CESifo Working Papers

www.cesifo.org/wp

Does Culture Cause Economic Development? A Reassessment of the Evidence from European Regions

Jeremy Edwards

CESIFO WORKING PAPER NO. 4015
CATEGORY 6: FISCAL POLICY, MACROECONOMICS AND GROWTH
NOVEMBER 2012

CESifo *Working Paper No. 4015*

Does Culture Cause Economic Development? A Reassessment of the Evidence from European Regions

Abstract

This paper shows that Tabellini's recent claim to have provided evidence that culture has a causal effect on economic development is unjustified. Tabellini's claim is based on an instrumental variables analysis in which two instruments are used to identify the supposed causal effect. One of these – past literacy – is an invalid instrument. The other – past political institutions – is a weak instrument. The estimates obtained using this second instrument are so imprecise that they cannot be used to support any conclusions about the effect of culture on economic development.

JEL-Code: C260, O400, O430, P100.

Keywords: culture, economic development, instrumental variables, weak instruments.

Jeremy Edwards
Faculty of Economics
University of Cambridge
Sidgwick Avenue
UK – Cambridge, CB3 9DD
je12@econ.cam.ac.uk

November 2012
I thank Tim Guinnane, Erin Hengel, Ajit Karnik, Sheilagh Ogilvie and Paul Seabright for very helpful comments and Guido Tabellini for making his data publicly available (at http://didattica.unibocconi.it/mypage/index.php?IdUte=48805&idr=13301&lingua=ita).

1. Introduction

Does culture have a causal influence on economic development? A recent paper by Tabellini (2010) argues that data for European regions support the hypothesis that cultural traits have such a causal effect, either directly, or indirectly through better-functioning institutions. Tabellini considers the possible beneficial effects of two broad cultural traits on economic performance. First, trust and respect for others may promote economic development by encouraging anonymous exchange and participation in public good provision as well as by improving the functioning of government institutions. Second, confidence in the virtues of individualism may improve economic performance because individuals then perceive a clear link between individual effort and economic success. However, although culture may influence development, economic development is also likely to influence culture.[1] In order to estimate a causal effect of these cultural traits on economic development it is therefore necessary to identify some exogenous source of variation in culture. To do this, Tabellini uses the literacy rate in European regions at the end of the nineteenth century and the past political institutions of these regions as instruments for culture. Controlling for various other influences, Tabellini concludes that the instrumental variable (IV henceforth) estimates of the effects of culture show that the component of culture explained by past literacy and past political institutions is an important determinant of regional economic performance in modern Europe.[2]

Tabellini is careful to note various qualifications concerning the identifying assumptions and tests of instrument validity used in his analysis. Nevertheless, as this paper shows, his instrumental variable analysis suffers from serious problems and so fails to identify a causal effect of culture on regional economic performance. Past literacy is not a valid instrument for culture. There are no obvious reasons to think that past political institutions is an invalid instrument, but it appears to be a weak instrument. The IV estimates of the effects of culture on economic performance obtained using past political

[1] Tabellini (2010), 678, 690-1.
[2] Tabellini (2010), 704, 710.

institutions as the instrument are so imprecise as to be completely uninformative. They are compatible with culture having either no effect on economic performance or an incredibly large effect. This lack of precision in the IV estimates means that there is no statistically significant difference between them and the OLS estimates of the effect of culture on economic performance. If the IV estimates were precise, the absence of such a difference would imply that the OLS estimates could be given a causal interpretation. But the IV estimates of the effect of culture are so imprecise that it is not possible to do so. The lack of precision of the IV estimates not only means that any possible causal effect of culture is very poorly determined, but also prevents the OLS estimates being given a causal interpretation despite the absence of a significant difference between them and the IV ones.

2. The identification strategy for establishing a causal effect of culture

Tabellini's empirical strategy for establishing a causal effect of culture on regional economic performance involves regressing measures of either the level or the growth rate of regional per capita gross value added on contemporaneous education, past urbanisation, country dummy variables, the initial level of economic development (when the dependent variable is the growth rate) and alternative measures of culture, with past literacy and past political institutions being used as instruments for the culture measures. Full details of the data and, in particular, the justification for the measures of cultural traits that are used are given in Tabellini (2010). In order to understand Tabellini's identification strategy I give in this section a brief account of how each variable is measured and the reasons for its inclusion in the regression equation, as well as outlining the justification of the two instruments used for the culture variables.

The level of regional economic development (*yp9500*) is measured as the average over the period 1995-2000 of per capita gross value added in international prices (adjusted for purchasing power) expressed as a percentage of the EU15 average. The rate of growth of regional economic development (*growth*) is measured as the log difference of per capita regional gross value added over the period 1977-2000. When *growth* is the dependent variable, the log of the level of regional per capita gross value added in 1977 (*lyp77*) is included as an exogenous regressor to allow for possible convergence effects in regional growth rates.

A measure of contemporaneous education is included as an exogenous regressor because human capital is expected to influence development and growth, and also because education is likely to influence cultural traits. Both the level of economic development and cultural traits were measured in the late 1990s, and much of the adult population in this period went to school in the 1960s and 1970s. The measure of regional education is thus regional school enrolment rates in 1960 (*school*).

In order to be valid instruments, past literacy and past political institutions must be uncorrelated with unobserved determinants of current economic performance. This requirement is more likely to be met if a measure of past regional economic development is included as an exogenous regressor. In the absence of data on past regional per capita gross value added, the level of past regional development is measured by the fraction of the regional population that lived in cities with more than 30,000 inhabitants around 1850 (*urb_rate1850*).

The regions in the dataset come from eight European countries, and country dummy variables are included as exogenous regressors to allow for country-specific effects on regional economic performance. Among other things, these country dummy variables reflect the effect of common national institutions on regional performance.

