

Franchising and Control Across Multiple Markets: The Case of the Convenience Store Industry*

Dennis Campbell^a, Srikant Datar^a, Tatiana Sandino^{b,*}

^a*Harvard Business School, Harvard University, Boston, MA 02163*

^b*Marshall School of Business, University of Southern California, Los Angeles, CA 90089*

March 1, 2007

Abstract:

Chain organizations operate units which are typically dispersed across different types of markets, and thus serve significantly diverging customer bases. Such market-type dispersion is likely to compromise the headquarters' ability to control its store managers' behavior and satisfy the divergent needs of different types of customers. In this paper we find evidence that franchising can be viewed as an organizational and contractual solution to the control problems resulting from market-type dispersion. Specifically, we show that market-type dispersion is related to the degree of franchising at the chain level as well as the probability of franchising a given store within a chain. Our results are robust to alternative definitions of market-type dispersion and to other determinants of franchising such as the stores' geographic distance from headquarters and geographic dispersion.

JEL Classification: D82, D86, L22, M41

Keywords: Control Systems, Franchising, Agency Costs, Market Dispersion, Retailing

* We thank Xiang Ao, Rajiv Lal, Kenneth Merchant, Taylor Randall, Scott Schaefer, William Simpson, Ingrid Vargas, Shelby Yu, and seminar participants at the University of Utah for their comments. We also thank David Huelsbeck and Deborah Strumsky for research assistance. All errors remain our own.

* Corresponding Author. Tel: 213.740.4842; Fax: 213.747.2815; email: sandino@marshall.usc.edu

Franchising and Control Across Multiple Markets: The Case of the Convenience Store Industry

1. Introduction

The chain organizational form has become one of the most prevalent forms of organizing service and retail industries. Baum et al. (2000) describe chains as “collections of service organizations, doing essentially the same thing, linked together into larger super-organizations.” According to Baum et al. (2000), geographic location is often the only difference among a chain’s units, and is therefore a critical element of an individual chain store’s success.

A key characteristic in chains is that different locations frequently lead particular stores to operate in widely different circumstances. We define such variation in location characteristics within a chain as *market-type dispersion*. The relevant location characteristics in retail chains and other service-provider chains are those affecting a store’s demand conditions (household average size, ethnicity, per capita income, etc.) as identified in the marketing literature (Hoch et al., 1995; Mulhern et al., 1998; Mittal et al., 2004). Market-type dispersion is not necessarily associated with geographic dispersion because (1) stores can be close to each other geographically but still have sharply diverging types of customers (e.g. convenience stores in the ethnic neighborhoods of Chicago serve customers of entirely different ethnicities, despite being just a few blocks apart), and (2) stores can be geographically distant among each other and yet have similar customer bases.

While store managers have authority over some operational decisions such as managing inventory levels, for most chains, significant strategic and operational

decisions are often retained at the chain level, such as advertising, purchasing, or selecting the product mix. It is more challenging to retain control over these overarching decisions at the chain level when market-type dispersion is high. Specifically, market-type dispersion leads to two control problems: First, from a corporate perspective, relative differences in local conditions make it more difficult for the chain headquarters to monitor and control individual stores (Landier et al., 2006). Second, from a demand perspective, a chain with wide ranging customer bases will have a harder time appealing to all its customers (Anderson and Mittal, 2000).

Previous literature in control systems in chain organizations, suggests franchising control mechanisms are used to provide incentives to the franchisees running the stores and to cope with higher monitoring costs in geographically dispersed stores (Brickley and Dark, 1987; Lafontaine, 1992). Interviews with chain managers also reveal that franchising mechanisms have other advantages over internal control systems when dealing with market-type dispersion. First, they attract franchisees with entrepreneurial and market-sensing skills that are superior to those possessed by company owned-store managers. Second, they attract entrepreneurs who are better able to bear risk. Third, they allow the chain to allocate risk to local franchisees who, thanks to their superior knowledge of the market and their flexibility to adapt their operating procedures to local conditions, perceive lower risk in operating the store than the risk perceived by the chain. As the Chief Operating Officer of a large convenience store chain in the Northeast U.S. explained

It makes sense to franchise stores where risk is higher. Franchisees are more entrepreneurial and would have more control over difficult situations such as adapting merchandising to local market conditions, and managing theft, shrinkage, and other aspects required to maintain performance.

In this paper we examine whether chains use franchising to minimize the control problems arising from market-type dispersion both at a chain and at a store level. In the first analyses, we use 2004 TDLinx data (obtained from A.C. Nielsen Company) from 420 convenience store chains¹ to explore the effect of market-type dispersion on chains' decisions to franchise all, some or none of its stores. Employing an ordinal logit model, we find that market-type dispersion is positively associated with the chains' decisions to franchise stores. This result is robust to other determinants previously explored in the literature including monitoring concerns arising from the chain's geographic dispersion (Brickley and Dark, 1987; Norton, 1988) as well as chain size—a proxy used to capture the maturity of the firm and its need to raise money—(Caves and Murphy, 1976; Lafontaine, 1992).

In the second part of our analyses, we focus on the 43 chains in our sample that own some stores and franchise others, and the 34,892 stores operated by those 43 chains. Conditional on a firm's decision to franchise only some of its stores, we use a fixed-effects logit model to explore the decision to franchise or not franchise individual stores as a function of the store-specific *market-type divergence* (i.e., the extent to which a store's market characteristics diverge from the most prevalent characteristics of the chain as a whole). We find evidence supporting an association between the decision to franchise and the store's market-type divergence that is robust to chain fixed effects as well as the store's demographic characteristics.

¹ We focus on the convenience store industry mainly because: (1) the heterogeneity of markets convenience stores serve varies widely across chains; (2) convenience stores essentially compete on location and are relatively undifferentiated in other dimensions, reducing the number of factors to be considered for empirical analysis.

Our study contributes to the literature related to franchising and the use of internal controls in geographically dispersed chain organizations. While several economics researchers have found empirical evidence that franchising agreements are used by chains to decrease agency costs in geographically dispersed locations (Caves and Murphy, 1976; Brickley and Dark, 1987; Norton, 1988; Brickley, Dark, and Weisbach, 1991), our study examines the decision to franchise in the context of market-type dispersion.² We show that the market-type dispersion in a chain plays an incremental role in the chain's decision to franchise its units, over and above geographic dispersion.

Our work also contributes to an emerging literature in control systems describing the relationship between organizational design and the availability of adequate control systems. Two emerging lines of research explore this relationship.

First, several papers have established a link between geographic expansion and control systems. For example, using data from 250 publicly traded U.S. firms in the years 1986-1999, Landier et al (2006) find that firms are geographically concentrated when information is more difficult to transfer over long distances. Similarly, Berger and DeYoung (2001) find that bank organizations that choose a greater geographic scope tend to be those with superior parental control. Sandino (2007) documents that young retail chains introducing control systems fitting their strategies tend to expand at a higher rate. This paper complements this line of research by suggesting franchising as an alternative control mechanism to deal with expansion into new markets, where agency costs are

² Interestingly, Fladmoe-Lindquist and Jacque (1995) find a result that is consistent with the notion that market-type dispersion is related to a firm's decision to franchise. They find that the propensity of a group of U.S. based service firms to franchise internationally is related to the "cultural" distance between the franchisor and its foreign franchisees. They capture "cultural distance" using a set of dummy variables indicating whether the foreign franchisees are located in one of six areas: Latin Europe, Nordic/Germany, Far East, Near East/Arabic, Latin America/Caribbean, and Independent country clusters. The Anglo area is utilized as the baseline.

exacerbated by less accessible information and by a greater need to rely on the store manager's efforts to adapt to the local circumstances.

