

EL RETORN

DEL CONEIXEMENT D'ANGLÈS EN EL MERCAT LABORAL CATALÀ

Antonio Di Paolo

AQR-IREA, Universitat de Barcelona

Hi ha un consens general al voltant del paper de l'anglès com a "lingua franca" en un món sempre més globalitzat, on la freqüència i la intensitat de les interaccions (econòmiques i socials) entre agents de països i llengües diferents s'incrementen de manera progressiva (Fidmuc, 2011). Així mateix, un nombre creixent de treballs empírics evidencien l'existència d'un rendiment econòmic positiu del domini de llengües estrangeres i de l'anglès en particular¹. En termes generals, el coneixement de llengües es pot entendre com un factor de capital humà. Aquesta afirmació no solament val per al cas dels immigrants que apre-

nen la llengua del seu país d'acollida (Chiswick i Miller, 2007, 2014), sinó que també es pot aplicar a l'aprenentatge de llengües estrangeres.

Més específicament, el coneixement de llengües diferents de la nativa pot tenir un valor econòmic positiu en el mercat laboral per diferents raons. En primer lloc, adquirir competències en diferents llengües pot augmentar la productivitat dels treballadors, no solament perquè incrementa les seves possibilitats comunicatives, sinó també gràcies a l'impacte positiu que el bilingüisme pot tenir en les capacitats cognitives (Adeosope *et al.*, 2010; Costa i Sebastián-Gallés, 2014). En segon lloc, el coneixement de l'anglès (o d'altres llengües) pot representar un dels factors que permet accedir a determinades ocupacions o categories professionals que estan més remunerades. En tercer lloc, el domini de l'anglès podria estar remunerat fins i tot sense haver de fer-lo servir a la feina. Això, perquè ser competents en una llengua estrangera constitueix un actiu que pot servir com a senyal positiu "als ulls dels ocupadors", reflectint altres característiques productives inobservades del treballador. De manera consistent amb aquestes prediccions teòriques, la literatura conté diferents estudis que reporten estimacions positives del retorn de les habilitats lingüístiques en llengües estrangeres (sobretot anglès)

1 Agraïments: l'autor agraeix el suport econòmic del Projecte I+D ECO2013-41022-R del Ministeri d'Economia i Competitivitat.

En aquest treball només es farà referència a l'anglès com a llengua estrangera, atès que representa la llengua estrangera més comunament coneguda a la Unió Europea (Unió Europea, 2012), a banda de la disponibilitat de dades per al cas català, que es limita al coneixement d'anglès (vegeu secció 2).

entre la població nativa de diferents països², desenvolupats i emergents. Pel que fa a la Unió Europea, els resultats dels treballs de Ginsburgh i Prieto-Rodríguez (2011) i Williams (2011) indiquen que ser competent en anglès té un rendiment positiu en termes de salaris, però aquesta relació canvia de manera significativa entre països, i és generalment més elevada on la fracció d'adults capaços de parlar (i escriure) en aquesta llengua és més baixa, com en el cas d'Espanya. Més específicament, Williams (2011) també exposa els efectes positius del coneixement d'altres llengües estrangeres diferents de l'anglès, tot i que no resulten ser significatius per a tots els països considerats. Saiz i Zoido (2005) analitzen el retorn del coneixement de llengües diferents de la nativa entre titulats universitaris als EUA, i indiquen l'existència d'un rendiment positiu però significativament més moderat respecte als resultats obtinguts per a altres casos (on l'anglès no és la llengua oficial). Evidències consistents també s'han obtingut per a diferents països en vies de desenvolupament, com ara Sud-àfrica (Casale i Posel, 2011), Índia (Azam *et al.*, 2013) o Turquia (Di Paolo i Tansel, 2015). De manera transversal, els treballs existents també apunten l'existència d'una relació indirecta entre llengües estrangeres i salaris, que funciona mitjançant el mecanisme de selecció ocupacional; no obstant això, no tota la correlació condicional entre les dues variables s'explica amb aquest mecanisme, atès que l'efecte estimat del coneixement d'anglès en els salaris es redueix però continua sent positiu també controlant pel tipus d'ocupació en l'equació d'ingressos.

Fent referència als resultats obtinguts en estudis existents sobre el valor econòmic del coneixement de llengües estrangeres, aquest treball presenta una aplicació al cas del mercat laboral de Catalunya. El cas català és especialment rellevant, atès que Catalunya és una comunitat autònoma amb una forta obertura al comerç exterior i als intercanvis internacionals, que acull un nombre important d'empreses multinacionals i rep un flux considerable de turistes estrangers, fets que poden estimular la demanda de treballadors amb competències en llengües estrangeres.

