

Primeiras evidências sobre os determinantes da duração dos cursos de mestrado em economia no Brasil

Anderson L. Schneider[§]

Leonardo Porto[§]

Fabiana Rocha[□]

RESUMO

O objetivo do artigo é avaliar, a partir da estimação de modelos de duração, quais as características dos alunos de mestrado em Economia que afetam a probabilidade de conclusão dos cursos. Os estimadores de Kaplan-Meyer, Cox estendido e de aceleração de saída apresentaram resultados bastante semelhantes. O estimador de Kaplan-Meyer apontou como estatisticamente significantes o centro, a instituição onde foi feita a graduação, o ano de ingresso e a nota relativa. Os modelos de Cox estendido e de aceleração de saída confirmaram, além do ano de ingresso, a influência da nota relativa no término mais rápido dos cursos. A redução no prazo institucional para conclusão dos cursos de mestrado parece ter tido um efeito positivo. Assim, existe uma indicação de que os centros devem efetivamente aumentar suas exigências tanto em termos de prazo quanto de desempenho. Vale observar que este trabalho constitui uma primeira tentativa de avaliar a questão, uma vez que muitos centros ainda não têm, de forma sistematizada, as informações necessárias para as estimações. Contudo, espera-se que ele estimule os demais centros que não fizeram parte da amostra a organizar as informações sobre seus alunos, de forma que as estimações possam ser refeitas no futuro usando uma amostra mais representativa. O melhor conhecimento dos determinantes da duração dos mestrados permitiria que os diferentes centros de pós-graduação redesenhassem suas políticas no sentido de reduzir o tempo de duração sem que a qualidade dos seus cursos fosse comprometida, permitindo uma melhor alocação dos recursos destinados ao mestrado, que são essencialmente públicos.

Palavras-chave: modelos de duração, mestrado.

ABSTRACT

The purpose of this paper is to determine, using duration models, which characteristics of Economics graduate students affect their probability of concluding their courses. The Kaplan-Meyer estimator pointed out the graduate institution, the undergraduate institution, the entry year, and the relative grade as statistically significant. The Cox model and the parametric model showed that besides the entry year, the relative grade is important for the duration of the Master program. There is, therefore, some indication that it would be important to increase the demand for performance and decrease the length of the courses. It is necessary to observe that this paper constitutes a first attempt to address the question. Many of the graduate institutions do not have organized data on their students, and this way could not be included in the sample. We hope, however, that the results obtained here stimulate these institutions to gather information on their students and that the estimations could be done again in the future using a more representative sample. A better understanding of the subject would allow the different graduate institutions to change their policies, reducing the duration of their courses without losing quality. The ultimate result would be a better allocation of the resources that are mainly public.

Key words: duration models, master's program, college education.

JEL classification: I2.

* Os autores gostariam de agradecer às seguintes pessoas por terem, com presteza e gentileza, fornecido os dados: Mauro Borges Lemos e Maria Cecília S. Neto do CEDEPLAR/UFMG, Marcos de Barros Lisboa e André Vidal da EPGE, Carlos Roberto Azzoni, José Paulo Chahad, Reinaldo Lourenço e Valéria Lourenção do IPE/USP, Márcio Garcia, Humberto Moreira e Denise Udderman da PUC-Rio e Ronald Hillbrecht da UFRGS. Gostariam ainda de agradecer a Paulo Picchetti pelos comentários e sugestões. Quaisquer erros são responsabilidade exclusiva dos autores.

§ Mestrados em Economia IPE/USP.

□ Departamento de Economia, USP

Recebido em dezembro de 2001. Aceito em fevereiro de 2002.

1 Introdução

Nos últimos anos houve uma mudança significativa na academia brasileira. Hoje vários centros têm programas de Mestrado, e os que ainda não instituíram um programa de doutorado têm demonstrado grande interesse em fazê-lo. Esta mudança quantitativa também foi acompanhada por uma mudança qualitativa. Hoje os professores devem possuir pelo menos o título de doutor, é importante que tenham pelo menos alguma experiência internacional e exige-se muito mais em termos de publicação, tanto em boas revistas nacionais quanto, e talvez principalmente, em revistas internacionais. Tudo isso é avaliado pela CAPES, que vem fazendo, sistematicamente, uma avaliação criteriosa dos cursos de Economia no País.

A sistematização e o acompanhamento da produção científica foram realizadas recentemente por Issler (2002). Falta, contudo, um acompanhamento de como a produção científica, digamos intermediária, está ocorrendo na academia brasileira. Por produção científica intermediária deve-se entender a finalização do Mestrado e do Doutorado com a conseqüente defesa de tese. Isto porque, uma das características dos cursos de pós-graduação no Brasil, segundo a própria CAPES, é o elevado tempo de conclusão dos cursos e, por este motivo, uma grande parte do conhecimento demora para ser difundido. É esta lacuna que o presente artigo pretende preencher. Para tanto, procura-se avaliar, usando modelos de duração, quais as características dos alunos de Mestrado que fazem com que eles tenham uma maior probabilidade de concluir com sucesso seus cursos de Mestrado. A amostra utilizada não é grande e corresponde aos cursos de Mestrados da PUC-Rio, IPE-USP, Universidade Federal de Minas Gerais e Universidade Federal do Rio Grande do Sul. A EPGE colocou gentilmente à disposição as informações sobre seus alunos, mas devido a algumas inconsistências estas tiveram que ser retiradas da amostra. Assim sendo, somente 4 dos 22 centros associados à ANPEC são avaliados, pelo fato de a maior parte dos centros não disporem de dados sistematizados necessários às estimações. Espera-se que com o passar do tempo mais centros tenham estas informações para que no futuro o estudo possa ser refeito incluindo todos os centros integrantes da ANPEC.

O trabalho está organizado em três seções além desta introdução. A segunda seção apresenta a amostra utilizada e discute algumas diferenças entre os centros. A terceira apresenta e discute os resultados dos diferentes modelos econométricos usados, quais sejam: o estimador de Kaplan-Meier, o modelo de Cox estendido e o modelo de aceleração da saída. A última seção resume as conclusões e aponta direções de pesquisa futura.

2 Descrição dos dados

Conforme antes indicado, a amostra é composta pelos alunos ingressantes nos cursos de Mestrado em Economia da PUC-Rio, do IPE/USP, da Universidade Federal de Minas Gerais e da Universidade Federal do Rio Grande do Sul no período 1992/99. Como se pode observar na Tabela 1, existe uma clara tendência de aumento do número de ingressos, o que vem indicar o aumento da importância dos cursos de pós-graduação no País.

Tabela 1
Número de Alunos Ingressantes - 1992/99

Ingressantes por ano	Número de alunos	Participação no total (%)
1992	23	8.0
1993	26	9.0
1994	30	10.4
1995	37	12.8
1996	37	12.8
1997	42	14.5
1998	49	17.0
1999	45	15.6
Total	289	100.00

A Tabela 2 traz a descrição das principais características dos alunos ingressantes. Observa-se que existe uma alta taxa de conclusão dos cursos. Do total de alunos ingressantes, 77,9% concluíram o Mestrado. Além disso, a grande maioria (73,7%) dos alunos é do sexo masculino e 41,5% possuíam entre 20 e 23 anos quando iniciaram o curso. Com relação à origem geográfica, 67,1% dos alunos são da região Sudeste, indicando uma forte concentração regional. A maioria dos alunos que faz Mestrado em Economia também fez graduação em Economia (69,9%). Pouco menos da metade (45%) dos alunos acaba cursando o Mestrado na mesma instituição na qual concluíram a graduação. Por fim, a maior parte dos alunos ingressa na pós-graduação tão logo termina a graduação (64,3% dos alunos ingressaram no Mestrado no máximo um ano após o término da graduação) e somente 10% demoram 6 ou mais anos entre a conclusão da graduação e o início do Mestrado.

Tabela 2
Participação Porcentual das Categorias no Número Total de Alunos
Ingressantes no IPE/USP - 1992/99

Categoria	Participação no total de alunos
Naturalidade	
Centro-Oeste	5.9
Estrangeiro	5.5
Nordeste	5.2
Norte	2.1
Sudeste	67.1
Sul	14.2
Graduação	
Instituição Estrangeira	4.8
Instituição Privada	22.8
Instituição Pública	72.3
Cursou graduação no mesmo Centro?	
Não	55.0
Sim	45.0
Migração	
Não migrou de Estado para fazer Mestrado	61.6
Migrou de Estado para fazer Mestrado	38.4
Anos entre fim da graduação e início do curso	
0	10.7
1	53.6
2	14.2
3 a 5	10.4
6 a 10	6.2
mais de 10	3.8
Sexo	
Feminino	26.3
Masculino	73.7
Idade quando ingressou no curso	
20-21 anos	5.9
22-23 anos	35.6
24-26 anos	30.1
27-30 anos	16.6
mais de 30 anos	11.8
Curso de Graduação	
Economia	69.9
Engenharia	4.5
Outros	4.9
Não Informado	20.8
Conclusão do Mestrado	
Sim	77.9
Não	22.2

Finalmente, a Tabela 3 mostra algumas estatísticas comparando algumas características selecionadas dos diferentes centros. A duração média dos cursos é de 44 meses. A PUC-Rio

e a UFMG apresentam uma duração menor que a média (cerca de 40 meses). A PUC-Rio tem também a duração mínima (19 meses), enquanto a UFMG tem a duração máxima (95 meses). A USP conta com o maior número de alunos (quase 50% superior ao dos demais centros).

Tabela 3
Estatísticas-padrão de Variáveis Seleccionadas para os Diferentes Centros

Centro	Duração do curso			Desvio Padrão	Número de alunos
	Média	Máxima	Mínima		
PUC	39.40	82.97	18.83	13.37	60
UFMG	39.72	95.13	23.70	12.99	66
UFRGS	45.59	82.97	23.40	12.70	68
USP	47.41	76.20	27.33	12.30	95
Total	43.56	95.13	18.83	13.21	289

3 Resultados econométricos

3.1 Visão geral

A literatura econômica associada a modelos de duração utilizou-se de técnicas estatísticas inicialmente elaboradas pela engenharia industrial e pelas ciências biomédicas para realizar estudos tais como a vida útil de uma máquina ou o tempo de sobrevivência de um paciente após um transplante. (Kiefer, 1988) Em geral, o conceito central nestes métodos estatísticos não é a probabilidade incondicional da ocorrência de determinado evento, mas sim sua probabilidade condicional (no nosso caso de interesse, por exemplo, a probabilidade do aluno terminar o curso em 30 meses, dado que ele cursou 29 meses).

Existem três categorias de modelos de duração: modelos paramétricos, semiparamétricos e não-paramétricos. Para todas essas três categorias as duas informações fundamentais são dadas pela função risco (*hazard function*), que fornece a probabilidade condicional de o aluno terminar o curso em dado período, dado que ele já o cursou até aquela data, e pela função sobrevivência (*survivor function*), que descreve a probabilidade de a duração ser igual ou superior a t .

É importante observar que a especificação de determinado modelo com base na função risco ou sobrevivência apenas enfatiza probabilidades condicionais, enquanto que

especificações em termos da função distribuição de probabilidade dão ênfase a probabilidades incondicionais. Para qualquer especificação em termos da função risco ou sobrevivência existe uma especificação matemática equivalente em termos de distribuição de probabilidades. Dessa forma, a abordagem de sobrevivência não identifica novos parâmetros. Geralmente, a abordagem escolhida pelo analista dependerá do foco do objeto de estudo em questão.

É comum em modelos de duração que a amostra seja formada por elementos para os quais somente um limite temporal inferior está disponível. Mais precisamente, aparecem na amostra elementos para os quais o tempo até a ocorrência do evento é desconhecido simplesmente porque o evento não é observado até o final do período amostral. Em suma, algumas informações sofrem o problema de “censura”. No nosso caso específico, a censura refere-se aos alunos que ainda não haviam terminado o curso no momento da coleta de informações para formar a amostra. Desta forma, o curso para tais alunos durou no mínimo o que foi observado na amostra, mas não é possível saber a duração efetiva do curso nesses casos. Todos os métodos de estimação desenvolvidos para as funções risco e sobrevivência levam em conta tal tipo de informação, tratando de modo diferenciado aquelas observações que não sofrem o problema de censura.

