

Não faz sentido ser exato em relação a algo se você
nem mesmo sabe o contexto do que esta falando
J. Von Neumann

Índice

I – Apresentação	1
II – Histórico	2
III – Metodologia	5
III. 1- A análise dos problemas.....	5
III. 2 – Uma metodologia de análise econométrica.....	8
IV – Aplicação.....	12
V – Bibliografia.....	27
Anexos (Planilha de dados).....	28

I - Apresentação

Investigando o uso da econometria encontra-se muita critica e pouca adesão. Em muito do segundo caso observa-se, também, uma fé cega, o que vem a diminuir ainda mais o adequado uso dos métodos e técnicas da econometria. No que diz respeito ao primeiro caso – a crítica severa, percebe-se a dificuldade na aceitação, por um lado, por ignorar-se a eficácia dos exercícios econométricos, e, por outro lado, pela complexidade da teoria estatística e da matemática que lhe dá suporte, que exige, principalmente, conhecimento de álgebra matricial e cálculo diferencial.

Tendo as mesmas dificuldades acima apontadas, uso este breve ensaio de forma a procurar o melhor possível esclarecer minhas dúvidas, dividindo o trabalho em três partes. A primeira buscando um breve histórico do nascimento e desenvolvimento da econometria; a segunda apontando uma metodologia e análise econométrica; e, em terceiro lugar, aplicando o método econométrico em estudo de caso sobre a demanda de álcool combustível no período de 85/89. São Três campos: o histórico, o metodológico e o aplicado.

Histórico:

A aplicação da análise econométrica possui varias dificuldades: o problema metodológico da especificação das variáveis independentes, da forma do modelo, da utilização adequada da estatística como método de análise, e, principalmente, a questão de que “(...) antes de responder o que é econométrica, deve-se colocar a questão “o que é economia”?

JOHNSTON [pág. 1]

Do ponto de vista da história do pensamento econômico, a econometria é considerada como resultado de uma evolução metodológica dentro da economia desde Adam Smith até nossos dias. A genealogia mais antiga encontrada foi apontada por WRIGHT [pág. 79], que considera *Von Thunen* como o formulador do primeiro modelo econométrico. A obra clássica deste autor - O ESTADO ISOLADO (1826), procura explicar a influência da distancia do mercado consumidor sobre os sistemas de cultivo da terra:

“(…) Thunen distingue-se pelo emprego da matemática na obtenção de valores precisos para conceitos econômicos teóricos, por ele desenvolvidos com rigor e clareza excepcionais. (...) Para ganhar experiência prática morou em uma propriedade rural, em 1799, e fez uma serie de visitas técnicas em 1804. Em 1810, na remota região de Mecklenburg, comprou a fazenda Tellow (458 ha). E iniciou os experimentos e medições essenciais à elaboração de *O Estado Isolado*.

Para KIRSTEN [pág.76] “(...) com o impacto da teoria de Cournot (1818) e dos famosos princípios de Marshall (1890), associados às contribuições de Jevons e Walras, houve quase uma separação entre Economia Pura, perfeitamente científica e a Economia Aplicada, que comportaria considerações de tempo e espaço, refletidas pelos dados numéricos de observação”.

E referindo-se especificamente a econometria observa que,

“(…) simultaneamente a essa evolução metodológica observada na economia, constataram-se algumas tentativas de alteração completa do método de estudo dos fenômenos econômicos. A primeira foi a Escola histórica (seg. metade do séc. XIX), cujo principal fundamento era o de que se poderia construir uma Ciência Econômica sem ser necessário apelar para hipóteses *a priori* (uma teoria sem medida). A segunda delas é representada pela Estatística Econômica, e que se colocou em evidencia nas primeiras 3 décadas do século XX; constituía –se na aplicação da análise estatística às séries de dados econômicos se recorrer à Teoria Econômica; era, pois, uma medida sem teoria.

Ressalte-se esta segunda posição, que levou, nos EUA do início do século, a uma radicalização, com a criação dos famosos *barômetros* de Harvard:

“(…) O grupo de Harvard rejeitou deliberadamente todas as teorias econômicas, afirmando que as mesmas são construções esquemáticas, não apoiadas em material empírico, sendo portanto inúteis para as pesquisas de flutuação da conjuntura. Somente as pesquisas empíricas podem proporcionar resultados que, no futuro, poderiam constituir uma base para a generalização teóricas” LANGE [pág. 24].

Com esta abordagem foram generalizadas as análises baseadas exclusivamente em medidas empíricas, que se mostraram, em curto espaço de tempo, pouco eficientes como instrumentos de previsão, bastando citar a imprevista “crise” que se iniciou em 1929. De acordo com este autor, as pesquisas relacionadas com a previsão das flutuações da conjuntura econômica deixaram muitas conquistas valiosas. Graças a elas, foram

elaborados métodos estatísticos para o isolamento das flutuações sazonais, tendências e flutuações cíclicas. Alguns desses métodos entraram para sempre na técnica estatística moderna.

A Econometria, segundo KIRSTEN [pág. 76], “utilizando os instrumentos da matemática, da estatística, e da Teoria Econômica, representa a “medida com teoria e a teoria com medida”. Isto é de suma importância, pois deve ser destacado com rigor a separação entre teoria Econômica e a Estatística”.

A Estatística vem a ser o elo de ligação entre a Teoria Econômica e a realidade. Nestes termos não se constitui a Estatística uma ciência propriamente dita, mas sim em um método de análise, procurando pelo uso da matemática, a determinação dos dados que representam certa realidade.

O estudo do método estatístico é feito didaticamente pela Estatística Matemática e a estatística Aplicada. A segunda é responsável pela aplicação dos processos de análise estatística ao tratamento dos dados de um determinado campo específico. Entretanto em cada campo de aplicação, os fatos a ela pertinentes estão sujeitos a determinadas estruturas características e que são responsáveis pelos mesmos. “(...) Essas estruturas poderão ou não teoricamente ser ignoradas; ignorá-las seria tentar construções puramente empíricas, que ficam condicionadas à manutenção do estado de regularidade existente nas observações estatísticas. Nesse sentido a aplicação do método estatístico na análise dos dados numéricos de observação pode ser feita – segundo orientações “medida sem teoria e medida com teoria”. KIRSTEN [pág. 78].

