap's trajectories offer additional substantive insights beyond static market structure maps. For example, monitoring an individual firm's trajectory can reveal whether and how its position has changed in the past. Panel A of Figure 7 displays the trajectory of Capital One (a large financial service provider). Its trajectory reveals that Capital One has indeed changed its position substantially over the past 20 years. Moreover, we observe a pronounced bend. This sudden change of direction occurred around 2004 and coincided with the expansion of Capital One's business model. In the early 1990s, Capital One solely focused on selling credit cards. Then, throughout the 1990s, Capital One expanded into auto loans. Finally, around 2005, Capital One expanded into consumer banking.

Likewise, the trajectory of American Express (Panel B of Figure 7) reveals its departure from the investment submarket, where investment banks such as Morgan Stanley, Oppenheimer or Jefferies are

positioned (and remain at their position), towards the financials submarket including retail banks. This observation coincides with the divesting of American Express's investment banking units, which started in the early 1990s. The aforementioned trajectories are thereby insightful to the respective firms themselves and other firms in the market because these trajectories may alert them to emerging competitors at an early stage.

Figure 7: Trajectories of Selected Firms
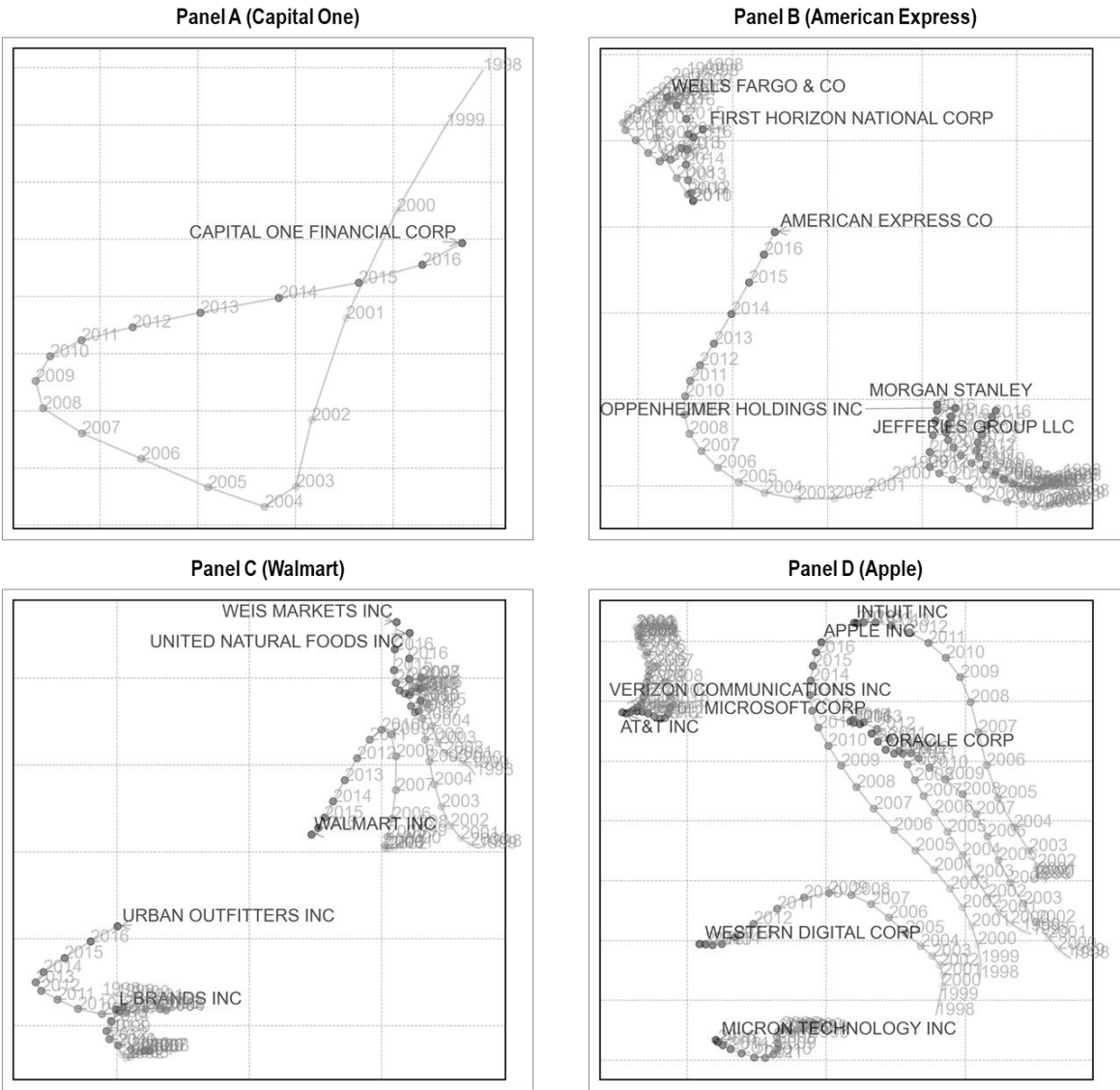
Notes: Trajectories of Capital One, American Express, Walmart, and Apple, including some of their peers.

Besides studying each firm's trajectory, market researchers can gain additional insights from studying firms' trajectories relative to each other. Panel C shows how the trajectory of Walmart diverges from grocery stores (such as Weis Markets or United Natural Foods) and starts to move closer towards apparel firms (such as Urban Outfitters or L Brands), revealing the growing expansion of its fashion assortments featuring both national and private brands.

Likewise, Panel D of Figure 7 shows the trajectories for Apple vs. some of its peers (such as Western Digital). The map shows that Apple and Western Digital were positioned closely in 1998. Both firms provided computer hardware, together with other hardware providers, such as HP. Over time, however, Apple's and Western Digital's trajectories diverged. While the trajectory of Western Digital converges with other hardware providers such as Micron, the trajectory of Apple points in a different direction. Eventually, Apple's trajectory converges with digital services providers, such as Intuit. The recovered trajectory for Apple reflects Apple's evolving business model, which evolved from mainly selling hardware (such as the MAC) towards software and digital services (such as its AppStore or iTunes). While a static market structure map (for instance, around 2010) would place the three firms at approximately equal distance, EvoMap's dynamic market structure map indicates the divergence of Apple's and Western Digital's positions. Indeed, a few years later, Apple and Western Digital faced very different competitors.

The above examples highlight the value of mapping firms' trajectories because they reveal firms' (re-)positioning efforts at an individual level and relative to one another. We make our empirical study's full results available to readers in an interactive browser-based online tool specifically developed for this article (visit www.evomap.io). Using our tool, readers can easily explore the trajectories of various firms of interest on their own.

Finally, we can also use EvoMap to obtain a fine-grained picture of a smaller set of firms in the market (using, for instance, EvoMap implemented for metric MDS rather than t-SNE). We provide the results for such an analysis of 10 competitors within the technology submarket in Online Appendix F. The results reflect similar evolution patterns, supporting our findings' robustness.

6   DISCUSSION AND CONCLUSION

This article presents a novel dynamic mapping framework called EvoMap that uses longitudinal data on firms' competitive relationships to reveal firms' trajectories within evolving market structures. Prior marketing research highlights the value of, and need for, monitoring the evolution of market structures. Although high-frequency, longitudinal data on competitive relationships among firms are increasingly available, market structure analysis has not yet fully unlocked the potential insights buried within these data.

The proposed EvoMap framework can unlock this potential by revealing the trajectories of firms' competitive positions over time. These trajectories contribute multiple new insights to market structure and

competitive analysis. First, by providing a forward-looking perspective on competition, they can help managers to detect competitive threats and opportunities early. Naturally, maintaining competitive advantages in dynamic environments is challenging (Eisenhardt 1989). Thereby, recognizing competitive threats early (i.e., while they are small) is critical to counter them (Peteraf and Bergen 2003). Doing so, however, is often a challenging task, as firms need to consider not only their closest peers but a broad range of potential competitors (Bergen and Peteraf 2002). Such broad-based competitor identification is especially challenging within a dynamic competitive landscape, as competition can come in many forms and from many directions (Levitt 1960; Peteraf and Bergen 2003). The trajectories estimated by EvoMap can help identify such threats by revealing other firms with converging trajectories. Thereby, EvoMap can jointly identify the trajectories of hundreds of firms (as demonstrated in our empirical application).