Com relação à diferença entre as três classes de modelos mencionadas, conforme destacam Menezes-Filho e Picchetti (2000), os modelos paramétricos exigem um grau crescente de hipóteses de especificação e de poder de explicação quando comparados com os modelos não-paramétricos, de modo que é sempre interessante considerar as três categorias como análises complementares. Por esta razão, optou-se por estimar a duração dos cursos de Mestrado usando as três categorias de modelo. A seção 3.2 apresenta as estimativas usando-se o estimador de Kaplan-Meier, a seção 3.3 inclui variáveis explicativas por meio do modelo de Cox e, finalmente, a seção 3.4 mostra os resultados do modelo de aceleração de saída.

Para uma exposição técnica dos diferentes modelos, ver Kiefer (1988) e Greene (1993).

3.2 Estimador de Kaplan-Meier

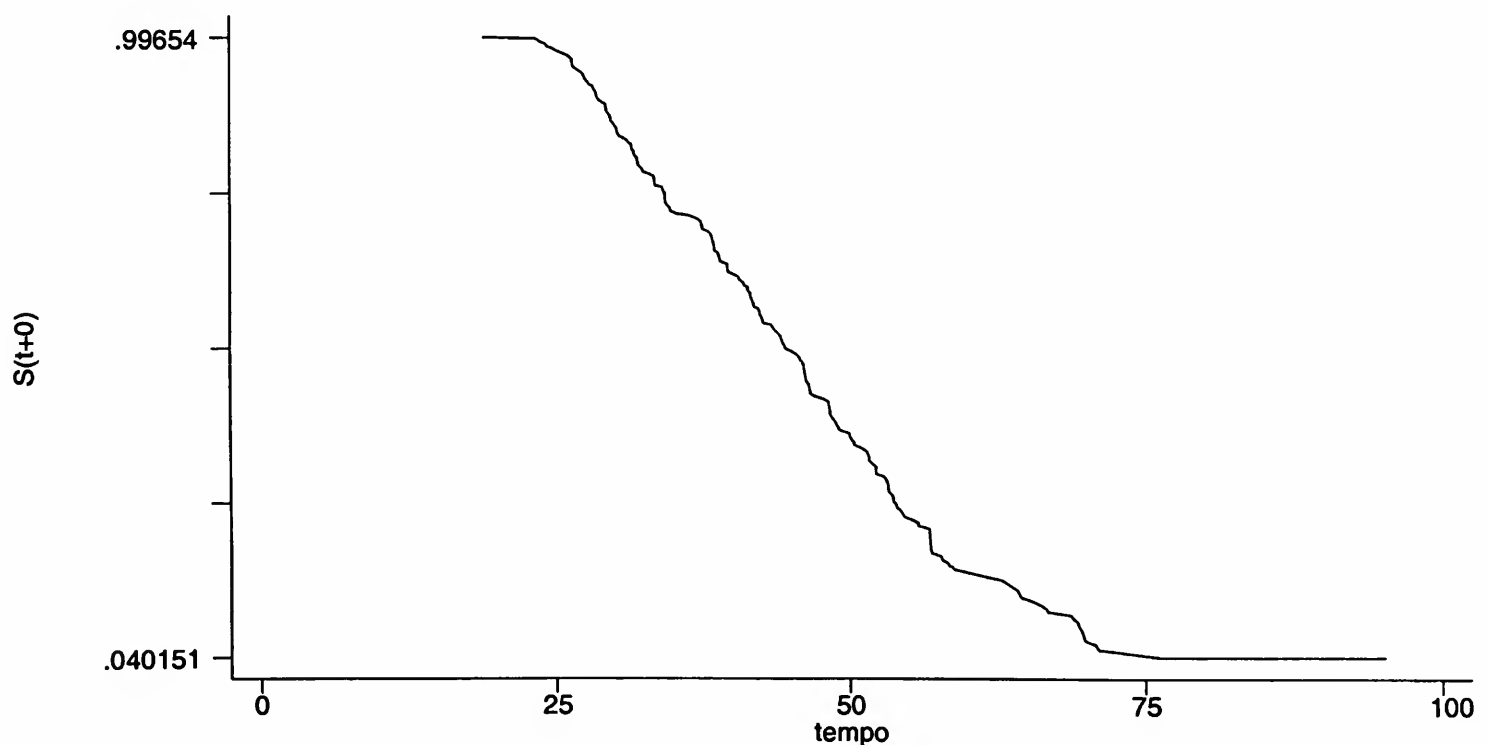
Dentro da classe de modelos não-paramétricos, o estimador de Kaplan-Meier se destaca por apresentar uma abordagem estritamente empírica para estimar as funções risco e sobrevivência, representando uma alternativa para a análise preliminar dos dados.

A Tabela A1, do Apêndice, contém, na primeira coluna, os instantes de tempo nos quais os alunos ou saem da amostra pelo fato de já terem concluído o curso, ou aparecem como tendo duração censurada pelo fato de ainda estarem matriculados no curso. Cabe destacar que, para

cada aluno, o tempo não guarda qualquer relação com o tempo calendário, pois este é mensurado tomando como origem o momento em que aluno iniciou o curso de Mestrado. Assim, por exemplo, dos 289 alunos que compõem amostra, somente 1 havia completado o curso em 18 meses, e assim sucessivamente. A terceira coluna apresenta o número de alunos que completaram o curso. A quarta coluna mostra o número de alunos que são censurados pelo fato de ainda estarem matriculados no curso no momento da coleta dos dados. Desta forma, por exemplo, dos 221 alunos cuja duração no Mestrado é igual ou superior a 34,17 meses, 5 aparecem com duração censurada e, conseqüentemente, registrados como tendo no mínimo 34,17 meses de duração no curso. Na quinta coluna encontra-se a estimativa de Kaplan-Meyer da função sobrevivência, enquanto que as duas últimas colunas mostram os limites do intervalo de confiança, ao nível de 95%, para a função sobrevivência estimada. A duração média do curso de Mestrado para o período 1992-1999 é de cerca de 41 meses.

O Gráfico 1 apresenta a estimativa de Kaplan-Meyer da função sobrevivência, com o eixo das abscissas representando o tempo de duração em meses. Observa-se que a probabilidade de um aluno permanecer no curso se reduz lentamente com o passar do tempo. Depois de 75 meses no curso, a probabilidade de permanecer no Mestrado cai praticamente a zero, enquanto a probabilidade de sobreviver em dois anos (24 anos) de curso é praticamente certa.

Gráfico 1
Estimativa de Kaplan-Meyer da Função Sobrevivência



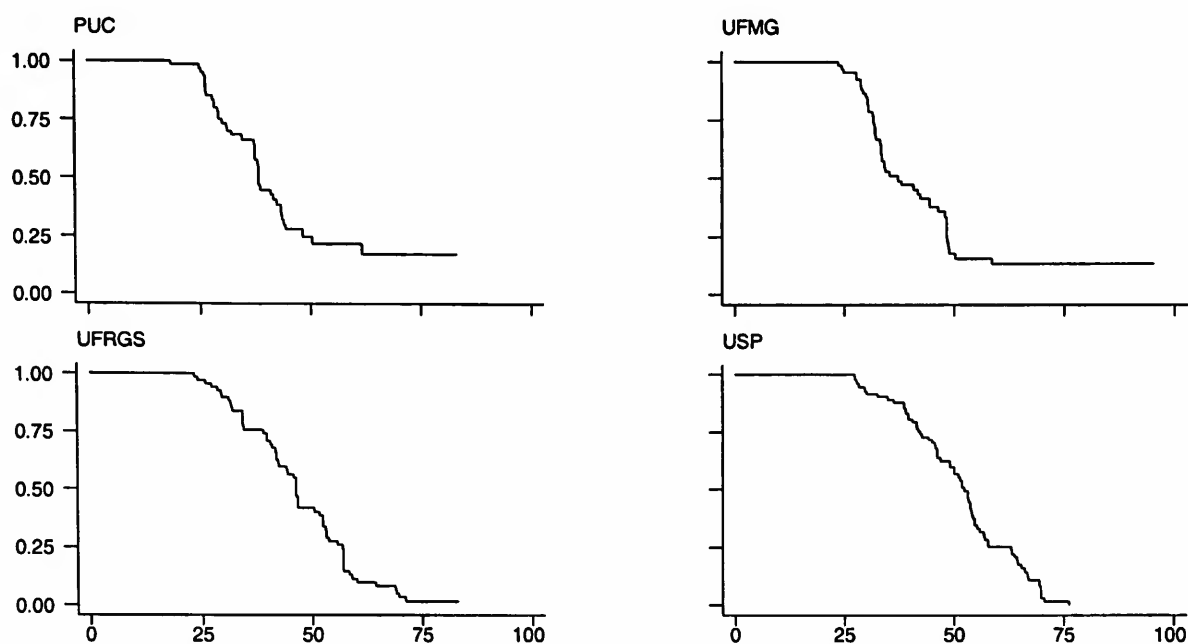
Com o objetivo de testar a igualdade das funções risco entre os grupos aplicou-se o teste de Savage para cada estratificação. A Tabela 4 mostra a síntese dos resultados para as várias estratificações. A hipótese nula de igualdade entre as funções risco não pode ser rejeitada para sexo, idade, naturalidade, para o fato de o aluno ter feito graduação no mesmo centro que está cursando o Mestrado, para migração e para o número de anos entre o final da graduação e o começo do Mestrado. Por outro lado, a hipótese nula de igualdade entre as funções risco pode ser rejeitada para centro, instituição onde foi feito o curso de graduação, ano de ingresso e nota relativa. Com relação a essa última variável, sua construção foi efetuada por meio da divisão em três grupos (percentil 33,33, acima do percentil 33.33 até o percentil 66,66 e acima do percentil 66.66) de alunos de cada centro de acordo com as respectivas notas nas disciplinas cursadas. Tal estratégia se justifica pelo fato de as disciplinas não serem perfeitamente comparáveis entre os diferentes centros. Portanto, o que posteriormente denominaremos de notas “alta”, “média” e “baixa” representam a ordenação do desempenho nas disciplinas cursadas pelos alunos em cada centro.