Second, other papers have evaluated the link between the decision to delegate and the availability of control systems. For example, Abernethy et al (2004) and Demers et al. (2004) find evidence that the decision to delegate authority in an organization is complementary to its control systems (performance measurement and incentives systems). More specifically, Moers (2006) finds that the decision to delegate authority to middle managers depends on the ability of the firm's internal controls to solve the incentive problem. Such ability is captured by Moers (2006) by measuring the quality of contractible performance measures (their sensitivity, precision, and verifiability). In our study, we assume the ability of the firm's internal controls to solve the incentive problem in a chain organization is negatively related to the diversity of markets served by the chain's stores, given that such diversity makes it more difficult and costly to monitor the performance of a store using internal controls. However, we find evidence that this inability of the firm's internal controls to solve the incentive problem is not resolved by centralizing authority (in which case the chain might be limited in its expansion) but instead it is solved by using an alternative organizational design in the form of franchising.

The remainder of the paper is organized as follows: Section 2 provides our hypotheses development while Section 3 describes the sample selection. Section 4 provides a description of the market-dispersion measures utilized in our analyses, and Section 5 focuses on the chain level and store level empirical analyses, respectively. Section 6 concludes.

2. Hypotheses Development

Expansion into different markets may allow efficient chain organizations to strengthen (and derive economic rents from) their brand name; leverage their expertise and managerial skills; take advantage of economies of scale; and exploit the benefits of risk diversification (Berger and DeYoung, 2001). However, serving multiple markets may also result in negative consequences for the chain. Expansion into diverse markets can reduce efficiency as coordination, communication, and monitoring problems arise. Specifically, an increase in a chain's market-type dispersion generates two major control problems from the headquarters' perspective:

Greater agency conflicts from monitoring

The headquarters³ of a chain confronts an agency problem if the store managers' actions do not reflect the best interests of the company as a whole. Monitoring store managers in highly dispersed locations is a challenging task, and as such, it impairs the headquarters' ability to limit opportunistic store managers from engaging in shirking or perquisite-taking (Jensen and Meckling, 1976) and from providing a lower-quality service while free-riding on the chain's brand name (Brickley and Dark, 1987). Chains with stores serving heterogeneous markets experience more information asymmetries between the store managers and the chains' headquarters (Landier et al., 2006) than chains that are geographically dispersed but serve similar types of customers, given that different types of customers call for distinct ways of operating the stores which are likely to be better understood by the local managers than by the headquarters.

³ We assume the headquarters' interests represent the interests of the owners of the chain, that is, the interest of maximizing the chain's value.

Greater difficulties serving customers in different markets

While market-type dispersion has the effect of limiting the information flows to headquarters, it also generates a greater need to adapt to the tastes of divergent customers, increasing the value of the store manager's input in the sales process (Lafontaine, 1992). Operating a store in a highly divergent market-type requires market-sensing and entrepreneurial skills that are different from the operational skills that chains develop in their store managers.

A number of studies in marketing and service operations provides evidence that different types of customers place different importance on the attributes that lead to customer satisfaction and consumption in a chain organization (Anderson and Mittal, 2000; Mittal et al., 2004). For instance, Mattila (1999) finds that the importance customers place on the physical environment attribute of luxury hotels is associated with their ethnicity; Mittal et al. (2004) finds that the customers' responsiveness to automobile dealership services varies with their income, age, education, and ethnicity; while Campbell and Frei (2006) suggest that local market characteristics across bank branches, such as income and competition, affect customers' perceptions with respect to service wait-times. Thus, a chain serving a diversity of markets requires effective controls that will motivate store managers to increase their efforts to adapt to the tastes of local customers while coordinating their actions with those of the chain headquarters. An uncontrolled expansion could result in an inconsistent quality of service across stores that could hurt the chain's brand.

The control problems in a chain with high levels of market-type dispersion could potentially be mitigated by (a) increasing the level of ratification and monitoring of the

decisions made by the store managers, or (b) by providing incentives that would align the store managers' interests to those of the chain as a whole, where the store managers would bear the wealth effect of their actions (Fama and Jensen, 1983). On one hand, increasing the level of monitoring through internal control mechanisms can be extremely costly in organizations that are not only geographically dispersed but also dispersed across different types of markets.⁴ On the other hand, an extreme form of the incentive solution, where a store manager purchases all the residual claims of its units is also unlikely, as a chain benefits from retaining certain decision rights, particularly those related to the development and protection of its brand (Brickley and Dark, 1987). Furthermore, store managers in company-owned stores are risk-averse so extreme forms of the incentive solution could be very inefficient.

We next develop arguments (and later test) for why franchising is an organizational solution to the information and control problems in stores with high levels of market-type divergence (i.e., stores in markets the chain does not traditionally serve). These arguments reflect discussions we have had with executives of convenience store chains. First, franchisees possess the entrepreneurial and market-sensing skills needed to operate stores with high market-type divergence to a much greater extent than own-store managers (Bradach, 1997). Second, franchisees are entrepreneurs and thus are better able to bear risk compared to the average store manager. Third, in many cases, franchisees view the risk of operating these stores to be much lower than the risk as seen by the chain. Chains see these stores as more risky because their standard operating procedures

⁴ Notice however, that technological progress has facilitated the ability to control a firm's expansion. Focusing on the banking industry, Berger and Young (2002) find evidence that banking organizations' control over its affiliates has increased over time, while agency costs associated to geographic distance from headquarters have decreased over time.

and control mechanisms are not suitable for the stores in these markets. Franchisees have no such issues.

Charging a franchise fee achieves efficient risk sharing, given the franchisee's willingness to bear risk, and provides the incentives that limit the franchisee's shirking behavior. It motivates franchisees to exert the necessary effort to adapt to local needs and provide a high-quality service, as the franchisee retains the residual profit from its unit. At the same time, franchising is a coordinating mechanism, since it allows the chain to retain the decision rights necessary to maintain the quality and consistency of the brand by imposing a set of guidelines to operate the brand (pricing, product mix selection, etc.), controlling the level of advertising, and retaining the right to terminate the franchise agreement if the franchisee violates the chain's conditions to operate the brand (Brickley and Dark, 1987). Thus, franchising serves as an efficient control mechanism in the face of market-type dispersion. Based on these arguments, we predict that:

Hypothesis 1: A chain's decision to franchise some or all of its stores will be positively associated with the extent to which its stores are dispersed across different types of markets.

