

## DESEMPENHO EDUCATIVO E IGUALDADE DE OPORTUNIDADES EM PORTUGAL E NA EUROPA: O PAPEL DA ESCOLA E A INFLUÊNCIA DA FAMÍLIA\*

Manuel Coutinho Pereira\*\*

### 1. INTRODUÇÃO

A economia da educação está enraizada nas teorias económicas que estudam o capital humano. Um dos seus ramos preocupa-se com a compreensão do que acontece nas escolas, ou seja, com o modo como o desempenho interage com uma multiplicidade de fatores como as características e o contexto familiar dos alunos, os recursos escolares e as características institucionais dos sistemas educativos. Tal relaciona-se com várias agendas de investigação; em primeiro lugar, a que tem a ver com a eficácia na utilização dos recursos. A educação representa uma parcela grande da despesa pública em quase todos os países. Assim, é importante saber se um aumento da despesa levará a uma melhoria do desempenho, ou se um determinado nível de desempenho pode ser alcançado gastando menos. A literatura sugere que as políticas de recursos puras não são eficazes, a menos que sejam acompanhadas por alterações nos incentivos. Neste contexto, é importante considerar as características institucionais, por exemplo, a autonomia das escolas, entre os determinantes dos resultados. Outra agenda de investigação decorre do facto estilizado segundo o qual os antecedentes familiares têm um papel-chave no sucesso escolar, repetidamente confirmado pela literatura desde a publicação do conhecido relatório Coleman (Coleman, 1966). Um elevado grau de dependência do contexto socioeconómico por parte dos resultados é, contudo, uma característica indesejável dos sistemas de ensino pois, nesse caso, a escolaridade não contribui suficientemente para atenuar a desigualdade social. No entanto, intervenções neste domínio só poderão funcionar se os mecanismos subjacentes forem bem compreendidos. Por exemplo, o nível socioeconómico pode não determinar diretamente os resultados, mas antes as escolas em que os alunos ingressam, e a desigualdade pode ocorrer principalmente entre as escolas.

O objetivo deste trabalho é reunir evidência empírica sobre estas questões para Portugal e para vários países da União Europeia, com base nos dados disponibilizados pelo *Programme for International Student Assessment (PISA)* da OCDE de 2006. O PISA compreende testes que avaliam os conhecimentos dos alunos em matemática, ciências e leitura no final da escolaridade obrigatória, nos países participantes. Os testes medem a capacidade dos alunos para utilizarem os conhecimentos adquiridos em situações que ocorrem no mundo real, ao invés da aprendizagem de currículos específicos. O PISA é um programa contínuo que tem sido administrado em ciclos de três anos;

\* A parte inicial da pesquisa que levou a este artigo foi um trabalho conjunto com Sara Moreira. O autor agradece os comentários de Nuno Alves, Mário Centeno, Jorge Correia da Cunha, Ana Cristina Leal e José Ferreira Machado. As opiniões expressas no artigo representam as opiniões do autor e não coincidem necessariamente com as do Banco de Portugal ou do Eurosistema. Eventuais erros e omissões são da exclusiva responsabilidade do autor.

\*\* Banco de Portugal, Departamento de Estudos Económicos.

até à data, foi realizado em 2000, 2003, 2006 e 2009. Este estudo utiliza os dados do último ciclo disponível aquando da sua redação - o de 2006<sup>1</sup>. Para além dos países da OCDE, um número crescente de países parceiros tem participado no Programa (em 2006, este estendeu-se a 57 países).

Os programas internacionais de avaliação dos alunos, como o PISA, são instrumentos importantes para controlar a qualidade dos sistemas educativos, na medida em que disponibilizam dados cujos processos de recolha e definições são comparáveis entre países. De facto, os resultados do PISA foram exaustivamente debatidos pelos meios de comunicação e os diferentes intervenientes no processo educativo, em alguns países participantes (ver Fuchs e Woessmann, 2007). Ao mesmo tempo, os dados resultantes destes programas são um elemento fundamental para a análise empírica em economia da educação (Hanushek e Woessmann, 2010). A base de dados do PISA inclui, além dos resultados dos testes, muita informação sobre os alunos e o seu contexto familiar, as características das escolas e os recursos escolares. Uma vantagem óbvia deste tipo de dados relativamente aos dados nacionais é a possibilidade de se avaliar a importância dos impactos estimados por comparação com as estimativas para outros países. Por exemplo, os coeficientes das variáveis relativas ao contexto familiar em regressões que explicam os resultados dos testes podem ser vistos como indicadores de igualdade de oportunidades. Além disso, para algumas variáveis, como as relacionadas com o desenho institucional dos sistemas de ensino, a variação dentro do país é normalmente escassa ou nula. Assim, as bases de dados internacionais são importantes para que os investigadores possam identificar a influência de tais variáveis.

Neste estudo são estimadas funções de produção de educação, regredindo o desempenho dos alunos - medido pela pontuação nos testes - sobre um amplo conjunto de variáveis explicativas. Estas compreendem ao nível do estudante, por exemplo, o género, a idade, o ano de escolaridade, as habilitações literárias e a profissão dos pais, a condição de imigrante, e indicadores de riqueza e recursos educativos em casa. Ao nível da escola, as variáveis disponíveis incluem o rácio aluno/professor, a dimensão das turmas, medidas de escassez de professores, a dimensão e localização das escolas, a sua natureza pública ou privada, e indicadores de autonomia. São estimadas separadamente funções de produção de educação para as pontuações em matemática e leitura. A fim de garantir uma maior homogeneidade socioeconómica e cultural dentro do grupo de países considerados e minimizar a importância dos fatores omitidos, o estudo confina-se aos países da OCDE participantes no programa que pertencem à União Europeia (com exceção da França, porque os seus dados para as escolas não estão disponíveis).

Consoante as conveniências da análise, este estudo centra-se em todos os países considerados ou em grupos dos mesmos. Por exemplo, as funções de produção de educação são estimadas para Portugal e, respetivamente, para os grupos de três países em que os alunos têm melhores e piores pontuações. Torna-se assim possível avaliar em que medida as diferenças no desempenho (tanto quanto o modelo o consegue explicar) têm a ver com diferenças na contribuição das diversas variáveis, ou seja, com os coeficientes da função de produção. Esta abordagem difere da seguida

(1) Os resultados do PISA 2009 foram disponibilizados após a conclusão do estudo.

em outros estudos (por exemplo, Woessmann, 2003), que estimam uma função de produção para todos os países, a fim de explorar a variabilidade entre-países nos regressores. Com regressões deste tipo, é também possível avaliar se, após controlar para um conjunto de variáveis explicativas, as diferenças iniciais no desempenho entre os países permanecem. Por exemplo, pode-se investigar até que ponto o baixo nível de escolaridade em Portugal explica a diferença para a média das pontuações dos alunos portugueses, dada a importância das habilitações dos pais para o sucesso escolar. O estudo termina com uma análise da variabilidade das pontuações e do papel que nesta desempenha a desigualdade social. Neste contexto, dado que os alunos estão agrupados por escolas, é importante verificar se a variabilidade do desempenho é um fenómeno entre-escolas ou intraescolas. O estudo inicia-se por uma digressão descritiva pela base de dados do PISA 2006.

## 2. A BASE DE DADOS DO PISA 2006 E ALGUNS RESULTADOS DESCRITIVOS

Os testes do PISA são realizados por uma amostra representativa da população de estudantes com cerca de 15 anos<sup>2</sup> que frequentam as escolas num determinado país e se encontram no 7.º ano de escolaridade ou acima. Em geral, o desenho da amostra assume a forma de dois níveis de amostragem estratificada, sendo as escolas sorteadas aleatoriamente numa primeira fase, e os estudantes em cada uma destas, numa segunda fase. Por exemplo, para Portugal, 173 escolas foram selecionadas em primeiro lugar, e 40 alunos (ou todos os alunos elegíveis, quando menos de 40) foram posteriormente escolhidos aleatoriamente em cada uma destas. Os dados do PISA, para além das pontuações nos testes, resultam essencialmente de dois questionários preenchidos, respetivamente, pelos alunos e pelas escolas. A base de dados de 2006 abrange 5 109 alunos em Portugal e 131 598 alunos no conjunto de países considerados.

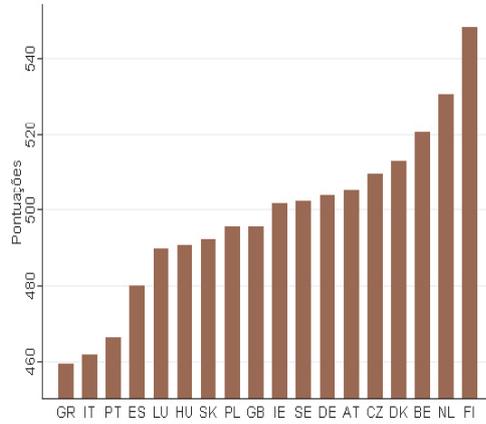
Os estudantes incluídos na amostra do PISA não são uniformemente representativos da população, pelo que a base de dados é fornecida com ponderadores finais por aluno, os quais refletem as probabilidades de amostragem e outros fatores, como as não-respostas. Por outro lado, as pontuações dos alunos no PISA são reportadas sob a forma de cinco *valores plausíveis* para cada matéria - leitura, matemática e ciências - os quais correspondem a extrações aleatórias da distribuição estimada de capacidades de cada aluno (ver OCDE, 2009, Capítulo 6). As estatísticas para a população, incluindo estatísticas descritivas e parâmetros de modelos, são calculadas como a média das estatísticas correspondentes obtidas separadamente para cada valor plausível. No cálculo destas últimas são utilizados os ponderadores por estudante. A variância das estatísticas para a população resulta de duas fontes: a variância amostral e a chamada variância de imputação que reflete o erro de medição nos testes (OCDE, 2009, Capítulo 8).

