

## 4 Base de Dados e Metodologia

### 4.1 Base de dados

Para o desenvolvimento do trabalho foi necessária a coleta de dados de empresas ligadas à indústria de energia ao redor do mundo. A amostra coletada inclui informações operacionais e econômico-financeiras do período compreendido entre 1989 e 2012. Os dados foram extraídos da base de dados da EVALUATE ENERGY, que consistiam de 473 empresas no total.

Foram excluídas as empresas que não apresentavam dados operacionais em nenhum dos segmentos de negócio de petróleo (produção, refino e *marketing*), resultando em uma amostra final de 339 empresas de 48 países (tabela 1). As informações constituem dados em painel não-balanceados, já que o número de informações varia por companhia. Dada a dificuldade em obter dados de qualidade para *midstream*, esse segmento não será objeto de estudo nesse trabalho.

Dados em painel são aqueles caracterizados por apresentarem observações em duas dimensões, geralmente, tempo e espaço, ou seja, é a combinação dos dados em séries temporais e os dados em *cross-section*. Os dados em série temporal são um conjunto de observações dos valores que uma variável assume em diferentes momentos. Já os dados em *cross-section*, são dados de uma ou mais variáveis coletadas no mesmo ponto do tempo (GUJARATI 2000, BROOKS 2008).

Pela tabela 1, pode-se observar que, embora a amostra englobe empresas de todos os continentes, há uma forte concentração de empresas em alguns países. Os cinco países com maior representatividade na amostra em número de empresas são; os Estados Unidos com 42,5% do total da amostra (144 empresas), seguidos de empresas do Canadá e Reino Unido com 13% cada (ou 44 empresas), Austrália com 3,8% das companhias da amostra (ou 13 empresas) e Rússia com 3% (ou 10 empresas).

**Tabela 1** - Amostra, por País

País	Nº de empresas
Estados Unidos	144
Canadá	44
Reino Unido	44
Austrália	13
Rússia	10
Europa (outros)	41
Ásia (outros)	23
América Latina (outros)	11
Oriente Médio (outros)	4
Oceania (outros)	3
África (outros)	2

Fonte: Elaboração própria, com base no Evaluate Energy

Da amostra, quase 90% das empresas atuam no segmento de exploração e produção de petróleo (E&P ou *upstream*), enquanto que somente 30% das empresas atuam nas atividades de refino e *marketing* (*downstream*), como observado na tabela 2. Isso tende a mostrar que as atividades de *upstream* tendem a ter uma maior concorrência entre as empresas quando comparadas as atividades de *downstream*.

**Tabela 2** - Estágios de Produção por Empresa

Estágio	Empresas	%
E&P	305	89,97%
Refino	104	30,68%
Marketing	98	28,91%

Fonte: Elaboração própria, com base no Evaluate Energy

No que diz respeito ao grau de integração das empresas da amostra, a tabela 3 mostra que 71% das empresas atuam em apenas um segmento da indústria de petróleo, ou seja, são não-integradas (especializadas). As empresas que apresentam algum tipo de integração vertical somam 29% das empresas da amostra, dos quais aproximadamente 18,5% das empresas têm atuação nos três segmentos da indústria (exploração/produção, refino e *marketing*), ou seja, têm atuação totalmente integrada, enquanto as empresas restantes atuam de forma semi-integrada (atuando em dois segmentos). Em princípio, essa grande concentração de empresas não-integradas tende a contrariar a percepção comum de que a indústria petrolífera se caracteriza principalmente por um alto nível de integração vertical. Isso decorre do fato de que as maiores empresas dessa

indústria (*majors*) tendem a apresentar uma estrutura integrada de atuação (do *upstream* ao *downstream*).

