

3 Metodologia

O objetivo desse capítulo é explicar a forma como a pesquisa foi realizada, classificando-a em termos da linha epistemológica e da classificação dela dentro do campo da estratégia. Além disso, aqui também será caracterizado o nosso universo e amostra, onde é mostrada e fundamentada a forma como a simulação dos dados foi realizada. Explica-se também o método de análise dos dados. O método é a modelagem de equações estruturais (MEE), apoiada por diversas técnicas que garantiram a consistência dos dados.

3.1. Linha epistemológica

A linha epistemológica seguida por esse projeto é fundamentalmente **positivista**. Essa linha foi escolhida, primeiramente, por causa da afinidade do tema com métodos quantitativos. Mais claramente ainda, podemos embasar a escolha de uma linha positivista pelas características apresentadas abaixo, que são consistentes com a descrição de Gil (1991) sobre a linha positivista:

- Em sua maior parte, o estudo se baseia em fatos observáveis (que são efetivamente simulados, mas que poderiam ser medidos), como as estratégias das empresas, o desempenho, os recursos que elas possuem e o ambiente;
- Os fenômenos são explicados pela relação entre eles. Tenta-se relacionar estratégia, desempenho, tipo de indústria, ambiente e recursos para explicar os fenômenos estudados, baseando-se no referencial teórico visto anteriormente;
- O estudo considera partes isoladas (cada empresa que terá os dados simulados) para tentar explicar a realidade. Tenta-se elucidar os fenômenos isoladamente _ cada um dos construtos _ considerando-os independentes para, depois, juntar as conclusões tiradas e explicar o

fenômeno mais abrangente, que é a relação estratégia, indústria, ambiente e recursos vs. desempenho;

- Procura-se uma unidade metodológica. Os dados serão simulados da mesma forma, utilizando-se o mesmo sistema de unidades;
- As técnicas utilizadas são estatísticas, utilizando-se variáveis, operacionalizando-se os conceitos, etc.

Observa-se claramente a utilização do método hipotético-dedutivo, pois os pressupostos para a explicação dos fenômenos foram propostos de acordo com uma teoria embasada em observação e pesquisa e esses pressupostos são testados para se chegar a uma conclusão.

3.2. Metodologia no campo da estratégia

No estudo da estratégia, têm-se vários autores que publicaram trabalhos com o objetivo de classificar estudos e modelos produzidos nessa área de conhecimento. Entre eles, destacam-se os trabalhos de Mintzberg *et al.* (2000), Hambrick (1980), Venkatraman & Camillus (1984), entre outros. Além de classificação, outros trabalhos importantes se preocupam em propor metodologias para fazer a validação dos modelos propostos no campo da estratégia. Entre outros trabalhos desse tipo, destacam-se os de Venkatraman & Grant (1986) e de Scandura & Williams (2000).

Mintzberg *et al.* (2000) propuseram dividir o estudo de estratégia em 10 grandes escolas (para uma breve descrição de cada uma das escolas, consulte a página 10 deste trabalho). O modelo proposto nesse trabalho possui características de várias escolas. Por isso, podemos classificá-lo como fazendo parte da **escola de configuração**.

Essa escola, como citado pelos autores, é a que integra as mensagens das outras escolas. Essa integração fica bem clara quando há no modelo elementos da **escola do posicionamento**, representada pelas teorias de Porter; da **escola de aprendizado**, pois a capacidade de aprendizado é considerada como parte dos recursos da organização; da **escola ambiental**, pois o ambiente é considerado de forma importante como agente da formação das estratégias entre outras escolas.

Além disso, encontra-se, no trabalho, as duas facetas principais dessa escola. Tem-se, ainda, as **configurações**, que são os estados da organização e do contexto que a cerca, representadas pelos construtos *Fatores Ambientais*, *Estrutura da Indústria* e *Processos e equipamentos proprietários*; e a **transformação**, que descreve o processo de geração da estratégia, representada pelo construto *Estratégia*.

Venkatraman & Camillus (1984) criaram uma classificação baseada no conceito de *fit*, ou seja, encaixe no estudo de gestão estratégica. Eles propõem uma tipologia para classificar os estudos de gestão estratégica baseada em duas dimensões: a **conceitualização de *fit*** e o **domínio do *fit***.

