International and Development Economics

B

ridging the Barriers:
Knowledge Connections,
Productivity, and Capital
Accumulation

R. Quentin Grafton,
Tom Kompas,
P. Dorian Owen

04–5

Asia Pacific School of Economics and Government
THE AUSTRALIAN NATIONAL UNIVERSITY

<http://apseg.anu.edu.au>

Abstract

The paper explains the large differences in cross-country productivity performance by modeling and testing the effects of social barriers to communication on productivity and capital accumulation. In an optimal growth model, social barriers to communication that impede the formation of knowledge connections are shown to reduce both transitory and steady-state levels of total factor productivity (TFP), per capita consumption, and reproducible capital. A 'bridging' parameter in the growth model that lowers the disutility of forming knowledge connections generates testable and dynamic implications about the effects of social barriers on capital, consumption, and productivity. Extensive empirical testing of the theoretical propositions yields a robust and theoretically consistent result — linguistic barriers to communication reduce productivity and capital accumulation. The findings provide a theoretical justification and a robust explanation for cross-country differences in TFP, and fresh insights into how productivity 'catch up' may be initiated.

Asia Pacific School of Economics and Government
THE AUSTRALIAN NATIONAL UNIVERSITY

<http://apseg.anu.edu.au>

Bridging the Barriers: Knowledge Connections, Productivity, and Capital Accumulation*

R. Quentin Grafton§

Asia Pacific School of Economics and Government
The Australian National University

Tom Kompas

Asia Pacific School of Economics and Government
The Australian National University

P. Dorian Owen

Department of Economics
University of Otago

Abstract

The paper explains the large differences in cross-country productivity performance by modeling and testing the effects of social barriers to communication on productivity and capital accumulation. In an optimal growth model, social barriers to communication that impede the formation of knowledge connections are shown to reduce both transitory and steady-state levels of total factor productivity (TFP), per capita consumption, and reproducible capital. A ‘bridging’ parameter in the growth model that lowers the disutility of forming knowledge connections generates testable and dynamic implications about the effects of social barriers on capital, consumption, and productivity. Extensive empirical testing of the theoretical propositions yields a robust and theoretically consistent result — linguistic barriers to communication reduce productivity and capital accumulation. The findings provide a theoretical justification and a robust explanation for cross-country differences in TFP, and fresh insights into how productivity ‘catch up’ may be initiated. (*JEL* O41, C61, C21)

Key Words: knowledge connections, productivity, economic growth

Running Title: Knowledge Connections and Productivity

§Contact Author:

R. Quentin Grafton
Asia Pacific School of Economics and Government
J.G. Crawford Building (Bldg 13)
Ellery Crescent
The Australian National University
Acton, ACT 0200
AUSTRALIA

tel +61-2-6125-6558

fax +61-2-6125-5570

email quentin.grafton@anu.edu.au

November 2004

I. Introduction

The existence of enormous differences in the levels of productivity and factor accumulation across countries constitutes one of the most perplexing issues in economics. Many explanations have been offered for the large disparities, including the initial level of capital stocks (physical, natural and human), human capital externalities, macroeconomic stability, quality of institutions, geography, trade openness, and rules over foreign investment. Increasingly, economists are exploring the ways that public and civic institutions, social mores and norms of behavior, and social networks influence economic activity. Such analysis recognizes that economic growth goes beyond factor accumulation and is also linked to social interactions between people.

In this paper we focus on the macroeconomic effects of social barriers to communication and their consequences for total factor productivity (TFP) and (human or reproducible) capital accumulation. In an optimal growth model, we show that social barriers impede knowledge communication links that otherwise make labor more productive. The model generates testable propositions, namely, that lower values of a ‘bridging’ parameter raise the disutility of forming knowledge connections across agents, which, in turn, reduces both transitory and steady-state levels of TFP, per capita consumption, and capital (physical or human). Extensive empirical testing of the theoretical propositions yields a robust and theoretically consistent result — linguistic barriers to communication reduce productivity and capital accumulation.

The theoretical and empirical results are consistent with a number of important stylized facts at a regional and global level and stress the importance of social barriers to communication on economic performance. The findings help explain the persistence of differences in cross-country TFP, the existence of country growth laggards and leaders, and provide fresh insights as to how countries might initiate productivity ‘catch up’.

The paper is organized as follows. Section II solves an optimal growth model of the effects of social barriers to communication and analyzes the implications for TFP and factor accumulation. Section III describes the data and the empirical model used to test the theoretical propositions. Section IV reports the empirical results. The economic implications of social impediments to communication are examined in section V. Concluding remarks are offered in section VI.

II. Knowledge Connections and Social Barriers to Communication

Our modeling of the macroeconomic effects of social barriers to communication on productivity and growth is novel, although Parente and Prescott (2000) also emphasize the importance of barriers to knowledge, in particular a lack of competition, that impede the adoption of new technologies. Our work has similarities to important contributions by Lazear (1999), Nettle (2000), Rauch (2001) and Grafton, Knowles and Owen (2004) that emphasize the importance of linguistic and cultural diversity for, respectively, exchange and trade between individuals, aggregate per capita GDP, international trade, and per capita income and productivity. We also owe a debt to those who have tested for the interaction between economic performance and various characterizations of social capital, social infrastructure or social capability (Easterly and Levine 1997; Hall and Jones 1999; Helliwell and Putnam 1995; Knack and Keefer 1997; Temple and Johnson 1998; Zak and Knack 2001). Others, such as Bénabou (1996), stress the importance of heterogeneity, especially with respect to inequality and school funding, while Gradstein and Justman (2002) examine the importance of social polarization in terms of human capital formation. None of the above approaches, however, develops a theoretical model of the effects of social barriers to communication on macroeconomic performance, nor has any previous study linked these effects to explain differences in both factor accumulation and productivity.

Using an optimal growth model, we posit that economy-wide output is increasing in the level of a reproducible capital stock (physical or human), labor, and the number of knowledge communication links between agents. Our interpretation of the model is that communication links help in the creation of productivity-enhancing ideas, and also in the transmission of tacit knowledge. Differences across agents make communication and interaction worthwhile via ‘cross-fertilization’ of knowledge and ideas — complementary knowledge — but social barriers that inhibit communication or interchange (such as linguistic differences) raise the cost of mutually beneficial and productivity-enhancing communications.

Our modeling implicitly incorporates three key ideas. One, cooperation and group interactions enable economies to use large amounts of specialized knowledge (Becker and Murphy 1992; Lucas 1988; Rivera-Batiz and Romer 1991). Two, although knowledge is inherently nonrival, the creation and transfer of tacit knowledge or ‘know-how’ is highly dependent on communication links within social groups (Brown and Duguid 2000; Coleman, Katz, and Menzel 1966; Marshall 1916; Powell 1990; Ryan and Gross 1943; Saxenian 1994; Calvó-Armengol and Jackson 2004) and also by ‘weak ties’ or ‘bridges’ (Granovetter 1973) across social groups (Rogers 1995; Meyer 1998; Valente 1995). Three, individuals communicate more easily the greater the similarity between them (Tarde 1895; Lazarsfeld and Merton 1954; Bertrand, Luttmer, and Mullainathan 2000), and communication and cooperation across social groupings, such as across linguistic barriers, is often much more limited than within groups (Bénabou 1996; Borjas 1992 and 1995; Burt 2002; Davis 1967; Schelling 1978; Sherif, Harvey, White, Hood, and Sherif 1961; Solo 1967).

A. The Model

To capture the effects of social barriers to communication we assume that a representative agent’s utility function, given by equation (1), depends positively on per capita consumption

at time t , $c(t)$, and negatively on the effort required to establish knowledge connections across agents defined by $\varepsilon(s(t))$, i.e.,

$$(1) \quad U(c, \varepsilon(s)) = \int_0^\infty \left[\frac{c(t)^{1-\theta}}{1-\theta} - \frac{\varepsilon(s(t))^z}{z\beta} \right] e^{-\rho t} dt$$

In (1), θ is the inverse of the intertemporal elasticity of substitution (assumed to lie between zero and one), z is a communication disutility coefficient that is greater than one, ρ is the rate of time preference and β is an economy-wide ‘bridging’ parameter that affects the ease of establishing knowledge connections. The bridging parameter is taken to be sufficiently positive to ensure that $U(c, \varepsilon(s))$ is jointly concave and is bounded from above by the assumption that, even in the absence of social barriers, establishing knowledge links between individuals is always costly. Effort in forming connections, $\varepsilon(s(t))$, is an implicit function of the number of connections, where $\varepsilon(\cdot)$ is the effort function and $s(t)$ is the number of knowledge connections. The number of connections has a lower bound of zero.

Equation (1) is consistent with an intertemporal consumption/leisure model of individual preferences, and the negative effect of $\varepsilon(s(t))$ on utility incorporates the notion that the time involved in making knowledge connections is privately costly. An increase in the bridging parameter β , which makes it easier for agents to form knowledge connections, lowers the ‘utility-cost’ of forming connections. The bridging parameter represents the initial conditions in the economy, such as the degree of linguistic diversity, that help determine the cost of establishing knowledge links with other people. Low levels of the bridging parameter would represent an economy where social barriers to communication, such as a lack of a common language, make it expensive to establish knowledge links in terms of the disutility of effort.

To complete the model, aggregate output is determined by

$$(2) \quad Y(t) = \alpha_0 (s(t)N(t))^{\alpha_1} K(t)^{\alpha_2}$$

where α_0 is economy-wide productivity, $N(t)$ is the size of the labor force, $s(t)N(t)$ is knowledge connections-augmented units of labor where the productivity of labor is increasing in the number of economy-wide knowledge connections, and $K(t)$ is the reproducible capital stock (physical or human). For convenience, we assume a one-to-one mapping between the effort from making knowledge connections and the number of connections, i.e., $\varepsilon(s(t)) = s(t)$, and that (2) exhibits constant returns to scale. Neither assumption, however, is essential to derive our results.