Específicament, es vol analitzar el retorn de dominar la llengua anglesa (saber parlar i escriure) en termes de salari per hora treballada, mantenint constants (*ceteris paribus*) altres variables que també poden afectar el salari mitjançant l'ús del mo-

2 També cal fer referència als treballs que analitzen l'efecte del coneixement de llengües autòctones, com ara el cas del català, en els resultats laborals (vegeu Rendon, 2007; Di Paolo, 2011; Di Paolo i Raymond, 2012, entre d'altres).

del de regressió. Les dades analitzades procedeixen de l'Enquesta de Condicions de Vida i Hàbits de la Població de Catalunya (ECVHP) del 2011, que conté informació sobre coneixement oral i escrit de l'anglès (entre altres variables). A nivell metodològic, primer s'estima un model de regressió bàsic que conté com a variables explicatives del (logaritme del) salari hora, el gènere, els anys d'estudi, l'experiència laboral i una variable dicotòmica que recull el coneixement d'anglès. Segon, s'afegeixen de manera progressiva característiques familiars (estat civil i nombre de fills), variables laborals (tipus de contracte i tipus d'empresa) i indicadors que recullen el tipus d'ocupació (efectes fixos). Aquesta incorporació progressiva de variables explicatives permet apreciar l'existència de mecanismes indirectes que relacionen el coneixement de l'anglès amb els salaris, com ara la via d'accés a determinades ocupacions. Tercer, atès que la variable dependent (el salari hora) s'observa només entre la submostra d'individus ocupats regularment –per la qual cosa podria no ser representativa del conjunt de la població–, s'aplica el model d'autoselecció endògena de Heckman, que permet obtenir estimacions consistents de l'equació de salari en cas de selecció mostral no aleatòria. Aplicar aquesta metodologia és interessant, no només perquè corregeix la possible inconsistència dels estimadors en cas de selecció endògena, sinó perquè també permet comprovar si el coneixement de l'anglès afecta directament la probabilitat d'estar ocupat. Aquesta anàlisi addicional és especialment rellevant, atès que pot evidenciar un element addicional (el coneixement de llengües estrangeres) que permet defensar-se en un moment de contracció del mercat laboral en un context de crisi (2011).

A banda d'això, també considerem la possible complementaritat entre educació formal i coneixement lingüístic, atès que és possible que saber anglès tingui un valor en el mercat laboral només si s'acompanya d'un nivell d'instrucció suficientment adequat per poder accedir a les ocupacions en què l'anglès és rellevant. Per tant, de manera semblant a Azam *et al.* (2013) i Di Paolo i Tansel (2015), s'estima l'equació de salari (augmentada amb l'indicador de coneixement d'anglès) separatament per a individus amb educació elevada (secundària postobligatòria o terciària) i baixa (secundària obligatòria o inferior). Obtenir un rendiment de l'anglès significativament més elevat entre els treballadors amb instrucció elevada indicaria que les competències lingüístiques han de ser complementades amb un cert nivell d'instrucció reglada per ser remunerades en el mercat laboral.

“El coneixement de l'anglès o d'altres llengües pot representar un dels factors que permet accedir a determinades ocupacions que estan més remunerades”

Cal remarcar des del començament que els resultats presents en aquest article referents a la relació entre coneixement d'anglès i resultats laborals s'han d'interpretar com a correlacions condicionals, que podrien no representar efectes causals, atesa la presència de variables inobservades (habilitat, motivació, etc.) que poden afectar tant el salari com la propensió a saber parlar i escriure en anglès. A més a més, possibles relacions de causalitat inversa i errors de mesura en la variable que recull el coneixement d'anglès³ representen factors addicionals que podrien desviar el coeficient estimat amb el model de regressió del paràmetre poblacional que es voldria estimar. Aquests problemes econòmics es podrien resoldre mitjançant l'aplicació del mètode de variables instrumentals (disposant d'un instrument vàlid per al coneixement lingüístic) o altres tècniques més sofisticades que permeten una interpretació causal dels resultats. Tot i que no és possible aplicar aquestes metodologies amb les dades disponibles per al cas català, considerem que l'anàlisi de correlacions condicionals encara representa una eina valuosa per entendre el valor econòmic del coneixement de llengües estrangeres a Catalunya.

La resta de l'article s'estructura d'aquesta forma: a la secció següent es descriuen les dades i algunes evidències descriptives. Successivament es mostren i interpreten els resultats obtinguts de les diferents estimacions. Finalment, es presenten les conclusions del treball.

1. Dades i estadístics descriptius

Aquest treball es basa en les dades procedents de l'Enquesta de Condicions de Vida i Hàbits de la Població de Catalunya

3 Atès que aquesta variable és una autovaloració de coneixement que du a terme l'individu entrevistat, és possible que s'estigui sobrevalorant el coneixement d'anglès, perquè generalment les persones tendeixen a inflar les seves competències lingüístiques quan responen a una enquesta. Això genera un error de classificació en la variable que defineix els individus que dominen l'anglès, que generalment produeix un biaix positiu en l'estimació de l'efecte del coneixement lingüístic en els salaris (vegeu Dustmann i Van Soest, 2004, per a més detalls sobre aquesta qüestió).

