

Segmentação de mercado: proposta de uma metodologia de associação entre clientes e produtos no contexto do mercado imobiliário

Market segmentation: a proposed methodology for associating clients and products in the real estate market

Ana Augusta Ferreira de Freitas
Luiz Fernando Mählmann Heineck

Resumo

O presente trabalho tem como objetivo propor uma metodologia para análise de segmentação de mercado. A técnica utilizada para segmentar o mercado, *Chi-Square Automatic Interaction Detection (CHAID)*, e a abordagem proposta permitem que a associação direta entre produtos e perfis seja feita de modo a dar suporte às decisões relacionadas às estratégias de marketing de maneira inovadora, à medida que consideram a associação entre estes de maneira direta. Uma ilustração da proposta é desenvolvida com base em um banco de dados contendo cerca de 3000 entrevistas com potenciais clientes do mercado imobiliário. Neste caso, a associação entre imóveis e famílias, fruto da aplicação da técnica e da metodologia proposta, traz resultados que se mostram muito mais úteis do que os relatados pela literatura existente sobre o assunto.

Palavras-chave: Segmentação de mercado. Mercado imobiliário. CHAID.

Abstract

This article proposes a new methodology to analyze market segmentation. Chi-Square Automatic Interaction Detection (CHAID) technique and the proposed approach allow to directly associate consumers and dwelling products, providing support for decision-making related to marketing strategies. An empirical example is analyzed using a database, which contains data obtained in interviews with around 3000 prospective clients while searching for a new home. The statistical association between the two groups, resulting from the application of the proposed methodology, provides more useful results than those previously reported in the literature.

Keywords: Market segmentation. Construction market. CHAID.

Ana Augusta Ferreira de
Freitas
Centro de Estudos Sociais
Aplicados
Universidade Estadual do Ceará
Av. Paranjana, 1700
Fortaleza - CE - Brasil
CEP 60740-903
Tel.: (85) 3101-9940
E-mail: freitas8@terra.com.br

Luiz Fernando Mählmann
Heineck
Departamento de Engenharia
Mecânica e Produção
Universidade Federal do Ceará
Fortaleza - CE - Brasil
CEP 60455-760
Tel.: (85) 3366-9600
E-mail:
heineck@pesquisador.cnpq.br

Recebido em 26/02/08
Aceito em 28/10/08

Introdução

Partindo do artigo seminal de segmentação de mercado produzido por Smith (1956), é possível documentar que empresas dos diversos setores da economia têm usado os conceitos de segmentação na definição das suas estratégias de marketing. A estratégia de segmentação de mercado tem como pressuposto a existência de um mercado heterogêneo, caracterizado pela existência de demandas diferenciadas, que compõem distintos segmentos de mercado. Sua adoção implica a identificação de classes de consumidores que se constituem em possíveis segmentos-alvo da empresa e que norteiam os esforços de marketing no sentido de satisfazer a demanda dos segmentos escolhidos.

A heterogeneidade do mercado atendido por empresas de diversos setores justifica a adoção de estratégias com base na segmentação. Tal heterogeneidade é consequência da forma diferenciada como os consumidores avaliam os diversos atributos, conforme discute Dibb (1994) para o mercado imobiliário. Conhecer o processo de decisão habitacional requer a geração e a análise de dados para identificação das diferentes características dos segmentos de mercado. González e Formoso (2006) lembram que, nesse contexto, as decisões de lançamentos imobiliários são frequentemente realizadas de forma subjetiva e sem base em dados reais.

Na verdade, essa não é a única peculiaridade desse mercado. Em relação à segmentação, Smith, Garbarino e Martini (1992) identificam que os analistas estratificam o mercado por tipologia habitacional, e não por benefícios ao consumidor. A propriedade imobiliária é então avaliada em termos dos atributos físicos nela incorporados, em vez da percepção do consumidor acerca do espaço, da atmosfera e das ligações sociais, o que seria mais adequado, já que parte das decisões tomadas são baseadas em aspectos não financeiros. Esses pressupostos têm profundo impacto na forma como as estratégias de marketing serão tratadas, com o produto sobrepondo-se às decisões relacionadas a serviços oferecidos e aos relacionamentos estabelecidos entre empresa e consumidor.

Sob a ótica da academia, a maioria dos estudos sobre mercado imobiliário é baseada na teoria econômica neoclássica que assume que as decisões imobiliárias são tomadas de forma a maximizar a utilidade do decisor. As influências humanas, segundo Gibling e Nelson (2003), são desconsideradas, e os pesquisadores continuam a abordar o mercado imobiliário a partir de uma

orientação para a produção, e não para o mercado. Ademais, de maneira geral, apesar da atenção acadêmica que a teoria sobre segmentação tem atraído e dos recursos nela investidos, as evidências sugerem que as empresas têm problemas com a operacionalização desses conceitos. Essas dificuldades originam-se de diversas fontes, que vão desde a escolha da variável mais apropriada para a distinção dos segmentos até a técnica de definição dos grupos a ser utilizada. Segundo Dibb (1998), a ausência de estudos que facilitem a implementação dos conceitos de segmentação resulta do fato de que o processo utilizado nem sempre alcança segmentos homogêneos de clientes, para os quais programas de marketing possam ser efetivamente desenvolvidos. Em outras palavras, respeitando-se o próprio conceito de segmentação, o processo de segmentar pode ser ineficiente e errôneo. Nesse sentido, o presente trabalho tem como objetivo a proposição de uma metodologia para análise de segmentação de mercado.

A justificativa para o desenvolvimento do trabalho ora apresentado reside na necessidade de exploração dos conceitos teóricos e, essencialmente, operacionais de segmentação, no sentido de preencher uma lacuna existente na literatura de marketing e de economia da construção, que traga uma discussão ampla sobre as técnicas disponíveis e suas respectivas características, vantagens e desvantagens. O trabalho alia-se a outras discussões que vêm sendo feitas no âmbito da Associação Nacional de Tecnologia do Ambiente Construído (ANTAC), na sua divisão de Economia da Construção, onde estão inseridos os trabalhos na área de marketing na construção. Entre eles, cita-se o trabalho de Miron (2002), que se destaca ao apresentar um sistema de gerenciamento de requisitos de clientes em empreendimentos imobiliários. A origem desses trabalhos se dá a partir do reconhecimento das especificidades do processo de compra de um imóvel, considerado como um bem de alto envolvimento, em consequência dos riscos percebidos pelo consumidor. Ademais, Bernardes e Carvalho (2001) afirmam que o comportamento do consumidor de bens duráveis é marcado por uma inércia comportamental, que faz com que o consumidor só reaja diante de uma discrepância real entre a situação atual e a desejada.