Com relação à **conceitualização do *fit***, eles dividem os estudos em duas possibilidades. A primeira teoriza que a estratégia envolve conciliar os vários componentes do mix da estratégia de forma que a estratégia seja vista como um padrão de interações para combinar diversos elementos tanto internos quanto externos à firma. A outra possibilidade, em contrapartida, foca no conteúdo da estratégia com o objetivo de especificar que ações devem ser tomadas para que a estratégia se encaixe com as mais diversas condições ambientais.

Com relação ao **domínio do *fit***, os autores analisam o foco de atuação na formulação da estratégia. O foco pode ser primariamente o *fit* entre fatores externos (relacionados à estrutura do mercado) e variáveis estratégicas, sem referência direta a fatores internos da firma. O foco principal também pode ser no *fit* entre os fatores internos da firma e as variáveis estratégicas, assim como na integração tanto dos fatores internos quanto externos com a formulação da estratégia.

Considerando, então, essas duas dimensões, tem-se a tipologia mostrada no quadro 2 com 6 diferentes possibilidades de classificação ou escolas para os estudos no campo da estratégia:

DOMÍNIO DO FIT	Externo	1 – Escola da formulação da estratégia	4 – Escola de redes de estratégia interorganizacionais
	Interno	2 – Escola da implementação da estratégia	5 – Escola da escolha da estratégia
	Integrado	3 – Escola da implementação de formulação integrada	6 – Escola <i>gestalt</i>
		Conteúdo do fit (elementos a serem alinhados com a estratégia)	Padrão de iterações (processo de chegar ao fit)
CONCEITUALIZAÇÃO DE FIT			

Quadro 2 Tipologia de classificação de estudos em estratégia (Venkatraman & Camillus, 1984)

De acordo com essa tipologia, esse estudo será classificado como parte da **escola da implementação de formulação integrada**, pois o modelo, no que diz respeito à conceitualização de *fit*, se propõe a analisar que estratégias melhor se adaptam ao ambiente e às condições da empresa para que ela tenha um melhor desempenho. Com relação ao domínio do *fit*, o modelo considera que os fatores internos e externos devem ser relacionados com as escolhas estratégicas.

Com o objetivo de operacionalizar o conceito do estudo de estratégias de negócio na pesquisa, Hambrick (1980) também criou uma tipologia para classificação dos trabalhos feitos para essa área. O foco do autor foi classificar de acordo com a forma de operacionalização do construto. A tipologia proposta possui quatro classificações, que são as seguintes:

- Descrições textuais da estratégia – a descrição da estratégia é feita de forma puramente textual. Esse método é basicamente o de estudo de

caso, onde casos são descritos e, a partir dele se chega a diagnósticos com relação à estratégia da firma e suas implicações;

- Medidas de partes da estratégia – a operacionalização do construto estratégia é feito focando-se somente em uma variável (mais notavelmente, *market-share*). Na maior parte dos casos, procura-se identificar uma relação entre essa variável que descreve a estratégia e o desempenho da firma (muitas vezes considerando como desempenho a lucratividade);
- Medida multivariada da estratégia – o construto estratégia, aqui, é representado por uma multivariada. Essa relação se dá procurando-se várias variáveis que fazem parte do “significado” de estratégia de negócio e procurando-se a relação que elas têm entre si e com outras variáveis que descrevem outros construtos (como desempenho, por exemplo);
- Tipologias de estratégia – essa forma de caracterizar o construto estratégia se dá por meio da criação de uma classificação onde as empresas se enquadram de acordo com as estratégias utilizadas. Elas são classificadas de acordo com a semelhança de comportamento quando se comparam diversas variáveis que caracterizariam, para o pesquisador, a estratégia de negócios utilizada.

O autor resumiu, no quadro 2 mostrado abaixo, os critérios a serem utilizados pelo pesquisador para a escolha do método a ser utilizado em sua pesquisa. Esses critérios consideram os objetivos da pesquisa, o papel do construto de estratégia e a visão do pesquisador com relação ao mesmo.