Beyond identifying competitors, the revealed trajectories can also help monitor competitors' actions and alert managers about their rivals' strategic intentions– such as changes in their positioning strategy. When formulating one's positioning strategy, a firm needs to consider its competitors' current positions and also their potential actions, e.g., whether they move in a similar direction or not (Ning and Villas-Boas 2021). Monitoring competitors' trajectories can offer such indications and thus inform marketing strategies. Finally, once a firm has decided to reposition itself, the firm can monitor its trajectory to evaluate the effectiveness of its actions.

The visual representation in maps provides an easy-to-understand common language that facilitates communication across different roles within a firm. For instance, one could integrate dynamic market structure maps into existing analytics dashboards, which are a primary basis for decision-making in the industry (Wedel and Kannan 2016). With such dashboards, decision-makers across a firm could easily and frequently monitor market developments.

As EvoMap is a flexible framework, one can further apply it to many other empirical settings beyond the one presented in this article. For instance, one could infer firms' trajectories from changes in their ad copies or changes in the content on their websites. Such data are regularly archived and often made accessible, for instance, by competitive intelligence platforms. Modern data science provides ample means to identify and measure relationships among firms within such (often unstructured) data. For instance, (pre-trained) embeddings allow market researchers to measure the similarity among texts, images, speech, or a combination

thereof. Paired with such techniques, EvoMap might be used to study, for example, evolving product designs, firm communication, or brand logos.

Beyond firm-generated content, consumers' digital footprints offer a plethora of suitable data sources to study evolving market structures. Some examples are search logs, social media posts, and customer reviews. Because such data are often available historically and in real-time, they create additional opportunities to study evolving relationships among firms, brands, products, and consumers. We also see potential applications of EvoMap at the intersection of marketing and other domains. For instance, political parties communicate their positioning towards society in manifestos or through members' speeches. Such communication is often archived historically and could reveal party trajectories. Researchers could also use EvoMap in bibliometrics to study the evolution of scientific fields or individual journals' positions.

Future research could also explore different ways to link map positions (and changes therein) to firm attributes (e.g., how they describe their products). Doing so could help to understand the underlying market's structure and its evolution even further. Another avenue might be to investigate why firms in some industries exhibit more dynamic positions than others. As our empirical study showed, there exists substantial variation in the length of firms' trajectories. Identifying the drivers and consequences of these dynamics could further help managers develop an effective (re-)positioning strategy.

Naturally, EvoMap is not free of limitations. First, EvoMap is subject to the limitations of the particular static mapping method used. To mitigate this limitation, we designed EvoMap as a flexible framework that can easily accommodate various static mapping methods and implemented it for three very different methods (i.e., metric MDS, Sammon Mapping, and t-SNE). Second, the quality of EvoMap's dynamic market structure maps is bounded by the quality of the used data. Although EvoMap offers several innovations such as "noise cancellation", it still depends on valid relationship data. And third, EvoMap introduces two hyperparameters to the mapping process (α and p) that market researchers need to consider in their analysis. However, we (1) investigate and document their impact on map quality and (2) offer a tuning scheme to set them appropriately for the particular empirical setting and data at hand.

Methodologically, EvoMap provides the foundation for further research into mapping the dynamics of competitive relationships. Such research could, for instance, pair EvoMap with additional mapping methods, with more sophisticated optimization procedures (such as majorization techniques), or embed the estimated

trajectories into models of consumer choice (for instance, to augment dynamic market structure maps with evolving consumer ideal points), or study how the trajectories of firms that compete in multiple submarkets (e.g., Amazon) evolve differently in those submarkets (Ringel 2022). While recent research in marketing shows a growing interest in modeling and understanding the dynamics in high-frequency data (Du and Kamakura 2012; Schweidel and Moe 2014; Xiong and Bharadwaj 2014; Ma et al. 2015; Puranam et al. 2017; Zhong and Schweidel 2020), the dynamics of relationships among market actors inherent in such data have received less attention. We hope to open a new avenue for further research in this direction.

APPENDIX

A. Data-Generating Process and Hyperparameter Selection for Simulation Study

Ideally, the simulated competitive relationship matrices should represent various market structures with different submarkets and a wide variety of possible inter- and intra-submarket relationships. To create such variety in our simulation study, we formulate the following data-generating process (DGP):

First, we assign each firm $i \in I$ with uniform probability to one of k submarkets, where $n = |I|$ and $k \in \mathbb{N}^+$. Then, for each submarket $m \in \{1, \dots, k\}$, we draw the positions in the d-dimensional attribute space from a Gaussian distribution $\mathcal{N}(\mu_m, \sigma_{sub}^2)$ centered at the submarket mean μ_m with standard deviation σ_{sub}. Within the attribute space, we place the submarkets' means with uniform probability within the edges of a d-dimensional hypercube with unit length. We thus generate pairs of closer and more distant submarkets and obtain the simulated positions $X_{sim} \in \mathbb{R}^{n \times d}$. Such a process is flexible enough to simulate competitive relationship matrices for a broad class of potential market structures (e.g., small/large markets with few/many strongly/weakly connected submarkets). At the same time, the proposed DGP is parsimonious because it requires only a small number of meaningfully interpretable simulation parameters.

Next, we modify the simulated positions according to one of the following scenarios:

Scenario I   Emergence of a new submarket: A fraction (ρ) of firms are randomly assigned to a newly formed submarket and, over T periods, change their positions towards the new submarket.

Scenario II   Shifts in positioning: A fraction (ρ) of firms are randomly assigned to a different yet existing submarket and, over T periods, gradually shift their positions towards their assigned submarket.

Scenario III   Market entry: A fraction (ρ) of firms enter the market at random positions in the attribute space.

Over T periods, each new entrant gradually shifts towards its assigned submarket.

To make our simulation more realistic, we implement two other phenomena that occur in real-world empirical data: non-gradual evolution and noise. Our intuition is the following. First, we do not expect changes in firm positions to occur gradually and simultaneously over time. We, therefore, vary the speed at which each firm changes its position during our simulation. Specifically, we draw T (the number of periods over which a firm changes its position) randomly for each firm from a zero-truncated Poisson distribution. As a result, some firms will adjust their positions within each simulation iteration over a few periods (i.e., quickly and less gradually), and others over more periods (i.e., slowly and more gradually). We set the Poisson parameter to 8 for all simulation iterations. Second, empirical measurements of competitive relationships can be subject to noise due to the nature of the data source or the collection process. We, therefore, add varying degrees of white noise drawn from $\mathcal{N}(0, \sigma_{temporal}^2)$ to each firm's simulated position at each period before taking distances. We estimate the maps using the noisy relationship matrices but evaluate them against the simulated positions without noise. Appendix Table A-1 provides an overview of simulation parameters, their underlying market structure characteristic, the values assigned to the parameters, and the total number of simulation iterations. We chose the parameter space such that it covers a broad spectrum of possible market structures.

Appendix Table A-1: Simulation Parameter Space

Simulation Parameter	Label	Market Structure Characteristic	Values		
n	Number of firms	Market size	50	100	250
d	Number of dimensions	Market complexity	4	8	16
k	Number of submarkets	Market fragmentation	4	8	12
σ_{sub}	Within submarket standard deviation	Degree of submarket separation	.050 (strong)	.075 (medium)	.150 (weak)
$\sigma_{temporal}$	Temporal noise	Smoothness of evolution	.010 (low)	.025 (medium)	.050 (high)
ρ	Share of firms affected	Market dynamism	.050 (low)	. 100 (medium)	.150 (high)
Number of combinations	$3^6 = 729$				
Number of evolution scenarios	3				
Number of simulation iterations	$729 \times 3 = 2{,}187$				

Notes: The table lists all simulation parameters. We simulate each scenario for each combination. In total, we conduct 2,187 simulation iterations (729 combinations for each of the three evolution scenarios).

As we outlined in Section 3.3, running EvoMap requires setting two hyperparameters (α and p). According

to our tuning strategy, we would usually first evaluate several combinations of these two hyperparameters and then manually fine-tune them. Unfortunately, such an elaborate procedure is not feasible for 2,187 simulation iterations. Therefore, we adopt the following approach: In each simulation iteration, we evaluate the following $3 \times 5 = 15$ hyperparameter combinations for EvoMap: $\alpha \in \{10^{-5}, 10^{-4}, 10^{-3}\}, p \in \{1, \ldots, 5\}$, and pick the combination that yields the best result across all metrics. To identify this best combination, we rank the 15 combinations according to our four metrics and then pick the one with the lowest (i.e., best) rank across all four metrics.