(ECVHP) del 2011, que proporciona una mostra representativa de la població de Catalunya⁴. Les dades ECVHP11 són especialment adequades per dur a terme els objectius d'aquest article, atès que no només contenen informació sociodemogràfica i laboral detallada dels individus enquestats, sinó que també inclouen informació sobre coneixements de català, castellà i anglès. En concret, es disposa d'informació sobre diferents competències lingüístiques de les tres llengües: cada individu autovalora (sí o no) la seva capacitat d'entendre, llegir, parlar i escriure les tres llengües. En aquest treball, considerem que una persona domina l'anglès si és capaç de parlar i escriure en aquesta llengua (tothom que compleixi aquestes dues condicions sap també llegir-hi). Per tant, la variable explicativa d'interès és un indicador que pren valor 1 si l'individu sap parlar i escriure en anglès i 0 en cas contrari. Amb aquesta definició estricta de coneixement lingüístic s'intenten minimitzar els possibles errors de classificació deguts a la naturalesa subjectiva de la variable.

La variable depenent principal d'aquest treball és una mesura de salari per hora treballada (net d'impostos). La base de dades conté informació sobre la renda mensual neta que els individus perceben de la seva activitat laboral principal. També es disposa de les hores treballades habitualment en una "setmana típica". Aleshores, el salari hora net es construeix dividint la remuneració mensual pel nombre d'hores de treball mensuals (és a dir, les hores setmanals multiplicades per un factor de 4,3). Com a variables de control, s'han seleccionat el gènere de l'enquestat, els anys d'estudi completats⁵, l'experiència laboral potencial (que equival a l'edat - anys d'estudis - 6), l'estat civil (casat o no) i el nombre de fills, treballar en el sector públic, tenir un contracte fix i ser autònom o empresari (respecte a ser treballador assalariat). A més a més, també es consideren les diferents categories ocupacionals, basades en la Classificació Catalana d'Ocupacions de 2011 a dos dígits.

La mostra que s'ha seleccionat per dur a terme l'anàlisi econòmica exclou totes les observacions d'individus inactius, és a dir,

4 Per a més informació sobre les dades ECVHP11, vegeu <http://www.enquestadecondicionsdevida.cat/>.

5 La variable anys d'estudi ha estat imputada a partir dels nivells educatius completats: 0 per a analfabets, 3 per a l'educació primària incompleta, 6 per a educació primària completa, 8 per a l'EGB, 10 per a l'ESO, 11 per al BUP, 12 per al COU, batxillerat i FP de grau mitjà, 14 per a l'FP de grau superior, 15 per als estudis universitaris de grau mitjà, 16 per al grau universitari, 17 per a la llicenciatura, 18 per als màsters universitaris, 20 per al doctorat.

Taula 1. Estadístics descriptius

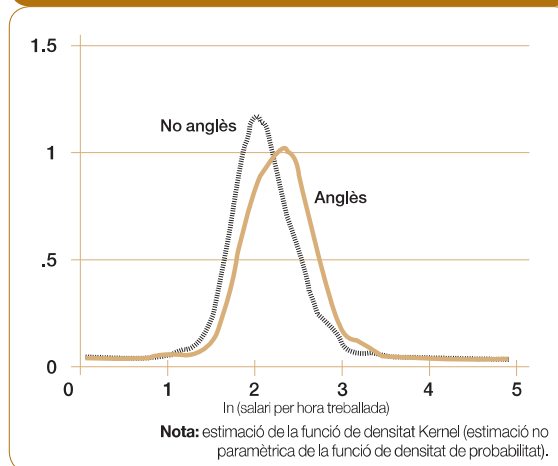
	Nom de la variable	Ocupats		no ocupats	
		Mitjana	Desviació estàndard	Mitjana	Desviació estàndard
Ingressos nets mensuals		1.389,7	652,31		
Hores de treball setmanals		39,505	10,285		
ln (salari per hora treballada)	ln_salarih	2,038	0,406		
Dona	dona	0,489	0,500	0,639	0,481
Anys d'estudi	anys_educ	11,95	3,700	9,561	3,577
Edat	edat	42,06	10,32	43,84	12,75
Experiència potencial	pot_exp	24,11	11,72	28,29	14,25
Parla i escriu anglès	angles	0,342	0,475	0,197	0,398
Casat	casat	0,578	0,494	0,567	0,496
Nombre de fills	nfills	1,192	0,999	1,360	1,142
Sector públic	public	0,203	0,402		
Contracte fix	contr_fix	0,695	0,460		
Autònom/empleador	auton_empr	0,149	0,356		
Atur provincial als 16 anys d'edat	atur 16	11,194	7,924	9,838	8,007
Nombre d'observacions		2.012		877	

totes les persones que no busquen feina de manera activa, els menors de 16 anys, les persones que estudien de manera exclusiva i les persones jubilades o invàlides; els individus majors de 65 anys també s'han eliminat de la mostra, encara que continuïn treballant. També s'exclouen les persones nascudes fora d'Espanya o en altres comunitats autònomes espanyoles però que han arribat a Catalunya després dels sis anys d'edat (d'aquesta manera tothom ha estat escolaritzat a Catalunya i es limiten diferències sociolingüístiques i culturals dins la mostra). Com ja s'ha comentat anteriorment, el salari i les altres característiques laborals s'observen només entre els individus que treballen regularment. Tot i que l'anàlisi principal es basa en aquesta submostra, també es mantenen les observacions dels individus que no treballen en el moment de l'enquesta però busquen feina activament (és a dir, els aturats seguint una definició estàndard), amb l'objectiu d'estimar l'equació de salari tenint en compte el possible biaix d'auto-selecció entre la població ocupada.