Apresenta-se em primeiro lugar a média das pontuações em matemática e leitura (Gráficos 1A e 1B) para o conjunto dos países considerados, o qual inclui a Alemanha, a Áustria, a Bélgica, a Dinamarca, a Eslováquia, a Espanha, a Finlândia, a Grécia, a Hungria, a Irlanda, a Itália, o Luxemburgo,

(2) Mais precisamente, os alunos têm entre 15 anos e 3 meses e 16 anos e 3 meses de idade. Conforme explicado abaixo, tais pequenas diferenças de idade têm um impacto sobre o desempenho, pois, juntamente com o ano de escolaridade, podem captar um efeito de repetência.

**Gráfico 1A**

**DESEMPENHO EM MATEMÁTICA POR PAÍS**  
Pontuação média

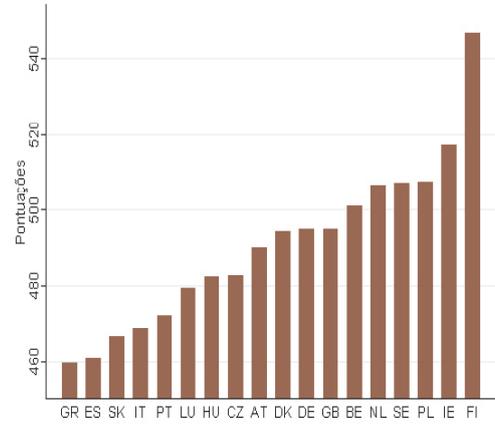


**Fonte:** Cálculos do autor com base nos dados do PISA 2006.

**Nota:** Média das médias ponderadas calculadas para cada valor plausível.

**Gráfico 1B**

**DESEMPENHO EM LEITURA POR PAÍS**  
Pontuação média



**Fonte:** Cálculos do autor com base nos dados do PISA 2006.

**Nota:** Ver gráfico anterior.

os Países Baixos, a Polónia, Portugal, a República Checa, a Suécia e o Reino Unido. Note-se que as pontuações dos testes são normalizadas para uma média de 500 e um desvio padrão de 100 nos países da OCDE. No PISA 2006, os alunos portugueses ficaram no 16. lugar em matemática e no 14. lugar em leitura entre os alunos destes 18 países; estes maus resultados não diferem muito dos obtidos em ciclos anteriores<sup>3</sup>.

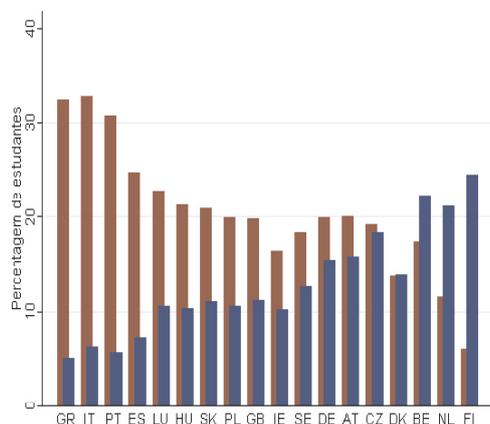
A fim de confrontar os resultados para Portugal com os de outros países, foram selecionados (dando peso igual a ambas as matérias consideradas), dois grupos com os três países, respetivamente, com o melhor e o pior desempenho. O primeiro grupo inclui a Bélgica, a Finlândia e os Países Baixos, e o segundo a Espanha, a Grécia e a Itália. Portugal tem níveis de desempenho semelhantes aos destes últimos países, e também partilha com eles alguns traços socioeconómicos e culturais. As distribuições das pontuações nos testes (não apresentadas) são quase coincidentes para Portugal e para o grupo com baixo desempenho. As distribuições para o grupo com alto desempenho encontram-se comparativamente deslocadas para a direita, mas a sua dispersão é comparável. Por exemplo, o intervalo interquartil dividido pela média é 0.27 em Portugal e nos países de baixo desempenho, para as pontuações em matemática, e 0.25 nos países de alto desempenho; para as pontuações em leitura, estes valores estão entre 0.26 e 0.28. O espetro de resultados no PISA encontra-se dividido em seis níveis sucessivos de proficiência, associados a uma maior dificuldade das tarefas que o aluno deve executar. Os Gráficos 2A e 2B apresentam a proporção de alunos, respetivamente, no nível 1 e abaixo e no nível de 5 e acima, com os países ordenados de acordo com a pontuação média. À medida que esta pontuação aumenta, a proporção de alunos nos níveis mais baixos de proficiência tende a descer, e nos níveis mais altos tende a subir. No entanto, alguns países, como a Alemanha, a Áustria, a Bélgica, e a República Checa, têm uma grande proporção

(3) No PISA 2009, o desempenho dos estudantes portugueses melhorou bastante, particularmente, em leitura, área em que a pontuação média não difere significativamente da média da OCDE em termos estatísticos. Portugal ocupa a 11. e a 15. posições, respetivamente, em leitura e matemática no mesmo grupo de 18 países, no PISA mais recente.

Gráfico 2A

## PROFICIÊNCIA EM MATEMÁTICA POR PAÍS

Alunos no nível 1 e abaixo (a vermelho) e no nível 5 e acima (a azul)



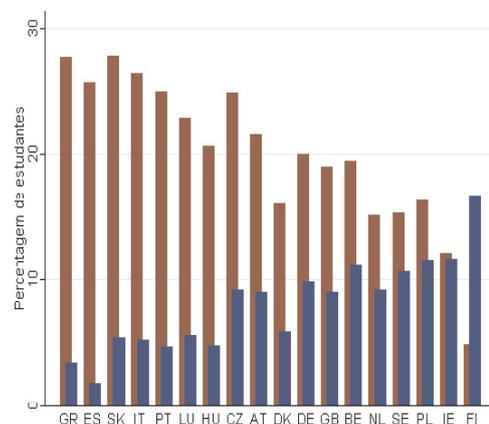
Fonte: Cálculos do autor com base nos dados do PISA 2006.

Nota: Média das percentagens para cada valor plausível.

Gráfico 2B

## PROFICIÊNCIA EM LEITURA POR PAÍS

Alunos no nível 1 e abaixo (a vermelho) e no nível 5 e acima (a azul)



Fonte: Cálculos do autor com base nos dados do PISA 2006.

Nota: Ver gráfico anterior.

de estudantes em níveis de proficiência baixos dado o desempenho médio do país, o que indica uma maior dispersão das pontuações (esta questão é retomada na secção 4).

O Quadro 1 mostra as características da população estudantil, das suas famílias e das escolas em Portugal, distinguindo entre escolas públicas e privadas, e nos dois grupos de países de referência. Estas constituem as variáveis incluídas nas funções de produção de educação que serão estimadas na próxima secção. Nos países onde a idade obrigatória para o início da escolaridade é aos seis anos, os alunos encontram-se maioritariamente distribuídos entre o 9. e o 10. ano (refletindo normas específicas quanto à data de nascimento). Tal é o caso de Portugal e de todos os países dos dois grupos, exceto a Finlândia, onde a escola começa aos sete anos e os alunos se concentram no 9. ano<sup>4</sup>. Um número razoavelmente elevado de alunos portugueses - cerca de 20 por cento - ainda está no 7. ou no 8. ano de escolaridade, devido a maiores taxas de retenção.

No que respeita ao contexto familiar, o hiato entre Portugal e os países com melhor desempenho é particularmente grande para as habilitações literárias e a profissão dos pais. Por exemplo, em Portugal cerca de 25 por cento alunos têm pelo menos um dos pais numa ocupação intelectual/especializada, e cerca de 20 por cento têm pelo menos um dos pais com o ensino superior completo. Estas percentagens são cerca de 60 por cento para os países onde os alunos têm melhor desempenho. Além disso, menos de 5 por cento dos estudantes nesses países reportam que o nível mais alto de habilitações dos pais é a educação primária, contra quase 40 por cento dos alunos portugueses. A mesma situação ocorre, embora em menor grau, relativamente aos estudantes dos países com pior desempenho. O indicador de riqueza (calculado a partir das respostas sobre

(4) Mais geralmente, no conjunto de países considerados, a idade obrigatória para o início da escola vai dos quatro ou cinco anos no Reino Unido aos sete anos na Dinamarca, na Finlândia, na Polónia e na Suécia.