A potential limitation of testing hypothesis 1 is that market-type dispersion is an endogenous variable. The decision to expand into new markets resides with the chain headquarters. Thus, a chain level analysis may suffer from endogeneity due to (a) correlated omitted variables (i.e., there may be one or more unobserved variables excluded from our empirical analysis, such as the value of the chain's brand, that may explain both the decision to franchise and the decision to expand into new markets, and thus, may be the reason why we find a correlation between franchising and market-type dispersion), or (b) reverse causality (i.e., it may not be clear whether market-type

dispersion leads to more franchising, or a chain's decision to franchise affects the headquarters' decision to expand into new markets). These endogeneity concerns are mitigated if we test franchising decisions at the store level while controlling for chain fixed effects. Control problems become more severe for stores that operate in markets with marked differences from the predominant market characteristics in the chain as a whole. Thus, conditional on a chain's decision to franchise some but not all of its stores, we test the following prediction:

Hypothesis 2: Chains franchising some but not all of their stores are more likely to franchise units in locations where market characteristics diverge more from the most prevalent location characteristics in each chain as a whole.

3. Sample Description

We focus our analyses on the convenience store industry for various reasons. First, for empirical analysis, the heterogeneity of markets that a convenience store chain faces varies widely across chains and can be proxied by the variations in the location demographic characteristics of the chains' stores. Second, convenience store chains essentially compete on location and are relatively undifferentiated in other dimensions, reducing the number of factors to be considered for empirical analysis. Third, focusing on a relatively homogeneous industry allows us to control for several industry-specific conditions such as the degree of regulation or the extent to which store managers engage in chain-specific investments that would affect governance and monitoring costs (Williamson, 1985).

We use two different data sources for our analyses: the convenience stores dataset from the TDLinx⁵ 2004 Channel Database licensed to and facilitated by the National Association of Convenience Stores (NACS), which includes the location (address) and ownership composition (whether or not each store is franchised) of all convenience stores in the United States, and the Environmental Systems Research Institute (ESRI) 2004 Business Location Data, which tracks the locations of all business entities in the United States, as well as demographic information from the 2000 U.S. Census. We analyze the extent to which market-type dispersion affects franchising decisions at the chain-level using data from 420 convenience store chains with at least 20 stores each (i.e., we exclude “mom and pop” convenience stores), and at a store-level using data from 34,892 stores from the subset of 43 chains (out of the 420) that own some stores and franchise others.

Notice that the latest U.S. Census was taken in 2000, while the TDLinx data corresponds to 2004. As a result, the store location characteristics observed are lagged measures with respect to chain characteristics. Since the four year difference between the datasets is a relatively short period of time, population movements have most likely introduced random, rather than systematic, noise. Thus, the time difference for the observations in the datasets should not bias the results of our analyses.

4. Measurement of Market-Type Dispersion and Divergence

Market-type variation can be measured both at the chain level and at the store level.

At the chain level, the degree of market-type *dispersion* is the amount of variation in

⁵ TDLinx is a division of A.C. Nielsen (U.S.), Inc. Its TD Channel Database covers the population of chain stores in the United States for several service and retail industries, one of which is convenience store chains.

location characteristics among stores within a chain. For example, a chain with all of its stores in Santa Monica, CA would have a lower level of market-type dispersion than a chain with stores in Downtown LA, Santa Monica, and Sherman Oaks. At the store level, the degree of market-type *divergence* is how much a store's location demographic characteristics differ from the average demographic characteristics in a chain. For example, if a chain has one store in Harvard Square and the rest of its stores in Roxbury, then the store in Harvard Square should have higher market-type divergence than the stores in Roxbury. In the next sub-sections we describe alternative measures of market-type dispersion at the chain level and market-type divergence at the store level.

4.1 Measurement of Chain Level Market-Type Dispersion

Earlier we defined market-type dispersion as the variation of certain location characteristics among stores within a chain. The relevant location characteristics are the ones that affect the stores' demand conditions. Therefore, to construct a measure for market-type dispersion, we first identify the location characteristics that are most relevant to the demand of a convenience store. Then, we discuss several possibilities for measuring dispersion.

The construction of market-type dispersion is particularly important because this measure has not been studied in previous literature. Therefore, we construct market-type dispersion in four different ways and compare the regression results obtained with the different measures.

The marketing literature has identified population density, average household size, per capita income, age, and ethnicity as some of the most relevant drivers of customer purchasing behavior in grocery and retail stores (Gupta and Chintagunta, 1994; Hoch et

al., 1995; Kalyanam and Putler, 1997; Mulhern et al., 1998). Examples of findings from previous research along these five dimensions include:

Population Density	(i) stores located in highly populated (urban) areas experience higher sales growth (Gómez et al., 2003); (ii) customers in less populated (rural) areas are more likely to travel beyond their local market, resulting in lost sales for the rural stores (Miller and Kean, 1997);
Household Size	(iii) individuals from larger households prefer generic products and larger brand sizes, they are more price sensitive, and have a higher propensity to use off-price coupons due to the large amount of disposable income they spend in groceries (Bawa and Shoemaker, 1987; Gupta and Chintagunta, 1994; Hoch et al., 1995)
Per capita income and age	(iv) customers that are younger, better educated ⁶ , and/or wealthier have higher opportunity costs (due to a high marginal value of time), and thus, are less price and promotion sensitive (Gupta and Chintagunta, 1994; Hoch et al., 1995); (v) wealthier customers are more responsive to displays and feature advertising, consistent with the customers' desire to decrease in-store (displays) and out-store (advertising) search costs (Boatwright et al., 2004);
Ethnicity	(vi) African-Americans spend relatively more money on generic grocery purchases than Whites (Wilkes and Valencia, 1985); (vii) the percentage of African-American and Hispanic consumers in an area is associated with higher price sensitivity (Mulhern and Williams, 1994; Hoch et al., 1995; Mulhern et al., 1998).

All of these location characteristics are available from the ESRI dataset (which contains data from the 2000 U.S. Census) at a zip-code level. We use the U.S. Census' definition of per capita income (in dollar terms) and average household size (in terms of number of persons). We compute population density as the number of inhabitants per square mile. For ethnicity, we calculate the percentage of white population, while we utilize the median age of the population to calculate age.

⁶ Although we did not have a measure for education, Boatwright, Dhar and Rossi (2004) find that wealth and education are highly correlated. Thus, including only one of the variables in our analysis should provide adequate results.

We discuss different ways of constructing dispersion measures along the five location characteristics described above. One set of measures is based on the variation of individual location characteristics, while a second set of measures is defined utilizing cluster analysis.

Market-Type Dispersion Based on Variation of Location Characteristics

One way to measure dispersion is to obtain the standard deviation of each location characteristic L across the stores in a given chain i . However, a disadvantage of the standard deviation measure is that it is measured in the same units as the location characteristic L (e.g., the standard deviation of per capita income is measured in dollars), and thus, it cannot be compared to other location characteristics. To address this issue, we calculate the dispersion of each location characteristic L by normalizing the standard deviation by the average value of the location characteristic across all stores within the chain i :

$$Normalized\ Dispersion_{Li} = \frac{Std\ Deviation_{Li}}{Mean_{Li}}$$

The normalized dispersion measure has no units and thus can be compared along different location characteristics. This measure gives a sense of the amount of variation among stores in a chain for a particular location characteristic, in proportion to the typical value of that location characteristic within the chain. To construct the chain's overall market-type dispersion we aggregate the normalized dispersion measures of each chain i either through addition or multiplication:

$$MDispersion_Sum_i = \sum_{L=1}^5 Normalized\ Dispersion_{Li}$$

$$MDispersion_Mult_i = \prod_{L=1}^5 Normalized\ Dispersion_{L_i}$$

The added dispersion measure would assign equal weights to the five location characteristics and treat them independently. The multiplied aggregated dispersion differs from the added aggregated dispersion because it contains interaction effects. Thus, *MDispersion_Mult* is more sensitive to extremely high or low dispersion values on any single location characteristic. For example, this term would place more (less) weight on the average household measure when income dispersion or some other location characteristic is also high (low).