	Descrições textuais da estratégia	Medidas de partes da estratégia	Medida multivariada da estratégia	Tipologias de estratégia
CRITÉRIO				
Objetivo da pesquisa				
Construção da teoria	Útil	Útil	Útil	Útil
Teste da teoria	Inútil	Útil	Útil	Uso limitado
Demonstração de generabilidade	Uso limitado	Útil	Útil	Útil

Melhora da confiabilidade das medidas	Inútil	Útil	Útil	Uso limitado
Papel do construto de estratégia na pesquisa				
Variável preditora	Útil	Útil	Útil	Útil
Variável moderadora	Útil	Útil	Uso limitado	Útil
Variável de critério	Útil	Útil	Inútil	Útil
Visão do pesquisador sobre o construto de estratégia				
Estratégia pretendida	Útil	Útil	Uso limitado	Uso limitado
Estratégia realizada	Útil	Útil	Útil	Útil
Estratégia é compreensiva	Útil	Inútil	Útil	Útil
Estratégia é integrativa (tem lógica interna)	Útil	Inútil	Inútil	Útil

Quadro 3 Critérios para escolha de método de operacionalização do construto estratégia (Hambrick, 1980)

Considerando-se a classificação mostrada acima, o presente estudo pode ser classificado como **Medida multivariada da estratégia**. Isso é bem claro, pois, para caracterizar a estratégia da firma, utilizamos um total de 16 variáveis (que já foram explicadas acima). As técnicas estatísticas que serão explicadas mais tarde procurarão explicar as relações entre essas variáveis e os outros construtos do modelo, com a intenção de confirmar o que é previsto pela teoria já apresentada.

O passo a ser dado após classificar o estudo no campo da estratégia é garantir a validade dos construtos que fazem parte do modelo. Com relação a isso, Venkatraman & Grant (1986) propuseram que, para um construto ser considerado válido, o pesquisador deve garantir que ele tenha validade, consistência interna, validação de convergência, validação de discriminante e validação nomológica. O quadro 3, retirado do mesmo artigo referenciado anteriormente e mostrado abaixo,

contém uma explicação do que é cada um desses conceitos e propõe técnicas para testá-los.

Componente	Definição	Técnicas relevantes/ <i>framework</i> analítico
Validação de conteúdo	O quanto medidas empíricas refletem um domínio específico do conteúdo.	Revisão por especialistas e análise do grau de consistência entre os construtos
Consistência interna	O quanto cada uma das variáveis refletem somente um conceito	
1. Unidimensionalidade	Ausência de erro de medida	• Análise fatorial confirmatória • Análise fatorial exploratória
2. Confiabilidade	O quanto múltiplas tentativas de se medir o mesmo conceito com diferentes métodos convergem para o mesmo resultado	• Alfa de Cronbach • Coeficiente de generabilidade dos modelos de equações estruturais
Validação de convergência	O quanto um conceito difere dos outros conceitos	• Análise de correlação • Matriz MTMM • Metodologia de equações estruturais – análise fatorial confirmatória
Validação de discriminante	O quanto as predições de um modelo teórico são confirmadas	• Análise de correlação • Matriz MTMM • Metodologia de equações estruturais – análise fatorial confirmatória
Validação (preditiva) nomológica		• Correlações • Regressões • Modelo causal

Quadro 4 Um resumo dos componentes chave da validação de construto (Venkatraman & Grant, 1986)

Além do trabalho de Venkatraman & Grant, Scandura & Williams (2000) também trabalharam no sentido de identificar tendências de como validar trabalhos na área de estratégia. Nesse sentido, eles consideraram a proposta feita

por Cook & Campbell (1976), que considera quatro tipos de validações a serem feitas:

- Validação interna – as relações de causa e efeito somente podem ser verificadas se existe uma correlação entre as variáveis. Os procedimentos usados para coletar os dados demonstram que a causa precede o efeito e explicações alternativas foram descartadas;
- Validação externa – a garantia de que o modelo proposto é válido através do tempo, para qualquer amostra da população ou para qualquer outra combinação de fatores ambientais que o pesquisador considere que esteja no escopo da pesquisa;
- Validação de construto – como as medidas empregadas são coerentes com a teoria na qual o modelo e o teste são baseados. Essa validação é a mesma considerada por Venkatraman & Grant (1986) e já falada anteriormente com mais detalhes.
- Validação das conclusões estatísticas – a habilidade de se chegar a conclusões baseadas nas evidências estatísticas encontradas.