Quadro 1

	Portugal			Países com baixo desempenho <sup>(b)</sup>	Países com alto desempenho <sup>(b)</sup>
	Total	Escolas públicas	Escolas privadas		
<b>VARIÁVEIS EXPLICATIVAS (MÉDIAS)<sup>(a)</sup></b>					
<b>Características do estudante</b>					
7. ano de escolaridade	6.6	6.9	3.9	0.2	0.2
8. ano de escolaridade	13.1	13.5	9.9	3.7	5.3
9. ano de escolaridade	29.5	29.7	27.6	21.0	47.5
10. ano de escolaridade	50.9	50.0	58.6	72.4	46.5
11. ano de escolaridade				2.8	0.5
idade (anos)	15.7	15.7	15.7	15.8	15.8
género feminino	51.7	52.1	48.3	50.0	48.8
<b>Contexto familiar</b>					
riqueza (índice) $\subset [-2.1, 2.3]$ <sup>(c)</sup>	-0.17	-0.20	0.11	-0.14	0.45
recursos educativos em casa (ind.) $\subset [0, 7]$	6.2	6.2	6.4	6.1	6.2
livros em casa < 25	38.9	40.0	29.2	24.7	28.0
livros em casa 25-200	45.6	45.6	46.3	53.2	48.3
livros em casa > 200	15.5	14.5	24.5	22.1	23.7
não imigrante	94.1	93.7	97.4	94.6	89.7
imigrante de segunda geração	3.5	3.8	1.3	4.6	4.0
imigrante de primeira geração	2.4	2.5	1.2	0.8	6.3
língua do teste em casa	97.7	97.6	98.1	85.9	90.9
outra língua nacional em casa	-	-	-	11.2	4.0
língua estrangeira em casa	2.3	2.4	1.9	2.9	5.1
<b>ocupação mais elevada dos pais</b>					
manual/não especializada	12.6	13.2	7.6	11.8	7.1
manual/especializada	24.2	24.9	18.7	19.0	9.4
intelectual/não especializada	36.3	36.6	33.4	23.7	22.4
intelectual/especializada	26.9	25.4	40.3	45.5	61.2
<b>escolaridade mais elevada dos pais</b>					
primária ou menor	38.6	39.2	33.5	7.9	3.6
secundária inferior	15.3	15.4	14.8	22.3	5.4
secundária superior	23.5	24.0	19.6	38.5	32.3
terciária	22.5	21.5	32.1	31.2	58.6
<b>Características da escola</b>					
dimensão da escola (1000 alunos)	0.957	0.922	1.268	0.694	0.824
proporção de alunos	50.9	51.0	49.7	49.7	48.9
localizada em localidade < 15 000 pessoas	42.5	40.8	57.9	24.9	25.6
localizada em localidade 15 000 - 100 000 pess.	35.9	38.2	15.1	42.0	51.0
localizada em localidade > 100 000 pessoas	21.6	21.0	27.1	33.1	23.4
amplitude graus escolaridade (grau max-min)	5.1	4.7	8.3	4.9	4.8
proporção de repetentes	14.6	15.3	7.2	10.5	4.7
escola enfrenta concorrência <sup>(d)</sup>	72.9	71.9	81.7	78.4	84.4
autonomia recursos (ind.) $\subset [-1.1, 2.0]$	-1.0	-1.1	-0.8	-0.6	0.1
autonomia programas/aval. (ind.) $\subset [-1.4, 1.3]$	-0.5	-0.5	-0.3	0.1	0.4
escola recebe pressão dos pais <sup>(e)</sup>	7.1	4.9	26.5	16.1	7.4
escola pública	91.1			84.4	44.2
escola privada	8.9			15.6	55.8
<b>Recursos escolares</b>					
dimensão das turmas (alunos)	24.0	23.7	26.3	27.0	22.3
rácio aluno/professor	8.9	8.4	13.3	10.4	12.9
propor. de computadores com internet	80.2	79.7	84.5	84.9	89.1
rácio computador/aluno	0.07	0.07	0.06	0.11	0.15
aulas normais - matemática (horas)	3.5	3.5	3.8	3.5	3.2
aulas normais - língua do teste (horas)	3.2	3.2	3.4	3.9	3.1
escassez de professores matemática	1.3	0.6	7.7	10.5	21.1
escassez de professores língua do teste	0.0	0.0	0.0	9.2	13.5
<b>Familiaridade com computadores</b>					
autoconf. tarefas internet (ind.) $\subset [-2.9, 0.8]$	-0.2	-0.2	-0.2	-0.4	0.4
tempo gasto intern./divers. (ind.) $\subset [-2.1, 3.2]$	0.2	0.2	0.1	-0.1	0.2

Fonte: Cálculos do autor com base nos dados do PISA 2006.

Notas: (a) Médias ponderadas; valores em percentagem do total exceto quando indicado de outra forma (no Anexo 1 são dados mais detalhes sobre a definição das variáveis). (b) Os países com baixo desempenho incluem a Espanha, a Grécia e a Itália; os países com alto desempenho incluem a Bélgica, a Finlândia e os Países Baixos. (c) Os intervalos mostram o percentil 1. e 99. do indicador. (d) A escola reporta que existe pelo menos outra escola na mesma área que compete pelos seus alunos. (e) A escola reporta que existe pressão da maioria dos pais para melhoria dos padrões escolares.

a posse de bens duradouros) tem um maior valor médio para os países com melhor desempenho, como seria de esperar dado o facto de estes serem mais ricos. Em contrapartida, o indicador de recursos educativos em casa tem um nível semelhante em todos os países. A proporção de alunos imigrantes é ligeiramente maior nos países com melhor desempenho do que em Portugal e no grupo com pior desempenho (cerca de 10 por cento contra 5 por cento). Em Portugal, os alunos das escolas privadas provêm de famílias mais favorecidas do que os seus colegas das escolas públicas, o que é visível no indicador de riqueza, e particularmente nos escalões superiores das habilitações e profissão dos pais.

Passamos agora para as variáveis da escola. A proporção de escolas privadas difere substancialmente entre os países com melhor desempenho (acima de 50 por cento)<sup>5</sup> e Portugal e países com piores resultados (cerca de 10 por cento). As escolas em Portugal são maiores, localizadas em localidades relativamente mais pequenas e têm uma maior proporção de repetentes do que as escolas em ambos os grupos de referência. Entre os indicadores de recursos, é particularmente notável o baixo rácio aluno/professor nas escolas públicas em Portugal. As escolas portuguesas têm comparativamente menos autonomia na gestão dos recursos e definição dos programas e métodos avaliação, em especial, em relação às dos países com melhor desempenho. Curiosamente, as escolas nos dois grupos de referência reportam alguma escassez de professores qualificados, em contraste com as portuguesas. Note-se que na base de dados do PISA 2006 não há nenhuma medida da experiência do pessoal docente, pelo que esta não pôde ser incluída na função de produção da educação<sup>6</sup>.

### 3. A FUNÇÃO DE PRODUÇÃO DE EDUCAÇÃO

#### 3.1. Especificação e possíveis causas de enviesamento

A relação entre o desempenho e seus determinantes é muitas vezes estudada através da função de produção da educação - ver, por exemplo, Hanushek (1979) e Todd e Wolphin (2003) para discussões detalhadas sobre a especificação e estimação dessas funções. Estas podem ser genericamente especificadas como

$$Score_{ij} = b_0 + b_1 St_{ij} + b_2 Fm_{ij} + b_3 Sc_j + b_4 Re_j + e_{ij},$$

onde o índice  $i$  diz respeito ao aluno e o índice  $j$  à escola.  $Score$  é a pontuação no teste,  $St$  é um vetor de características do aluno,  $Fm$  inclui medidas do contexto familiar,  $Sc$  engloba diversas características da escola e  $Re$  inclui variáveis ligadas aos recursos escolares. As variáveis explicativas do desempenho dos alunos que consideramos neste estudo são enumeradas no Quadro 17. O principal problema que afeta a estimação das funções de produção de educação é a en-

(5) Note-se que a maioria dessas escolas tem uma gestão privada, mas financiamento público.

(6) A base de dados inclui a proporção de professores com ensino superior, mas esta variável tem variabilidade reduzida (a mediana está perto dos 90 por cento). Como, além disso, a mesma não existe para alguns países, não foi tida em conta nas regressões.

(7) Não existem observações em falta para as pontuações dos testes; pelo contrário, a maioria das variáveis explicativas tem uma pequena quantidade de observações em falta. A fim de evitar uma grande perda de informação, estas observações foram imputadas previamente à estimação, à semelhança de estudos anteriores (por exemplo, Fuchs e Woessmann, 2007). Os detalhes são apresentados no Anexo 2.

dogeneidade de alguns regressores, muitas vezes decorrente de sua correlação com variáveis não observáveis, nomeadamente as capacidades cognitivas dos estudantes. Na verdade, medidas dessas capacidades deveriam aparecer na equação acima, mas isso quase nunca acontece dada a sua natureza. Um dos regressores que consideramos - o indicador que mede a autoconfiança em tarefas associadas à *internet* (ver familiaridade com computadores no Quadro 1) - é de particular interesse a este respeito, porque pode ser visto como um indicador das capacidades dos estudantes<sup>8</sup>, embora possa refletir também as posses da família e/ou os recursos da escola, na medida em que estes condicionam o acesso dos alunos à *internet*. Um segundo problema diz respeito ao facto de os dados do PISA serem seccionais, possibilitando a estimação de modelos em níveis, mas não de especificações de «valor acrescentado», que requerem dados em painel.

O ano de escolaridade capta a exposição do aluno a programas mais ou menos avançados e, combinado com a idade do estudante, também um efeito de repetição do ano (ver a próxima secção). Note-se que a situação do estudante em termos de repetência não aparece na nossa regressão. Esta variável refletiria o desempenho passado, que está relacionado com o desempenho atual. Desta forma, o ano de escolaridade é, em parte, endógeno ao desempenho atual, podendo esperar-se uma sobreestimação do seu efeito (positivo) relativamente a uma regressão que também controlasse para a situação de repetência.