**Tabela 3** - Graus de Integração por Empresa

Estágio	Empresas	%
Especializada	241	71,09%
Semi-integrada	35	10,32%
Integrada	63	18,58%

Fonte: Elaboração própria, com base no Evaluate Energy

As tabelas 4, 5 e 6 mostram o quão relevante é nossa amostra das atividades de produção, refino de petróleo e *marketing* de produtos, respectivamente, em relação ao total da indústria de petróleo. A tabela 4 apresenta que a amostra utilizada corresponde a 23,51% da produção mundial de petróleo em 1989, crescendo para 54,74% em 2012. A tabela 5 considera que a amostra utilizada no trabalho responde por 36,06% da carga processada mundial em 1989 e 73,24% em 2012, já para o *marketing* a amostra responde por 44,07% em 1989 crescendo para 84,15% em 2012 em relação ao total global, mostrado pela tabela 6.

Para as atividades de *downstream* (refino e *marketing*), a amostra contempla grande parte do mercado mundial, enquanto que para o *upstream* (produção de petróleo), apesar da amostra responder por pouco mais da metade da produção global (em 2012), ainda assim permanece sendo uma amostra relevante da produção mundial de petróleo. Com isso, pode-se concluir que a amostra corresponde a parte relevante das empresas participantes da indústria de petróleo, e que há uma quantidade maior de empresas que atuam em *upstream* do que em *downstream*.

**Tabela 4** - Relevância da Amostra entre no Upstream mundial (Produção Petróleo - mil bbl/dia)

	1989	1990	1995	2000	2005	2010	2012
Amostra	15.047	16.229	22.916	37.235	48.588	51.537	47.156
Total Mundo	64.002	65.385	67.990	74.955	82.014	83.272	86.152
%	23,51%	24,82%	33,70%	49,68%	59,24%	61,89%	54,74%

Fonte: Elaboração própria, com base no Evaluate Energy e BP Statistical Review

**Tabela 5 - Relevância da Amostra no Refino (Refino - mil bbl/dia)**

	1989	1990	1995	2000	2005	2010	2012
Amostra	22.099	24.885	31.440	44.994	54.498	60.248	55.831
Mundo Total	61.276	61.311	62.815	68.429	73.976	75.293	76.233
%	36,06%	40,59%	50,05%	65,75%	73,67%	80,02%	73,24%

Fonte: Elaboração própria, com base no Evaluate Energy e BP Statistical Review

**Tabela 6 - Relevância da Amostra no Marketing (Marketing - mil bbl/dia)**

	1989	1990	1995	2000	2005	2010	2012
Amostra	27.002	31.208	39.127	59.000	64.832	66.588	64.150
Mundo Total	61.276	61.311	62.815	68.429	73.976	75.293	76.233
%	44,07%	50,90%	62,29%	86,22%	87,64%	88,44%	84,15%

Fonte: Elaboração própria, com base no Evaluate Energy e BP Statistical Review

## 4.2 Metodologia

No intuito de testarmos se existe uma relação positiva entre a integração vertical e um melhor desempenho operacional e/ou econômico/financeiro nas empresas envolvidas na indústria de petróleo, como afirma alguns autores, utiliza-se um modelo semelhante ao utilizado por Barrera-Rey (1995). O modelo utilizado é uma função de produção Cobb-Douglas.

Segundo Gujarati (2000), a função de produção de Cobb-Douglas na forma estocástica é expressa a seguir;

$$Y_i = \beta_1 X_{2i}^{\beta_2} X_{3i}^{\beta_3} e^{u_i} \quad (1)$$

onde Y é o produto, X<sub>2</sub> é o trabalho, X<sub>3</sub> o capital, u o termo de perturbação estocástico e e a base do logaritmo natural. Observa-se que a relação entre a produção e os insumos é não-linear. Com o objetivo de transformarmos essa relação em linear utiliza-se o logaritmo, assim, obtém-se:

$$\ln Y_i = \ln \beta_1 + \beta_2 \ln X_{2i} + \beta_3 \ln X_{3i} + u_i \quad (2)$$

Agora, o modelo se torna linear, ou seja, um modelo de regressão linear para os parâmetros  $\beta$ 's, embora seja não-linear para as variáveis X e Y, e sim, linear nos logs dessas variáveis, sendo um modelo log-log, log duplo ou log-linear.