In per capita form, and suppressing t , the economy's aggregate production function is given by

$$(3) \quad y = \alpha_0 s^{\alpha_1} k^{\alpha_2}$$

where $y = Y/N$ and $k = K/N$. The change in the reproducible capital stock with respect to time is governed by

$$(4) \quad \dot{k} = y - c$$

B. Theoretical Results

To solve the optimization problem we maximize (1) subject to (4), the initial condition $k(0) = k_0$ and the necessary feasibility constraints. We note that c and s are both control variables, and define λ as the co-state variable.

Along the optimal path, both (4) and the following necessary conditions must be satisfied for all t :

$$(5) \quad c^{-\theta} = \lambda$$

$$(6) \quad \frac{s^{z-1}}{\beta} = \lambda \alpha_1 \alpha_0 s^{\alpha_1-1} k^{\alpha_2}$$

$$(7) \quad \frac{\dot{\lambda}}{\lambda} = \rho - \alpha_0 \alpha_2 s^{\alpha_1} k^{\alpha_2-1}$$

Equation (6) shows that along the optimal growth path the representative agent will ensure that the instantaneous marginal disutility from making knowledge connections equals the instantaneous marginal benefit from production. Higher effort today, and thus lower current utility, generates more knowledge connections, greater capital accumulation, higher output and, ultimately, larger future consumption.

Both output and the effort from making connections are increasing in the number of connections. For any number of connections *less* than the optimal steady-state s^* the marginal utility from consumption from an extra connection exceeds the disutility of effort, leading to an increase in the desired number of connections. Higher levels of the economy-wide bridging parameter β reduce the disutility of effort from making connections and, thus, increase both transitional and steady-state number of connections. An increased number of connections, in turn, has dynamic implications because it raises both the capital-labor ratio and per capita consumption along the optimal growth path, and also at the steady state.

The intuition for the dynamic effort-output relationship can be shown with equations (5) and (6) that, together, imply

$$(8) \quad s = \left(c^{-\theta} \alpha_0 \alpha_1 \beta k^{\alpha_2} \right)^{1/(z-\alpha_1)}.$$

Thus a once-and-for-all increase in the bridging parameter β , which reduces the disutility associated with making connections, raises the number of knowledge connections along the optimal growth path. Equation (8), along with the necessary conditions, can be used to derive the following transition paths:¹

$$(9) \quad \dot{c} = \frac{c}{\theta} \left[\alpha_0^{\frac{z}{z-\alpha_1}} \alpha_1^{\frac{\alpha_1}{z-\alpha_1}} \beta^{\frac{\alpha_1}{z-\alpha_1}} k^{\frac{z\alpha_2}{z-\alpha_1}-1} c^{\frac{-\theta\alpha_1}{z-\alpha_1}} \alpha_2 - \rho \right]$$

and

$$(10) \quad \dot{k} = \alpha_0^{\frac{z}{z-\alpha_1}} \alpha_1^{\frac{\alpha_1}{z-\alpha_1}} \beta^{\frac{\alpha_1}{z-\alpha_1}} k^{\frac{z\alpha_2}{z-\alpha_1}} c^{\frac{-\theta\alpha_1}{z-\alpha_1}} - c.$$

At the steady state, because

$$(11) \quad s^{\alpha_1} = \left(\frac{\rho}{\alpha_0 \alpha_2} \right) k^{1-\alpha_2},$$

per-capita consumption (c^*) is a function of the steady-state reproducible capital (k^*) and is expressed as follows:

$$(12) \quad c^* = k^* \left[\frac{\rho}{\alpha_2} \right].$$

Given (8) and the steady-state value for consumption given by (12) it follows that

$$(13) \quad \alpha_0 \alpha_2 \left[\left(k \frac{\rho}{\alpha_2} \right)^{-\theta} \alpha_0 \alpha_1 \beta k^{\alpha_2} \right]^{\frac{\alpha_1}{z-\alpha_1}} k^{\alpha_2-1} = \rho.$$

Thus the steady-state values for consumption and reproducible capital can be written as

$$(14) \quad c^* = \left(\frac{\rho}{\alpha_2} \right) \left[\alpha_0^z \alpha_1^{\alpha_1} \alpha_2^{z-\alpha_1+\theta\alpha_1} \rho^{\alpha_1-z-\theta\alpha_1} \beta^{\alpha_1} \right]^{\frac{1}{\alpha_1(z+\theta-1)}}$$

and

$$(15) \quad k^* = \left[\alpha_0^z \alpha_1^{\alpha_1} \alpha_2^{z-\alpha_1+\theta\alpha_1} \rho^{\alpha_1-z-\theta\alpha_1} \beta^{\alpha_1} \right]^{\frac{1}{\alpha_1(z+\theta-1)}}.$$

These results yield the following proposition.

PROPOSITION 1: *A lower value of the bridging parameter β reduces both the transitory and steady-state levels of per capita consumption and capital.*

It follows immediately that, if $z > 1$, which is required for convexity in the effort-disutility relationship, and $0 < \theta < 1$, proposition 1 holds true. The significance of this result is that the initial conditions, or policy actions, that influence the cost of forming knowledge connections have both transitory and steady-state implications. The implication is that actions successful at overcoming social barriers to communication will increase the transmission and diffusion of tacit knowledge that, in turn, will increase both the growth and steady-state levels of capital and consumption.

The intuition for our results is that higher levels of the bridging parameter lower the costs of forming knowledge connections and, therefore, increase the *knowledge connections-augmented* rate of return given by $\alpha_0^{\frac{z}{z-\alpha_1}} \alpha_1^{\frac{\alpha_1}{z-\alpha_1}} \beta^{\frac{\alpha_1}{z-\alpha_1}} k^{\frac{z\alpha_2-1}{z-\alpha_1}} c^{\frac{-\theta\alpha_1}{z-\alpha_1}} \alpha_2$ in equation (9). A higher rate of return on capital induces factor accumulation and raises the steady-state values of both per capita consumption and capital. By contrast to a comparable Ramsey model, where the steady-state value of capital depends only on the rate of time preference and is also policy invariant, we find that the level of the bridging parameter affects both the transition paths and steady-state values of capital and consumption.

We can derive the theoretical implications of the bridging parameter for TFP by first substituting (12), or the expression for per-capita consumption as a function of steady-state reproducible capital, into (8) — the derived expression for the number of knowledge connections — to obtain,

$$(16) \quad s = (\alpha_0 \alpha_1)^{\frac{1}{z-\alpha_1}} \beta^{\frac{1}{z-\alpha_1}} \left(\frac{\rho}{\alpha_2} \right)^{\frac{-\theta}{z-\alpha_1}} k^{\frac{\alpha_2-\theta}{z-\alpha_1}}.$$

By substituting (16) into (3) — the expression for per capita output — and multiplying by N , we can derive a closed-form solution for aggregate output given by,

$$(17) \quad Y = AN^{1-\frac{\alpha_1(\alpha_2-\theta)}{z-\alpha_1}+\alpha_2} K^{\frac{\alpha_1(\alpha_2-\theta)}{z-\alpha_1}+\alpha_2},$$

where A is TFP and derived to be

$$(18) \quad A = \alpha_0 (\alpha_0 \alpha_1)^{\frac{\alpha_1}{z-\alpha_1}} \left(\frac{\rho}{\alpha_2} \right)^{\frac{-\theta \alpha_1}{z-\alpha_1}} \beta^{\frac{\alpha_1}{z-\alpha_1}}.$$

An intertemporal version of TFP can also be derived that shows that the time path for productivity is increasing in the bridging parameter. This result, and also equation (18), yield the following proposition.

PROPOSITION 2: *A lower value of the bridging parameter β reduces both the transitory and steady-state levels of total factor productivity.*

Our result provides a causal explanation for cross-country differences in TFP, and also implies that policy actions that can overcome social barriers to forming knowledge connections can initiate productivity ‘catch up’. Both propositions 1 and 2 can be tested using cross-country data and measures of TFP, factor accumulation (physical and human), social barriers to communication, and other variables.

III. Tests of the Propositions

Alesina, Devleeschauwer, Easterly, Kurlat and Wacziarg (2003) is the source of data on three measures of social barriers to communication based on linguistic, ethnic and religious fractionalization. Fractionalization, especially linguistic fractionalization, provides a measure of the level of an economy-wide bridging parameter where lower levels of the parameter are represented by *higher* levels of fractionalization. In our testing we also include explanatory variables that may mitigate the effects of social barriers to communication on productivity, such as the level of social infrastructure, trade openness, measures of mass communications, and population density.

Each fractionalization measure is calculated as one minus a Herfindahl index of ethnolinguistic group share and represents the probability of two randomly selected individuals being from a different social group, i.e.,

$$(19) \quad FRAC_i = 1 - \sum_j^n f_{ji}^2$$

where f_{ji} is the share of (linguistic, ethnic or religious) group j in country i . The primary source of the Alesina *et al.* (2003) fractionalization measures is the Encyclopedia Britannica (2000), which provides data on 1,055 major linguistic groups for 201 countries or dependent territories. Using other data sources for cross-checking, Alesina *et al.* (2003) calculate, for the early to mid 1990s, measures of ethnic (*Ethnic*), linguistic (*Language*) and religious (*Religion*) fractionalization for up to 215 countries. The three fractionalization measures we use have been investigated by Alesina *et al.* (2003) and Alesina and La Ferrara (2004) as possible determinants of long-run growth. We emphasize, however, that our paper is the first

to examine the effects of the Alesina *et al.* (2003) fractionalization measures on productivity and capital accumulation.

