The Effect of a Change in Language of Instruction on the Returns to Schooling in Morocco

Author(s): Joshua D. Angrist and Victor Lavy

Source: *Journal of Labor Economics*, Jan., 1997, Vol. 15, No. 1, Part 2: Essays in Honor of Yoram Ben-Porath (Jan., 1997), pp. S48-S76

Published by: The University of Chicago Press on behalf of the Society of Labor Economists and the NORC at the University of Chicago

Stable URL: http://www.jstor.com/stable/2535401

The Effect of a Change in Language of Instruction on the Returns to Schooling in Morocco

Joshua D. Angrist, *Massachusetts Institute of Technology, Hebrew University, and National Bureau of Economic Research*

Victor Lavy, *Hebrew University*

Until 1983, the language of instruction for most subjects in grades 6 and above in Moroccan public schools was French. Beginning in 1983, the language of instruction for new cohorts of Moroccan sixth graders was switched to Arabic. We use this policy change to estimate the effect of French language skills on test scores and earnings. The estimates suggest that the elimination of compulsory French instruction led to a substantial reduction in the returns to schooling for Moroccans affected by the change. This reduction appears to be largely attributable to a loss of French writing skills.

Many of the countries that gained independence in the 1950s and 1960s have only recently implemented programs designed to actively promote the national language at the expense of the colonial language. Although

This article was prepared for the October 1993 conference held at Hebrew University to commemorate the 1-year anniversary of the death of Yoram Ben-Porath. Thanks go to the conference participants and to Katherine Kleynhoff, Jamil Salmi, and seminar participants at the World Bank for helpful advice and comments. Thanks also go to the Direction de la Statistique in Morocco for providing us with Labor Force Survey data and to Doris Damian, Alex Mogel, and Dan Nabot for expert research assistance. The views expressed here are ours alone and are not meant to reflect the views of the World Bank or any government agency.

[*Journal of Labor Economics*, 1997, vol. 15, no. 1, pt. 2]

0734-306X/97/1501-0010$01.50

Morocco became independent of France in 1956, it was only in the 1980s that French was replaced by Arabic as the main language of instruction in Moroccan secondary schools. Recently, Algeria and Tunisia have also embarked on similar reform programs. In 1991, for example, an Algerian law was implemented requiring the use of Arabic as the exclusive language of instruction at all levels of the educational system.

The long delay in switching the language of instruction in the Maghreb has been attributed to the costs and social upheavals associated with a language transition. In many North African countries, even after independence, teachers were often French citizens or natives who had been trained in French schools. For example, replacement of French primary school teachers was not completed in Morocco until 1974 (Spencer 1980). Even after the switch to native-born teachers, the use of textbooks written in French has persisted. Moreover, educated labor force participants have traditionally found employment in the French-trained civil service or in an economic sector dominated by trade with French-speaking countries (Elbiad 1991).[1] Finally, Berber and other North African minority groups often prefer French to Arabic in official business and have resisted policies that make the use of Arabic compulsory (Heggoy 1990).

The belated but ongoing efforts to promote native languages in the Maghreb suggest that language is an important symbol of independence and national sovereignty.[2] Governments appear to be willing to tolerate the many social costs of a language transition in return for the perceived social benefits. But the consequences of a language transition for *individual* labor market outcomes are unclear. School quality and the economic returns to schooling could be either increased or decreased by a change in language of instruction or increased for some students but decreased for others. On the one hand, students must fit into the existing economic and social systems, dominated by French language and culture. On the other hand, the insistence on instruction in a foreign language may be a barrier to poor or rural students. In general, students may be able to learn more effectively when school instruction is in the native language. Also, strong native language skills may prove to be of more enduring value in the local labor market than French language skills. In the short term, however, language reforms appear to have contributed to the popular

[1] Elbiad (1991) reports that, even after the language reform, French was still used almost exclusively in public administration, ministerial offices, and technical and scientific sectors of the Moroccan economy.

[2] See, e.g., Weinstein (1990). In a discussion of English teaching in Moroccan high schools, Aarim (1992, p. 53) notes, "The leadership of the Moroccan National Movement at that time viewed the mastering of Arabic and the instruction of the population along patriotic lines as the only means of achieving Morocco's complete independence."

impression that Moroccan schools "are now turning out bilingual illiterates" (Moatassime 1992, p. 178).

The effect of language transition is of interest not only because of its immediate policy relevance in the Maghrebi countries but also because of long-standing interest in the economic value of the well-defined human capital embodied in language skills. Research on the economic value of language includes work by McManus, Gould, and Welch (1983), Chiswick (1991, 1993), Grenier (1984), Kossoudji (1988), Tainer (1988), and Dustmann (1994). Of particular concern in the economics literature is the possibility that individuals with better language skills may earn more for reasons other than those skills (see, e.g., Chiswick and Miller 1993). For example, if the best French speakers in Morocco always come from wealthy families, then the positive correlation between language skills and earnings is not necessarily indicative of a causal relationship. In this study, we use the sharp change in the language of instruction to generate instrumental variables that identify the relationship between language skills and earnings.

This article is organized as follows. Section I presents background information on the Moroccan school system and the Arabization program. The two micro data sets used are described in Section II. In Section III, we present ordinary least squares (OLS) estimates of the effect of language training on earnings. Section IV reports OLS estimates of the link between Arabization and test scores as well as instrumental variables estimates that link differences in test scores to differences in earnings. Overall, the estimates suggest that elimination of compulsory French instruction led to substantial reductions in earnings and in the returns to schooling for young Moroccans with more than a primary school education. The main avenue for this effect appears to be a sharp reduction in French writing skills among students who were taught in Arabic instead of French.

I. Schooling and Language in Morocco

In the 1960s, the Moroccan Ministry of Education initiated a policy designed to replace foreign teachers in Moroccan schools (primarily French citizens) with Moroccan teachers. This policy was followed by the replacement of French by Arabic as the language of instruction in middle and secondary schools. This latter change was called "Arabization" and was implemented in 1983.[3] Until the fall of 1983, the language

[3] The details of the policy change and its implementation are described in several internal documents (dated from 1980 to 1985) of the Ministry of Education in Morocco and are summarized in several World Bank reports (see, e.g., World Bank 1985).

of instruction for all subjects taught in Moroccan postprimary schools, except those directly tied to Arabic (such as Arabic literature or religion), was French.[4] The language of instruction in primary school was and remains Arabic.

As in many French-speaking systems, until 1990 Moroccan primary schools included grades 1–5, middle school included grades 6–9, and secondary school included grades 10–12.[5] Arabization was first implemented in the 1983–84 school year for all students entering sixth grade, the first year of middle school. The change in the language of instruction was then rolled ahead one additional grade in each new school year, so that by 1986–87 all middle school grades were affected by the new regime. By 1989–90, a similar change had been accomplished in all secondary schools. In the fall of 1990, members of the first cohort of students who had studied exclusively in Arabic entered universities.[6] Thus, under the Arabization reform schedule, a student entering sixth grade in September 1983 was affected by the Arabization program; he or she was taught most subjects in Arabic. Any student entering sixth grade after September 1983 was also affected, while any student completing sixth grade before September 1983 was taught in French.

School attendance in Morocco is compulsory from ages 7 to 16, corresponding to grades 1–9. In practice, however, some children enter first grade at age 6. For example, in 1983–84, 16% of first graders were 6 years old, and in 1986–87, 25% of first graders were aged 6. Many children do not attend school at all or drop out before ninth grade (Khandker, Lavy, and Filmer 1993). Data from the Living Standard Measurement and Literacy Survey (LSMS) analyzed here also suggest that at least half of enrolled

[4] Students in the Koranic (religious) school system study all subjects in Arabic. Among students in our survey population, roughly 5% studied in the Koranic system.

[5] Starting in September 1990, sixth grade was moved to primary school, and middle school reduced to cover grades 7–9. Also since 1990, the parallel system of Koranic education, which traditionally covered literacy and religious studies through the university level, has for the most part been subsumed in special tracks within the public school system. Koranic schools, however, still represent the principal form of preschool education in Morocco.