Likewise, running dynamic t-SNE requires setting its hyperparameter λ, which performs a similar task as EvoMap's hyperparameter α (i.e., both α and λ balance the static vs. temporal component of the respective cost functions). Thus, we also run dynamic t-SNE for the same three values $\lambda \in \{10^{-5}, 10^{-4}, 10^{-3}\}$ and pick the best result for each simulation iteration accordingly. For larger values of λ, we find that static goodness-of-fit (measured via Hit-Rate) drops substantially for dynamic t-SNE[4]. We set the remaining parameters of dynamic t-SNE (e.g., its optimization settings) to the default values suggested by its authors.

Appendix Figure A-1 shows the hyperparameter distributions for all 2,187 simulation iterations. Notably, for most cases ($> 98\%$), the best solution is obtained with p greater than one. In these cases, EvoMap's additional smoothing penalties lead to better solutions.

Appendix Figure A-1: Selected Hyperparameter Values in Simulation Study (t-SNE)

---
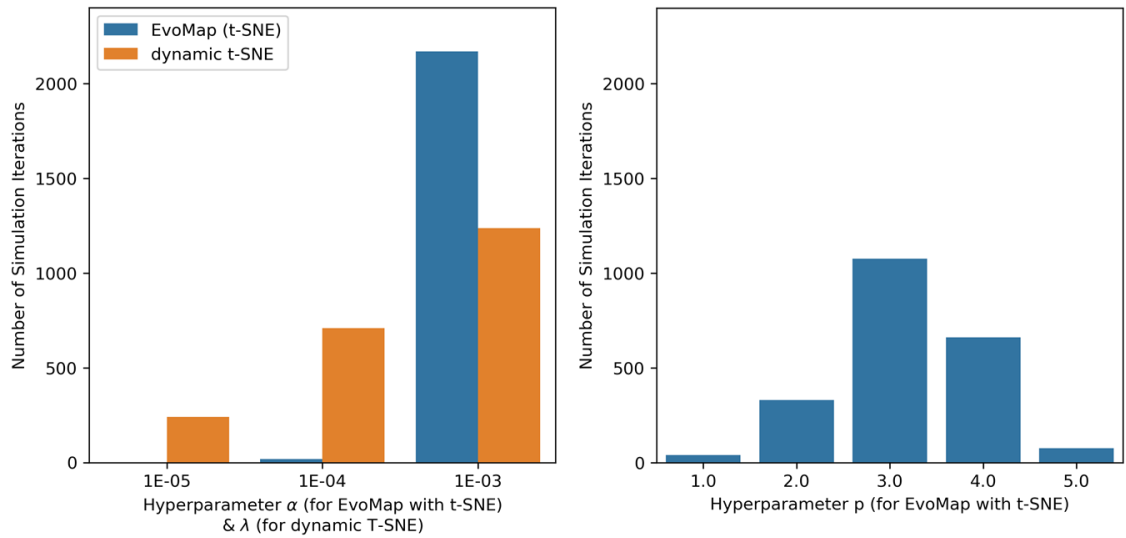
[4] For $\lambda = 10^{-2}$, the average (adjusted) Hit-Rate of dynamic t-SNE drops to ~13% (based on a test run with 50 randomly selected simulation iterations).

## B. Simulation Results for Metric MDS and Sammon Mapping

We also evaluate EvoMap in our simulation study paired with alternative static mapping methods and present the results in Appendix Table B-1. We exclude t-SNE-specific alternatives in this comparison (i.e., Dynamic t-SNE, sequential and fixed initialization). As for EvoMap paired with t-SNE, we also run each additional alternative of EvoMap for multiple hyperparameter combinations. Specifically, we run EvoMap (with MDS / Sammon) for $3 \times 3 = 9$ combinations: $\alpha \in \{10^{-2}, 10^{-1}, 5 \times 10^{-1}\}$ and $p \in \{1, ..., 3\}$. Appendix Figure B-1 shows the distribution of the selected hyperparameter values. Note that the range of suitable values for EvoMap's hyperparameters depends on the chosen mapping method (and its static cost function). The results in Appendix Table B-1 show that in line with our previous findings for EvoMap paired with t-SNE, pairing EvoMap with MDS or Sammon Mapping also produces much better solutions than other approaches.

Appendix Table B-1: Simulation Results for MDS & Sammon Mapping

Dynamic Mapping Approach	MIS-ALGN	PERS	C-CORR	HIT-RATE
Panel A: MDS				
(I) EvoMap (MDS)	**0.067** **(0.078)**	**0.894** **(0.099)**	**0.748** **(0.318)**	**0.546** **(0.137)**
(II) Ex-post Aligned MDS	0.505 (0.548)	-0.389 (0.201)	0.253 (0.373)	0.525 (0.140)
(III) Independent MDS	2.428 (1.360)	-0.480 (0.102)	-0.130 (0.201)	0.525 (0.140)
Panel B: Sammon Mapping				
(I) EvoMap (Sammon)	**0.094** **(0.155)**	**0.868** **(0.137)**	**0.718** **(0.317)**	**0.598** **(0.135)**
(II) Ex-Post Aligned Sammon	5.023 (11.320)	-0.488 (0.114)	0.082 (0.298)	0.586 (0.140)
(III) Independent Sammon	6.335 (11.647)	-0.488 (0.089)	-0.090 (0.209)	0.586 (0.140)

Notes: Reported metrics are averages over 2,187 simulations as outlined in Appendix Table A-1 (729 combinations and three evolution scenarios). Standard deviation in parentheses. Best result per metric marked in bold. MIS-ALGN (Misalignment): Misalignment of successive maps, measured as the average length of all movement in positions. PERS (Persistence): Persistence of identified trajectories, measured as the average Pearson correlation coefficient of successive changes in positions. C-CORR (Change Correlation): Average Pearson correlation coefficient of trajectory length on the map vs. trajectory length in the simulated positions. HIT-RATE (Hit-Rate): Average 10-NN hit-rate of nearest neighbor recovery, adjusted for random agreement.

Appendix Figure B-1: Selected Hyperparameter Values in Simulation Study (MDS & Sammon)
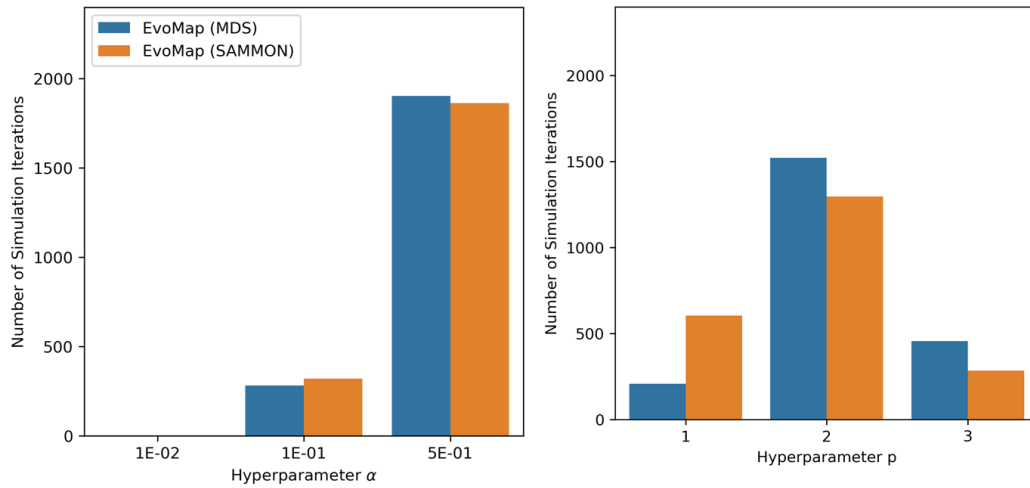
C. Assessing Overfitting

EvoMap seeks to identify underlying trends in the changes in firms' competitive positions. Doing so requires separating trends from noise or, stated differently, avoiding overfitting individual periods. Overfitting naturally arises when fitting maps to each period independently, as the resultant maps will adjust to all (even minor) changes in every period, including temporary noise. For EvoMap, increasing the hyperparameter p sets stronger constraints on the resultant trajectories, linking the estimated positions across an increasing number of successive periods. Thereby, the resultant maps should be less sensitive to changes in individual periods, thus preventing overfitting.