As variáveis explicativas respeitantes à condição sócioeconómica deverão ser, em larga medida, exógenas ao sucesso escolar<sup>9</sup> e ter uma contribuição constante para os conhecimentos adquiridos pelos estudantes ao longo dos anos, que é bem captada por uma modelação em nível. As características básicas da escola como a localização e a dimensão, e as características institucionais como a autonomia, podem também ser consideradas exógenas na função produção. Contudo, a pressão dos pais - incluída neste grupo - pode não o ser e captar, para além da influência exercida sobre as escolas para melhorarem os padrões, um efeito de seleção de melhores escolas pelos pais mais empenhados na educação dos seus filhos. Isso pode levar a uma sobreestimação de um eventual impacto positivo da variável sobre o desempenho. Dado que se controla extensivamente para a situação familiar e de imigração, mesmo que os alunos de famílias favorecidas predominem nas escolas privadas, isto não deverá causar um enviesamento na medida do efeito-escola privada.

As variáveis que medem a utilização de recursos pelas escolas, por exemplo, a dimensão das turmas e as horas de aulas normais, são claramente menos adequadas para a especificação em níveis usada. Com efeito, tais variáveis normalmente mudam de um ano para o outro, e o nível corrente de aprendizagem depende também dos valores que assumiram no passado. Com os dados do PISA esta dificuldade não pode ser ultrapassada. A quantidade de recursos pode, ela própria, responder ao desempenho dos alunos: por exemplo, os alunos com insucesso escolar podem ter aulas suplementares<sup>10</sup>. No entanto, dados os indicadores de recursos que utilizamos e o facto de

(8) Note-se que esta variável mede a competência em tarefas pouco relacionadas com o uso do computador como instrumento de aprendizagem (ver Anexo 1). Caso contrário, o seu poder explicativo seria menos interessante. Outra variável disponível na base de dados do PISA que mede a autoconfiança em tarefas informáticas gerais (não utilizada) parece mais propensa a este problema.

(9) Supondo que as capacidades cognitivas não estão correlacionadas entre as gerações, caso contrário os alunos mais capazes poderiam estar associados com as famílias favorecidas.

(10) Um exemplo de um indicador de recursos na base de dados do PISA fortemente afetado por este tipo de endogeneidade são as horas de aprendizagem fora da escola. Tentou-se incluí-lo na função de produção, mas o efeito positivo que poderá ter sobre o desempenho é mais do que compensado pela seleção de alunos de baixo desempenho que o mesmo também capta.

estes serem definidos em termos de médias por escola (exceto para as horas de aulas normais), é improvável que isso afete muito as nossas estimativas.

Ao estimar funções de produção de educação para vários países, há que contar com fatores que têm um impacto sobre os resultados escolares e que são específicos a cada país, por exemplo, as atitudes sociais em relação à escolaridade. Nas funções estimadas para os dois grupos de referência, tal é acomodado pela inclusão de variáveis binárias ao nível do país.

### 3.2. Determinantes do desempenho em Portugal e nos países de referência

#### *Características do aluno*

O Quadro 2 apresenta as estimativas das funções de produção de educação para Portugal e para os dois grupos de países onde os alunos têm, respetivamente, os piores resultados (Espanha, Grécia, e Itália) e os melhores (Bélgica, Finlândia e Países Baixos). Começamos pelo impacto das características dos alunos. Em Portugal, existe um efeito positivo do ano de escolaridade no desempenho, claramente significativo (o mesmo se verifica para os grupos de países de alto e baixo desempenho). Uma comparação com os coeficientes das outras variáveis binárias na regressão mostra que sua magnitude é muito grande. Para os estudantes portugueses do PISA 2006, o ano de escolaridade e a idade interagem do seguinte modo. No conjunto dos estudantes com uma dada idade, aqueles que se encontram no grau de escolaridade mais alto - que pode ser o 9. ou o 10. ano dependendo da data de nascimento<sup>11</sup> - nunca repetiram um ano, aqueles que se encontram um grau abaixo já repetiram um ano, e assim por diante (isso verifica-se para todos os alunos, exceto os nascidos entre 15 de setembro e 31 de dezembro, cujos pais podiam e decidiram adiar a entrada na escola por um ano). Assim, como se disse, a variável capta não só o impacto dos programas que o aluno está a aprender, mas também um efeito associado à repetência.

O poder explicativo da idade, com o grau de escolaridade constante, tem a ver com os alunos nascidos entre 15 de setembro e 31 de dezembro: um número crescente destes, à medida que as datas de nascimento se aproximam do final do ano, esperaram mais um ano para entrar na escola. Assim, com a diminuição da idade dentro deste grupo, e para cada grau de escolaridade exceto o 10. ano, a proporção de crianças que ingressaram na escola com seis anos aumenta e a dos alunos que repetiram uma vez (9. ano) ou mais um ano (graus abaixo) diminui, e daí a relação negativa com o desempenho. O género tem uma influência clara sobre as pontuações, tendo os alunos um melhor desempenho em matemática e as alunas um melhor desempenho em leitura. Este efeito é estimado com precisão e, como seria de esperar, similar entre países.

(11) De acordo com as regras que regem o início da escolaridade obrigatória, os alunos que nunca repetiram o ano encontram-se no 10. ano de escolaridade, se completaram seis anos de idade até 15 de setembro de 1996, no 9. ano, se completaram seis anos após 31 de dezembro, e num desses anos, dependendo da decisão dos pais, se completaram seis anos entre essas datas.

## Quadro 2 (continua)

FUNÇÃO DE PRODUÇÃO DE EDUCAÇÃO, ESTIMATIVAS <sup>(a)</sup>						
	Portugal		Países com baixo desempenho <sup>(b)</sup>		Países com alto desempenho <sup>(b)</sup>	
	Matemática	Leitura	Matemática	Leitura	Matemática	Leitura
<b>Características do estudante</b>						
ano de escolaridade (7.) <sup>(c)</sup>						
8.	39.4	53.2	12.1	54.4	51.4/73.0	90.2/67.7 <sup>(d)</sup>
	3.5	4.1	12.6	18.0	7.7/12.6	14.8/19.0
9.	83.1	95.7	58.8	93.5	86.0/124.6	132.1/111.8
	3.8	4.4	13.0	17.2	6.7/13.1	15.5/18.3
10.	144.1	158.7	115.2	144.1	142.2/-	186.1/-
	2.9	3.7	13.1	17.5	6.7/-	15.6/-
11.	-	-	119.7	150.3	202.6/-	245.2/-
			13.8	18.7	10.3/-	15.7/-
idade						
	-17.1	-23.4	5.5	0.5	-17.1	-18.8
	2.1	2.6	1.2	1.6	1.4	1.6
género feminino (masculino)						
	-26.9	20.9	-26.8	18.7	-24.3	16.9
	1.6	1.4	0.8	1.0	1.0	1.1
<b>Contexto familiar</b>						
riqueza						
	-3.5	0.3	-4.2	-7.9	-2.0	-3.8
	1.3	1.1	0.5	0.8	0.7	0.9
recursos educativos em casa						
	2.2	2.0	8.5	8.7	7.8	8.0
	0.7	1.0	0.4	0.4	0.5	0.6
livros em casa (menos do que 25)						
entre 25 e 200	19.6	15.4	21.8	23.2	22.2	21.4
	1.5	1.6	1.0	0.9	1.0	1.1
mais do de 200	35.4	22.8	46.2	39.9	49.0	42.7
	2.1	2.6	1.2	1.2	1.1	1.4
situação de imigração (não imigrante)						
imigrante de segunda geração						
	-17.4	-14.5	-7.1	-4.6	-24.0	-9.6
	3.3	6.5	2.0	2.1	2.9	3.1
imigrante de primeira geração						
	-15.1	-9.8	-15.9	8.8	-27.7	-18.5
	6.6	5.6	4.0	5.0	1.9	3.4
língua em casa (a do teste)						
outra língua nacional						
	-	-	-0.8	-3.2	26.3	28.1
			1.3	1.6	2.0	2.2
língua estrangeira						
	23.8	-10.1	12.2	-9.6	-1.2	-18.3
	5.2	4.9	3.4	3.2	3.1	2.9
ocup. mais elev. dos pais (man./não esp.)						
manual/especializada						
	1.2	2.6	4.0	0.4	4.5	7.2
	2.8	3.1	1.4	1.5	2.3	2.6
intelectual/não especializada						
	2.3	10.0	7.4	7.8	10.5	14.6
	2.8	3.2	1.1	1.3	1.7	2.3
intelectual/especializada						
	18.1	23.7	12.6	11.9	21.3	26.5
	2.7	2.9	1.2	1.5	2.2	2.5
escol. mais elev. dos pais (prim. ou menor)						
secundária inferior						
	-0.4	3.6	13.5	17.8	13.2	16.3
	2.3	2.6	1.8	2.1	5.2	3.7
secundária superior						
	2.4	3.0	17.6	25.2	10.2	18.6
	1.3	2.7	1.8	1.9	3.9	2.8
terciária						
	0.3	5.8	15.2	20.0	10.5	19.2
	2.6	2.9	1.7	1.6	3.9	2.9