Noções teóricas sobre segmentação

Segundo Green e Krieger (1991), juntamente com o posicionamento de produtos, a segmentação de mercados é um dos mais pesquisados conceitos de marketing, tendo evoluído significativamente desde a pesquisa inicial de Smith (1956), que a considerava essencialmente uma estratégia de *merchandising*. No trabalho de Wind (1978), o autor mostra essa evolução através do tempo. A partir dessa revisão, outros autores fizeram importantes contribuições para o amadurecimento da literatura na área (DICKSON, 1982; FINE, 1980; RUDELIUS *et al.*, 1987).

Existem duas questões importantes para que a segmentação se transforme numa arma de competição por parte das empresas que a utilizam. A primeira é a correta definição da variável que deve ser utilizada como base para a formação dos grupos, juntamente com a técnica estatística mais recomendada; a segunda está relacionada à forma como esses resultados podem ser transformados em definição das estratégias de marketing. Em relação à base de segmentação, as variáveis mais freqüentemente utilizadas são classificadas como geográficas, demográficas, psicográficas, aspectos relacionados com o produto e variáveis comportamentais (KOTLER, 1998). Woo (1998) defende a idéia de que as variáveis demográficas, embora sejam as mais freqüentemente utilizadas para segmentação, não são as mais indicadas. Isso porque os grupos formados a partir de variáveis demográficas são assumidos como reflexo de padrões homogêneos de consumo, o que não é necessariamente verdade. Além disso, a partir dos resultados gerados não é possível se tomar nenhuma decisão direta sobre a implementação das estratégias de marketing, já que não se faz uma associação clara entre preferências e consumidores. Corroborando com o pensamento de outros pesquisadores, como Haley (1978) e Dibb (1998), acredita-se que o melhor sistema de segmentação é aquele que mais claramente embasa as decisões a serem tomadas pelo departamento de marketing. A vinculação entre segmentação e tomada de decisão gerencial é até hoje a preocupação central dos teóricos e práticos que discutem o assunto (WÜBBEN; WANGENHEIM, 2008).

Após definir a base para a divisão dos grupos, o processo de segmentação, considerado de uma maneira simplista, tem uma definição clara e bem estabelecida. Isso porque, na maioria dos mercados, é possível imaginar o desenrolar das três etapas, quais sejam: segmentação, escolha do segmento-alvo e posicionamento. Algumas

técnicas conseguem operacionalizar as três etapas propostas de maneira integrada, como é o caso da aplicação da análise conjunta sugerida por Green e Krieger (1991). No entanto, a maioria das técnicas disponíveis considera apenas a primeira etapa do processo. Segundo Wind (1978), essa primeira etapa pode ser operacionalizada por meio de duas abordagens. São elas:

- (a) segmentação *a priori*, em que o pesquisador escolhe uma ou mais variáveis de interesse (e.g. idade, benefícios procurados) e então classifica os compradores de acordo com essa(s) variável(is). Nas segmentações *a priori*, o número de segmentos, seu tamanho e sua descrição são geralmente conhecidos no início do processo; e
- (b) segmentação *post hoc* ou segmentação baseada em *cluster*, em que o pesquisador escolhe uma série de variáveis inter-relacionadas (e.g. características psicográficas, preferências por benefícios do produto), e as pessoas são agrupadas de acordo com a similaridade entre elas. Nas segmentações *post hoc*, o número de segmentos, seu tamanho e sua descrição são evidenciados depois do término da operação. Esta última abordagem tem recebido maior atenção nas últimas décadas.

A análise de *cluster*, aplicada para fins de exemplificação no trabalho de Dibb (1998), é uma das técnicas multivariadas mais amplamente utilizadas pelos autores da área de marketing e de segmentação (ALDENDERFER; BLASHFIELD, 1984), inclusive em trabalhos na área de mercado imobiliário (e.g. LEITE; FORMOSO, 2004). O princípio da técnica é a divisão baseada em uma medida de similaridade ou distância para criação de grupos homogêneos ou *clusters*. Para uma revisão detalhada sobre a técnica sugere-se Everitt (1980). Os problemas relacionados à utilização da análise de *cluster* são, similarmente, bem documentados na literatura (e.g. ALDENDERFER; BLASHFIELD, 1984; CHATURVEDI *et al.*, 1997; EVERITT, 1980). Sem dúvida o mais importante deles é assegurar que os grupos formados são genuínos, já que, independentemente da situação ou dos dados analisados, a técnica termina gerando grupos, que podem não ter significado mercadológico. Além disso, os grupos formados pela técnica nem sempre são facilmente identificados, bem como não é direta a indicação do grupo ao qual um novo elemento agregado à amostra deve pertencer.

Em muitos casos, procede-se a uma análise fatorial antes da aplicação da técnica de *cluster*, a fim de reduzir as dimensões das variáveis originais (e.g. BROMERCHENKEL, 1996; WOO, 1998). Esse procedimento é criticado por alguns autores (e.g.

ALDENDERFER; BLASHFIELD, 1984), que defendem que informações relevantes podem ser descartadas nessa fase inicial e que a estrutura de *cluster* gerada pode não ser válida. Do mesmo modo, a análise discriminante é também muitas vezes utilizada de maneira complementar à técnica de *cluster*. Nesse caso, o objetivo é identificar quais as dimensões que discriminam de forma mais efetiva os *clusters* gerados e procuram estereotipar os grupos formados.

Em todos os casos em que as técnicas de análise multivariada são utilizadas (e.g. análise discriminante, análise fatorial, correlação canônica, análise de componentes principais), o objetivo é a redução do número de variáveis que representam a situação desejada. Uma comparação entre os diversos métodos pode ser vista em Richardson e Thalheimer (1982).

Alguns estudos também foram conduzidos a fim de testar a eficiência da análise conjunta (tradução do termo inglês *conjoint analysis*) na segmentação de mercados. Green e Krieger (1991) propuseram a utilização dessa técnica para implementação da segmentação, teste do mercado-alvo selecionado e conseqüente posicionamento do produto, considerando assim todas as fases do processo, de maneira conjunta. Mais recentemente, Sattler e Hartmann (2008) reuniram em uma publicação as principais contribuições dessa técnica, considerando seu uso comercial nas práticas de análises de preferências e de segmentação.