Primeiramente, utilizaremos o modelo definido pela equação 3 para verificarmos como se comportariam as vendas em resposta à alterações em determinadas variáveis. Assim, a equação utilizada para isso é:

$$y_{it} = c + \gamma y_{it-1} + \alpha l_{it} + \beta k_{it} + \phi_1 s_{it} + \phi_2 b_{it} + \phi_3 p_{it} + \rho u_t + \omega t_t + \varphi_1 \text{Dummy\_Integração}_{it} + \varphi_2 \text{Dummy\_Estatal}_{it} + \epsilon_{it} \quad (3)$$

onde  $y$  é o log das vendas,  $l$  é o trabalho,  $k$  é o capital,  $s$  é uma medida de poder de mercado,  $p$  um índice de desempenho relativo da empresa,  $b$  é um índice de alavancagem,  $u$  um índice de fatores cíclicos da indústria,  $i$  é a identidade da empresa,  $t$  o período de tempo que é usado como uma variável independente. Ainda são utilizadas duas variáveis *dummy*, referentes à integração e à propriedade das empresas, respectivamente:  $\epsilon$  é o erro com variância  $\sigma^2$  e média 0.

A variável dependente utilizada são as vendas ( $y$ ) devido à maior disponibilidade de obtenção dos dados quando comparados aos dados necessários para se calcular o valor adicionado (que tende a ser uma medida um pouco mais precisa). As vendas ( $y$ ) foram deflacionadas pelo PPI (*Producer Price Index* – Índice de Preços do Produtor) do segmento de extração de petróleo dos Estados Unidos, já que é o país que tem o maior número de empresas da amostra, bem como o segmento de *upstream* ser o que engloba a maioria das empresas da amostra, além de ser o que obtém a maior parcela de investimentos da indústria de petróleo.

Utilizamos como variável independente as vendas das empresas com defasagem de um período ( $y_{it-1}$ ), refletindo a dinâmica das vendas ao longo do tempo. A principal razão em se fazer isso é que existem custos de ajustamento no processo de produção em outras variáveis que se refletem nas vendas (como ajustes em capital e trabalho).

A variável trabalho ( $l$ ) está expressa pelo ln do número de empregados da firma, enquanto que a variável capital ( $k$ ) são os ln dos ativos fixos das empresas (em milhões de dólares). Ambas variáveis são expressas em ln.

A variável de poder de mercado ( $s$ ) foi calculada como sendo a participação da empresa (vendas) em relação ao total da indústria (vendas). Como a amostra corresponde a um painel não-balanceado, além de não considerar a totalidade da indústria, a medida é calculada em termos relativos, considerando-se apenas as

vendas totais da amostra disponível. Considera-se que uma firma com maior participação no mercado tende a obter melhores resultados, graças, por exemplo, às economias de escala.

A variável  $p$  representa uma medida de desempenho relativo, onde se considera a diferença entre o desempenho da empresa, medida pela margem líquida da empresa, e o desempenho médio global da amostra (margem líquida média de toda a amostra), ou seja, representa os retornos anormais obtidos por uma determinada empresa em relação à média da indústria, resultando em quanto uma empresa pode ser mais (ou menos) eficiente que a indústria.

A variável alavancagem ( $b$ ) se relaciona com a hipótese da “disciplina de capital”, considerada por Jensen (1988), onde quanto mais alta a alavancagem da firma, menor será o fluxo de caixa disponível que pode ser usado pelos gestores em atividades que possam reduzir a riqueza dos acionistas. Portanto, quanto maior a exposição à dívida, maior seria a eficiência da firma. Nesse sentido, é considerada no modelo como alavancagem o índice definido pela razão entre o total das dívidas (de curto e longo prazos) e o total do capital dos acionistas.