Several measures of cultural traits are available from two waves of the World Value Surveys, carried out in 1990-91 and 1995-97. *Trust* is the percentage of respondents in each region who answer ‘most people can be trusted’ to the question ‘Generally speaking, would you say that most people can be trusted or that you can’t be too careful in dealing with people?’ *Respect* is the percentage of respondents in each region who mention ‘tolerance and respect for other people’ when asked ‘Here is a list of qualities that children can be encouraged to learn at home. Which, if any, do you consider to be especially important? Please choose up to five.' *Obedience* is the percentage of respondents in each region who mention 'obedience' in response to this same question. *Control* is the average response in each region (multiplied by 10) to the question 'Some people feel they have completely free choice and control over their lives, while other people feel that what they do has no real effect on what happens to them. Please use this scale (from 1 to 10) where 1 means "none at all" and 10 means "a great deal" to indicate how much freedom of choice and control you feel you have over the way your life turns out.' Tabellini argues that *trust* and *respect* promote economic development because they encourage anonymous exchange and public good provision as well as improving the functioning of government institutions. A larger value of *control* is interpreted as registering a more widespread conviction that individual effort will pay off, with correspondingly beneficial effects on regional economic performance. A larger value of *obedience*, by contrast, is interpreted as an indicator of greater distrust of the benefits of individualism, with correspondingly harmful effects on regional economic performance.

As well as using these four measures of specific cultural traits, Tabellini also combines them into summary measures based on the first principal component of some or all of the specific measures. The first principal component of all four measures (*pc_culture*) is positively correlated with *trust*, *respect* and *control*, and negatively correlated with *obedience*, and is taken to be a net measure of the aspects of regional culture that favour development. The first principal component of the three measures that are expected to have positive effects on economic performance (*trust*, *respect* and *control*) is *pc_culture_pos*, while the first principal component from the two responses on the desirable qualities of children (*obedience* and *respect*) is *pc_children*. All these principal components are multiplied by 100 so that they can be interpreted as percentages.

Tabellini describes these seven measures as unconditional measures of culture. To remove some of the endogenous components of culture, he also calculates conditional measures of regional culture. Using the dataset of individual responses from the World Values Survey, each of the seven unconditional culture measures is regressed on a set of regional dummy variables, marital status, gender, age group, self reported social class, and two categorical variables for health condition and years of education. The estimated coefficients on the regional dummy variables are then used to construct measures of conditional culture for each of the regions. These measures control for characteristics of individual respondents in the World Values Survey and thus can be seen as more accurate indicators of culture that is specific to regions. When conditional culture measures are used in the regression analysis, regional observations are weighted by the inverse of the standard errors of these estimated measures of regional conditional culture to allow for different measurement errors across regions.

The fact that there may be reverse causation between regional culture and regional economic performance means that IV rather than OLS estimation of the regression equation is required if the coefficient of the culture variable is to be given a causal interpretation. The two instruments used for the culture measures are *literacy*, the literacy rate in each region around 1880, and *pc_institutions*, the first principal component of a measure of constraints on the executive in each region at five dates: 1600, 1700, 1750, 1800 and 1850. These variables are expected to be correlated with current culture because current culture is influenced by contemporaneous social interactions and cultural traditions inherited from earlier generations. Contemporaneous social interactions are represented in the estimated regression equation by the *school* variable, which measures the education of the currently adult population, and the country dummies, which reflect, among other things, current national institutions. There are no measures of inherited cultural traditions, but the culture of earlier generations is influenced by past social interactions, and hence by the historical analogues of the contemporaneous social interaction variables. These historical analogues – *literacy*, a measure of past education, and *pc_institutions*, a measure of past political institutions – are thus possible instruments for current culture.

In order to be valid instruments, it must be the case that *literacy* and *pc_institutions* can be excluded from the structural regression equation that relates regional economic performance to culture and other variables: in other words, *literacy* and *pc_institutions* must be uncorrelated with the error term in this equation. The justification for these exclusion restrictions is that the regression equation controls for contemporaneous education (*school*), contemporaneous national institutions (country dummy variables) and past economic development (*urb_rate1850*). The identifying assumption is thus that past education and past political institutions do not have a direct effect on current economic performance once current education, current institutions and past economic development are included in the structural equation. Tabellini notes that this is a rather strong assumption, but points out that the exclusion of these historical variables is more plausible when the measure of regional economic performance is *growth*, in which case the regression equation also controls for initial regional per capita gross value added (*lyp77*). Before considering whether *literacy* and *pc_institutions* are valid instruments in more detail in section 4, I first present the main results which underlie Tabellini's claim that culture has a causal effect on economic performance.

3. Tabellini's main results

The main results on which Tabellini bases his conclusion that culture is a causal influence on regional economic performance are shown in Tables 1 and 2. I present results only for the three summary measures of culture (*pc_culture*, *pc_culture_pos*, and *pc_children*) because the *F* statistics reported in Tables 5 and 6 of Tabellini (2010) suggest that there is a serious weak instrument problem when the four specific measures of cultural traits (*trust*, *respect*, *control* and *obedience*) are used. Furthermore, I present results only for the conditional measures of culture since they do not differ in any important respect from those obtained when unconditional measures of culture are used.