Finalmente, outro grupo de técnicas estatísticas aparece com o objetivo de agrupar conjuntos de dados, utilizando também a simplificação do número de variáveis discriminantes que melhor explicam os segmentos gerados. São elas a Redução da Proporção na Incerteza (PRU - *Proportional Reduction in Uncertainty*) e a Detecção de Interação Automática do Qui-Quadrado (CHAID - *Chi-Square Automatic Interaction Detection*). No primeiro caso, uma série de estatísticas de qui-quadrado é conduzida para avaliar quais são os melhores regressores de uma variável quantitativa dependente, dentro de uma perspectiva de predição de resultados. No segundo caso, procede-se a análises de variância a fim de explicar a heterogeneidade de uma variável dependente, buscando a classificação de resultados. Na área de marketing, as duas técnicas podem ser utilizadas na geração de grupos de consumidores com preferências homogêneas (FREITAS, 2000a; HAWKINS; KASS, 1982).

No caso específico do CHAID, este se caracteriza como um algoritmo que aplica árvore de classificação para dividir um conjunto de dados em subconjuntos que discriminam de maneira

diferenciada a variável resposta (dependente) e para combinar categorias que não diferem significativamente entre si. O processo usa variáveis discretas. O método de interação foi originalmente desenvolvido por Kass (1980) e posteriormente aplicado em pesquisas em setores diversos, como na área de saúde (FRANK *et al.*, 2008) e turismo (HSU; KANG, 2007). A suposição para uso do modelo é a de que existe uma variável dependente que pode ser categorizada e associada a variáveis independentes que expliquem sua variabilidade. Essa associação é preliminarmente indicada a partir da revisão da literatura. No modelo é possível testar ainda a associação com variáveis não indicadas na literatura, variáveis estas que serão inclusas no modelo a partir de uma avaliação de significância estatística.

Para identificar as variáveis independentes com maior poder de explanação são aplicados testes de qui-quadrado. Uma visão do funcionamento *bottom-up* da sistemática do CHAID pode ser apreciada a seguir. O método testa a associação de duas a duas variáveis independentes e a variável dependente, verificando a homogeneidade do grupo assim formado. O par de variáveis independentes que mostrou homogeneidade é então entendido como mais uma variável independente, de um nível hierárquico maior, que passa a ser testado de modo combinatório com todas as outras variáveis independentes isoladas ou em pares já formados. Esse processo continua de forma recursiva, cada vez agregando um número maior de variáveis independentes, seus pares, trincas ou qualquer outro agrupamento, sempre que isso resultar em um novo grupo homogêneo, em níveis hierárquicos cada vez maiores da segmentação. No caso, um nível hierárquico maior da segmentação corresponde a grupos mais amplos de clientes, ou seja, um mercado menos segmentado.

Entendida a lógica estatística do procedimento, é possível identificar os seguintes passos metodológicos para o uso da técnica: seleção da variável dependente categórica (passo 1); criação de categorias para essa variável (passo 2); definição das variáveis independentes (passo 3); e estabelecimento das opções de segmentação (passo 4), em que se definem as regras de finalização da divisão e o tamanho mínimo dos grupos formados. É também nesta etapa que se estabelece o nível de significância estatística que irá impactar na junção dos níveis das variáveis independentes e na sua elegibilidade, para serem incluídas no processo. A cada passo podem ser visualizados testes estatísticos de associação entre as variáveis predictoras e a variável dependente (passo 5). Duas

opções de qui-quadrado estão disponíveis: Likelihood Ratio qui-quadrado, para variáveis dependentes nominais, e Pearson qui-quadrado, para variáveis dependentes ordinais.

Na prática, os resultados apresentados são visualizados na forma de um diagrama em árvore (chamado de dendograma), com divisões progressivas de ramos cada vez mais específicos. Por exemplo, no caso de um estudo de segmentação, a medida dependente pode ser o benefício procurado por um grupo de clientes; e as medidas independentes, as características socioeconômicas desses. Nesse caso, a mais importante das características socioeconômicas será a primeira divisora da população. A seguir cada novo grupo gerado irá ser subdivido, utilizando uma nova variável independente, que melhor explique o benefício procurado.

A técnica é limitada a variáveis categorizadas (nominais, ordinais ou quantitativas), mas, diferentemente de outros processos de segmentação, utiliza toda a distribuição da variável dependente, não é restringida a divisões binárias e não faz suposição de normalidade para as variáveis. Além dessas vantagens, podem ser elencadas outras:

- (a) a possibilidade de utilização de variáveis dependentes e independentes em diversas formas (nominais, ordinais e categóricas);
- (b) a possibilidade de os valores faltantes serem tratados pelos softwares disponíveis, sendo agregados ao grupo mais próximo em termos de homogeneidade da distribuição da variável dependente ou isolados em uma categoria distinta;
- (c) a não-necessidade de estabelecer-se *a priori* uma ou mais variáveis como base para segmentação; e
- (d) a possibilidade de utilização de várias bases para segmentação, em qualquer ordem, tendo

como critério tão-somente a homogeneização dos segmentos criados.

A metodologia proposta neste trabalho utiliza a técnica de CHAID para segmentar os grupos, dadas as vantagens discutidas anteriormente. A metodologia ora proposta aborda a segmentação vista por dois ângulos: do produto e das pessoas. No primeiro caso, os produtos serão definidos conforme uma gama de variáveis importantes para explicação da base de segmentação. A associação entre grupos de produtos (segmentos de produtos) e grupos de pessoas (segmentos de mercado) será então realizada por meio de uma tabulação cruzada. O esquema dessa proposta é ilustrado na Figura 1.

Segundo a proposta apresentada neste trabalho, os dois grupos são associados após a escolha de uma base de segmentação comum. Essa base deve preferencialmente estar relacionada com o comportamento de compra futuro, ou seja, trabalhar com preferências manifestadas pelos clientes que redundem posteriormente em escolhas reais no mercado.

O banco de dados que dará suporte à análise contém cerca de 3.000 entrevistas com consumidores que procuravam um imóvel para comprar nos anos de 1995 a 1998 em diversas cidades do país. A forma de captação desses dados pode ser vista com profundidade em Freitas (2000b). No entanto, é importante registrar algumas informações. As pesquisas foram feitas em feiras de imóveis. Apenas as pessoas que declaravam estar interessadas em comprar um imóvel (para uso próprio ou para investimento) foram entrevistadas. As entrevistas eram feitas preferencialmente com a família, salvo as ocasiões em que apenas um membro da família se colocava à disposição para participar da entrevista.