Tradicionalmente, como medidas de movimentos cíclicos das indústrias, variável  $u$  (ou pressão da demanda), são utilizados os índices de crescimento da própria indústria, no entanto, a indústria de petróleo apresenta outra medida que reflete a pressão da demanda que é o índice de capacidade utilizada do refino. Optou-se por usar essa última medida que é calculada pela razão entre a carga de petróleo processada e a capacidade nominal de refino existente (ambas em mil bbl/dia), para esse cálculo, os dados foram coletados do *BP Statistical Review 2013*.

As variáveis *dummy* consideradas são as seguintes; variável integração e variável acionista estatal. Para a primeira se considera o valor 1 no caso de que a companhia tenha atividades em mais de um segmento na indústria (sendo portanto, semi ou totalmente integrada) e assume valor zero se ela atua em apenas um segmento (especializada). Já para a segunda variável *dummy*, se atribui o valor 1 se a firma tem participação acionária estatal relevante, e valor zero caso contrário. No caso da variável *dummy* de propriedade estatal, considerou-se como empresa com acionista relevante estatal aquelas que constam no banco de dados EVALUATE ENERGY como NOC's (National Oil Companies – Empresas Nacionais de Petróleo).

As variáveis *dummy* são utilizadas para se verificar se uma empresa obtém um desempenho melhor se tem atuação integrada, e/ou se são empresas estatais. Via de regra considera-se que uma empresa integrada apresenta melhores resultados que uma não-integrada pela razões explicadas anteriormente. Também é considerado como norma que uma empresa privada alcança melhores resultados que uma companhia estatal visto que a primeira não apresenta ingerências políticas nem são usadas como meio de realizar políticas governamentais, como costuma acontecer com as empresas estatais.

Utilizou-se como base da Equação 3, o modelo utilizado por Barrera-Rey (1995), que segue abaixo;

$$y_{it} = \lambda_i + \gamma y_{it-1} + \alpha l_{it} + \beta k_{it} + \delta VI_{it} + \phi_1 s_{it} + \phi_2 b_{it} + \phi_3 p_{it} + \rho u_t + \omega t_t + \epsilon_{it} \quad (4)$$

onde as diferenças se encontram na variável efeitos da firma ( $\lambda$ ), que não foi utilizada na Equação 3, e a variável que representava a medida física de integração vertical (VI), que na Equação 3 foi substituída pela variável *dummy* de integração. Uma variável que não estava presente no modelo de Barrera –Rey e se encontra nesse trabalho se refere a variável *dummy* de propriedade estatal. Além disso, os dados utilizado por Barrera-Rey estão restritos à empresas dos Estados Unidos, enquanto nesse trabalho se procurou utilizar dados de empresas de todo o mundo e de período distinto (1974 à 1993 versus 1989 à 2012).

Com relação ao modelo original utilizado por Barrera-Rey (1995), a tabela 7 abaixo resume as principais diferenças e semelhanças em relação ao modelo utilizado nesse trabalho.