## Quadro 2 (continuação)

FUNÇÃO DE PRODUÇÃO DE EDUCAÇÃO, ESTIMATIVAS <sup>(a)</sup>						
	Portugal		Países com baixo desempenho <sup>(b)</sup>		Países com alto desempenho <sup>(b)</sup>	
	Matemática	Leitura	Matemática	Leitura	Matemática	Leitura
<b>Características da escola</b>						
dimensão da escola	4.2	8.5	-2.6	-1.9	24.7	20.0
	3.1	2.9	2.2	2.6	3.5	4.1
proporção de alunas	88.9	98.5	-5.0	34.3	34.7	52.7
	23.2	25.2	4.7	5.7	5.5	6.8
localizada em (localidade < 15 000 pessoas)						
localidade 15 000 - 100 000 pessoas	1.5	-3.7	3.1	6.2	-3.1	-1.2
	2.6	2.7	2.1	2.7	2.5	3.3
localidade > 100 000 pessoas	6.4	10.9	11.7	14.5	-6.5	1.0
	2.9	3.1	2.1	2.5	3.2	3.7
amplitude dos graus de escolaridade	0.0	1.3	2.1	1.0	1.3	1.3
	0.6	0.7	0.6	0.6	0.3	0.4
proporção de repetentes	-26.2	-9.8	-60.0	-107.8	-75.0	-120.6
	11.6	14.0	14.3	13.9	17.1	23.1
escola enfrenta concorrência (não)	6.8	1.3	-0.1	3.7	0.5	4.3
	2.3	2.2	2.0	2.3	3.4	4.5
autonomia na alocação de recursos	-8.9	44.5	1.1	2.1	3.4	3.1
	12.1	13.6	1.4	2.0	1.5	2.0
autonomia nos programas/avaliação	-2.7	-7.6	-2.5	-1.6	-1.6	-2.8
	1.5	1.5	0.9	1.1	1.2	1.6
escola recebe pressão dos pais (não)	6.0	9.4	14.8	15.0	11.5	12.9
	3.5	4.9	2.4	2.2	3.4	3.0
escola privada (pública)	13.4	-12.3	-35.2	-24.9	7.4	6.2
	4.7	6.4	4.7	4.2	2.1	2.0
<b>Recursos escolares</b>						
dimensão das turmas	0.5	0.3	-0.3	-0.1	1.5	1.8
	0.3	0.3	0.1	0.1	0.4	0.5
rácio aluno/professor	0.1	-0.4	3.7	2.8	4.4	4.6
	0.5	0.6	0.4	0.3	0.5	0.8
propor. de computadores com internet	-0.5	11.5	17.2	13.1	15.5	3.1
	4.0	5.3	2.9	4.1	5.2	5.5
rácio computador/aluno	23.9	17.9	13.6	-10.6	26.3	-2.2
	29.4	29.5	7.4	9.8	10.2	12.5
horas de aulas língua do teste/matemática	6.8	5.4	9.0	7.4	9.4	3.5
	0.3	0.5	0.3	0.2	0.2	0.3
escassez de profess. língua/matem. (não)	-	-	3.5	11.5	-12.0	-19.1
			2.9	3.2	2.4	3.5
<b>Familiaridade com computadores</b>						
autoconfiança em tarefas da internet	10.3	16.2	14.8	18.2	13.1	17.5
	1.1	0.9	0.4	0.5	0.7	1.0
tempo gasto na internet ou em diversão	-5.2	-8.1	-12.6	-14.3	-7.2	-9.2
	0.9	0.9	0.4	0.6	0.5	0.6
coeficiente de determinação	0.56	0.56	0.36	0.34	0.48	0.45
observações na amostra	4981	4981	45660	45660	18319	18319

Fonte: Cálculos do autor.

Notas: (a) Média dos coeficientes estimados por regressões dos mínimos quadrados ponderados dos cinco valores plausíveis em matemática e leitura, respetivamente, sobre as variáveis no quadro e variáveis binárias para os países (não mostradas); os desvios-padrão aparecem em itálico. A variância depende da variância amostral, calculada pela variante de Fay do método de *balanced repeated replication*, e da variância de imputação. (b) Os países com baixo desempenho incluem a Espanha, a Grécia e a Itália; os países com alto desempenho incluem a Bélgica, a Finlândia e os Países Baixos. (c) Categoria omitida entre parêntesis, para as variáveis binárias. (d) O efeito do ano de escolaridade é estimado separadamente para Finlândia (valores à direita) onde a idade de início da escola é aos 7 anos, e nos outros países com alto desempenho (valores à esquerda) onde esta é aos 6 anos.

### Contexto familiar

A função de produção de educação estimada inclui várias medidas do contexto socioeconómico do estudante, e os resultados confirmam que estas têm um forte impacto nas pontuações dos testes. Aqui destaca-se a contribuição do regressor livros em casa, o que não é uma surpresa já que este é muitas vezes o melhor preditor do desempenho em regressões similares (Hanushek e Woessmann, 2010). Naturalmente, não é o número de livros em casa por si que está causalmente associado com o desempenho, mas esta variável capta muito bem um ambiente familiar propício à aprendizagem. Os impactos medidos para Portugal são menores do que para os dois grupos de países de referência, especialmente na categoria superior (mais de 200 livros). Nas três regressões, conforme se sobe na escala de profissões dos pais, emerge uma influência positiva sobre o desempenho, particularmente marcada para as ocupações intelectuais/especializadas. Quanto à educação formal dos pais, a sua contribuição em Portugal, ou é estatisticamente não significativa, ou está no limiar da significância, o que contrasta com o forte impacto nos dois grupos de referência. Entre as variáveis relativas ao contexto familiar, as qualificações académicas poderão estar especificamente associadas ao acompanhamento pelos pais das tarefas escolares dos seus filhos. Estes resultados podem indicar uma menor capacidade ou disponibilidade dos pais portugueses a este respeito.

Um segundo conjunto de variáveis neste grupo refere-se à nacionalidade e à língua falada em casa. A condição de imigrante geralmente implica uma desvantagem em termos de desempenho, que é maior nos países onde os alunos têm melhores resultados. Para Portugal, os estudantes imigrantes de segunda geração têm piores resultados do que seus colegas de primeira geração em termos de estimativas pontuais (embora a diferença não seja estatisticamente significativa nos níveis habituais), implicando que o impacto negativo deste fator não se atenua à medida que aqueles e as suas famílias permanecem mais tempo no país. É interessante constatar que, controlando para a condição de imigrante, falar uma língua estrangeira em casa tem um impacto positivo e significativo nas pontuações em matemática em Portugal (na leitura este impacto é negativo). Tal resultado pode refletir um empenhamento muito forte na escola de certos grupos de estudantes imigrantes, superando o dos estudantes com ascendência portuguesa.

A contribuição dos recursos educativos em casa para o desempenho dos alunos portugueses é positiva e significativa, mas fica aquém da contribuição nos dois grupos de referência. O coeficiente do indicador de riqueza é, ou não significativo, ou negativo e significativo, o que indica que este não tem impacto próprio quando são tidos em conta outros aspetos da situação sócio-económica dos alunos.

Os impactos das variáveis do contexto familiar podem ser interpretados numa outra dimensão relacionada com a equidade educacional - um assunto explorado em mais detalhe na secção 4. A relação entre uma medida sumária do estrato socioeconómico (por exemplo, o número de livros em casa) e uma medida do desempenho - por vezes designada como declive do *gradiente socioeconómico* - é frequentemente utilizada como um indicador de igualdade de oportunidades educacionais. Um gradiente socioeconómico com forte declive implica resultados escolares mais desiguais para crianças de diferentes estratos, mantendo tudo o resto constante. Na função de produção de edu-

cação estimada para Portugal, os coeficientes das variáveis que medem vários aspetos do contexto familiar sugerem um menor impacto no desempenho, em particular em comparação com o grupo de países com melhores resultados. Dois fatores poderão ser responsáveis por isso. O primeiro é uma atitude mais passiva dos pais em relação à educação e um menor envolvimento na vida escolar dos seus filhos. O segundo é um sistema de ensino que tende a contrabalançar mais a situação desigual das crianças de diferentes classes sociais.

### **Características da escola**

Na função de produção de educação foram consideradas diversas características da escola. As estimativas pontuais da influência da dimensão da escola são positivas para Portugal, embora apenas significativas no caso dos testes de leitura. Tal indica a existência de economias de escala, em consonância com Pereira e Moreira (2007). Nos grupos de referência, o mesmo tipo de evidência está confinado aos países onde os estudantes têm um alto desempenho. Em contrapartida, a localização das escolas em Portugal aparece menos importante do que naquele estudo, pois somente a categoria superior - localização numa cidade com mais de 100 000 pessoas - faz uma diferença (positiva) significativa para o desempenho. Tais resultados devem ser explicados pelo conjunto alargado de variáveis de controlo aqui utilizado. Uma maior proporção de repetentes tem o esperado impacto negativo sobre o desempenho, enquanto uma proporção maior de alunas contribui para um ambiente propício ao sucesso escolar.

Do ponto de vista teórico, o efeito da autonomia das escolas no sucesso escolar é ambíguo. Por um lado, este pode ser positivo porque os decisores ao nível da escola tendem a ter melhor informação. No entanto, a autonomia também pode ser utilizada pelos decisores para prosseguirem os seus próprios objetivos, os quais podem não coincidir com a melhoria do desempenho dos alunos (Hanushek e Woessmann, 2010). Na literatura empírica tem-se verificado que a existência de margem de manobra na definição do orçamento (dado o montante global), incluindo a contratação e remuneração dos professores, bem como na escolha dos manuais e métodos de ensino, tende a melhorar o desempenho. Em contrapartida, a autonomia das escolas no que respeita à dimensão do seu orçamento, e a autonomia dos professores sobre as matérias a serem abordadas nas aulas, aparecem negativamente associadas com o desempenho, possivelmente porque estas se prestam mais a um comportamento oportunista. Deve acrescentar-se que as regressões realizadas não são a forma ideal de estudar tais efeitos, uma vez que não exploram plenamente a variação entre-países, a qual facilita esse estudo. Não obstante, as nossas estimativas pontuais geralmente encaixam neste tipo de evidência. A autonomia na afetação de recursos tem uma contribuição positiva para o desempenho (em Portugal isto acontece apenas para as pontuações em leitura) e a autonomia na escolha das matérias e na avaliação uma contribuição negativa, embora nem sempre significativas. Note-se que, no caso português, as escolas públicas não têm praticamente nenhuma autonomia na alocação e remuneração dos professores pelo que o primeiro daqueles indicadores está próximo do mínimo global; são as escolas privadas que lhe conferem alguma variabilidade.