O quadro 4 mostra métodos verificados pelos autores que são feitos para se conseguir cada uma das validações descritas acima que foram utilizados na literatura:

Validação interna	Validação externa	Validação de construto	Validação das conclusões estatísticas
• Frames temporais • Pesquisa em estratégia	• Tipo de amostra • Ocupação dos sujeitos • Pesquisa em estratégia	• Análise fatorial confirmatória • Análise fatorial confirmatória • Validade preditiva / discriminante / convergente • Generabilidade	• Tamanho da amostra • Número de variáveis dependentes • Métodos de análises de dados

Quadro 5 Métodos utilizados na literatura para se fazer validação de estudos em estratégia (Scandura e Williams, 2000)

O presente estudo conta com uma grande preocupação com relação à validade de seu modelo. Por isso, usaram-se os *frameworks* acima descritos para se fazerem as validações necessárias. A validação interna, assim como a externa, foram feitas através da pesquisa na literatura de estratégia que já foi mostrada acima. A validação de construto foi realizada com os seguintes métodos:

- Validação de conteúdo – revisão dos construtos por especialistas em estratégia de negócios;
- Validação de consistência / unidimensionalidade – Análise fatorial exploratória e confirmatória (SEM);
- Validação de consistência / generabilidade – Alfa de Cronbach e coeficiente de generabilidade do modelo de equações estruturais;
- Validação de convergência – Análise fatorial confirmatória;
- Validação de discriminante – Análise fatorial confirmatória;
- Validação nomológica – como o modelo proposto não tem a pretensão de fazer previsões, essa validação não é feita.

A validação das conclusões estatísticas é feita garantindo-se um tamanho da amostra, número de casos por variável etc. necessários para que os métodos utilizados tenham validade.

O item 3.4 desse trabalho, que fala sobre os métodos utilizados para a análise dos dados, irá mais fundo na explicação de cada uma das técnicas estatísticas utilizadas e citadas acima para garantir as validações necessárias.

3.3. Metodologia para análise dos dados

Para operacionalizar o modelo, foram usadas duas técnicas estatísticas: análise de fator e modelagem de equações estruturais.

A primeira técnica foi usada para se reduzir o número de variáveis associadas a cada um dos construtos, pois, inicialmente, tinha-se 71 variáveis para representar os 7 construtos.

A segunda técnica foi utilizada para encontrar as relações entre as variáveis que representavam cada um dos construtos e as relações entre os construtos. Essa técnica foi utilizada porque esta é a melhor forma de

analisar sistemas de equações onde a variável dependente em uma equação é independente em outra.

3.3.1. Análise de Fator

Segundo Hair *et al.* (1998), análise de fator é um conjunto de técnicas estatísticas multivariadas que tentam encontrar relações latentes entre variáveis. Elas fazem isso reduzindo as variáveis existentes em fatores que são explicados por meio de equações onde as variáveis independentes são as iniciais, como podemos ver no exemplo abaixo:

$$\textbf{Fator 1} = 0,49x_1 + 0,32x_2 + 0,91x_3 - 0,11x_4 + 0,57x_5 \quad \text{Equação (1)}$$

A análise de fator tem dois objetivos principais:

- Explicar as relações latentes entre variáveis – descobrir que relações escondidas as variáveis possuem. Isso é feito quando a análise de fator é a técnica principal de um estudo e é utilizada para comprovar alguma teoria que diz que as variáveis se comportam de determinada forma à luz do conhecimento produzido sobre o assunto;
- Redução dos dados – quando se tem muitas variáveis e o objetivo é reduzi-las, na maioria das vezes, para aplicá-las em outra técnica estatística.

Dependendo do método utilizado, os fatores produzidos implicam numa perda do poder explicativo, se comparados às variáveis iniciais, pois nem toda a variância da variável é usada. Se for utilizada a técnica de **análise componente**, toda a variância é considerada. Em contrapartida, se for usada a **análise de fator comum**, existe uma perda, pois, nessa técnica, separam-se três tipos de variância de cada variável:

- Variância comum – é a parte da variância que a variável compartilha com outras variáveis da análise;

- Variância específica – a variância que é específica daquela variável;
- Variância do erro – é a variância obtida devido a imprecisões na coleta dos dados, erros de média, etc.

A técnica, então, monta os fatores considerando somente a variância comum e desconsiderando as outras, ocorrendo, então, a perda.

