ISSN 0871-2573

Professores benevolentes ou exigentes?*

Pedro Pita Barros, Universidade Nova de Lisboa
Centre for Economic Policy Research, London

Inês Cabral, Universidade Nova de Lisboa

Dezembro 1995
Working Paper nº 265

Professores benevolentes ou exigentes?*

Pedro Pita Barros
Universidade Nova de Lisboa
Centre for Economic Policy Research, London

Inês Cabral
Universidade Nova de Lisboa

Dezembro de 1995

Abstract

Duas justificações são frequentemente apresentadas para a existência de um teste intermédio nas normas de classificação em disciplinas do ensino superior. Por um lado, a realização de um teste intermédio permite reduzir a incerteza associada à nota final – *teoria do professor benevolente*. Por outro lado, constitui uma forma de induzir um maior esforço de estudo nos alunos – *teoria do professor exigente*.

No presente artigo são analisadas as condições em que cada uma das teorias é consistente com a observação frequente de uma maior ponderação dada ao exame final. No caso da teoria do professor benevolente, estas condições assumem uma forma facilmente testável. Realiza-se assim uma análise empírica simples. Esta revela que consoante o curso se verifica consistência com uma ou com outra teoria. Não se detecta, na pequena amostra considerada, uma predominância clara de qualquer das teorias.

Endereço para correspondência:
Faculdade de Economia
Universidade Nova de Lisboa
Travessa Estêvão Pinto
1070 Lisboa
Tel: (01) 383 36 24
Fax: (01) 388 60 73
Email: ppbarros@fe.unl.pt

*São devidos agradecimentos a Ana B. Reis, Catarina Roseta e Vasco Santos pelas sugestões e comentários realizados. Os erros e omissões são da exclusiva responsabilidade dos autores.

1 Introdução

A análise económica tem-se dedicado a muitos e variados problemas. Curiosamente, uma das áreas que tem recebido menos atenção na literatura diz respeito à própria actividade académica, seja de ensino, seja de investigação. É nosso propósito olhar para a justificação de um facto estilizado correntemente observado: a existência de testes intermédios como elemento de avaliação em cursos de ensino superior. Tipicamente, o aluno é classificado mediante um exame final ou através de uma média ponderada de vários elementos de avaliação, por exemplo, *exame final* e *teste intermédio* (ou *trabalho*). A observação das características usuais deste último sistema revela que maior ponderação é atribuida ao *exame*, na generalidade dos casos.

Duas justificações são frequentemente ouvidas. Por um lado, a existência de um teste intermédio favorece os alunos pois permite-lhes reduzir a incerteza associada à nota final. Esta é a teoria do *professor benevolente*. Por outro lado, induz um maior esforço de estudo por parte dos alunos. Esta justificação é denominada teoria do *professor exigente*.

A teoria do professor benevolente é desenvolvida no quadro de decisões em contexto de incerteza para os alunos, em que a diversificação de elementos de avaliação pode, na presença de aversão ao risco, aumentar a utilidade esperada da nota. A ponderação óptima dada ao *exame* é determinada como se o professor estivesse na situação do aluno, visando maximizar a utilidade da nota final.

A teoria do professor exigente tem uma fundamentação substancialmente distinta. O professor procura maximizar o esforço dispendido pelo aluno por considerar ser esta a melhor forma de garantir que o aluno aprenda. Dado que é o professor quem decide qual a ponderação a atribuir a cada elemento de avaliação, assume-se que esse é o único mecanismo de incentivos detido para estimular o aluno a atingir o objectivo pretendido. A teoria económica desenvolveu na última década uma extensa literatura sobre relações idiossincráticas entre agentes económicos, conhecida pela designação genérica de problemas de "principal-agent".[1]

[1]Uma possível tradução para a lingua portuguesa seria problemas de mandante-mandatário. Devido à consagração do termo inglês na literatura optou-se pela sua manutenção na exposição.

Este tipo de problemas analisa situações em que um agente económico, o "principal", procura induzir outro agente económico, o "agent", a realizar determinada acção. O problema torna-se não trivial pelo facto de essa acção não ser verificável, e como tal não poder figurar em contratos realizados entre as duas partes. A indução de esforço de estudo do aluno por parte do professor é um exemplo deste tipo de relação.

As soluções de equilíbrio encontradas nas duas teorias permitem tirar ilações sobre a razoabilidade dos objectivos, assim como concluir sobre o comportamento tipo mais frequente. A análise revelou, por um lado, que as condições para a existência de testes intermédios são mais restritivas do que a ligeireza com que são usualmente referidas faria supor. Tal verifica-se em cada uma das explicações apresentadas. Uma simples investigação descritiva a partir dos resultados de uma amostra de cursos de ensino superior sugere que existe, em geral, suporte quer para a teoria do professor benevolente quer para a teoria do professor exigente.

O modelo desenvolvido é extensível a mais elementos de avaliação bem como a outros parâmetros considerados importantes nos ensinos básico e secundário: assiduidade, participação e comportamento. Por outro lado, se *exame* e *teste* não forem os elementos mais relevantes para a classificação final podem ser facilmente substituídos por outros como os acima descritos. Este esquema pode mesmo aplicar-se a qualquer avaliação de *performance* que envolva um nível de esforço, de empenho ou de dedicação não observável pelo agente avaliador.