Market-Type Dispersion Based on Cluster Analysis

A different approach for measuring market-type dispersion is to segment the markets using cluster analysis and then to observe the number of clusters (or “market-types”) served by each of the chain stores. In line with the clustering approach of most segmentation studies in marketing research (Punj and Stewart, 1983; Chatuverdi et al., 1997), we cluster the 29,827 zip-codes in the ESRI database into groups with similar location characteristics (i.e. population density, average household size, per capita income, ethnicity, and age). We construct these clusters following three steps (Punj and Stewart, 1983, Sharma, 1996):

1. We standardize each location variable across the whole sample of zip-codes by mean centering it and dividing it by the location characteristic’s standard deviation. Thus each location variable is measured in terms of the number of standard deviations it deviates from the mean.
2. We estimate an *a priori* cluster solution using a hierarchical centroid clustering technique⁷, to (a) determine a candidate number of clusters, and (b) identify an

⁷ The hierarchical centroid clustering technique aggregates the zip-codes into groups by progressively grouping (in each iteration) the two closest zip-codes or groups of zip-codes, based on the euclidean distance among their location characteristics (or the average location characteristics in a zip-code group already formed, also known as the centroid of the group). This clustering technique suggests that the

initial partition of the clusters that would not be affected by outliers (i.e. our initial partition does not consider small groups with less than 20 zip-codes).

3. Using the *a priori* cluster solution obtained in step 2 as a starting point, we utilize a non-hierarchical (K-means) technique⁸ to obtain a final cluster solution. Based on this procedure, we obtain 25 groups of zip-codes with an R-square⁹ of 81.1%, and R-squares that range between 77% (for the average household size) and 85% (for the degree of ethnicity) for the five location characteristics, suggesting our cluster solution has a reasonable level of homogeneity within groups and the clusters are well separated based on all the location variables.

Using the 25 zip-code clusters (or “types of markets”), we construct two additional market-type dispersion measures defined as the number of clusters where the stores in chain *i* operate (*MDispersion_NClust_i*), and a dispersion measure based on a herfindahl index defined as:

$$MDispersion_HHI_i = \left[1 - \sum_{x=1}^{25} \left(\frac{\text{number of stores in cluster } X \text{ in chain } i}{\text{total number of stores in chain } i} \right)^2 \right] \times 100$$

The *MDispersion_NClust_i* values vary between 1 and 25, while the *MDispersion_HHI_i* measure has a minimum value of zero, if all stores in a chain are located in a single type of market (cluster), and a maximum value of 96 if a chain has exactly 1/25th of its stores in each of the 25 clusters.

The market dispersion measures based on clusters differ from the dispersion measures obtained from the variation of each of the location characteristics in different ways. First,

number of clusters should be determined at the point at which grouping the next two closest zip-codes or groups of zip-codes would significantly decrease the homogeneity within groups. A reason to refine the hierarchical cluster solution in step 3 is that once an observation is assigned to a cluster, hierarchical clustering does not allow that it be reassigned to another cluster.

⁸ The *a priori* cluster solution from the hierarchical clustering consists of 25 clusters and their respective centroids (average location characteristics within the cluster). The non-hierarchical K-means technique reassigns the zip-codes in the *a priori* cluster solution to the cluster whose centroid is closest and then recalculates the new centroids of the 25 new clusters. Reassignment continues until every case is assigned to the cluster with the nearest centroid and no further changes occur in the clusters.

⁹ The R-square is the ratio between the [standard deviation between groups] and [the sum of the standard deviation between groups and the standard deviation within groups].

MDispersion_NClust and *MDispersion_HHI* are less sensitive to outliers in location characteristics. For instance, having a store location with average per capita income much higher than the rest of the chain would significantly increase the standard deviation of per capita income within the chain, but would have little effect on the *NClust* or Herfindahl calculations of market-type dispersion. Second, *MDispersion_NClust* and *MDispersion_HHI* lose information on how stores are distributed within each cluster. Finally, *MDispersion_NClust* and *MDispersion_HHI* do not discriminate among which clusters the stores fall under. For instance, if there are two chains A and B, where chain A has half of its stores in Cluster 1 and half of its stores in Cluster 2, while chain B has half of its stores in Cluster 1 and half of its stores in Cluster 3, then both chains would have the same *MDispersion_NClust* and *MDispersion_HHI* values. However, if the difference between the location characteristics of Clusters 1 and 2 is greater than the difference in location characteristics in Clusters 1 and 3, then under *MDispersion_Sum* and *MDispersion_Mult*, chain A would be more dispersed than chain B.

4.2 Measurement of Store Level Market-Type Divergence

As defined earlier, market-type divergence at the store level is the extent to which the store's location characteristics differ from the average location characteristics in the chain. To differentiate these measures from those at the chain level, we refer to store-level dispersion measures as "market divergence" measures. As in the case of the chain level dispersion measures, we use population density, average household size, per capita income, ethnicity, and age, as the relevant location characteristics.

We construct two sets of market divergence measures: the first is based on the divergence between the store's individual location characteristics and the chain's average

characteristics, while the second is defined in terms of the clusters identified for each chain.

Market-Type Divergence Based on Individual Location Characteristics

To calculate market divergence at the store level, we calculate the absolute value of the difference between the value of a location characteristic L for store j , and the average value of location characteristic L across all stores in chain i , that is, $\left|L_{ij} - \bar{L}_i\right|$. This measure has different units for different location characteristics (e.g. per capita income divergence is measured in dollar terms, while average household size is measured in number of persons). To make the measures of different location characteristics comparable, we normalize this divergence measure by the standard deviation of the L values for all stores in chain i . Thus, we define for each location characteristic L :

$$Normalized\ Divergence_{Lij} = \frac{\left|L_{ij} - \bar{L}_i\right|}{std\ deviation_{L_i}}$$

This measure indicates how many standard deviations store j 's location characteristic L is away from the chain i 's average. We aggregate this normalized divergence metrics, through either multiplying or adding the *Normalized Divergence* measure across the five different location characteristics:

$$MDivergence_Sum_{ij} = \sum_{L=1}^5 Normalized\ Divergence_{Lij}$$

$$MDivergence_Mult_{ij} = \prod_{L=1}^5 Normalized\ Divergence_{Lij}$$

As in the case of the chain level dispersion measures, the aggregated divergence metric constructed through addition would treat the five market characteristic dimensions

as independent of one another, while the aggregated divergence measure constructed through multiplication would contain interaction effects for the different location characteristics.