Com relação às premissas, precisa-se de, pelo menos, 5 observações para cada variável a ser analisada, sendo um mínimo de 10 observações ideal. Além disso, é desejável um bom nível de **multicolinearidade** entre as variáveis. Pode-se identificar se elas existem de modo que a aplicação da análise de fator é apropriada pela análise da matriz de correlação das variáveis. Dois testes podem ser utilizados:

- Teste de esfericidade de Bartlett – testa a hipótese nula de que existem correlações entre as variáveis. Um valor próximo de 1 nesse teste é desejável.
- Medida de adequação da amostra – vai de 0 a 1, sendo que um valor razoável seria acima de 0,7. Valores abaixo de 0,5 são inaceitáveis.

Após rodar o método, deve-se escolher quantos fatores serão utilizados. Alguns critérios para isso podem ser utilizados. O critério mais utilizado é o do **autovalor**. Todo fator tem um autovalor a ele associado e estabelece-se um valor de corte, que geralmente é 1, pois indica que o fator contabiliza a variância de, pelo menos, uma variável inicial.

Pode-se também estabelecer um critério de escolha a priori, onde o número de fatores é escolhido, pois a teoria indica um determinado número deles.

A porcentagem de variância explicada também pode ser o critério para se escolher o número. Os fatores que contabilizam maior variância ficam até que a soma das variâncias seja satisfatória.

Outro critério é o do **scree test**. Gera-se um gráfico das variáveis versus seu autovalor associado. Quando o gráfico demonstra uma queda brusca na variação do autovalor de uma variável até a próxima com maior autovalor, faz-se o corte.

3.3.1.1. Interpretação dos resultados

Os fatores resultantes da análise podem não ser facilmente interpretados por que a variância das variáveis iniciais pode ser distribuída entre os fatores de forma que uma interpretação dos mesmos não seja fácil. Para facilitar a interpretação dos resultados e a explicação do significado dos fatores, existem recursos que podem ser utilizados. Pode-se fazer uma rotação de fatores para que eles possam ser melhor explicados de acordo com as variáveis iniciais que têm peso significativo nele. As seguintes técnicas podem ser usadas:

- Varimax – método de rotação ortogonal que tenta fazer com que as variáveis iniciais tenham peso 1 ou 0 em cada fator;
- Quartimax – método também ortogonal que tenta fazer com que uma variável tenha peso grande em um fator e pequeno nos outros;
- Equimax – esse método tenta fazer o que os dois anteriores fazem ao mesmo tempo;
- Outras rotações oblíquas, como Promax, Oblimin, etc.

Com relação aos pesos que as variáveis iniciais têm em cada fator, pode-se considerar que um peso acima de 0,5 ou abaixo de -0,5 indica que aquela variável é representativa naquele fator. Sendo assim, ao interpretá-lo, isso deve ser levado em consideração.

Para se testar a confiabilidade das variáveis que, à luz da teoria, deveriam medir o mesmo conceito, pode-se usar o **alpha de Cronbach**. Se o seu valor for acima de 0,7, tem-se uma boa confiabilidade a ponto de se poder fazer uma média dessas variáveis para representar o conceito. Isso é útil, pois, se há um fator com altos pesos de algumas variáveis, o teste pode ser feito com elas para ver se o fator será usado ou se será possível abrir mão do mesmo e fazer a média, o que facilita a interpretação.

3.3.2. Modelagem de Equações Estruturais (MEE)

Modelagem de Equações Estruturais (MEE) é uma técnica que expande a habilidade explanatória do pesquisador, pois ela permite analisar estatisticamente vários relacionamentos de uma vez (Hair *at al.*, 1998). Ao contrário da análise de fator, regressão múltipla, correlação canônica ou MANOVA, essa técnica combina vários métodos estatísticos e permite estabelecer relações onde se possui várias variáveis dependentes, independentes e relacionamentos diferentes entre elas.