La taula 1 conté els estadístics descriptius bàsics de les variables utilitzades per a l'anàlisi empírica, per a les submostres seleccionades d'individus ocupats i aturats, respectivament. Comparant les dues columnes, és possible apreciar que entre els aturats la proporció de dones és més elevada, la mitjana d'instrucció més baixa i l'edat (experiència) més alta. Si bé el 34% dels individus ocupats (de la mostra seleccionada) declara saber parlar i escriure en anglès, aquest percentatge baixa al voltant del 20% entre els aturats.

Amb l'objectiu de donar una primera visió de les diferències d'ingressos associades al coneixement d'anglès, el gràfic 1 conté les estimacions de la distribució del salari per hora treballada (en logaritmes) d'acord amb el coneixement d'anglès. És evident que hi ha fortes diferències salarials associades a les competències lingüístiques, atès que la distribució del salari hora entre els que sí dominen l'anglès està clarament més desplaçada cap a la dreta respecte dels que no saben parlar i escriure en aquesta llengua. De fet, el salari hora mitjà entre

Gràfic 1. Coneixement d'anglès i distribució del salari per hora treballada (en logaritmes)



els que saben anglès és de 9,23 euros, mentre que per als que no saben anglès la mitjana de salari hora baixa a 7,94 (una reducció del 16% aproximadament). Sembla evident que els individus que són competents en llengua anglesa estan més representats en la part alta de la distribució salarial. Cal esbrinar si aquestes diferències no condicionals també es mantenen tenint en compte altres variables de control, tant sociodemogràfiques com laborals, així com el possible biaix d'autoselecció en ocupació.

2. Resultats econòmics

En aquesta secció es descriuen els resultats de les estimacions que s'han dut a terme. L'equació 1 descriu el model lineal que explica el logaritme natural del salari per hora de l'individu "*i*" (w_i) en funció d'un conjunt de variables de control (X_i) i d'una *dummy* que pren valor 1 si l'individu sap parlar i escriure anglès i 0 en cas contrari:

$$\ln(w_i) = \alpha + \beta'X_i + \gamma I(\text{anglès}_i = 1) + \varepsilon_i$$

El coeficient d'interès és γ , que mesura la diferència percentual en salari per hora entre els individus que saben anglès i els que no, mantenint constants les altres variables de control incloses en el model. La taula 2 conté els resultats de les diferents estimacions. A la columna 1 hi apareixen els coeficients estimats del model bàsic estimat per mínims quadrats ordinaris (MQO), en què el vector de variables de control només conté un indicador de gènere, els anys d'estudi i l'experiència potencial (amb terme quadràtic), més l'indicador de coneixement d'anglès.

Les estimacions indiquen que les dones obtenen un salari per hora treballada del 10% inferior respecte als homes i que per cada any addicional d'estudi el salari augmenta el 5,5%; a més a més, l'experiència laboral (potencial) té el típic perfil quadràtic. Més important, l'efecte del coneixement de l'anglès en els salaris és positiu i significatiu, indicant que el fet de saber parlar i escriure en anglès està associat a un salari més elevat del 6,2%. A la segona columna s'afegeixen característiques familiars (indicador d'estar casat i nombre de fills) al vector de variable explicatives (X_i), permetent efectes diferents entre homes i dones mitjançant interaccions. La inclusió d'aquestes variables redueix lleugerament el diferencial de gènere i el rendiment de l'educació i de l'experiència, a banda d'obtenir un efecte (típic) positiu del fet d'estar casats sobre els salaris. No obstant això, les interaccions amb la *dummy* de

dona no són estadísticament diferents de zero. El resultat especialment important per a aquest treball és que la inclusió d'aquestes variables familiars pràcticament no altera el coeficient estimat associat al coneixement d'anglès, que roman positiu i estadísticament significatiu.