Market-Type Divergence Based on Cluster Analysis

Using the 25 market clusters identified at the chain level, we construct two measures of store level divergence. The first one, *MConcentration_sameclust_{ij}*, indicates the percentage of stores in chain *i* belonging to the same cluster as the store *j*. The second captures the divergence between the store's cluster and the most frequent cluster in the chain, weighted by the frequency of the main cluster. The measure is calculated as follows:

$$MDivergence_mainclust_{ij} = \%mainclust_i * distance_{ij}$$

where,

%mainclust = is the percentage of stores in the most frequent cluster in chain *i*

distance = is the distance between the centroids of the store *j*'s cluster and the chain's

most frequent cluster, calculated as $\sum_{L=1}^5 \left| \bar{L}_{cj} - \bar{L}_{cm} \right|$, where \bar{L}_{cj} is the average value of

location characteristic *L* in store *j*'s cluster, while \bar{L}_{cm} is the average value of location

characteristic *L* in the chain's most frequent cluster.¹⁰

To summarize, the two methods for calculating dispersion at the chain level and divergence at the store level are: (1) measures based on the variation of individual location characteristics that lead to the *MDispersion_Sum* and *MDispersion_Mult* measures at the chain level, and to the *MDivergence_Sum* and *MDivergence_Mult*

¹⁰ Whenever the chain had more than one "most frequent" cluster, then the distance is calculated relative to the closest "most frequent" cluster.

measures at the store level; and (2) measures based on 25 different clusters (types of markets), which include *MDispersion_NClust* and *MDispersion_HHI* at the chain level, and *MDivergence_mainclust* and *MConcentration_sameclust* at the store level.

5. Research Methodology and Empirical Results

5.1 Chain-Level Methodology

To test Hypothesis 1—i.e., to investigate whether a positive association exists between a chain’s decision to franchise its stores and market-type dispersion—we perform multivariate tests, where the dependent variable captures the chain’s decision to franchise none, some, or all of its stores and the main explanatory variable consists of the chain’s market-type dispersion (*MDISPERSION* measured as *MDispersion_Sum*, *MDispersion_Mult*, *MDispersion_NClust*, or *MDispersion_HHI* as described in Section 4). A general specification of the model is described by the following equation:

$$FRANCHISING_i = f(MDISPERSION_i, CHAIN_STATEHHI_i, CHAIN_NSTORES_i)$$

Our multivariate approach consists of estimating an Ordinal Logit Model where the franchising variable is a categorical variable equal to 0 if the chain owns all of its stores, 1 if it franchises some but not all of its stores, and 2 if it franchises all of its stores. We include two control variables, identified by the previous literature as determinants of a chain’s decision to franchise its stores. The first measure, *CHAIN_STATEHHI* is defined as:

$$\left[1 - \sum_{x=1}^{50} \left(\frac{\text{number of stores in state } X \text{ in chain } i}{\text{total number of stores in chain } i} \right)^2 \right] \times 100$$

This measure varies between 0 (if all stores in a chain are located in one state) and 98 (if a chain has 1/50th of its stores in each of the 50 states). Monitoring costs (and thus, the

propensity to franchise) tend to be higher if stores are spread across multiple locations and if stores are isolated, instead of surrounded by other stores from the same chain (Brickley and Dark, 1987; Brickley et al., 1991). *CHAIN_STATEHHI* captures both of these dimensions as it measures the extent to which a chain's stores are dispersed across multiple states.¹¹ Consequently, we expect that *CHAIN_STATEHHI* will be positively related to franchising.. The second measure, *CHAIN_NSTORES* is the number of stores in the chain. On one hand, this measure captures the level of maturity of the firm and should indicate a smaller need for franchising as a means to raise capital (Carney and Gedajlovic, 1991; Lafontaine, 1992). Such interpretation would lead us to expect a negative relation to franchising. On the other hand, franchising and the chain's number of stores are endogenously determined since chains that franchise more might be able to grow larger, leading to a positive relation between both measures. Thus we make no prediction on this control variable.

5.2 Store-Level Methodology

Conditional on a chain's decision to own some of its stores and franchise others, we examine whether a positive association exists between franchising and market-type dispersion at a store level (i.e., hypothesis 2). We run the following logit regression model on the sub-sample of stores j belonging to chains i that franchise some, but not all of their stores:

$$Pr (FRANCHISE_{ij}) = f (MDIVERGENCE_{ij}, DISTANCE HQ_{ij}, \\ NSTORES_ZIPCHAIN_{ij}, CHAIN FIXED EFFECTS_{ij})$$

¹¹Measuring geographic dispersion of a chain at the state-level is consistent with previous research on chain-level franchising decisions (Scott 1995). All of our results are robust to measuring geographic dispersion using more granular geographic areas such as counties or zip codes.

The dependent variable of this model is a dummy equal to 1 if the store is a franchisee and zero otherwise. The main explanatory variable in this model is the *MDIVERGENCE* measure which can be measured using any of the store-level market divergence measures: *M Divergence_Sum*, *M Divergence_Mult*, *M Divergence_mainclust*, and *M Concentration_sameclust* (see Section 4). We expect the first three of these measures to be positively related and the fourth measure to be negatively related to the store's likelihood of being franchised.

We control for geographic dispersion using two control measures: *DISTANCE_HQ*, which is the distance in miles between the store's zip code and the headquarters' zip code, and *NSTORES_ZIPCHAIN* which is the number of same chain stores within the store's zip-code. The former measure is expected to be positively related, while the latter measure is expected to be negatively related, to the store's likelihood of being franchised. Finally, we control for chain fixed effects. Controlling for chain fixed effects allows us to mitigate omitted variables concerns at the chain-level, such as the chain's need to raise money or the chain's value of the brand (Brickley and Dark, 1987; Lafontaine, 1992).¹²

¹² For our store-level analyses, we estimate logits with chain-indicators. It is well known that coefficient estimates from logit models which include dummy variables to directly estimate fixed effects are biased (Madalla 1987). However, holding the number of cross-sectional groups constant, the coefficient bias in these models tends to 0 as the number of within-group observations increases (Wooldridge 2002). For our store-level analyses, we have only 43 chains with an average of 811 stores per chain and only 4 of these chains operate less than 50 stores (and more than 20). Thus, our panel dataset is characterized by a small number of groups, each with a large number of within-group observations and is ideally suited for estimation by logit with group (e.g chain) dummies. We also tested the robustness of our results to the fixed-effects logit model of Chamberlain (1980). However, this model is computationally difficult for datasets with very large numbers of within group observations (Greene 1993; Baltagi 1995). We were not able to get this model to converge when including the largest chains in our sample (e.g. chains with greater than 1,000 stores). We estimated the Chamberlain fixed-effects logit model on the sample of chains with less than 1,000 stores each and found substantively similar results to those reported in Table 7. For the sample of chains with over 1,000 stores each, we also estimated a basic logit specification with chain indicators and, again, found substantively similar results. This latter sample would be highly unlikely to suffer from biased coefficient estimates as it consists of only 10 chains with an average of 2,594 stores per chain.

5.3 Summary Statistics

Summary statistics for our sample of 420 chains with at least 20 stores are provided in Table 1. Convenience store chains in our sample vary widely on all measures of market dispersion. Chains compete in approximately 11 different "market-types" (*MDISPERSION_NCLUST*) on average and range from competing across only two different market types to competing across 24 of the 25 market types identified in our cluster analysis of zip code demographics. Chains in our sample also show considerable variation in geographic dispersion ranging from concentrating all stores within a single state (*CHAIN_STATEHHI=0*) to coming close to equally dispersing stores across all 50 states (*CHAIN_STATEHHI=96*). The average chain in our sample has 152 stores.