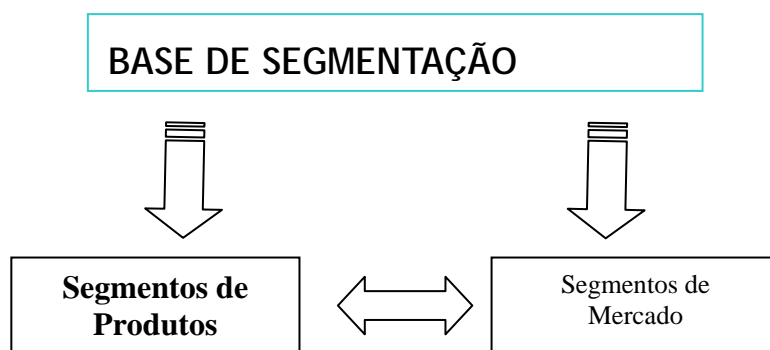


Figura 1 - Proposta de segmentação com a associação de produtos e pessoas

Grupos de pesquisas distintos em todo o Brasil gerenciaram a aplicação e condução das entrevistas. Os questionários aplicados tinham estruturas semelhantes e eram compostos de duas partes: características socioeconômicas do entrevistado e preferências habitacionais. As pesquisas foram realizadas nas cidades de Belém (PA), Blumenau (SC), Caxias do Sul (RS), Florianópolis (SC), Natal (RN), Passo Fundo (RS), Pelotas (RS), Porto Alegre (RS), Recife (PE), Santa Maria (RS) e Vitória (ES). Não houve nenhum critério sistemático para a escolha dessas cidades.

Campo empírico: uma aplicação no mercado de bens duráveis

De acordo com a metodologia proposta, a primeira etapa para sua implementação é a escolha da base de segmentação que será utilizada. Pela revisão da literatura apropriada sobre o assunto (e.g. ADAIR; MCGREAL, 1994; SCHOLTEN; HOOIMEIJER, 1984), a base escolhida foi o preço esperado a ser pago pelo futuro imóvel e suas cinco categorias ou níveis, conforme será visto a seguir (passo 1 e passo 2). Numa primeira etapa, já documentada pela literatura (FREITAS, 2000a), procedeu-se à segmentação do mercado com base no produto (segmentação do produto), utilizando a técnica de CHAID. Na ocasião, os resultados mostraram a existência de 12 tipologias de imóveis. Essas tipologias foram formadas pelas variáveis que têm o maior poder de explicação acerca da variabilidade do preço dos imóveis que estavam sendo procurados (Tabela 1).

Em seguida, procedeu-se à definição dos segmentos de mercado com base nas características dos clientes (segmentação de mercado). Para tal, definiu-se uma lista de variáveis independentes (variáveis preditoras) passíveis de explicarem a criação dos distintos grupos de consumidores (Tabela 2). Aquelas com maior poder de explicação da variabilidade dos preços dos imóveis procurados foram escolhidas pelo CHAID e utilizadas na criação dos grupos (Figura 2). Essa etapa é o passo 3 da técnica.

Foi estipulado o tamanho mínimo de 50 elementos para cada subgrupo (passo 4). As categorias que não apresentaram diferenças estatísticas na explicação da variabilidade do preço dos imóveis foram agrupadas em um único segmento. Para que a variável fosse elegível para compor a árvore, ela

deveria ter um *p-value* menor ou igual a 0,05. Esse valor significa a probabilidade de atestar que há uma relação entre a variável predita e a variável dependente na população quando, na verdade, elas não estão relacionadas na população.

A variável dependente preço é categórica e possui os seguintes níveis: até R\$ 42.500 (1); de R\$ 42.500 a R\$ 55.000 (2); de R\$ 55.000 a R\$ 75.000 (3); de R\$ 75.000 a R\$ 120.000 (4); e mais de R\$ 120.000 (5). Os números que aparecem no lado direito do nome das variáveis correspondem aos seus códigos, que estão representados na Tabela 2. Em alguns casos, existe um ponto ao lado desses números, o que significa que os valores faltantes foram agregados àquele grupo, pois a variabilidade dos preços associada a essa variável faltante não era estatisticamente diferente daquela encontrada no grupo.

Conforme descrito nas etapas do método, alguns níveis foram agregados por não apresentarem diferenças estatísticas para $p=0,05$ (passo 5). É o caso, por exemplo, dos níveis 2 e 3 da variável preditora valor patrimonial (bens) e dos níveis 1, 2 e 3 da variável filhos.

Conforme explicado na seção teórica, a técnica do CHAID também oferece estatísticas que ajudam a julgar a adequação das divisões feitas. O Pearson qui-quadrado das variáveis preditoras elegíveis no modelo apresentou valores entre 7,8 e 15,06, todos significativos (*p-value* menor que 0,05).

Os retângulos em negrito no final das ramificações representam os segmentos de mercado. A Tabela 3 mostra suas características, que foram definidos pelo CHAID.

A partir da análise do dendograma, pode-se tecer algumas considerações sobre os resultados obtidos. Em linhas gerais, nota-se que poucas variáveis independentes foram escolhidas para explicar o preço procurado e que a principal é a renda mensal familiar. Além disso, os segmentos de mercado são definidos utilizando-se distintas variáveis, dando livre prioridade de escolha àquelas que melhor explicam a variação do preço dentro de cada um deles. A vantagem conceitual da técnica foi aqui utilizada, não se atendo a classificações rígidas ou padronizadas.

Pode-se notar que as famílias com renda superior foram subdividas, preferencialmente, pelo tamanho da família ou idade do chefe da família, e as de menor renda, pelo valor patrimonial. Além disso, a idade do filho mais velho é o fator distintivo nas famílias com menor renda.