To test this hypothesis, we randomly split the data in each simulation iteration into two parts: estimation and hold-out data. We split along the time axis, such that the full relationship matrix in each period between the first and last period either belongs to the hold-out or the estimation data. We then estimate the sequence of maps using only the relationship matrices in the estimation data. To test for overfitting, we predict the positions on the map for the hold-out data using the estimated trajectories. Specifically, to derive firm i's position in any period in the hold-out data, we use linear interpolation between the nearest preceding and succeeding map position in the estimation data. Doing so predicts the map position along the estimated trajectories and tests the extent to which these trajectories capture the underlying evolution trends. For evaluation, we compute the average (adjusted) Hit-Rates for the predicted and estimated maps, respectively (see Appendix Table C-1). As expected, independent mapping (even when aligned ex-post via Procrustes Analysis) overfits most heavily.

EvoMap, in contrast, exhibits no sign of overfitting: The average difference between the Hit-Rate on the estimation and hold-out data is below one percentage point.

Appendix Table C-1: Hold-Out Validation Results

Dynamic Mapping Approach	HIT-RATE (Estimation Data)	HIT-RATE (Hold-Out Data)	HIT-RATE % Change
(1) EvoMap (t-SNE)	66.093 (10.588)	65.495 (10.658)	-0.595 (1.955)
(2) Dynamic t-SNE (Rauber et al. 2016)	44.109 (13.944)	42.809 (14.083)	-1.296 (2.822)
(3) Ex-post Alignment (t-SNE)	63.780 (12.039)	60.447 (13.416)	-3.330 (3.179)
(4) Sequential Initialization (t-SNE)	63.106 (12.301)	60.900 (13.414)	-2.202 (2.719)
(5) Fixed Initialization (t-SNE)	63.795 (12.032)	60.366 (13.513)	-3.426 (2.970)
(6) Independent Mapping (t-SNE)	63.780 (12.039)	58.443 (13.146)	-5.333 (3.354)

Notes: Average Adjusted 10-NN Hit-Rate values in percentage points across 2,187 simulation iterations. Standard deviation in parentheses.

REFERENCES

Bergen M, Peteraf MA (2002) Competitor identification and competitor analysis: A broad-based managerial approach. *Managerial and Decision Economics*. 23(4-5):157-169.

Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E (2008) Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*. 2008(10):P10008.

Borg I, Groenen PJ (2005) *Modern multidimensional scaling: Theory and applications*. (Springer Science & Business Media, New York).

Carroll JD, Chang J-J (1970) Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition. *Psychometrika*. 35(3):283-319.

Chen L, Buja A (2009) Local multidimensional scaling for nonlinear dimension reduction, graph drawing, and proximity analysis. *Journal of the American Statistical Association*. 104(485):209-219.

Cooper LG (1988) Competitive maps: The structure underlying asymmetric cross elasticities. *Management Science*. 34(6):707-723.

Culotta A, Cutler J (2016) Mining brand perceptions from Twitter social networks. *Marketing Science*. 35(3):343-362.

D'Aveni RA (1994) *Hypercompetition: Managing the dynamics of strategic maneuvering*. (Free Press, New York).

Day GS (1984) Strategic marketing planning: The pursuit of competitive advantage. *St. Paul, MN: West Publishing*.

Day GS, Shocker AD, Srivastava RK (1979) Customer-oriented approaches to identifying product-markets. *Journal of Marketing*. 43(4):8-19.

DeSarbo WS, Carroll JD (1985) Three-way metric unfolding via alternating weighted least squares.

Psychometrika. 50(3):275-300.

DeSarbo WS, Grewal R, Wind J (2006) Who competes with whom? A demand-based perspective for identifying and representing asymmetric competition. *Strategic Management Journal*. 27(2):101-129.

DeSarbo WS, Manrai AK, Manrai LA (1993) Non-spatial tree models for the assessment of competitive market structure: An integrated review of the marketing and psychometric literature, Eliashberg J, Lilien G, eds. *Handbooks in operations research and management science* (North-Holland, Amsterdam), 193-257.

Du RY, Kamakura WA (2012) Quantitative trendspotting. *Journal of Marketing Research*. 49(4):514-536.

Eisenhardt KM (1989) Making fast strategic decisions in high-velocity environments. *Academy of Management journal*. 32(3):543-576.

Elrod T, Russell GJ, Shocker AD, Andrews RL, Bacon L, Bayus BL, Carroll JD, Johnson RM, Kamakura WA, Lenk P (2002) Inferring market structure from customer response to competing and complementary products. *Marketing Letters*. 13(3):221-232.

France S, Carroll D (2007) Development of an agreement metric based upon the RAND index for the evaluation of dimensionality reduction techniques, with applications to mapping customer data (Springer Berlin Heidelberg, Berlin, Heidelberg), 499-517.

France SL, Akkucuk U (2021) A review, framework, and R toolkit for exploring, evaluating, and comparing visualization methods. *The Visual Computer*. 37(3):457-475.

France SL, Ghose S (2016) An analysis and visualization methodology for identifying and testing market structure. *Marketing Science*. 35(1):182-197.

Fruchterman TM, Reingold EM (1991) Graph drawing by force-directed placement. *Software: Practice and Experience*. 21(11):1129-1164.

Gabel S, Guhl D, Klapper D (2019) P2V-MAP: Mapping market structures for large retail assortments. *Journal of Marketing Research*. 56(4):557-580.

Gower JC (1975) Generalized procrustes analysis. *Psychometrika*. 40(1):33-51.

Hoberg G, Phillips G (2010) Product market synergies and competition in mergers and acquisitions: A text-based analysis. *Review of Financial Studies*. 23(10):3773-3811.

Hoberg G, Phillips G (2016) Text-based network industries and endogenous product differentiation. *Journal of Political Economy*. 124(5):1423-1465.

Hoberg G, Phillips G, Prabhala N (2014) Product market threats, payouts, and financial flexibility. *Journal of Finance*. 69(1):293-324.

Hunt SD (1983) General theories and the fundamental explananda of marketing. *Journal of Marketing*. 47(4):9-17.

Kim JB, Albuquerque P, Bronnenberg BJ (2011) Mapping online consumer search. *Journal of Marketing Research*. 48(1):13-27.

Kim K, Gopal A, Hoberg G (2016) Does product market competition drive CVC investment? Evidence from the US IT industry. *Information Systems Research*. 27(2):259-281.

Lambkin M, Day GS (1989) Evolutionary processes in competitive markets: beyond the product life cycle.

Journal of Marketing. 53(3):4-20.

Lee TY, Bradlow ET (2011) Automated marketing research using online customer reviews. *Journal of Marketing Research*. 48(5):881-894.

Levitt T (1960) Marketing myopia. *Harvard Business Review*. 38:45–56.

Li S, Zhan X (2019) Product market threats and stock crash risk. *Management Science*. 65(9):4011-4031.

Lilien GL, Rangaswamy A (2004) *Marketing engineering: Computer-assisted marketing analysis and planning*. (Trafford Publishing, Victoria, BC, Canada).

Liu L, Dzyabura D, Mizik N (2020) Visual listening in: Extracting brand image portrayed on social media. *Marketing Science*. 39(4):669-686.

Ma L, Sun B, Kekre S (2015) The squeaky wheel gets the grease - An empirical analysis of customer voice and firm intervention on Twitter. *Marketing Science*. 34(5):627-645.

Maaten Lvd, Hinton G (2008) Visualizing data using t-SNE. *Journal of Machine Learning Research*. 9(Nov):2579-2605.

Mela CF, Gupta S, Jedidi K (1998) Assessing long-term promotional influences on market structure. *International Journal of Research in Marketing*. 15(2):89-107.

Moore WL, Winer RS (1987) A panel-data based method for merging joint space and market response function estimation. *Marketing Science*. 6(1):25-42.

Nam H, Joshi YV, Kannan P (2017) Harvesting brand information from social tags. *Journal of Marketing*. 81(4):88-108.

Netzer O, Feldman R, Goldenberg J, Fresko M (2012) Mine your own business: Market-structure surveillance through text mining. *Marketing Science*. 31(3):521-543.

Ning ZE, Villas-Boas JM (2021) Following the customers: Dynamic competitive repositioning. *Management Science*. 68(2):1002-1018.

Peteraf MA, Bergen ME (2003) Scanning dynamic competitive landscapes: A market-based and resource-based framework. *Strategic Management Journal*. 24(10):1027-1041.

Puranam D, Narayan V, Kadiyali V (2017) The effect of calorie posting regulation on consumer opinion: A flexible latent dirichlet allocation model with informative priors. *Marketing Science*. 36(5):726-746.

Rao VR, Sabavala DJ (1986) *Measurement and use of market response functions for allocating marketing resources*. (Marketing Science Institute, Boston).