A pressão dos pais tem um impacto positivo no desempenho, embora, como mencionado acima,

isto também possa refletir o efeito da escolha de melhores escolas por pais mais cuidadosos e informados, além da pressão colocada sobre as escolas. O coeficiente do indicador de escola privada é negativo, mas no limiar da não significância para a leitura, e positivo e significativo para a matemática. Neste último caso, o impacto medido na média da variável dependente é de cerca de 3 por cento. Pereira e Moreira (2007) - que usaram a nota média no 12. ano nos exames nacionais para todas as disciplinas, em 2003/04 e 2004/05 - estimaram um efeito-escola privada de 7 a 8 por cento. Neste estudo obtém-se um efeito mais fraco para a matemática (e de sinal oposto para a leitura). Para além de um conjunto de dados diferente, este resultado pode ser explicado pela ausência de controlos familiares no estudo mencionado, levando a um enviesamento do coeficiente em alta. Embora seja possível que a inclusão desses controlos esvazie aproximadamente o conteúdo explicativo do indicador de escola privada, estes resultados não devem ser vistos como definitivos. Dado que as escolas privadas representam uma pequena parte (cerca de 10 por cento) do universo dos estudantes em causa, poderão existir enviesamentos causados pelo processo de amostragem. A comparação das médias das pontuações nas escolas públicas e privadas na base de dados do PISA 2006 e nos exames nacionais do 9. ano (ano letivo 2006/07), sugere essa possibilidade<sup>12</sup>.

### **Recursos escolares**

As medidas tradicionais de recursos escolares, tais como a dimensão média das turmas e o rácio aluno/professor, não entram significativamente, em termos estatísticos, na função de produção de educação para Portugal. Nos dois grupos de países de referência, o coeficiente do rácio aluno/professor tem um sinal positivo (contraintuitivo) e é estatisticamente significativo - o que também se verifica para a dimensão das turmas nos países com alto desempenho. Hanushek (1986) sumaria as conclusões de muitos estudos sobre a contribuição dos recursos escolares e conclui que esta é muitas vezes não significativa e, em alguns casos, são encontrados efeitos contrários ao esperado. Isto está em linha com o resultado bem conhecido de que essas variáveis têm uma influência muito mais fraca (se é que a têm) no desempenho do que as relativas ao contexto familiar. A quantidade de horas de aulas contrasta com este tipo de evidência, tendo um impacto positivo e estatisticamente significativo sobre o desempenho, nas três regressões. Esta parece ser a única variável onde intervenções ao nível dos recursos poderão trazer resultados positivos.

### **Familiaridade com computadores**

Argumentou-se acima que a variável autoconfiança em tarefas associadas à *internet* poderia ser uma medida das capacidades dos alunos. O facto de o seu impacto ser positivo e significativo, e não muito diferente de país para país, corrobora esta interpretação. Em contrapartida, o tempo despendido com entretenimento e navegação na *internet* está negativamente relacionado com o desempenho.

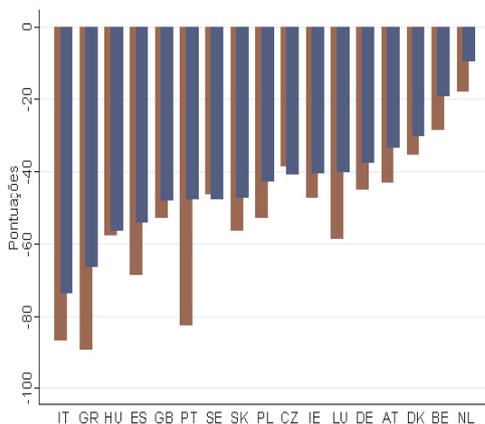
(12) No PISA 2006, as notas médias das escolas privadas são superiores, respetivamente, em 5.5 por cento em matemática e 3.2 por cento na leitura. Os valores correspondentes nos exames nacionais do 9. ano, 2006/07, em matemática e português são 22.6 e 7.8 por cento (Jornal Público, 2007).

### 3.3. Diferenças no desempenho dos países mantendo as características dos pais constantes

Nos *rankings* de países construídos com base nos resultados dos testes do PISA, como os apresentados nos gráficos 1A e 1B, obviamente não se controla para os fatores que determinam tais resultados. Ao mesmo tempo, determinantes como as habilitações literárias e a estrutura ocupacional da população, captados pelas variáveis do contexto familiar, são uma herança do passado e pouco afetados pelas atuais políticas de educação. Uma questão que se coloca neste contexto é avaliar como os *rankings* mudariam se existisse uma distribuição semelhante das características parentais nos diversos países. Tal é particularmente pertinente para Portugal que tem uma das menores dotações entre os países da União Europeia em termos do nível de escolaridade da população. A fim de se investigar esta questão, estimou-se uma função de produção de educação para todos os países, incluindo os regressores socioeconómicos (ver Quadro 1) e variáveis binárias específicas a cada país<sup>13</sup>. Os coeficientes destas últimas podem ser interpretados como as pontuações médias mantendo o contexto familiar constante, e comparados com as médias não condicionadas apresentadas nos Gráficos 1A e 1B. Este exercício exige alguma cautela, uma vez que o nosso modelo é apenas uma aproximação à realidade. Na prática, se o nível de determinadas variáveis, digamos, num país com baixo desempenho, fosse igual ao nível vigente num país com alto desempenho, a alteração efetiva nas pontuações diferiria da decorrente do modelo. Os resultados são apresentados nos Gráficos 3A e 3B, em termos da distância de cada país face ao país com desempenho mais

Gráfico 3A

DESEMPENHO EM MATEMÁTICA POR PAÍS,  
DIFERENÇA PARA A FINLÂNDIA  
Média não condicionada (a vermelho) e média com  
o contexto familiar constante (a azul)

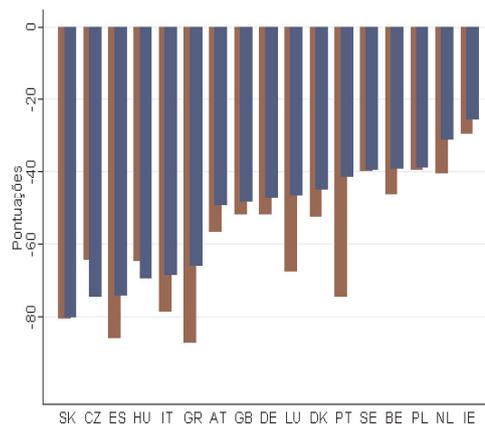


Fonte: Cálculos do autor.

Nota: Baseado em regressões dos mínimos quadrados ponderados das pontuações em matemática sobre constantes específicas a cada país (média não condicionada), e essas constantes e as variáveis socioeconómicas (média com o contexto familiar constante). O gráfico mostra a diferença entre os coeficientes para cada país e para a Finlândia.

Gráfico 3B

DESEMPENHO EM LEITURA POR PAÍS,  
DIFERENÇA PARA A FINLÂNDIA  
Média não condicionada (a vermelho) e média com  
o contexto familiar constante (a azul)



Fonte: Cálculos do autor.

Nota: O mesmo que o gráfico anterior, mas para as pontuações em leitura.

(13) Os restantes regressores, tais como características das escolas e os recursos escolares, não são incluídos uma vez que não queremos controlar para os mesmos. Os regressores socioeconómicos irão captar o efeito dos regressores omitidos na medida em que exista correlação entre ambos. No entanto, esta deverá ser relativamente baixa (exceto, porventura, para a natureza pública ou privada das escolas) e causar uma distorção pequena.

elevado, que é a Finlândia.

Portugal é o país onde o hiato para a Finlândia mais se encurta quando se considera a média condicionada das pontuações, em vez da média não condicionada, o que confirma um forte impacto negativo da composição sócioeconómica da população. A redução daquele hiato é mais acentuada do que em outros países de baixo desempenho, como a Espanha e a Itália, que, conforme demonstrado no Quadro 1, têm uma situação mais favorável relativamente à escolaridade e estrutura ocupacional do que Portugal. Os Gráficos 3A e 3B indicam uma situação menos sombria para Portugal em termos do sucesso escolar do que Gráficos 1A e 1B. Em matemática, embora ainda na metade inferior do *ranking*, Portugal está perto do grupo de países com níveis médios de desempenho. Em leitura, a mudança é mais acentuada, e os alunos portugueses aparecem na metade superior.

#### 4. ALGUNS ASPETOS DA VARIABILIDADE NO DESEMPENHO DOS ALUNOS

Considera-se agora a variabilidade no desempenho dos alunos, ou seja, a desigualdade académica, complementando os resultados anteriores que procuraram explicar o respetivo nível. As análises costumam colocar muita ênfase na desigualdade social como fonte de desigualdade académica. Tal justifica-se pela importância da condição sócioeconómica do estudante para o sucesso escolar. Além disso, os regressores deste grupo têm geralmente maior variância do que, especialmente, os relacionados com os recursos escolares (em particular no contexto de análises dentro de cada país ou envolvendo países com níveis semelhantes de desenvolvimento). Ao mesmo tempo, a desigualdade académica sustenta as desigualdades sociais, pois o nível de escolaridade das gerações mais velhas é o fator mais importante por trás da distribuição dos trabalhadores por ocupações e níveis salariais que prevalece atualmente. De facto, um dos principais objetivos dos sistemas educativos é atenuar progressivamente essas desigualdades, garantindo que a distribuição das competências dos alunos quando terminam a escola é menos desigual do que a dos seus pais. Outras fontes importantes de variabilidade nos resultados educacionais são as capacidades cognitivas dos alunos e a contribuição dos professores para o sucesso da aprendizagem.