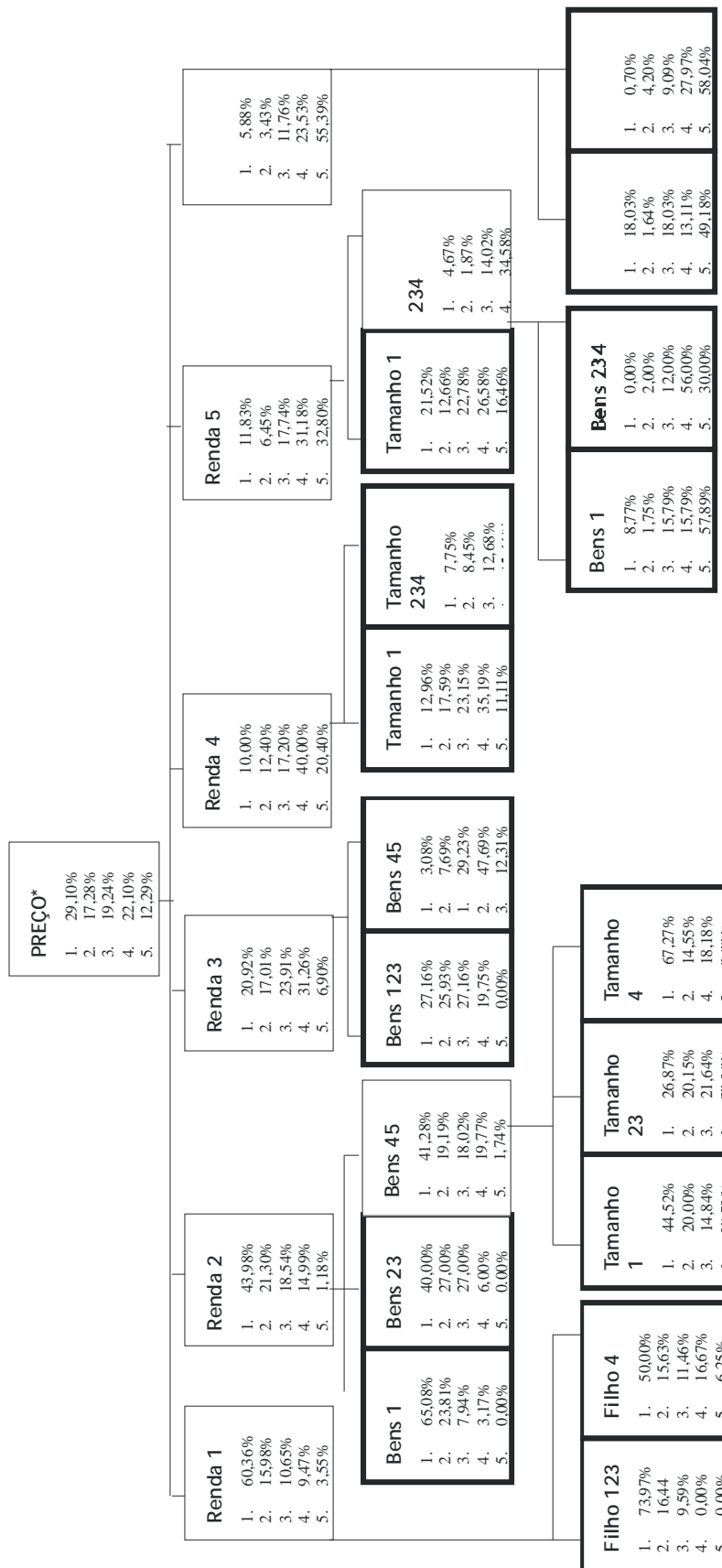
Segmentos de produtos	Características
Grupo 1	Um quarto
Grupo 2	Dois quartos, baixa qualidade do bairro
Grupo 3	Dois quartos, média qualidade do bairro
Grupo 4	Dois quartos, alta qualidade do bairro
Grupo 5	Três quartos, uma garagem, baixa qualidade do bairro
Grupo 6	Três quartos, uma garagem, média qualidade do bairro
Grupo 7	Três quartos, uma garagem, alta qualidade do bairro
Grupo 8	Três quartos, duas ou três garagens, sem suíte ou com uma suíte
Grupo 9	Três quartos, duas ou três garagens, duas ou mais suítes
Grupo 10	Quatro quartos, baixa ou média qualidade do bairro
Grupo 11	Quatro quartos, alta qualidade do bairro, sem suíte ou com uma suíte
Grupo 12	Quatro quartos, alta qualidade do bairro, duas ou mais suítes

Fonte: Freitas (2000a)

Tabela 1 - Caracterização dos segmentos de produtos

Variáveis Independentes	Categorias
Estado civil	Casado 2. Solteiro 3. Outros
Idade do chefe da família	1. até 25 anos 2. de 26 a 35 anos 3. de 36 a 45 anos 4. mais de 45 anos
Idade do filho mais velho (filho)	nenhum 2. criança 3. jovem 4. adulto
Estágio do ciclo de vida (em apenas 3 cidades)	solteiro casado sem filhos com menos que 45 anos casado com filho criança casado com filho mais velho na adolescência casado com filho mais velho em idade adulta casado sem filho com idade acima de 45 anos divorciado ou viúvo
Tempo de residência	1. menos de 1 ano 2. de 1 a 3 anos 3. de 4 a 6 anos 4. de 7 a 9 anos 5. mais de 10 anos
Tipo de habitação atual	casa 2. apartamento
Tamanho da família	de 1 a 2 pessoas 2. de 3 a 4 pessoas 3. de 5 a 6 pessoas 4. 7 ou mais pessoas
Fonte de renda	empresa privada 2. funcionário público 3. profissional autônomo 4. rendas variadas 5. outros
Condição de propriedade atual	própria 2. alugada 3. cedida
Renda familiar	até R\$ 1.000 de R\$ 1.000 a R\$ 2.000 de R\$ 2.000 a R\$ 3.000 de R\$ 3.000 a R\$ 4.000 de R\$ 4.000 a R\$ 5.000 mais de R\$ 5.000
Valor patrimonial (bens)	até R\$ 13.000 de R\$ 13.000 a R\$ 27.000 de R\$ 27.000 a R\$ 41.000 de R\$ 41.000 a R\$ 78.000 mais de R\$ 78.000

Tabela 2 - Variáveis independentes para categorização dos segmentos de produtos



*Os valores nos quadros são os percentuais de imóveis nos níveis de preço
 Figura 2 - CHAID dendrograma para definição dos segmentos de mercado