O texto encontra-se organizado do seguinte modo. Na secção 2 é descrita a teoria do professor benevolente e as suas implicações. Na secção seguinte, a terceira, é apresentada a teoria do professor exigente.[2] A secção 4 é dedicada à apresentação de alguns dados exemplificativos. Finalmente, a quinta secção conclui o trabalho.

[2]Estas secções são descritas com algum pormenor que poderá parecer excessivo ao leitor mais familiarizado com as questões de assimetria de informação. Esse grau de pormenor poderia ser sacrificado à custa de se atingir uma menor audiência. Tomou-se no presente texto a perspectiva de facilitar o entendimento do trabalho ao maior número possível de leitores.

2 A teoria do professor benevolente

Apesar de se procurar manter alguma generalidade, é conveniente a consideração de algumas hipóteses que permitem simplificar cálculos e traduzir comportamentos razoáveis dos agentes intervenientes no problema.

Considera-se que o aluno é avesso ao risco. As suas preferências pela nota final de um curso, N, são representadas adequadamente por uma função de utilidade que respeita os axiomas de Von Neumann-Morgenstern. Esta função de utilidade, U, é contínua, diferenciável e exibe aversão absoluta ao risco constante, sendo o grau de aversão ao risco representado por γ.[3]

A existência de incerteza aparenta ser o factor responsável pela introdução de mais do que um elemento de avaliação, dado o interesse em reduzir a variabilidade da nota final. Assim poder-se-ia pensar que a determinação da ponderação óptima dada ao exame (representada por α) teria como objectivo minimizar a variância da nota para o aluno, pelo menos para um mesmo valor esperado. Deste modo, um sistema de avaliação que incluisse mais do que um elemento de avaliação facultaria ao aluno a possibilidade de diversificar o risco associado à sua nota final.

Vamos, de acordo com o exposto, admitir que o objectivo do professor é determinar a ponderação óptima que mais beneficia o aluno. O benefício para o aluno é medido pela utilidade esperada da nota final. Como forma de simplificar a exposição, em lugar de se utilizar o valor esperado da utilidade, é empregue a aproximação ao equivalente certo descrita por

$$\mathcal{E}[U(N)] \simeq \mathcal{E}[N] - \frac{\gamma}{2}\sigma_N^2 \tag{1}$$

sendo $\mathcal{E}[\cdot]$ o operador valor esperado e σ_N^2 a variância da nota final.

Admita-se, por um momento, que o valor esperado da nota de *exame* e *teste* é idêntico. O sistema de avaliação é definido por

$$N = \alpha E + (1-\alpha)T \tag{2}$$

[3] Esta última hipótese visa apenas facilitar a resolução algébrica do modelo.

com $\alpha \in [0,1]$. A maximização do equivalente certo na ponderação a dar à nota final reduz-se à minimização da variância da nota final. O problema da escolha óptima da ponderação a dar ao exame pode ser escrito como:

$$\min_{\alpha} \sigma_N^2 \tag{3}$$

A variância da nota final é dada pela seguinte relação:

$$\sigma_N^2 = \alpha^2\sigma_E^2 + (1-\alpha)^2\sigma_T^2 + 2\alpha(1-\alpha)\sigma_E\sigma_T\rho \tag{4}$$

sendo ρ o coeficiente de correlação entre as notas do *exame* e do *teste* e σ_i^2 a variância das notas do teste ($i = T$) e exame ($i = E$).

A condição de primeira ordem para o problema de minimização da variância da nota final origina como solução:[4]

$$\alpha^* = \frac{\sigma_T^2 - \sigma_E\sigma_T\rho}{(\sigma_E^2 + \sigma_T^2 - 2\rho\sigma_E\sigma_T)} \tag{6}$$

Como simplificação, defina-se $\delta = \sigma_E/\sigma_T$. Pode-se então reescrever a expressão da ponderação óptima como:

$$\alpha^* = \frac{1-\delta\rho}{1+\delta^2-2\delta\rho} \tag{7}$$

Se as variâncias dos dois instrumentos de avaliação forem iguais, ou seja $\delta = 1$, verifica-se que

$$\alpha^* = 1/2, \qquad \forall\rho. \tag{8}$$

Com $\delta \neq 1$ e independência entre os dois elementos de avaliação ($\rho = 0$), tem-se

$$\alpha^* = \frac{\sigma_T^2}{\sigma_E^2 + \sigma_T^2} \tag{9}$$

Obtém-se assim o resultado intuitivo de que a ponderação dada ao exame é maior se a sua variância de nota for inferior à do teste. É natural que tal ocorra. Quanto mais elementos

[4] A condição de segunda ordem para mínimo é

$$\frac{\partial^2\sigma_N^2}{\partial\alpha^2} = 2\sigma_E^2 + 2\sigma_T^2 - 4\sigma_E\sigma_T\rho > 0. \tag{5}$$

para $\rho \in (-1,1)$.

de avaliação já realizados menores as "surpresas" (idiossincracias do professor) e menor a variância nas notas. As restantes situações, caracterizadas por uma correlação não nula ($\rho \neq 0$), originam o mesmo tipo de resultado. Tomando $\rho \in (0,1)$, a seguinte desigualdade tem que ser verificada para que $\alpha > 1/2$ seja óptimo:

$$\frac{\sigma_T^2 - \rho\sigma_E\sigma_T}{\sigma_E^2 + \sigma_T^2 - 2\sigma_E\sigma_T\rho} > \frac{1}{2} \qquad (10)$$

que pode ser simplificado para a solução anterior $\sigma_T^2 > \sigma_E^2$.