Table 1 shows, for the three alternative summary measures of culture, the OLS and IV estimates of the coefficient of culture in the regression of *yp9500* on *school*, *urb_rate1850*, country dummy variables and culture. The instruments used are *literacy* and *pc_institutions*. Table 2 does the same in the case where *growth* is the dependent variable and *lyp77* is added to the regressors. Both tables report the *p* value of Hansen's *J* statistic for testing the overidentifying restriction, which is the appropriate statistic in the case of a heteroskedastic or clustered covariance matrix. This is a test of the joint null hypotheses that the regression model is correctly specified and that, conditional on one instrument being valid, the other is uncorrelated with the regression error term and thus correctly excluded from the equation. Although the dataset contains 69 regions, observations for *literacy* are available for only 67 of them, so all results in Tables 1 and 2 are obtained from regressions with 67 observations. I follow Tabellini in presenting results for both heteroskedasticity-robust standard errors (which allow for heteroskedastic errors that are independent across observations) and standard errors that are clustered by countries (these are robust and also allow for arbitrary correlation of errors within countries while continuing to assume independence of errors across countries). It is natural to expect that the errors in the relationship between regional economic performance and the regressors will be correlated within countries, which creates a *prima facie* case for paying more attention to the results based on clustered standard errors.

The IV estimates of the culture variables in Tables 1 and 2 correspond to those reported by Tabellini in Tables 5 and 6 of his paper. All the IV estimates in Table 1 and many of those in Table 2 are significantly different from zero on the basis of inference using the standard asymptotic approximation. This leads Tabellini to conclude that there is clear evidence from European regions that culture has a causal effect on economic performance. Before this conclusion can be accepted, however, it is necessary to consider whether *literacy* and *pc_institutions* are good instruments for the culture variables.

An unavoidable problem with the IV estimator of the coefficient of an endogenous regressor (one that is correlated with the regression error term) is that it is biased towards the inconsistent OLS estimator in finite samples even though it is asymptotically consistent. Furthermore, the finite-sample distribution of the IV estimator can differ substantially from the asymptotic distribution on which IV standard errors and confidence intervals are usually based. Both the bias towards the OLS estimator and the difference between the finite-sample and asymptotic distributions can be very large if the instruments are weak.

Table 1: OLS and IV estimates of the relationship between culture and the level of economic development

Culture measure	Standard errors	OLS estimate	95% OLS confidence interval	IV estimate	95% standard IV confidence interval	95% AR confidence interval	*J* test *p* value	First-stage *F* statistic
pc_culture	Robust	0.59	0.33 – 0.86	1.11	0.57 – 1.66	0.59 – 1.92	0.2129	10.83
	Clustered	0.59	0.14 – 1.05	1.11	0.19 – 2.03	0.53 – 3.92	0.4004	12.89
pc_culture_pos	Robust	0.77	0.42 – 1.11	1.16	0.54 – 1.78	0.68 – 2.02	0.1015	17.47
	Clustered	0.77	0.46 – 1.07	1.16	0.28 – 2.03	0.64 – 3.45	0.2957	16.63
pc_children	Robust	0.58	0.20 – 0.95	1.40	0.63 – 2.16	0.82 – 2.18	0.0838	6.75
	Clustered	0.58	-0.10 – 1.25	1.40	0.25 – 2.54	0.78 – 4.11	0.1831	22.65

Notes.
This table compares OLS and IV estimates of the effects of alternative measures of conditional culture on *yp9500* in regressions estimated on 67 observations. The estimation procedure weights observations by the inverse of the standard errors of the estimated measures of regional conditional culture to allow for different measurement errors across regions. The regressions also include *school*, *urb_rate1850* and country dummy variables as exogenous regressors, the estimated coefficients of which are not reported. In all cases the instruments used for the IV estimates are *literacy* and *pc_institutions*.

There is no single unambiguous indicator of whether the instruments used for IV estimation are weak. A common rule of thumb, derived from Staiger and Stock (1997), is that weak instruments are not a matter for concern if, in the first-stage regression of the endogenous variable on the instruments and the other regressors in the structural regression equation, the *F* statistic for the test of the null hypothesis that the coefficients of the instruments are zero is greater than 10. Tables 1 and 2 show that, in most cases, the value of this *F* statistic for *literacy* and *pc_institutions* in the first-stage regressions for the culture variables is above 10. However, it is incorrect to regard a *F* statistic larger than 10 as meaning that weak instruments are not a problem: Hahn and Hausman (2003) point out that the magnitude of the correlation between the errors in the first-stage and structural regression equations is also an important influence on the size of the bias of the IV estimator. Unfortunately the estimated residuals of the first-stage and structural equations cannot be used to provide an unbiased estimate of this correlation, because any bias in the IV estimates will translate into bias in the estimated residuals of the structural equation. The natural way of assessing whether weak instruments are a problem in the present context is to consider confidence intervals for the IV estimates of the effect of culture on economic performance that have correct coverage without requiring any information about the relationship between the endogenous variable and the instruments. Such confidence intervals are valid whether the instruments are strongly or weakly correlated with the culture variables. Dufour (2003) argues that the Anderson-Rubin (AR henceforth) procedure for computing such confidence intervals has several desirable features.[3] Tables 1 and 2 therefore report AR confidence intervals for the IV estimates of the effect of the culture variables using the Chernozhukov-Hansen (2008) extension of the AR procedure, which makes it robust to heteroskedastic or clustered errors. One drawback of the AR procedure is that it may use information inefficiently when there are more instruments than endogenous regressors. However, the simulation results in Chernozhukov and Hansen (2005) suggest that the AR procedure has correct size, and that the loss of power is small when there are only a few more instruments than endogenous regressors, as is the case here.

[3] Anderson and Rubin (1949).