MEE tem a capacidade de fazer análises quando se tem múltiplas equações onde a variável dependente de uma equação é independente em outras, como mostra-se no exemplo abaixo, tirado de Hair *at al.*(1998), pp. 597:

$$Y_1 = b_1X_1 + b_2X_2$$

$$Y_2 = b_3X_2 + b_4X_3 + b_5Y_1 + b_6Y_3$$

$$Y_0 = b_7Y_1 + b_8Y_2$$

Os objetivos da MME são basicamente 3, segundo Garson (2006):

- Análise confirmatória – quando se tem um modelo teórico que precisa ser testado pelo pesquisador. Nesse tipo de estudo, o *fit* do modelo é testado por testes que determinam se os padrões de variância e covariância dos dados são consistentes com o modelo desenvolvido;
- Modelos alternativos – quando se tem mais de um modelo teórico e se quer chegar à conclusão de qual se encaixa melhor aos dados possuídos pelo pesquisador. Os testes são feitos da mesma forma que acima para todos os modelos e o *fit* é comparado.
- Desenvolvimento de um modelo – um modelo, a priori definido, é testado e modificado segundo a sugestão de índices de modificações da MEE. Esse método pode gerar um modelo com um excelente encaixe nos dados, mas como ele não foi definido com base na teoria anteriormente, corre-se o risco de não se adaptar bem a outros dados. No caso desse objetivo, é sempre desejado que análises confirmatórias sejam feitas com novos dados.

Os modelos gerados pelo pesquisador que podem ser testados pela MEE são diagramas de caminhos (que podem ser representados por conjuntos de equações do tipo mostrado acima). Tanto os construtos do diagrama quanto as direções das setas devem ser definidas a priori. A figura abaixo mostra um exemplo de um diagrama de caminhos que pode ser operacionalizado por essa técnica.

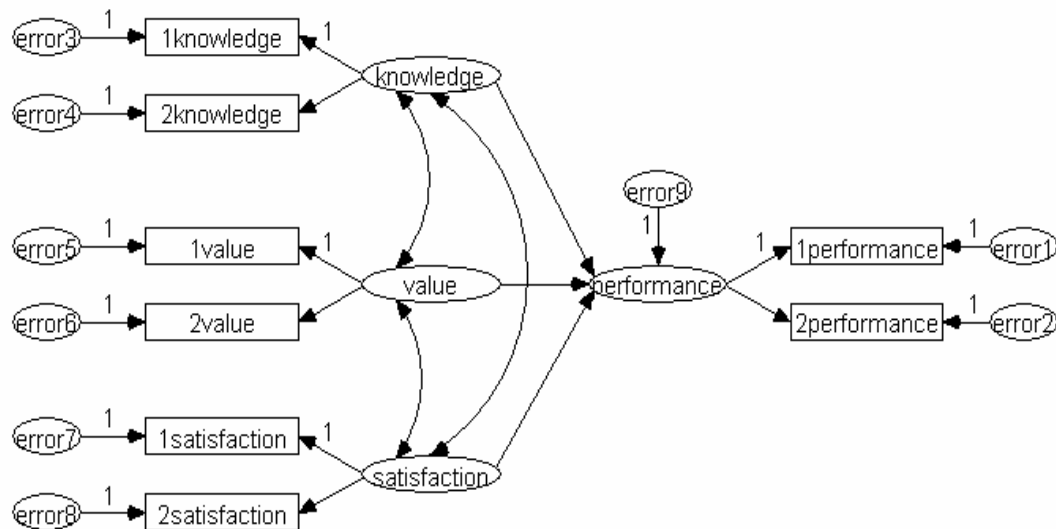


Figura 15 Diagrama de caminho (Arbuckle *et al.*, 1999 – pp. 125)

As elipses maiores são os construtos do modelo ou variáveis latentes ou variáveis endógenas que representam no sistema de equações as variáveis que são tanto dependentes quanto independentes em alguma equação. Os retângulos são os indicadores ou variáveis exógenas e são somente variáveis independentes nas equações. As elipses menores são os erros associados a cada uma das variáveis.

A entrada para MEE é a matriz de correlação ou de covariância entre as variáveis. As premissas de MEE dependem do procedimento de estimação escolhido (alguns dos procedimentos serão explicados à frente). Mas, basicamente, eles variam mais na questão de tamanho da amostra e do quão rigorosos são para umas ou outras premissas. Algumas das mais importantes citadas por Garson (2006) são as seguintes:

- Normalidade multivariada dos indicadores (variáveis independentes de cada construto) – a ausência de normalidade multivariada dos

indicadores pode diminuir a utilidade do modelo, pois tende a inflar o valor *chi-square* (que indica o quão diferente o modelo é das matrizes de covariância, sendo mais diferentes quando o valor é maior);