[6] The language of instruction in Moroccan universities is still French. In Algeria, where Arabization has been more abrupt than in Morocco, Arabic instruction was extended to universities in 1991 (Moatassime 1992). In Morocco, the language switch in middle and secondary schools was accompanied by steps designed to ease the transition. For example, special Arabic training was offered to teachers of subjects that had previously been taught in French. Science teachers maintained the use of international scientific notation, and an effort was made to prepare textbooks in Arabic. Since the policy change, French language is taught as a subject from the third grade of primary school onwards for 2 hours a week and for 6 hours a week in middle and secondary school.

children repeat at least one primary grade. The reason for the prevalence of grade repetition appears to be that the number of places in middle school (entering in grade 6) is limited.[7] In the past, many parents kept their children in primary school until space in middle school became available. Beginning in 1984, however, opportunities to repeat grades were severely limited. This policy change also has implications for our estimation strategy, which is discussed below.

Our cross-section data sets do not have information on age at school entry. To identify cohorts likely to have been affected by the Arabization policy, we assume that most students enter first grade at age 7. We then use a fixed-age cutoff based on the assumption that most students repeat one or two grades of primary school.[8] The timeline in table 1 illustrates this cutoff. A child entering sixth grade in the fall of 1983, who progressed through primary school without skipping a grade, would have entered first grade in the fall of 1978. Assuming that the child was 7 when entering first grade, he or she will have turned 20 by the end of 1991, the year in which most of the survey interviews were conducted. More realistically, the child is likely to have repeated at least one primary school grade. As shown in row 2 of table 1, a child entering sixth grade in 1983, who began first grade at age 7 and repeated fifth grade once, will have entered sixth grade at the age of 13. This child would have turned 21 by the end of 1991. As shown in row 3, a child who repeated fifth grade twice would have entered sixth grade at age 14 and turned 22 by the end of 1991.

In the empirical work, we assumed that individuals 21 and younger who were interviewed in 1991 and 20 and younger who were interviewed in 1990 were potentially affected by the Arabization policy. In principle, a child who turned 7 in the last quarter of 1976 and waited until the fall of 1977 to enter first grade could have reached the age of 22 if he or she skipped a single primary grade and was interviewed very late in 1991. A cutoff indicator based on age 21 in 1991, however, turns out to be more highly correlated with variables likely to be affected by the Arabization policy. The treatment dummy for Arabization is the interaction of two dummy variables: one to indicate being below the 21-in-1991 age cutoff and one to indicate having 6 or more years of school enrollment.

[7] Many high school students also repeat twelfth grade in multiple attempts to pass the high school baccalaureate exam.

[8] The LSMS data set does have information on the number of years enrolled at each schooling level. Grade repetition, however, may be correlated with unobserved variables so that coding treatment status using information on grade repetition could lead to misleading estimates of policy effects.

Table 1
Schooling Timeline

	Year															
	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991
1. No grade repetition:																
Grade	...	...	1	2	3	4	5	6	7	8	9	10	11	12	...	...
Age at last birthday	...	...	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2. Repeat fifth grade:																
Grade	...	1	2	3	4	5	5	6	7	8	9	10	11	12	...	...
Age at last birthday	...	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
3. Repeat fifth grade twice:																
Grade	1	2	3	4	5	5	5	6	7	8	9	10	11	12	...	...
Age at last birthday	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22

NOTE.—School begins in September and age is calculated assuming a student's birthday falls in the first 3 quarters of the year. School is compulsory beginning with first grade at age 7.

II. Data and Descriptive Analysis

Data from two sources are used to estimate the relationship between language of instruction and earnings. The first is a sample drawn from the Moroccan Labor Force Surveys (LFS) conducted in 1990 and 1991. The LFS is an annual cross-section survey in which basic labor market and demographic information is collected. The sample used here consists of 4,426 observations on men aged 15–45 years in 1991 pooled with 2,614 observations on men aged 15–45 years in 1990 for a total of 7,040 observations. We restrict the sample to men under 45 because there are few female wage earners and because the estimation strategy involves a comparison of workers aged 21 or less with a group of older workers.

The second sample is drawn from the 1991 Living Standard Measurement and Literacy Survey. The LSMS is a comprehensive household survey conducted by the Moroccan Direction de la Statistique with technical assistance from the World Bank. Unlike the LFS, the LSMS is not restricted to urban households. Another important feature of the LSMS is that it contains a subsample of individuals who were given tests to assess proficiency in French and Arabic reading and writing and in mathematics. Questions for this "literacy subsample" also include self-assessments of basic skills. The LSMS extracts used here consist of 1,166 observations on men aged 15–45 in 1991 who had earnings and 1,377 observations on men with valid test scores. There are 581 observations on men aged 15–45 with information on both earnings and test scores. Additional details on both the LFS and LSMS are provided in the data appendix.

Descriptive Statistics

Descriptive statistics for the LFS and LSMS extracts are reported in table 2. These statistics show that the average age in both data sets is around 30. The LSMS extract is 62.3% urban, although the full sample (which is not limited to wage earners) is only about half urban. The average potential experience (age − highest grade completed − 7) is 17–18 years. The LFS contains a variable recording the respondent's age at first job. Using this variable, we estimate an average of 13 years labor market experience. Differences between the two experience measures are probably explained by the fact that, because of a high rate of grade repetition, the highest grade enrolled undercounts actual years in school. Also, the potential experience variable does not allow for a period of unemployment between leaving school and starting work.

A large fraction of both samples, 20% in the LFS and 28% in the LSMS, have no formal schooling of any kind. Between 32% and 36% have some primary schooling, and roughly 19% have been to middle school (grades 6–9). Between 12% and 15% attended the second level of middle school (high school). Roughly 11% of the LFS sample has 1 or

Table 2
Descriptive Statistics

Variable	LFS (1)	LSMS	
		Full Sample (2)	Urban (3)
Age	31.0	30.3	31.6
	(7.6)	(8.2)	(7.8)
Urban	1.0	.623	1.0
Experience:			
Potential experience	17.4	17.4	17.5
	(8.4)	(9.1)	(8.6)
Years since first job	13.3	...	...
	(8.5)		
Highest level of schooling enrolled:			
None	.20	.281	.171
Primary (1–5)	.357	.323	.293
Secondary I (6–9)	.186	.190	.237
Secondary II (10–12)	.150	.125	.175
Some college (13–15)	.107	.041	.061
College graduate (16+)	...	.040	.0634
Young:			
Age less than 22 in 1991[a]	.133	.188	.132
Age less than 21 in 1990[b]	.110	...	...
Young*Schooling:			
None	.014	.054	.015
Primary (1–5)	.081	.089	.073
Secondary I (6–9)	.025	.040	.040
Secondary II (10–12)	.004	.004	.003
Some college (13–15)	.0001	.001	.001
Hourly wage	9.14	9.88	12.74
	(14.1)	(17.3)	(20.1)
Log hourly wage	1.76	1.69	1.97
	(.88)	(1.03)	(1.06)
Log monthly wage	7.08	...	...
	(.72)		
No. of observations	7,040	1,166	726

NOTE.—LFS = Labor Force Survey and LSMS = Living Standard Measurement and Literacy Survey. Standard errors are in parentheses. Samples include men aged 15–45 with wage and salary earnings, who were not educated in religious (Koranic) institutions. The samples also exclude men with an hourly wage above 400 dirams. The LFS sample includes 4,426 observations interviewed in 1991. The potential experience variable in the LFS is age − education − 7, where education is years of schooling assigned to schooling levels. The potential experience variable in the LSMS is calculated using information on years enrolled at each schooling level. Statistics are unweighted.

[a] Percentage of 1991 observations.

[b] Percentage of 1990 observations.

more years of college education, and about 8% of the LSMS sample has some college education.

An average 12% (13.3% in 1991 and 11% in 1990) of the LFS and 19% of the LSMS are defined here to be "young," that is, subject to a change in language of instruction while in school. Sixty-five percent of the young sample in the LFS (.081/.124) and 47% of the young sample in the LSMS (.089/.188) have a primary school education. These groups

were less likely to be affected by the Arabization policy because the change in language of instruction affected only those students who went to middle school or beyond. The most important group affected by Arabization consists of those in the young samples who have at least a middle school education. There are 205 such individuals in the LFS and 52 such individuals in the LSMS sample with earnings data. There are 63 young men with at least a middle school education in the larger LSMS subsample with valid test scores (1,377 observations total).