A “medida com teoria” toma como ponto de partida a existência de uma teoria explicativa, racionalmente elaborada, para o mecanismo a que os dados se referem, e o procedimento metodológico seria:

- considerar os dados estatísticos que se pretende analisar;
- considerar uma teoria (um conjunto de hipóteses) que se proponha a explicar o mecanismo responsável pelos dados estatísticos observado, e;
- verificar, por meio de processos estatísticos adequados, a diferença entre o comportamento apresentado pelos dados observados e o comportamento que esses mesmo deveriam apresentar se a teoria fosse adequada para descrevê-los.

Assim os objetivos da Econometria seriam:

- Efetuar medidas das variáveis e de agregados econômicos,
- Estimar parâmetros pertencentes às relações construídas pela Teoria Econômica;
- Formular hipóteses a respeito do comportamento da realidade;
- Submeter a prova, com base na realidade, de teorias fornecidas pela economia;
- Construir novas teorias (conjunto de hipóteses).

O método de investigação econométrica é, portanto, dedutivo e indutivo ao mesmo tempo, constituindo –se, mesmo, na aplicação das regras da lógica à economia. O

que se conhece por econometria, portanto, pode ser definido como aplicação das técnicas da estatística e da matemática aplicada à análise dos problemas econômicos.

“(...) portanto, parece recomendável que não se deve tentar formar econométricos, e sim economistas que conhecem adequadamente técnicas econométricas” WRIGHT [pág.80].

III – Metodologia

Dividimos este item em duas partes: a análise do problema e sugerimos um método de abordagem na segunda parte.

No primeiro caso – a análise do problema, ressaltamos que antes de tudo o estudo econométrico tem que ter uma justificativa baseada em uma análise de custo X benefício. A construção de um modelo tem, geralmente, alto custo. A utilização da abordagem econométrica tem de ter forte justificativa, principalmente pela organização que pretende utilizar os resultados. Isto direciona os esforços e somente uma real necessidade justifica a abordagem.

III.1 – Análise do problema

No que diz respeito ao problema em si, várias dificuldades surgem. JOHNSTON [pág.3] sintetiza da seguinte maneira:

“Antes de tudo o “(...) comportamento dos construtores de modelo é muito pragmático. Tudo é relativo: tudo depende do problema à mão (...). Várias questões não são respondidas de forma definitiva:

- Forma funcional: Questões teóricas sozinhas usualmente não podem especificar a forma funcional.
- Definição de dados e medição. A teoria é algumas vezes precisa e algumas vezes não.
- Estrutura defasada: Novamente, a teoria econômica não pode especificar sobre apropriadas estruturas defasadas...Na prática o mundo sai sempre de um desequilíbrio para outro; então os dados atuais refletem ajustamento de processo ao invés de posições de equilíbrio.
- Implicações quantitativas x qualitativas: A dificuldade da obtenção e adequação dos dados.
- Escolha entre teorias. Provavelmente existem tantas teorias quanto teóricos. Alguns modelos teóricos diferem em grau, mas não em espécie.

A econometria se envolve com todas as cinco questões. Sua tarefa básica é “colocar carne e sangue *empíricos* nas estruturas teóricas”. Isto envolve vários passos cruciais. Primeiro deles, a teoria ou modelo deve ser explicitada com a forma funcional, portanto, começa com a forma funcional mais simples consistente com a especificação *a priori*.

As dificuldades se iniciam com a caracterização do problema e a própria disponibilidade de um grande elenco de técnicas. LANGE [pág.102] chama a atenção para dois exemplos básicos: da caracterização da função demanda e da função oferta.

Características da demanda – Uma simples média móvel, simples “alisamento” exponencial, ou técnicas de regressão causais, normalmente, não serão adequados quando o problema tratado apresenta fortes sazonalidades. De outro lado, um grande número de métodos temporais como alisamento exponencial sazonal, método de Fourier ou Box-Jenkins, poderiam ser usados. Este exemplo ilustra a fundamental importância do conhecimento da natureza do que se pretende analisar.

Outro problema é o de se estimar uma função demanda usando –se dados de vendas (Vendas = demanda + estoques). As implicações são óbvias. Se usarmos dados históricos para projetar demanda, a previsão será quase sempre subestimada.

Um segundo problema comum é a necessidade de adequação dos dados. O número de automóveis à álcool, entre 85 a 90, cresceu a uma taxa anual de 20%, enquanto o consumo médio é relativo e caiu no mesmo período. Outro problema é a inflação e a necessidade de adequação através de métodos de deflação.

As colocações acima também mostram que só é possível interpretar devidamente as estatísticas de mercado, e determinar a curva de procura ou de oferta, a base dessas estatísticas, quando essa curva de procura, ou de oferta, não seja sujeita a modificações de maior importância no período em estudo. Se a curva de procura e de oferta variarem simultaneamente, será impossível, por meio de métodos acima descrito, deslindar os dados estatísticos contido nas séries temporais de variação nos preços de equilíbrio e na procura.

(O problema de identificação).

Vários exemplos poderiam ser citados mostrando como devemos ser cautelosos em pesquisas desse tipo. Pode ser relativamente fácil imaginar um processo estocástico apropriado, mas certas premissas econômicas de natureza teórica são necessárias, a fim de que se possam tirar conclusões adequadas. Tais premissas derivam, usualmente, não do material estatístico de que dispomos, porém de uma alguma fonte adicional, como por exemplo, de informações históricas quanto à ocorrência de ciclos econômicos, revoluções tecnológicas, variações da renda nacional, etc. Portanto, nem mesmo as séries temporais dão-nos um quadro completo da situação do mercado, e somente a análise técnica baseada em fontes adicionais de informações nos podem salvaguardar de conclusões errôneas.