- Normalidade multivariada das variáveis latentes (variáveis dependentes ou construtos) – a normalidade das variáveis latentes também é necessária para MEE;
- Linearidade – MEE assume uma relação linear entre os indicadores e as variáveis latentes;
- Múltiplos indicadores – devem ser usados, pelo menos, três indicadores para cada variável latente;
- Problemas de identificação - devem-se ter mais elementos de covariância (duplas de variáveis latentes) do que setas no modelo indicando relações;
- Covariâncias maiores do que zero – correlações entre variáveis são necessárias para se utilizar MEE;
- Tamanho da amostra – depende muito do procedimento de estimação, mas recomenda-se entre 200 e 400 observações para modelos entre 10 e 15 indicadores.

Existem vários procedimentos de estimação diferentes para MEE. Alguns deles são citados abaixo:

- Maximum likelihood estimation (MLE) – o procedimento mais comum de estimação. Tem uma amostra mínima de 100 a 150 observações (Hair *et al.*, 1998), com um ideal de 200 (mais do que isso pode dar super sensibilidade ao modelo). Esse procedimento é muito sensível à normalidade multivariada;
- Generalized least squares (GLS) – requer muito menos recursos computacionais do que a MLE e é mais suscetível à convergência. Quando a amostra é grande (maior do que 2500), trabalha bem mesmo sem a normalidade multivariada dos dados (Garson, 2006);
- Ordinary least squares (OLS) – é a forma de regressão tradicional. Ela estima as equações separadamente e, por isso, outros métodos são preferenciais. Esse método não é tão sensível à normalidade multivariada quanto MLE (Garson, 2006).

- Two-stages least squares (2SLS) – esse método é uma adaptação do OLS para suportar erros correlacionados. Também estima as equações separadamente e, por isso, é pior do que MLE (Garson, 2006).
- Weighted least squares (WLS) – o método costuma dar estimativas exageradamente otimistas e tem piores resultados do que MLE e GLS (Garson, 2004). Precisa de uma amostra maior do que 2000.
- Unweighted least squares (ULS) – não gera coeficiente *chi-square* e não assume normalidade multivariada (Garson, 2006).
- Asymptotically distribution-free (ADF) – essa técnica não assume normalidade multivariada e necessita de uma amostra entre 200 e 500. A entrada de dados para essa técnica não é a matriz de correlação, mas os dados em si (Garson, 2006).
- Elliptical distribution theory (EDT) – não necessita de normalidade multivariada e precisa de uma amostra maior do que 2500 observações.

O GFI (*goodness-of-fit*) é o indicador usado para verificar se o modelo tem uma boa utilidade. Esse indicador mede o grau em que as matrizes de entrada reais e observadas (matrizes de covariância e correlação) são previstas pelo modelo estimado (Hair *et al.*, 1998). O indicador AGFI (*adjusted goodness-of-fit*) ajusta o GFI com os graus de liberdade do modelo. Ambos os indicadores vão de 0 a 1 e é indicado ter um AGFI maior do que 0,8.

Nesse trabalho, o método Maximum likelihood estimation (MLE) de estimação será usado.

3.4. Caracterização da amostra

Os dados relativos às firmas que foram usados nesse trabalho foram simulados. Primeiramente, foram simuladas todas as variáveis do modelo citadas na seção 2.1.3.5 desse trabalho para 3000 firmas. Após isso, foi feita uma análise de fator para reduzir o número de variáveis. Posteriormente, percebeu-se que, com a simulação simples feita anteriormente, o modelo operacionalizado com MEE não tinha bom encaixe nos dados. Dessa forma, escolheu-se utilizar os fatores selecionados no momento anterior e simulá-los diretamente para que fossem encontrados dados que atendessem melhor às premissas da técnica utilizada. A

simulação final foi atingida após diversas tentativas para se chegar a uma amostra final com dados que dessem bom encaixe com o modelo.

A primeira amostra foi gerada utilizando-se o gerador de números normais da ferramenta SPSS. Para cada uma das 71 variáveis iniciais do modelo, foram gerados 3000 casos em que a distribuição deles desse normal com média 0 e desvio padrão 1. Dessa forma, gerou-se dados que, se fossem reais, teriam sido passados por uma transformada Z, sendo $z = (X - \mu) / \sigma$.