Tabela 4
Teste para Igualdade das Funções Risco para os Diferentes Grupos

	Estatística χ^2	P-Valor
Sexo	0.07	0.796
Centro	9.09	0.028
Idade	0.93	0.819
Naturalidade	8.15	0.148
Instituição do curso de graduação	4.69	0.096
Graduação no mesmo Centro	0.21	0.649
Ano de ingresso	28.68	0.000
Nota relativa	28.44	0.000
Migração	0.11	0.741
Anos entre final da graduação e início do curso	2.47	0.290

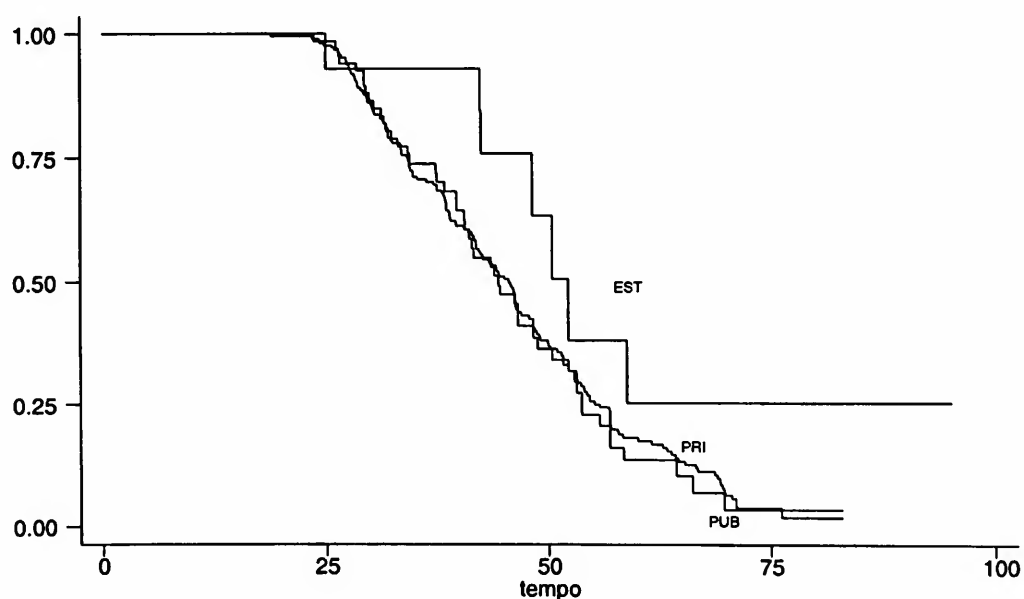
O Gráfico 2 mostra a função sobrevivência de acordo com o Centro. Como se vê, para todos os centros a probabilidade de permanecer no Mestrado é de praticamente 100% nos dois primeiros anos de curso. Contudo, quando o aluno está há 4 anos no Mestrado, a probabilidade de terminar o curso na PUC-Rio e na UFMG é bem menor do que na UFRGS e na USP (menos de 25% contra cerca de 35%). Por outro lado, passados os 4 anos, a probabilidade de permanecer na PUC-Rio e na UFMG é a mesma, enquanto que na UFRGS e na USP tal probabilidade cai lentamente até atingir praticamente zero depois de 75 meses.

Gráfico 2
Estimativa de Kaplan-Meier da Função Sobrevivência Segundo o Centro



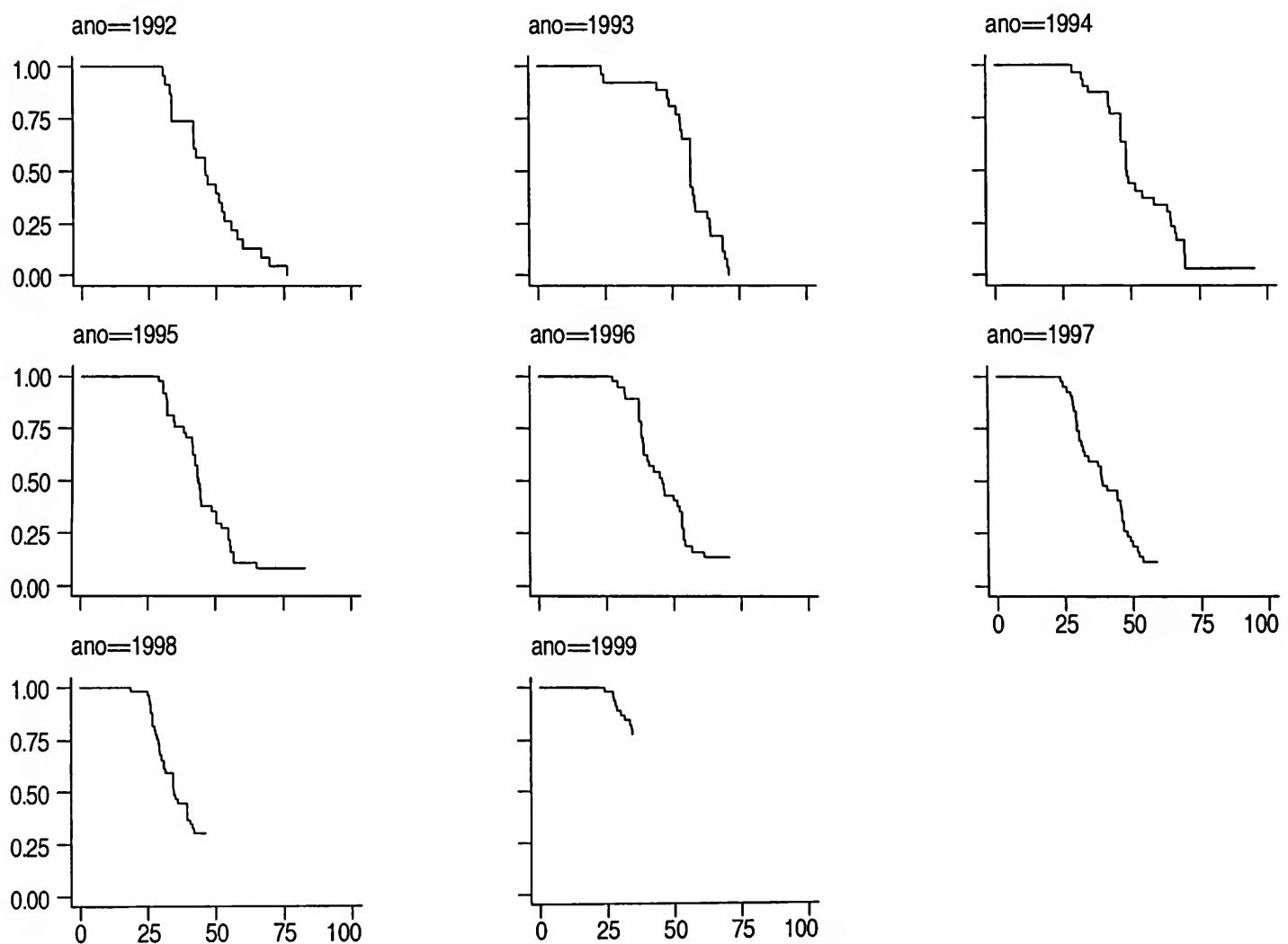
No Gráfico 3 visualiza-se a função sobrevivência segundo o “tipo” de instituição na qual o aluno cursou a graduação (universidade pública, privada ou “estrangeira”, caso o aluno tenha cursado a graduação no exterior). Como se observa, o fato do aluno ter cursado a graduação em universidade privada ou pública não interfere, de maneira significativa, no comportamento da função sobrevivência, ao passo que a realização da graduação no exterior aumenta a probabilidade de este permanecer no Mestrado, comparativamente aos dois grupos anteriores.

Gráfico 3
Estimativa de Kaplan-Meier da Função Sobrevivência Segundo o Tipo de Instituição na Qual o Aluno Cursou a Graduação



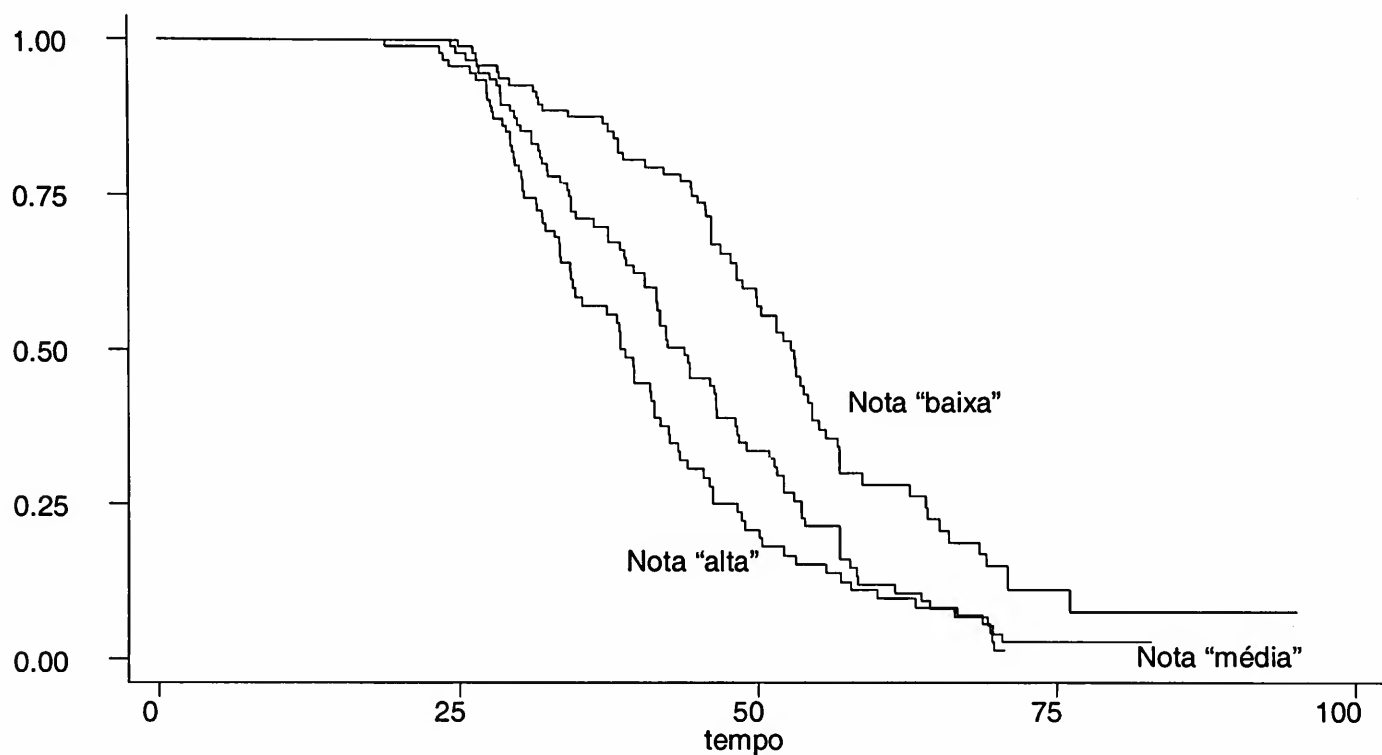
No Gráfico 4 encontram-se as funções sobrevivência estimadas segundo o ano de ingresso dos alunos no Mestrado. Observa-se a ocorrência de comportamentos diferenciados da função sobrevivência, o que vem indicar a possível existência de um efeito coorte entre os alunos. Tal efeito pode ser resultado, entre outros fatores, do ambiente macroeconômico, institucional e acadêmico diferenciado que cada coorte tem ao longo de seu respectivo período de Mestrado. Mais importante ainda parece ser a mudança de regras nos diferentes centros reduzindo o período máximo de conclusão de curso nos anos recentes. Em outros termos, se do ponto de vista institucional é possível permanecer mais tempo no Mestrado, o aluno fica. Caso contrário, ele finaliza o Mestrado no prazo legal mais curto. Assim sendo, parece bastante positiva a discussão de alguns centros no sentido de reduzir ainda mais o tempo máximo para a conclusão de seus cursos. Isto implicaria, simultaneamente, uma disseminação mais rápida do conhecimento e uma economia de recursos.

Gráfico 4
Estimativa de Kaplan-Meier da Função Sobrevivência
Segundo o Ano de Ingresso do Aluno



O Gráfico 5 expõe a função sobrevivência de acordo com a nota relativa do aluno. Como se pode observar, os alunos de melhor desempenho dentro de cada centro possuem uma probabilidade de permanecer no Mestrado inferior aos demais estudantes, enquanto que os alunos de pior desempenho são aqueles que possuem a maior probabilidade de permanecer no Mestrado.

Gráfico 5
Estimativa de Kaplan-Meier da Função Sobrevivência
Segundo a Nota Relativa do Aluno



Como observado anteriormente, este tipo de resultado é importante como uma análise preliminar dos dados. O estimador de Kaplan-Meier, por sua simplicidade e flexibilidade, é útil como ponto de partida quando não se possui uma informação precisa sobre a forma funcional do modelo. Até o presente momento, considerou-se que a função risco depende unicamente do período em que o aluno está no curso. No entanto, pode-se imaginar que uma série de variáveis, tais como o sexo, idade, curso de graduação, tempo entre o fim da graduação e ingresso no Mestrado, entre outras, provavelmente interfiram na taxa instantânea de saída. A inclusão de tais variáveis será feita nas duas próximas seções.