In addition to the Alesina *et al.* (2003) indexes, we also use Fearon's (2003) cultural fractionalization measure, *Culture*, which is based on the structural distance between languages. For example, *Culture* accounts for the fact that linguistic barriers (e.g., in Cyprus) between Greek and Turkish are much greater, because they are structurally unrelated languages, than (e.g., in Ukraine) between Russian and Ukrainian which are Indo-European, Slavic and East Branch languages (Fearon 2003, pp. 211-212). All four fractionalization indexes reflect the number and relative sizes of distinct social groups within a country. Cross-country summary statistics of the fractionalization measures and other key variables are provided in Table 1.

A. Proposition 1

Proposition 1 implies that the higher is the *initial* level of social barriers to communication (lower β), the lower will be the transitory and steady-state levels of reproducible capital (physical or human). We test this proposition by estimating the following equations:

$$(20) \quad \Delta AYS_i = \delta_0 + \delta_1 FRAC_i + \delta_2 AYS60_i + \delta_3 \ln RGDPW60_i + \mu_i$$

$$(21) \quad \Delta \ln KAPW_i = \gamma_0 + \gamma_1 FRAC_i + \gamma_2 \ln KAPW65_i + \gamma_3 \ln RGDPW60_i + \nu_i$$

where ΔAYS is the change in Barro and Lee's (2001) measure of the average years of schooling in the total population aged 15 years and over between 1960 and 1999, $\Delta \ln KAPW$ is the change in the natural log of real physical capital stock per worker between 1965 and 1990 (from the Penn World Tables) and subscript i denotes observations for country i .

For *FRAC*, as well as the measures constructed by Alesina *et al.* (2003) and Fearon (2003), we also use an ethnolinguistic fractionalization index for 1960, *ELF*, obtained from La Porta *et al.* (1999). Although Alesina *et al.* (2003) argue that fractionalization measures exhibit considerable time persistence, *ELF* is, on balance, our preferred regressor to test proposition 1 because it is dominated by estimates dated around the base-period, thus providing more of an initial measure of the social barriers to communication. *AYS60* and *lnKAPW65* are base-period values for the respective capital stock proxies. *lnRGDPW60* is (the natural log of) real gross domestic product per worker in international prices in 1960. The appended error terms, μ_i and ν_i , are country specific and assumed to be independently and normally distributed. Consistency with proposition 1 requires that the estimated coefficient on the base-period fractionalization measure be negative and statistically significant.

B. Proposition 2

To test whether higher social barriers to communication (lower β) have a negative effect on TFP, as predicted, we estimate variants of the following reduced-form model:

$$(22) \quad \ln TFP_i = \pi_0 + \pi_1 Ethnic + \pi_2 Language + \pi_3 Religion_i + \psi Control_i + \xi_i.$$

In (22), $\ln TFP$ is the Hall-Jones proxy for the natural log of TFP and ξ_i is the country-specific error term. The Hall-Jones measure for TFP is solved as a labor-augmenting measure of productivity from a Cobb-Douglas production function, given estimates of output per worker, physical capital stock, labor input and years of schooling. Hall and Jones (1999) assume that the relative efficiency of labor is a piecewise linear function of years of schooling and that the capital share is equal to one third.²

The term ψ in (22) is a vector of parameters to be estimated and *Control* is a vector of regressors to control for variables such as institutional quality, population density, trade openness, and measures of mass communication that may influence TFP. If social barriers to communication do affect productivity, as predicted by proposition 2, then we would expect the estimated coefficients for at least some of the fractionalization regressors, especially linguistic fractionalization, to be negative and statistically significant.

IV. Empirical Results

The tests for proposition 1 and 2 are presented separately because they require different data. Our primary focus is on the effects on productivity of social barriers to communication because we hypothesize that it is knowledge links that make labor more productive, which, in turn, induces capital accumulation.

A. Capital Accumulation

Table 2 provides the ordinary least squares (OLS) estimates of (20) and (21) that test proposition 1 using alternative fractionalization indexes to proxy the effects of social barriers to communication. The reported diagnostics include Doornik and Hansen's (1994) χ^2 test for normality of the errors (denoted *Normality*) and an *F*-form of an asymptotic test for heteroskedasticity (denoted *White-Hetero*) based on regressing the squared residuals on the original regressors and their (non-redundant) squares (White 1980). For two of the models estimated for $\Delta \ln KAPW$ the heteroskedasticity test (in columns (5) and (6)) is statistically significant; however, the use of heteroskedasticity-consistent standard errors has little effect on the statistical significance of the coefficients.

In all models, the relevant base-period capital stock measure has a significant negative coefficient at the 5-percent level of significance, or better. Base-period real GDP per worker

has a positive coefficient that is statistically significant at the 10-percent level for the model in column (3) and at the 5-percent level or better in the other models. As hypothesized, the estimated coefficients for *ELF* are negative and statistically different from zero at the 1-percent level of significance in both the human capital and physical capital equations.

To test the robustness of our results to different fractionalization measures, we also include the three Alesina *et al.* (2003) fractionalization indexes as regressors in variants of (20) and (21). The results indicate that ethnic, but not linguistic or religious, fractionalization has a negative coefficient that is statistically significant at the 5-percent level in both the ΔAYS and $\Delta \ln KAPW$ equations. However, the Fearon fractionalization index, *Culture*, which is also a measure of linguistic fractionalization, does have a negative coefficient that is statistically significant in the human capital equation at the 5-percent level. Although measurement of human and physical capital stocks is problematical, our results do support the hypothesis that the larger the economy-wide social barriers to communication, the lower the rate of capital accumulation.

B. Total Factor Productivity: OLS Results

Table 3 provides OLS estimates of variants of equation (22) that tests proposition 2. As far as we are aware, Table 3 provides the first reported test of the effects of the Alesina *et al.* (2003) measures of fractionlization on cross-country TFP.³ In column (1), which includes only the fractionalization measures and no control regressors, the coefficients on *Ethnic* and *Language* have the predicted negative signs and are both statistically significant at the 5-percent level.

In addition to the fractionalization measures, other factors are also likely to influence TFP. Consequently, column (2) gives the results of a model that includes, separately, the two components of Hall and Jones' (1999) social infrastructure index. The two components are

GADP, an index of government antidiversion policies, which incorporates equally-weighted measures of law and order, bureaucratic quality, corruption, risk of expropriation and government repudiation of contracts, and *YrsOpen*, an index of the extent to which countries are open to international trade.⁴ In the model in column (2), the coefficient on *Ethnic* is no longer statistically significant, but the results for *Language* are robust to the addition of these controls. Given our hypotheses about the nature of the transmission of productivity-enhancing ideas, we would expect linguistic differences to be the most important barriers to communication across networks.

Diagnostic tests suggest the presence of heteroskedasticity (with the *White-Hetero* test statistically significant at the 5-percent level for the models in columns (2), (3) and (4)). Heteroskedastic-consistent standard errors are also reported, although these give qualitatively similar results to the conventional standard errors.

Given the hypothesized importance of linguistic barriers, we re-estimated the initial model, but included only a measure of linguistic differences (Fearon's fractionalization index, *Culture*) along with the controls *GADP* and *YrsOpen*. Column (3) reports these results; the coefficient on Fearon's index is negative, as predicted, and statistically significant at the 5-percent level.

The results in columns (4) to (6) provide further evidence on the robustness of the initial results. Studentized residuals and leverage statistics were calculated for the model in column (2) in order to identify potential outliers and/or influential observations.⁵ Column (4) presents the results from re-estimating the model, but with the observations identified by the above statistics removed from the sample, in order to check the sensitivity of the results to the omission of outliers and/or influential observations. While the overall goodness of fit improves and the coefficient on *Language* increases in absolute size, the results are qualitatively unchanged.

To test whether the effects of fractionalization vary between rich and poor countries, we also re-estimated the model in column (2) excluding OECD countries; these results are given in column (5) and are very similar to those in columns (2) and (4). As a further check on the sensitivity of the results, column (6) provides estimates using an alternative measure of $\ln TFP$ obtained from Islam (1995). As predicted by proposition 2, the coefficient for *Language* is negative and statistically significant at the 5-percent level throughout.

C. Total Factor Productivity: Robustness Results

As a check on the robustness of the results in Table 3, we applied a general-to-specific (Gets) algorithm implemented in PcGets (Hendry and Krolzig 2001) to select a preferred model for TFP. The essence of Gets modelling is to start from a general unrestricted model that is ‘congruent’ with the data, i.e., displays no evidence of misspecification. Variables with coefficients that are not statistically significant are eliminated in order to obtain a simpler congruent model that encompasses rival models in the sense that no important information is lost (e.g., Hendry 1995, p. 365).⁶

The general-to-specific approach has been significantly enhanced by Hendry and Krolzig (1999) and Krolzig and Hendry (2001). Their innovations include: examining multiple search paths, considering only model reductions that do not fail diagnostic tests in order to retain congruence, employing ‘pre-search simplification’, using overlapping sub-sample testing to aid in the overall assessment of the ‘reliability’ of the significance of the coefficients, implementing encompassing tests to distinguish between competing candidate congruent models that emerge from different search paths, and using an information criterion to make a final selection if encompassing tests fail to pick a unique dominant final model.⁷

Monte Carlo evidence to date (e.g., Krolzig and Hendry 2001; Hendry and Krolzig 2001, 2004) suggests that the different elements of the overall algorithm combine to give

impressive properties: the size of the model selection process is close to the nominal size of the tests used in the search such that the power approaches that obtained if the process started from the data generating process.⁸ In particular, Hoover and Perez (2004), in a Monte Carlo study designed to reflect the ‘realistic’ setting of cross-country growth regressions, show that a cross-section version of a Gets algorithm outperforms Levine and Renelt’s (1992) and Sala-i-Martin’s (1997) versions of Leamer’s (1983) extreme-bounds approach to model selection.⁹

Table 4, column (1) reports results for the model specified in equation (22) where, in addition to *GADP* and *YrsOpen*, the control variables include measures of mass communication, population density and interaction effects. Given that social barriers to communication impede the exchange of productivity-enhancing ideas, we hypothesize that physical infrastructure that aids in communications may mitigate the negative impact on TFP. We also test whether increased proximity between people, as measured by population density (*Popn Density*) and road density (*Road Density*), reduces the effect of social communication barriers. Interaction effects are included to test the hypothesis that increases in mass communications or population density reduce the negative partial effect of linguistic fractionalization on TFP. Due to the heavily parameterized nature of the model given in column (1) of Table 4, it is not surprising that few of the individual coefficients are statistically significant at conventional levels.¹⁰ Nevertheless, we use this initial model as a starting point for the application of a general-to-specific simplification process.