**Tabela 7** - Principais diferenças entre o modelo de Barrera-Rey (1995) e o utilizado neste trabalho

Variável	Significado	Barrera-Rey (1995)	Neste Trabalho
$y$	vendas	Log das vendas deflacionadas ou valor adicionado	Somente log das vendas deflacionadas
$\lambda$	efeitos das firmas (" <i>firm effects</i> ")	Captura os fatores específicos das firmas não observados que influenciam a produtividade. Eliminado da equação posteriormente.	Não utilizado
$l$	trabalho	Número total de empregados.	log do número total de empregados
$k$	capital	Ativos fixos ponderado por um deflator específico da indústria.	log do ativo fixo
$VI$	medida física de integração vertical	Composição de 4 medidas quantitativas de integração vertical	Não-utilizado, substituído pela <i>Dummy</i> Integração
$s$	medida de poder de mercado	Participação de cada empresa nas vendas da indústria. Proxy considerando o total da amostra como total da indústria	Participação de cada empresa nas vendas da indústria. Proxy considerando o total da amostra como total da indústria
$b$	medida de alavancagem	Total de empréstimos e dívida a serem pagos em um ano sobre total do capital do acionista mais impostos diferidos menos goodwill	Total das dívidas sobre capital total dos acionistas
$p$	medida de competitividade relativa	Medida que representa o quanto a margem de lucro da empresa se desvia de todas as outras margens de lucro das empresas. Média do somatório do diferencial das margens de lucro	Diferença entre a margem líquida da empresa menos a margem líquida média da indústria
$u$	índice de fatores cíclicos	Índice de capacidade utilizada de refino (carga processada sobre capacidade nominal de refino).	Índice de capacidade utilizada de refino (carga processada sobre capacidade nominal de refino)
$t$	período de tempo	Período de tempo, utilizado como tendência.	Período de tempo, utilizado como tendência
<i>Dummy_Integração</i>	variável dummy de integração	Não-utilizado	Medida qualitativa de integração: 1, se empresa integrada ou semi-integrada; 0, caso contrário
<i>Dummy_Estatal</i>	variável dummy de propriedade estatal	Não-utilizado	Medida qualitativa de propriedade da empresa: 1, se empresa com capital relevante estatal (NOC's); 0, caso contrário

Fonte: Elaborado pelo autor



No intuito de observar se diferentes amostras resultam em resultados semelhantes, estabelecemos 10 amostras de dados, sendo eles, os seguintes:

Amostra 1 – onde são considerados os dados de todas as 339 empresas;

Amostra 2 – considera-se apenas as empresas privadas, ou seja, que não sejam consideradas como tendo relevante participação acionária estatal, resultando em 295 empresas;

Amostra 3 – englobam as empresas com relevante participação acionária estatal, de acordo com a EVALUATE ENERGY que nesse trabalho consiste em 44 empresas;

Na amostra completa estão consideradas empresas muito pequenas, que foram excluídas das amostras seguintes, visto que elas são pouco significativas na produção da indústria como um todo. Foram, portanto, excluídas todas as empresas com atividades em *upstream* que produziam menos de 10 mil barris por dia de petróleo como média de seus dados disponíveis, bem como empresas com uma capacidade de refino menor que 40 mil barris de petróleo por dia como média dos anos disponíveis. Com isso, as amostras ficaram assim;

Amostra 4 – composta de 225 empresas, sendo a amostra total ao se excluírem as empresas com as restrições acima;

Amostra 5 – sendo as empresas privadas da Amostra 6 resultando em 181 empresas.

Amostra 6 – são as empresas estatais, sendo 44 empresas.

Um resumo das amostras podem ser observadas na tabela 8.

**Tabela 8** - Amostra de Dados

Amostra	Tipo	Nº de empresas	Amostra	Tipo	Nº de empresas
1	Total	339	4	Total s/micro	225
2	Privadas	295	5	Privadas s/micro	181
3	Estatais	44	6	Estatais s/micro	44

Fonte: Elaboração própria, com base no Evaluate Energy

A exclusão das empresas com pouca representatividade na produção de petróleo pode ser confirmada pela tabela 9, onde observa-se que ao se desconsiderar essas empresas há uma redução na produção de petróleo de apenas 11 mil bbl/dia em 1989, crescendo até 447 mil bbl/dia em 2012, respondendo por menos de 1% da produção da amostra total.

**Tabela 9** - Diferenças entre a Amostra Total e Reduzida (Produção de Petróleo - mil bbl/dia)

	1989	1990	1995	2000	2005	2010	2012
Amostra Total	15.047	16.229	22.916	37.235	48.588	51.537	47.156
Amostra Reduzida	15.036	16.207	22.862	37.184	48.481	51.211	46.709
Diferença	11	22	54	52	107	326	447

Fonte: Elaboração própria, com base no Evaluate Energy

Vemos que para as amostras que consideram apenas as empresas estatais não ocorrem mudanças quando se consideram as restrições às pequenas empresas, portanto para a regressão desconsideraremos a Amostra 6.