Esta é uma hipótese que pode ser refutada empiricamente. No caso de tal suceder, é necessário procurar uma justificação alternativa para que seja dada uma maior ponderação à nota do exame. Essa justificação alternativa pode ser fornecida pela teoria do professor exigente, desenvolvida na próxima secção. Contudo, ainda neste contexto, é possível apresentar uma outra razão para que a ponderação atribuída ao teste seja inferior à do exame, mesmo que esta última prova possua uma maior variância de resultados.

A justificação é a existência de médias/valores esperados diferentes em cada uma das provas. Uma maior variância pode ser preferida se a ela corresponder um maior valor esperado. Em termos formais, tal significa que o exame deve exibir uma média de notas superior à do teste.

A intuição do argumento pode ser formalizada no enquadramento do modelo proposto definindo

$$\mathcal{E}(T) = \mu_T, \qquad \mathcal{E}(E) = \mu_E \qquad (11)$$

como as notas esperadas no teste e exame, respectivamente.

Nestas circunstâncias, o valor óptimo da ponderação do exame é:

$$\alpha^* = \frac{\mu_E - \mu_T + \gamma(\sigma_T^2 - \rho\,\sigma_E\sigma_T)}{\gamma(\sigma_E^2 + \sigma_T^2 - 2\rho\,\sigma_E\sigma_T)} \qquad (12)$$

Verifica-se assim que $\mu_E > \mu_T$ é uma condição necessária para que a ponderação atribuida à nota do exame seja superior (no caso de $\sigma_T^2 < \sigma_E^2$). Essa ponderação será tanto maior quanto menor for o grau de aversão absoluta ao risco. Note-se que no caso de igual valor esperado entre provas, o grau de aversão absoluta ao risco não é relevante pois não existe qualquer compensação de a maior risco poder também corresponder um maior valor esperado.

3 A teoria do professor exigente

A avaliação escolar envolve um problema de *assimetria de informação* dado que o professor desconhece o nível de esforço desenvolvido pelo aluno ou os conhecimentos que este adquiriu. Por outro lado, existe incerteza nessa avaliação pois a nota do(s) elemento(s) de avaliação não depende(m) apenas do esforço ou conhecimento do aluno, mas também de uma componente aleatória, não se podendo inferir da nota obtida o nível exacto de conhecimentos adquiridos pelo aluno.

O desenvolvimento da teoria do professor exigente requere um grau de formalização do problema diferente. O objectivo atribuido ao professor exigente é o estímulo de esforço por parte do aluno ao longo do curso. Na verdade, será provavelmente mais realista supor que o professor valoriza o conhecimento adquirido e não o esforço desenvolvido *per se*. Admite-se assim que existe uma forte correlação positiva entre o esforço que o aluno desenvolve e o volume de conhecimentos que adquire.[5]

O professor tem como único instrumento disponível para levar os alunos a estudar o sistema de avaliação (contrato) que anuncia ao(s) aluno(s). Este pode aceitá-lo ou não (caso em que desiste do curso). Após a decisão de aceitação, ou não, do sistema de avaliação, o aluno desenvolve um determinado nível de esforço, que condiciona a nota. Finalmente, a Natureza junta um ruído aleatório à nota do aluno. Deste modo, constata-se que o problema apresenta as propriedades típicas de um problema de risco moral: a informação é simétrica à altura do contrato, tornando-se assimétrica posteriormente. O professor considera que os alunos são idênticos e exibem preferências semelhantes às descritas na secção anterior.

[5]Em certo sentido, existe uma função produção de conhecimentos que tem como factor produtivo primordial o volume de estudo do aluno. De um modo mais rigoroso, a função aquisição de conhecimentos adquire a forma $F(e)+\varepsilon$, sendo ε um termo aleatório e $F(\cdot)$ uma função não decrescente no esforço de estudo. Nestas condições a maximização do conhecimento esperado do aluno ou a soma dos esforços desenvolvidos origina a mesma solução para o problema.

O professor não observa o nível de esforço, representado por e, mas conhece as notas de teste e exame, T e E, respectivamente. No entanto, T e E são funções do esforço desenvolvido para cada uma dessas provas, e_1 e e_2 respectivamente, bem como de factores aleatórios ε_1 e ε_2. Isto é, $T = T(e_1, \varepsilon_1)$ e $E = E(e_2, \varepsilon_2)$. Assim, um determinado resultado, E ou T, pode provir de diferentes combinações de nível de esforço e de pura sorte (ou falta dela). Em termos mais técnicos, as funções são ***não invertíveis***.