The results in column (2) of Table 4 are the final specific model selected using the Gets model selection algorithm applied to the model in column (1), Table 4. Two measures of social barriers to communication, *Language* and *Religion*, are selected and have coefficients that are statistically significant at the 1-percent level and have the hypothesized negative sign. One of the measures of mass communications, the number of telephones per capita (*Telephones*), has a coefficient with the hypothesized positive sign that is also statistically

significant at the 1-percent level. Another mass communication measure is included in the selected interaction term *Language*Radios*. Its coefficient is positive and statistically significant at the 5-percent level, implying that the negative effects of linguistic fractionalization are reduced with improvements in mass communication, proxied by the number of radios per capita.

Further robustness tests are provided in columns (3) and (4) in Table 4. Column (3) contains median regression (least absolute errors) estimates for the final selected model to assess the robustness of the results to potential outliers. Point estimates and standard errors based on the design-matrix-bootstrapping estimator (Buchinsky 1998) produce qualitatively similar conclusions to column (2) with the estimated coefficients for linguistic and religious fractionalization both negative and statistically significant at the 1-percent level. Column (4) presents the results of the final model selected from a general-to-specific search applied to a model of the form in column (1) of Table 4, except that Fearon's (2003) *Culture* index replaces the three Alesina *et al.* (2003) measures and the *Language* variable in the interaction terms. Again, the linguistic diversity measure (*Culture*) is selected in the final model and has a negative coefficient that is statistically significant at the 5-percent level. In addition, both the trade openness measure and telephones per capita are also selected in the final model.

Overall, the robustness tests indicate that the estimated coefficients for the linguistic fractionalization indexes have a negative and statistically significant on TFP. These results are consistent with proposition 2, namely, that higher economy-wide social barriers to communication have a negative impact on productivity.

D. Total Factor Productivity: IV Results

A possible concern with the estimates reported in Tables 3 and 4 is that, while it may be reasonable to treat the fractionalization measures as exogenous, several of the controlling

variables may be endogenous. If this is the case, then OLS estimates will be inconsistent. To address this issue, we use instrumental variables that should be uncorrelated with ξ_i in (22), but strongly correlated with the potentially endogenous variables.

Table 5 presents results obtained using instrumental variables (IV) estimation in which all variables other than the fractionalization measures are treated as potentially endogenous. We follow Hall and Jones (1999) in including Frankel and Romer's (1999) (natural log) predicted trade share (based on a trade model including exogenous gravity variables), *lnFraRom*, and the fraction of the population speaking a European language, *EurFrac*, in the instrument set. Hall and Jones (1999) also use distance from the equator as an instrument, but, following Sachs's (2003) argument that this is a poor proxy for geographical factors such as climate, we instead use mean annual temperature, *MeanTemp*, which provides better fits for the first-stage regressions, as well as the proportion of land area within 100km of the coast, *LT100km*, and total land area, *LandArea*. In addition, we include a measure of 'state antiquity', *StateHist*, constructed by Bockstette, Chanda, and Putterman (2002), which their empirical results suggest is a significant predictor of Hall and Jones' (1999) composite social infrastructure measure.¹¹ We also include the interactions between linguistic fractionalization and a subset of the geographical instruments in some of the instrument sets to allow for the endogeneity of interaction terms involving fractionalization and the other right-hand-side variables, such as *Language*Radios*.

Table 5 provides evidence on the suitability of the sets of instruments used. To check on the explanatory power of the instrument sets, the values of R^2 for the first-stage regressions of each right-hand-side endogenous variable on the instruments, including a constant, were calculated. We also calculated p -values of the F -statistics for the joint null hypothesis that the coefficients on all the instruments (including the exogenous regressors) are zero; these are not reported in Table 5 because they are all 0.000. These p -values, reflecting the high R^2

values, indicate that the instrument sets are strongly associated with the endogenous right-hand-side variables.

To check on the correlation between the residuals and the instruments we calculated Sargan's (1964) general misspecification test for instrumental variables estimation of over-identified models. The test statistic, denoted Sargan χ^2 in Table 5, is obtained as NR^2 from the regression of the IV residuals on the set of all instruments and is asymptotically distributed as a central chi-square with degrees of freedom equal to the number of over-identifying restrictions. The hypothesis that the over-identifying instruments are independent of the error terms is not rejected for any of the models. We also report a Hausman test of the consistency (Hausman 1978) of the OLS estimates by comparison with IV based on the selected instrument set(s); under the null that the OLS estimates are consistent, the test is asymptotically distributed as a central chi-square with degrees of freedom equal to the number of potentially endogenous right-hand-side variables. The results imply that OLS estimates are not significantly affected by endogeneity for the models in columns (1) and (3) of Table 5, but are inconsistent when compared to the IV estimates in column (5), using a 5-percent significance level, and more marginally, at the 10-percent significance level, for columns (2) and (4).

Columns (1) and (2) in Table 5 report the IV estimation results for the models corresponding to the OLS estimates in column (2) and (3) in Table 3. Again, both sets of results are consistent with the hypothesis that linguistic fractionalization has a negative impact on TFP. The results presented in column (3) in Table 5 correspond to the model estimated in column (2) of Table 4, i.e., including those variables retained in the final model from the OLS-based general-to-specific selection process. Apart from a reduction in the statistical significance of the coefficient on *Telephones*, the IV results are very similar to those obtained using OLS, an interpretation supported by the non-rejection of the Hausman

test. Column (4) in Table 5 is the final model obtained by commencing with the general model in Table 4, column (1) and applying the general-to-specific simplification, but based throughout on IV estimation, using the specified instrument set, rather than OLS. One component of Hall and Jones' (1999) social infrastructure proxy, *YrsOpen*, and *Road Density* are selected, in place of *Telephones*, but linguistic and religious fractionalization continue to have a significant negative effect.¹² In addition, the role of communications, proxied by *Radios*, in reducing the effect of linguistic fractionalization remains significant through the interaction term.

To illustrate the robustness of the results for the fractionalization and communications variables to the inclusion of social infrastructure proxies, column (5) of Table 5 reports the results obtained by again applying the general-to-specific simplification based on IV estimation commencing from a general model excluding *GADP* and *YrsOpen*. The variables selected are, apart from the excluded *YrsOpen* variable, identical to those in column (4) of Table 5, reinforcing the robustness of these results.

An important feature of both the OLS and IV results is that, despite using an 'agnostic' model selection approach, linguistic fractionalization is consistently selected among the set of relevant explanatory variables. Overall, the empirical results provide strong statistical support for proposition 2.

E. Economic Significance of Total Factor Productivity Results

To assess the economic significance of the effect of social barriers to communication, we carried out a simple simulation. Taking the results from Table 4, column (2) as representative, the coefficients, which being statistically significant at the 5-percent level or better are all relatively precisely estimated, were used to predict the values of $\ln TFP$ for each country and these were transformed into levels. The 110 countries in the sample were then

sorted in ascending order on the basis of their values for *Language*. The means of the predicted values of TFP in levels for the lower and upper quartile countries (defined as the bottom 27 and top 27 countries in terms of the ranking with respect to *Language*) were then calculated.

The ratio of the mean predicted TFP values for the quartile with the lowest measure of linguistic fractionalization, relative to the mean predicted TFP values for the quartile with the highest measure of linguistic fractionalization, is greater than two (2.293). This implies that the effects of social barriers to communication are economically as well as statistically significant in explaining cross-country variation in TFP levels. If taken at face value, and given that all other causal factors between the two sets of countries are accounted for by our model, the results suggest that *if* countries with the highest levels of linguistic fractionalization were to ‘bridge’ the language barriers to the same extent as nations with the lowest levels of fractionalization, they could initiate a very large and positive productivity jump.

Together with other explanatory factors, such as measures of institutional quality and openness to trade, our results provide a plausible explanation for the large disparity in productivity across countries, and why these differences may not necessarily decline over time.

V. Economic Effects of Social Barriers to Communication

Our model emphasizes the *social* dimension of cross-country economic differences rather than simply differences in levels of capital (human and physical). It also explains or supports a number of important stylized facts, and thus goes further than the literature on social cohesion and polarization (Bénabou 1996; Gradstein and Justman 2002), or existing explanations for cross-country differences in TFP (Parente and Prescott 2000).

Our results address three key features of economic performance. One, the on-going high performance of leading industrialized countries; two, the ability of a few countries to initiate ‘catch up’ with economic leaders; and three, the reason why some countries remain growth laggards (Easterly and Levine 2001; Pritchett 1997). To the extent that increased knowledge connections contribute to higher levels of trust and cooperation between individuals, our results also provide a possible explanation for the positive empirical relationship between social capital and human capital accumulation (Glaeser, Laibson, and Sacerdote 2002).