A few other descriptive facts are relevant to this project. First, distributions of schooling by single year (not reported in table 2) show that men in the Young group have more schooling at low levels of schooling. For example, the proportion of Young workers with 5 years of schooling (which means completion of primary school) is 22.8% versus 19.1% for Old workers. The transition rate from primary to middle school went up for younger cohorts as well: from 0.15 (2.96/19.11) for the Old group to 0.26 (5.94/22.84) for the Young group.[9] Second, the overall share of employment of educated workers was rising in the 1980s (Kingdom of Morocco 1990, table 1.1) while the share of employment in the public sector, which traditionally employs more educated workers, increased slightly (Kingdom of Morocco 1990, table 3.5). Thus, sectors where educated workers were most likely to be employed appear to have experienced stable or growing employment shares at the time Arabization cohorts entered the labor market.

Differences-in-Differences Estimates

A differences-in-differences analysis offers the simplest strategy for the evaluation of the impact of Arabization on earnings and test scores. For illustration, we use the log hourly wage variable in the LSMS and a subsample of 418 observations on men aged 15–45 with at least some formal schooling. Among men aged 22 or more in this group, the difference in log earnings between those with 6 or more years of schooling and those with less than 6 years of schooling is 1.07. Among men aged 21 or less in this sample, the difference in log earnings between those with 6 or more years of schooling and those with less than 6 years of schooling is only .137. The difference in these two differences, −.933 (SE = .234), is an estimate of the effect of being subject to the Arabization program that controls for secular age and schooling group effects.[10]

[9] A similar pattern emerges when we do not condition on being a wage earner in the LSMS data; it is also visible in administrative data from the Moroccan Ministry of Education (summarized in World Bank [1989, 1991]).

[10] The differences-in-differences technique can be motivated by the regression model

$$y_i = X_i\beta + d_i\alpha_1 + s_i\alpha_2 + (d_i \cdot s_i)\gamma + \varepsilon_i,$$

A major problem with this calculation is that among younger men, the loss of labor market experience associated with more schooling will be more important than the loss of experience from schooling in the older group. Thus, it is unlikely that the only reason the returns to schooling are lower for the young is a change in the language of instruction.

A simple check on whether the differences-in-differences estimate has *any* connection with a change in the language of instruction is to see whether the decline in the economic returns to schooling across age groups is paralleled by a decline in test scores. For illustration, we use the French writing index (the index takes on values 0, 1, and 2 and is described below and in the data appendix). Among older men, the average difference in the writing index by schooling group is 1.44, while among younger men the average difference in the writing index by schooling group is only .635. The difference in these two differences, −.805 (SE = .171), probably reflects the effect of the Arabization program on French writing skills.

III. Earnings Equations

In this section we discuss regression estimates of the effect of Arabization on earnings that control for experience and other covariates. Interaction terms for being aged 21 or younger in 1991 and having been enrolled in grade 6 or above indicate the group of students most likely to have been affected by the Arabization program. The effects of interest are estimated using the following equation:

$$y_i = X_i\beta + d_i\alpha_1 + \sum_j s_{ji}\alpha_{2j} + \sum_k z_{ki}\gamma_k + \varepsilon_i, \qquad (1)$$

where j indexes schooling dummies for highest grade enrolled, s_{ji}, and j = 1 for grades 1–5 (primary school), 2 for grades 6–9 (middle school), 3 for grades 10–12 (high school), 4 for grades 13–15 (some college), 5 for grades 16+ (college graduates). The variables $z_{ki} = s_{ki} \cdot d_i$ are interactions between the schooling dummies and a dummy for age less than 21, d_i. Individuals affected by the change in language of instruction have either z_{2i}, z_{3i}, or z_{4i} equal to one (there are no young college graduates). Other terms in equation (1) include a vector of additional covariates, X_i, and an error, ε_i. The vector X_i includes various controls for age and experience.

---

where y_i is the dependent variable, d_i is a dummy for being young, s_i is a dummy for having 6 or more years of schooling, and ε_i is an error term. The OLS estimates of the coefficient γ gives the differences-in-differences estimate.

A. Results from the LFS

The results of estimating equation (1) using LFS data are reported in table 3 for two dependent variables: log hourly wages and log monthly wages. Column 1 reports coefficients from an equation that includes a quadratic in age, an urban dummy, schooling-group dummies, a dummy for being in the young sample, and Young*Schooling interaction terms. The omitted schooling group has no formal education. The schooling coefficients in column 1 show an hourly wage premium of .17 for men with a primary school education, increasing to 1.65 for men with some college. The coefficients are somewhat smaller in column 2, which shows results for monthly wages.

In the LFS, labor market experience can be measured as years since the respondent's first job. Results from replacing the age quadratic with an experience quadratic are reported in columns 3 and 4. As is usual in regressions of this sort, schooling coefficients increase when experience replaces age in the list of covariates.

The coefficients of interest in table 3 are the interaction terms between the dummy for being young and the dummies for schooling groups. Results in the first two columns show a marginally significant reduced premium for young men with a primary school education. The middle school interaction term shows that the schooling premium is substantially and significantly reduced for young men with 6–9 years of schooling. The reduced primary school effects in columns 1 and 2 are probably not directly attributable to the Arabization program, which affected students enrolled in grade 6 or higher.

One possible explanation for the negative Young*Primary interaction term is the failure to control for the reduced labor market experience of the young primary-educated cohort. The results in column 3 and 4 show that once labor market experience (years since first job) is included in the regression, the Young*Primary interaction term is reduced to zero. This result favors the interpretation of other interaction terms as measuring the earnings effect of the Arabization program. Note also that the interaction term identifying young men with 6–9 years of schooling falls by only about 25% when experience replaces age in the list of covariates and is significantly different from zero. The results in column 3 suggest that the Arabization program reduced the middle school wage premium by roughly 27% for hourly wages and by 19% for monthly wages. Almost all of the other Young*schooling-group interaction terms are negative as well.

Alternate summary statistics measuring the earnings effect of the Arabization program are presented in column 5. These results are from equations where schooling-group dummies have been replaced by years of schooling (highest-grade enrolled). Note that we do not have a variable

Table 3
Wage Equations, Moroccan Men Aged 15–45 in the LFS

Regressor	Hourly Wage (1)	Monthly Wage (2)	Hourly Wage (3)	Monthly Wage (4)		Hourly Wage (5)	Hourly Wage (6)
Year = 91	.055	.077	.071	.091		.068	.054
	(.015)	(.012)	(.015)	(.012)		(.016)	(.015)
Experience	...	...	.057	.052		.052	.076
			(.003)	(.003)		(.003)	(.004)
Experience^2	...	...	−.001	−.001		−.001	−.001
			(.0001)	(.0001)		(.0001)	(.0001)
Age	.127	.119	...	...		...	...
	(.011)	(.009)					
Age^2	−.001	−.001	...	...		...	...
	(.0002)	(.0001)					
Schooling groups:							
1–5	.171	.111	.185	.133	Linear schooling	.062	.086
	(.021)	(.018)	(.022)	(.019)		(.005)	(.005)
6–9	.471	.339	.584	.453	Linear schooling	.117	.135
	(.024)	(.020)	(.025)	(.021)		(.002)	(.002)
10–12	.882	.653	1.07	.840		...	...
	(.025)	(.020)	(.027)	(.022)			
13+	1.65	1.16	1.93	1.43		...	...
	(.027)	(.022)	(.030)	(.025)			
Young	.058	.020	−.387	−.378		−.280	−.092
	(.068)	(.057)	(.063)	(.052)		(.063)	(.063)
Young × Schooling:							
1–5	−.116	−.090	−.011	−.004	Young × linear	−.026	−.011
	(.066)	(.055)	(.068)	(.057)		(.014)	(.014)
6–9	−.359	−.263	−.269	−.190	Young × linear	−.065	−.045
	(.077)	(.063)	(.079)	(.066)		(.009)	(.009)
10–12	−.329	−.085	−.218	.005		...	...
	(.131)	(.109)	(.135)	(.112)			
13+	−.510	−.406	−.500	−.409		...	...
	(.590)	(.489)	(.605)	(.502)			
R^2	.56	.55	.54	.523		.50	.53

NOTE.—LFS = Labor Force Survey. There are 7,040 observations in the LFS sample. Standard errors are in parentheses. In cols. 1–5, experience is measured as years since first job; in col. 6, experience is measured as age − imputed years of schooling − 7.

measuring actual years of schooling. As an approximation, we imputed years of schooling based on the types of school enrolled. To construct the estimates reported in column 5, the relationship between years of schooling and wages was allowed to differ according to whether the respondent has more or less than 6 years of schooling and whether the respondent is Young. As before, the regressions include an experience quadratic, where experience is measured as years since first job.