Tal como é usualmente referido, considera-se o *exame* mais difícil que o *teste*, ou seja, o exame exige maior esforço para se obter uma mesma nota na prova. Sem perda de generalidade, usa-se como escala de avaliação o intervalo de 0 a 20. As notas do *exame* e do *teste* são determinadas de acordo com:

$$E(e_2, \varepsilon_2) = 20\left(1 - e^{-\beta e_2}\right) + \varepsilon_2 \tag{13}$$

$$T(e_1, \varepsilon_1) = 20\left(1 - e^{-e_1}\right) + \varepsilon_1 \tag{14}$$

As formas funcionais admitidas garantem que a nota atribuída se situa, em valor esperado, no intervalo desejado. A aditividade do efeito aleatório implica que a variância da nota seja independente do esforço exercido pelo aluno. Este pode apenas afectar a média da distribuição de notas, permanecendo a dispersão inalterada.[6] O parâmetro $\beta < 1$ reflecte a dificuldade relativa do exame. Para um mesmo esforço o aluno obtém uma menor nota no exame que no teste.

O desenvolvimento de actividades de estudo tem custos para os alunos. Admite-se que os custos de desenvolver esforço para *exame* e *teste* são, respectivamente:

$$C(e_2) = \varphi e_2 \tag{15}$$

$$C(e_1) = \theta e_1 \tag{16}$$

com $\varphi \geq \theta > 0$ constantes. Ou seja, o esforço de estudo para o *exame* é pelo menos tão

[6]São ignoradas as tecnicalidades impostas por as notas serem truncadas em 0 como limite inferior e 20 como limite superior. Uma forma de solucionar o problema seria considerar o termo aleatório aditivo ao esforço do aluno. Contudo, a complexidade adicional dessa opção é evitada a favor da simplicidade de exposição. Os resultados centrais não dependem qualitativamente desta hipótese.

custoso como o mesmo volume de estudo dedicado ao *teste*, reflectindo, por exemplo, o maior volume de matéria que é normalmente abarcada pelo *exame*.

Caracterizando o problema por analogia com os modelos correntes de 'principal-agent' defina-se como "principal" o professor e como 'agent' o aluno.[7] A ordem do jogo é a seguinte:

1. O professor oferece o esquema de avaliação (o 'contrato'), definido por α. Admite-se que este é o único esquema de avaliação possível.

2. o aluno decide aceitar ou não o contrato (no caso de rejeição, tem $\alpha = 1$ se o teste for facultativo, e muda de curso se for obrigatório);

3. o aluno exerce esforço de estudo para o teste (e_1) e para o exame (e_2).

4. a nota final (esperada) é dada por $N(e_1, e_2) = \alpha E(e_2) + (1-\alpha)T(e_1)$.[8]

O conjunto de escolhas é, para o professor, $A_P = \alpha \in [0,1]$. Para o aluno, $A_A = \{e_1, e_2\}$, com $e_i \in [0, +\infty)$. A utilidade do aluno é dada por $V(N)$ menos o custo de desenvolver esforço. A utilidade do professor é simplesmente a soma do esforço desenvolvido pelo aluno para cada uma das provas.

Assim, o objectivo do aluno é o de maximizar, através do nível de esforço escolhido para cada uma das provas, a utilidade esperada da nota final, líquida de custos (ou desutilidade) de estudar. Dado que o empenho no estudo proporciona também prazer e conhecimento acumulado, pressupõe-se que à função custos foi já deduzido esse efeito. Por seu lado, o objectivo do professor consiste na maximização, através do parâmetro α, do esforço dispendido pelo aluno na aprendizagem da disciplina. Crê-se que o nível de esforço, por exemplo, medido em horas de estudo, é uma boa aproximação para o que o aluno fica efectivamente a saber.

O professor reconhece que o aluno é livre de rejeitar o método de avaliação proposto e que o contrato deve dar ao aluno incentivos para optar pelo nível de esforço desejado. Estas duas constatações não são mais do que as habituais restrições de participação e compatibilidade de incentivos, respectivamente, desta classe de problemas.

[7] Para uma apresentação destes modelos veja-se, por exemplo, Rasmusen (1989) ou Laffont (1988).

[8] Note-se que a nota final é aleatória devido aos elementos ε_1 e ε_2.

O equilíbrio do jogo é encontrado seguindo o esquema de indução retrospectiva. Assim, começa-se por determinar os valores óptimos para as decisões dos jogadores da última para a primeira etapa do jogo. Para resolver algebricamente o problema vai-se inicialmente maximizar em e_1 e e_2 a utilidade esperada do aluno. Tal como na secção anterior, o valor esperado da utilidade da nota é simplificado pelo recurso à aproximação ao equivalente certo. Definindo $\mathcal{W}$ como sendo o valor esperado da utilidade da nota líquido dos custos do esforço dispendido pelo aluno, tem-se a seguinte expressão:

$$\begin{aligned}\mathcal{W} &= \mathcal{E}[U(N)] - C(e_1) - C(e_2) \qquad (17)\\ &= (1-\alpha)T(e_1) + \alpha E(e_2) - \frac{1}{2}\gamma\left[\alpha^2\sigma_E^2 + (1-\alpha)^2\sigma_T^2 + 2\alpha(1-\alpha)\rho\sigma_T\sigma_E\right] - \theta e_1 - \varphi e_2\end{aligned}$$

Deste modo, é possível determinar explicitamente os níveis óptimos das três variáveis de decisão. Substituindo as funções propostas para as notas de teste e de exame e procedendo à maximização da expressão acima, encontram-se os valores óptimos para os níveis de esforço que irão funcionar como restrições ao problema do professor na escolha do ponderador α.