E no caso da curva de oferta? LANGE [pág.103] teve a idéia de aproveitar para esse fim, os mesmos dados estatísticos que o levaram às curvas de procura. Entretanto, o preço do produto no ano dado deveria ser confrontado, não com a oferta desse mesmo ano, mas com a oferta do ano seguinte. Com efeito, a decisão dos agricultores, quanto ao aumento ou redução eventual de sua produção, por exemplo, de cereais, é tomada sob a influência do preço vigente no ano dado, enquanto que o efeito da decisão só se concretiza no ano seguinte. Este é exemplo significativo da utilização de *defasagem* como meio de interpretar o problema.

O gráfico que ilustra a correlação entre os preços de um ano dado e a produção desse mesmo ano foi interpretada no estudo acima com função procura, e o gráfico representado pelo preço num ano dado com a produção do ano seguinte como função oferta. A primeira dessas curvas era decrescente enquanto a segunda era ascendente.

O método em questão é o correto apenas para produtos agrícolas; para produtos industriais não nos parece possível, pois que o processo de variação na produção industrial se dá de maneira inteiramente diferente.

Conseqüentemente o processo proposto por LANGE para determinar as curvas de procura e de oferta requer em cada caso um controle econômico. Não basta estudar o

material estatístico. Além desse é necessário empregar investigações auxiliares não contidas nas series estatísticas relativa à produção e preço.

Existem outras discussões de suma importância e vale citar aqui o caso entre variáveis aleatórias. Em regressão o *regressando* é sempre tratado como aleatório, enquanto os regressores podem ser ou não, de acordo com as circunstâncias. Cito o exemplo da agricultura, onde é feita a regressão da produção (aleatória) com a utilização de adubos (variável controlada e não aleatória). Se estivermos tratando com variáveis não experimentais (caso da renda familiar, por exemplo), a regressão será tratada inteiramente como aleatória.

II . 2 – Uma metodologia de análise econométrica

BUFFA [pág. 181] propõe os seguintes passos na preparação da projeção com o uso de modelos de regressão (causais ou *cross section*):

- 1- Especificar as variáveis independentes;
- 2- Especificar a forma do modelo;
- 3- Estimar os parâmetros do modelo;
- 4- Testar modelo;
- 5- Testar as premissas dos “mínimos quadrados”;
- 6- Estimar valores para as variáveis.

1- Especificar as variáveis independentes

De maneira a diminuir a ambigüidade, MADDALA [pág.90] propõe jamais usar os termos independentes e dependentes, mas sim, na ordem, variáveis de causa (ou variável explanatória) e variáveis de efeito, ou mais especificamente *regressor e regressando*.

Neste passo são feitas hipóteses sobre os tipos de relações causais entre variáveis dependentes e independentes. Para BUFFA [pág. 190] este passo na formulação de modelos de projeção é mais uma arte que uma ciência, desde que requer entendimento do ambiente o qual afeta a variável dependente a ser projetada (o entendimento a *priori*, já comentado). Neste ponto, o projetista deve também querer especificar as variáveis independentes as quais estão relacionadas com o problema. Daí a existência de especificações múltiplas.

2 – Especificar a forma do modelo

Questões teóricas sozinhas usualmente não podem especificar a forma funcional. Isto é um típico exemplo do fato que restrições qualitativas derivadas da teoria econômica não servem para delimitar as formas funcionais mais perto.

Por exemplo, pode ser do tipo $Y=A.B^x$, onde A e B são constantes. Se este for o caso, deve-se fazer uma transformação de maneira a converter equações não-lineares em lineares. Uma maneira de verificar esta necessidade é preparar um gráfico plotando a variável dependente contra as independentes, caso a caso.

3 – Estimar os parâmetros

A discussão é ampla e vários métodos são disponíveis. O método dos mínimos quadrados é o mais utilizado; inclusive quando se trata de modelos não lineares.

A equação dos mínimos quadrados estabelece a fórmula matemática da linha reta que passa através dos dados da “melhor” forma possível. O “melhor” no caos é dado pela linha que minimiza a soma dos quadrados soa desvios dos pontos em relação a reta estimada.

4- Testar o modelo

Para aplicar o modelo neste estágio, requer –se responder duas questões: a) quão confiável é a projeção com o uso deste modelo? b) o que se pode falar a respeito da significância do modelo? Existe grande numero de teste estatístico aptos a responder estas questões; três deles são de particular importâncias.

a – coeficiente de determinação R^2 - Estabelece a proporção da variação da variável dependente que é explicada pelo modelo.

b) – erro padrão da estimativa – Este dado indica o campo esperado de variações que encontraremos se usarmos o modelo para fazer uma projeção.

c) – teste de significância – É possível que o resultado ocorreu por “chance”? É possível que o resultado ocorreu por “chance”? É possível que o atual R^2 é zero, mas estimamos 0,74 por pegar um conjunto de dados “ruins”? por especificar um mal modelo? ou por causa de ser pequena a amostra? O teste significância F indica qual é o caso. Em um exemplo podemos ter o valor $F = 40,4$ e a probabilidade que o modelo não seja significante é de 0,0001. Portanto estamos seguros que o nosso modelo é significante, e explica a variações na variável que estamos projetando.

d) – erros de especificações – A especificação de um modelo de regressão consiste numa formulação da equação de regressão e de proposição ou pressupostos referentes aos regressores e ao termo de perturbação. Um erro de especificação, no sentido amplo do termo, ocorre todas as vezes que a formulação de uma equação de regressão ou um de seus pressupostos forem incorretos. No sentido menos amplo do termo, o erro de especificação só se refere a erros na formulação

da equação de regressão adequada, e esta é a interpretação que aqui adotamos. Consideramos aqui:

- omissão de uma variável independente;
- não consideração de uma mudança qualitativa em uma das variáveis independentes;
- inclusão de uma variável independente não importante;
- forma matemática incorreta da equação de regressão;
- especificação incorreta da maneira pela qual a perturbação entra na equação de regressão.

5 – Testando as premissas do modelo de regressão usando o “Método dos Mínimos Quadrados”.