3.3 Modelo de Cox estendido

Diferentemente do modelo de regressão linear clássico, no qual as variáveis explicativas afetam a média da distribuição da variável dependente, o efeito dos regressores no modelo de Cox estendido ocorre por meio da multiplicação da função risco por um escalar, deslocando-a na mesma proporção ao longo de toda a duração do curso de Mestrado, no caso em estudo.

Inicialmente, considerou-se como potenciais variáveis explicativas a naturalidade, o sexo, o ano de entrada, a idade quando o aluno ingressou no curso, o tempo entre o final da graduação e início do Mestrado, a nota relativa, se o aluno migrou ou não para cursar o Mestrado, se o aluno fez ou não a graduação no mesmo centro no qual cursou o Mestrado, se o aluno cursou economia ou outro curso, se a instituição cuja graduação foi cursada é pública, privada ou estrangeira, o centro no qual o aluno cursou o Mestrado e as interações do centro de Mestrado com as demais variáveis. Seguindo a estratégia de Menezes-Filho e Picchetti (2000), contudo, foram eliminadas seqüencialmente as variáveis cujos coeficientes não apresentaram significância estatística. A especificação final é apresentada na Tabela 5. Ao invés dos coeficientes, a tabela expõe a razão de risco para cada variável explicativa. Caso seu valor seja superior a um, a variável em questão desloca a função de risco básico para cima, elevando a probabilidade de o aluno terminar o curso, caso contrário, diminui a referida probabilidade.

Tabela 5
Estimativas do Modelo de Cox Estendido para o Modelo Selecionado

Variável	Razão de Risco	Erro Padrão	Teste "z"	P-valor
Ingresso em 1998	2.461606	0.478224	4.637	0.000
Ingresso em 1993	0.562668	0.104722	-3.09	0.002
Centro_UFRGS	1.551901	0.25905	2.633	0.008
PUC_anos_entre_grad	1.089461	0.016647	5.607	0.000
UFMG_nota_relativa	1.468253	0.116416	4.844	0.000
Nota_relativa	1.528946	0.136958	4.74	0.000

Número de observações: 289

Número de observações não-censuradas: 225

Wald $\chi^2(6)$: 128.01

Log Verossimilhança: -1029.11

O fato de o aluno ter ingressado no curso em 1998, controlado pelas demais variáveis, aumenta a probabilidade de o mesmo terminar o curso com sucesso, enquanto o fato de ele ter ingressado em 1993 diminui esta probabilidade.

Para a UFRGS, o centro de origem do candidato aumenta a probabilidade de término. Para a PUC, quanto mais tempo passar entre a graduação e a entrada no Mestrado, maior é a probabilidade de término. Em todos os centros, a nota relativa é uma variável fundamental, com os alunos de melhor desempenho possuindo uma probabilidade mais elevada de terminar o curso com sucesso.

Os Gráficos 7, 8, 9 e 10, a seguir, mostram as funções sobrevivência para um aluno típico da amostra para as diferentes variáveis explicativas selecionadas no modelo final.

Gráfico 7

Estimativa do Modelo de Cox Estendido para a Função Sobrevivência de um Aluno Típico da Amostra Ingressante nos Anos de 1993 e 1998

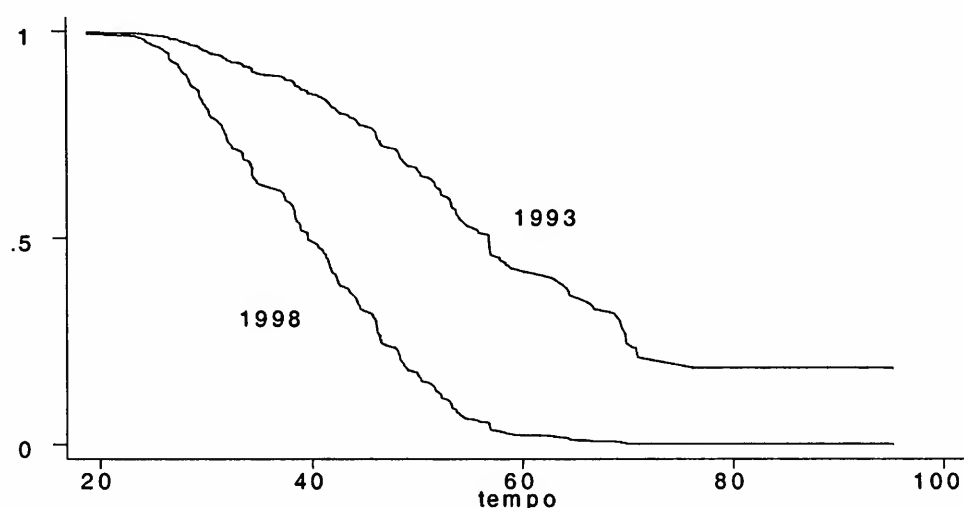


Gráfico 8

Estimativa do Modelo de Cox Estendido para a Função Sobrevivência de um Aluno Típico da UFMG Segundo a Nota Relativa

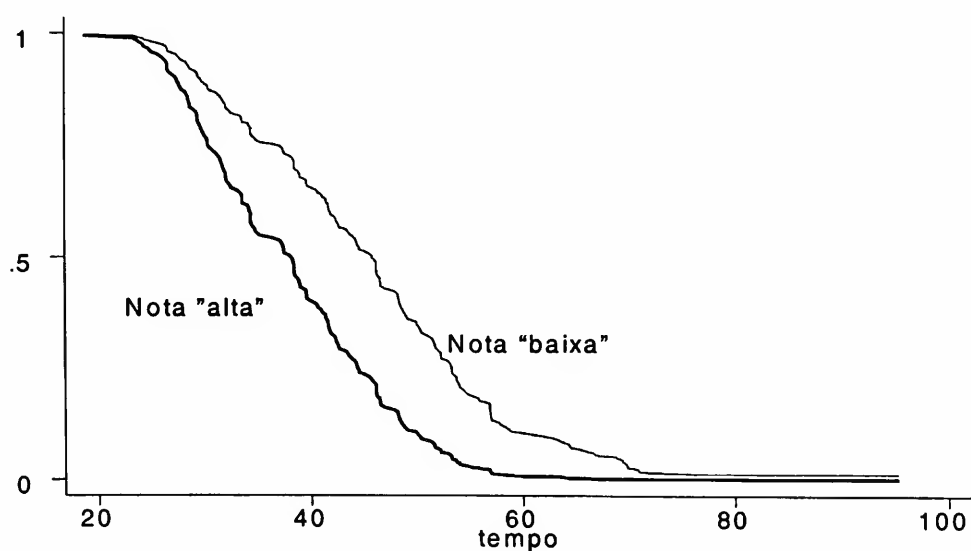


Gráfico 9
Estimativa do Modelo de Cox Estendido para a Função Sobrevivência de um Aluno Típico Segundo o Centro (UFRGS ou não)

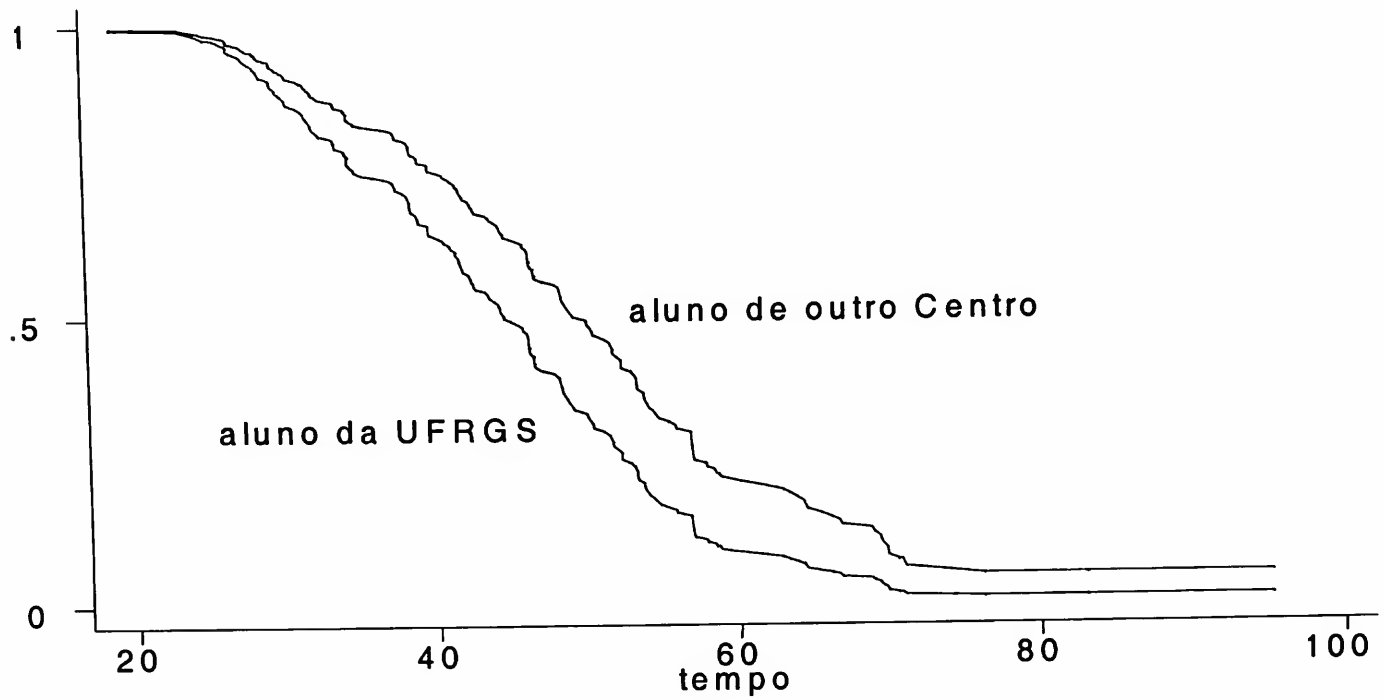
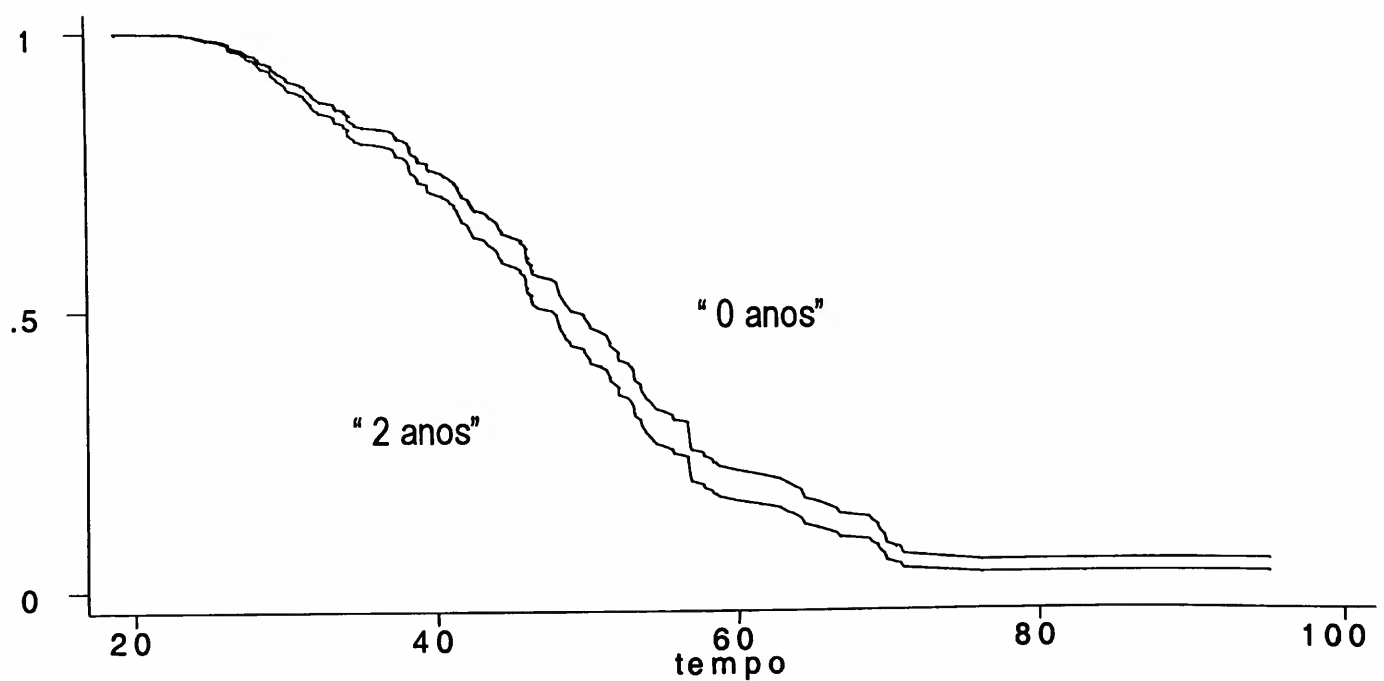
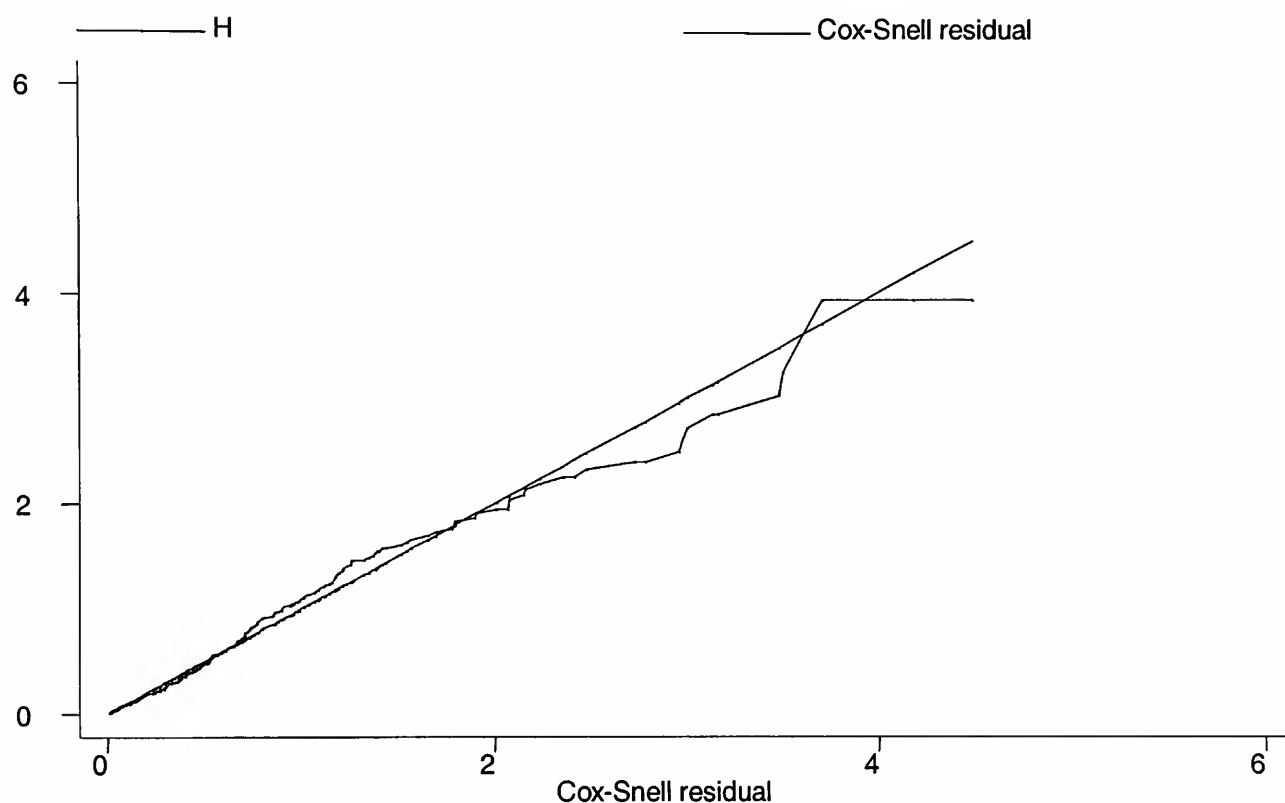


