**LISTA ECONOMETRIA II – Prof. Daniel dos Santos**

Você possui um banco de dados de crianças de 6 a 17 anos, com as seguintes informações:

- Indicador (dummy) de matrícula na escola: Y

- Indicador de recebimento de benefício do Bolsa Família por algum membro da família: T

- Variáveis de localização geográfica do domicílio (unidade da federação em que se localiza, indicador de que o domicílio está em área urbana): **L**

- Variáveis de composição demográfica da família (presença da mãe e do pai no domicílio, número de irmãos): **F**

- Características individuais (sexo, raça, idade): **P**

- Variáveis de inserção dos pais no mercado de trabalho (se estão ocupados, se são trabalhadores formais, se estão desempregados): **M**

- Status socioeconômico da família (educação da mãe e do pai, renda familiar per capita): **S**

1. A. Discuta as limitações existentes em um modelo que relacione linearmente Y com seus determinantes.

B. Formule um modelo de utilidades aleatórias em que as famílias comparem custos e benefícios de matricular seus filhos na escola e decidam de forma ótima, expressa no indicador Y, que revela a preferência dos agentes por determinada ação. Para cada ação, considere que os agentes tomam como dado T, **L, F, P, M, S** na hora de decidir se matriculam seus filhos.

C. Suponha que a parte não observável do benefício de matricular os filhos, ε, tenha distribuição normal-padrão e seja independente das características observáveis. Qual a expressão para Pr (Y = 1 | T, **L, F, P, M,S**) implicada pelo modelo descrito na pergunta anterior?

D. Descreva como faria para obter o efeito que receber o programa bolsa-família tem sobre a probabilidade de frequência à escola para um indivíduo médio da população (isto é, aquele que possui $L=\overbar{L};G=\overbar{F};P=\overbar{P};M=\overbar{M};S=\overbar{S}$).

E. Descreva como faria para obter o efeito médio produzido pelo bolsa-família se cada indivíduo de sua amostra passasse da situação de não-recipiente para recipiente. Sabendo que apenas famílias de baixo status socioeconômico (**S=S0)** constituem o público-alvo do programa, como você estimaria o efeito esperado do programa sobre os elegíveis?

1. Críticos dirão que as famílias decidem conjuntamente matricular ou não seus filhos e trabalhar ou não, de modo que **M** seria endógeno.
2. Reformule o modelo descrito na parte B da questão anterior para ter como variável dependente uma decisão policotômica da família (matricular e trabalhar, matricular e não trabalhar, não matricular e trabalhar, e não matricular e não trabalhar). Se facilitar sua explicação, trabalhe com a hipótese de que os benefícios líquidos não observáveis das diferentes alternativas se distribuem conjuntamente segundo uma logística.
3. Como você estimaria, neste caso, o efeito esperado do programa bolsa-família sobre a taxa de matrícula das famílias pertencentes ao público-alvo do programa?
4. Como sua resposta aos itens A e B mudaria se você agora supusesse que as decisões da família de trabalhar e de matricular os filhos na escola ocorram sequencialmente?
5. Como pesquisador perspicaz, você sabe que o programa bolsa-família não apenas exige como contrapartida que as crianças entre 6 e 17 anos estejam matriculadas na escola, como também que frequentam ao menos 80% das aulas.
6. Suponha que em sua base de dados conste informação sobre o número de dias em que a criança foi à escola na última semana. Proponha um modelo de contagem em que o número de dias frequentados seja sua variável dependente.
7. Como você computaria o efeito que o bolsa-família teve sobre a probabilidade de que uma criança média da população vá à escola ao menos 80% das vezes?
8. Escreva a função de verossimilhança do modelo proposto no item A
9. Reescreva a função de verossimilhança considerando que nos dados há uma inflação de zeros resultante do fato de que uma fração α dos indivíduos entre 6 e 17 anos não estão matriculadas na escola
10. Nosso objetivo é estimar a demanda por ensino superior privado no país. Se facilitar sua formulação, suponha que o custo de oportunidade do tempo é nulo (ou seja, todos os indivíduos em sua amostra gostariam de continuar estudando). A heterogeneidade (não-observável) dos indivíduos é dada pela qualidade do ensino recebido, ε, que por sua vez é proporcional ao salário (w) futuramente recebido. Considere que existe na economia um serviço público gratuito homogêneo de qualidade Q, e que você observe em seus dados o nível de endividamento estudantil, Di, além da escolha realizada por cada sujeito, Yi. Os agentes são maximizadores de renda líquida, entendida como a diferença entre salário e remuneração da dívida contraída (w – rD).
	1. Construa o problema de decisão entre ir para uma universidade pública ou privada, mostrando como estima-lo.
	2. Como você estimaria a taxa de juros que cada agente paga sobre sua dívida?
	3. O governo agora anuncia um programa chamado Pró-uni 2, onde indivíduos que moram na região z = “N” terão metade de suas dívidas perdoadas ao final do curso. Como isto afeta a proporção dos moradores de “N” que optam por uma universidade pública? Como isso afeta a proporção total de estudantes matriculados em universidades privadas?
	4. Você consegue propor uma forma de a partir deste exercício estimar o quanto os alunos das universidades públicas estariam dispostos a pagar pelos seus respectivos cursos?