Do problema do aluno, obtém-se como soluções óptimas:

$$\begin{aligned}\frac{\partial\mathcal{W}}{\partial e_2} = 0 &\implies e_2^* = \frac{1}{\beta}\log\left(\frac{20\alpha\beta}{\varphi}\right)\\ \frac{\partial\mathcal{W}}{\partial e_1} = 0 &\implies e_1^* = \log\left(\frac{20(1-\alpha)}{\theta}\right)\end{aligned}$$

Tendo em conta estas soluções para o nível de esforço, o problema do professor é

$$\begin{aligned}\max_{\alpha} S &= e_1 + e_2\\ s.a. \quad & e_2^* = \frac{1}{\beta}\log\left(\frac{20\alpha\beta}{\varphi}\right)\\ & e_1^* = \log\left(\frac{20(1-\alpha)}{\theta}\right)\\ & \mathcal{W} \geq \bar{\mathcal{W}}\end{aligned}$$

sendo $\bar{W}$ o valor da utilidade de reserva.[9]

Da resolução deste problema, o valor óptimo para o factor de ponderação é

$$\alpha^* = \frac{1}{1+\beta} \tag{18}$$

Sendo $0 \leq \beta \leq 1$, o valor de α será sempre superior ou igual a $1/2$. Assim, verifica-se que para maximizar o nível de esforço total dispendido pelo aluno, o professor escolherá sempre $1/2 \leq \alpha^* \leq 1$ pois é a forma de conferir maior peso ao elemento de avaliação que exige maior esforço.

O facto de o valor de α, determinado de forma óptima, não depender dos custos marginais do esforço (θ e φ) deve-se à simplicidade da função de utilidade considerada assim como à linearidade das funções custo de esforço, ambas escolhidas com intuitos de simplificação de cálculo.

Através da expressão encontrada infere-se que a melhor maneira de conseguir incentivar os alunos a desenvolver mais esforço na disciplina, usando o sistema de avaliação descrito por α como mecanismo de incentivos, será fixando um valor para α tanto mais elevado quanto maior for o grau de dificuldade relativa do exame face ao teste (descrita por β).

Apenas na situação em que o grau de dificuldade é idêntico existe igual ponderação nos dois elementos de avaliação. Caso contrário, ter-se-á sempre a atribuição de maior peso ao exame. A dependência do ponderador α relativamente ao parâmetro β é o factor gerador da assimetria deste peso a favor do exame, corroborando o que usualmente se observa nas normas de classificação escolar.

Na realidade, o objectivo dos professores poderá ser mais correctamente descrito por uma combinação dos dois casos extremos considerados até este ponto, professor benevolente e professor exigente. Esta extensão pode ser facilmente realizada, por exemplo, através de uma função objectivo para o professor do tipo:

$$S = (1-\psi)(e_1+e_2) + \psi\left[\mathcal{E}[N] - \frac{1}{2}\gamma\sigma_N^2\right], \qquad \psi \in [0,1] \tag{19}$$

[9] Admite-se que $\bar{W}$ é suficientemente baixo para que a restrição de participação não seja activa. No caso de cursos obrigatórios, esta é uma hipótese sem grandes implicações. No caso de cursos facultativos, o aluno poderá dispor da possibilidade de frequentar outros cursos.

O tratamento do problema nesta forma geraria uma solução que teria aspectos de ambas as teorias, estando mais próxima de uma ou de outra consoante o valor de ψ estivesse mais próximo da unidade ou de zero. Dado que a apresentação deste caso geral não fornece intuição adicional para o entendimento do problema tratado, não é realizada.

4 Alguma evidência empírica

Nas secções anteriores identificaram-se duas hipóteses teóricas concorrentes para a existência de testes intermédios no sistema de avaliação de cursos ministrados no ensino superior.[10] A estes testes intermédios é atribuida tipicamente menor ponderação que ao exame final.

O passo seguinte da análise é, naturalmente, a confrontação das duas hipóteses em termos empíricos. Esta tarefa é particularmente difícil, devido ao desconhecimento das funções esforço dos alunos.

Vamos por este motivo avaliar apenas se a primeira hipótese, denominada *professor benevolente*, recebe algum suporte empírico ou não. Esta opção resulta da facilidade de falsificação das premissas que dão apoio à teoria do professor benevolente: maior variância na nota do teste que na nota do exame.