The linear-schooling parameterization of the hourly wage equation generates schooling coefficients of .062 for those with less than 6 years of schooling and .117 for those with 6 or more years. We cannot reject the hypothesis of no Young*Schooling interaction term for those with less than 6 years of schooling. This interaction term is analogous to the Young*Primary interaction term discussed previously. In contrast, there is a large and very precisely estimated age difference in the linear returns to schooling for those with 6 or more years of schooling (−.065 with an SE of .009).

The last column in table 3 reports the results of estimation of the linear-schooling model with the modification that experience is computed as potential experience. Potential experience is measured as age − imputed years of schooling − 7. The results in column 6, like those in column 5, show no difference in schooling coefficients by age for those with less than 6 years of schooling. For those with more than 6 years of schooling, the return to schooling is estimated to have fallen by .045 (from .135 to .09) for younger cohorts. Thus, the Arabization program can be interpreted as cutting the economic return to a single year of postprimary schooling by one-third to one-half.[11]

B. Results from the LSMS

The results of estimating equation (1) using LSMS data are reported in columns 1 and 2 of table 4. Schooling coefficients and interaction terms from the LSMS are considerably larger than those computed using the LFS. This may be because the schooling variable is of higher quality in the LSMS data. When equation (1) is estimated in the LSMS with an age quadratic, the schooling-group interaction with a Young dummy is −.364 (SE = .124) for men with a primary education and substantially larger and negative for men with 6–9 and 10–12 years of education.

[11] We also estimated models analogous to those in cols. 1–4 using a dummy-variable parameterization of schooling combined with imputed potential experience. These models generate insignificant interaction terms between schooling and Young at all schooling levels. Note, however, that the LFS schooling variable is a poor measure of years of completed schooling so that the potential experience control for labor market experience is problematic in the LFS data. Problems with LFS schooling variables are discussed in the data appendix.

Table 4
Reduced Forms for Moroccan Men Aged 15–45 in the LSMS

	Hourly Wages		Test Scores				
Regressor	(1)	(2)	French Writing (3)	French Reading (4)	Arabic Writing (5)	Arabic Reading (6)	Math (7)
Urban	.150	.156	.160	.135	.062	.134	.193
	(.046)	(.045)	(.031)	(.041)	(.039)	(.040)	(.033)
Experience	...	.105	...	...	...	...	...
		(.012)					
Experience^2	...	−.002	...	...	...	...	...
		(.0003)					
Age	.073	...	.021	.088	.066	.090	.019
	(.030)		(.015)	(.021)	(.021)	(.021)	(.017)
Age^2	−.0005	...	−.0002	−.001	−.001	−.001	−.0003
	(.0005)		(.0002)	(.0003)	(.0003)	(.0003)	(.0003)
Schooling groups:							
1–5	.285	.382	.118	.897	1.15	.897	.658
	(.059)	(.062)	(.027)	(.052)	(.052)	(.052)	(.037)
6–9	.800	1.06	1.17	1.62	1.85	1.63	1.23
	(.069)	(.073)	(.075)	(.053)	(.038)	(.051)	(.054)
10–12	1.07	1.49	1.59	1.84	1.92	1.84	1.60
	(.074)	(.083)	(.072)	(.037)	(.029)	(.037)	(.050)
13–15	1.80	2.36	1.89	1.86	1.91	1.86	1.80
	(.110)	(.120)	(.026)	(.032)	(.030)	(.032)	(.030)
16+	2.18	2.90	1.84	1.84	1.90	1.84	1.77
	(.110)	(.124)	(.026)	(.035)	(.033)	(.034)	(.031)
Young	.212	.082	.168	.264	.207	.271	.086
	(.132)	(.107)	(.054)	(.074)	(.078)	(.074)	(.061)
Young × Schooling:							
1–5	−.364	−.153	−.142	−.346	−.240	−.358	−.171
	(.124)	(.129)	(.031)	(.080)	(.088)	(.079)	(.059)
6–9	−.923	−.604	−.776	−.052	−.0005	−.060	−.032
	(.147)	(.157)	(.117)	(.088)	(.525)	(.087)	(.088)
10–12	−1.003	−.621	−.908	−.369	−.414	−.371	−.307
	(.324)	(.329)	(.367)	(.289)	(.299)	(.290)	(.291)
13+	−.190	.154	−.029	.021	−.002	.023	−.016
	(.693)	(.695)	(.020)	(.029)	(.033)	(.029)	(.029)
N	1,166		1,377				
R^2	.57	.58	.69	.65	.68	.66	.67

NOTE.—Experience is measured as age − (highest grade completed) − 7. Standard errors are in parentheses; those in cols. 3–7 are heteroscedasticity-consistent.

Replacing age with a potential experience quadratic generates estimates of the Young*Primary interaction term that are not significantly different from zero. In contrast, interaction terms for the two secondary school groups remain large, negative, and significantly different from zero. The interaction term for young men with 6–9 years of a schooling shows a reduction of −.604 in the schooling premium and the interaction term for young men with 10–12 years of schooling shows a reduction of −.621 in the schooling premium. The latter estimate should probably be discounted, however, because there are only five young men with 10–12 years of schooling in the LSMS extract. Note that, even though the LSMS schooling coefficients are much larger than the LFS schooling coefficients, the significant Young*Schooling interaction terms in models that control for experience amount to roughly half the returns to schooling in both data sets.[12]

As a check on the wage equations reported in tables 3 and 4, we also estimated equations that allow for additional interaction terms between experience and schooling. Some of these additional results are reported in table 5. Columns 1–3 of the table report results from the LSMS, and columns 4–6 report results from the LFS.

The specification reported in columns 1 and 4 of table 5 is the same as the specification with an experience quadratic reported in tables 3 and 4, with the addition of an interaction between experience and years of schooling. The second variation, reported in columns 2 and 5, adds age and an age*experience interaction to the first variation. The third variation drops the schooling*experience interaction but adds age, age^2, and a full set of quadratic age*experience interactions (age*experience, age^{2}experience, age*experience^2). In all three specifications, the size of the interactions between Young and schooling dummies are reduced while the SEs are less affected. However, the basic finding of a lower average return to schooling for young individuals "treated" by the Arabization program is unchanged. For example, the coefficient of the interaction between Young and 6–9 years of schooling is negative and marginally significant in the first two variations using the LSMS and the last variation using the LFS.

[12] The LSMS experience coefficients of .097 and −.002 in col. 2 of table 4 are remarkably similar to the experience coefficients estimated by Angrist (1990, table 5) using social security earnings records for white American men in their thirties. These estimates are .1022 and −.0027. The absence of any significant Young dummy in col. 2 of table 4 suggests that the potential experience model captures all relevant age effects in the LSMS. The LFS experience coefficients in the potential experience model are somewhat smaller than the corresponding LSMS coefficients (col. 6 of table 3 for the LFS). This may be due to the lower-quality schooling variables in the LFS.