A. High Productivity Performance

We emphasize that diversity across individuals, *per se*, is not detrimental to productivity because differences provide the basis for mutually beneficial exchanges and the ‘cross-fertilization’ of knowledge and ideas — a point made by John Stuart Mill (1848, p. 594) over 150 years ago. Rather, it is the associated higher costs of and barriers to group-to-group communication that act as an impediment to increases in productivity and factor accumulation that diversity would otherwise bring. Indeed, radial, spanning or bridging connections at an individual level, are strongly associated with early adoption of technologies (Valente, 1995, p. 42; Meyer, 1998). Our results support this finding on a national level with evidence that factors that inhibit radial or bridging links, such as linguistic barriers, lower economy-wide productivity.

We speculate that a comparative lack of social barriers to communication may, in part, explain the high productivity of the United States (US), which has a common language and is a multicultural and pluralistic society with a geographically and socially mobile population (Borjas 1992). Thus countries, like the US, that have a common language and a unifying culture can reap the benefits from complementary knowledge sets inherent in different social groups. By contrast, countries that are less socially diverse and mobile than the US, or that

are diverse but have major impediments (social, physical and institutional) to group-to-group communications, may be productivity ‘laggards’ because of less effective radial or bridging links across groups.

B. Productivity ‘Catch Up’

Our modeling offers insights as to how countries might engineer a ‘catch up’ in terms of productivity by fostering approaches that mitigate barriers to communication across social groups. For example, the offering of common national curricula to reduce social distance (Gradstein and Justman 2002), subsidizing citizenship and native language classes for immigrants, promoting a common official language (Lazear 1999), and investing in mass communications (such as internet access and communication links) are all approaches that may raise productivity by reducing the costs of establishing knowledge links across individuals.

To some extent, such measures have been adopted to varying degrees by countries, but without a full recognition of their economic benefits for both productivity and factor accumulation. In sum, national policies could positively influence economic growth provided they lower the social communication costs that impede the creation and diffusion of productivity-enhancing ideas.

C. Stylized Facts

We have explored the concept and consequences of social barriers to communication at an economy level, but our results are also consistent with a number of important findings at a regional and global level. These stylized facts provide additional support for our interpretation as to why social, and especially linguistic, barriers have a negative affect on productivity and factor accumulation.

Lazear (1999), Rauch (2001), Rauch and Trindade (2002), and others, have identified the importance of a common language in trade, but their explanation is that ethnic networks alleviate the difficulties of enforcing contracts, provide information on trade opportunities and help match buyers and sellers. Our work also emphasizes the importance of trade and migration flows, but we stress the importance of the transmission of ideas whereby knowledge connections across social groups provide a basis for productivity gains.

Our thesis that economy-wide productivity is positively affected by knowledge connections across agents also has empirical support in the spillover literature. For instance, Park (2004) finds, using OECD data on cross-country student flows, that the return of foreign-educated workers is an empirically important channel for research and development spillovers. Our findings are also consistent with empirical work by Javorcik (2004); she uses individual firm-level data to show positive productivity spillovers from contacts between domestic firms and their foreign affiliates. Moretti (2004) also finds empirical evidence for human capital spillovers in manufacturing plants within the same city that are increasing in the level of interactions between workers across industries. An individual example of knowledge spillovers of the type we hypothesize is reported by Easterly (2002, pp. 145-148); he describes how a single person played a lead role in developing the garment industry in Bangladesh following the transfer of tacit knowledge via South Korea.

At a local or regional level a number of distinguished thinkers, including Schelling (1978) and Tarde (1895), have observed the tendency for ‘like-with-like’ interactions, known as homophily (Lazarsfeld and Merton 1954, p. 23). This is consistent with our model where establishing knowledge links with different agents is costly. Locations where people ‘connect’ also exemplify how a lowering of the average costs of connecting across agents, equivalent to an increase in the bridging parameter in our model, promotes knowledge transfer and innovation. Thus our model is consistent with the existence of localized

productivity effects, and is supported by empirical evidence of localized and spatial patterns of patents (Jaffe, Trajtenberg, and Henderson 1993; Bottazzi and Peri 2003).

Finally, we emphasize that our hypothesis that increased agent-to-agent connections raises labor productivity and capital accumulation is supported by evidence that cities promote the formation of human capital (Borjas 1995; Glaeser and Maré 2001; Lucas 1988; Marshall 1916, p. 271; Moretti 2004). Our finding of a knowledge connections-augmented rate of return for capital (human or physical) in locations where people have lower ‘connection’ costs, such as in cities, provides an explanation that goes beyond the matching of skilled workers (Kremer 1993) and human capital externalities (Lucas 1988) as to why factors of production agglomerate, and why capital might flow from poor to rich countries (Lucas 1990). In other words, lower social costs of communication in cities, and also in highly-productive countries, generate knowledge spillovers that augment the rate of return to factors in such locations and induce accumulation.

VI. Concluding Remarks

This paper addresses the question: what explains the huge variation in productivity across countries? In an optimal growth model that incorporates social barriers to communication, measured by a ‘bridging’ parameter, we derive dynamic implications for both transitional and steady-state levels of productivity, per-capita consumption and capital. The model generates testable propositions: greater social barriers to communication reduce economy-wide productivity, and also lower transitory and steady-state levels of per-capita consumption and capital.

Theoretical propositions are tested using cross-country data from up to 118 countries. The empirical results obtained from OLS and instrumental variable estimation, and with an extensive set of diagnostic and robustness tests, are statistically and economically significant. These regressions provide strong support for the theoretical result that lower levels of a

‘bridging’ parameter, as measured by linguistic fractionalization, reduce total factor productivity. Some evidence is also found that the effects of social barriers to communication may be mitigated by improvements in mass communications. The empirical findings also show that the greater the initial social barriers to communication, as measured by a base-period ethnolinguistic fractionalization index, the smaller is the increase in the stock of human capital and physical capital.

Our findings and interpretation of the economic effects of social barriers to communication are broadly consistent with a number of stylized facts including the importance of spillovers in research and development, in human capital formation, and in localized productivity effects, and the flow of capital to places where people ‘connect’. The theory and empirical evidence together provide an important explanation for the large cross-country differences in total factor productivity, and also generate fresh insights as to how countries might initiate productivity ‘catch up’.

APPENDIX A: COUNTRIES INCLUDED IN SAMPLE

The following countries are included in the sample for the regressions in Table 4, columns (1) to (3): Algeria, Argentina, Australia, Austria, Bangladesh, Barbados, Belgium, Benin, Bolivia, Botswana, Brazil, Burkina Faso, Burundi, Cameroon, Canada, Central African Republic, Chad, Chile, China, Colombia, Comoros, Congo, Costa Rica, Côte d'Ivoire, Cyprus, Denmark, Dominican Republic, Ecuador, Egypt, Fiji, Finland, France, Gabon, The Gambia, Germany, Greece, Guatemala, Guinea, Guinea-Bissau, Guyana, Honduras, Hong Kong, Hungary, Iceland, India, Indonesia, Iran, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kenya, South Korea, Lesotho, Madagascar, Malawi, Malaysia, Mali, Malta, Mauritania, Mauritius, Mexico, Morocco, Mozambique, Myanmar, Namibia, Netherlands, New Zealand, Nicaragua, Niger, Nigeria, Norway, Oman, Pakistan, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Poland, Portugal, Saudi Arabia, Senegal, Seychelles, Sierra Leone, Singapore, Somalia, South Africa, Spain, Sri Lanka, Sudan, Suriname, Swaziland, Sweden, Switzerland, Syria, Thailand, Togo, Trinidad and Tobago, Tunisia, Turkey, Uganda, United Kingdom, United States of America, Uruguay, Venezuela, Zambia and Zimbabwe.

APPENDIX B: DATA SOURCES AND DEFINITIONS

lnTFP: Hall and Jones measure of total factor productivity (in natural logs) in 1998. Source: Hall and Jones (1999)

Ethnic, Language, Religion: Fractionalization indexes for ethnic, linguistic and religious groups. Source: Alesina *et al.* (2003)

Culture: Cultural fractionalization index accounting for cultural distances between groups based on language. Source: Fearon (2003)

GADP: index of 'government antidiversion policies' calculated as the average of five International Country Risk Guide measures (1985-1995) law and order, bureaucratic quality, corruption, risk of expropriation, government repudiation of contracts, [0-1] range. Source: Hall and Jones (1999)

YrsOpen: Sachs and Warner (1995) index of fraction of years open during 1950 to 1994 period. [0, 1] range. Source: Hall and Jones (1999)

Telephones: Telephone mainlines (per 1,000 people) in 1988. Source: World Bank (2000).

Popn Density: Population density (people per sq km) in 1988. Source: World Bank (2000).

Radios: Radios (per 1,000 people) 1989. Source: World Bank (2000).

Road Density: Roads/Land Area in 1988 or nearest year. Source: Total roads (kms) in 1988, or nearest year, from Canning (1998); Land Area (in sq km) from World Bank (2000).

MeanTemp: Mean annual temperature (degrees Celsius) in 1987. Source: McArthur and Sachs (2001, Appendix)

LT100km: Proportion of land area within 100km of the seacoast. Source: McArthur and Sachs (2001, Appendix)

LandArea: Land area (sq km). Source: World Bank (2000).