É necessário contudo alguma clarificação da metodologia usada. Em particular, deve-se reconhecer que os alunos diferem, sendo que a comparação das variâncias de notas de teste e exame deve, de algum modo, ser controlado para essas diferenças. Em concreto, considera-se que a nota do aluno é formada por cinco elementos. Primeiro, uma média geral da escola. A essa média geral é adicionado um termo específico ao aluno que representa a sua capacidade individual. Se esse efeito for positivo, o indivíduo é um aluno acima da média e se for negativo, encontra-se abaixo da média. Este termo é distribuído aleatoriamente pelos alunos, segundo uma distribuição com média nula e variância finita. Para além destes dois efeitos, é adicionado um efeito específico ao curso, que pode ser interpretado como a sua dificuldade relativa, bem como um elemento associado com a natureza de cada prova em particular (teste

[10] Embora os argumentos sejam directamente aplicáveis a outros graus de ensino, o facto de a aplicação se reportar ao ensino superior leva-nos a focar nesta interpretação.

versus exame). Este efeito possui também média nula e variância finita. Finalmente, existe um elemento aleatório puro em cada prova de cada cadeira que cada aluno realiza.

Formalmente, admita-se que a nota do aluno i na prova j do curso k é:

$$N_{ik}(j) = \xi + \mu_j + \eta_i + \nu_k + \varepsilon_{kij} \tag{20}$$

sendo ξ a nota esperada dos alunos, μ_j um efeito específico à prova, η_i um efeito (aleatório) específico ao aluno i, ν_k um efeito (aleatório) específico ao curso k e ε_{kij} um efeito aleatório puro, cuja variância depende da prova em questão.

Admitem-se como hipóteses técnicas média nula para os efeitos específicos e para o termo aleatório ε. Os efeitos μ_j, η_i, ν_k e ε_{kil} têm variâncias finitas $\sigma_\mu^2, \sigma_\eta^2, \sigma_\nu^2$ e $\sigma_\varepsilon^2(j)$ respectivamente. Não existem correlações cruzadas entre qualquer par de efeitos.

Para um determinado curso k a variância das notas atribuídas em teste ($j = T$) e exame ($j = E$) são:

$$\sigma_k^2(T) = \sigma_\eta^2 + \sigma_\varepsilon^2(T) \tag{21}$$

$$\sigma_k^2(E) = \sigma_\eta^2 + \sigma_\varepsilon^2(E) \tag{22}$$

Como é aparente, a variância das notas depende não só do efeito aleatório puro como da heterogeneidade entre alunos. Apenas o primeiro efeito é relevante em termos de confronto de dados com a teoria do professor benevolente. Contudo, a diferença entre as variâncias das duas provas é independente do efeito específico de cada aluno:

$$\Delta = \sigma_N^2(T) - \sigma_N^2(E) = \sigma_\varepsilon^2(T) - \sigma_\varepsilon^2(E) \tag{23}$$

Deste modo, a hipótese que suporta a teoria do professor benevolente é $\Delta > 0$. A rejeição desta condição levaria a atribuir a estrutura de ponderações de teste e exame observadas à teoria do professor exigente. Se os dados disponíveis não entrarem em contradição com a condição $\Delta > 0$, ficamos nas condições iniciais em termos de interpretação. Isto é, a observação empírica é compatível com qualquer das duas teorias.

Para além da avaliação em termos das variâncias, que corresponde ao modelo mais simples apresentado, é conveniente fazer igualmente uma avaliação das diferenças de médias de notas entre provas. A este respeito, não oferece qualquer dificuldade mostrar que

$$\mathcal{E}[N_{ikE}] - \mathcal{E}[N_{ikT}] = \mu_E - \mu_T \tag{24}$$

Com base nos valores das médias e variâncias para o teste e exame e usando a ponderação atribuída ao teste, infere-se o grau de aversão absoluta ao risco (γ) compatível com a situação observada corresponder a um equilíbrio da teoria do professor benevolente.

Em relação à teoria do professor exigente, embora não seja realizável um teste formal, é possível a apresentação de evidência informal. Dos valores das notas de teste e de exame, admitindo como válidas as formas funcionais especificadas, podem-se recuperar os valores implícitos de θ e φ/β da situação de equilíbrio. Com base na hipótese de que o custo de esforço de estudo para exame é pelo menos igual ao de estudo para o teste ($\varphi \geq \theta$) e maior dificuldade intrínseca do exame ($\beta < 1$), uma implicação imediata é $(\varphi/\beta)/\theta > 1$.

A análise empírica é concretizada do seguinte modo. Admita-se que a ponderação dada ao exame (α) é escolhida de forma óptima. Considerem-se de seguida os valores dos parâmetros estruturais de cada uma das teorias que são implicados pelo facto de α observado ser óptimo. Se resultar um grau de aversão ao risco, γ, negativo, então não há consistência com a teoria do professor benevolente. De modo similar, se o valor implícito de $(\varphi/\beta)/\theta$ for inferior a 1, não se verifica consistência com a teoria do professor exigente. Os valores de γ são obtidos por uma inversão da expressão (12) . Os valores de $(\varphi/\beta)/\theta$ resultam da substituição dos valores óptimos de esforço na expressão da nota esperada do aluno. Concretamente, no caso do professor benevolente,

$$\gamma = \frac{\mu_E - \mu_T}{\alpha\sigma_E^2 - (1-\alpha)\sigma_T^2} \tag{25}$$

sob a hipótese de independência das notas entre provas.[11] Para a teoria do professor exigente.

$$\frac{\varphi/\beta)}{\theta} = \frac{\alpha(20 - \mu_E}{(1-\alpha)(20-\mu_T)} \tag{26}$$

[11]Note-se que se realizam hipóteses adicionais em relação às necessárias para o teste simples $\Delta > 0$.