Gráfico 10
Estimativa do Modelo de Cox Estendido para a Função Sobrevivência de um Aluno Típico da PUC Segundo o Tempo Entre o Final da Graduação e Início do Mestrado



O Gráfico 11 mostra o procedimento proposto do Cox e Snell (1968) para examinar o ajustamento da especificação selecionada. No eixo das ordenadas encontra-se o valor estimado dos resíduos de Cox e Snell, enquanto que no eixo das abscissas encontram-se os valores observados da função risco integrada, obtidos a partir do estimador de Kaplan-Meier. Caso o modelo especificado seja o correto, espera-se que os pontos se concentrem ao longo da linha de 45 graus, ao passo que desvios sistemáticos indicam uma especificação inadequada do modelo. Como se pode observar, o modelo de Cox se ajusta consideravelmente aos dados.

Gráfico 11
Procedimento de Cox e Snell para Testar a Especificação
do Modelo de Cox Estendido



Por outro lado, a hipótese mais importante do modelo de Cox, cuja validade merece maior atenção, refere-se à proporcionalidade das funções risco entre os diversos grupos ao longo do tempo. O teste formal, cujos resultados são expostos na Tabela 6, é baseado em Grambsch e Thernau (1994). A hipótese alternativa, considerada pelo referido teste, postula que os coeficientes no modelo de riscos proporcionais variam no tempo.

Tabela 6
Resultados do Teste de Grambsch e Thernau

	rho	χ^2	df	P-Valor
Ingresso em 1998	-0.0916	1.80	1	0.180
Ingresso em 1993	0.0743	0.94	1	0.332
Centro_UFRGS	0.0850	1.71	1	0.191
PUC_anos_entre	-0.0485	0.12	1	0.726
UFMG_nota_relativa	0.0204	0.10	1	0.751
Nota_relativa	-0.0781	1.48	1	0.223
Teste global		9.29	6	0.158

Não é possível rejeitar a hipótese nula da proporcionalidade das funções risco entre os diversos grupos ao longo do tempo, tanto para o conjunto das variáveis explicativas (teste global) quanto para as variáveis individuais.

3.4 Modelo de aceleração da saída

Se, por um lado, o modelo de Cox não necessita de hipóteses sobre a forma funcional da função de risco básico, por outro, restringe a maneira pela qual as variáveis explicativas interagem com essa última função. O modelo de aceleração da saída, por sua vez, permite a adoção de uma distribuição paramétrica para a duração, além de tornar a função risco dependente de um conjunto de regressores, o que equivale a transformar a unidade de mensuração do eixo relativo ao tempo. Mais especificamente, no modelo de aceleração da saída o logaritmo natural do tempo de sobrevivência é expresso como uma função linear das covariadas.

Ao permitir a parametrização da duração, esta estratégia de modelagem possibilita que se estude com maior detalhamento a existência de dependência de duração. Nesse sentido, o fato de um aluno ter maior ou menor probabilidade de terminar o curso em determinado período de tempo, dado que ele durou até aquele momento, pode ser colocado em termos de dependência de duração e, conseqüentemente, a resposta dependerá, em grande parte, da distribuição paramétrica que o analista esteja disposto a aceitar.

Quanto aos resultados, inicialmente realizamos várias estimativas do tempo de saída do curso com base nas distribuições de Weibull, exponencial, log-logística, lognormal e gamma generalizada. Com relação às variáveis explicativas incluídas, adotou-se a mesma especificação do modelo de riscos proporcionais, com exceção da inclusão de uma constante, que no modelo anterior não podia ser estatisticamente identificada. A Tabela 7 apresenta uma síntese

dos resultados encontrados para as diferentes estimações. Como é possível observar, o sinal dos coeficientes estimados é bastante semelhante entre as diferentes distribuições paramétricas para o tempo de duração.

Tabela 7
Síntese dos Resultados Obtidos Por Meio da Estimação do Modelo de Aceleração da Saída Segundo as Distribuições Seleccionadas

	Exponencial	Weibull	Lognormal	Log-logística	Gamma Generalizada
Centro_UFRGS	-0.232	-0.116	-0.068	-0.065	-0.054
Ingresso em 1993	0.007	0.130	0.171	0.197	0.174
Ingresso em 1998	-0.127	-0.247	-0.224	-0.209	-0.226
PUC_anos_entre_grad	-0.032	-0.022	-0.021	-0.019	-0.022
UFMG_nota_relativa	-0.112	-0.099	-0.091	-0.091	-0.090
Nota_relativa	-0.195	-0.117	-0.111	-0.107	-0.110
Constante	4.548	4.262	4.098	4.079	4.062
Log-Verossimilhança	-287.62	-85.41	-63.81	-63.61	-63.02
Critério de informação de Akaike	587.25	184.81	141.62	141.21	142.04

Resultados mais detalhados da função sobrevivência e da função risco usando-se a distribuição log-logística que, como será visto mais adiante, se ajusta bastante bem aos dados, serão apresentados a seguir.

Tabela 8
Estimativas do Modelo de Aceleração da Saída Baseado na Distribuição Log-logística

Variável	Razão de Risco	Erro Padrão	Teste "z"	P-valor
Centro_UFRGS	-0.06504	0.035033	-1.856	0.063
Ingresso em 1993	0.196769	0.045648	4.311	0.000
Ingresso em 1998	-0.20929	0.05226	-4.005	0.000
PUC_anos_entre_grad	-0.01891	0.008745	-2.162	0.031
UFMG_nota_relativa	-0.09078	0.017341	-5.235	0.000
Nota_relativa	-0.10697	0.023245	-4.602	0.000
Constante	4.079351	0.052713	77.388	0.000
Gamma	0.151953	0.008484		