As premissas são:

- a variável dependente é linearmente relacionada a independentes, ou as variáveis foram transformadas.
- a variância dos erros é constante, onde os erros são definidos como a diferença entre os valores atuais da variável dependente e as estimativas produzidas pelo modelo.
- os erros que ocorrem entre um caso e outro são independentes entre si? Em outras palavras, o erro produzido pela diferença entre o valor atual e o predito, é independente dos erros produzidos entre a atual predição e outras instancias?
- os erros são normalmente distribuídos.

Normalmente o primeiro é fácil de verificar se cuidado foi feito no *plotting* e se o tamanho da amostra e se o tamanho da amostra e grande. O problema é no segundo e terceiro caso. No segundo tem-se o caso da heterocedasticidade e terceiro da autocorrelação. Uma maneira fácil é plotar os erros residuais correspondentes aos valores da variável dependente. Se o gráfico aparece com uma dispersão aleatória dos pontos, então as chances são de que nem um problema existe. Se, entretanto, os pontos no gráfico mostram qualquer relação, isto é evidencia de que heterocedasticidade existem; daí nosso modelo, e todos os testes específicos, são inválidos.

Um segundo teste eo DURBIN WATSON (Ver :MADDALA [pag.89]. Se a estatística DW for maior que 2.5 e menor que 1.5 geralmente existem problemas.

Se os teste propostos provarem que não estamos seguros das premissas dos mínimos quadrados MQO. e melhor voltar ao com~ ou colocar hipótese adicionais ou variáveis independentes alternativas. fazer algumas transformações ou mesmo. formular um modelo alternativo.

O que é heterocedasticidade?

Um dos pressupostos básicos da utilização do método dos mínimos quadrados ordinários é a de que a variância das variáveis explicativas X_i tem de ser constante; isto garante a propriedade de *homocedasticidade*. Caso contrário (variância não constante), implica a característica de *heterocedasticidade*. Isto é fácil de se observar ao estimarmos, por exemplo, a demanda em função da renda; quanto maior a renda, obviamente, haverá maior dispersão dos dados (baixa renda tem pouca flexibilidade de consumir, portanto com pequena variância; alta renda possui mais flexibilidade entre poupar e consumir - alta variância).

A maneira de superar esta dificuldade, é a utilização do método dos mínimos quadrados generalizados. Técnica que busca corrigir, justamente, esta questão.

O que é autocorrelação?

Outro pressuposto do método dos mínimos quadrados ordinários é a independência dos distúrbios estocásticos. O modelo generalista é o seguinte:

$$Y = \beta \cdot X + U_i$$

Y é o vetor $[y_1, y_2, \dots, y_n]$

X é a matriz $\begin{matrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{matrix}$

$$U_i = [u_1, u_2, \dots, u_n]$$

β é o vetor $[\beta_1, \beta_2, \beta_3, \dots, \beta_n]$

O fenômeno da autocorrelação ocorre quando os U_i 's possuem correlação.

O que é multicolinearidade?

É o fenômeno onde as variáveis explicativas X_i são de alguma maneira correlacionadas. É o terceiro fenômeno importante que traz distorções na análise de regressão, devendo de alguma forma ser superado. Um entendimento “geométrico” seria o seguinte. Imagine-se uma regressão com uma variável dependente (demanda, por exemplo) e duas variáveis “explicativas” (renda e produto interno bruto, por exemplo); caso haja correlação entre renda e pib, o “plano” que passa pelos pontos no espaço formado pelas duas variáveis não possui estabilidade; qualquer informação adicional pode modificar completamente a inclinação deste plano em relação a quaisquer das duas variáveis. Também é fenômeno que exige correção.

6 -Estimar valores para as variáveis independentes e projetar

Devemos, para finalizar, estimar valores para as variáveis independentes. Este é um dos dilemas dos modelos de projeção. Entretanto é uma potencia da técnica, já que se pode usar análise "*what-if*". Dai uma das fraquezas do *approach* e uma de suas maiores forças.

IV – Aplicação (Estudo de caso)

O problema analisado: Análise de consumo do álcool combustível no Brasil-1985/89

Com todas as dificuldades que o problema apresenta, a maior de todas é a forte intervenção que tem o governo, principalmente na formação dos preços dos combustíveis, aí incluído, o álcool etílico hidratado carburante (AEHC). Apesar dessas dificuldades, apresentamos a seguir os resultados do exercício proposto.

Utilizaremos a metodologia de análise econométrica proposta no capítulo III por BUFFA.

1 -Especificar as variáveis

variável dependente (*regressando*): Por não conseguirmos os estoques de passagem referentes ao estudo em questão, estabelecemos como meta o estudo do consumo ao invés da demanda uma medida indireta.

CC - Consumo médio de álcool combustível no período de Jan'85 a Jun'89 - Dividimos o consumo mensal de álcool etílico (anidro + hidratado) pela frota de veículos a álcool, descontado o sucateamento. Dados do CNP -Conselho Nacional do Petróleo ANFAVEA - Associação Nacional dos fabricantes automotores; CENAL -Conselho Nacional do Álcool. Curva de sucateamento dada pelo CENAL em estudo do DENATRAN.

variáveis independentes (*regressores*)

S - Salário real mensal, dado pela deflação do índice de salário médio da ANDIB - Associação Nacional das Industrias de Base -o defletor utilizado foi o IGP-DI da FGV.

P - Preço do álcool na bomba -com o IVV (Imposto sobre vendas a varejo); o preço foi deflacionado para Jan'85 = 100, pelo IGP-DI da FGV, e ponderado pelos dias de vigência.

CD -Consumo médio defasado do CC, por um período

T - foi considerada a variável tempo, para testar a melhoria tecnológica dos veículos a álcool combustível.