Table 5
Variations on the Wage Equations

	Data Set					
	LSMS			LFS		
Regressor	(1)	(2)	(3)	(4)	(5)	(6)
Experience	.021	−.028	.417	.031	.153	.015
	(.018)	(.023)	(.149)	(.005)	(.006)	(.006)
Experience^2	−.0002	−.001	.0006	−.0001	−.0001	.002
	(.0004)	(.0004)	(.007)	(.0001)	(.0002)	(.0002)
Age	...	.098	.023	...	.062	.160
		(.029)	(.113)		(.005)	(.028)
Age^2	...	...	.005	...	...	−.002
			(.003)			(.001)
Schooling:						
1–5	−.133	−.281	−.039	−.006	−.086	−.095
	(.101)	(.109)	(.140)	(.035)	(.035)	(.022)
6–9	.163	−.159	.100	.304	.092	.361
	(.157)	(.183)	(.208)	(.046)	(.047)	(.027)
10–12	.381	−.132	−.011	.723	.416	.765
	(.190)	(.242)	(.249)	(.055)	(.058)	(.029)
13–15	1.139	.427	.316	1.549	1.129	1.536
	(.223)	(.305)	(.305)	(.060)	(.066)	(.039)
16+	1.578	.633	.129	...	...	...
	(.239)	(.365)	(.384)			
Young	−.296	−.074	−.197	−.557	−.254	.002
	(.121)	(.137)	(.154)	(.067)	(.067)	(.071)
Young × Schooling:						
1–5	−.026	−.050	.125	.091	.089	−.014
	(.189)	(.128)	(.137)	(.069)	(.067)	(.067)
6–9	−.294	−.371	−.088	−.110	−.038	−.208
	(.161)	(.162)	(.188)	(.081)	(.079)	(.079)
10–12	−.215	−.230	.093	−.004	.086	−.174
	(.329)	(.329)	(.354)	(.137)	(.133)	(.130)
13+	.606	.804	1.715	−.263	−.036	−.332
	(.686)	(.686)	(.726)	(.604)	(.585)	(.586)
Schooling × Experience	.005	.002	...	.0019	.002	...
	(.001)	(.001)		(.0003)	(.0003)	
Age × Experience	...	...	.005	...	.0011	−.001
			(.003)		(.0003)	(.003)
N		1,166			7,040	
R^2	.59	.60	.61	.55	.57	.57

NOTE.—LSMS = Living Standard Measurement and Literacy Survey and LFS = Labor Force Survey. The regressions reported in cols. 3 and 6 also include as regressors the following interactions: Age^2 × Experience, Age × Experience^2, and Age^2 × Experience^2. These coefficients, as well as the coefficients on an urban dummy in cols. 1–3 and a year = 91 dummy in cols. 4–6, are not reported in the table. In the LSMS sample, the effect of the Age*Experience interaction cannot be estimated separately since experience is computed as age − schooling − 7. Standard errors are in parentheses.

The principal finding from the analysis of earnings appears to be roughly consistent across both the LFS and the LSMS data sets and two dependent variables: controlling for labor market experience, there is no difference across age cohorts in the hourly or monthly wage premium for having a primary school education. But for men who went to middle

school after the Arabization program was initiated, there is a large and often statistically significant decline in the economic payoff to having 6–9 or 10–12 years of schooling. The estimates that are most comparable across data sets are those in column 3 of table 3 (hourly wages, controlling for experience in the LFS) and those in column 2 of table 4 (hourly wages, controlling for experience in the LSMS). The Young*Schooling interaction terms amount to about half the returns to middle school (6–9 grades) in both data sets.[13] The estimates for 10–12 years of schooling amount to 20% of the returns to schooling in the LFS and 40% in the LSMS. This is not much of a difference given the sampling variance associated with these estimates.

We believe the pattern of schooling interactions in tables 3 and 4 is probably attributable to the Arabization program. It should be noted, however, that use of selected samples restricted to economically active wage earners raises the possibility that the findings are attributable to selection bias. That is, the changes in schooling coefficients that we have attributed to Arabization may actually be the result of cross-cohort changes in labor force participation or wage earner status among middle school and high school graduates. The labor force participation rate (i.e., the fraction economically active) for men aged 15–45 years is about 75%. Although labor force participation rates increase with schooling overall, there is a clear negative association for younger men who have not necessarily finished their schooling. The resulting selection bias could be positive or negative depending on the selection rule determining labor force participation for school leavers.

On the plus side, there is little evidence for selection problems resulting from our restriction of the economically active sample to wage earners. When we put a dummy for wage-earner status on the left-hand side of the model used to produce the estimates in column 2 of table 4, the results support the null hypothesis of no significant Young*Schooling interaction terms in the sample of 2,694 economically active men aged 15–45 in the LSMS. The p-value for the joint test (4 df) is .31. The largest t-statistic for any single Young*Schooling interaction term is 1.75 (for 6–9 years of schooling).

IV. Additional Results from the LSMS

In this section, we use the LSMS literacy subsample to explore whether the change in language of instruction can be linked to direct measures of skills. The skill measures are a set of six test score indices, each taking

[13] Compare .269/.584 (LFS) to .604/1.06 (LSMS). It is not surprising that the schooling coefficients themselves are very different. Neither data set has simple Current Population Survey–type highest-grade-enrolled questions.

Table 6
Distribution of LSMS Test Scores

Test	Index Value		
	0	1	2
A. Test score subsample (N = 1,377):			
French writing	77.9	4.3	17.8
French reading	55.8	11.9	32.2
Arabic writing	50.3	8.9	40.8
Arabic reading	55.8	12.0	32.2
Math written	51.3	29.8	18.9
Math reasoning	34.9	27.8	37.3
B. Earnings subsample (N = 581):			
French writing	62.3	5.5	32.2
French reading	39.9	11.7	48.4
Arabic writing	36.1	7.6	56.3
Arabic reading	39.8	11.9	48.4
Math written	37.2	29.6	33.2
Math reasoning	26.5	23.2	50.3

NOTE.—The table shows the authors' tabulations from the Literacy Supplement to the Living Standards Measurement Survey (LSMS), which is described in the data appendix.

on the values 0, 1, and 2. A value of 0 indicates little or no ability, a value of 1 indicates some ability, and a value of 2 indicates "functional competence." Additional information on the tests and indices is provided in the data appendix.

Table 6 shows the distribution of test score indices in the extract of men aged 15–45 and in the subsample with data on wages. French writing appears to be the most difficult skill measured by the tests, with 78% of the sample indicating little or no ability. Math reasoning (the ability to think through simple arithmetic calculations out loud) is the most common skill; only 35% of men tested show no competence. Both French and Arabic reading skills are more common than writing skills. The lower part of table 6 shows the distribution of indices for the 581 men with data on earnings as well as test scores. Men with earnings are more likely than men in the larger sample to demonstrate competence in every subject.

A. Test Score Equations

Columns 3–7 of table 4 report coefficient estimates from a regression of test score indices on the same set of covariates as in equation (1).[14] Estimates from a regression of the French writing index on covariates

[14] The reported test score equations include an urban dummy and an age quadratic. The results are very similar when an experience quadratic replaces the age quadratic.

show little difference in writing scores between men with a primary school education and men with no formal schooling. This is not surprising because any French training at the primary school level probably focuses on reading and speaking. Regressions for the other indices show a substantial increase in test scores at the primary level. The biggest increase in French writing scores is between primary grades and grades 6–9.

The bottom rows of columns 3–7 in table 4 report the interaction terms between Young and schooling groups in the test score equations. Results for the French writing equation show a small (though statistically significant) Young*Primary interaction term, with much larger interaction terms for grades 6–9 and 10–12. These findings are consistent with the notion that the Arabization program reduced French language skills.

Results from each of the test score equations show negative Young*Primary interactions. Except in the French writing equation, interaction terms other than the primary interaction term are not significantly different from zero. This suggests that Arabization is associated largely with a substantial reduction in French writing skills. The negative Young*Primary interactions in equations other than the French writing equation are probably not attributable to Arabization. These negative estimates could be attributable to a decline in the quality of primary schools occurring at the same time as Arabization or to changes in student quality.

One likely explanation for negative Young*Primary interaction terms in test score equations other than that for French writing is that in the academic year 1984–85 the Ministry of Education introduced less restrictive requirements for promotion from primary to middle school (World Bank 1986). For example, students who repeated more than 4 years were barred from continuing on in public schools.[15] Between the academic years 1985–86 and 1987–88, the total number of grade repeaters decreased by 26% (World Bank 1989), and the share of repeaters in total enrollment declined from 0.32 in 1983–84 to 0.22 in 1984–85 and to 0.17 in 1987–88. This change means that the youngest members of the young primary-educated group will have had fewer years of school enrollment than older members. The reduction in the number of years of primary schooling could be expected to have little effect on French writing scores (because writing is not yet taught) but would be reflected in other skills for primary school graduates.

[15] Until 1984, only 24% of students completed the 5-year primary cycle in 5 years; 21% dropped out before the fifth year. Of those students who remained, many repeated at least three classes before completing primary school (World Bank 1986). An average 8.6 years of schooling was required for each graduate of the 5-year primary cycle. Similarly, in grade 12 the repetition rate was almost 30%. This appears to be related to repeated attempts to pass the baccalaureate screening exam for university admissions.