Segmentos de Mercado	Características
Família 1	Renda até R\$ 1.000, sem filhos ou com o filho mais velho criança ou adolescente
Família 2	Renda até R\$ 1.000, com o filho mais velho adulto
Família 3	Renda entre R\$ 1.000 e R\$ 2.000 e patrimônio até R\$ 13.000
Família 4	Renda entre R\$ 1.000 e R\$ 2.000 e patrimônio entre R\$ 13.000 e R\$ 41.000
Família 5	Renda entre R\$ 1.000 e R\$ 2.000, patrimônio maior que R\$ 41.000, com famílias compostas de até 2 pessoas
Família 6	Renda entre R\$ 1.000 e R\$ 2.000, patrimônio maior que R\$ 41.000, com famílias compostas de 3 a 7 pessoas
Família 7	Renda entre R\$ 1.000 e R\$ 2.000, patrimônio maior que R\$ 41.000, com famílias compostas de mais de 7 pessoas
Família 8	Renda entre R\$ 2.000 e R\$ 3.000 e patrimônio até R\$4 1.000
Família 9	Renda entre R\$ 2.000 e R\$ 3.000 e patrimônio maior que R\$ 41.000
Família 10	Renda entre R\$ 3.000 e R\$ 4.000 com famílias compostas de até 2 pessoas
Família 11	Renda entre R\$ 3.000 e R\$ 4.000 com famílias compostas de mais de 2 pessoas
Família 12	Renda entre R\$ 4.000 e R\$ 5.000 com famílias compostas de até 2 pessoas
Família 13	Renda entre R\$ 4.000 e R\$ 5.000 com famílias compostas de mais de 2 pessoas e patrimônio até R\$ 13.000
Família 14	Renda entre R\$ 4.000 e R\$ 5.000 com famílias compostas de mais de 2 pessoas e patrimônio maior que R\$ 13.000
Família 15	Renda superior a R\$ 5.000 cujo chefe da família tem até 35 anos
Família 16	Renda superior a R\$ 5.000 cujo chefe da família tem mais que 35 anos

Tabela 3 - Caracterização dos segmentos de mercado definidos pelo CHAID

Os segmentos de mercado delineados pelo CHAID são uma visão sintética dos agrupamentos de famílias, à medida que reduz as suas possibilidades combinatoriais. Considerando, por exemplo, que as famílias poderiam variar em seu tamanho (de 1 até mais de 7 pessoas), renda (desde até R\$ 1.000 até mais de R\$ 5.000), ter-se-ia um grande número de potenciais compradores, caso se quisesse varrer todas as possibilidades de combinação de renda e tamanho da família, por exemplo. Um número excessivo de segmentos seria formado mesmo explorando-se somente os casos encontrados no banco de dados de combinações dessas duas variáveis. Em vez disso, a técnica só cria um novo grupo se houver razões estatísticas para tal. De outra forma, mantém em um grupo só compradores que poderiam ser segmentados ainda mais, mas sem efeitos práticos quanto as suas escolhas relativas à variável dependente (preço do imóvel).

Ao examinar a distribuição do preço que se espera pagar pelo próximo imóvel, pode-se notar que existe uma tendência crescente desde o segmento de mercado 1 (famílias com renda até R\$ 1.000, sem filhos ou com o filho mais velho criança ou adolescente) até o segmento 16 (famílias com renda entre R\$ 4.000 e R\$ 5.000 cujo chefe da família tem mais de 35 anos).

A observação anterior é apenas uma indicação das possibilidades de seleção do mercado-alvo que uma empresa pode querer atingir, baseada apenas no critério do preço que os clientes estão dispostos a pagar. No entanto, se o objetivo é a definição das características do produto (uma das variáveis que compõem o composto mercadológico quando da definição das estratégias de marketing), é preciso que o incorporador tenha informações a respeito de quais são as características residenciais importantes, que explicam a formação daquele preço desejado.

A redução no número de combinações das características pessoais para apenas 16, ou seja 16 segmentos de mercado, torna exequível a ligação destes com as características dos imóveis que esses segmentos procuram (ou seja, os segmentos de produtos procurados, em número de 12). Existem 192 (16*12) distribuições estatisticamente diferentes para a variável preço do imóvel procurado, cabendo analisar, posteriormente, em termos práticos, a conveniência de entender mercadologicamente essas distribuições. Restringindo-se os níveis de significância do teste do qui-quadrado e aumentando-se o número mínimo de casos em cada agrupamento, pode-se chegar a um menor número de combinações de segmentos de mercado e de produto, em busca da

praticidade de aplicação do modelo. É interessante observar que ainda é possível construir modelos como os de regressão logística para a explicação da variação do preço dentro de cada par segmento de mercado * segmento de produto (ver FREITAS, 2000a), o que amplia a discussão sobre a necessidade de posteriormente chegar-se a um número manejável de agrupamentos de mercado.

Apresenta-se abaixo um exemplo de associação final entre segmentos de produtos e segmentos de mercado. A Tabela 4, extraída do cruzamento completo de 16 grupos familiares e de 12 grupos de imóveis, traz um resumo dessa relação, para 7 segmentos de mercado e 5 segmentos de produtos, de onde é possível destacar ilustrativamente algumas observações.

A Tabela 4 apresenta menos da metade dos tipos de famílias e tipos de imóveis. Foram escolhidos porque apresentam algum destaque na porcentagem apresentada, para explicar as preferências por tipos de imóveis. Por exemplo, a família 1, apesar de representar apenas 4,9% do total de famílias, é o 3º grupo mais representativo para os imóveis do tipo 1 e o 4º para o grupo 4. A família 2 é a segunda mais representativa no grupo 1 e também a segunda no grupo 4, justificando assim sua inclusão na tabela acima. Por outro lado, os grupos de imóveis ali presentes são exemplos de porcentagens pequenas de demanda (2,5%, 4,2% e 4,1% para os grupos 1, 10 e 12 respectivamente), e porcentagens médias (como os grupos 4 e 7, com respectivamente 12,5% e 10,6%). Os grupos de imóveis escolhidos, em número de 5 em um total de 12, representam cerca de um terço dos imóveis demandados. Fora dessas considerações nada de especial existe nesses grupos.

Para cada grupo de imóveis, as sete famílias apresentadas explicam a preferência de cerca de

70% dos imóveis existentes em cada grupo, o que ilustra o poder do método apresentado para conduzir a segmentação do mercado imobiliário. Um grupo de sete famílias em 16 (menos da metade das famílias) é responsável por mais de dois terços das preferências dentro dos grupos de apartamentos aqui escolhidos. Os esforços de marketing e de desenvolvimento do produto podem agora ser direcionados de maneira mais intensa para esses segmentos reduzidos de agrupamentos familiares.