Número de observações: 289

Número de observações não-censuradas: 225

Wald $\chi^2(6)$: 215.27

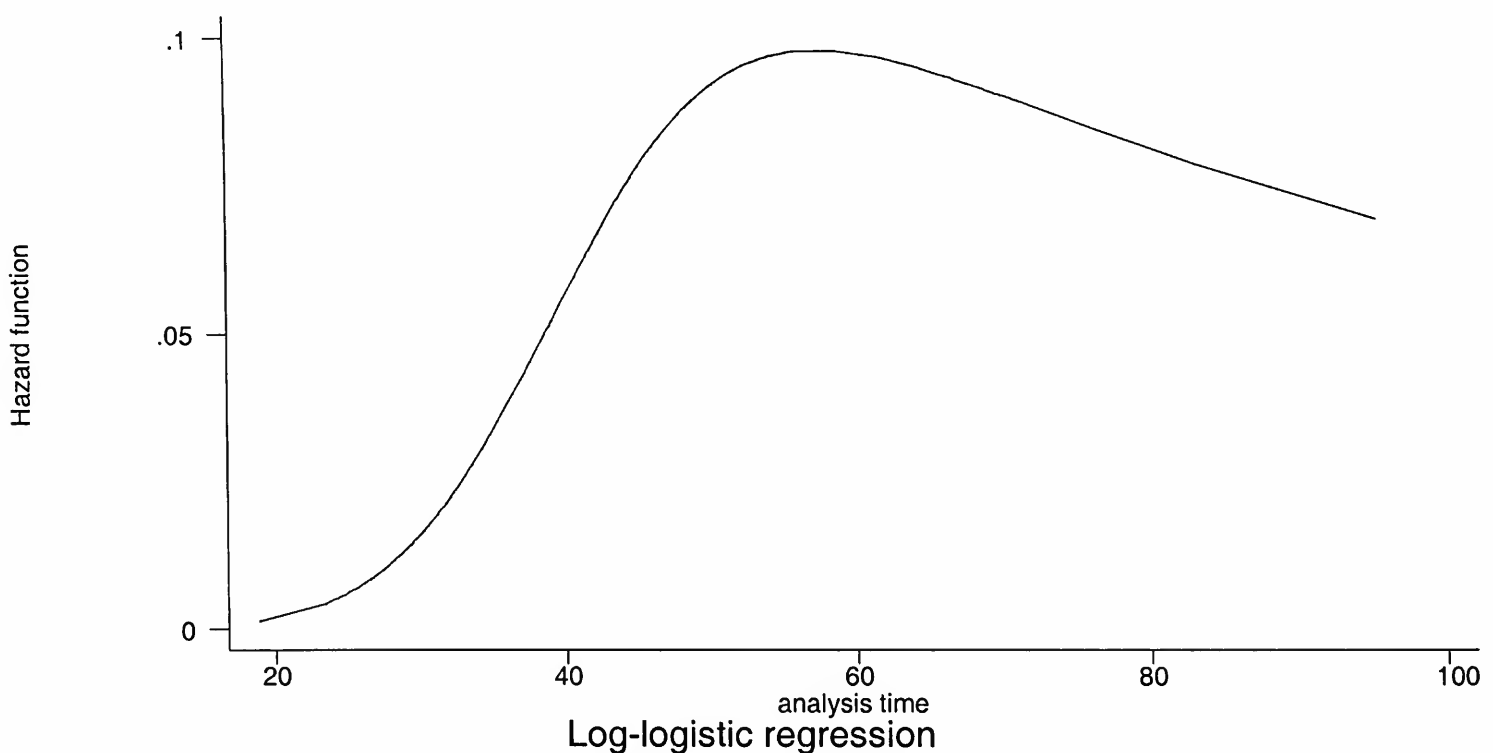
Log Verossimilhança: -63.6075

Apesar da métrica diferente com relação ao modelo de risco proporcional, os resultados qualitativos dos efeitos das covariadas sobre o tempo de duração do curso são idênticos àqueles encontrados para o modelo de Cox estendido. No entanto, a vantagem do modelo paramétrico reside na possibilidade de estimação direta das funções risco e sobrevivência.

O Gráfico 12 mostra a função risco para um aluno típico da amostra. Observa-se que a probabilidade condicional de o aluno terminar o curso aumenta rapidamente até o quadragésimo mês de curso, quando então essa probabilidade cresce de forma mais lenta, alcançando um valor próximo à unidade um pouco antes do sexagésimo mês de curso, quando então essa probabilidade começa a apresentar comportamento decrescente.

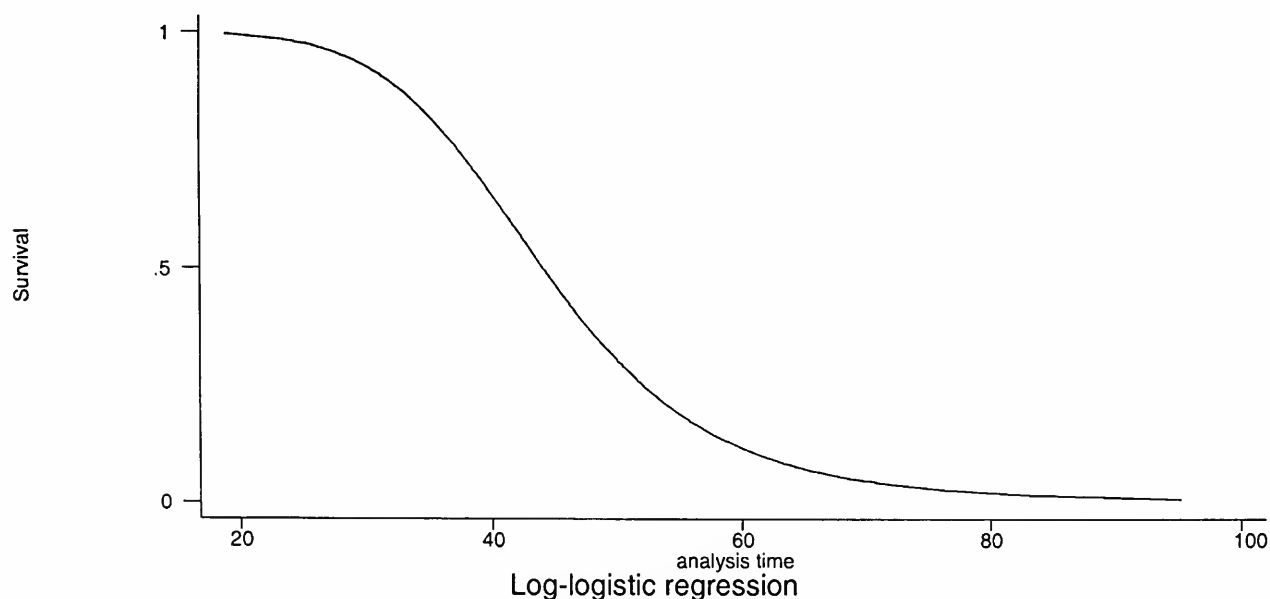
Gráfico 12

Estimativa da Função de Risco para um Aluno Típico da Amostra



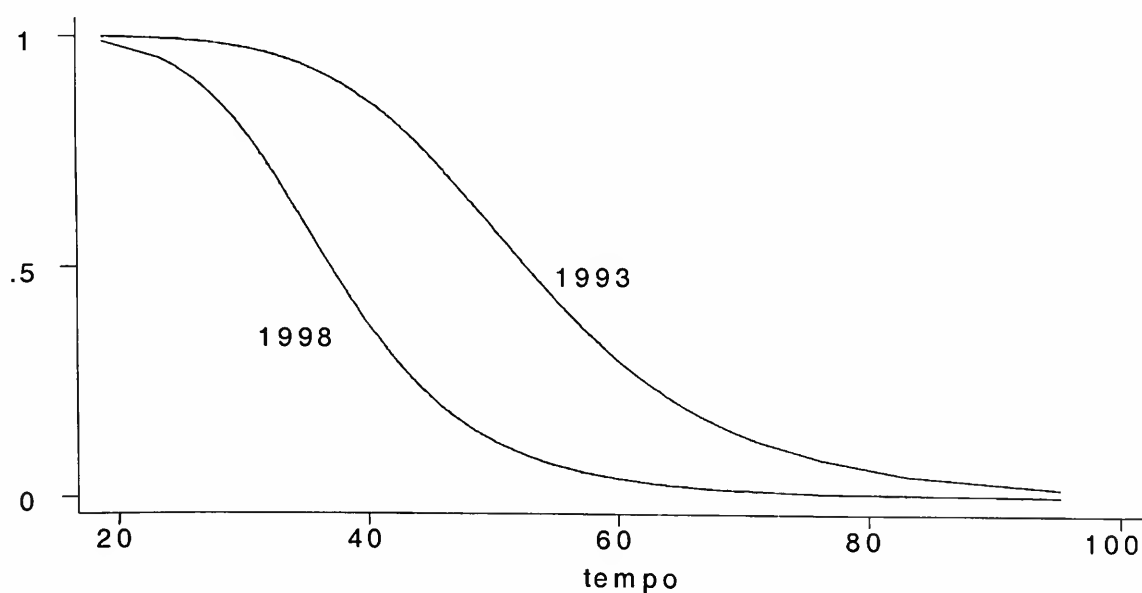
O Gráfico 13, por sua vez, mostra a função sobrevivência para um aluno típico da amostra. Ele fornece a mesma informação dada pela função risco, ou seja, a probabilidade do aluno típico da amostra continuar no curso decresce lentamente até aproximadamente o quadragésimo mês, quando começa a apresentar um comportamento de queda mais acentuado, alcançando praticamente zero depois do sexagésimo mês de curso.

Gráfico 13
Estimativa da Função Sobrevivência para um Aluno Típico da Amostra



O aluno típico que ingressa em 1998 tem uma função sobrevivência bastante distinta da do aluno que ingressa em 1993, como se observa no Gráfico 14. Primeiro, a probabilidade de continuar no curso cai muito mais rapidamente. Antes mesmo de 40 meses a probabilidade de sobrevivência já é menor do que 50%, enquanto que para o aluno ingressante em 1993 só perto dos 60 meses é que este ponto é atingido. Além disso, para um ingressante em 1998, a probabilidade de sobrevivência cai a zero já a partir do sexagésimo mês, enquanto que para os ingressantes em 1993 isto só ocorre depois do octagésimo mês.

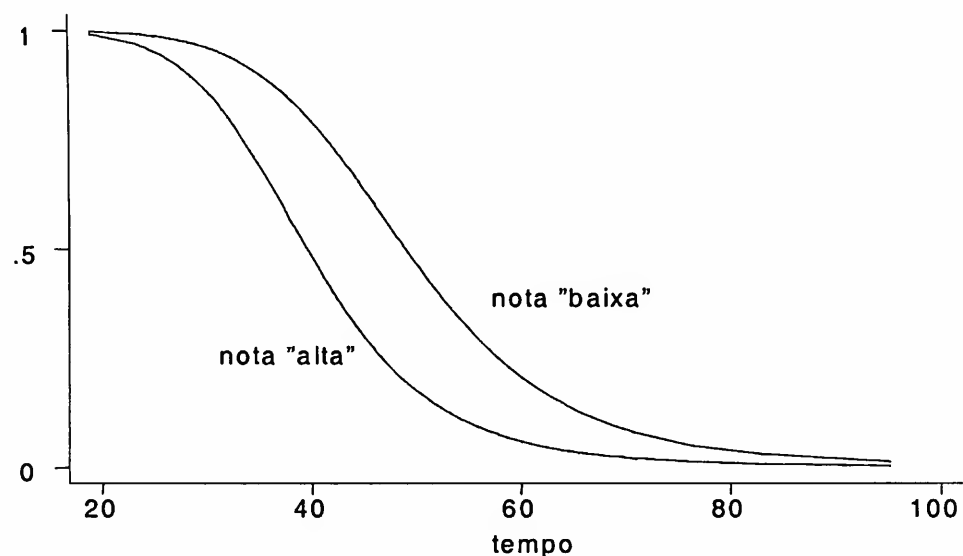
Gráfico 14
Estimativa do Modelo de Cox Estendido para a Função Sobrevivência de um Aluno Típico da Amostra Ingressante nos Anos de 1993 e 1998



O Gráfico 15 mostra a estimativa da função sobrevivência de um aluno típico da amostra de acordo com a nota relativa. O aluno com nota alta tem uma probabilidade de sobrevivência de 50% no quadragésimo mês, enquanto que um aluno de nota baixa tem esta mesma probabilidade até quase o sexagésimo mês.

Gráfico 15

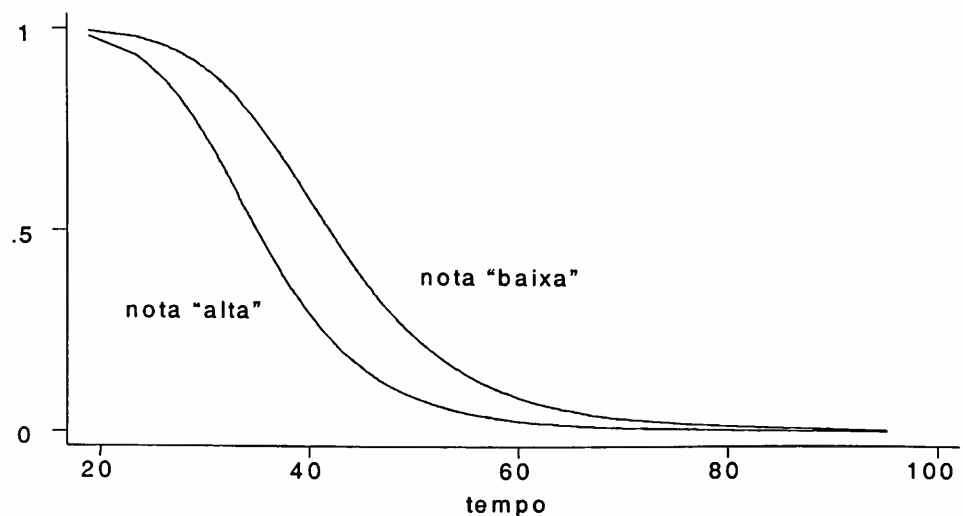
Estimativa do Modelo de Cox Estendido para a Função Sobrevivência de um Aluno Típico da Amostra Segundo a Nota Relativa



O Gráfico 16 mostra a estimativa para a função sobrevivência de um aluno típico da UFMG segundo a nota relativa, mostrando o mesmo efeito qualitativo do gráfico anterior.