Table 7
Additional Test Score Equations

Regressor	Dependent Variable: French Writing (1)	French Reading (2)	Arabic Writing (3)	Arabic Reading (4)	Math Writing (5)	Math Reasoning (6)
Schooling:						
1–5	.116	.904	1.16	.904	.662	.667
	(.027)	(.052)	(.052)	(.052)	(.037)	(.057)
6–9	1.16	1.63	1.85	1.63	1.23	1.11
	(.075)	(.052)	(.038)	(.051)	(.054)	(.064)
10–12	1.59	1.85	1.93	1.85	1.60	1.28
	(.072)	(.036)	(.029)	(.036)	(.050)	(.057)
13–15	1.88	1.88	1.93	1.88	1.81	1.35
	(.027)	(.030)	(.029)	(.030)	(.029)	(.048)
16+	1.84	1.87	1.93	1.87	1.79	1.31
	(.027)	(.034)	(.032)	(.034)	(.031)	(.049)
Young	.199	.112	.014	.112	−.018	.136
	(.065)	(.074)	(.073)	(.074)	(.064)	(.113)
Young × Primary (1–5 years of school) × Age:						
15	−.074	−.628	−.758	−.628	−.338	−.370
	(.038)	(.102)	(.159)	(.101)	(.104)	(.142)
16	−.110	−.528	−.434	−.587	−.389	−.307
	(.034)	(.126)	(.149)	(.116)	(.099)	(.134)
17	−.161	−.504	−.239	−.504	−.273	−.139
	(.036)	(.143)	(.198)	(.142)	(.125)	(.148)
18	−.070	−.349	−.222	−.349	−.017	−.063
	(.060)	(.182)	(.206)	(.182)	(.113)	(.146)
19	−.191	−.302	.305	−.302	.063	−.174
	(.037)	(.177)	(.181)	(.177)	(.136)	(.165)
20	−.207	.081	.122	.081	−.109	.116
	(.040)	(.181)	(.180)	(.181)	(.103)	(.171)
21	−.202	−.100	−.014	−.101	.036	−.385
	(.080)	(.216)	(.205)	(.216)	(.151)	(.209)
Young × Schooling:						
6–9	−.777	−.046	.008	−.053	−.027	−.083
	(.116)	(.089)	(.052)	(.088)	(.088)	(.119)
10–12	−.916	−.330	−.364	−.323	−.280	−.060
	(.368)	(.288)	(.299)	(.288)	(.291)	(.198)
13+	−.024	−.007	−.038	−.007	−.036	.317
	(.022)	(.026)	(.031)	(.026)	(.028)	(.060)
R^2	.69	.66	.69	.66	.68	.43

NOTE.—LSMS = Living Standard Measurement and Literacy Survey. There are 1,377 observations in the LSMS sample with test scores. Robust standard errors in parentheses. The equations also include an urban dummy and age quadratic.

As a partial check on the "grade repetition" explanation for negative Young*Primary interaction terms in the test score equations, table 7 reports results from equations where the Young*Primary interaction term is allowed to vary by single years of age. The French writing equation shows small negative interactions at younger ages, with the largest Young

*Primary interaction terms at ages 19–21. In contrast, results from the other test score equations show no negative effects of the Young*Primary interaction terms at older ages (20–21) but large and significant negative effects for the youngest members of the young primary-educated cohort. The size of these negative effects generally increases with decreasing age. This finding is consistent with the notion that the youngest of the young primary-educated cohort spent fewer years in primary school.

B. Ordinary Least Squares and Two-Stage Least Squares Estimates of the Effect of Language Skills on Earnings

Columns 1–4 of table 8 report OLS estimates of equations for hourly wages where test score indices replace the schooling-group interaction terms in equation (1). Columns 5–6 combine the earnings and test score "reduced forms" in table 4 to produce two-stage least squares (2SLS) estimates of the effect of French test scores on earnings.

When the score indices are entered in pairs for reading and writing, as in columns 1–3, OLS estimates of coefficients on writing scores have statistically significant coefficients. For example, the French writing coefficient in column 1 is .172 (SE = .059). When all three writing score indices are combined, as in column 4, only the French and Math score indices have significant coefficients. Including three scores reduces the schooling coefficients somewhat.

Columns 5 and 6 report 2SLS estimates of equations where the French writing index is the single endogenous regressor. The excluded instruments used to compute the estimates reported in column 5 are interactions of a Young dummy with four schooling-group dummies for grades 1–5, 6–9, 10–12, and 13 or more years of schooling. These 2SLS estimates effectively attribute the interaction terms in the earnings reduced form solely to differences in French writing test scores by age and schooling group. In particular, the causal interpretation of 2SLS estimates requires that the instruments be correlated with earnings solely because of their correlation with the endogenous regressor. The 2SLS estimate of the French writing coefficient in column 5 is .527 (SE = .272).

It is unlikely that French writing skills are the only reason the switch to Arabic instruction had an effect. The 2SLS estimates should, therefore, be interpreted as capturing one important reason for the effect of Arabization on the returns to schooling. Moreover, as noted above, the Young*Primary interaction term is almost certainly correlated with test scores for reasons other than the Arabization program. Results are, therefore, presented in column 6 from an equation that is estimated excluding the Young*Primary interaction term from the instrument list. These results differ little from those in column 5. Overidentification test statistics for instrument-error orthogonality are reported in the last row of table 8. Both test statistics have marginal significance levels around .05.

Table 8
Test Scores and Wages in the LSMS

	Estimation Method							
Regressor	OLS (1)	OLS (2)	OLS (3)	OLS (4)	2SLS (5)	2SLS (6)	2SLS (7)	2SLS (8)
Urban	.091	.108	.066	.057	.029	.035	.410	.230
	(.069)	(.069)	(.070)	(.070)	(.086)	(.086)	(.442)	(.122)
Experience	.106	.116	.113	.095	.087	.090	.100	.100
	(.017)	(.017)	(.017)	(.014)	(.022)	(.022)	(.030)	(.020)
Experience^2	−.002	−.002	−.002	−.001	−.001	−.001	−.002	−.001
	(.0004)	(.0004)	(.0004)	(.0004)	(.0005)	(.0005)	(.0006)	(.0005)
Schooling:								
1–5	.327	.217	.207	.204	.281	.286	1.26	1.55
	(.095)	(.106)	(.094)	(.107)	(.094)	(.094)	(1.26)	(.743)
6–9	.778	.791	.709	.580	.372	.411	2.20	2.49
	(.141)	(.145)	(.129)	(.153)	(.340)	(.344)	(1.93)	(1.04)
10–12	1.16	1.27	1.12	.956	.587	.642	2.83	2.80
	(.170)	(.162)	(.157)	(.182)	(.474)	(.480)	(2.26)	(.997)
13–15	2.01	2.17	1.98	1.76	1.32	1.38	3.85	3.59
	(.217)	(.202)	(.204)	(.226)	(.576)	(.583)	(2.49)	(.981)
16+	2.67	2.83	2.64	2.41	1.98	2.04	4.47	4.23
	(.211)	(.197)	(.198)	(.221)	(.568)	(.575)	(2.45)	(.970)
Young	−.123	−.114	−.092	−.157	−.127	−.127	−.341	−.290
	(.115)	(.115)	(.114)	(.108)	(.118)	(.117)	(.336)	(.180)
French writing	.172	...	...	.133	.527*	.494*	.545	.365
	(.059)			(.059)	(.272)	(.275)	(.511)	(.137)
French reading	.011	...	...	...	...	...	...	...
	(.057)							
Arabic writing	...	.164	...	.058	...	...	...	−1.08*
		(.076)		(.061)				(.654)
Arabic reading	...	−.049	...	...	...	...	...	...
		(.075)						
Math written	...	...	.203	.151	...	...	−1.47*	...
			(.066)	(.069)			(2.01)	
Math reasoning	...	...	.033	...	...	...	...	...
			(.050)					
χ^2 (df)	...	...	...	...	7.3 (3)	6.9 (2)	4.0 (3)	2.8 (3)
R^2	.61	.61	.61	.61	.59	.59	.43	.48

NOTE.—LSMS = Living Standard Measurement and Literacy Survey. Standard errors are in parentheses. There are 581 observations in the LSMS subsample with test scores and wages. Excluded instruments in col. 2 are interactions of Young and schooling-group dummies. Excluded instruments in col. 5, 7, and 8: Young × primary schooling interaction, Young × 6–9 years of schooling interaction, Young × 10–12 years of schooling interaction, Young × 13–15 years of schooling interaction. Excluded instruments in col. 6 are the same except that Young × primary schooling interaction is dropped.