Os alunos encontram-se distribuídos por escolas, e a dispersão no desempenho pode materializar-se em certa medida através da existência de escolas que diferem substancialmente quanto ao mesmo. Assim, é importante considerar, para além da variabilidade total no desempenho, a proporção avaliada *entre- e intraescolas*. Se a componente entre-escolas for grande relativamente à componente intraescolas, tal significa que os estudantes com níveis mais baixos de sucesso se concentram em algumas escolas e os estudantes com níveis mais altos em outras. Este fenómeno pode acontecer por diversas razões, por exemplo, as escolas podem ser frequentadas por alunos provenientes predominantemente de famílias favorecidas ou desfavorecidas. Quando existe uma grande assimetria entre regiões dentro de um país, dado que a mobilidade dos estudantes é limitada, a composição social das escolas localizadas nas regiões mais pobres irá diferir significativamente da composição das suas homólogas nas regiões mais ricas. O mesmo se verifica para as assimetrias entre bairros dentro das grandes cidades, sobretudo na ausência de normas que

obrigam os alunos a frequentar a escola mais próxima, uma vez que os pais tendem a matricular os seus filhos em escolas frequentadas por estudantes com características afins.

A separação dos estudantes por escolas que oferecem currículos específicos, por exemplo, geral ou profissional, numa fase inicial do seu percurso (digamos, entre os 10 e os 12 anos de idade) introduz uma diferenciação entre as escolas. Este mecanismo - designado na literatura por *early tracking* - existe nos sistemas educativos de alguns dos países que consideramos (Brunello e Checchi, 2007). Tal separação pode ser feita com base em critérios como testes formais, recomendação dos professores e auto-seleção. Os testes formais introduzem uma segmentação dos alunos de acordo com sua origem social e capacidades cognitivas, visto que estas são determinantes importantes das pontuações nos testes de colocação. Uma segmentação de acordo com a origem social poderá ocorrer mesmo no caso de auto-seleção, pois os pais com profissões manuais podem achar mais natural matricular os seus filhos em escolas que oferecem currículos profissionais, e os pais com profissões intelectuais, nas escolas que oferecem programas que dão acesso à universidade.

Como primeiro exercício, consideramos a associação entre a decomposição da variância das pontuações em matemática e de uma medida sumária do contexto familiar, o índice da situação económica, social e cultural (índice ESCS). Este índice, disponível na base de dados do PISA, sintetiza várias dimensões do contexto familiar (ver o Anexo 1 para os detalhes sobre a sua construção), exceto a condição de imigrante. A decomposição da variância para cada uma das variáveis é obtida a partir da estimação de um modelo multinível sem regressores, incluindo interceções específicas à escola modeladas como parâmetros aleatórios<sup>14</sup> (ver Goldstein, 2010). O Gráfico 4A mostra a associação entre a variância total das pontuações em matemática e do índice ESCS (em ambos os casos normalizada para uma média igual a 100) e o Gráfico 4B a associação entre a respetiva proporção entre-escolas. Nos países onde essas proporções são maiores, as escolas diferem mais substancialmente no que se refere ao desempenho dos alunos e à composição social.

Começamos por examinar a dispersão das pontuações em matemática. Não há uma relação óbvia entre o desempenho (Gráfico 1A), por um lado, e a variabilidade das pontuações e a sua decomposição nestes gráficos, por outro, nos diversos países. O país com melhor desempenho - a Finlândia - tem a menor variância global e proporção entre-escolas; no entanto, a Bélgica, também no topo do ranking de desempenho, tem a maior variância e uma das proporções entre-escolas mais elevadas (os Países Baixos, igualmente no topo, têm a mais alta). Baixos níveis de desempenho são também compatíveis com situações bastante diferentes em termos de variabilidade das pontuações, conforme ilustrado pelos casos da Itália e da Espanha. Estas conclusões estendem-se em geral à dispersão das pontuações em leitura (não mostrada), que é maior do que a das pontuações em matemática na maioria dos países.

Quanto à dispersão do índice ESCS, Portugal surge como um caso especial, com um valor quase 75 por cento acima da média no conjunto dos países (Gráfico 4A). No entanto, esta dispersão é

(14) Este modelo explica a variável dependente como a soma dessas interceções (iguais a uma média fixa mais um desvio aleatório desta, específico a cada escola) e de uma variável residual ao nível do estudante. As estimações foram realizadas utilizando o programa GLLAMM (Rabe-Hesketh et al., 2004) que é executado no STATA. A ponderação foi feita de acordo com o primeiro dos dois esquemas de ponderação propostos por Pfeffermann et al. (2008). No caso das pontuações em matemática tomou-se o primeiro valor plausível.

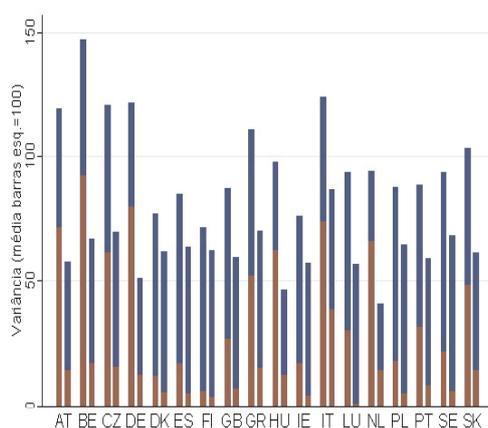


Concluimos esta secção com a estimação de um modelo multinível incluindo o conjunto completo de regressores socioeconómicos como preditores da variabilidade das pontuações, ao nível do estudante e (tomando as médias por escola) ao nível da escola<sup>17</sup>. Estes últimos captam as externalidades para a escola como um todo associadas à sua composição sócioeconómica. A variância total e sua decomposição neste modelo são apresentadas no Gráfico 5, que mostra também as quantidades correspondentes no modelo sem regressores (medidas no eixo dos Ys nos Gráficos 4A e 4B acima), a fim de facilitar a comparação dos resultados. A parte inferior de cada barra, a vermelho, indica a componente entre-escolas da variância. Todas as quantidades são normalizadas pela variância média no conjunto dos países no modelo sem regressores para que as alturas das barras sejam comparáveis.

Os regressores socioeconómicos explicam uma parte importante da variabilidade das pontuações, pelo que a variância estimada no segundo modelo é bastante menor do que no primeiro. Além disso, a diminuição da componente entre-escolas supera claramente a da componente intra-escolas, em especial nos países com separação precoce dos alunos. A variância entre-escolas que permanece é, todavia, um pouco maior para este conjunto de países do que para os restantes, com exceção da Itália e da Grécia, o que presumivelmente reflete os outros aspetos da diferenciação entre as mesmas induzidos por aquele mecanismo. Note-se que a Itália e, em menor grau, a Grécia são casos especiais em que o desempenho dos alunos parece diferir substancialmente de umas

**Gráfico 5**

**VARIÂNCIA DAS PONTUAÇÕES EM MATEMÁTICA, TOTAL E ENTRE-ESCOLAS**  
Modelo sem regressores (barras à esquerda) e controlando para o contexto familiar (barras à direita)



**Fonte:** Cálculos do autor.

**Nota:** Com base na estimação de modelos multinível, respetivamente, sem regressores (o mesmo modelo que na nota ao Gráfico 4A), e incluindo os regressores socioeconómicos, quer ao nível do aluno (centrados em torno das médias por escola), quer como médias por escola. A variância captada por este último modelo diz respeito à parte não explicada pelos regressores.

(17) Os coeficientes de todas as variáveis (ou seja, os declives) são modelados como parâmetros constantes, enquanto as interceções continuam a ser aleatórias e específicas à escola.

escolas para as outras para além do que é explicado pela desigualdade social. A variabilidade entre-escolas das pontuações em Portugal é comparável à que se verifica nos países sem separação precoce dos alunos (com as duas exceções mencionadas) e, dentro deste grupo, superior, por exemplo, à dos países do Norte da Europa e também da Espanha.

A evidência apresentada mostra que nos países onde a influência da desigualdade social é maior, tal influência é sentida sobretudo através dos efeitos da composição sócioeconómica das escolas e da diferenciação entre os respetivos níveis de desempenho. A variância remanescente após controlar para a desigualdade social, que como vimos é na sua maior parte um fenómeno intra-escolas, deverá ser principalmente explicada por variáveis não observáveis. Estas poderão incluir, por exemplo, as capacidades do estudante e aspetos da qualidade do ensino como a organização das aulas e os métodos utilizados pelos professores. Note-se que as variáveis nas regressões de produção de educação estimadas na secção 3, que agora foram omitidas, deverão apenas explicar uma pequena parte da variância remanescente, na medida em que dizem sobretudo respeito à escola. Esta leitura é também sugerida pela parte considerável da variabilidade das pontuações que permanece inexplicada nessas regressões, como mostra o valor relativamente baixo dos coeficientes de determinação no Quadro 2. Carneiro (2006) conclui, da mesma forma, que variáveis análogas às incluídas nas funções de produção neste estudo não conseguem explicar uma parte considerável da variância das pontuações no PISA 2003 para Portugal.