2 -Especificar a forma do modelo;

Utilizamos especificação de dois modelos:

i.) EV- modelo de elasticidade variável: $CC = \beta_1 + \beta_2 \cdot CD = \beta_3 \cdot S + \beta_4 \cdot P + \beta_5 \cdot T + u_i$

(Observar que os parâmetros β_1 , β_2 , β_4 e β_5 não são coeficientes de elasticidades)

i.i.) EC -modelo de elasticidade constante: $CC = \beta_1 \cdot CD^{\beta_2} \cdot S^{\beta_3} \cdot P^{\beta_4} \cdot T^{\beta_5} \cdot u_i$ (1)
onde $u_i \Rightarrow$ distúrbio estocástico.

(Observar que, neste caso, os parâmetros β_2 , β_3 , β_4 e β_5 são coeficientes de elasticidade)

No caso do modelo de elasticidade constante, usamos o método dos mínimos quadrados -MMQ, na forma transformada:

$$\ln CC = \ln \beta_1 + \beta_2 \cdot \ln CD + \beta_3 \cdot \ln S + \beta_4 \cdot \ln P + \beta_5 \cdot \ln T + \ln u_i \quad (2)$$

Ressalte-se, portanto) que os resíduos u_i entram na equação original (1), de forma multiplicativa, a partir das regressões efetuadas em (2). Ver MADDALA [pag.88], e que os parâmetros da reta, β_2 , β_3 , β_4 e β_5 são as próprias elasticidades

O gráfico 1. abaixo. mostra o diagrama de dispersão entre consumo medio (l/veic/mês) e o preço do etanol carburante.

CONS. MÉIO X PREÇOS AEC

GRÁFICO

Parece haver uma correlação negativa; abaixo fazemos a regressão entre consumo e preço.

O gráfico 2, abaixo, mostra o diagrama de dispersão entre o consumo médio mensal (1/ veic / mês) e o salário real S.

CONS. MÉDIO X SALÁRIO REAL

É evidente a correlação positiva entre salário e consumo médio (curva de Engel)

O gráfico 3, abaixo, amostra o diagrama de dispersão em escala logarítima entre o consumo médio (1/veic/mês) e o preço real (Jan'85=100)

CONS.MÉDIO X PREÇO

LN PREÇO

Aqui se vê, a grande dispersão entre LN (Preço) e LN (Consumo), mostrando haver baixa correlação entre as duas variáveis.

O gráfico 4, abaixo mostra o diagrama de dispersão em escala logaritmica do consumo médio mensal (1/veic/mês) e o salário real S.

Repetindo o caso de elasticidade - variável, o gráfico acima mostra a significativa correlação entre consumo médio e preço.

3 -Estimar os Parâmetros dos modelos

A tabela abaixo representa as combinações de regressões efetuadas, apontando o grau de explicação (corrigido), as variáveis independente consideradas e o t calculado para o específico parâmetro β_k . (Todas as regressões. calculo de parâmetros e testes estatísticos foram feitos com o *software* TSP .

Modelo de elasticidade constante -EC

Eq. N°	R2 (%)	Consumo defasado 1/v/mês	Salário Real (índice)	Preço Real J85=100	Tempo	Parâmetros
1	71,9	0,634 6,48	0,191 5,587	-0,04 -1,304	-0,012 -1,978	DW=1,834 F=34,079 F*4,48=2,6
2	71,5	0,635 6,445	0,193 2,590	-	-0,0116 -1,8650	DW=1,767 F=44,24 F*3,49=2,8
3	70,1	0,696 7,276	0,1597 2,146	-0,036 -1,115	-	DW=1,76 F=41,66 F*3,49=2,8
4	68,1	0,798 9,97	-	-	-0,008 -1,256	DW=1,75 F= 56,53 F*2,50=3,2
5	67,7	0,822 10,5	-	-	-	DW=1,735 F=110,23 F*1,51=4,0
6	48,2	-	0,50 6,48	-	-0,02 -3,06	DW=0,98 F=25,18 F*3,49=2,8
7	39,7	-	0,494 5,93	-	-	DW=0,811 F=35,21 F*1,51=4,0
8	0,1	-	-	-0,04 -0,681	-	DW=0,35 F= 0,465 F*1,51=4,0

A tabela 2, abaixo, mostra as combinações de regressões para o modelo de elasticidade variável -EV.

Eq. N°	R2 (%)	Consumo defasado 1/v/mês	Salário Real (índice)	Preço Real J85=100	Tempo	Parâmetros
1	73,9	0,59 5,878	0,358 2,703	-0,131 -1,586	-0,2 -2,29	DW=1,77 F=37,84 F*4,48=2,6
2	73,1	0,597 5,86	0,385 2,66	-	-0,18 -2,05	DW=1,68 F=44,82 F*3,49=2,8
3	70,1	0,696 7,276	0,1597 2,146	-0,036 -1,115	-	DW=1,76 F=41,66 F*3,49=2,8
4	69,8	0,761 8,87	-	-	-0,153 -1,66	DW=1,69 F= 61,19 F*2,50=3,2
5	68,8	0,828 10,75	-	-	-	DW=1,70 F=115,66 F*1,51=4,0
6	55,2	-	0,90 6,50	-	-0,407 -4,0	DW=0,96 F=33 F*2,50=3,2
7	42	-	1,026 6,21	-	-	DW=0,78 F=38,59 F*1,51=4,0
8	0,1	-	-	-0,09 -0,57	-	DW=0,34 F= 0,32 F*1,51=4,0

Listamos os resultados em ordem decrescente do grau de explicação R2 F => F calculado

Para as variáveis explicativas (independentes) são apontados os parâmetros e, abaixo, o valor da estatística t de Student. Também são apontados o valor do teste de Durbin-Watson e da estatística F.

$$F * = F \text{ tabelado} \Rightarrow F_{k-1, n-k} (\alpha=5\%)$$

Resultados:

.verificamos a inexistência de relação entre o preço real do combustível e o consumo médio por veículo (confirmado pelos testes abaixo descritos);

.A escolha entre os modelos de elasticidade-constante (EC) e elasticidade variável (EV) devem ser estudadas frente aos testes de verificação de erros de especificação, no caso o teste de verificar a forma matemática incorreta na equação de regressão.