EurFrac: Fraction of population speaking a major Western European language: English, French, German, Portuguese, or Spanish. Source: Hall and Jones (1999)

lnFraRom: Natural log of the Frankel-Romer predicted trade share (computed from a gravity model based on population and geography). Source: Hall and Jones (1999)

StateHist: Measures the length and coverage of formal states in current geographical borders from 1 to 1950. Source: *Statehist5* from Bockstette, Chanda, and Putterman (2002)

ELF: Ethnolinguistic Fractionalization – Average value of five different indices (range 0 to 1). Source: La Porta *et al.* (1999, Appendix B).

lnRGDPW60: Real GDP (chain) per worker (1996 international prices) (in natural logs). Source: Penn World Tables 6.1

lnKAPW: Real non-residential capital stock per worker (1985 international prices) (in natural logs). Source: Penn World Tables 5.6

AYS: Average schooling years in the total population (aged 15 years and over). Source: Barro and Lee (2001)

References

- Alesina, Alberto; Devleeschauwer, Arnaud; Easterly, William; Kurlat, Sergio and Wacziarg, Romain. "Fractionalization." *Journal of Economic Growth* 8 (June 2003): 155-194.
- Alesina, Alberto and La Ferrara, Eliana. "Ethnic Diversity and Economic Performance." Working Paper no. 10313. Cambridge, Mass.: National Bureau of Economic Research, 2004.
- Barro, Robert J. and Lee, Jong-Wha. "International Data on Educational Attainment: Updates and Implications." *Oxford Economic Papers* 53(3) (July 2001): 541-563.
- Becker, Gary S. and Murphy, Kevin M. "The Division of Labor, Coordination Costs, and Knowledge." *Quarterly Journal of Economics* 107(4) (1992): 1137-1160.
- Belsley, David A.; Kuh, Edwin and Welsch, Roy E. *Regression diagnostics: identifying influential data and sources of collinearity*. New York: John Wiley, 1980.
- Bénabou, Roland. "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance." *American Economic Review* 86(3) (1996): 584-609.
- Bertrand, Marianne; Luttmer, Erzo F.P. and Mullainathan, Sendhil. "Network Effects and Welfare Cultures." *Quarterly Journal of Economics*, 115(3) (2000): 1019-1056.
- Bockstette, Valerie; Chanda, Areendam and Putterman, Louis "States and Markets: The Advantage of an Early Start." *Journal of Economic Growth* 7(4) (December 2002): 347-369.
- Borjas, George J. "Ethnic Capital and Intergenerational Mobility." *Quarterly Journal of Economics* 107(1) (1992): 123-150.
- Borjas, George J. "Ethnicity, Neighborhoods, and Human Capital Externalities." *American Economic Review* 85(3) (1995): 365-390.
- Bottazzi, Laura and Peri, Giovanni. "Innovation and Spillovers in Regions: Evidence from European Patent Data." *European Economic Review* 47 (2003): 687-710.
- Brock, William A. and Durlauf, Steven N. "Growth Empirics and Reality." *The World Bank Economic Review*, 15(2) (2001): 229-272.
- Brown, John S. and Duguid, Paul. *The social life of information*. Boston, MA: Harvard Business School Press, 2000.
- Buchinsky, Moshe. "Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research." *Journal of Human Resources*, 33(1) (1998): 88-126.
- Burt, Ronald S. "Bridge Decay." *Social Networks* 24 (2002): 333-363.
- Calvó-Armengol, Antoni and Jackson, Mathew O. "The Effects of Social Networks on Employment and Inequality." *American Economic Review* 94 (2004): 426-454.
- Canning, David. "A Database of World Infrastructure Stocks, 1950-95." Policy Research Working Paper no. 1929. Washington DC: The World Bank, 1998.
- Coleman, James S.; Katz, Elihu and Menzel, Herbert. *Medical innovation: a diffusion study*. Indianapolis: The Bobbs-Merrill Company, Inc., 1966.
- Davis, James A. "Clustering and Structural Balance in Graphs." *Human Relations* 20 (1967): 181-187.
- Doornik, Jurgen A. and Hansen, Henrik. "An Omnibus Test for Univariate and Multivariate Normality." Working Paper. Oxford: Nuffield College, Oxford, 1994.
- Durlauf, Steven N. "On the Empirics of Social Capital." *Economic Journal* 112 (2002): F459-F479.
- Easterly, William. *The elusive quest for growth: economists' adventures and misadventures in the tropics*. Cambridge, Mass.: MIT Press, 2002.
- Easterly, William and Levine, Ross. "Africa's Growth Tragedy: Policies and Ethnic Divisions." *Quarterly Journal of Economics* 111(4) (1997): 1203-1250.

- Easterly, William and Levine, Ross. "It's Not Factor Accumulation: Stylized Facts and Growth Models." *World Bank Economic Review* 15(2) (2001): 177-219.
- Encyclopedia Britannica. *Encyclopedia Britannica 2000*. Chicago: Encyclopedia Britannica, 2000.
- Fearon, James D. "Ethnic and Cultural Diversity by Country." *Journal of Economic Growth* 8(2) (June 2003): 195-222.
- Frankel, Jeffrey A. and Romer, David. "Does Trade Cause Growth?" *American Economic Review* 89(3) (June 1999): 379-399.
- Glaeser, Edward L.; Laibson, David and Sacerdote, Bruce. "An Economic Approach to Social Capital." *Economic Journal* 112 (2002): F437-F458.
- Glaeser, Edward L. and Maré, David C. "Cities and Skills." *Journal of Labor Economics* 19(2) (April 2001): 316-342.
- Gradstein, Mark, and Justman, Moshe. "Education, Social Cohesion, and Economic Growth." *American Economic Review* 92(4) (September 2002): 1192-1204.
- Grafton, R. Quentin; Knowles, Stephen and Owen, P. Dorian "Total Factor Productivity, Per Capita Income and Social Divergence." *Economic Record* 80(250) (September 2004): 302-313.
- Granovetter, Mark S. "The Strength of Weak Ties." *American Journal of Sociology* 78(6) (May 1973): 1360-1380.
- Hall, Robert E. and Jones, Charles I. "The Productivity of Nations." Working Paper no. 5812. Cambridge, Mass.: National Bureau of Economic Research, 1996.
- Hall, Robert E. and Jones, Charles I. "Why Do Some Countries Produce so Much More Output Per Worker than Others?" *Quarterly Journal of Economics* 114(1) (February 1999): 83-116.
- Hausman, Jerry A. "Specification Tests in Econometrics." *Econometrica* 46(6) (1978): 1251-1271.
- Helliwell, John F. and Putnam, Robert D. "Economic Growth and Social Capital in Italy." *Eastern Economic Journal* 21(3) (Summer 1995): 295-307.
- Hendry, David F. *Dynamic econometrics*. Oxford: Oxford University Press, 1995.
- Hendry, David F. and Krolzig, Hans-Martin. "Improving on 'Data mining reconsidered' by K.D. Hoover and S.J. Perez." *Econometrics Journal* 2 (1999): 202-219.
- Hendry, David F. and Krolzig, Hans-Martin. *Automatic econometric model selection using PcGets*. London: Timberlake Consultants Ltd, 2001.
- Hendry, David F. and Krolzig, Hans-Martin. "The Properties of Automatic Gets Modelling." *Economic Journal* (2004), forthcoming.
- Hoover, Kevin D. and Perez, Stephen J. "Truth and Robustness in Cross-Country Growth Regressions." *Oxford Bulletin of Economics and Statistics* (2004), forthcoming.
- Islam, Nazrul. "Growth Empirics: A Panel Data Approach." *Quarterly Journal of Economics* 110(4) (1995): 1127-1170.
- Jaffe, Adam B.; Trajtenberg, Manuel and Henderson, Rebecca. "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." *Quarterly Journal of Economics*, 108(3) (1993): 577-598.
- Javorcik, Beata S. "Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers Through Backward Linkages." *American Economic Review* 94 (2004): 605-627.
- Knack, Stephen and Keefer, Phillip. "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation." *Quarterly Journal of Economics* 112(4) (November 1997): 1251-1288.
- Kremer, Michael, O. "The O-ring Theory of Economic Development." *Quarterly Journal of Economics* 108(3) (1993): 551-575.