Only 14% of chains in our sample operate any of their stores as franchises. Further information on the distribution of franchise activity in our sample is provided in Table 2 which shows that 10.2% of sample chains organize by franchising some, but not all stores while 3.3% operate all stores as franchises. Chains that do franchise some, but not all, stores vary widely in their franchising decisions, operating approximately 53% of their stores as franchises on average, with the proportion of stores franchised ranging from 0.2% to 99.6%. Table 3 demonstrates that our measures of chain-level market-type dispersion show only small to moderate correlations with measures of geographic dispersion.

Summary statistics for the sample of 34,892 stores belonging to the 43 chains which franchise some, but not all of their stores are provided in Table 4 which shows that 69% of such stores are franchised. Stores in the sample vary widely on all measures of market-type divergence with the average store operating in a "market-type"

(*MCONCENTRATION_SAMECLUST*) in which its parent chain concentrates 7.6% of its stores. Not surprisingly, there is also wide variation across all the demographic characteristics used to construct the market-type divergence measures. The average store is 658 miles away from chain headquarters and operates in a zip-code with 2 stores from the same chain. Table 5 demonstrates that our measures of store-level market-type divergence show only small correlations with measures of geographic dispersion.

5.4 Chain-Level Empirical Results

The ordinal-logit specifications in Table 6 demonstrate results on the association between the likelihood of franchising and geographic dispersion that are consistent with prior literature on the determinants of franchising. Consistent with higher monitoring costs associated with geographic dispersion of business units (Brickley and Dark, 1987; Scott, 1995), chains that are more geographically disperse (*CHAIN_STATEHHI*) are more likely to franchise some or all of their stores ($p < .05$ in all specifications). The results in Table 6 provide strong evidence in support of Hypothesis 1. When measured by *MDISPERSION_SUM*, *MDISPERSION_MULT*, or *MDISPERSION_HHI*, market dispersion is positively and significantly related to the likelihood of franchising ($p < .10$ in all specifications).¹³

To gauge the economic significance of these results, our coefficient estimates suggest that, holding size and geographic dispersion constant at their mean values, a chain with the mean level of market dispersion measured as *MDISPERSION_HHI* has a 96.3%, 3.2%, and 0.5% probability of franchising no, some, or all stores respectively. By contrast, holding all other variables at their mean values, a chain for which

¹³ We replicate our analysis by substituting the dependent variable with the percentage of stores franchised by each chain (including 0% for those that franchised none, and 100% for those that franchised all stores) and find essentially the same results.

MDISPERSION_HHI is one standard deviation above the mean has a 71.2%, 24.1%, and 5.6% probability of franchising no, some, or all stores respectively. Overall, these results support the hypothesis that, after controlling for geographic dispersion, chains are more likely to use franchising when operating across diverse market-types (Hypothesis 1).

5.5 Store-Level Empirical Results

The logit specifications in Table 7 demonstrate results on the association between the likelihood of franchising a given store and measures of geographic distance and concentration which are consistent with prior literature (e.g., Brickley and Dark 1987). Geographic distance from headquarters (*DISTANCE_HQ*) is positively and significantly ($p < .001$) related to the probability of franchising a store while the number of stores in the same zip code operated by the chain with which the store is affiliated (*NSTORES_ZIPCHAIN*) is negatively and significantly ($p < .001$) related to the probability of franchising a store in all specifications.

The results in Tables 7 provide some evidence in support of Hypothesis 2. Contrary to our prediction, the proportion of the chains stores operating in the same cluster as a particular store (*MCONCENTRATION_SAMECLUST*) is *positively* related to the probability of franchising that store. However, when measured by *MDIVERGENCE_SUM*, *MDIVERGENCE_MULT*, or *MDIVERGENCE_MAINCLUST*, market-type divergence is positively and significantly related to the likelihood of franchising as predicted ($p < .001$ for all measures).

To gauge the economic significance of these results, our coefficient estimates suggest that, holding distance to headquarters and the number of same-chain stores in the zip

code constant at their mean values, the probability of a store with the mean level of market-type divergence being franchised is approximately 42% regardless of the measure of market-type divergence used. By contrast, holding all other variables at their mean values, the probability of the most divergent store in a chain being franchised is approximately 61%, 67%, or 50% on average when market-type divergence is measured by *MDIVERGENCE_SUM*, *MDIVERGENCE_MULT*, or *MDIVERGENCE_MAINCLUST* respectively. Overall, these results support the hypothesis that, after controlling for geographic dispersion, chains are more likely to franchise stores located in markets with characteristics that diverge most significantly from the predominant characteristics of markets in which the chain operates (Hypothesis 2).

6. Conclusions

Chain organizations operate units which are typically dispersed across different types of markets, and thus serve significantly diverging customer bases. Such market-type dispersion is likely to compromise the headquarters' ability to control its stores for two reasons: First, relative differences in local conditions make it difficult to monitor the store manager's behavior. Second, a chain with wide-ranging customer bases will have a harder time serving its customers and will need to rely more heavily on the store managers' ability to adapt to local needs.

In this paper we show that chains experiencing higher levels of market-type dispersion are more likely to use franchising, presumably to cope with the control problems described. We show that a chain's decision to franchise some or all of its stores is related to market-type dispersion, measured in several alternative ways, and that this

result is robust to other determinants previously explored in the literature including the chain's geographic dispersion. We also examine the decision to franchise at a store level for the subset of chains that own some stores and franchise others. We predict and find that stores whose location characteristics are more divergent from the most prevalent location characteristics of the chain as a whole, are more likely to be franchised.

Our study is relevant to the academic literature since it identifies an important factor (market-type dispersion) that seems to lead companies to franchise their stores, and to practitioners, such as managers and consultants, dealing with control challenges related to a chain's geographic expansion into different markets. However, our study is limited in the extent to which franchising is explained, as we are unable to observe differences in the franchising contracts utilized by the chains analyzed (e.g., fixed fees, royalties).