.Ressalte-se as baixas elasticidades encontradas. No caso do modelo de elasticidade- variável -EV. e necessário calcular a elasticidade no ponto médio, através da fórmula ($\eta = X/Y$); fazendo-se os cálculos, os resultados mostram as baixas elasticidades encontradas. Portanto, o consumo médio é pouco elástico em relação às variáveis consideradas.

.Fazendo-se o teste de melhoramento da regressão, vê-se que para a inclusão da variável preço (da eq.2 para a eq.1), aceitamos a hipótese nula (Ver teste 4.3.) A inclusão da var. preço real não melhora a regressão (apesar do pequeno crescimento do R2 corrigido).

.No caso da inclusão do salário real (da eq. 4 para a eq. 2), há melhoria da regressão, com contribuição significativa da inclusão da variável salário real -S.

.Também há melhoria quando se inclui o consumo médio por veículo defasado (da eq.6 para a eq. 2).

.Adotamos para os exercícios sobre testes,que virão a seguir, o modelo expresso pela equação 2 -EV, onde os três parâmetros β_2 , β_3 e β_4 , individualmente, apresentam hipótese nula rejeitada $\beta_k = 0$. O teste F (reconhecimento da regressão) também mostra hipótese nula rejeitada, mostrando haver regressão, conforme veremos a seguir..

Adaptação da equação estimada

O gráfico abaixo mostra a adaptação da equação estimada versus o gráfico observado, utilizando-se a equação 2, modelo de elasticidade variável -EV.

FIGURA

4 - Testar o modelo

4.1.) O teste t e o teste F :

O teste t e usado para os parâmetros estimados individualmente. O teste indica a precisão dos estimadores de cada um dos coeficientes da regressão considerados separadamente.

Se desejarmos testar a hipótese nula, no caso de uma regressão simples:

$H_0: \beta_1 = 0$ (ou $\beta = k$) contra a hipótese alternativa $H_a: \beta_1 \neq 0$ (ou $\beta_1 \neq k$).

a estatística $\frac{\beta_k - 0}{s_{\beta_k}} \sim t(n-k)$ (t de Student com (n-k) graus de liberdade)

onde n - tamanho da amostra e k - no parâmetros

4.2.) Teste para reconhecimento da regressão

Uma hipótese mais ampla e a de que nenhuma das variáveis independentes tem influência sobre a media de Y_i : portanto desejamos testar a hipótese nula:

$H_0: \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$

Se H_0 for verdadeira, a variação de Y, de observação para observação, não será afetada por mudanças em qualquer das variáveis independentes, mas será puramente aleatória.(...) o valor observado de SQR (Soma dos quadrados devidos a regressão) só diferirá de zero por causa da amostragem.

A estatística:

$\frac{SQR(k-1)}{SQR(n-k)} \sim F(k-1, n-k)$ (S de Snedecor, com k-J graus de liberdade no numerador

SQR(n-k) e n-k graus de liberdade no denominador)

Note-se que se qualquer dos coeficientes da regressão estimada $\beta_2, \beta_3, \beta_4$ e β_5 for significativamente diferente de zero, conforme o teste t, então o valor de F será também significativamente diferente de zero, contanto que os testes sejam efetuados ao mesmo nível de significância e contra a mesma alternativa.

"Mas e possível acharmos nenhum dos β 's significativamente diferentes conforme o teste t, e ao mesmo tempo, rejeitarmos a H_0 de que $\beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$ pelo teste F. Isso poderia acontecer no caso de as variáveis independentes estarem altamente correlacionadas entre si. Em tal situação, as influências separadas de cada uma das variáveis independentes sobre y pode ser fraca, enquanto sua influência conjunta pode ser forte". KMENTA [423].

Nas tabelas acima, apresentamos o F calculado e o F tabelado = F^* , mostrando diretamente o resultado deste teste em cada uma das simulações.

4.3.) Teste para melhoramento da regressões

Um teste um tanto diferente refere-se a influencia das variáveis adicionais sobre a média de Y_i . Dada a regressão:

$$Y_i = \beta_1 + \beta_2 .X_{i2} + \beta_3 .X_{i3} + \dots + \beta_k .X_{ik} + u_i$$

A adição de variáveis a esta regressão proporciona melhor resultado?

$$Y_i = \beta_1 + \beta_2 .X_{i2} + \beta_3 .X_{i3} + \dots + \beta_k .X_{ik} + \beta_{k+1} .X_{i,k+1} + \dots + \beta_q .X_{iq} + u_i$$

O teste é o seguinte:

$$\frac{(SQR_g - SQR_k) / (g - k)}{(SQE_q) / (n - q)} \sim F (q - k, n - q)$$

Usando o R^2 temos:

$$\frac{[R^2 - q - R^2 - k]}{1 - R^2_q} \quad \frac{[n - q]}{q - k} \quad (*)$$

Fazemos esses testes para verificar melhoria quando da inclusão de variáveis no Modelo EV.

1 e 2 -testar a inclusão da variável preço - P:

F calc = 0.674 (conforme calculo 4.3.(*))

$F^* (1,48) = 4.08 \Rightarrow$ aceito H_0 .

A inclusão do preço não contribui para a melhoria da regressão.

4 e 2 - testar a inclusão da variável Salário real -S:

F calc = 5.846

$F (1,50) = 4.08 \Rightarrow$ rejeito H_0

O S contribui para a melhoria da regressão

8 e 2 -testar a inclusão da variável Consumo médio/veic. defasado -CD:

F calc = 48,18

$F^* (1, 50) = 4.08 \Rightarrow$ rejeito H_0

O CD contribui para a melhoria da regressão.

5 -Testar as premissas do Modelo

Dada a escolha da equação 2 - modelo EV. fazemos abaixo testes de multicolinearidade, heterocedasticidade e autocorrelação dos distúrbios naquela equação.