La tercera i quarta columna contenen els resultats de l'estimació del model de Heckman de selecció mostral endògena, que bàsicament consisteix a estimar de manera simultània (mitjançant la tècnica de versemblança) l'equació de salari (observat només si l'individu està ocupat),

$$\ln(w_i) = \alpha + \beta'X_i + \gamma I(\text{anglès}_i = 1) + \varepsilon_i$$

$$\text{si } ocupat_i = 1$$

i una equació que determina la propensió (latent) d'estar ocupat ($ocupat_i^* = 1$) i, per tant, determina la mostra amb salari observat, en funció d'un conjunt de factors explicatius (inclouent el coneixement d'anglès) mitjançant un model Probit (és a dir, suposant una distribució normal estàndard del terme d'error):

$$ocupat_i^* = 1 = \mu + \theta'Z_i + \delta I(\text{anglès}_i = 1) + u_i$$

$$\Pr(ocupat_i = 1) = \Pr(ocupat_i^* \geq 0) = \phi(\mu + \theta'Z_i + \delta I(\text{anglès}_i = 1))$$

Amb l'estimació simultània de les equacions de salaris i ocupació es permet que els errors de les dues equacions estiguin correlacionats. D'aquesta forma, es captura la possible auto-selecció endògena en la submostra d'ocupats: certes característiques inobservades que afecten la probabilitat d'estar ocupats (i, per tant, de pertànyer a la submostra per a la qual s'observa el salari hora) també podrien estar relacionades amb el potencial salarial (és a dir, el terme d'error de l'equació de salaris). A més a més, obtenir un coeficient de correlació entre els errors no significatius indicaria que estimar l'equació de salari per MQO no atorgaria coeficients inconsistents a causa de l'autoselecció en ocupació. Cal notar també que el vector de variables explicatives de l'equació d'ocupació (Z_i) conté, a banda de les mateixes variables de control de l'equació de salaris, una restricció d'exclusió; aquesta és una variable que representa un determinant significatiu de la probabilitat d'estar ocupat però no afecta directament els salaris. En aquest estudi, s'utilitza com a restricció d'exclusió per a l'equació d'ocupació la taxa d'atur a la província de residèn-

Taula 2. Resultats de les estimacions

Variable dependent	(1) ln_salarh	(2) ln_salarh	(3) Pr(ocupat)	(4) ln_salarh	(5) ln_salarh	(6) ln_salarh
Constant	1,090*** (0,049)	1,111*** (0,050)	-1,007*** (0,188)	0,912*** (0,275)	1,159*** (0,052)	1,618*** (0,079)
Dona	-0,099*** (0,016)	-0,058** (0,016)	0,039 (0,088)	-0,053** (0,025)	-0,065*** (0,025)	-0,022 (0,026)
Anys educ	0,055*** (0,003)	0,052*** (0,003)	0,095*** (0,009)	0,059*** (0,010)	0,044*** (0,003)	0,023*** (0,004)
Pot exp	0,019*** (0,049)	0,014*** (0,003)	0,056*** (0,010)	0,018*** (0,007)	0,011*** (0,003)	0,012*** (0,003)
Pot exp2	-0,000*** (0,000)	-0,000** (0,000)	-0,001 (0,000)	-0,000* (0,000)	-0,000* (0,000)	-0,000** (0,000)
Anglès	0,062*** (0,020)	0,065*** (0,019)	0,115* (0,063)	0,071*** (0,021)	0,064*** (0,019)	0,043** (0,018)
casat		0,116*** (0,027)	0,514*** (0,114)	0,149*** (0,052)	0,106*** (0,027)	0,098*** (0,027)
nfills		0,021 (0,014)	0,002 (0,050)	0,021 (0,013)	0,024* (0,014)	0,012 (0,014)
dona*casat		-0,041 (0,038)	-0,672*** (0,137)	-0,086 (0,077)	-0,049 (0,037)	-0,060* (0,036)
dona*nfills		-0,018 (0,019)	-0,054 (0,062)	-0,025 (0,020)	-0,020 (0,019)	-0,013 (0,019)
public					0,175*** (0,019)	0,159*** (0,021)
contr fix					0,082*** (0,022)	0,062*** (0,023)
auton empr					0,049 (0,039)	0,036 (0,039)
atur 16			-0,010** (0,004)			
corr(au)			0,462 (0,677)			
Efectes fixos d'ocupació	no	no	no	no	no	sí
R ² ajustat	0,222	0,236	--	--	0,265	0,353
nombre d'observacions	2.012	2.012	2.889	2.012	2.012	2.012

Nota: ***significatiu a l'1%, **significatiu al 5%, *significatiu al 10%; errors estàndards robustos a l'heteroscedasticitat entre parèntesis

cia⁶ quan l'individu tenia 16 anys d'edat (variable "atur_16"), que suposadament afecta les decisions ocupacionals sense influir directament en els ingressos percebuts en el moment de l'enquesta. Els resultats obtinguts⁷, que apareixen a les columnes 3 i 4 de la taula 1 indiquen que el coneixement d'anglès no

6 Les dades sobre atur provincial provenen de De la Fuente (2012), que proporciona series enllaçades d'atur provincial fins a l'any 1964.