Em suas respostas, tome o cuidado de explicitar as hipóteses e cálculos feitos.

1. Suponha agora que você possui uma amostra fornecida pelo sindicato das universidades particulares, contendo informações sobre as trajetórias profissionais dos egressos de universidades desse tipo no Brasil. Na base de dados há ainda informações detalhadas sobre salários dos ex-alunos, e indicadores de qualidade do ensino recebido durante o ensino superior, tais como a proporção de professores com título de doutorado, existência ou não de biblioteca, e tamanho médio das turmas.
	1. Como você estimaria a importância de cada um destes itens na determinação da qualidade do ensino?
	2. Como você testaria a hipótese de que a existência de biblioteca na universidade é tão importante quanto ter 100% (ao invés de nenhum) de professores com doutorado?
	3. Como sua resposta ao item anterior mudaria se você também tivesse acesso a informações sobre indicadores de qualidade do ensino (mas não de salários) de egressos do ensino público?
2. Em uma base de dados você tem informações sobre a escolaridade de um grupo de indivíduos (a partir da qual você pode construir um indicador T que separe os que decidiram cursar ensino superior dos que optaram por não fazê-lo), características individuais, **X**, e salário obtido, w. Seu objetivo é estimar o ganho salarial associado a frequentar o ensino superior. É sabido que o mercado de trabalho contém ocupações para trabalhadores qualificados (entendidos como aqueles que têm diploma universitário) e não qualificados, e que estas ocupações remuneram as características individuais de modo distinto.
	1. Construa um modelo teórico de vantagens comparativas, que indique como os sujeitos tomam a decisão de frequentar faculdade baseados em seus diferenciais de salário potencial em postos de trabalho qualificados e não-qualificados.
	2. Como você estimaria esse modelo se seu computador não tivesse capacidade de maximizar a função de verossimilhança? Como você estimaria os elementos matriz de covariâncias das variáveis não-observáveis desse modelo?
	3. Se tivesse que informar o impacto da passagem pelo ensino superior sobre o salário, como faria? Este impacto varia de indivíduo para indivíduo? Qual o impacto médio percebido por aqueles que frequentaram o ensino superior? E qual o impacto médio que eventual experiência deste tipo teria sobre aqueles que efetivamente optaram por parar de estudar ao final do ensino médio?
3. Em nosso livro-texto, Wooldridge apresenta os resultados de uma estimação de determinantes do número de filhos por mulher, baseados em duas especificações distintas do modelo econométrico: uma linear e outra de Poisson.



* 1. Discuta as vantagens e desvantagens de cada especificação.
	2. Conforme lê-se na tabela, a variância estimada do resíduo na regressão de Poisson sugere sub-dispersão. Proponha uma forma alternativa de estimar este modelo, onde o problema de sub-dispersão esteja endereçado.
	3. Se houver no modelo inflação de zeros, como este problema poderia ser endereçado? Mostre o modelo estatístico que você estimaria e a respectiva função de verossimilhança.
	4. Como você construiria, a partir de cada uma das estimações acima, uma estimativa do impacto de um ano adicional de escolaridade sobre o número médio de filhos por mulher sobre pessoas com características médias da população, e o impacto de se elevar um ano de escolaridade para todas as pessoas da amostra sobre a média de filhos.
	5. Suponha que a variável número de filhos fosse transformada para as seguintes categorias: “nenhum”, “um a três”, “4 a 6”, “7 ou mais”. Construa um modelo teórico para o qual o número de filhos por mulher possa ser estimado através de um logit ordenado, e mostre como os parâmetros descritos no item (d)