Quadro 1: Dados sobre avaliação em cursos do ensino superior.

Curso	Variância			Média			α	N	γ	$\frac{\varphi/\beta}{\theta}$	tipo do professor
	Teste	Exame	Dif.	Teste	Exame	Dif					
100	24.44	14.02	10.43	12.05	13.26	-1.31	0.6	13	> 0	> 1	benev/exig
110*	6.42	2.53	3.88	13.57	10.07	3.50	0.5	14	< 0	> 1	exigente
170*	15.94	21.17	-5.23	10.74	11.54	-0.80	0.5	13		< 1	benev/exig
305*	21.66	11.43	10.23	12.48	9.33	3.14	0.6	21	< 0	> 1	exigente
310	3.88	16.69	-12.81	14.97	11.10	3.87	0.8	201	< 0	> 1	exigente
501	1.33	2.04	-0.71	14.28	15.04	-0.77	0.5	47	> 0	< 1	benevolente
722	3.92	7.83	-3.91	12.53	12.97	-0.43	0.5	43	> 0	< 1	benevolente
724	9.59	7.66	1.93	15.24	13.11	2.12	0.5	21		> 1	benev/exig
742	4.36	13.86	-9.50	12.00	9.96	2.04	0.5	23	< 0	> 1	exigente
784	2.97	6.85	-3.89	13.72	12.42	1.31	0.6	90	< 0	> 1	exigente

* – teste (ou trabalho) facultativo; Dif – diferença entre os valores de teste e de exame;
N – número de observações; α – ponderação atribuida ao exame.

No quadro seguinte, são apresentadas as variâncias e as médias de teste e exames numa amostra de cursos com teste, bem como os valores para o grau de aversão absoluta ao risco, γ, e para $(\varphi/\beta)/\theta$.[12] Na última coluna indica-se qual o tipo de professor consistente com os valores apresentados nas duas colunas imediatamente anteriores. Os cursos considerados respeitam ao segundo semestre lectivo do ano 1994/95 da Licenciatura em Economia da Faculdade de Economia da Universidade Nova de Lisboa. Os cursos encontram-se identificados pelo seu número de código.[13]

Numa primeira observação aos valores do Quadro 1, pode-se dizer que não há um padrão decisivo relacionando a ponderação dada ao exame com o diferencial entre teste e exame de médias ou variâncias.[14]

Como forma de tentar ultrapassar esta ambiguidade, são calculados valores para os dois

[12]Esta amostra foi determinada com base no critério de o curso ter teste intermédio e de a nota final ser facilmente replicável com base em elementos objectivos. Foram, por esta razão, omitidos os cursos com elementos de avaliação como a 'participação nas aulas' e os 'trabalhos de grupo'. No caso de cursos com testes facultativos, a escolha de ir ou não ao teste constitui um factor de selecção da amostra (concretamente, apenas os alunos com valor η_i elevado escolhem ir ao teste no caso de aversão ao risco decrescente).

[13]Os valores numéricos são facilmente calculados a partir dos elementos constantes do Quadro 1. Podem ser obtidos por pedido aos autores.

[14]A definição de um padrão é dificultada pelo pequeno número de cursos nas condições definidas. Veja-se a nota 12.

indicadores mensuráveis de cada uma das teorias com base nos dados apresentados.

Os resultados deste exercício de calibragem revelam uma situação quase paritária no suporte (não contradição) da teoria do professor benevolente e da teoria do professor exigente. Verifica-se também que a maioria dos cursos são consistente apenas com uma das explicações para a ponderação atribuida ao exame. O leitor mais atento poderia pensar que a própria estrutura dos modelos levaria a que os dados não seriam incompatíveis com ambos os modelos. Contudo tal não é o caso. Existem valores para as médias e variâncias que originam consistência ou inconsistência com ambas as teorias. Por exemplo, o curso 100 não entra em contradição com nenhuma das teorias. Não é, pois, trivial que não respeitando as condições impostas por uma teoria se respeitem as condições de consistência interna da outra teoria.

Sendo a hipótese de partida para a análise empírica a de α ser escolhido de forma óptima, há um factor adicional a considerar. Existe na Faculdade de Economia da Universidade Nova de Lisboa uma regra de avaliação que impede uma ponderação de exame inferior a 50%.

Do ponto de vista das implicações testáveis da teoria do professor exigente, esta regra não tem qualquer relevância. Já quanto à teoria do professor benevolente é necessário um pouco mais de cuidado, pois não se pode ignorar a possibilidade de ser óptimo estabelecer uma ponderação de exame inferior a 50%. Nestes casos, a regra imposta constitui uma restrição activa ao problema e tende a actuar contra encontrar-se evidência de professor benevolente. Assim, se mesmo com a possibilidade de a restrição ser activa ($\alpha = 0.5$), o professor é considerado benevolente, não há alteração da classificação do professor. Porém, nos casos em que há consistência com a teoria do professor exigente e se observa $\alpha = 0.5$ (cursos 724 e 110), coloca-se a questão de os dados também serem consistentes com uma interpretação de professor benevolente limitado pela regra geral da escola.