Após serem gerados os 3000 casos acima, as variáveis foram reduzidas a 38 por análise de fator. Esse procedimento será melhor explicado na seção 4. A partir daí, diversas tentativas de gerar fatores que tivessem um encaixe aceitável nos três modelos a serem analisados foram feitas. A simulação final considerou as seguintes relações entre as variáveis:

- Fatores FA1 a FA10, AI1 a AI3 e AE1 a AE3 - esses fatores são as variáveis que descrevem os construtos exógenos do modelo proposto. Os modelos menores baseados na escola posicionamento e *resource-based* são “pedaços” do modelo proposto. Sendo assim, essas variáveis são simuladas aleatoriamente da mesma forma que foram gerados os dados antes da análise de fator, ou seja, cada um dos fatores é gerado por um gerador com distribuição normal, média zero e desvio padrão 1.
- Fatores PE1 a PE5 – esses fatores são as variáveis descritoras do construto Processos e Equipamentos Proprietários. Esse construto é endógeno no modelo principal e recebe impacto dos construtos Aprendizado Interno, Aprendizado Externo e Fatores Ambientais. Dessa forma, essas variáveis foram geradas como uma combinação linear das variáveis AI1 a AI3, AE1 a AE3 e FA1 a FA10 mais um erro normal (erro gerado com média 0 e desvio padrão 1) multiplicado por uma constante (4). Os multiplicadores das variáveis geradoras foram gerados aleatoriamente dentro de um intervalo entre -1 e 1.
- Fatores E1 a E8 – esses fatores são as variáveis descritoras do construto Estratégia. Esse construto é endógeno no modelo principal e recebe impacto dos construtos Processos e Equipamentos Proprietários e Fatores Ambientais. Dessa forma, essas variáveis foram geradas como uma combinação linear das variáveis PE1 a PE5 e FA1 a FA10

mais um erro normal (erro gerado com média 0 e desvio padrão 1) multiplicado pela mesma constante (4). Os multiplicadores das variáveis geradoras foram gerados aleatoriamente dentro de um intervalo entre -1 e 1.

- Fatores EI1 a EI3 – esses fatores são as variáveis descritoras do construto Estrutura da Indústria. Esse construto é endógeno no modelo principal e recebe impacto dos construtos Estratégia e Fatores Ambientais. Dessa forma, essas variáveis foram geradas como uma combinação linear das variáveis E1 a E8 e FA1 a FA10 mais um erro normal (erro gerado com média 0 e desvio padrão 1) multiplicado por uma constante (4). Os multiplicadores das variáveis geradoras foram gerados aleatoriamente dentro de um intervalo entre -1 e 1.
- Fatores D1 a D6 – esses fatores são as variáveis descritoras do construto Desempenho. Esse construto é endógeno no modelo principal e recebe impacto dos construtos Estrutura da Indústria, Estratégia e Processo e Equipamentos Proprietários. Dessa forma, essas variáveis foram geradas como uma combinação linear das variáveis PE1 a PE5, E1 a E8 e EI1 a EI3 mais um erro normal (erro gerado com média 0 e desvio padrão 1) multiplicado por uma constante (4). Os multiplicadores das variáveis geradoras foram gerados aleatoriamente dentro de um intervalo entre -1 e 1.

Esse método de geração dos dados garante que as variáveis dos modelos tenham uma relação entre si, mas não tenham uma correlação muito alta devido ao erro. Se as variáveis forem geradas sem nenhuma correlação, o modelo vai ter peso zero em todas as suas setas. Se tiverem correlação alta, a técnica Maximum likelihood estimation (MLE), que é usada para operacionalizar o modelo nesse trabalho, tem dificuldade de atingir convergência. Sendo assim, a simulação deve garantir convergência com significância na relação entre os construtos.

Com relação à quantidade de casos a serem utilizados em cada um dos modelos, para o modelo proposto (modelo 1), que possui 38 variáveis e 7 construtos, são utilizados os 3000 casos gerados. Para o modelo baseado na escola posicionamento (modelo 2), que possui 17 variáveis e 3 construtos; e para o modelo baseado na escola resource-based (modelo 3), que possui 16 variáveis e 4 construtos, 400 casos escolhidos aleatoriamente são utilizados para evitar *over-fit*.