Gráfico 16

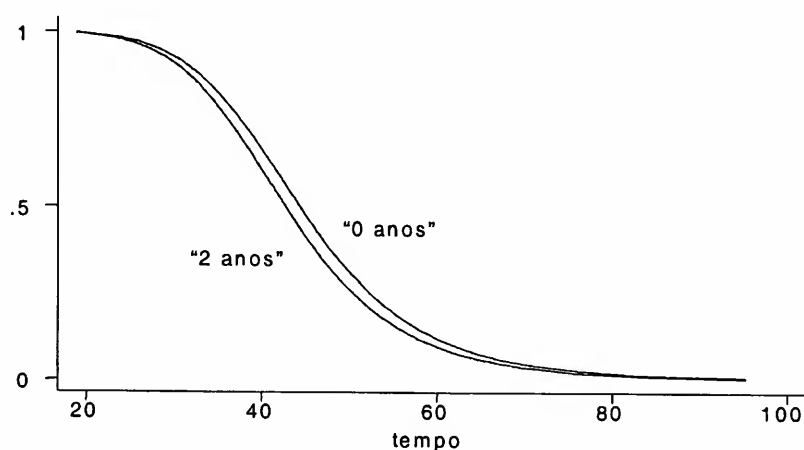
Estimativa do Modelo de Cox Estendido para a Função Sobrevivência de um Aluno Típico da UFMG Segundo a Nota Relativa



No Gráfico 17 visualiza-se o comportamento da função sobrevivência de um aluno típico da PUC segundo o tempo entre o final da graduação e início do Mestrado. Para esse centro, o aluno típico que possui dois anos entre o final da graduação e início no curso possui uma probabilidade ligeiramente menor de permanecer nesse, quando comparado ao aluno típico que sai da graduação diretamente para o Mestrado.

Gráfico 17

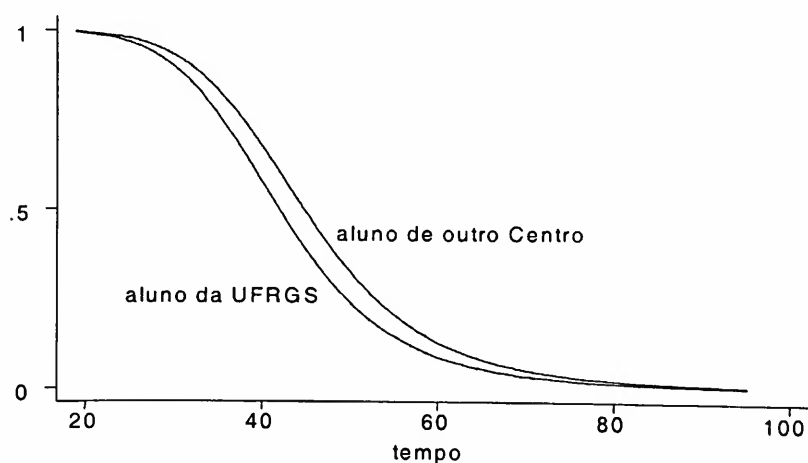
Estimativa do Modelo de Cox Estendido para a Função Sobrevivência de um Aluno Típico da PUC Segundo o Tempo Entre o Final da Graduação e Início do Mestrado



O Gráfico 18 mostra a função sobrevivência para um aluno típico da UFRGS comparativamente ao aluno típico dos demais centros. Observa-se que, controlado pelas demais variáveis, o fato do aluno cursar o Mestrado na UFRGS diminui a probabilidade de permanência no curso.

Gráfico 18

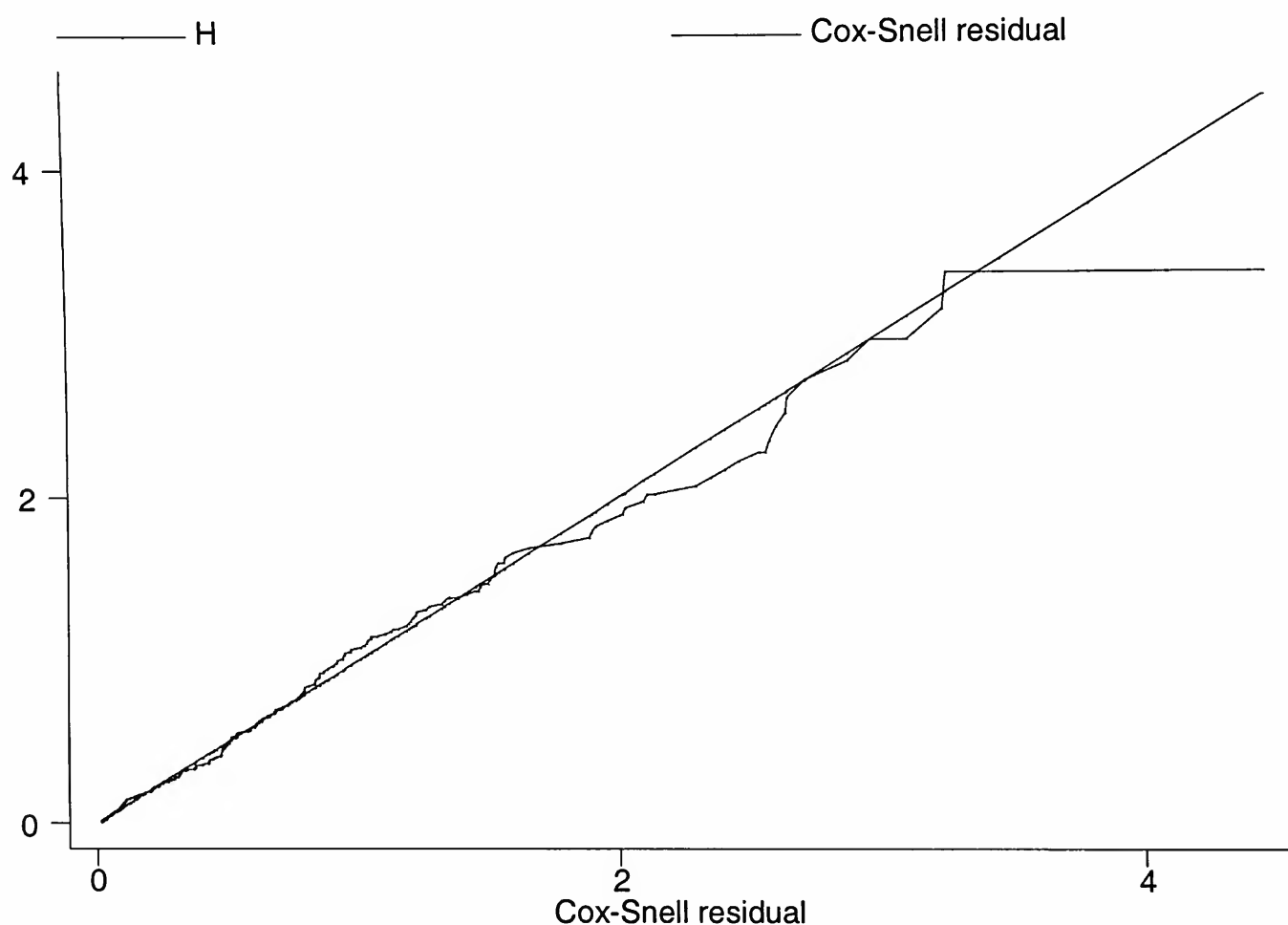
Estimativa do Modelo de Cox Estendido para a Função Sobrevivência de um Aluno Típico da Amostra Segundo o Centro (UFRGS ou não)



Para analisar a adequação do modelo escolhido usa-se novamente o modelo de Cox e Snell. Como observado, o modelo de aceleração da saída com a especificação baseada na função log-logística se ajusta bastante bem aos dados, conforme mostra o Gráfico 19.

↑

Gráfico 19
Procedimento de Cox e Snell para Testar a Especificação do Modelo
Baseado na Distribuição Log-logística



4 Conclusão

Conforme a análise precedente, os estimadores de Kaplan-Meier, Cox estendido e de aceleração da saída apresentaram resultados bastante semelhantes, com ligeiras distinções.

Quanto ao estimador de Kaplan-Meier, é possível asseverar que foram confirmadas estatisticamente pelo teste de Savage o centro, a instituição onde foi feita a graduação, o ano de ingresso e a nota relativa.

No que se refere ao modelo de Cox estendido, além da confirmação do ano de entrada, a influência da nota relativa no término mais rápido do curso poderia ser acrescentada à análise. Com relação ao modelo de aceleração da saída, os resultados encontrados são praticamente os mesmos do modelo semiparamétrico, na medida em que os efeitos qualitativos das covariadas sobre o tempo de duração do curso mostrados são idênticos àqueles encontrados para o modelo de Cox estendido.

Em suma, os três estimadores utilizados no estudo não apresentaram efeitos contraditórios das variáveis explicativas sobre o tempo de duração do Mestrado, fato este que parece legitimar os resultados encontrados.

Para finalizar, é importante observar novamente que este é apenas um estudo preliminar sobre o assunto. A maior parte dos centros não tem informação sistematizada sobre as características dos seus alunos. Assim, espera-se que com o passar do tempo a amostra aumente e os resultados se tornem mais claros e mais gerais.

De qualquer forma, o exercício é válido como uma primeira abordagem para um problema que é fundamental, qual seja, o de verificar quais são os determinantes da duração dos cursos de Mestrado em Economia no Brasil. Isto porque, espera-se que a partir do conhecimento desses determinantes seja possível redesenhar as políticas dos diversos centros de pós-graduação, no sentido de reduzir o tempo de duração sem que a qualidade dos mesmos seja comprometida, permitindo, assim, uma melhor alocação dos recursos destinados ao Mestrado, que são essencialmente públicos.

Bibliografia

- Avelino, R. *Os determinantes da duração do desemprego em São Paulo*. 2001. Dissertação (Mestrado em Economia). IPE/USP. São Paulo.
- Cox, D. R.; Snell, E. J. A general definition of residuals. *Journal of the Royal Statistical Society*, B 30. p. 248-275, 1968.
- Grambsch, P.; Therneau, T. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*, v. 81, p. 515-526, 1994.
- Greene, W. *Econometric analysis*. Macmillan, 1993.
- Issler, J. V.; Pillar, T. Mensurando a produção científica de pesquisadores e instituições brasileiras. *Texto de Seminários Acadêmicos*, IPE/USP, 2002.

Kiefer, N. Economic duration data and hazard functions. *Journal of Economic Literature*, v. 26, p. 646-679, 1988.

Menezes-Filho, N. A., Picchetti, P. Os determinantes da duração do desemprego em São Paulo. *Pesquisa e Planejamento Econômico*, v. 30, n. 1, abril 2000.