Nos dois casos de interesse (cursos 110 e 724), $\sigma_T^2 - \sigma_E^2 > 0$ e $\mu_T - \mu_E > 0$ que geram um valor máximo de γ positivo (e portanto admissível),[15] que é bastante superior aos valores implícitos de γ detectados nas restantes situações de acordo com a teoria do professor benevolente. Há, portanto, concordância com a interpretação do professor benevolente restrito

[15]os valores críticos são 1.8 para o curso 110 e 2.2 para o curso 724.

pela regra de $\alpha \geq 0.5$. Esta interpretação origina uma solução em que o multiplicador de Lagrange associado à restrição ser activa é dado por:[16]

$$\lambda = \mu_T - \mu_E - \frac{\gamma}{2}\left(\sigma_T^2 - \sigma_E^2\right) > 0 \tag{27}$$

o que permite identificar restrições sobre o parâmetro γ.

A evidência apresentada sugere assim que ambas as justificações para a presença de um teste intermédio estão presentes na realidade em maior ou menor grau em cada curso (consoante as características do curso e/ou do professor).

5 Conclusões

O presente artigo procurou analisar os fundamentos da prática observada da existência de testes intermédios como instrumento de avaliação em cursos superiores. Embora o texto esteja escrito em termos de avaliação em cursos de ensino superior, a análise é igualmente aplicável a outros graus de ensino.

As duas teorias concorrentes identificadas foram denominadas *teoria do professor benevolente* e *teoria do professor exigente*.

No primeiro caso, a existência de testes intermédios é determinada pelo desejo de diminuir o risco defrontado pelos alunos através da diversificação de elementos de avaliação. A condição a impor para que se observe o resultado de testes intermédios motivarem uma redução no risco na nota do aluno é facilmente testável empiricamente: a variância nas notas de teste devem ser inferiores à existente nas notas de exame.

A teoria alternativa, do professor exigente, explica a existência de testes intermédios como forma de induzir um maior nível de estudo por parte dos alunos. O teste empírico desta hipótese é claramente mais difícil de executar, pois envolve a necessidade de conhecer o custo de esforço a realizar pelo aluno. É contudo possível apresentar um indicador simples de consistência interna.

[16] Consultar o anexo para a derivação deste resultado.

Como forma parcial de avaliar o que os dados de uma amostra de cursos nos transmite, foi investigado se a condição da teoria do professor benevolente se verifica. O cálculo do grau de aversão absoluta ao risco, média e variância das notas de teste e exame revelou que os fundamentos da teoria do professor benevolente para a prática de uma ponderação superior do exame na nota final estão em contradição com uma parte da (pequena) amostra de resultados de cursos do ensino superior. De modo similar, a teoria do professor exigente é consistente com uma subamostra (complementar à que não contradiz a teoria do professor benevolente).

Em suma, apesar de não se ter concretizado um teste de validade de cada uma das teorias justificativas da existência de um teste intermédio como elemento de avaliação (que tem peso inferior ao exame final), a evidência que foi possível reunir sugere a presença de elementos de ambas as teorias.

Identificada a relevância qualitativa dos elementos de aversão ao risco por parte do aluno e de indução do esforço de estudo, fica como desafio para investigação futura uma análise pormenorizada sobre a importância quantitativa destes elementos. A sua compreensão e avaliação quantitativa poderão sugerir a adopção (ou não) de testes intermédios de uma forma mais generalizada, bem como a ponderação que lhes deverá estar associada.

Referências

LAFFONT, Jean-Jacques (1988). *Economics of Uncertainty and Information.* MIT Press.

RASMUSEN, Eric (1989). *Games and Information – An introduction to game theory.* MacMillan.

Anexo

O problema do professor com a inclusão da restrição de a ponderação a atribuir ao exame não poder ser inferior a 0.5 é:

$$\max_{\{\alpha,\lambda\}} \mathcal{L} = \alpha\mu_E + (1-\alpha)\mu_T - \frac{1}{2}\gamma\left(\alpha^2\sigma_E^2 + (1-\alpha)^2\sigma_T^2 + 2\rho(1-\alpha)\alpha\sigma_E\sigma_T\right) + \lambda\left(\alpha - \frac{1}{2}\right) \quad (28)$$

As condições de primeira ordem deste problema admitindo que a restrição é activa são:[17]

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \alpha} &= \mu_E - \mu_T - \gamma\left(\alpha\sigma_E^2 - (1-\alpha)\sigma_T^2 + \rho(1-2\alpha)\sigma_E\sigma_T\right) + \lambda = 0 \\
\frac{\partial \mathcal{L}}{\partial \lambda} &= \alpha - \frac{1}{2} = 0
\end{aligned}$$

De onde resulta

$$\lambda = \mu_T - \mu_E - \frac{\gamma}{2}\left(\sigma_T^2 - \sigma_E^2\right) \tag{29}$$

Para o problema fazer sentido, $\lambda > 0$, o que tem implicações testáveis para o valor de γ.

[17] O caso de a restrição não ser activa já foi tratado.