Table 2: OLS and IV estimates of the relationship between culture and the growth rate

Culture measure	Standard errors	OLS estimate	95% OLS confidence interval	IV estimate	95% standard IV confidence interval	95% AR confidence interval	*J* test *p* value	First-stage *F* statistic
pc_culture	Robust	0.0083	0.0010 – 0.0156	0.0169	0.0043 – 0.0295	0.0032 – 0.0460	0.4374	11.09
	Clustered	0.0083	-0.0011 – 0.0178	0.0169	-0.0037 – 0.0374	-0.0050 – 0.0560	0.4437	16.53
pc_culture_pos	Robust	0.0121	0.0014 – 0.0229	0.0174	0.0035 – 0.0312	0.0035 – 0.0467	0.2626	16.23
	Clustered	0.0121	-0.0041 – 0.0284	0.0174	-0.0050 – 0.0397	-0.0050 – 0.0673	0.2953	16.24
pc_children	Robust	0.0068	-0.0008 – 0.0144	0.0213	0.0044 – 0.0383	0.0054 – 0.0611	0.1753	8.29
	Clustered	0.0068	-0.0027 – 0.0162	0.0213	-0.0017 – 0.0443	-0.0081 – 0.0752	0.1400	15.51

Notes.
This table compares OLS and IV estimates of the effects of alternative measures of conditional culture on *growth* in regressions estimated on 67 observations. The estimation procedure weights observations by the inverse of the standard errors of the estimated measures of regional conditional culture to allow for different measurement errors across regions. The regressions also include *school*, *urb_rate1850*, *lyp77* and country dummy variables as exogenous regressors, the estimated coefficients of which are not reported. In all cases the instruments used for the IV estimates are *literacy* and *pc_institutions*.

Comparison of the standard confidence intervals in Tables 1 and 2 based on the asymptotic distribution of the IV estimator with the corresponding AR confidence intervals suggests that Tabellini's IV estimates of the effects of culture are subject to weak-instrument concerns. In Table 1, the heteroskedasticity-robust AR confidence intervals are only a little wider than the standard heteroskedasticity-robust ones, but the cluster-robust AR confidence intervals are between 45 and 84 per cent greater than the cluster-robust ones based on the asymptotic distribution of the IV estimator. Furthermore, the cluster-robust AR confidence intervals contain considerably larger values than do the standard ones, so that they are compatible with much greater effects of culture on the level of development than are the standard confidence intervals. In Table 2, all the AR confidence intervals are substantially wider and contain larger values than the corresponding standard ones. This evidence that Tabellini's IV estimates of the effect of culture are in most cases subject to weak-instrument problems makes it particularly important that *literacy* and *pc_institutions* should be valid instruments, since even a very small correlation between a weak instrument and the structural equation error can result in IV estimates being highly inconsistent. The validity of these instruments is one of the subjects of the next section.

4. In search of better IV estimates

A fundamental difficulty with IV estimation is that it is necessary to assume that one instrument is valid. This is a maintained hypothesis, which cannot be tested. When two or more instruments are available, it is possible to investigate whether the use of different subsets of the instruments results in significant differences in the estimated coefficient of the endogenous variable. If this estimate does depend significantly on which subset of instruments is used in the IV estimation, then serious doubt is raised about the validity of some or all of the instruments. However, if all subsets of the instruments tell a similar story about the coefficient of the endogenous variable, then there is no obvious reason to think that some instruments are invalid, although it is still necessary to maintain the hypothesis that one of them is valid.

Tabellini adopts this approach to the investigation of instrument validity by using each of *literacy* and *pc_institutions* separately as a single instrument for *pc_culture* in a regression explaining *yp9500* and adding the variable not used as an instrument as an extra regressor. The results show that the estimated coefficient of *pc_culture* is substantially different according to which of the two variables is used as the single instrument.[4] When *literacy* is the single instrument, the estimated coefficient of *pc_culture* is 1.74, but when *pc_institutions* is the single instrument, it is 0.75. These different point estimates correspond to dramatic differences in the estimated effect of culture on the level of regional economic development. The value of *yp9500* for the Italian region of Calabria is 69.17, i.e., its per capita gross value added is about 69 per cent of the EU15 average. The corresponding value for the Italian region of Lombardy is 151.67. A coefficient of 0.75 for *pc_culture* implies that if the value of this conditional culture variable in Calabria increased to its value in Lombardy, cet. par., then the value of *yp9500* in Calabria would increase to 106.25. If the coefficient of *pc_culture* was 1.74 then the corresponding increase in Calabrian *yp9500* would be to 155.21. Thus a coefficient of 0.75 means that, if Calabria had Lombardy's culture, the gap between Calabrian and Lombard *yp9500* would be reduced by about 45%, but a coefficient of 1.74 means that this gap would be more than completely eliminated. The two different instruments therefore yield two enormously different estimates of the effect of culture. Tabellini notes that these differences raise doubts about the validity of the instruments, but although they lead him to consider the power of the overidentification test based on the *J* statistic, he does not explore what might lie behind them.[5]

In order to understand these differences in the estimated effects of the culture variables, Tables 3 and 4 present comparisons of IV and OLS estimates for the regression models obtained using first *literacy* and then *pc_institutions* as a single instrument for the culture variables. When *literacy* is used as the single instrument *pc_institutions* is included as an exogenous regressor, and similarly *literacy* is included as an exogenous

[4] Equations 1 and 2 in Table 7 of Tabellini (2010)

[5] One interpretation of the *J* test in this context is that it tests whether the difference between the two just-identified estimators is small relative to sampling variance.

Table 3: OLS and just-identified IV estimates of the relationship between culture and the level of economic development