7 També els coeficients de les variables de control de l'equació d'ocupació són interessants: com calia esperar, més instrucció i més experiència laboral augmenten la probabilitat d'estar ocupat. No hi ha diferències de gènere entre persones no casades i sense fills. No obstant això, els homes casats tenen més propensió a estar ocupats mentre que entre les dones l'efecte d'estar casada és negatiu. A més a més, la variable que recull la taxa d'atur provincial quan l'individu tenia 16 anys d'edat té el signe negatiu esperat i és estadísticament diferent de zero.

només incrementa la remuneració salarial, sinó que també afecta positivament la probabilitat d'estar ocupat (efecte marginal, +7 punts percentuals). No obstant això, el fet de tenir en compte el possible biaix de selecció en ocupació només incrementa lleugerament el valor estimat de l'efecte del domini de l'anglès en els salaris (de 0,065 a la columna 2 fins a 0,071 de la columna 4). De fet, l'estimació del coeficient de correlació entre els errors de les equacions de salari i d'ocupació no és estadísticament diferent de zero; per tant, es pot estimar de manera consistent l'efecte de l'anglès en els salaris simplement treballant amb la submostra d'individus ocupats (i, per aquest motiu, es pot seguir aplicant MQO com a mètode d'estimació).

Les columnes 5 i 6 contenen els resultats obtinguts afegint característiques laborals al conjunt de regressors. En primer lloc

(columna 5), s'inclouen indicadors per i) treballar al sector públic, ii) tenir un contracte indefinit (o de funcionari) i iii) ser autònom o empresari (*versus* treballador assalariat). Els resultats referents a aquestes variables laborals addicionals són els esperats: els treballadors amb contracte indefinit i, encara més, els que treballen al sector públic obtenen un salari hora major (a paritat d'altres condicions), però no es detecten diferències significatives entre assalariats i no assalariats. No obstant això, controlar pel tipus de contracte i pel tipus d'empresa no té cap efecte rellevant en el valor estimat del retorn del domini de l'anglès. A la columna 6, en canvi, es mostren els coeficients estimats que s'obtenen incloent en el model de regressió un conjunt d'indicadors (efectes fixos) per tipus d'ocupació⁸: en aquest cas, sí que es produeix una reducció notable en el rendiment salarial del coneixement d'anglès, que passa de ser del 6,4% (columna 5) fins al 4,3% (una reducció del 33% aproximadament). Tot i ser encara significatiu, i econòmicament rellevant, la reducció en el coeficient estimat associat al coneixement d'anglès és substancial i ens indica que part de la correlació condicional entre coneixement lingüístic i salaris es deu al mecanisme ocupacional, és a dir, individus competents en anglès accedeixen a ocupacions més remunerades (gràcies a les seves competències lingüístiques o a les característiques inobservades que aquestes capturen). Aquest resultat confirma l'evidència prèvia, obtinguda per diferents autors que han considerat com el mecanisme ocupacional intervé en la relació entre anglès i salaris (per exemple, Ginsburgh i Prieto-Rodríguez, 2011, Azam *et al.*, 2013, i Di Paolo i Tansel, 2015).

Finalment, amb l'objectiu de valorar l'existència de complementaritats entre coneixement de llengües estrangeres i educació formal, s'estimen els models separatament per a individus amb un nivell educatiu alt (és a dir, estudis secundaris postobligatoris o universitaris) i per a individus amb un nivell educatiu baix (estudis secundaris obligatoris o inferiors), respectivament. Els coeficients seleccionats (referents al coneixement d'anglès) es mostren a la taula 3 (els resultats complets es mostren a les taules 1A i 1B de l'annex).

Com es pot apreciar, hi ha un retorn salarial del coneixement d'anglès només per als individus amb un nivell d'estudis elevat. De fet, els coeficients estimats per a aquest grup són pràctica-

ment iguals als que s'havien obtingut per a tota la mostra⁹, mentre que per als individus amb un nivell d'estudis baix els coeficients estimats referents al coneixement d'anglès no són en cap cas estadísticament diferents de zero. Això indica que l'efecte que s'havia detectat per a tota la mostra es genera només entre els individus amb més instrucció formal. Per tant, aquesta evidència indica l'existència de complementaritats entre educació i competències lingüístiques: és a dir, el coneixement de llengües estrangeres (d'anglès) genera un retorn en el mercat laboral només si es combina amb un nivell d'estudis suficientment elevat per poder accedir a ocupacions on tenir aquest tipus de competències es valora positivament i es retribueix a nivell salarial. Aquest resultat és consistent amb el que s'ha obtingut en referència amb el coneixement d'anglès a l'Índia (Azam *et al.*, 2013) i que també es produeix en relació amb el coneixement de català de la població immigrada a Catalunya (Di Paolo i Raymond, 2012).