References

- Abernethy, M.A., Bowens, J. and L. van Lent. 2004. Determinants of Control System Design in Divisionalized Firms. *The Accounting Review* 79 (3): 545-570.
- Anderson, E.W. and V. Mittal. 2000. Strengthening the Satisfaction-Profit Chain. *Journal of Service Research* 3 (2): 107-120.
- Baltagi, B. 1995. *Econometric Analysis of Panel Data*. New York, NY: John Wiley and Sons.
- Baum, J.A., Li, S.X. and J.M. Usher. 2000. Making the Next Move: How Experiential and Vicarious Learning Shape the Locations of Chains' Acquisitions. *Administrative Science Quarterly* 45 (4): 766-801.
- Bawa, K. and R. Shoemaker. 1987. The Coupon-Prone Consumer: Some Findings Based on Purchase Behavior Across Product Classes. *Journal of Marketing* 51 (October): 99-110.
- Berger, A. and R. DeYoung. 2001. The Effects of Geographic Expansion on Bank Efficiency. *Journal of Financial Services Research* 19 (2,3): 163-184
- Berger, A. and R. DeYoung. 2002. Technological Progress and the Geographic Expansion of the Banking Industry. Working Paper, Federal Reserve System.
- Boatwright, P., Dhar, S. and P. E. Rossi. 2004. The Role of Retail Competition, Demographics and Account Retail Strategy as Drivers of Promotional Sensitivity. *Quantitative Marketing and Economics* 2: 169-190.
- Bradach, J. L. 1997. Using the plural form in the management of restaurant chains. *Administrative Science Quarterly* 42: 276-303.
- Brickley, J.A. and F.H. Dark. 1987. The Choice of Organizational Form. The case of Franchising. *Journal of Financial Economics* 18: 401-420.
- Brickley, J.A., Dark, F.H. and M.S. Weisbach. 1991. An Agency Perspective on Franchising. *Financial Management* 20: 27-35
- Campbell, D. and F. X. Frei. 2006. Optimal Local Capacity Decisions in Services: Theoretical Analysis and Empirical Evidence. Working Paper, Harvard Business School, Harvard University.
- Carney, M. and E. Gedajlovic. 1991. Vertical Integration in Franchise Systems: Agency Theory and Resource Explanations. *Strategic Management Journal* 12: 607-629.

- Caves, R. E. and W. F. Murphy. 1976. Franchising: Firms, Markets, and Intangible Assets. *Southern Economic Journal* 42: 572-586.
- Chamberlain, G., 1980. "Analysis of Covariance with Qualitative Data," *Review of Economic Studies*, 47, pp. 225-238.
- Chatuverdi, A.J.; Carroll, D.; Green, P.E. and J.A. Rotondo. 1997. A Feature-Based Approach to Market Segmentation Via Overlapping K-Centroids Clustering. *Journal of Marketing Research* 34 (3): 370-377.
- Demers, E.A., M.B. Shackell, and S.K. Widener. 2004. Incentives, performance measurement, and delegation: An empirical analysis of interdependencies in organizational architecture. Working paper, University of Rochester.
- Fama, E.C. and M.C. Jensen. 1983. Separation of Ownership and Control. *Journal of Law and Economics* 26 (2): 301-326.
- Fladmoe-Lindquist, K. and L.L. Jacque. 1995. Control Modes in International Service Operations: The Propensity to Franchise. *Management Science* 41 (7): 1238-1249.
- Gómez, M.I., McLaughlin, E.W. and D.R. Wittink. 2003. Do Changes in Customer Satisfaction Lead to Changes in Sales Performance in Food Retailing? Working Paper. Yale School of Management.
- Greene, W., *Econometric Analysis*, 2nd ed., Macmillan, New York, 1993.
- Gupta, S. and P.K. Chintagunta. 1994. On Using Demographic Variables to Determine Segment Membership in Logit Mixture Models. *Journal of Marketing Research* 31 (1): 128-136.
- Hoch, S.J., Kim, B., Montgomery, A. L. and P.E. Rossi. 1995. Determinants of Store-Level Price Elasticity. *Journal of Marketing Research* 32 (1): 17-29.
- Jensen, M.C. and W.H. Meckling. 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics* 3: 305-360.
- Kalyanam, K. and D.S. Putler. 1997. Incorporating Demographic Variables in Brand Choice Models: An Indivisible Alternatives Framework. *Marketing Science* 16 (2): 166-181
- Lafontaine, F. 1992. Agency Theory and Franchising: Some Empirical Results. *The Rand Journal of Economics* 23 (2): 263-283.
- Landier, A., Nair, V.B. and J. Wulf. 2006. Tradeoffs in Staying Close: Corporate Decision Making and Geographic Dispersion. Working Paper. Stern School of Business. New York University.

- Miller, N.J. and R.C. Kean. 1997. Factors Contributing to Inshopping Behavior in Rural Trade Areas: Implications for Local Retailers. *Journal of Small Business Management* 35: 80-94.
- Maddala, G., "Limited Dependent Variable Models Using Panel Data," *Journal of Human Resources*, 22, 3, 1987, pp. 307-338.
- Mattila, A.S. 1999. The role of culture in the service evaluation process. *Journal of Service Research* 1 (3): 250-261.
- Mittal, V., Kamakura, W. and R. Govind. 2004. Geographic Patterns in Customer Service and Satisfaction: An Empirical Investigation. *Journal of Marketing* 68 (July): 48-62.
- Moers, F. 2006. Performance Measure Properties and Delegation. *The Accounting Review* 81 (4): 897-924.
- Mulhem, F.J. and J.D. Williams. 1994. A Comparative Analysis of Shopping Behavior in Hispanic and Non-Hispanic Market Areas. *Journal of Retailing* 70 (Fall): 231-252.
- Mulhern, F. J., Williams, J.D. and R.P. Leone. 1998. Variability of Brand Price Elasticities across Retail Stores: Ethnic, Income and Brand Determinants. *Journal of Retailing* 74 (3): 427-446.
- Norton, S.W. 1988. An Empirical Look at Franchising as an Organizational Form. *Journal of Business* 61(2): 197-218.
- Punj, G. and D.W. Stewart. 1983. Cluster Analysis in Marketing Research: Review and Suggestions for Application. *Journal of Marketing Research* 20 (2): 134-148.
- Sandino, T. 2007. Introducing the First Management Control Systems: Evidence from the Retail Sector. *The Accounting Review*. forthcoming.
- Sharma, S. 1996. Applied Multivariate Techniques. New York, NY: John Wiley & Sons.
- Scott F., 1995. Franchising vs. Company Ownership as a Decision-Variable of the Firm. *Review of Industrial Organization*. 10: 69-81
- Wilkes, R. E. and H. Valencia. 1985. A Note on Generic Purchase Generalizations and Subcultural Variations. *Journal of Marketing* 49 (Summer): 114-120.
- Williamson, O. 1985. *The Economic Institutions of Capitalism*. New York, New York: The Free Press.
- Wooldridge, J. (2002). *Econometric Analysis of Cross-Section and Panel Data*. Cambridge, MA: The MIT Press.

Table 1
Summary Statistics for the Chain-Level Dataset (N=420 Chains)

Variable	Mean	Std. Dev.	Minimum	Maximum
Market Dispersion Measures				
<i>MDISPERSION_NCLUST</i>	10.8	5.1	2	24
<i>MDISPERSION_SUM</i>	1.97	0.58	0.56	5.16
<i>MDISPERSION_MULT</i>	0.001	0.002	0.000002	0.01
<i>MDISPERSION_HHI</i>	78.6	12.0	27.15	94.20
Geographic Dispersion Measures				
<i>CHAIN_STATEHHI</i>	26.7	30.5	0	96.4
Chain Characteristics				
<i>CHAIN_NSTORES</i>	152	463.5	20	5758
<i>CHAIN_FRANCHISOR</i>	0.14	0.34	0	1

Notes:

MDISPERSION_SUM: Market-type dispersion measured as the aggregate sum of individual normalized dispersion measures based on location characteristics (population, income, age, ethnicity, household size).