* Endogenous regressors.

In an alternative approach to possible problems of omitted variables bias in 2SLS estimates of the earnings equation, we explored a strategy imposing restrictions on unobserved variance components. This approach provides a partial solution to problems of omitted variables bias that might arise if, for example, Arabization is associated with a change in the quality of students at different schooling levels.

Assume that two test scores, for example, French writing and math,

belong in the earnings equation. Assume also that each test score can be decomposed as

$$w_{1i} = X_i\delta_{11} + d_i\delta_{12} + \sum_j s_{ji}\delta_{13j} + [\theta_i\lambda_1 + \eta_{1i}]$$

and (2)

$$w_{2i} = X_i\delta_{21} + d_i\delta_{22} + \sum_j s_{ji}\delta_{23j} + [\theta_i\lambda_2 + \eta_{2i}],$$

where θ_i is a mean-zero variance component representing unobserved individual heterogeneity or ability, λ_1 and λ_2 are score-specific factor loadings, and η_{1i} and η_{2i} are variance components orthogonal to θ_i.

Additional orthogonality assumptions are made regarding the variance components in score 1 (French writing) and score 2. In particular, for each instrument, z_{ki} ($k = 1, 2, 3, 4$), we assume that

$$E[z_{ki}\theta_i] \neq 0; \tag{3a}$$

$$E[z_{ki}\eta_{1i}] \neq 0; \tag{3b}$$

and

$$E[z_{ki}\eta_{2i}] = 0. \tag{3c}$$

These assumptions capture the notion that the instruments are correlated with the second score, w_{i2}, solely because of individual heterogeneity. The instruments are correlated with the French writing score, w_{1i}, because of heterogeneity and because of the policy change, the effect of which is subsumed in the variance component, η_{1i}.

The equation of interest includes covariates, test scores, and the heterogeneity component, θ_i:

$$y_i = X_i\beta + d_i\alpha_1 + \sum_j s_{ji}\alpha_{2j} + w_{1i}\phi_1 + w_{2i}\phi_2 + [\theta_i + \varepsilon_i], \tag{4}$$

where θ_i is correlated with the test scores and ε_i is a component that is uncorrelated with all the regressors in (4). As before, the excluded instruments include the interaction terms z_{1i}, z_{2i}, z_{3i}, z_{4i}. These instruments are assumed to be uncorrelated with ε_i but cannot be used to identify the parameters in equation (4) because of the component θ_i in the error term.

Given restrictions (2) and (3), the parameters in equation (4) are identified. To see this, solve the second score equation for θ_i and use this to substitute for θ_i in (4). The resulting equation is

$$y_i = X_i[\beta - \delta_{21}/\lambda_2] + d_i[\alpha_1 - \delta_{22}/\lambda_2] + \sum_j s_{ji}[\alpha_{2j} - \delta_{23j}/\lambda_2] + w_{1i}\phi_1 + w_{2i}[\phi_2 + 1/\lambda_2] + [\varepsilon_i - \eta_{2i}/\lambda_2]. \quad (5)$$

This equation cannot be estimated by ordinary least squares because w_{2i} is necessarily correlated with η_{2i}. But the set of z_{ki} can be used as instrumental variables because these are uncorrelated with both ε_i and η_{2i}. The only regressor that must be treated as endogenous is the second test score, w_{i2}.

The 2SLS estimates of equation (5) are reported in columns 7 and 8 of table 8. In column 7, the second score (w_{i2}) is taken to be written math and in column 8 the second score is taken to be Arabic writing. The estimates in both columns have standard errors on most of the coefficients that are higher than for the other specifications. This is probably a consequence of the fact that the instruments are not highly correlated with either math or Arabic writing.

The coefficient on w_{1i} is an estimate of ϕ_1. The estimates in columns 7 and 8 are .545 and .365. These estimates are not very precise and are not significantly different from the 2SLS estimates in columns 5–6. The 2SLS coefficient on w_{2i} is an estimate of $[\phi_2 + 1/\lambda_2]$. The estimate in column 7 is -1.47 with a standard error of 2.01. The estimate in column 8 is -1.08 with an SE of .654. Based on these estimates, we cannot reject the null hypothesis that unobserved heterogeneity of the sort modeled in equations (2) and (4) is not actually a problem for the 2SLS estimates. In other words, there is little evidence for a spurious relationship between Arabization (as captured by the Young*Schooling instruments) and earnings.

V. Summary and Conclusions

Results from two data sets suggest that the switch from French to Arabic instruction in Moroccan secondary schools had a negative effect on the French language skills and earnings of young Moroccan men. The estimates suggest that Arabization reduced the economic premium to postprimary education in young cohorts by as much as one-half. The negative effect of Arabization on earnings and the returns to schooling appears to be largely explained by a decline in French writing skills. The size of this decline is between one-half and one index unit, equivalent to going from functional competence to only some ability or from some ability to no ability. Our earnings results are roughly consistent with those of Dustmann (1994), who reports that the ability to write in German appears to be the most important language skill for the earnings of migrant workers in Germany.

The OLS estimates of the relationship between language skills and earnings among Moroccans suggest that going from minimal skills to

functional competence (one index unit) in written French raises earnings by about 17%. This is smaller than Chiswick's (1991) estimates, which imply that reading fluency increases the hourly wages of immigrants in the United States by 30%. In contrast, our 2SLS estimates of the effect of test scores on earnings are much larger, implying that going from some ability to functional competence raises earnings by more than 50%.[16] Of course, the loss of French writing skills is probably not the sole reason for the reduction in returns to schooling experienced by students subject to the Arabization program. The findings suggest, however, that policies designed to reduce the private costs associated with the switch in language of instruction would be most effective if they focused on additional training in written French.

The findings reported here also have implications for economic research on the returns to well-defined language skills. Using the Arabization program as a kind of natural experiment, the results presented here provide strong evidence that foreign language skills raise earnings in a developing country and that the loss of these skills would reduce earnings. There is little evidence that OLS estimates of the economic value of French language skills in Morocco are inflated by unobserved ability or the failure to account for some other form of heterogeneity bias. In fact, as in much of the recent literature on the returns to schooling (see, e.g., Angrist and Krueger 1991), OLS estimates of the returns to language skills are lower than the instrumental variables estimates.

Data Appendix

The 1990 and 1991 Labor Force Surveys

The first sample used in this study is drawn from the 1990 and 1991 Labor Force Surveys. These surveys are administered annually by the Moroccan Direction de la Statistique in the urban areas of Morocco to roughly 30,000 randomly selected households. (The LFS is documented in Kingdom of Morocco [1993*a*].) The LFS collects demographic and economic information on labor force participation, employment, schooling, labor market experience, hours of work, earnings, and unemployment.

The 1990 and 1991 LFS each included about 100,000 individuals aged 15 and over of which nearly 60,000 respondents were economically active and 45% of the actively employed were employed as salaried workers. The labor force participation rate in the whole sample is roughly 75% for men and 25% for women. The labor force participation rate for men

[16] Because the reduced forms in table 4 suggest that Arabization reduced French writing test scores by between 3/4 and 1 unit, a one-unit change seems most relevant for assessing the policy effect.

aged 25–45 is 95%. For women in the same age group, the participation rate is 33%.

Our data set includes two random samples of 8,700 observations on salaried workers (roughly one-third of the total) in each of the 1990 and 1991 surveys. Because the LFS is an annual rotating panel (half the sample is replaced every year), our working subsample is selected from all 1991 observations plus those 1990 observations not interviewed in 1991 (roughly 13,000 observations in total). In our empirical analysis, we use only men aged 15–45 with nonmissing schooling and experience variables, who did not attend Koranic schools. Eliminating 334 wage outliers leaves a sample of 7,040 observations.