Apêndice

Tabela A1
Resultados do Estimador de Kaplan-Meyer

Tempo	Total de Observações	Saídas	Perda Líquida	Função Sobrevivência	Função Sobrevivência IC 95% Lim Inferior	Função Sobrevivência IC 95% Lim Superior
18.83	289	1	0	0.9965	0.9757	0.9995
23.4	288	1	0	0.9931	0.9726	0.9983
23.7	287	1	0	0.9896	0.9682	0.9966
24.23	286	1	0	0.9862	0.9635	0.9948
24.37	285	1	0	0.9827	0.9589	0.9928
24.8	284	1	0	0.9792	0.9544	0.9906
25	283	1	0	0.9758	0.9499	0.9884
25.6	282	1	0	0.9723	0.9454	0.9861
26	281	1	0	0.9689	0.941	0.9837
26.17	280	1	0	0.9654	0.9366	0.9812
26.47	279	1	0	0.9619	0.9323	0.9787
26.5	278	1	0	0.9585	0.928	0.9762
26.53	277	2	0	0.9516	0.9196	0.971
26.67	275	1	0	0.9481	0.9154	0.9684
27.33	274	2	0	0.9412	0.9071	0.963
27.4	272	1	0	0.9377	0.903	0.9603
27.63	271	2	0	0.9308	0.8948	0.9548
27.73	269	1	0	0.9273	0.8907	0.952
27.93	268	1	0	0.9239	0.8867	0.9492
28.17	267	1	0	0.9204	0.8827	0.9464
28.23	266	1	0	0.917	0.8787	0.9436
28.4	265	1	0	0.9135	0.8747	0.9407
28.43	264	1	0	0.91	0.8707	0.9378
28.57	263	2	0	0.9031	0.8628	0.9321
28.67	261	1	0	0.8997	0.8588	0.9292
29	260	1	0	0.8962	0.8549	0.9262
29.23	259	1	0	0.8927	0.851	0.9233
29.3	258	2	0	0.8858	0.8432	0.9174
29.33	256	1	0	0.8824	0.8393	0.9145
29.43	255	1	0	0.8789	0.8354	0.9115
29.6	254	1	0	0.8754	0.8315	0.9085
29.67	253	1	0	0.872	0.8277	0.9055
29.7	252	1	0	0.8685	0.8238	0.9025
29.87	251	1	0	0.8651	0.82	0.8995
30	250	1	0	0.8616	0.8162	0.8965
30.17	249	1	0	0.8581	0.8123	0.8935
30.27	248	1	0	0.8547	0.8085	0.8905
30.3	247	2	0	0.8478	0.8009	0.8844
30.43	245	1	0	0.8443	0.7971	0.8813
31.07	244	2	0	0.8374	0.7895	0.8752
31.23	242	1	0	0.8339	0.7858	0.8721
31.43	241	1	0	0.8304	0.782	0.869
31.53	240	2	0	0.8235	0.7745	0.8629
31.67	238	1	0	0.8201	0.7707	0.8598
31.7	237	1	0	0.8166	0.767	0.8567
31.83	236	1	0	0.8131	0.7632	0.8536
31.93	235	1	0	0.8097	0.7595	0.8504

Tempo	Total de Observações	Saídas	Perda Líquida	Função Sobrevivência	Função Sobrevivência IC 95% Lim Inferior	Função Sobrevivência IC 95% Lim Superior
31.97	234	1	0	0.8062	0.7558	0.8473
32	233	1	0	0.8028	0.752	0.8442
32.03	232	1	0	0.7993	0.7483	0.8411
32.2	231	1	0	0.7958	0.7446	0.8379
32.37	230	1	0	0.7924	0.7409	0.8348
32.47	229	1	0	0.7889	0.7372	0.8316
33	228	1	0	0.7855	0.7335	0.8285
33.37	227	1	0	0.782	0.7298	0.8253
33.43	226	1	0	0.7785	0.7261	0.8222
33.47	225	1	0	0.7751	0.7224	0.819
33.53	224	2	0	0.7682	0.7151	0.8127
34.1	222	1	0	0.7647	0.7114	0.8095
34.17	221	1	5	0.7612	0.7078	0.8063
34.2	215	1	0	0.7577	0.704	0.803
34.27	214	0	30	0.7577	0.704	0.803
34.33	184	1	0	0.7536	0.6995	0.7993
34.37	183	3	0	0.7412	0.6862	0.7881
34.5	180	1	0	0.7371	0.6817	0.7844
34.73	179	1	0	0.733	0.6773	0.7807
34.8	178	1	0	0.7289	0.6729	0.7769
35.27	177	1	0	0.7248	0.6685	0.7731
36.3	176	1	0	0.7206	0.6641	0.7694
37.07	175	1	0	0.7165	0.6597	0.7656
37.33	174	1	0	0.7124	0.6553	0.7618
37.47	173	2	0	0.7042	0.6465	0.7542
37.5	171	1	0	0.7001	0.6422	0.7504
37.97	170	1	0	0.6959	0.6378	0.7466
38.23	169	1	0	0.6918	0.6335	0.7428
38.37	168	2	0	0.6836	0.6248	0.7352
38.4	166	1	0	0.6795	0.6205	0.7313
38.47	165	2	0	0.6712	0.6118	0.7236
38.5	163	1	0	0.6671	0.6075	0.7198
38.8	162	1	0	0.663	0.6032	0.7159
38.87	161	1	0	0.6589	0.5989	0.7121
38.93	160	1	0	0.6548	0.5946	0.7082
38.97	159	1	0	0.6506	0.5903	0.7043
39.53	158	1	0	0.6465	0.5861	0.7005
39.6	157	3	0	0.6342	0.5733	0.6888
40.53	154	2	0	0.6259	0.5647	0.681
40.6	152	1	0	0.6218	0.5605	0.6771
40.97	151	1	0	0.6177	0.5563	0.6732
41.07	150	1	0	0.6136	0.552	0.6693
41.33	149	1	0	0.6095	0.5478	0.6654
41.37	148	1	0	0.6053	0.5436	0.6614
41.53	147	1	0	0.6012	0.5394	0.6575
41.57	146	1	0	0.5971	0.5351	0.6536
41.63	145	1	0	0.593	0.5309	0.6496
41.8	144	2	0	0.5848	0.5225	0.6417
41.87	142	1	0	0.5806	0.5184	0.6378
42.2	141	1	0	0.5765	0.5142	0.6338
42.3	140	1	0	0.5724	0.51	0.6299
42.33	139	1	0	0.5683	0.5058	0.6259
42.5	138	1	0	0.5642	0.5016	0.6219
42.53	137	1	0	0.56	0.4975	0.6179
42.63	136	1	0	0.5559	0.4933	0.614
43.33	135	1	0	0.5518	0.4892	0.61
43.47	134	1	0	0.5477	0.485	0.606
43.6	133	1	0	0.5436	0.4809	0.602
43.9	132	1	0	0.5395	0.4767	0.598
44.1	131	1	0	0.5353	0.4726	0.594
44.2	130	1	0	0.5312	0.4684	0.59
44.3	129	2	0	0.523	0.4602	0.5819
44.47	127	1	0	0.5189	0.4561	0.5779
44.5	126	1	0	0.5147	0.452	0.5739
45	125	1	0	0.5106	0.4479	0.5699
45.47	124	1	0	0.5065	0.4438	0.5658
45.7	123	1	0	0.5024	0.4397	0.5618
45.73	122	1	0	0.4983	0.4356	0.5577
45.93	121	1	0	0.4942	0.4315	0.5537
46.03	120	1	0	0.49	0.4274	0.5496
46.17	119	4	0	0.4736	0.4111	0.5334
46.23	115	2	0	0.4653	0.403	0.5252
46.4	113	1	0	0.4612	0.3989	0.5211
46.43	112	0	15	0.4612	0.3989	0.5211

Tempo	Total de Observações	Saídas	Perda Líquida	Função Sobrevivência	Função Sobrevivência IC 95% Lim Inferior	Função Sobrevivência IC 95% Lim Superior
46.53	97	2	0	0.4517	0.3894	0.5119
46.57	95	1	0	0.4469	0.3847	0.5072
46.97	94	1	0	0.4422	0.3799	0.5026
47.83	93	1	0	0.4374	0.3752	0.4979
48.13	92	1	0	0.4327	0.3705	0.4932
48.3	91	3	0	0.4184	0.3563	0.4792
48.33	88	1	0	0.4137	0.3517	0.4745
48.47	87	1	0	0.4089	0.347	0.4698
48.63	86	1	0	0.4042	0.3423	0.4651
48.77	85	1	0	0.3994	0.3376	0.4603
48.97	84	1	0	0.3946	0.333	0.4556
49.1	83	1	0	0.3899	0.3283	0.4509
49.93	82	1	0	0.3851	0.3237	0.4461
50	81	1	0	0.3804	0.3191	0.4414
50.13	80	1	0	0.3756	0.3144	0.4366
50.33	79	1	0	0.3709	0.3098	0.4319
50.37	78	1	0	0.3661	0.3052	0.4271
51.03	77	1	0	0.3614	0.3006	0.4223
51.43	76	1	0	0.3566	0.296	0.4175
51.63	75	2	0	0.3471	0.2869	0.4079
51.67	73	1	0	0.3423	0.2823	0.4031
52.2	72	2	0	0.3328	0.2732	0.3935
52.23	70	2	0	0.3233	0.2642	0.3838
52.83	68	1	0	0.3186	0.2596	0.3789
53.07	67	1	0	0.3138	0.2551	0.3741
53.17	66	1	0	0.3091	0.2506	0.3692
53.2	65	1	0	0.3043	0.2461	0.3643
53.23	64	1	0	0.2995	0.2417	0.3594
53.3	63	1	0	0.2948	0.2372	0.3545
53.67	62	1	0	0.29	0.2327	0.3496
53.7	61	1	0	0.2853	0.2283	0.3447
53.73	60	1	0	0.2805	0.2238	0.3398
53.9	59	1	0	0.2758	0.2194	0.3349
54	58	1	0	0.271	0.215	0.3299
54.27	57	1	0	0.2663	0.2106	0.325
54.67	56	2	0	0.2568	0.2018	0.3151
55.23	54	1	0	0.252	0.1974	0.3101
55.73	53	1	0	0.2472	0.193	0.3051
55.8	52	1	0	0.2425	0.1887	0.3001
56.77	51	1	0	0.2377	0.1843	0.2951
56.93	50	7	0	0.2045	0.1542	0.2598
57	43	1	0	0.1997	0.15	0.2547
57.77	42	1	0	0.1949	0.1457	0.2496
57.8	41	1	0	0.1902	0.1415	0.2445
58.33	40	1	0	0.1854	0.1373	0.2393
58.43	39	1	0	0.1807	0.1331	0.2342
58.6	38	0	5	0.1807	0.1331	0.2342
58.83	33	1	0	0.1752	0.1281	0.2284
60.07	32	1	0	0.1697	0.1232	0.2226
61.6	31	1	0	0.1643	0.1183	0.2168
62.87	30	1	0	0.1588	0.1135	0.2109
63.33	29	1	0	0.1533	0.1086	0.2051
63.8	28	1	0	0.1478	0.1038	0.1992
64.23	27	1	0	0.1424	0.0991	0.1932
64.4	26	1	0	0.1369	0.0943	0.1873
64.5	25	1	0	0.1314	0.0896	0.1813
65.37	24	1	0	0.1259	0.085	0.1753
66.13	23	1	0	0.1205	0.0803	0.1692
66.6	22	1	0	0.115	0.0757	0.1631
66.77	21	1	0	0.1095	0.0712	0.157
68.73	20	1	0	0.104	0.0667	0.1509
68.9	19	1	0	0.0986	0.0622	0.1447
69.3	18	1	0	0.0931	0.0578	0.1384
69.33	17	1	0	0.0876	0.0534	0.1322
69.53	16	1	0	0.0821	0.0491	0.1258
69.73	15	1	0	0.0767	0.0448	0.1194
69.8	14	1	0	0.0712	0.0406	0.113
69.87	13	1	0	0.0657	0.0365	0.1065
70.53	12	1	0	0.0602	0.0325	0.0999
70.77	11	0	5	0.0602	0.0325	0.0999
71.03	6	1	0	0.0502	0.0239	0.091
76.2	5	1	0	0.0402	0.0163	0.0812
82.97	4	0	3	0.0402	0.0163	0.0812
95.13	1	0	1	0.0402	0.0163	0.0812