- Krolzig, Hans-Martin and Hendry, David F. "Computer Automation of General-to-Specific Model Selection Procedures." *Journal of Economic Dynamics and Control* 25 (2001): 831-866.
- La Porta, Rafael; Lopez-de-Silanes, Florencio; Shleifer, Andrei and Vishny, Robert. "The Quality of Government." *Journal of Law, Economics and Organization* 15(1) (1999): 222-282.
- Lazarsfeld, Paul F. and Merton, Robert K. "Friendship as Social Process: A Substantive and Methodological Analysis." In *Freedom and control in modern society*, edited by Monroe Berger, Theodore Abel and Charles H. Page. New York: D. Van Nostrand Company, Inc., 1954: 18-66.
- Lazear, Edward P. "Culture and Language." *Journal of Political Economy* 107(6), Part 2 (December 1999): S95-S126.
- Leamer, Edward E. "Let's Take the Con Out of Econometrics." *American Economic Review* 73(1) (1983): 31-43.
- Levine, Ross and Renelt, David. "A Sensitivity Analysis of Cross-Country Growth Regressions." *American Economic Review* 82(4) (1992): 942-963.
- Lucas, Robert E. Jr. "On the Mechanics of Economic Development." *Journal of Monetary Economics* 22 (1988): 3-42.
- Lucas, Robert E. Jr. "Why Doesn't Capital Flow from Rich to Poor Countries?" *American Economic Review* 80(2) (May 1990): 92-96.
- Marshall, Alfred. *Principles of economics: an introductory volume*. London: Macmillan and Co., Limited, 1916.
- McArthur, John W. and Sachs, Jeffrey D. "Institutions and Geography: Comment on Acemoglu, Johnson and Robinson (2000)." Working Paper no. 8114. Cambridge, Mass.: National Bureau of Economic Research, 2001.
- Meyer, David R. "Formation of Advanced Technology Districts: New England Textile Machinery and Firearms, 1790-1820." *Economic Geography* 74 (1998): 31-45.
- Moretti, Enrico. "Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions." *American Economic Review* 94 (2004): 656-690.
- Mill, John S. *Principles of political economy with some of their applications to social philosophy*, with an introduction by V.W. Bladen, Toronto: University of Toronto Press, Routledge and Kegan Paul, (first published in 1848), 1965.
- Nettle, Daniel. "Linguistic Fragmentation and the Wealth of Nations: The Fishman-Pool Hypothesis Reexamined." *Economic Development and Cultural Change* 49(2) (January 2000): 335-348.
- Parente, Stephen L. and Prescott, Edward C. *Barriers to riches (Walras-Pareto lectures)*. Cambridge, Mass.: MIT Press, 2000.
- Park, Jungsoo. "International Student Flows and R&D Spillovers." *Economics Letters* 82 (2004): 315-320.
- Penn World Tables. Philadelphia: Center for International Comparisons, University of Pennsylvania. http://pwt.econ.upenn.edu/php_site/pwt_index.php
- Powell, Walter W. "Neither Market nor Hierarchy: Network Forms of Organization." *Research in Organizational Behavior* 12 (1990): 295-336.
- Pritchett, Lant. "Divergence, Big Time." *Journal of Economic Perspectives* 11(3) (Summer 1997): 3-17.
- Rauch, James E. "Business and Social Networks in International Trade." *Journal of Economic Literature* 39 (December 2001): 1177-1203.
- Rauch, James E. and Trindade, Vitor. "Ethnic Chinese Networks in International Trade." *Review of Economics and Statistics* 84 (February 2002): 116-130.

- Rivera-Batiz, Luis A. and Romer, Paul M. "Economic Integration and Endogenous Growth." *Quarterly Journal of Economics* 106(2) (May 1991): 531-555.
- Rogers, Everett M. *Diffusion of innovations*. Fourth Edition, New York: The Free Press, 1995.
- Ryan, Bryce and Gross, Neal C. "The Diffusion of Hybrid Seed Corn in Two Iowa Communities." *Rural Sociology* 8(1) (1943): pp. 15-24.
- Sachs, Jeffrey D. "Institutions Don't Rule: Direct Effects of Geography on Per Capita Income." Working Paper no. 9490. Cambridge, Mass.: National Bureau of Economic Research, 2003.
- Sachs, Jeffrey D. and Warner, Andrew M. "Economic Reform and the Process of Global Integration." *Brookings Papers on Economic Activity* 1 (Spring 1995): 1-95.
- Sala-i-Martin, Xavier X. "I Have Just Run Two Million Regressions." *American Economic Review* 87(2) (May 1997): 178-183.
- Sargan, J. Denis. "Wages and Prices in the United Kingdom: A Study in Econometric Methodology." In *Econometric analysis for national economic planning*, edited by Peter E. Hart; G. Mills and John K. Whitaker. London: Butterworths, 1964: 25-63.
- Saxenian, Annalee. *Regional advantage: culture and competition in Silicon Valley and Route 128*. Cambridge, Mass.: Harvard University Press, 1994.
- Schelling, Thomas C. *Micromotives and macrobehavior*. New York: W.W. Norton and Company, 1978.
- Sherif, Muzafer; Harvey, O.J.; White, B. Jack; Hood, William R. and Sherif, Carolyn W. *Intergroup conflict and cooperation: the Robbers Cave experiment*. Norman, Oklahoma: Institute of Group Relations, University of Oklahoma, 1961.
- Solo, Robert A. *Economic organizations and social systems*. Indianapolis: The Bobs-Merrill Company, Inc., 1967.
- Tarde, Gabriel. *Les lois de l'imitation*. Second edition, Félix Alan: Paris, 1895.
- Temple, Jonathan and Johnson, Paul A. "Social Capability and Economic Growth." *Quarterly Journal of Economics* 113(3) (August 1998): 965-990.
- Valente, Thomas W. *Network models of the diffusion of innovations*. Cresskill, New Jersey: Hampton Press, Inc., 1995.
- White, Halbert. "A Heteroskedastic-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica* 48(4) (May 1980): 817-838.
- World Bank. *World development indicators on CD-ROM*. Washington DC: The World Bank, 2000.
- Zak, Paul J. and Knack, Stephen. "Trust and Growth." *Economic Journal* 111 (2001): 295-321.

TABLE 1 – SUMMARY STATISTICS FOR KEY VARIABLES

	<i>N</i>	Mean	Standard Deviation	Minimum	Maximum
<i>lnTFP</i>	110	7.9570	0.7195	6.2845	9.0154
<i>Ethnic</i>	110	0.4424	0.2763	0.0000	0.9302
<i>Language</i>	110	0.3771	0.3028	0.0021	0.9227
<i>Religion</i>	110	0.4217	0.2500	0.0028	0.8603
<i>Culture</i>	106	0.2951	0.2156	0.0000	0.7330
<i>GADP</i>	110	0.6167	0.1958	0.3080	1.0000
<i>YrsOpen</i>	110	0.3581	0.3453	0.0000	1.0000
<i>Telephones</i>	110	128.19	176.86	0.6224	663.94
<i>Popn Density</i>	110	189.00	680.06	1.5527	5683.4
<i>Radios</i>	110	379.25	344.95	0.2517	2119.3
<i>Road Density</i>	110	0.5450	0.9323	0.0043	4.7438
<i>ELF</i>	82	0.3451	0.2975	0.0000	0.8902
ΔAYS	82	2.7558	1.2827	−0.8050	6.5910
<i>AYS60</i>	82	3.8318	2.4689	0.1160	9.7260
$\Delta \ln KAPW$	57	0.8840	0.6248	−0.5495	3.0909
<i>lnKAPW65</i>	57	8.3797	1.3194	4.6347	10.536
<i>lnRGDPW60</i>	82	8.8213	0.9268	6.5403	10.376

N is the number of observations. *N* = 110 corresponds to the sample used in Table 4, columns (1)–(3), *N* = 106 to Table 4, column (4), *N* = 82 to Table 2, column (1), and *N* = 57 to Table 2, column (4).

TABLE 2 – CHANGES IN CAPITAL STOCKS AND SOCIAL BARRIERS TO COMMUNICATION

Dependent variable	(1) ΔAYS	(2) ΔAYS	(3) ΔAYS	(4) $\Delta \ln KAPW$	(5) $\Delta \ln KAPW$	(6) $\Delta \ln KAPW$
Constant	-1.586 (2.097)	-1.842 (1.865)	0.062 (2.051)	1.955 (0.965)	0.837 (0.948)	2.309 (0.924)
<i>Ethnic</i>	-1.672 (0.741)			-0.843 (0.355)		
<i>Language</i>	0.160 (0.654)			-0.166 (0.342)		
<i>Religion</i>	0.283 (0.615)			-0.034 (0.312)		
<i>Culture</i>		-1.168 (0.697)			-0.504 (0.428)	
<i>ELF</i>			-1.263 (0.566)			-1.187 (0.340)
<i>AYS60</i>	-0.286 (0.091)	-0.242 (0.082)	-0.204 (0.082)			
<i>lnKAPW60</i>				-0.382 (0.105)	-0.374 (0.114)	-0.426 (0.101)
<i>lnRGDPW60</i>	0.678 (0.247)	0.664 (0.227)	0.443 (0.243)	0.278 (0.158)	0.366 (0.171)	0.272 (0.149)
Diagnostics						
R^2	0.185	0.168	0.158	0.301	0.198	0.331
Regression SE	1.208	1.214	1.200	0.547	0.590	0.521
N	82	79	82	57	54	56
<i>Normality</i>	0.527	2.279	2.048	3.516	3.014	5.565
[p -value]	[0.769]	[0.320]	[0.359]	[0.172]	[0.222]	[0.062]
<i>White-Hetero</i>	1.043	1.064	1.659	1.520	2.843	2.470
[p -value]	[0.418]	[0.392]	[0.143]	[0.163]	[0.019]	[0.036]

Notes: Standard errors are in parentheses and p -values for diagnostic tests in square brackets. *Normality* is the Doornik-Hansen test of normal errors and *White-Hetero* is White's test for heteroskedasticity.

TABLE 3 – DETERMINANTS OF TFP: OLS RESULTS

Dependent variable: $\ln TFP$	(1)	(2)	(3)	(4)	(5)	(6)
Constant	8.533 (0.138) [0.142]	7.206 (0.260) [0.300]	7.237 (0.213) [0.183]	7.297 (0.273) [0.243]	7.342 (0.369) [0.423]	5.905 (0.300) [0.278]
<i>Ethnic</i>	–0.755 (0.301) [0.274]	0.182 (0.283) [0.290]		0.311 (0.311) [0.283]	0.148 (0.335) [0.341]	–0.233 (0.311) [0.339]
<i>Language</i>	–0.567 (0.278) [0.251]	–0.532 (0.229) [0.211]		–0.763 (0.244) [0.259]	–0.560 (0.260) [0.231]	–0.652 (0.251) [0.297]
<i>Religion</i>	–0.087 (0.254) [0.279]	–0.417 (0.220) [0.223]		–0.465 (0.211) [0.229]	–0.502 (0.272) [0.281]	–0.070 (0.270) [0.260]
<i>Culture</i>			–0.618 (0.244) [0.213]			
<i>GADP</i>		1.310 (0.395) [0.364]	0.952 (0.353) [0.280]	1.273 (0.407) [0.367]	1.190 (0.603) [0.614]	2.293 (0.463) [0.432]
<i>YrsOpen</i>		0.644 (0.206) [0.189]	0.853 (0.199) [0.180]	0.655 (0.201) [0.200]	0.588 (0.235) [0.206]	0.672 (0.231) [0.225]
Diagnostics						
R^2	0.243	0.494	0.470	0.575	0.336	0.722
Regression SE	0.644	0.531	0.527	0.479	0.578	0.516
N	118	118	113	108	96	88
<i>Normality</i>	7.936 [0.019]	2.467 [0.291]	0.037 [0.982]	2.365 [0.307]	1.920 [0.383]	0.679 [0.712]
<i>White-Hetero</i>	1.351 [0.241]	2.739 [0.005]	5.073 [0.0001]	3.285 [0.001]	1.714 [0.091]	0.919 [0.521]

Notes: Conventional standard errors are in parentheses and heteroskedastic-consistent standard errors in square brackets. *Normality* is the Doornik-Hansen test of normal errors and *White-Hetero* is White's test for heteroskedasticity. The sample used in column (4) omits influential observations and/or outliers, and in column (5) omits OECD countries. In column (6) the dependent variable is Islam's (1995) measure of $\ln TFP$.