The raw earnings variable in the LFS is monthly wages in the survey month. We computed hourly wages by dividing monthly wages by weekly hours worked in the reference week times 4.5. Years of schooling are not reported in the LFS. Rather, the LFS survey records responses to questions on highest level of school enrolled (e.g., primary school) and diploma or certificate of completion. We coded schooling-group dummy variables according to the highest level of enrollment. In practice, there were many inconsistencies between the two schooling variables. For example, some respondents reporting only primary school enrollment claimed to have middle school certification. There were also inconsistencies across years for respondents who could be linked across rotation groups. Inconsistencies in schooling variables are worse for older respondents, probably because respondents' recall of childhood schooling experiences deteriorates with time.

The Living Standard Measurement and Literacy Surveys

The second sample used in this article is taken from the 1990–91 Moroccan Living Standards Measurement and Literacy Survey (documented in Kingdom of Morocco [1993*b*]). The LSMS is a comprehensive household survey that was conducted by the Moroccan Direction de la Statistique (with technical assistance from the World Bank) on a random sample of households throughout Morocco. The LSMS fieldwork started in October 15, 1990, and ended in October 31, 1991. The survey consists of a broad household questionnaire that includes information about schooling, employment, labor and nonlabor income, consumption and expenditures, health, fertility, and savings. The sample is stratified over the seven economic regions of Morocco and contains 20 primary clusters of 24 households within each region. Each primary cluster is further divided into three secondary clusters of eight households each. The sample includes 3,360 households containing almost 20,000 individuals.

Out of a total sample size of 19,618, 6,618 individuals were found to be "economically active" (by International Labor Organization definitions). The number of economically active men in the 15–45 age-group with nonmissing schooling variables who did not attend Koranic schools is 2,694. The labor force participation rate for men aged 15–24 is 68%, and for men in the 25–45 age-group it is 97%. Roughly 43% of the economically active men in the sample are wage earners. The fraction of females

who are wage earners is 20%. Restricting the 15–45-year-old male sample to wage earners and eliminating a few wage outliers leaves a sample of 1,166.

The raw LSMS earnings variable is weekly wages paid for the main job in the reference week. We computed hourly wages by dividing weekly wages by weekly hours worked in the reference week. The LSMS reports information on the highest level of enrollment (primary, secondary, etc.), whether this level of schooling was completed, and the number of years enrolled at each level.

A battery of tests was given to the LSMS literacy subsample. A random sample of two-thirds of all LSMS households were targeted for testing, encompassing 9,600 individuals aged 9–69. For a variety of reasons, however (described in Lavy, Spratt, and Leboucher 1995), only 8,000 individuals were actually tested. This is roughly half of the sample in the age-groups eligible for testing. Out of 2,694 total observations on economically active men aged 15–45, our extract includes 1,377 observations with valid test scores. Out of 1,166 observations on men with earnings data, our extract includes 581 observations with valid test scores.

Tests were given only to individuals with at least some formal education. Subject areas included basic reading and writing skills in Arabic and French (when some French ability was reported) and mental and written math.[17] French reading assessment included two levels: word decoding and word-picture matching at level 1, and text comprehension at level 2. French writing assessment also included two levels: word dictation at level 1, and sentence dictation at level 2. Other information was collected on educational experiences, self-judgments on language-specific reading and writing abilities, math abilities, and reading and writing practice at home and at the work place.

Raw test scores from each level were coded by World Bank researchers into a set of six indices for French reading and writing, Arabic reading and writing, and mental and written mathematics (again, see Lavy et al. [1995] for details.) The indices range from 0 to 2 for Arabic and French writing and from 0 to 3 for Arabic and French reading and for written mathematics because the raw scores for these tests exhibited greater dispersion. We recoded the 0–3 indices to 0–2 indices by subtracting one from each positive value. An index value of 0 indicates little or no ability, an index of 1 indicates some ability, and an index value of 2 indicates functional competence. The test score indices were recoded in this way so as to make the six indices as comparable as possible. We note, however, that the results are very similar when the original indices are used.

References

Aarim, Naji. "Approaches to TEFL in Moroccan High Schools." *Curriculum and Teaching* 7, no. 1 (1992): 51–61.

[17] High school graduates with a baccalaureate or its equivalent were exempted from the French tests and later assigned the highest score possible.

Angrist, J. D. "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records." *American Economic Review* 80 (June 1990): 313–36.

Angrist, J. D., and Krueger, A. "Does Compulsory School Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics* 106 (November 1991): 979–1014.

Chiswick, Barry R. "Speaking, Reading, and Earnings among Low-Skilled Immigrants." *Journal of Labor Economics* 9 (April 1991): 149–70.

———, ed. *Immigration, Language and Ethnicity, Canada and the United States.* Washington, DC: American Enterprise Institute, 1993.

Chiswick, Barry R., and Miller, P. "The Endogeneity between Language and Earnings: An International Analysis." Paper presented at the Center of Economic Policy Research conference on the "Economics of International Migration: Econometric Evidence" at Konstanz, February 1993.

Dustmann, Christian. "Speaking Fluency, Writing Fluency, and the Earnings of Migrants." Discussion Paper no. 905. London: Center for Economic Policy Research February 1994.

Elbiad, Mohamed. "The Role of Some Population Sectors in the Progress of Arabization in Morocco." *International Journal of the Sociology of Language* 87 (1991): 27–44.

Grenier, Giles. "The Effects of Language Characteristics on the Wages of Hispanic American Males." *Journal of Human Resources* 19 (Winter 1984): 35–52.

Heggoy, Andrew. "Arabization and the Kabyle Language and Cultural Issues in Algeria." *Africana Journal* 15 (1990): 292–304.

Khandker, Shahidur; Lavy, Victor; and Filmer, Deon. "Schooling of Children in Morocco: Can the Government Improve Outcomes?" Washington, DC: World Bank, Policy and Research Department, 1993.

Kingdom of Morocco, Ministry of Education, Direction de l'Enseignment Secondaire. *Langue d'enseignement.* Rabat: Ministry of Education, n.d.

Kingdom of Morocco. *Population active urbaine: 1990.* Rabat: Ministere Des Affaires Economiques et Sociales, Direction de la Statistique, 1990.

Kingdom of Morocco. *Population active urbaine, 1990: Resultats detailles.* Rabat: Ministere des Affaires Economiques et Sociales, Direction de la Statistique, 1993. (*a*)

Kingdom of Morocco. *Niveaux de vie des menages, 1990–91: Priemiers resultats.* Vol. 1. Rabat: Ministere des Affaires Economiques et Sociales, Direction de la Statistique, 1993. (*b*)

Kossoudji, Sherrie A. "English Language Ability and the Labor Market Opportunities of Hispanic and East-Asian Immigrant Men." *Journal of Labor Economics* 6 (April 1988): 205–28.

Lavy, Victor; Spratt, Jennifer; and Leboucher, Nathalie. "Changing Patterns of Illiteracy in Morocco: Assessment Methods Compared." Living Standard Measurement and Literacy Survey Working Paper no. 120. Washington, DC: World Bank, Policy and Research Department, 1995.

McManus, Walter; Gould, William; and Welch, Finis. "Earnings of Hispanic Men: The Role of English Language Proficiency." *Journal of Labor Economics* 1 (April 1983): 101–3.

Moatassime, Ahmed. "Cultural Pluralism and Education in the Maghreb." *Prospects* 22 (1992): 171–84.

Spencer, William. *Historical Dictionary of Morocco.* African Historical Dictionaries no. 24. Metuchen, NJ: Scarecrow, 1980.

Tainer, Evelina. "English Language Proficiency and Earnings among Foreign-born Men." *Journal of Human Resources* 23 (Winter 1988): 108–22.

Weinstein, B., ed. *Language Policy and Political Development.* Norwood, NJ: Ablex, 1990.

World Bank. *The Moroccan Education System: A Sector Report.* Washington, DC: World Bank, Policy and Research Department, 1985.

———. *Kingdom of Morocco: Education Sector Reform Program.* Report no. P-4240-MOR. Washington, DC: World Bank, 1986.

———. *Staff Appraisal Report, Kingdom of Morocco: Rural Primary Education Project.* Report no. 7434-MOR. Washington, DC: World Bank, 1989.

———. *Staff Appraisal Report, Kingdom of Morocco: Rural Basic Education Development Project.* Report no. 8936-MOR. Washington, DC: World Bank, 1991.