**Universidade de São Paulo Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto Departamento de Economia Programa de Pós-Graduação em Economia - PPGE Econometria**

**Segunda lista de exercícios – Prof. Daniel Santos 2016**

1. 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)
1. Duração
	1. Defina função de sobrevivência.
	2. Suponha que você tenha acesso a uma base de dados coletada em outubro de 2015 por uma agência de empregos, em que no cadastro de um amplo grupo de desempregados seja perguntado a quanto tempo a pessoa se encontra nesta situação. Além do tempo de desemprego, você tem também acesso a características socioeconômicas e demográficas **X** do indivíduo. Descreva como você estimaria a probabilidade de que um indivíduo obtenha emprego em dezembro como função do seu tempo de desemprego e demais características. Em particular, descreva detalhadamente a função de verossimilhança (você é livre para escolher o modelo de duração que achar melhor).
	3. Como você calcularia as chances de que um indivíduo de sua amostra esteja empregado durante todo o mês de dezembro?
2. Seu objetivo é estimar os determinantes do desempenho educacional no Brasil. Mais especificamente, quer descobrir por que alguns alunos têm desempenho superior a outros em testes de matemática. Sua base de dados consiste em um painel com 4 anos consecutivos de notas de matemática, e características socioeconômicas e demográficas **X** dos alunos.
	1. Escreva um modelo empírico que gostaria de estimar para responder se existem diferenças sistemáticas por sexo nas provas e matemática
	2. Descreva uma forma de estimar o modelo acima, enunciando as hipóteses de que precisaria para que seus estimadores fossem consistentes.
	3. Descreva possíveis limitações no estimador proposto acima. Discuta por que sua missão seria mais fácil se o objetivo fosse saber se existem diferenças de desempenho segundo a renda familiar per capita.
	4. Você percebe em seus dados que existe forte correlação entre o desempenho nas provas entre períodos consecutivos de tempo. Imediatamente, surge a suspeita de que quaisquer impactos da renda familiar sobre o desempenho podem estar contaminados pelo fato de que jovens com renda familiar mais alta naturalmente obteriam notas mais altas por terem também uma série de outras características favoráveis em relação aos demais jovens. Explique quais seriam as potenciais causas de uma correlação sistemática entre notas ao longo do tempo. Qual a diferença neste caso entre uma correlação intertemporal induzida por heterogeneidade intrínseca dos indivíduos ou por dependência de estado?
	5. Explique como você faria para detectar se existe dependência de estado (no caso linear e markoviano, representado por um modelo de painel dinâmico com variáveis dependentes defasadas), e estimar o modelo endereçando este problema.
	6. Como sua resposta mudaria se a dependência de estado induzir a uma auto-correlação parcial próxima de 1 entre as notas de matemática ao longo do tempo? Explique sua resposta.
3. Apresente e discuta as características dos estimadores OLS, efeitos ﬁxos, efeitos aleatórios e primeiras-diferenças para dados em painel.
4. Suponha que sua consultoria seja contratada pelo Clube dos Treze para estimar o montante de receitas arrecadadas com vendas de ingressos em função da divisão em que um time atua. Para tal, a entidade lhe disponibiliza uma base de dados referentes ao Campeonato Brasileiro, para os anos de 2000 a 2013, com os seguintes times.
* Avaí
* Grêmio
* Corinthians
* Figueirense
* Flamengo
* Palmeiras
* Portuguesa
* Santo André
* São Paulo
* Ipatinga
* Figueirense
* Sport
* Vitória
* Santos

Com esta estrutura, você pode montar um painel com os times sendo as unidades seccionais e cada uma das temporadas sendo a dimensão temporal. Com base neste procedimento de estimação, discuta quais seriam as características (benefícios e limitações) dos métodos de efeitos aleatórios, de efeitos ﬁxos e de primeiras diferenças. Em sua resposta, procure enfatizar o papel das variações within, between e ﬁrst-diﬀerence, assim como discutir a validade da hipótese E[α|X] = 0 assumida pelo método de efeitos aleatórios.

1. Com base na metodologia de primeiras diferenças para dados em painel, demonstre porque as estimações tendem a ser ineﬁcientes, enfatizando tanto o papel da variação $∆X\_{it}=X\_{it}- X\_{it-1} ,$ como o formato da matriz de variância-covariância do resíduo diferenciado $∆ϵ\_{it}=ϵ\_{it}-ϵ\_{it-1}$
2. Com relação à estimação de painel com variável dependente defasada (painel dinâmico):
	1. Explique porque os estimadores de efeitos aleatórios e de efeitos ﬁxos tendem a ser inconsistentes;
	2. Justiﬁque e explique o procedimento de instrumentalização da variável dependente defasada em um estimador de primeiras diferenças, explicando os métodos de Anderson & Hsiao, de Arellano & Bond e ﬁnalmente, Blundell & Bond. Procure enfatizar tanto a questão da instrumentalização como a questão do método GMM. Explique também o propósito dos testes de Sargan / Hansen e de autocorrelação de primeira e de segunda ordem (m1 e m2) do resíduo diferenciado.
3. Com relação ao modelo de variável dependente binária em painel:
	1. Explique o procedimento de Mundlak (1978) e de Chamberlain (1980) no modelo probit em painel.
	2. Explique o procedimento Fixed eﬀects logit (ou logit condicional).
	3. Explique o procedimento Random eﬀects logit e probit.