Culture measure	Instrument	Standard errors	OLS estimate	95% OLS confidence interval	IV estimate	95% AR confidence interval	*C* test *p* value	First-stage *F* statistic
pc_culture	*literacy*	Robust	0.47	0.14 – 0.81	1.74	0.66 – 5.88	0.0361	8.72
		Clustered	0.47	-0.14 – 1.09	1.74	0.35 – 180	0.1260	5.72
pc_culture	*pc_institutions*	Robust	0.40	0.10 – 0.70	0.75	-0.08 – 1.71	0.2005	11.43
		Clustered	0.40	0.04 – 0.76	0.75	-0.14 – 3.86	0.4124	17.29
pc_culture_pos	*literacy*	Robust	0.67	0.21 – 1.12	2.16	0.87 – 5.07	0.0357	14.56
		Clustered	0.67	-0.05 – 1.39	2.16	0.46 – 16.30	0.1228	7.69
pc_culture_pos	*pc_institutions*	Robust	0.54	0.17 – 0.91	0.75	-0.07 – 1.70	0.4531	13.89
		Clustered	0.54	0.27 – 0.82	0.75	-0.15 – 3.24	0.6183	23.25
pc_children	*literacy*	Robust	0.45	0.01 – 0.89	3.19	1.04 – +∞	0.0199	2.36
		Clustered	0.45	-0.32 – 1.21	3.19	0.80 – +∞	0.1146	2.41
pc_children	*pc_institutions*	Robust	0.43	0.06 – 0.81	0.89	-0.09 – 1.81	0.1768	13.11
		Clustered	0.43	-0.17 – 1.04	0.89	-0.19 – 3.00	0.3398	45.11

Notes.
This table compares OLS and just-identified IV estimates of the effects of alternative measures of conditional culture on *yp9500* in regressions estimated on 67 observations. The estimation procedure weights observations by the inverse of the standard errors of the estimated measures of regional conditional culture to allow for different measurement errors across regions. The regressions also include *school*, *urb_rate1850* and country dummy variables as exogenous regressors, the estimated coefficients of which are not reported. The just-identified IV estimates are obtained using either *literacy* or *pc_institutions* as the single instrument for the culture variable.

regressor when *pc_institutions* is the single instrument. As in the previous section, only results for conditional measures of culture are reported. The results obtained when unconditional measures of culture are used are very similar, and in what follows the one case where there are differences worth remarking upon is noted.

Since the IV estimates reported in Tables 3 and 4 come from just-identified models, *J* statistics cannot be calculated. However, the *p* values of another statistic based on the value of the objective function in generalised method of moments estimation are reported in these tables. This is the *C* statistic for testing the null hypothesis that the regressor being treated as endogenous (in the present case, one of the culture variables) is actually an exogenous regressor (Hayashi (2000, 218-221), Baum et al. (2003)). It is the appropriate statistic for such a test with a heteroskedastic or clustered covariance matrix. The *C* statistic amounts to a test of whether there is a statistically significant difference between the OLS and IV estimates of the coefficient of the culture variables in the various regressions. If the null hypothesis that culture is an exogenous regressor is not rejected, then there is no need to resort to the IV estimator, and the OLS estimator, which has smaller asymptotic variance, is appropriate.

Since my focus in this section is on IV estimates that are robust to weak instruments, Tables 3 and 4 do not report the standard confidence intervals based on the asymptotic distribution of the IV estimator. As the IV estimates in these tables are obtained from just-identified models, there is no loss of information in using the AR procedure to construct robust confidence intervals for the estimated effects of the culture variables.

Several points stand out from Tables 3 and 4. The first is that *literacy* is clearly a weak instrument. In all but one of the cases in which *literacy* is used as the single instrument for the culture variables the first-stage *F* statistic is below 10. By contrast, the first-stage *F* statistics when *pc_institutions* is used as the single instrument are always greater than 10. However, the AR confidence intervals in this case are wider and contain larger values than do the (unreported) standard IV confidence intervals, so the just-

Table 4: OLS and just-identified IV estimates of the relationship between culture and the growth rate

Culture measure	Instrument	Standard errors	OLS estimate	95% OLS confidence interval	IV estimate	95% AR confidence interval	*C* test *p* value	First-stage *F* statistic
pc_culture	*literacy*	Robust	0.0066	-0.0017 – 0.0150	0.0255	0.0016 – 0.1818	0.1023	5.21
		Clustered	0.0066	-0.0040 - 0.0173	0.0255	-∞ – +∞	0.1368	3.54
pc_culture	*pc_institutions*	Robust	0.0068	-0.0008 – 0.0145	0.0137	0.0011 – 0.0453	0.1877	12.90
		Clustered	0.0068	-0.0016 – 0.0153	0.0137	-0.0027 – 0.0740	0.3162	14.10
pc_culture_pos	*literacy*	Robust	0.0109	-0.0017 – 0.0235	0.0336	0.0022 – 0.1263	0.1298	6.88
		Clustered	0.0109	-0.0105 – 0.0323	0.0336	-∞ – +∞	0.1920	3.81
pc_culture_pos	*pc_institutions*	Robust	0.0105	-0.0008 – 0.0218	0.0136	0.0010 – 0.0431	0.5650	16.02
		Clustered	0.0105	-0.0064 – 0.0274	0.0136	-0.0027 – 0.0656	0.6887	17.12
pc_children	*literacy*	Robust	0.0049	-0.0034 – 0.0133	0.0701	0.0039 – +∞	0.0491	0.70
		Clustered	0.0049	-0.0045 – 0.0143	0.0701	-∞ – +∞	0.1152	0.88
pc_children	*pc_institutions*	Robust	0.0059	-0.0014 – 0.0132	0.0166	0.0016 – 0.0450	0.0812	16.48
		Clustered	0.0059	-0.0019 – 0.0137	0.0166	-0.0038 – 0.0567	0.1544	26.30

(Continued on next page)

Table 4 (continued)

Notes.
This table compares OLS and just-identified IV estimates of the effects of alternative measures of conditional culture on *growth* in regressions estimated on 67 observations. The estimation procedure weights observations by the inverse of the standard errors of the estimated measures of regional conditional culture to allow for different measurement errors across regions. The regressions also include *school*, *urb_rate1850*, *lyp77* and country dummy variables as exogenous regressors, the estimated coefficients of which are not reported. The just-identified IV estimates are obtained using either *literacy* or *pc_institutions* as the single instrument for the culture variable.

identified IV estimates of the effects of culture obtained using *pc_institutions* are not free of weak-instrument concerns.