5.1) Multicolinearidade:

Pelo pressuposto de ausência de multicolinearidade, exige-se que nenhuma das variáveis independentes esteja perfeitamente correlacionada com qualquer outra variável independente ou com qualquer combinação linear de variáveis independentes. KMENTA [43 7]

a multicolinearidade é uma questão de grau e não de natureza. A distinção significativa não esta entre a presença ou a ausência de multicolinearidade, mas entre seus vários graus.

como a multicolinearidade se refere a condições de variáveis independentes que se presume não- estocásticas, e uma característica da amostra e não da população.

Portanto, “não se faz teste de multicolinearidade”, mas pode-se, medir o seu grau em qualquer amostra particular.

Várias são as possibilidades de se detectar multicolinearidade. Uma delas e "Se regredimos cada uma das variáveis independentes sobre todas as variáveis independentes remanescentes, podemos obter uma medida do "grau de ajustamento" calculando o valor de R2 em cada caso. Se qualquer destes R2's estiver próximo da unidade, o grau de multicolinearidade será alto. KMENTA [448]

Fizemos as seguintes regressões auxiliares:

i.) CD x S R2 (corrigido) = 0,37 o β (S) = 0,98 com t = 5,69 F calc = 32,32 F* = 4,08

i.i.) CD x T R2 (corrigido) = 0,21 o β (T) = -0,505 com t = (-3,81); Fcalc = 14,48 F.=4, 08.

i.i.i.) S x T R2 (corrigido) = 0,03 o β (T) = -0,146 com t = (-1,56); F calc = 2,43 F. = 4, 08.

i.v.) CD = f(S,T) R2(corrigido) = 0,49; β (S) = 0,86 (t=7,34); β (T) = -0,38 (t=-3,4); F = 25, 73

v.) S = f(CD,T) R2 (corrigido) = 0,37 ; β (CD) = 0,49 (t=5,38); β (T) = 0,07 (t=-0,82); F = 16,39

v.i.) $T = f(S, CD)$ R^2 (corrigido) = 0,26; $\beta(S) = 0,19$ ($t = 4,53$); $\beta(CD) = -0,51$ ($t = -3,48$); $F = 3,05$.

Sobre a conclusão, continuando com KMENTA [448] : "Dado que alguma multicolinearidade quase sempre existe, a questão é saber em que "altura" o grau de multicolinearidade deixa de ser normal e torna-se "prejudicial". Esta questão não tem sido resolvida satisfatoriamente. Conforme um critério às vezes usado na prática, a multicolinearidade é considerada como prejudicial se, digamos, no nível de significância de 5%. o valor da estatística F for significativamente diferente de zero mas nenhuma das estatísticas t para os coeficientes da regressão (diversos da constante de regressão) o for. Neste caso rejeitaríamos hipótese de que não há relação entre Y_i de um lado em X_{i2} , X_{i3} e X_{i4} de outro, mas não rejeitaríamos a hipótese de que alguma das variáveis independentes é irrelevante em influenciar Y_i . Tal situação indica que a influência separada de cada uma das variáveis independentes é fraca relativamente a sua influência conjunta sobre Y_i . Isto é sintomático de um alto grau de multicolinearidade.

Pela discussão acima, diagnosticamos alguma grau de multicolinearidade: vêm-se as regressões i.) e i.v) com alto R^2 {corrigido}.

5.2) Heterocedasticidade

Uma possibilidade de análise é a visual, observando-se o gráfico dos resíduos: (Forma de funil?):

A partir da suposição de existência de heterocedasticidade, procede-se o cálculo do TESTE DE GOLDFELD & QUANDT:

H_0 : os resíduos u_i 's são HOMOCEDASTICOS

H_a : os resíduos U_i 's são HETEROCEDASTICOS

$n = 53$

Fazemos um corte de 5 observação centrais para possibilitar a execuções de 2 regressões separadamente:

Obtém-se para a primeira regressão (as primeiras 24 observações)

$$\sum u_i^2 = 1696$$

e, para a segunda regressão (as últimas 24 observações):

$$\sum u_i^2 = 972,61$$

Teste de Godfeld & Quandt:

$$\sum u_i^2 / \sum u_i^2 = F(22,22) = 972,1/1696 = 0,6$$

$$F^*(22,22) = 2,10 \text{ (5\% de nível de significância)}$$

Portanto, aceitamos a H0: os u_i 's são HOMOCEDASTICOS

5.3) Auto-correlação

Aplica-se mais particularmente a séries temporais -MADDALA [274])

Pelo pressuposto do modelo linear clássico, temos: KMENTA [317]

$$E(u_i u_j) = 0 \text{ para todo } i \neq j$$

Como se pressupões que a media de u_i e u_j seja zero, podemos escrever:

$$\text{Cov}(u_i, u_j) = E[u_i - E(u_j)] = E(u_i, u_j) = 0$$

Esta característica das perturbações de regressão é conhecida como não-auto-regressão, alguns autores referem-se a ela como a ausência da correlação serial. Isto implica que a perturbação que ocorre num ponto da observação não está correlacionada com qualquer outra perturbação. Isto significa que, quando as observações são feitas ao longo do tempo, o efeito da perturbação que ocorre, num período, não afeta o período seguinte. Por exemplo, num estudo da relação entre insumo e produto de uma firma ou industria feito através de observações mensais, a não-auto-regressão da perturbação implica que o efeito das quebras de máquinas é estritamente temporário, no sentido de que só a produção mensal corrente é afetada. No caso de observações seccionais, como as referentes a rendas e gastos de diferentes famílias, o pressuposto de não-auto-regressão significa que se o comportamento relativo aos gastos de uma família for "perturbado"- por exemplo, pela visita de um parente - isto não afeta o comportamento relativo a gastos, de outras famílias.

Um teste de ausência de auto-regressão "(...) que tem sido amplamente utilizado em aplicações econométricas, e conhecido como *teste de Durbin Watson*. (...) Devemos notar, de passagem, que o teste não é aplicável a equações de regressão mas quais o lugar da variável independente é tornado pelo valor defasado da variável dependente". KMENTA A [346]. Isto invalida o teste para o nosso caso, onde usamos CD como variável defasada (um período) para CC (repressor).