Rauber PE, Falcão AX, Telea AC (2016) Visualizing time-dependent data using dynamic t-SNE. *Proceedings of the Eurographics / IEEE VGTC Conference on Visualization: Short Papers* (Eurographics Association, Groningen), 73-77.

Ringel DM (2022) Multimarket Membership Mapping. *Journal of Marketing Research*. forthcoming.

Ringel DM, Skiera B (2016) Visualizing asymmetric competition among more than 1,000 products using big search data. *Marketing Science*. 35(3):511-534.

Sabnis G, Grewal R (2012) Competition and its implications for marketing strategy, Shankar V, Carpenter G,

eds. *Handbook of Marketing Strategy* (Edward Elgar Publishing, Cheltenham, UK), 58-77.

Sammon JW (1969) A nonlinear mapping for data structure analysis. *IEEE Transactions on computers*. 100(5):401-409.

Schönemann PH (1966) A generalized solution of the orthogonal procrustes problem. *Psychometrika*. 31(1):1-10.

Schönemann PH, Carroll RM (1970) Fitting one matrix to another under choice of a central dilation and a rigid motion. *Psychometrika*. 35(2):245-255.

Schweidel DA, Moe WW (2014) Listening in on social media: A joint model of sentiment and venue format choice. *Journal of Marketing Research*. 51(4):387-402.

Smelcer JB, Carmel E (1997) The effectiveness of different representations for managerial problem solving: comparing tables and maps. *Decision Sciences*. 28(2):391-420.

Tirunillai S, Tellis GJ (2014) Mining marketing meaning from online chatter: Strategic brand analysis of big data using latent dirichlet allocation. *Journal of Marketing Research*. 51(4):463-479.

Tversky A, Hutchinson J (1986) Nearest neighbor analysis of psychological spaces. *Psychological Review*. 93(1):3.

Urban GL, Johnson PL, Hauser JR (1984) Testing competitive market structures. *Marketing Science*. 3(2):83-112.

Valta P (2012) Competition and the cost of debt. *Journal of Financial Economics*. 105(3):661-682.

van Heerde HJ, Mela CF, Manchanda P (2004) The dynamic effect of innovation on market structure. *Journal of Marketing Research*. 41(2):166-183.

Wedel M, Kannan P (2016) Marketing analytics for data-rich environments. *Journal of Marketing*. 80(6):97-121.

Wei YM (2020) The similarity network of motion pictures. *Management Science*. 66(4):1647-1671.

Xiong G, Bharadwaj S (2014) Prerelease buzz evolution patterns and new product performance. *Marketing Science*. 33(3):401-421.

Xu KS, Kliger M, Hero AO (2013) A regularized graph layout framework for dynamic network visualization. *Data Mining and Knowledge Discovery*. 27(1):84-116.

Yang Y, Zhang K, Kannan PK (2021) Identifying market structure: A deep network representation learning of social engagement *Journal of Marketing*. forthcoming.

Zhong N, Schweidel DA (2020) Capturing changes in social media content: A multiple latent changepoint topic model. *Marketing Science*. 39(4):827-846.

ONLINE APPENDIX

A. Gradient Derivation and Implementation

Herein, we formally derive the gradient of EvoMap's cost function to facilitate the implementation of our

framework for different methods, optimization procedures, and various programming languages. We estimate

the sequence of maps $\left(\hat{X}_t\right)_{t=1,\dots,T}, \hat{X}_t \in \mathbb{R}^{n \times 2}$ by optimizing the following cost function:

$$C_{total}(X_1, \dots, X_T) = \sum_{t=1}^{T} C_{static}(X_t) + \alpha \cdot C_{temporal}(X_1, \dots, X_T) \tag{A-1}$$

Therefore, the gradient $\operatorname{grad} C_{total}(X_1, \dots, X_T)$ consists of the two independent components

$\operatorname{grad} C_{static}(X_t)$ and $\operatorname{grad} C_{temporal}(X_1, \dots, X_T)$. The chosen mapping method provides $\operatorname{grad} C_{static}(X_t)$.

We derive $\operatorname{grad} C_{temporal}(X_1, \dots, X_T)$ subsequently. Recall that we defined the temporal cost function as

$$C_{temporal}(X_1, \dots, X_T) = \sum_{i=1}^{n} f_w(i) \sum_{k=1}^{p} \sum_{t=k+1}^{T} 1_{[i \in I_{t,k}]} \left\| \nabla^k x_{i,t} \right\|^2 \tag{A-2}$$

Fix a period $\tau \in \{1, \dots, T\}$ and a firm $j \in \{1, \dots, n\}$ and let $x_{j,\tau} \in \mathbb{R}^2$ denote firm j's map position at time

τ. $\operatorname{grad} C_{temporal}$ consists of all partial derivatives of $C_{temporal}$ w.r.t. $x_{j,\tau} \ \forall j \in \{1, \dots, n\}, \tau \in \{1, \dots T\}$,

expressed by

$$\frac{\partial C_{temporal}}{\partial x_{j,\tau}} = f_w(j) \sum_{k=1}^{p} \sum_{t=k+1}^{T} 1_{[j \in I_{t,k}]} \frac{\partial \left\| \nabla^k x_{j,t} \right\|^2}{\partial x_{j,\tau}} \tag{A-3}$$

This partial derivative of the temporal component for firm j does not depend on any other firm's position.

Suppose that $j \in I_{t,k}$ (else, the respective term in the sum is zero). $f_w(j)$ only depends upon the input data

and is thus a scalar independent of $x_{j,\tau}$. The derivation of the partial derivative of $C_{temporal}$ therefore consists

of deriving the partial derivative of $\left\| \nabla^k x_{j,t} \right\|^2$ w.r.t $x_{j,\tau}$ for all $k \in \{1, \dots, p\}$ and $t \in \{1, \dots, T\}$. The partial

derivative of the outer norm depends on its choice. We implement EvoMap for the Euclidean norm, such that

$\frac{\partial \|\mathbf{v}\|^2}{\partial \mathbf{v}} = 2\mathbf{v}$ for any real-valued vector $\mathbf{v}$. Therefore,

$$\frac{\partial \left\| \nabla^k x_{j,t} \right\|^2}{\partial x_{j,\tau}} = 2 \cdot \nabla^k x_{j,t} \cdot \frac{\partial \nabla^k x_{j,t}}{\partial x_{j,\tau}} \tag{A-4}$$

Recall that the inner part of the norm corresponds to the k-th order difference of firm j's map position at

time t, formally defined as

$$\nabla^k x_{j,t} := \begin{cases} \left(\nabla^{k-1} x_{j,t} - \nabla^{k-1} x_{j,t-1} \right) & k \geq 1 \wedge t \geq k+1 \\ x_{j,t} & \text{for} \quad k = 0 \\ 0 & else \end{cases} \tag{A-5}$$

Thus, we can derive the partial derivative of the inner part of the norm, $\frac{\partial \nabla^k x_{j,t}}{\partial x_{j,\tau}}$ as follows. Assume that $k \geq 1 \wedge t \geq k + 1$ (else, the partial evaluates to 1, if $k = 0$, or 0). From the definition of $\nabla^k x_{j,t}$ in (A-5), it follows that:

$$\frac{\partial \nabla^k x_{j,t}}{\partial x_{j,\tau}} = \left(\frac{\partial \nabla^{k-1} x_{j,t}}{\partial x_{j,\tau}} - \frac{\partial \nabla^{k-1} x_{j,t-1}}{\partial x_{j,\tau}} \right) \tag{A-6}$$

such that we can derive $\frac{\partial \nabla^k x_{j,t}}{\partial x_{j,\tau}}$ recursively, starting with $k = 1$:

$$\frac{\partial \nabla^1 x_{j,t}}{\partial x_{j,\tau}} = \left(\frac{\partial \nabla^0 x_{j,t}}{\partial x_{j,\tau}} - \frac{\partial \nabla^0 x_{j,t-1}}{\partial x_{j,\tau}} \right) \tag{A-7}$$

where

$$\frac{\partial \nabla^0 x_{j,t}}{\partial x_{j,\tau}} = \frac{\partial x_{j,t}}{\partial x_{j,\tau}} = \begin{cases} 1 \ if \ t = \tau \\ 0 \ else \end{cases} \tag{A-8}$$

and

$$\frac{\partial \nabla^0 x_{j,t-1}}{\partial x_{j,\tau}} = \frac{\partial x_{j,t-1}}{\partial x_{j,\tau}} = \begin{cases} 1 \ if \ t = \tau + 1 \\ 0 \ else \end{cases} \tag{A-9}$$