The second salient feature of these tables is the difference between the IV point estimates of the effect of culture, on both the level and the growth rate of regional economic development, depending on which variable is used as the single instrument. In these tables, the estimate obtained when *literacy* is the single instrument is between 86 and 320 per cent larger than the corresponding estimate when *pc_institutions* is the single instrument.

The third notable point in Tables 3 and 4 is the difference between the AR confidence intervals for the IV estimates depending on which variable is used as the single instrument. These confidence intervals are always much wider for the estimates using *literacy* than for the corresponding estimates using *pc_institutions*. Indeed there are some cases where the AR confidence intervals using *literacy* are unbounded above, and others in which they are unbounded both above and below. Even when the AR confidence intervals are bounded, the IV estimates of the effect of culture on regional economic performance obtained using *literacy* are highly imprecise.

The final notable feature of the results in these tables is that, although the just-identified IV point estimates of the effect of culture on regional economic performance are always greater than the corresponding OLS estimates (dramatically so when *literacy* is the single instrument), there is little evidence that the differences between the IV and OLS estimates are statistically significant. Although the IV point estimates are substantially larger than the OLS ones, there is considerable overlap between the OLS and AR 95 per cent confidence intervals in most cases, and it is only for the heteroskedasticity-robust estimates using *literacy* as the instrument in Table 3 that the OLS point estimate does not lie within the AR confidence interval. The IV estimates are not sufficiently precisely estimated for it to be possible to say that they are significantly different from the OLS ones. This conclusion is broadly supported by the p values of the C statistic reported in Tables 3 and 4. The null hypothesis that the culture variables are exogenous cannot be rejected in most cases. In Table 3 this null hypothesis is rejected at the 0.05 level for the three cases in which the estimates are obtained for heteroskedasticity-robust standard errors using *literacy* as the single instrument. In Table 4 it is rejected once at the 0.05 level (when *pc_children* is the culture measure and the estimates are for heteroskedasticity-robust standard errors using *literacy* as the instrument) and once at the 0.10 level (in the corresponding case when *pc_institutions* is the instrument). For the estimates using cluster-robust standard errors, which, as noted in section 3, are probably the more relevant ones, the null hypothesis that the culture variables are exogenous is never rejected, even at the 0.10 level, irrespective of which variable is used as the single instrument and which measure of economic performance is the dependent variable.

Although it is clear from Tables 3 and 4 that *literacy* is an extremely weak instrument, this in itself does not mean that the rejections of the null hypothesis that culture is exogenous which occur in some cases when *literacy* is used as the single instrument require no explanation. After all, if an instrument is weak the finite-sample bias of the IV estimator towards the OLS estimator is accentuated. So when in Tables 3 and 4 *literacy* is used as the instrument and the *C* test fails to reject the null hypothesis that culture is exogenous, it may do so incorrectly because of this accentuated bias. Furthermore, the point estimates reported in Tables 3 and 4 when *literacy* is the single instrument are much larger than the corresponding OLS ones. A proper understanding of the results in Tables 3 and 4 requires an explanation of these features.

For the just-identified estimates using *pc_institutions* as the instrument, the OLS point estimate always lies within the AR confidence interval and the *C* test never rejects the null hypothesis that culture is exogenous at the 0.05 level. Thus there is no reason to incur the efficiency costs of IV estimation of the regression model which includes *literacy* as an exogenous regressor as well as a measure of culture: the evidence shows that the OLS estimator is appropriate in this case. Table 5 shows the estimated coefficients of *literacy* and the culture variables obtained from OLS estimation of this model. Heteroskedasticity-robust standard errors are shown as the upper figure in

Table 5: OLS estimates of the effects of culture and past literacy on economic performance

Regressors	Dependent variable					
	yp9500	*yp9500*	*yp9500*	*growth*	*growth*	*growth*
pc_culture	0.400 (0.149)*** (0.151)**			0.0068 (0.0038)* (0.0036)*		
pc_culture_pos		0.544 (0.185)*** (0.118)***			0.0105 (0.0056)* (0.0071)	
pc_children			0.433 (0.188)** (0.256)			0.0059 (0.0036) (0.0033)
literacy	0.680 (0.285)** (0.288)**	0.637 (0.274)** (0.276)*	0.759 (0.268)*** (0.279)**	0.0075 (0.0048) (0.0046)	0.0067 (0.0047) (0.0050)	0.0093 (0.0044)** (0.0048)*

Notes.
This table shows the effects of alternative measures of culture and *literacy* on different measures of economic performance in OLS regressions estimated on 67 observations. The estimation procedure weights observations by the inverse of the standard errors of the estimated measures of regional conditional culture to allow for different measurement errors across regions. When the dependent variable is *yp9500* the regressions also include *school*, *urb_rate1850* and country dummy variables as exogenous regressors, and when the dependent variable is *growth* the regressions include these regressors and *lyp77*. The estimated coefficients of these additional regressors are not reported. The upper figures in parentheses are heteroskedasticity-robust standard errors and the lower ones are cluster-robust standard errors. * denotes significance at the 0.10 level, ** denotes significance at the 0.05 level and *** denotes significance at the 0.01 level.

parentheses below each coefficient estimate, and cluster-robust standard errors as the lower such figure. It is clear that *literacy* is statistically significantly associated with the level of regional economic development even when controlling for the effects of culture, contemporaneous education, past development and country fixed effects. This raises serious doubts about the exclusion of *literacy* from the regression models explaining the level of economic development and hence about its validity as an instrument. It should be emphasised that, although in general it is not appropriate to use OLS regressions to judge the suitability of variables as instruments, in this case the OLS estimator has been tested against an IV alternative and shown to be preferable.