V -BIBLIOGRAFIA

- BUFFA, Elwood S. -Production / Inventory Systems -Planning and Control -Chap.7
Production and Inventory Control Theory -Illinois -Richard D.Irwin Inc. 1979.
- JOHNSTON, J -Econometric methods McGraw - Hill -1990.
- KIRSTEN, J. T. -Manual de Economia -Org. Diva Benevides Pinho -Cap.4 Metodologia
quantitativa na Pesquisa Econômica. Ed. Saraiva / 2a.ed. 1992.
- KMENTA. Jan -Elementos de Econometria /Vol 2 -ATLAS 1990.
- LANGE, Oskar -Introdução a econometria -Ed. Fundo de Cultura -1961.
- MANDALA, G.S. -Econometrics -McGraw-Hill- 1987.
- WRIGHT, Charles -Método Econométrico: Algumas reflexões sobre a obra pioneira de
Van Thunenn; in Revista de Econometria -Ano II / Numero 2 -Nov'82 -Editado pela
Sociedade brasileira de econometria.

Dados utilizados

	C	S	P	InC	InS	InP
Mês	Cons med	Ind sal real	Preço real			
0	Móvel	(Jan85 =100)	(Jan85 =100)			
Jan / 85	216,32	100	100	5,376758793	4,605170186	4,605170186
Fev/ 85	216,32	102,04	91,62	5,376758793	4,625364893	4,517649588
Mar/ 85	220,51	92,71	92,71	5,396215105	4,529476342	4,62771415
Abr/ 85	210,98	110,4	95,41	5,351763342	4,704110134	4,558183395
Mai /85	229,78	118,68	88,51	5,437122329	4,776430795	4,48311554
Jun/85	238,74	113,74	82,1	5,475375094	4,733915142	4,407938016
Jul/85	250,02	116,5	79,33	5,521540915	4,757891273	4,373616367
Ago/85	241,12	112,88	75,26	5,485294735	4,726325308	4,320948785
Set/85	243,24	105,57	74,79	5,49404861	4,65937424	4,314684186
Out/85	234,53	114,22	76,61	5,457583511	4,738126413	4,338727617
Nov/85	257,67	119,91	72,94	5,551679696	4,786741461	4,289637185
Dez/85	260,77	109,54	72,34	5,563638793	4,696289779	4,281377227
Jan/86	244,65	110,51	70,47	5,499828618	4,705106015	4,255187087
Fev/86	225,01	109,1	68,47	5,416144846	4,692264893	4,226395693
Mar/86	218,77	123,02	70,03	5,38802095	4,812346944	4,248923722
Abr/86	237,15	125,35	70,44	5,468692852	4,831109825	4,254761284
Mai/86	244,37	126,67	70,22	5,49868347	4,841585279	4,251633171
Jun/86	252,87	126,63	69,85	5,532875523	4,841269448	4,246350086
Jul/86	249,44	130,09	73,76	5,519218405	4,868226519	4,300816579
Ago/86	256,88	132,72	87,5	5,54860905	4,888241646	4,471638793
Set/86	256,54	133,18	86,56	5,547284598	4,891701597	4,460837815
Out/86	258,52	134,65	85,36	5,554997306	4,902678819	4,446877607
Nov/86	264,15	139,91	100,03	5,576517124	4,940999359	4,605470141
Dez/86	256,29	131,35	124,11	5,546309616	4,877865516	4,821168269

Jan/87	248,74	127,22	110,81	5,516408174	4,845917871	4,707817023
Fev/87	227,74	120,74	98,21	5,428204627	4,793639473	4,587108043
Mar/87	222,73	126,08	97,75	5,405960276	4,836916626	4,582413199
Abr/87	219,74	124,85	95,63	5,392445029	4,827113017	4,560486578
Mai/87	218,49	115,74	102,86	5,386740247	4,751346296	4,63336884
Jun/87	215,52	109,87	103,59	5,373053713	4,699297849	4,6404408
Jul/87	218,38	104,38	105,87	5,386236665	4,648038086	4,662211926
Ago/87	225,76	102,16	101,31	5,419472488	4,626540212	4,618185123
Set/87	223,44	103,31	100,34	5,409142922	4,637734177	4,608564419
Out/87	218,09	101,42	99,82	5,384907822	4,61927031	4,603368564
Nov/87	217,23	113,9	98,93	5,3809567	4,73532087	4,594412529
Dez/87	212,85	109,1	100,73	5,360587692	4,692264893	4,61244367
Jan/88	212,03	102,77	95,37	5,356727774	4,632493482	4,557764064
Fev/88	216,68	94,72	93,32	5,378421611	4,550925171	4,536034447
Mar/88	216,12	103,98	89,8	5,375833809	4,644198573	4,497584975
Abr/88	220,05	113,19	84,46	5,393854793	4,729067823	4,43627805
Mai/88	217,21	114,25	84,75	5,380864628	4,73838903	4,439705746
Jun/88	228,15	112,75	85,86	5,430003307	4,725172978	4,452718063
Jul/88	217,54	109,52	85,47	5,382382742	4,696107181	4,448165437
Ago/88	212,91	110,27	86,2	5,360869541	4,702931904	4,456670178
Set/88	201,22	110,48	83,16	5,304398837	4,704834509	4,420766463
Out/88	221,67	108,11	77,38	5,401189789	4,683149227	4,348728349
Nov/88	223,27	117,78	69,11	5,408381801	4,768818478	4,235699438
Dez/88	218,98	117,39	65,61	5,388980401	4,765501725	4,183728123
Jan/89	202,76	107,9	70,61	5,312023014	4,681204872	4,257171778
Fev/89	208,14	101,47	70,02	5,33821093	4,619763188	4,248780916
Mar/89	221,5	104,95	67,18	5,400422589	4,653484046	4,207375584
Abr/89	229,52	116,57	63,88	5,435990172	4,758491951	4,157006323
Mai/89	226,47	117,62	68,03	5,422612486	4,767459089	4,219948784
Jun/89	227,11	105,25	61,93	5,425434482	4,656338473	4,126004715
Jul/89	0					