We derive the second term in (A-6) from the first term by shifting the time indices by one period:

$$\frac{\partial \nabla^k x_{j,t-1}}{\partial x_{j,\tau}} = \frac{\partial \nabla^k x_{j,t}}{\partial x_{j,\tau+1}} \tag{A-10}$$

By inserting (A-8) and (A-9), we can then express (A-7) as

$$\frac{\partial \nabla^1 x_{j,t}}{\partial x_{j,\tau}} = \begin{cases} 1, \text{if } t = \tau \\ -1, \text{if } t = \tau + 1 \\ 0 \ \text{else} \end{cases} \tag{A-11}$$

which yields the partial $\frac{\partial \nabla^k x_{j,t}}{\partial x_{j,\tau}}$ for all t and $k = 1$. From that, we derive the expressions for $k = 2$ analogously to (A-7):

$$\frac{\partial \nabla^2 x_{j,t}}{\partial x_{j,\tau}} = \left(\frac{\partial \nabla^1 x_{j,t}}{\partial x_{j,\tau}} - \frac{\partial \nabla^1 x_{j,t-1}}{\partial x_{Ij\tau}} \right) \tag{A-12}$$

Where we have already derived the first term in (A-11) and we can obtain the second term from (A-11) after shifting time indices by one period:

47

$$\frac{\partial \nabla^2 x_{j,t-1}}{\partial x_{j,\tau}} = \begin{cases} 1, & if\ t = \tau + 1 \\ -1, & if\ t = \tau + 2 \\ 0 & else \end{cases} \tag{A-13}$$

Such that the partial $\frac{\partial \nabla^k x_{j,t}}{\partial x_{j,\tau}}$ for all t and $k = 2$ can be expressed as

$$\frac{\partial \nabla^2 x_{j,t}}{\partial x_{j,\tau}} = \begin{cases} 1, & if\ t = \tau \\ -2, & if\ t = \tau + 1 \\ 1, & if\ t = \tau + 2 \\ 0, & else \end{cases} \tag{A-14}$$

We can derive the remaining partials for k higher than two similarly by repeating the steps between (A-12) and (A-14). We obtain the final gradient by inserting all results into (A-4) and (A-3).

Based on these derivations, researchers can implement EvoMap for a given static mapping method as follows. We assume that the chosen method has a non-negative cost function C_{static} (i.e., lower values correspond to better solutions) and can be optimized iteratively via gradient-based methods. Provided that method, one needs to proceed as follows:

First, adopt (or derive) the static gradient grad C_{static} entailing the partial derivatives $\begin{bmatrix} \frac{\partial C_{static}}{\partial x_{1,\tau}} \\ \vdots \\ \frac{\partial C_{static}}{\partial x_{n,\tau}} \end{bmatrix}$ for any

map layout $X_\tau \in (X_t)_{t=1,\dots,T}$. If any firm is not present at time τ, set its respective entry in the gradient to

zero. Then, combine these gradients for all periods: grad $\sum_{\tau=1}^{T} C_{static} = \begin{bmatrix} grad\ C_{static}(X_1) \\ \vdots \\ grad\ C_{static}(X_T) \end{bmatrix}$. Here, we

simply stacked all map layouts in a temporal order.

Second, derive the partial derivatives $\begin{bmatrix} \frac{\partial C_{temp}}{\partial x_{1,\tau}} \\ \vdots \\ \frac{\partial C_{temp}}{\partial x_{n,\tau}} \end{bmatrix}$ of $C_{temporal}$ with respect to X_τ for any map layout $X_\tau \in$

$(X_t)_{t=1,\dots,T}$ using (A-3) and the expressions that follow it. Analogously to the static gradients, stack them in temporal order to obtain the final temporal gradient grad $C_{temporal}$.

Finally, combine the two (stacked) gradients according to (A-1). One can then use iterative optimization techniques to find the cost-minimizing sequence of map layouts (e.g., using the same optimization routine commonly used for the given static mapping method).

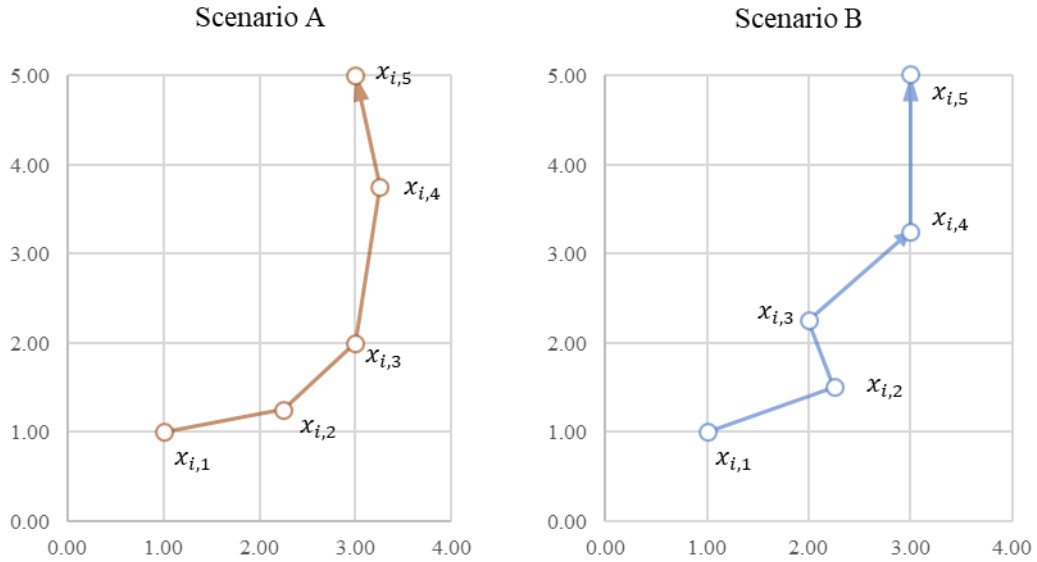## B. Numerical Examples for Penalty Derivations

Herein, we provide a numerical example for the derivation of the penalty terms in EvoMap's cost function. Specifically, we consider the same firm i under two hypothetical scenarios:

- Scenario A: Firm i moves gradually (left graph in Online Appendix Figure B-1)

- Scenario B: Firm i moves more erratically (right graph in Online Appendix Figure B-1)

We use these hypothetical trajectories to demonstrate how our cost function penalizes the solution in scenario B stronger than in scenario A (assuming everything else is equal). For both scenarios, Online Appendix Table B-1 displays the positions $x_{i,t}$, the resultant k-th order differences $\nabla^k x_{i,t}$ for $k \in \{1,2\}$, and the corresponding values of their norm $\|\nabla^k x_{i,t}\|^2$ which enter the cost function. The numerical example demonstrates that the cost function will take higher values (≈ 13.00 vs. ≈ 9.75) under scenario B. Thus, all else being equal (for instance, the static cost function values), the cost function favors the solution in scenario A. The example also demonstrates the added value of incorporating higher-order differences: When considering only $\|\nabla^k x_{i,t}\|^2$ for $k = 1$ both trajectories yield equivalent cost function values of 7.50.

Online Appendix Figure B-1: Two Different Trajectories

Notes: Axes correspond to the two map dimensions.