If, as Table 5 suggests is likely, *literacy* is positively correlated with the unobservable error in the structural equation for regional economic performance, it is possible to explain why the point estimates in Tables 3 and 4 are much larger when *literacy* is used as the single instrument and hence why, in this case, there are some rejections of the null hypothesis that culture is an exogenous regressor. As emphasised by Bound et al. (1995), if an instrument is only weakly correlated with a potentially endogenous variable, then it is possible for even a small correlation between the instrument and the error in the structural equation to result in a very large inconsistency in the IV estimate of the coefficient of the potentially endogenous variable. The inconsistency of the IV estimate may well be greater than the inconsistency of the OLS estimate. In the present context, where *literacy* is a weak instrument for the culture variables and seems to be positively correlated with the error in the structural equation, this possibility means that the probability limit of the IV estimate of the effect of culture on regional economic performance using *literacy* as the single instrument is likely to be substantially larger than the true value of this effect. Thus the difference in the magnitudes of the IV estimates in Tables 3 and 4 using *literacy* and *pc_institutions* as instruments can be explained as the consequence of *literacy* being a weak instrument which is positively correlated with the error in the structural equation. Note that this explanation applies to the regressions in which the dependent variable is the regional growth rate as well as to those in which it is the level of regional development. It is certainly true that the results in Table 5 show much less evidence of a statistically significant association between the regional growth rate and *literacy* than between the level of regional development and *literacy*. But when an instrument is weak, as *literacy* is, even a very small correlation between the instrument and the error in the structural equation can lead to a large inconsistency in the IV estimate.

Literacy thus appears not to be a valid instrument for the culture variables used in Tabellini's analysis, from which it follows that neither the IV estimates using it as the single instrument in Tables 3 and 4 nor the IV estimates in Tables 1 and 2 (when *literacy* and *pc_institutions* are jointly used as instruments) are consistent. Is *pc_institutions* a valid instrument? No definite answer to this question is possible in the absence of another instrument that can be correctly excluded from the regression equations explaining regional economic performance, and such an additional instrument is conspicuous by its absence. The *J* tests reported in Tables 1 and 2 cannot be used to argue that *pc_institutions* is a valid instrument, because they test whether it is valid conditional on *literacy* being a valid instrument. However, although I recognise that treating *pc_institutions* as a valid instrument is a maintained hypothesis, there are some grounds for regarding it as such an instrument. If *pc_institutions* is added as a regressor to the models reported in Table 5, it is not statistically significantly associated with regional economic performance, while *literacy* continues to have such an association with the level of economic development. This provides some reason for regarding the just-identified IV estimates in Tables 3 and 4 using *pc_institutions* as being consistent and hence the ones on which attention should be focused in order to establish whether culture has a causal effect on economic performance.

The IV estimates using *pc_institutions* as the single instrument are unfortunately not very informative about the effects of culture on regional economic performance. For all three culture measures, the AR 95 per cent confidence intervals for the IV estimates of the effect of culture on the level of economic development in Table 3 extend from small negative to large positive values. The upper bounds of the heteroskedasticity-robust AR confidence intervals are more than double the corresponding point estimates, while for the cluster-robust AR confidence intervals these upper bounds are between three and five times as large as the corresponding point estimates. On the basis of these confidence intervals it is not possible to reject at the 0.05 level the null hypothesis that culture has no effect on the level of regional development. But these confidence intervals are so wide that they are also compatible with culture having a small negative effect on regional development as well as with its having an unbelievably large positive effect.[6] The same general conclusion applies to the AR 95 per cent confidence intervals in Table 4 for the effect of culture on the regional growth rate using *pc_institutions* as the instrument. The heteroskedasticity-robust confidence intervals in that table, though wide, do allow rejection at the 0.05 level of the null hypothesis that the effect of culture on the growth

[6] Continuing with the example used on page 12, a coefficient of 0.75 for *pc_culture* implies that if the value of this conditional culture variable in Calabria increased to its value in Lombardy, cet. par., then the value of *yp9500* in Calabria would increase to 106.25. If the coefficient of *pc_culture* was 3 then the corresponding increase in Calabrian *yp9500* would be to 217.51. A coefficient of 3 or more therefore implies such large effects of culture on the level of regional development as to be incredible.

rate is zero. However, this null hypothesis cannot be rejected on the basis of the cluster-robust confidence intervals, which are *prima facie* the more relevant ones. But the AR confidence intervals in Table 4 are also compatible with incredibly large positive effects of culture on regional growth rates.[7] Overall it is clear that the IV estimates using *pc_institutions* as the instrument are so imprecise that they provide very little evidence about the effects of culture on regional economic performance.

The lack of precision in the IV estimates using *pc_institutions* as the single instrument makes it unsurprising that they are not statistically significantly different from the corresponding OLS estimates. These IV point estimates are always larger than the OLS ones, though to nothing like the same extent as those obtained using *literacy* as the single instrument. But the *C* test never rejects (at the 0.05 level) the null hypothesis that the culture variables can be treated as exogenous in the estimation of their effects on regional economic performance. The absence of evidence that culture is an endogenous variable means that there is no need to incur the efficiency costs of IV rather than OLS estimation, and hence the OLS estimates of the coefficients of culture in the regional economic performance regressions shown in Table 5 are preferable on statistical grounds to the corresponding IV estimates shown in Table 4. These OLS results show that the three summary measures of culture are statistically significantly associated with the level of regional economic development when controlling for the effects of past education, contemporaneous education, past development and country fixed effects. However, the evidence in support of a statistically significant association between the three culture measures and the regional growth rate (when adding the initial level of development to the list of control variables) is less strong. The OLS estimates of the coefficients of culture in the regional growth regressions are only ever significant at the 0.10 level. But there is somewhat stronger evidence of such an association between culture and the

[7] Continuing with the example on page 12 and in footnote 6, the value of *growth* for Calabria is equivalent to a increase of about 81 per cent in per capita gross value added over the period 1977-2000. The corresponding figure for Lombardy is 128.5 per cent. A coefficient of 0.0137 for *pc_culture* implies that if Calabria had had Lombardy's culture, cet. par., then the increase in per capita gross value added over this period would have been 256 per cent, which is already such a large figure as to strain credulity. If the coefficient was instead 0.0400 then the implied increase in Calabrian per capita gross value added would have been 1208 per cent, which is quite unbelievable.

regional growth rate if unconditional rather than conditional measures of culture are used, in which case some of the estimated coefficients are significant at the 0.05 level.