Apesar dessa contribuição, o fator que distingue as vantagens do método proposto em relação a resultados anteriores é de outra natureza, o que é mais bem explicado com um exemplo. Primeiro, pode-se agora proceder a análises delineando-se segmentos de mercado, e não mais variáveis isoladas. No trabalho anterior de Freitas (2000b), a autora observava que os imóveis de um quarto (grupo 1) são demandados essencialmente por famílias jovens (solteiros ou casados sem filhos). Ora, os solteiros ou casados sem filhos poderiam estar presentes em praticamente todas as 16 famílias, com exceção das 2, 6, 7, 11, 13 e 14, que fazem menção explícita a agrupamentos familiares maiores. A primeira etapa da classificação das famílias em grupos levou em consideração principalmente renda e patrimônio, sendo claro que os solteiros ou casados sem filhos podem ocupar todas as faixas dessas variáveis econômicas. Através da análise agora apresentada, pode-se mostrar com maior precisão (Tabela 4) que os dois segmentos que mais procuram esse tipo de apartamento são as famílias com renda de até R\$ 1.000 com o filho mais velho adulto (19%), representadas pela categoria família 2, e as famílias com renda entre R\$ 1.000 e R\$ 2.000, patrimônio maior que R\$ 41.000, compostas de até 2 pessoas (33%), representadas pela categoria família 5.

	Grupo 1	Grupo 4	Grupo 7	Grupo 10	Grupo 12
Família 1	7,1	6,6	0,6	2,8	0,0
Família 2	19,0	15,6	7,3	4,2	0,0
Família 5	33,3	22,3	11,2	2,8	0,0
Família 7	2,4	1,4	3,9	4,2	0,0
Família 11	2,4	4,3	14,5	28,2	14,3
Família 16	2,4	1,9	4,5	15,5	41,4
Outro grupo familiar de destaque	Família 6 9,5	Família 10 11,4	Família 6 18,4	Família 9 9,9	Família 15 17,1
Outras famílias	23,9	36,5	39,6	32,4	27,2
Total	100,0	100,0	100,0	100,0	100,0

Tabela 4 - Associação entre segmentos de produtos e segmentos de mercado (%)

Segundo, esta análise não é somente inovadora por considerar as variáveis de maneira conjunta, mas também por revelar quais são aquelas mais importantes na explicação do preço desejado. Corroborando o parágrafo anterior, o fato de pertencer a famílias aparentemente jovens (solteiros ou casados sem filhos) não é a única nem a melhor explicação para apontar a maior preferência por imóveis do tipo 1, sendo necessário incluir as variáveis de natureza econômica, renda e patrimônio familiar. Neste caso, tanto a família do tipo 2 como a do tipo 5 têm rendas relativamente baixas, o que explica a preferência pelo imóvel aparentemente de menor porte de todos cogitados pelos potenciais clientes (1 quarto). No entanto, essas famílias têm uma característica distinta em relação ao que se poderia esperar de famílias jovens. No tipo de família 1 tem-se a presença de um filho adulto, enquanto no agrupamento familiar do tipo 5, apesar da renda baixa, tem-se um patrimônio familiar mais significativo. Para esses dois grupos está-se diante de características de famílias velhas, e não jovens, como afirmado anteriormente na análise de Freitas (2000b), a partir de variáveis isoladas. Essas afirmações poderiam ser corroboradas por estatísticas descritivas a partir do banco de dados existente, identificando os casos correspondentes a essas células da tabela.

Dito de outra maneira, a concentração anterior do esforço de marketing para comercialização de imóveis de 1 quarto em famílias jovens estaria perdendo a oportunidade de explorar conjuntamente as questões de renda e patrimônio, que este novo modelo mostrou importantes para definir segmentos de mercado. Por um lado, as famílias jovens poderiam almejar imóveis maiores, adequando suas pretensões dentro das 5 categorias de preços de imóveis existentes para cada célula. Por outro lado, poder-se-ia identificar melhor as preferências de famílias mais velhas em relação a imóveis pretensamente menores. Se esse resultado for analisado no âmbito do departamento de *marketing* da empresa, tem-se uma informação relevante para a elaboração das estratégias do produto.

A diferença essencial desses resultados para outros advindos de estudos já feitos no mercado imobiliário é que apenas se consideram as variáveis que são realmente relevantes, sem forçar a explanação de agrupamentos familiares e de produtos. Essa característica distingue a proposta de modelagem do mercado imobiliário aqui apresentada da análise de *cluster*, já que esta agrega rigidamente variáveis do mercado (relativas aos clientes, aos produtos ou a ambas) e estereotipa os agrupamentos assim formados. Aqui

são indicados os cruzamentos que têm maior ocorrência, mas ainda abrindo caminho para a análise de todos os segmentos de mercado associados a um segmento de produto e vice-versa.

Conclusões

Da revisão teórica da área de segmentação de mercado é possível afirmar que há uma dificuldade de utilização dos resultados para fins mais concretos, como a implementação de estratégias de marketing. Entre as técnicas disponíveis, o CHAID apresenta-se como uma alternativa, consideradas as vantagens e particularidades de uso citadas ao longo deste texto.

Os resultados gerados através da aplicação do CHAID mostraram com clareza quais as principais bases de segmentação na criação de classes de consumidores no mercado imobiliário. Para explicar o preço procurado para o futuro imóvel, a renda é a principal variável, sendo o valor patrimonial e o tamanho da família as outras duas variáveis mais significativas. Estas duas variáveis irão definir diferentes grupos, sendo a primeira mais relacionada às famílias de maior poder aquisitivo, e a última, às de menor renda.

A partir da criação de segmentos de mercado, a associação com os imóveis experimentou uma redução de variabilidade, consequência da consideração das variáveis de forma conjunta. Além disso, a interpretação dos resultados ajuda na escolha do segmento-alvo e na implementação das estratégias de marketing.

Uma limitação que pode ser apontada no estudo é o fato de não ter sido considerada a disponibilidade de crédito nos mercados habitacionais pesquisados e o efeito dela nas preferências habitacionais. Entende-se também como limitação que o esgotamento de todas as variáveis independentes para criar segmentos de produtos e de mercados termina deixando o analista sem variáveis livres, verdadeiramente independentes, para conduzir análises dentro de cada célula do modelo; por exemplo, através de regressores logísticos. É o caso de explorar como se dá a preferência pelas cinco faixas de preços dentro de cada célula: é fácil se imaginar que a preferência por valores maiores de preços deve estar associada a variáveis econômicas, como renda e patrimônio familiar. Ocorre que essas variáveis já foram utilizadas, neste trabalho, para criar a maioria dos segmentos de mercado. Ressalta-se ainda que a escolha do preço como variável categórica deve ser interpretada com precaução, já que todas as outras variáveis influenciam o preço, enquanto variável de análise.