TABLE 4 – DETERMINANTS OF TFP: ROBUSTNESS RESULTS

Dependent variable: $\ln TFP$	(1)	(2)	(3)	(4)
Constant	8.079 (0.502)	8.072 (0.118)	8.292 (0.152)	7.706 (0.107)
<i>Ethnic</i>	0.122 (0.305)			
<i>Language</i>	-1.331 (0.908)	-0.755 (0.219)	-0.981 (0.311)	
<i>Religion</i>	-0.501 (0.258)	-0.507 (0.217)	-0.705 (0.328)	
<i>Culture</i>				-0.570 (0.245)
<i>GADP</i>	-0.171 (0.922)			
<i>YrsOpen</i>	0.206 (0.393)			0.722 (0.203)
<i>Telephones</i>	0.002 (0.001)	0.002 (0.0004)	0.001 (0.0006)	0.001 (0.0004)
<i>Popn Density</i>	0.00001 (0.00003)			
<i>Radios</i>	-0.0001 (0.0005)			
<i>Road Density</i>	0.017 (0.117)			
<i>Language*Telephones</i>	-0.002 (0.003)			
<i>Language*Radios</i>	0.002 (0.002)	0.002 (0.0006)	0.002 (0.0007)	
<i>Language*Popn Density</i>	0.0003 (0.001)			
<i>Language*Road Density</i>	-0.025 (0.407)			
<i>Language*GADP</i>	1.032 (1.859)			
<i>Language*YrsOpen</i>	0.103 (0.887)			
Diagnostics				
R^2	0.533	0.509	0.490	0.464
Regression SE	0.530	0.514	0.528	0.524
N	110	110	110	106
<i>Normality</i>	3.273	3.867		0.067
[p -value]	[0.195]	[0.145]		[0.967]
<i>White-Hetero</i>	1.524	1.412		1.639
[p -value]	[0.071]	[0.114]		[0.049]

Notes: Standard errors are given in parentheses and p -values for diagnostic tests in square brackets. Results in columns (1), (2) and (4) are obtained using OLS. Results in column (3) are median regression estimates.

TABLE 5 – DETERMINANTS OF TFP: IV RESULTS

Dependent variable: $\ln TFP$	(1)	(2)	(3)	(4)	(5)
Constant	7.735 (0.529)	7.798 (0.501)	8.228 (0.132)	8.003 (0.203)	8.308 (0.127)
<i>Ethnic</i>	-0.001 (0.373)				
<i>Language</i>	-0.540 (0.268)		-1.056 (0.272)	-0.838 (0.267)	-1.142 (0.219)
<i>Religion</i>	-0.299 (0.333)		-0.558 (0.237)	-0.525 (0.268)	-0.750 (0.246)
<i>Culture</i>		-0.800 (0.326)			
<i>GADP</i>	0.112 (1.136)	-0.468 (1.078)			
<i>YrsOpen</i>	1.479 (0.657)	1.999 (0.687)		0.871 (0.458)	
<i>Telephones</i>			0.001 (0.0006)		
<i>Road Density</i>				0.147 (0.158)	0.362 (0.113)
<i>Language*Radios</i>			0.002 (0.0009)	0.002 (0.0009)	0.003 (0.0007)
Diagnostics					
R^2	0.476	0.376	0.523	0.536	0.511
Regression SE	0.544	0.610	0.506	0.499	0.510
N	91	91	99	88	88
Sargan χ^2	1.887	2.216	5.374	2.556	6.084
[p -value]	[0.596]	[0.529]	[0.146]	[0.923]	[0.638]
Hausman χ^2	3.386	5.819	1.990	6.849	6.612
[p -value]	[0.184]	[0.055]	[0.370]	[0.077]	[0.037]
R^2 for first-stage regressions					
<i>GADP</i>	0.741	0.667			
<i>YrsOpen</i>	0.512	0.522		0.571	
<i>Telephones</i>			0.703		
<i>Road Density</i>				0.475	
<i>Language*Radios</i>			0.663	0.730	

Notes: Asymptotic standard errors are given in parentheses and p -values in square brackets. R^2 for IV regressions is calculated as the squared correlation- between the observed and predicted values of the dependent variable. Sargan χ^2 is Sargan's misspecification test for IV estimation and Hausman χ^2 is a test for the consistency of the corresponding OLS estimates.

Instrument sets: Column (1): *Ethnic*, *Language*, *Religion*, *MeanTemp*, *LT100km*, *StatHist*, *EurFrac*, *lnFraRom*; Column (2): *Culture*, *MeanTemp*, *LT100km*, *StatHist*, *EurFrac*, *lnFraRom*; Column (3): *Language*, *Religion*, *Meantemp*, *LT100km* and the interaction of *MeanTemp*, *LT100km* and *LandArea* with *Language*; Columns (4) and (5): *Ethnic*, *Language*, *Religion*, *StatHist*, *EurFrac*, *lnFraRom*, *MeanTemp*, *LT100km*, *LandArea* and the interaction of each of the last three variables with *language*.

Endnotes:

* The authors are especially grateful for helpful discussions with Steve Dowrick, Stephen Knowles and Warwick McKibbin in the preparation of this paper. We have also benefited from the comments of David Fielding, Kevin Fox, Tue Gorgens, Daniel Léonard, Jack Pezzey, Richard Pomfret, Ludovic Renou and Rhema Vaithianathan. We thank participants at seminars/presentations at the Centre for Applied Economic Research (CAER), The University of New South Wales in November 2002, School of Economics, The University of Queensland in May 2003, Department of Finance, Ottawa, Canada in July 2003, Conference of the New Zealand Association of Economists in July 2003, Research School of Pacific and Asian Studies, The Australian National University in July 2003, Australian Conference of Economists in Canberra in October 2003, Econometric Society Australasian Meeting in July 2004, Department of Economics, The University of Auckland in August 2004, and School of Economics, University of Adelaide in October 2004. Grafton and Owen acknowledge financial support from the Marsden Fund, administered by the Royal Society of New Zealand, and the research assistance of Clayton Weatherston. Kompas acknowledges the financial support of the Australian Research Council. Grafton and Kompas acknowledge the research assistance of Tuong Nhu Che and Tuan Ghee Yew.

¹ Substitution of (9) and (10) into (8) also allows us to derive the transition path for s .

² Hall and Jones note that their estimates are very similar to those obtained in Hall and Jones (1996) where "...the production function is not restricted to Cobb-Douglas, and factor shares are allowed to vary across countries" (Hall and Jones 1999, p. 93).

³ Alesina *et al.* (2003) include their three measures of fractionalization as regressors in models where the regressand is per capita income or various quality-of-government indexes, but do not test for the effects on TFP.

⁴ Given the way the components are measured, high values of $GADP$ are conducive to supporting production.

⁵ The cut off values used were 2 for the studentized residuals and $2k/N$ for the leverage statistics (Belsley, Kuh, and Welsch 1980).

⁶ The diagnostic tests implemented in the search algorithm were the *Normality* and *White-Hetero* tests, discussed above, plus F -tests for parameter constancy for breakpoints at the sample mid-point and 90th percentile. For the diagnostic tests, a 1-percent significance level was used throughout to help control the overall null-rejection probability, as suggested by the Monte Carlo evidence in Krolzig and Hendry (2001).

⁷ A complete listing of the PcGets algorithm is available in Hendry and Krolzig (2001, Appendix A1) or Krolzig and Hendry (2001, Tables 1 and 2).

⁸ In this context, power and size relate to the probabilities of retaining in the final model variables that are, respectively, included and not included in the data generating process.

⁹ Use of a general-to-specific modeling approach also helps address the issue of model uncertainty (Brock and Durlauf 2001; Durlauf 2002).

¹⁰ Excluding the constant, only the coefficient on *Religion* is statistically significant at the 10-percent level (on a two-tailed test), with the coefficients on *Language* and *Telephones* significant at the 15-percent level.

¹¹ This index rates the territory of the current geographical boundaries of a country in terms of whether the government is above tribal level, is colonial or locally based, and the territorial coverage of the government for 50 year sub-periods from 0 to 1950. A single observation for each country is obtained by discounting the effect of past values. We use the preferred measure of Bocksette, Chanda, and Putterman (2002) corresponding to a discount rate of five percent.

¹² Note that although *Road Density* is retained in the final model in column (4), it is not statistically significant at conventional significance levels; at each stage of the simplification process, the Gets algorithm retains variables whose exclusion would lead to lack of congruence (as judged by significant values for any of the diagnostic tests).