Online Appendix Table B-1: Numerical Example for Penalty Derivations

Symbolic Expression	Scenario A (gradual trajectory)					Scenario B (more erratic trajectory)				
	$x_{i,t}$ (Positions)	$\nabla^1 x_{i,t}$ (1st-Order Differences)	$\|\nabla^1 x_{i,t}\|^2$ (Sq. Eucl. Norm)	$\nabla^2 x_{i,t}$ (2nd-Order Differences)	$\|\nabla^2 x_{i,t}\|^2$ (Sq. Eucl. Norm)	$x_{i,t}$ (Positions)	$\nabla^1 x_{i,t}$ (1st-Order Differences)	$\|\nabla^1 x_{i,t}\|^2$ (Sq. Eucl. Norm)	$\nabla^2 x_{i,t}$ (2nd-Order Differences)	$\|\nabla^2 x_{i,t}\|^2$ (Sq. Eucl. Norm)
Time t = 1	$\begin{bmatrix}1.00\\1.00\end{bmatrix}$	--	--	--	--	$\begin{bmatrix}1.00\\1.00\end{bmatrix}$	--	--	--	--
Time t = 2	$\begin{bmatrix}2.25\\1.25\end{bmatrix}$	$\begin{bmatrix}1.25\\0.25\end{bmatrix}$	≈ 1.63	--	--	$\begin{bmatrix}2.25\\1.50\end{bmatrix}$	$\begin{bmatrix}1.25\\0.50\end{bmatrix}$	≈ 1.81	--	--
Time t = 3	$\begin{bmatrix}3.00\\2.00\end{bmatrix}$	$\begin{bmatrix}0.75\\0.75\end{bmatrix}$	≈ 1.13	$\begin{bmatrix}-0.50\\0.50\end{bmatrix}$	$= 0.50$	$\begin{bmatrix}2.00\\2.25\end{bmatrix}$	$\begin{bmatrix}-0.25\\0.75\end{bmatrix}$	≈ 0.63	$\begin{bmatrix}-1.50\\0.25\end{bmatrix}$	≈ 2.31
Time t = 4	$\begin{bmatrix}3.25\\3.75\end{bmatrix}$	$\begin{bmatrix}0.25\\1.75\end{bmatrix}$	≈ 3.13	$\begin{bmatrix}-0.50\\1.00\end{bmatrix}$	$= 1.25$	$\begin{bmatrix}3.00\\3.25\end{bmatrix}$	$\begin{bmatrix}1.00\\1.00\end{bmatrix}$	$= 2.00$	$\begin{bmatrix}1.25\\0.25\end{bmatrix}$	≈ 1.63
Time t = 5	$\begin{bmatrix}3.00\\5.00\end{bmatrix}$	$\begin{bmatrix}-0.25\\1.25\end{bmatrix}$	≈ 1.63	$\begin{bmatrix}-0.50\\-0.50\end{bmatrix}$	$= 0.50$	$\begin{bmatrix}3.00\\5.00\end{bmatrix}$	$\begin{bmatrix}0.00\\1.75\end{bmatrix}$	≈ 3.06	$\begin{bmatrix}-1.00\\0.75\end{bmatrix}$	≈ 1.56
$\sum_t \|\nabla^k x_{i,t}\|$	--	--	$= 7.50$	--	$= 2.25$	--	--	$= 7.50$	--	$= 5.50$
$\sum_t \sum_k \|\nabla^k x_{i,t}\|$	--	$= 9.75$				--	$= 13.0$			

Notes: $\|\cdot\|^2$ denotes the Squared Euclidean Norm (Sq. Eucl. Norm).

C. Additional Simulation Results

In what follows, we extend our simulation study to (1) investigate the benefits of EvoMap's adaptive regularization and smoothing properties, (2) test alternative specifications of two dynamic mapping metrics, and (3) improve Dynamic t-SNE's volatile performance by using EvoMap's optimization procedure.

Specifically, we introduce the following two variants of EvoMap to our simulation study: EvoMap (adaptive regularization only), which introduces firm-specific weights but ignores higher-order differences, and EvoMap (smoothing only), which considers higher-order differences but does not include any firm-specific weights. Considering the goodness-of-fit measures (HIT-RATE and C-CORR), we find that adaptive regularization increases dynamic goodness-of-fit (C-CORR increases), while static goodness-of-fit remains similar (HIT-RATE does not change). Thereby, the descriptive metrics reveal that the resultant maps are slightly less aligned, and trajectories become slightly less gradual (PERS decreases).

These findings show that adaptive regularization increases flexibility and allows EvoMap to recover actual market structure changes better. In contrast, smoothing results in more aligned positions (MIS-ALIGN decreases) and more gradual trajectories (PERS increases). Moreover, it also increases goodness-of-fit (recall that we evaluate goodness-of-fit against the simulated positions before adding any noise). Combining smoothing with adaptive regularization yields the highest goodness-of-fit of all alternatives.

Next, we introduce two additional metrics to our simulation study to demonstrate that our results are robust to alternative specifications of the evaluation metrics. Specifically, we propose the following alternative metrics:

- ALIGN (Alignment): An alternative measure to misalignment, using the average cosine similarity of successive positions across all periods and firms. Alignment varies between -1 and 1, where positive values indicate high alignment (and vice versa).

- CPA (Change Prediction Accuracy): An alternative measure to change correlation. Rather than simply using the correlation of simulated vs. recovered movement path lengths, we estimate a logistic regression model with a binary dependent variable "change" (1 if a firm was affected by one of the evolution scenarios within the simulation, 0 else) and a single independent variable "trajectory length" (total length of the movement path on the estimated market structure map). We evaluate its predictive accuracy. Since our sample is imbalanced with a large fraction of static and

a small fraction of dynamic positions, we use the F_1 score based on precision and recall as a measure of predictive accuracy.

Consistent with MIS-ALIGN and C-CORR, EvoMap also outperforms extant dynamic mapping approaches on ALIGN (Alignment) and CPA (Change Prediction Accuracy). Naturally, as the ALIGN metric is bounded between -1 and 1, it varies less than the unbounded MIS-ALIGN metric across different methods. See Online Appendix Table C-1 for details.

Finally, we improve the relatively poor performance of Dynamic t-SNE (Rauber et al. 2016), which they provide at https://github.com/paulorauber/thesne (latest version as of 03/22/2016), by using EvoMap's optimization procedure. Specifically, we run EvoMap without its adaptive regularization and smoothing components (such that its cost function equals Dynamic t-SNE's cost function up to a constant). Our optimization procedure improves all mapping quality metrics (see Online Appendix Table C-1). However, these improvements still fall short of EvoMap's mapping quality.

Online Appendix Table C-1: Simulation Results with Additional Evaluation Metrics

Mapping Approach	ALIGN	MIS-ALIGN	PERS	C-CORR	CPA	HIT-RATE
(I) EvoMap (t-SNE)	0.997 (0.004)	0.155 (0.079)	0.875 (0.079)	0.835 (0.145)	0.750 (0.288)	0.665 (0.105)
(II) EvoMap (t-SNE) - smoothing only	0.998 (0.004)	0.146 (0.076)	0.888 (0.071)	0.817 (0.147)	0.737 (0.289)	0.665 (0.105)
(III) EvoMap (t-SNE) - adaptive regularization only	0.993 (0.019)	0.240 (0.167)	0.512 (0.249)	0.716 (0.266)	0.656 (0.343)	0.658 (0.109)
(IV) Dynamic t-SNE - via EvoMap	0.994 (0.018)	0.224 (0.166)	0.539 (0.247)	0.693 (0.266)	0.637 (0.338)	0.659 (0.109)
(V) Dynamic t-SNE - Rauber et al. (2016)	0.935 (0.100)	20.636 (33.605)	0.370 (0.310)	0.291 (0.299)	0.228 (0.361)	0.401 (0.152)
(VI) Ex-post Alignment (t-SNE)	0.728 (0.291)	12.238 (14.496)	-0.379 (0.191)	0.129 (0.305)	0.069 (0.218)	0.642 (0.120)
(VII) Sequential Initialization (t-SNE)	0.965 (0.043)	2.474 (1.857)	0.044 (0.275)	0.287 (0.292)	0.214 (0.287)	0.634 (0.123)
(VIII) Fixed Initialization (t-SNE)	0.832 (0.199)	4.808 (3.992)	-0.317 (0.186)	0.078 (0.247)	0.064 (0.194)	0.643 (0.120)
(IX) Independent Mapping (t-SNE)	0.003 (0.125)	21.920 (16.093)	-0.484 (0.098)	-0.098 (0.215)	0.018 (0.119)	0.642 (0.120)

Notes: Reported metrics are averages over 2,187 simulation iterations corresponding to the parameter space reported in Appendix A. Standard deviation in parentheses. Dynamic t-SNE (Rauber et al. 2016): Dynamic t-SNE, as provided by its authors on GitHub. Dynamic t-SNE (via EvoMap): Dynamic t-SNE optimized via EvoMap's optimization procedure. We set hyperparameters as described in Appendix A. Their distributions are similar to the ones reported in Appendix A. For Dynamic t-SNE (Rauber et al. 2016), we set its remaining parameters to the default values provided by its authors.

D. Sensitivity Analysis

We investigate the sensitivity of EvoMap's dynamic mapping quality to the various market structure characteristics. To do so, we regress the evaluation criteria on the simulation parameters listed in Appendix Table A-1 (using linear regression with dummy-coded independent variables). We exclude one level of each simulation parameter in each regression model as its reference point.

Online Appendix Table D-1 reports the results for each of the four estimated regression models (1) to (4).