Finalmente, propõe-se que a metodologia seja seguida por outros estudos da mesma natureza, nos quais se deve discutir principalmente a escolha da base de segmentação e a interpretação dos dados para uso de definição de estratégias de marketing. Esta última discussão visa aproximar os estudos produzidos no meio acadêmico às necessidades encontradas nas relações práticas das empresas com o mercado.

Referências

ADAIR, A. S.; MCGREAL, W. S. Local housing markets, structure and buyer behaviour: The East Belfast case. **Scandinavian Housing & Planning Research**, v. 11, p. 27-38, 1994.

ALDENDERFER, M. S.; BLASHFIELD, R. K. **Cluster analysis**. Beverly Hills, CA: Sage, 1984.

BROMERCHENKEL, M. N. Comportamento do Consumidor Adolescente: Um Estudo de Segmentação do Mercado de Entretenimento. In: ENCONTRO NACIONAL DA ANPAD, 1996, **Anais...** 1996. p. 209-224.

CHATURVEDI, A.; CAROLL, J. D.; GREEN, P. E.; ROTONDO, J. A. A feature-based approach to market segmentation via Overlapping K-Centroids Clustering. **Journal of Marketing Research**, v. XXXIV, p. 370-377, Aug. 1997.

DIBB, S. Modeling in new housing choice: an application. **International Journal of Management Science**, v. 22, n. 6, p. 589-600, 1994.

DIBB, S. Market segmentation: strategies for success. **Marketing Intelligence & Planning**, p. 394-406, 1998.

DICKSON, P. R. Person-situation: segmentations's missing link. **Journal of Marketing**, v. 46, p. 56-64, Fall 1982.

EVERITT, B. **Cluster analysis**. New York, NY: Haltesd, 1980.

FINE, S. H. Toward a theory of segmentation by objectives in social marketing. **Journal of Consumer Research**, v. 7, p. 1-13, June 1980.

FRANK, L. D. *et al.* A hierarchy of sociodemographic and environmental correlates of walking and obesity. **Preventive Medicine**, v. 47, n. 2, p. 172-178, Aug. 2008.

FREITAS, A. A. F. **Segmentação do mercado imobiliário utilizando dados de preferência declarada**. 2000. 150 f. Tese (Doutorado em Engenharia de Produção) – Programa de Pós-

Graduação em Engenharia de Produção, UFSC, Florianópolis, 2000a.

FREITAS, A. A. F. Análise de segmentação utilizando o procedimento de Chi-Square Automation Interaction Detection (CHAID). In: ENCONTRO NACIONAL DA ANPAD, 1996, **Anais...**, Florianópolis, set. 2000b. CD-ROM.

GIBLER, K. M.; NELSON, S. L. Consumer behavior applications to real estate education. **Journal of Real Estate Practice and Education**, v. 6, n. 1, 2003.

GONZÁLEZ, M. A. S.; FORMOSO, C. T. Construção de modelos do mercado imobiliário para análise de viabilidade com regressão e sistemas de regras difusas. **Ambiente Construído**, Porto Alegre, v. 6, n. 4, p. 19-31, 2006.

GREEN, P. E.; KRIEGER, A. M. Segmenting markets with conjoint analysis. **Journal of Marketing**, v. 55, p. 20-31, Oct. 1991.

HALEY, R. Benefit segmentation: a Decision-Oriented Research Tool. In: McCARTHY, E. J. *et al.* **Readings in basic marketing**. Homewood: Irwin, 1978. p. 138-147.

HAWKINS, D. M.; KASS, G. V. Topics in Applied Multivariate Analysis. **Cambridge University Press**, p. 269-302, 1982.

HSU, C. H. C.; KANG, S. K. CHAID-Based Segmentation: international visitors trip characteristics and perceptions. **Journal of Travel Research**, v. 26, n. 2, p. 207-216, 2007.

LEITE, F. L.; FORMOSO, C. T. **Avaliação da satisfação do cliente final de empreendimentos do programa de arrendamento residencial**. Relatório de Estudo de Caso. NORIE, UFRGS, Porto Alegre, Mar. 2004. Disponível em: <<http://www.cpgec.ufrgs.br/norie/requali>>. Acesso em: 30 jul. 2008.

KASS, G. An exploratory technique for investigating large quantities of categorical data. **Applied Statistics**, v. 29, n. 2, p. 119-127, 1980.

KOTLER, P. **Administração de marketing: análise, planejamento, implementação e controle**. 5. ed. São Paulo: Atlas, 1998.

MIRON, L. **Proposta de diretrizes para o gerenciamento dos requisitos do cliente em empreendimentos da construção**. 2002. 150 f. Dissertação (Mestrado em Engenharia Civil) – Escola de Engenharia, Universidade Federal do Rio Grande do Sul, Porto Alegre

RICHARDSON, D. H.; THALHEIMER, R. On the use of grouping methods in the analysis of residential housing markets. **Regional Science and Urban Economics**, v. 12, p. 285-304, 1982.

RUDELIUS, W.; WALTON, J. R.; CROSS, J. C. **Improving the managerial relevance of market segmentation studies**. In: HOUSTON, Michael J. **Review of marketing**. Chicago: American Marketing Association, 1987. p. 385-404.

SATTLER, H.; HARTMANN, A. Commercial use of conjoint analysis. In: HÖCK, M.; VOIGT, K. **Operations management in theory and practice**. Wiesbaden: Gabler Edition Wissenschaft, 2008. 250 p.

SCHOLTEN, H. J.; HOOIMEIJER, P. The use of vacancy chains and loglinear models to analyse relationships among housing sub-markets. **Tijdschrift voor Economische en Sociale Geografie**, v. 75, n. 3, p. 202-213, 1984.

SMITH, W. R. Product differentiation and market segmentation as alternative marketing strategies. **Journal of Marketing**, p. 3-8, July 1956.

SMITH, C. A.; GARBARINO, L. N.; MARTINI, J. Analyzing the leasing criteria of retail tenants. **Journal of Property Management**, v. 40, n. 3, 1992.

WIND, Y. Issues and advances in segmentation research. **Journal of Marketing Research**, v. 15, p. 317-337, Aug. 1978.

WOO, K. Using quality perception to segment customers in service. **Marketing Intelligence & Planning**, p. 418-424, 1998.

WÜBBEN; M.; WANGENHEIM, F. V. Instant customer base analysis: managerial heuristic often "get it right". **Journal of Marketing**, v. 72, n. 3, p. 82-93, 2008.