## 5. CONCLUSÕES

Este estudo apresenta uma análise do nível e da variabilidade do desempenho escolar em Portugal e nos países da União Europeia utilizando a base de dados do PISA 2006. As principais conclusões são as seguintes:

- Os alunos portugueses ocupam consistentemente posições na metade inferior do *ranking* do desempenho no grupo de países considerados, tanto em matemática e como em leitura. Tais resultados são em parte provocados por um contexto familiar desfavorável, nomeadamente no que se refere às habilitações e ocupações dos pais.
- Na função de produção de educação para Portugal, à semelhança do acontece para ambos os grupos de países de referência, as variáveis socioeconómicas são os principais determinantes do desempenho, enquanto os recursos escolares têm uma contribuição pouco importante (exceto as horas de aulas normais).
- As variáveis socioeconómicas têm, no entanto, contributos mais fracos em Portugal do que, em particular, nos países de alto desempenho. Isto deverá ser explicado pelo menor envolvimento dos pais na vida escolar dos seus filhos e um papel mais equalizador do sistema de ensino relativamente ao sucesso escolar das crianças provenientes de diferentes estratos sociais.
- Algumas das variáveis na função de produção educação para Portugal, nomeadamente, a localização das escolas e a sua natureza pública ou privada aparecem, uma vez controlando extensiva-

mente para o contexto familiar, com menor impacto do que em estudos anteriores.

- Não existe uma relação óbvia entre o nível e a dispersão do desempenho nos diferentes países, sendo níveis altos e baixos compatíveis com diferentes graus de dispersão.
- A desigualdade social é uma importante fonte de variabilidade no desempenho, principalmente nos países cujos sistemas educativos têm mecanismos de seleção precoce dos alunos. Esta característica também introduz uma diferenciação importante no desempenho entre as escolas, devido a efeitos de triagem dos alunos de acordo com o contexto familiar e as suas capacidades.
- Portugal tem uma grande dispersão nos indicadores do contexto familiar, mas, dado um gradiente socioeconómico mais plano, esta é relativamente menos transmitida aos resultados nos testes, cuja variância está próxima da média no conjunto dos países considerados.

## BIBLIOGRAFIA

- Brunello, G. e Checchi, D. (2007). "Does school tracking affect equality of opportunity? New international evidence". *Economic Policy* October 2007, 781–861.
- Carneiro, P. (2006). "Equality of opportunity and educational achievement for Portugal." *Proceedings of the 2006 Banco de Portugal Conference on Economic Development in Portugal*. Banco de Portugal.
- Coleman, J., E. Campbell, C. Hobson, J. McPartland, A. Mood, F. Weinfeld e R. York (1966). *Equality of Educational Opportunity*. Washington, D.C.; U.S. Government Printing Office.
- Fuchs, T. e Woessmann, L. (2007). "What accounts for international differences in student performance? A reexamination using PISA data". *Empirical Economics* 32 (2/3), 433–464.
- Goldstein, H. (2010). *Multilevel Statistical models* (4th edition). Cichester, UK: Wiley.
- Hanushek, E. (1979). "Conceptual and empirical issues in the estimation of educational production functions". *Journal of Human Resources* 14 (3), 351-388.
- Hanushek, E. (1986). "The economics of schooling: Production and efficiency in public schools". *Journal of Economic Literature* 24 (3), 1141-1177.
- Hanushek, E., e L. Woessmann (2006). "Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries". *Economic Journal* 116 (510), C63-C76.
- Hanushek E. e Woessmann, L., (2010). "The economics of international differences in educational achievement". NBER *Working Paper* n. 15949. Cambridge, MA: National Bureau of Economic Research.
- Jornal Público (2007). *Ranking, exames do ensino básico e secundário*. Suplemento à edição de 2 de novembro de 2007.
- Kalton, G. e Kasprzyk, D. (1982). "Imputing for missing survey responses". *Proceedings of the Survey Research Methods Section*. American Statistical Association.
- OECD (2009). *PISA Data Analysis Manual*. Paris: Organization for Economic Cooperation and Development.
- Pereira, M. C. e Moreira, S., (2007). "A stochastic frontier analysis of secondary education output in Portugal". Banco de Portugal *Working Paper* n. 6/2007.
- Pfeffermann, D., C. J. Skinner, D. J. Holmes, H. Goldstein, e J. Rasbash (1998). "Weighting for Unequal Selection Probabilities in Multilevel Models". *Journal of the Royal Statistical Society Series B* 60 (1), 23-40.
- Rabe-Hesketh, S., Skrondal A., e Pickles, A. (2004). *GLLAMM Manual*. Paper n. 160, U.C. Berkeley Division of Biostatistics Working Paper Series. Berkeley, CA: University of California Berkeley.
- Todd, P., e Wolphin, K. (2003). "On the specification and estimation of the production function for cognitive achievement". *The Economic Journal* 113 (485), F3-F33.
- Woessmann, L. (2003). "Schooling resources, Educational Institutions and Student Performance: the International Evidence". *Oxford Bulletin of Economics and Statistics*, 65 (2), 117-170.

## ANEXO 1

### Definição de algumas variáveis explicativas

Riqueza (base de dados do PISA). Índice calculado com base nas respostas dos alunos sobre a posse de bens duradouros, como televisões, carros ou telemóveis.

Recursos educativos em casa (calculado pelo autor). Índice calculado somando as variáveis binárias sobre a posse de: uma mesa para estudo, um lugar tranquilo para estudar, um computador para trabalhos escolares, *software* educativo, calculadora própria, livros que auxiliem nos trabalhos escolares e um dicionário.

Condição de imigrante (base de dados do PISA). Variáveis binárias para: estudantes não imigrantes - nascidos no país, bem como pelo menos um dos pais; estudantes imigrantes de segunda geração - nascidos no país, mas os pais nasceram fora do país; alunos imigrantes de primeira geração - nascidos fora do país.

Amplitude dos graus de escolaridade (calculado pelo autor). Calculado como a diferença entre o grau máximo e o grau mínimo em cada escola.

Concorrência de outras escolas (base de dados do PISA). Variável binária para as escolas que reportaram a existência de pelo menos uma outra escola na mesma área competindo pelos seus alunos.

Autonomia na alocação de recursos (base de dados do PISA). Índice calculado com base nas respostas da escola sobre a responsabilidade na gestão dos recursos, por exemplo, contratação, despedimento e remuneração dos professores, e na formulação do orçamento da escola.

Autonomia de programas e avaliação (base de dados do PISA). Índice calculado com base nas respostas da escola sobre quem tem a responsabilidade nos métodos de avaliação dos alunos, conteúdo dos programas e livros usados.

Pressão dos pais (base de dados do PISA). Variável binária para as escolas que reportaram a existência de pressão constante de muitos pais em relação aos padrões escolares.

Autoconfiança em tarefas da internet (base de dados do PISA). Índice calculado com base nas respostas dos alunos sobre a perícia com que estes executam tarefas como conversar em rede, descarregar arquivos ou música da *internet*, e enviar correio eletrónico.

Tempo gasto na internet e entretenimento (base de dados do PISA). Índice calculado com base nas respostas dos alunos sobre a frequência com que estes usam o computador para tarefas como navegar na *internet*, jogar, descarregar música, enviar correio eletrónico e conversar em rede.

Índice da situação económica, social e cultural (ESCS) (base de dados do PISA). Índice que resume as respostas sobre a posse de bens duradouros, bens culturais, recursos educativos em casa, o número de livros em casa, a escolaridade mais elevada dos pais e a ocupação mais elevada dos pais.

## ANEXO 2

### Imputação de dados

A imputação de dados foi realizada utilizando o método de *predicted regression imputation* (ver Kalton e Kasprzyk, 1982). As variáveis com valores em falta foram regredidas sobre um conjunto “fundamental” de regressores incluindo o grau de escolaridade, a idade, o género, a localização da escola e o país (estas regressões foram estimadas para o conjunto dos países). As observações para as quais pelo menos um destes regressores fundamentais não tinha valores não foram consideradas. De notar que as variáveis ao nível do aluno neste grupo têm poucas ou nenhuma observações em falta. A inclusão da localização da escola permite eliminar da amostra as escolas que preencheram o respetivo questionário de forma muito esparsa, com a maioria das variáveis - incluindo a localização - em falta.

No processo de imputação, distinguiu-se entre as variáveis do estudante e da escola. As primeiras foram imputadas através de regressões ao nível do aluno, estimadas por mínimos quadrados ponderados, no caso das variáveis contínuas, e regressão *probit* (ordenada) ponderada, no caso das variáveis ordinais e binárias. Quando a variável a imputar era contínua, os valores ajustados foram preenchidos para os dados em falta; quando a variável era ordinal ou binária, a categoria prevista foi preenchida. As variáveis escolares foram imputadas através de regressões ao nível da escola, com as variáveis do estudante definidas como médias por escola, estimadas pelas versões não ponderadas dos mesmos métodos.

Apenas uma pequena proporção das observações teve de ser imputada para cada variável: em média, cerca de 2 por cento da amostra utilizada. A extensão máxima de imputação foi de 7 por cento. A comparação das estatísticas descritivas antes e depois da imputação mostra alterações muito pequenas. No entanto, sem o procedimento de imputação, considerando as regressões da secção 3, cerca de 34 por cento das observações teriam sido perdidas.