MDISPERSION_MULT: Market-type dispersion measured as the aggregate product of individual normalized dispersion measures based on location characteristics (population, income, age, ethnicity, household size).

$$MDISPERSION_HHI: \left[1 - \sum_{x=1}^{25} \left(\frac{\text{number of stores in cluster } X \text{ in chain } i}{\text{total number of stores in chain } i} \right)^2 \right] \times 100$$

$$CHAIN_STATEHHI: \left[1 - \sum_{x=1}^{50} \left(\frac{\text{number of stores in state } X \text{ in chain } i}{\text{total number of stores in chain } i} \right)^2 \right] \times 100$$

CHAIN_NSTORES: Total number of stores operated by the chain

CHAIN_FRANCHISOR: Indicator for whether a chain franchises any of its stores

Table 2
Distribution of Franchise Activity across Chains

	Number of Chains	% Sample	Mean % of Stores Franchised	Minimum % of Stores Franchised	Maximum % of Stores Franchised
<i>No Franchised Stores</i>	363	86.4	0	0	0
<i>Some Franchised Stores*</i>	43	10.2	52.5	0.2	99.6
<i>All Franchised Stores</i>	14	3.3	100	100	100

***Some Franchised Stores: 0 < % of Chain's Stores Franchised < 1**

Table 3
Chain-Level Correlations between Measures of Geographic and Market-Type Dispersion

	<i>CHAIN_STATEHHI</i>	<i>MDISPERSION_SUM</i>	<i>MDISPERSION_MULT</i>
<i>CHAIN_STATEHHI</i>	1		
<i>MDISPERSION_SUM</i>	0.3161*	1	
<i>MDISPERSION_MULT</i>	0.2757*	0.4129*	1
<i>MDISPERSION_HHI</i>	0.3470*	0.2602*	0.3434*

* Significant (p<.10)

Table 4
Summary Statistics for the Store-Level Dataset (N=34,892)

Variable	Mean	Std. Dev.	Minimum	Maximum
Franchise Choice				
<i>FRANCHISE</i>	0.69	0.46	0.00	1.00
Market Divergence Measures				
<i>MDIVERGENCE_SUM</i>	3.6	2.0	0.3	25.4
<i>MDIVERGENCE_MULT</i>	0.87	7.4	3.81x10 ⁻⁸	548.6
<i>MDIVERGENCE_MAINCLUST</i>	48.6	34.5	0.0	307.9
<i>MCONCENTRATION_SAMECLUST</i>	7.6	4.9	0.07	41.7
Geographic Divergence Measures				
<i>DISTANCE_HQ</i>	658	624	0	4,701
<i>NSTORES_ZIPCHAIN</i>	2.0	1.8	1.0	23.0
Zip Code Demographic Characteristics				
<i>% White Population</i>	73.9	21.8	0.5	100
<i>Average Household Size</i>	2.6	0.4	1.3	5.4
<i>Median Income</i>	52,954	18,573	10,830	192,006
<i>Median Age</i>	36.4	5.4	19.1	74.8
<i>Population per Square Mile</i>	2,700	4,065	0.003	96,990

Notes:

FRANCHISE: Indicator for whether or not store is franchised

MDIVERGENCE_SUM: Market-type divergence measured as the sum of the absolute values of normalized differences on each location characteristic between the store and the average value of the location characteristic for the chain.

MDISPERSION_MULT: Market-type divergence measured as the product of the absolute values of normalized differences on each location characteristic between the store and the average value of the location characteristic for the chain.

MDISPERSION_MAINCLUST: Cluster distance between the store's cluster and the most frequent cluster in the chain, weighted by the proportion of the chain's stores in the main cluster

MCONCENTRATION_SAMECLUST: The proportion of same-chain stores operating in the same market-type cluster as the store

DISTANCE_HQ: Distance in miles from chain headquarters to store

NSTORES_ZIPCHAIN: Number of same-chain stores in the zip code (stores which are operated by the same chain with which the store is affiliated)

Table 5
Store-Level Correlations between Measures of Geographic and Market-Type Divergence

	<i>DISTANCE_HQ</i>	<i>NSTORES_ZIPCHAIN</i>	<i>MDIVERGENCE_SUM</i>	<i>MDIVERGENCE_MULT</i>	<i>MDIVERGENCE_MAINCLUST</i>
<i>DISTANCE_HQ</i>	1				
<i>NSTORES_ZIPCHAIN</i>	0.1297*	1			
<i>MDIVERGENCE_SUM</i>	0.0924*	-0.0574*	1		
<i>MDIVERGENCE_MULT</i>	0.0218*	-0.0241*	0.4441*	1	
<i>MDIVERGENCE_MAINCLUST</i>	0.0697*	-0.1055*	0.4328*	0.0978*	1
<i>MCONCENTRATION_SAMECLUST</i>	0.1481*	0.0011	0.5713*	0.2060*	0.4990*

* Significant ($p < .10$)

All variables are defined in Table 4.

Table 6
Ordinal Logit Regressions Linking the Likelihood of Franchising No, Some, or All
Stores to Measures of Market Dispersion, Geographic Dispersion, and Chain
Characteristics

	Predicted	<i>Ordinal Logits</i>			
		<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>	<i>Model 4</i>
<i>CHAIN_STATEHHI</i>	+	0.03	0.03	0.03	0.015
<i>p-value</i>		0.00	0.00	0.000	0.011
<i>CHAIN_NSTORES</i>	?	0.001	0.001	0.0003	0.0002
<i>p-value</i>		0.009	0.006	0.084	0.15
<i>MDISPERSION_SUM</i>	+		0.45		
<i>p-value</i>			0.07		
<i>MDISPERSION_MULT</i>	+			318.3	
<i>p-value</i>				0.007	
<i>MDISPERSION_HHI</i>	+				0.20
<i>p-value</i>					0.00
<i>R-square (N)</i>		0.18 (420)	0.19 (420)	0.22 (420)	0.27 (420)

Note: Dependent variable in all specifications = 1 if chain does not franchise any stores; 2 if chain franchises some, but not all stores; 3 if chain franchises all stores
Explanatory variables are described in Table 1.

Table 7
Logit Regressions Linking the Probability of Franchising an Individual Store to
Measures of Market Divergence and Geographic Concentration

	Predicted	<i>Logits</i>		
		<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>
<i>Constant</i>	+	0.88	0.87	0.61
<i>p-value</i>		<0.001	<0.001	<0.001
<i>DISTANCE_HQ</i>	+	0.0008	0.0009	0.0009
<i>p-value</i>		<0.001	<0.001	<0.001
<i>NSTORES_ZIPCHAIN</i>	-	-0.03	-0.025	-0.02
<i>p-value</i>		<0.05	<0.05	<0.05
<i>MDIVERGENCE_SUM</i>	+	0.05		
<i>p-value</i>		<0.001		
<i>MDIVERGENCE_MULT</i>	+		0.008	
<i>p-value</i>			<0.001	
<i>MDIVERGENCE_MAINCLUST</i>	+			0.002
<i>p-value</i>				<.05
<i>MCONCENTRATION_SAMECLUST</i>	-			0.02
<i>p-value</i>				<.001
<i>Chain Indicators</i>		Yes	Yes	Yes
<i>R-square (N)</i>		0.37 (34,892)	0.37 (34,892)	0.37 (34,892)

Notes: Dependent variable in all specifications = 1 store is franchised; 0 if store is company owned
Chain indicators are jointly significant (p<.001).
Explanatory variables are described in Table 4.