Também com o objetivo de obter conclusões à respeito da rentabilidade das empresas, que não se encontra em Barrera-Rey (1995), utilizamos um modelo alternativo onde a variável dependente utilizada é o ROA (return on assets – retorno sobre ativos), que é uma medida que reflete a rentabilidade da firma. Esse modelo tem por objetivo encontrar quais são as variáveis que mais impactam positiva (ou negativamente) nos retornos das empresas de petróleo. O modelo com a descrição das variáveis segue abaixo:

$$ROA_{it} = c + \gamma y_{it} + \alpha l_{it} + \beta k_{it} + \phi_1 s_{it} + \phi_2 b_{it} + \phi_3 p_{it} + \rho u_t + \omega t_t + \phi_1 Dummy\_Integração_{it} + \phi_2 Dummy\_Estatal_{it} + \epsilon_{it} \quad (5)$$

Nessa equação as alterações em relação à equação 3 são: a variável dependente que agora é o ROA que é o retorno sobre os ativos totais (representando o desempenho econômico-financeiro das empresas); a variável de vendas que não está defasada. O ROA é calculado com sendo o EBIT dividido pelo total dos ativos fixo (média anual). Esse índice também foi obtido através do banco de dados da EVALUATE ENERGY.

A primeira das variáveis independentes presente são as vendas das empresas ( $y_{it}$  - log das vendas deflacionadas), refletindo a importância das vendas no alcance de bom desempenho econômico-financeiro. A principal razão em incluir essa variável se dá por que acredita-se que uma empresa com elevadas vendas (e com crescimento das vendas) tende a alcançar melhores resultados financeiros, considerando que receitas maiores (resultado de maior produção) podem se refletir em melhores resultados.

A variável trabalho está expressa pelo número de empregados da firma, enquanto que a variável capital são os ativos fixos das empresas (em milhões de dólares). Ambas variáveis são expressas em log.

A variável de poder de mercado (s) foi calculada como sendo a participação da empresa (vendas) em relação ao total da indústria (vendas). Como explicado anteriormente, a amostra corresponde a um painel não-balanceado, além de não considerar a totalidade da indústria, a medida é calculada em termos relativos como uma aproximação considerando-se a amostra disponível. Considera-se que uma firma com maior participação tende a obter melhores resultados, graças, por exemplo, as economias de escala e poder de mercado.

A medida de desempenho relativo, representada pela variável  $p$ , se considera como a diferença entre o desempenho da empresa, medida pela margem líquida da empresa, e o desempenho médio global da amostra (margem líquida média de toda a amostra), representando os retornos anormais obtidos por uma determinada empresa em relação à média da indústria, resultando em quanto uma empresa pode ser mais (ou menos) eficiente que a indústria.

A variável alavancagem (b) que no modelo é considerada como o índice definido pela razão entre o total das dívidas (de curto e longo prazos) e o total do capital dos acionistas, se relaciona com a hipótese da “disciplina de capital”, considerada por Jensen (1988), e espera-se que apresente uma correlação positiva com o ROA visto que a disciplina de capital tenderia a resultar em operações mais eficientes resultando em melhores desempenhos financeiros.

Assim como na equação 3, utiliza-se como medida de movimentos cíclicos da indústria, a variável  $u$ , que é o índice de capacidade utilizada do refino, calculada pela razão entre a carga de petróleo processada e a capacidade nominal de refino existente (ambas em mil bbl/dia), para esse cálculo, os dados foram coletados do *BP Statistical Review 2013*.

As variáveis *dummy* consideradas obedecem aos mesmos critérios da equação 3. As variáveis *dummy* são utilizadas para se verificar se uma empresa obtém um desempenho melhor se tem atuação integrada, e/ou se são empresas estatais. Via de regra considera-se que uma empresa integrada apresenta melhores resultados que uma não-integrada. Também é considerado como norma que uma empresa privada alcança melhores resultados que uma companhia estatal visto que a primeira não sofre de ingerências políticas.