Online Appendix Table D-1: Regression Results for Relationships between EvoMap's Dynamic Mapping Quality and Simulation Parameters

Market structure characteristic	Simulation parameter	Value	(1) MIS-ALIGN	(2) C-CORR	(3) PERS	(3) HIT-RATE
		Constant	0.0965*** (0.0045)	0.7719*** (0.0108)	0.8195*** (0.0058)	0.6417*** (0.0059)
Evolution	Scenario	II (Shifts in positions)	0.0505*** (0.0029)	0.1075*** (0.0068)	0.0028 (0.0036)	-0.0082** (0.0037)
		III (Market entry)	0.0272*** (0.0029)	0.0792*** (0.0068)	0.0149*** (0.0036)	-0.0168*** (0.0037)
Number of firms	n	100	0.0449*** (0.0029)	0.0398*** (0.0068)	0.0341*** (0.0036)	0.0602*** (0.0037)
		250	0.0681*** (0.0029)	0.0336*** (0.0068)	0.0629*** (0.0036)	-0.0705*** (0.0037)
Number of dimensions	d	8	-0.014*** (0.0029)	0.0146** (0.0068)	0.0175*** (0.0036)	-0.0163*** (0.0037)
		16	-0.0271*** (0.0029)	0.0357*** (0.0068)	0.033*** (0.0036)	-0.0296*** (0.0037)
Number of submarkets	k	8	-0.0527*** (0.0029)	-0.007 (0.0068)	0.0133*** (0.0036)	0.0989*** (0.0037)
		12	-0.071*** (0.0029)	-0.0283*** (0.0068)	0.0037 (0.0036)	0.1177*** (0.0037)
Within submarket standard deviation	σ_{sub}	0.075	-0.0056* (0.0029)	-0.0114* (0.0068)	-0.0006 (0.0036)	0.0092** (0.0037)
		0.150	-0.0142*** (0.0029)	-0.0298*** (0.0068)	-0.0026 (0.0036)	-0.0116*** (0.0037)
Temporal noise	$\sigma_{temporal}$	0.025	0.0245*** (0.0029)	-0.026*** (0.0068)	-0.0168*** (0.0036)	-0.0104*** (0.0037)
		0.050	0.0582*** (0.0029)	-0.0769*** (0.0068)	-0.0441*** (0.0036)	-0.0413*** (0.0037)
Share of firms affected	ρ	0.100	0.0331*** (0.0029)	0.0227*** (0.0068)	0.018*** (0.0036)	-0.0043 (0.0037)
		0.150	0.0524*** (0.0029)	0.0364*** (0.0068)	0.0296*** (0.0036)	-0.0066* (0.0037)
Observations			2,187	2,187	2,187	2,187
R-squared			0.53	0.20	0.23	0.55

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p <0.10., std.: standard deviation

E. Further Implementation Details and Simulation Runtime Estimates

For this paper, we implement EvoMap using Python version 3.7.2. We use the following additional packages:

Online Appendix Table E-1: Core Implementation Details

Package Name	Version Used	Use-Case
numpy	1.21.2	Numerical calculations
scipy	1.7.1	Efficient calculation of Euclidean distances and norms
numba	0.53.0	Just-in-time compilation for speed optimization

Our results (not shown here) show that EvoMap's runtime increases approximately quadratically with the number of firms and linearly with the number of periods[5]. Thus, running EvoMap once takes roughly the same time as running an existing mapping method independently for each period (assuming both implementations are equally efficient in their computations). The only difference is the more complex calculation of the gradient. Online Appendix Table E-2 presents the average runtime for all methods used in our simulations and shows that EvoMap's more complex gradient does not substantially affect runtime in our simulations.

Online Appendix Table E-2: Runtime Comparison across Methods

Method	Average Runtime in Seconds
EvoMap (t-SNE)	140.17 (159.47)
Dynamic t-SNE (via EvoMap)	128.79 (156.42)
Dynamic t-SNE (Rauber et al. 2016)	250.32 (234.39)
Independent t-SNE	130.88 (157.04)

Notes: Averages across 2,187 simulation iterations. Standard deviation in parentheses.

F. Extended TNIC Analysis using EvoMap with MDS

The main article presents EvoMap's dynamic market structure map for the full sample of 1,092 firms. For our analysis of the full sample, we paired EvoMap with t-SNE. Here, we aim to extend this analysis using EvoMap paired with an alternative mapping method (specifically, metric MDS). Our objective is twofold: First, to

---

[5] Our implementation of EvoMap, used throughout our simulations, uses the following default parameters: For t-SNE, we set perplexity to 15, the initial learning rate to 20, initial momentum factor to .5, final momentum factor to .8, the switch iteration to 250, the early exaggeration factor to 4, and the maximum number of iterations to 2000. For the Stress-based alternatives (MDS and Sammon), we set the initial learning rate to 0.1 and the maximum number of iterations to 2000. We automatically adjust the learning rate (via a factor of 0.1) for all methods in case the gradient diverges during optimization.

demonstrate the benefits of EvoMap's ability to easily accommodate alternative static mapping methods for practical competitive analysis applications. And second, to test the robustness of our empirical findings towards the choice of mapping method. To do so, we generate a second dynamic market structure map using EvoMap paired with metric MDS. As metric MDS, however, does not scale well to many firms, we focus this analysis on a smaller sample of firms from the technology sector. We then test if the resultant (small) map for EvoMap paired with metric MDS resembles the same patterns as the corresponding area in the (large) map for EvoMap paired with t-SNE.

We motivate the analysis of market structure evolution with two different static mapping methods (i.e., t-SNE and metric MDS) as follows: Imagine a market analyst required a birds-eye view on how a market comprised of many firms and submarkets evolved. Because mapping quality tends to be superior with t-SNE in such large N settings, the analyst chooses to first pair EvoMap with t-SNE (similar to the analysis in the main article). Having identified some high-level trends and interesting firm trajectories, the analyst now wants to study a specific area of the dynamic market structure map more thoroughly (e.g., a specific submarket, a sector, or the set of a firm's imminent competitors). Because metric MDS provides some advantages in smaller N settings over t-SNE (e.g., better interpretability of the resultant map distances), the analyst pairs EvoMap with metric MDS to generate a second dynamic market structure map for the subsample of firms in the area of interest.

We replicate the above-outlined scenario for our empirical application as follows. We first select a small, technology-focused subsample of all firms. Our subsample includes the following 10 firms from the technology sector: AT&T, US Cellular, Comcast, Microsoft, Western Digital, HP, Oracle, eBay, Intuit, Apple. We then estimate a dynamic market structure map for this subsample using EvoMap paired with metric MDS (we set $\alpha = 9.1 \times 10^{-1}$ and $p = 2$).

Online Appendix Figure F-1: Dynamic Market Structure Map for 10 Technology Firms 1998 – 2017 (estimated using EvoMap paired with metric MDS)
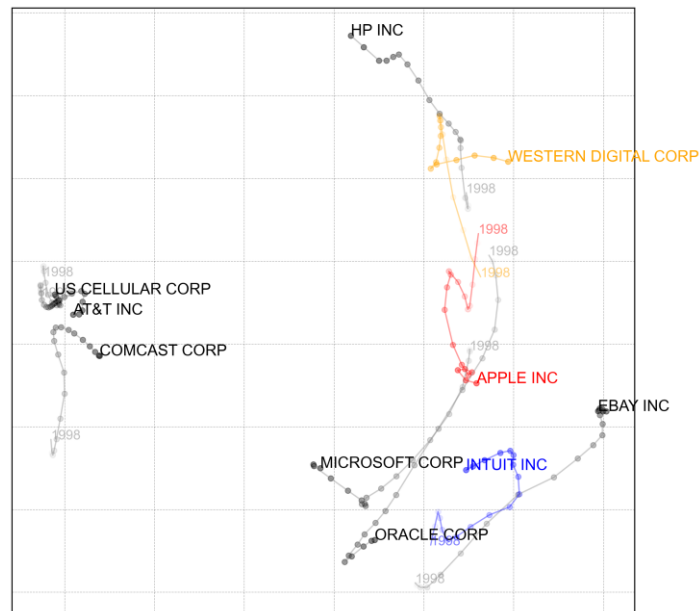
Notes: Trajectories correspond to 20 successive years between 1998 and 2017 (firm name labels last period).

While the smaller map reveals additional nuances, the overall insights do not change much. For example, Apple's trajectory diverges from Western Digital and converges with Intuit (as is the case in the larger dynamic market structure map in Figure 7). As such, our findings remain robust across two different samples and two different static mapping methods used with EvoMap.