Els resultats obtinguts mostren clarament que el coneixement d'anglès està associat a millors resultats laborals a Catalunya, confirmant l'evidència obtinguda per a altres realitats. Cal tornar a mencionar, però, que aquestes correlacions condicionals no necessàriament representen efectes causals, atès que hi pot haver factors inobservats que afecten tant la propensió a saber anglès com els salaris (per exemple, habilitat innata, motivació, qualitat de la instrucció, etc.). Dit d'una altra manera, és possible que les persones que sàpiguen anglès cobrin salaris més elevats no realment gràcies al seu coneixement de la llengua, sinó mercès a les seves capacitats latents (almenys des de la perspectiva de l'investigador). També és possible que el component indirecte de la relació entre llengua i salari, mitjançant les decisions ocupacionals, estigui afectat per l'existència de variables (inobservades) omeses; és a dir, podria ser l'heterogeneïtat individual inobservada que porta els individus a obtenir una millor ocupació (i no pas el domini de l'anglès *per se*) el que genera l'associació estadística condicional entre llengua, ocupació i salaris. Així mateix, els individus amb més nivell educatiu podrien ser els únics capaços de beneficiar-se laboralment del domini de l'anglès només gràcies a les seves millors capacitats innates, que els han permès progressar més en el procés d'acumulació de capital humà. D'altra banda, també seria necessari tenir en compte

⁸ Els valors estimats dels efectes fixos d'ocupació no es mostren per raons d'espai, però estan disponibles sota petició a l'autor.

⁹ En canvi, l'efecte sobre la probabilitat d'estar ocupat s'estima amb menys precisió i és menys pronunciat (en termes d'efecte marginal sobre la probabilitat predita).

Taula 3. Resultats de les estimacions segons nivell educatiu (coeficients seleccionats)

Variable dependent	(1) ln_salarih	(2) ln_salarih	(3) Pr(ocupat)	(4) ln_salarih	(5) ln_salarih	(6) ln_salarih
nivell educatiu alt (educació secundària postobligatòria o terciària)						
anglès	0,063*** (0,021)	0,069*** (0,021)	0,071 (0,069)	0,069*** (0,021)	0,070*** (0,020)	0,048** (0,020)
R ² ajustat	0,241	0,255	--	--	0,294	0,383
Efectes fixos d'ocupació	no	no	no	no	no	si
nombre d'observacions	1.345	1.345	1.699	1.345	1.345	1.345
nivell educatiu baix (educació secundària obligatòria o inferior)						
anglès	-0,023 (0,055)	-0,021 (0,055)	0,026 (0,146)	-0,021 (0,052)	-0,026 (0,056)	0,022 (0,053)
R ² ajustat	0,069	0,076	--	--	0,078	0,159
Efectes fixos d'ocupació	no	no	no	no	no	si
nombre d'observacions	667	667	1.190	667	667	667

Nota: ***significatiu a l'1%, **significatiu al 5%, *significatiu al 10%; errors estàndards robustos a l'heteroscedasticitat entre parèntesis

les altres possibles causes d'inconsistència de les estimacions per MQO, com la possible causalitat inversa o els errors de classificació de la variable que recull el coneixement d'anglès. En definitiva, caldria explorar amb més deteniment l'existència d'una relació causa-efecte entre coneixement d'anglès i resultats laborals, mitjançant tècniques econòmiques més sofisticades que permetin una interpretació causal dels resultats (que, malauradament, no es poden aplicar amb les dades disponibles per al cas català). No obstant això, els (pocs) treballs relacionats existents que tracten el tema de la causalitat de manera rigorosa (per exemple, Bleakley i Chin, 2004) indiquen que els diferents factors d'inconsistència tendeixen a compensar-se l'un amb l'altre, fent que els coeficients estimats amb simple MQO siguin indicatius de l'existència d'un efecte positiu del coneixement lingüístic en els resultats laborals (almenys a nivell qualitatiu).

3. Conclusions

En aquest article s'ha analitzat el valor econòmic del coneixement de llengua anglesa en el mercat laboral català. Les estimacions obtingudes confirmen l'existència d'una clara correlació condicional positiva entre el domini de la llengua anglesa i els resultats laborals a Catalunya. El coeficient estimat indica que saber parlar i escriure en anglès està associat a un major salari, amb un increment del 6,2-6,5%; suposant l'exogeneïtat de les variables explicatives, aquest resultat indica que ser competent en anglès té aproximadament el mateix efecte que una mica més d'un any d'educació formal addicional.

A més a més, resulta que l'impacte positiu sobre els salaris és en certa mesura indirecte, atès que el coneixement de llengües estrangeres afavoreix l'accés a determinades ocupacions que estan més ben remunerades. No obstant això, tot i controlant pel tipus d'ocupació, encara s'obté un coeficient positiu i significatiu associat al domini de la llengua (+4,3%), tot i que més moderat respecte al que s'obté sense condicionar pel tipus d'ocupació. Finalment, les estimacions mostren una forta complementaritat entre educació formal i coneixement d'anglès en el mercat laboral, atès que saber parlar i escriure en anglès genera un rendiment salarial positiu només per als individus amb un nivell d'estudi alt. Això vol dir que sense assolir primer un nivell de formació suficientment elevat per accedir a determinades ocupacions, en què el coneixement d'anglès pot ser rellevant, conèixer aquesta llengua no genera directament un retorn econòmic.