This raises the question of whether the OLS estimates of the coefficients of the culture variables should be given a causal interpretation, given that the *C* tests do not reject the null hypothesis that the culture variables can be treated as exogenous in the regression analysis. The answer is no. If the IV estimates of the effects of culture were precise, and they were not statistically significantly different from the OLS ones, then it would indeed be justifiable to give a causal interpretation to the OLS coefficient estimates. But the IV estimates of the effects of culture are not precise: quite the contrary. This lack of precision means that the null hypothesis that the OLS estimates are the same as the IV estimates cannot be rejected, and hence the OLS estimates are preferable on statistical grounds. But the lack of precision in the IV estimates implies that the test of this null hypothesis has low power. Consider, for example, the estimates of the coefficient of *pc_culture* in Table 3 using *pc_institutions* as the instrument. The AR 95% confidence interval shows that it is not possible to reject at the 0.05 level the null hypotheses that the IV estimate is equal either to 0.4, the value of the OLS point estimate, or to any value in the 95 per cent OLS confidence interval. But the AR confidence interval also shows that it is not possible to reject the null hypotheses that the IV estimate is equal to 0, or that it is equal to 1.6. The IV estimates are so uninformative about the possible causal effect of culture on economic performance that they are compatible with a very wide range of possible values for this effect, including those resulting from the OLS estimator. Consequently there are no statistical grounds for preferring the IV to the OLS estimates. But this does not mean that the OLS estimates identify a causal effect of culture on economic performance. The fact that a poorly-determined IV estimate is not significantly different from the OLS estimate is evidence that the instruments used are unable to identify the possible causal effect with any degree of precision, not that the OLS estimate can reasonably be given a causal interpretation. As Tabellini's discussion makes clear, there are good theoretical reasons to expect that economic development will have a causal effect on culture as well as possibly being causally influenced itself by culture. In the absence of well-determined IV estimates, these reasons for not giving the

OLS estimates a causal interpretation remain compelling, despite their being statistically indistinguishable from the IV ones.

5. Conclusion

This paper has shown that Tabellini's claim to have provided evidence that culture has a causal effect on economic performance does not stand up to careful scrutiny. Tabellini's claim is based on the use of past literacy and past political institutions as instruments to identify a causal effect of culture on economic performance. But neither of these is a satisfactory instrument. Past literacy is a weak instrument and furthermore does not appear to be a valid instrument. Past political institutions cannot definitely be said to be a valid instrument, but there are no obvious grounds for thinking that it is an invalid one. However, it too seems to be a weak instrument, because the confidence intervals for the effect of culture on economic performance obtained using it as the single instrument are very wide. These IV estimates are so imprecise that they provide hardly any information about the possible causal effect of culture, and consequently are not statistically significantly different from OLS estimates. But the lack of precision in the IV estimates also makes it impossible to give the OLS estimates a causal interpretation, despite the fact that they are not significantly different from the IV ones. In order to make any statements at all about possible causal influences of culture on economic performance it is necessary to have instruments that provide more precise estimates of these effects. Such estimates would not only mean that the causal effect of culture was well-determined but also that, if there were no statistically significant difference between the IV and OLS estimates, the latter could be given a causal interpretation.

Tabellini's data for European regions do not, therefore, provide any evidence of a causal influence of culture on economic performance. What these data do provide, however, is unambiguous evidence of statistically significant associations between summary measures of culture and the level of regional economic development. There is also some evidence of an association between these measures of culture and regional growth rates, although this is less clear-cut. Whether these associations reflect a causal influence of culture on economic performance or a causal influence of economic performance on culture, or indeed both influences, remains to be established. The identification of a causal effect of culture on regional economic performance in Europe requires better instruments than those used by Tabellini.

References

Anderson, T.W. and H. Rubin (1949). 'Estimation of the parameters of a single equation in a complete system of stochastic equations', Annals of Mathematical Statistics, 20, 46-63.

Baum, C.F., M.E. Schaffer and S. Stillman (2003). 'Instrumental variables and GMM: estimation and testing', The Stata Journal, 3, 1-31.

Bound, J., D.A. Jaeger and R.M. Baker (1995). 'Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak', Journal of the American Statistical Association, 90, 443-50.

Chernozhukov, V. and C. Hansen (2005). 'The reduced form: a simple approach to inference with weak instruments', working paper, available at SSRN: http://ssrn.com/abstract=937943 or http://dx.doi.org/10.2139/ssrn.937943

Chernozhukov, V. and C. Hansen (2008). 'The reduced form: a simple approach to inference with weak instruments', Economics Letters, 100, 68-71.

Dufour, J-M. (2003). 'Identification, weak instruments, and statistical inference in econometrics', Canadian Journal of Economics, 36, 767-808.

Hahn, J. and J.A. Hausman (2003). 'Weak instruments: diagnosis and cures in empirical econometrics', American Economic Review, Papers and Proceedings, 93, 118-25.

Hayashi, F. (2000). Econometrics. Princeton: Princeton University Press.

Staiger, D. and J.H. Stock (1997). 'Instrumental variables regression with weak instruments', Econometrica, 65, 557-86.

Tabellini, G. (2010). 'Culture and institutions: economic development in the regions of Europe', Journal of the European Economic Association, 8, 677-716.