Tot i que no és possible interpretar directament els resultats com a relacions de causalitat, els resultats són prou robustos per considerar que (almenys a nivell qualitatiu) hi ha un efecte positiu del coneixement d'anglès en els resultats laborals, sobretot en termes de millores salarials. Les implicacions d'aquest treball en termes de polítiques públiques són clares: millorant les competències en anglès dels treballadors se'n podria afavorir l'ocupabilitat i promoure uns millors resultats laborals quant al tipus d'ocupació i a la remuneració obtinguda. Segurament la forma més eficient d'assolir aquest objectiu és mitjançant la formació en idiomes estrangers al llarg del procés d'escolaritza-

Taula 1A. Resultats de les estimacions per nivell educatiu alt (secundari postobligatori o terciari)

Variable dependent	(1) ln_salaris	(2) ln_salaris	(3) Pr(ocupat)	(4) ln_salaris	(5) ln_salaris	(6) ln_salaris
Constant	0,627*** (0,074)	0,676*** (0,075)	-1,095*** (0,287)	0,694*** (0,115)	0,754*** (0,075)	1,392*** (0,106)
Dona	-0,064*** (0,019)	-0,043 (0,028)	0,101 (0,119)	-0,043 (0,029)	-0,051* (0,028)	-0,003 (0,028)
Anys educ	0,081*** (0,005)	0,078*** (0,005)	0,102*** (0,016)	0,078*** (0,006)	0,067*** (0,005)	0,033*** (0,006)
Pot exp	0,025*** (0,004)	0,019*** (0,004)	0,083*** (0,014)	0,018*** (0,005)	0,015*** (0,004)	0,014*** (0,004)
Pot exp2	-0,000*** (0,000)	-0,000*** (0,000)	-0,002*** (0,000)	-0,000** (0,000)	-0,000** (0,000)	-0,000** (0,000)
Anglès	0,063*** (0,021)	0,069*** (0,021)	0,071 (0,069)	0,069*** (0,021)	0,070*** (0,020)	0,048** (0,020)
casat		0,117*** (0,037)	0,296* (0,151)	0,116*** (0,035)	0,105*** (0,036)	0,101*** (0,036)
nfills		0,015 (0,018)	0,133* (0,079)	0,015 (0,015)	0,021 (0,018)	0,017 (0,017)
dona*casat		-0,055 (0,049)	-0,401** (0,175)	-0,053 (0,051)	-0,066 (0,048)	-0,083* (0,047)
dona*nfills		0,004 (0,025)	-0,161* (0,092)	0,005 (0,023)	-0,000 (0,024)	0,002 (0,023)
public					0,154*** (0,020)	0,156*** (0,024)
contr fix					0,122*** (0,025)	0,090*** (0,025)
auton empr					0,051 (0,045)	0,069 (0,046)
atur 16			-0,018*** (0,006)			
corr(au)			-0,044 (0,215)			
Efectes fixos d'ocupació	no	no	no	no	no	sí
R ² ajustat	0,241	0,255	--	--	0,294	0,383
nombre d'observacions	1.345	1.345	1.699	1.345	1.345	1.345

Nota: ***significatiu a l'1%, **significatiu al 5%, *significatiu al 10%; errors estàndards robustos a l'heteroscedasticitat entre parèntesis

ció. A més a més, possiblement s'hauria de fomentar relativament més l'aprenentatge de llengües estrangeres durant l'escolarització obligatòria, atès que l'exposició primerenca a llengües diferents de la nativa n'afavoreix molt més l'aprenentatge (fins i tot arribant a assolir competències semblants als nadius) que si l'exposició es produeix després de passar "l'edat crítica" per a l'assimilació lingüística (que se situa al voltant dels 12-14 anys d'edat, Bleakley i Chin, 2004). Pel que fa a la població adulta, també seria interessant fomentar l'adquisició de competències en llengües estrangeres mitjançant cursos de formació contínua, fins i tot fets dins de la mateixa empresa. Així mateix, l'Estat i les institucions locals podrien incentivar la participació de la població aturada en cursos d'idiomes gratuïts (o semigratuïts), atès que el coneixement de llengües sembla millorar també les possibilitats de tenir feina. No obstant això, els resultats referents a la complementaritat entre coneixement de llengües i educació formal també s'han de tenir en compte

a l'hora de plantejar polítiques orientades a millorar l'ocupabilitat dels treballadors mitjançant cursos d'idiomes: si primer no es promou l'assoliment d'un nivell educatiu suficientment elevat, reduint l'abandonament escolar i incrementant la participació en educació postobligatòria (ja sigui via formació professional o universitària), difícilment cap política orientada a l'aprenentatge de llengües podrà ser efectiva i eficient per afavorir l'èxit en el mercat laboral.

Referències

- Adesope O. O., Lavin T., Thompson, T., Ungerleider C.** (2010). "A systematic review and meta-analysis of the cognitive correlates of bilingualism". *Review of Educational Research*, 80 (2): 207-245.
- Azam, M., Chin, A., Prakash, N.**, (2013). "The returns to english-language skills in India". *Economic Development and Cultural Change*, 61(2): 335-367.
- Bleakley, H., Chin, A.** (2004). "Language skills and earnings: Evidence from childhood immigrants". *Review of Economics and Statistics*, 86(2), 481-496.