J Prod Anal
DOI 10.1007/s11123-007-0054-6

Bridging the barriers: knowledge connections, productivity and capital accumulation

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

Abstract The paper contributes to the explanation of the large differences in cross-country productivity performance by modelling and testing the effects of social barriers to communication on productivity and capital accumulation. In an optimal growth model, social barriers to communication, which 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. Empirical testing yields a robust and theoretically consistent result: linguistic barriers to communication reduce productivity and capital accumulation. The findings provide an explanation for cross-country differences in TFP, and fresh insights into how productivity 'catch up' may be initiated.

Keywords Knowledge connections · Productivity · Economic growth

JEL Classifications O41 · C61 · C21

R. Q. Grafton (✉) · T. Kompas
Crawford School of Economics and Government, The Australian National University, Canberra ACT 0200, Australia
e-mail: quentin.grafton@anu.edu.au

T. Kompas
e-mail: tom.kompas@anu.edu.au

P. D. Owen
Department of Economics, University of Otago, Dunedin, New Zealand
e-mail: DOwen@business.otago.ac.nz

1 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 and trade openness. 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.

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 paper is organized as follows. Section 2 solves an optimal growth model of the effects of social barriers to communication and analyzes the implications for TFP and capital accumulation. Section 3 describes the data, outlines the empirical models used to test the theoretical propositions and reports the empirical results. Section 4 reports checks on the robustness and economic significance of the empirical results. Concluding remarks are offered in Sect. 5.

2 Knowledge connections and social barriers to communication

Our modeling focuses on the macroeconomic effects of social barriers to communication on productivity and growth. Our work has similarities to contributions by Lazear (1999), Nettle (2000), Rauch (2001) and Grafton et al. (2002, 2004), among others, that emphasize the importance of diversity for, respectively, exchange and trade between individuals, aggregate per capita GDP, international trade, and per capita income and productivity. It also relates to studies that 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 capital accumulation and productivity.

Using an optimal growth model, we posit that aggregate 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 et al. 1966; Marshall 1890; 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 et al. 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 et al. 1961; Solo 1967).

2.1 The model

To capture the effects of social barriers to communication we assume that a representative agent's utility function, given by Eq. 1, depends positively on per capita consumption at time t, $c(t)$, and negatively on the effort required to establish knowledge connections across agents given by $\varepsilon(s(t))$, i.e.,

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

In Eq. 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 knowledge connections has a lower bound of zero.

Equation 1 is consistent with an intertemporal consumption/leisure model of individual preferences (Che et al. 2001) where the negative effect of $\varepsilon(s(t))$ on utility incorporates an implicit trade-off between leisure and forming knowledge connections such that the time spent making 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

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

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 Eq. 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

$$y = \alpha_0 s^{\alpha_1} k^{\alpha_2} \tag{3}$$

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

$$\dot{k} = y - c \tag{4}$$

2.2 Theoretical results

To solve the optimization problem we maximize utility in Eq. 1 subject to Eq. 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, Eq. 4 and the following necessary conditions must be satisfied for all t:

$$c^{-\theta} = \lambda \tag{5}$$

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

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

Eq. 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, higher 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 the 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 Eqs. 5 and 6 that, together, imply

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

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. Eq. 8, along with the necessary conditions, can be used to derive the following transition paths:[1]

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

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

At the steady state, because

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

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

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

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

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

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

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

and

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

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

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, which, 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 $\alpha_0^{\frac{z}{z-\alpha_1}} \alpha_1^{\frac{\alpha_1}{z-\alpha_1}} \beta^{\frac{\alpha_1}{z-\alpha_1}} k^{\frac{z\alpha_2}{z-\alpha_1}-1} c^{\frac{-\theta\alpha_1}{z-\alpha_1}} \alpha_2$ in Eq. 9. A higher rate of return on capital induces factor accumulation and raises the steady-state values of both per capita consumption and capital. This result is important because, 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 Eq. 12, or the expression for per-capita consumption as a function of steady-state reproducible capital, into Eq. 8, the derived expression for the number of knowledge connections, to obtain

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

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

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

where A is TFP and derived to be

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

An intertemporal version of TFP can also be derived showing that the time path for productivity is increasing in β. This result, and Eq. 18, yields 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 results provide a causal explanation for cross-country differences in TFP not found in the existing literature, and also imply that policy actions that can overcome social barriers to forming knowledge connections can initiate productivity 'catch up'. Both our propositions can be tested using cross-country data and measures of TFP, physical and human capital, social barriers to communication, and other variables.

3 Tests of the propositions

Social barriers to communication are proxied by measures of ethnic (*Ethnic*), linguistic (*Language*) and religious (*Religion*) fractionalization for the early to mid 1990s, calculated by Alesina et al. (2003). Lower levels of the economy-wide bridging parameter are represented by higher levels of fractionalization, especially linguistic fractionalization. Each fractionalization measure represents the probability of two randomly selected individuals being from a different social group, i.e.,

$$FRAC_i = 1 - \sum_j^n f_{ji}^2 \tag{19}$$

where f_{ji} is the share of (linguistic, ethnic or religious) group j in country i. The three fractionalization measures we use have been investigated by Alesina et al. (2003) and Alesina and La Ferrara (2005) as possible determinants of long-run growth. However, our paper is the first to examine the effects of these measures on productivity and physical and human capital. Our approach also differs from earlier empirical work in that we include variables, such as social infrastructure (Hall and Jones 1999) and measures of mass communication, that may mitigate the effects of social barriers to communication on productivity.

In addition, we provide estimates using Fearon's (2003) cultural fractionalization measure, *Culture*, which is based on the structural distance between languages.[2] The fractionalization indexes reflect the number and relative sizes

[2] 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).

of distinct social groups within a country. Cross-country summary statistics of the fractionalization measures and other key variables are provided in Table 1.

3.1 Proposition 1

Proposition 1 implies that the higher are social barriers to communication (the 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:

$$HCW_i = \delta_0 + \delta_1 FRAC_i + \delta_2 GADP_i + \delta_3 YrsOpen + \mu_i, \tag{20}$$

$$\ln\, KAPW_i = \gamma_0 + \gamma_1 FRAC_i + \gamma_2 GADP_i + \gamma_3 YrsOpen_i + v_i. \tag{21}$$

HCW represents human capital per worker. Our primary measure is Barro and Lee's (2001) estimate of the average years of schooling in the total population aged 15 years, either averaged over the period 1960–1999 (denoted *AYS*), or its value in 1999 (*YS*99). For consistency with the implied human capital measure underlying our chosen estimate of TFP, we also examine Hall and Jones' human capital measure, ln*HL*, which is based on a Mincerian-style piecewise linear function of years of schooling. ln*KAPW* represents the natural log of real physical capital stock per worker, either averaged over the period 1965–1990 (ln*KAV*) or its value in 1990 (ln*K*90).[3] 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*, a base-period measure of the social barriers to communication, may be more appropriate when the regressand is the average value of the human or physical capital stock over a long period.

In the spirit of the literature on the fundamental determinants of cross-country income levels (Hall and Jones 1999; Acemoglu et al. 2001; Rodrik et al. 2004) we include as controls the components of Hall and Jones' (1999) social infrastructure index: *GADP*, an index of government antidiversion policies (incorporating 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. Higher values of *GADP* indicate better social infrastructure while higher levels of *YrsOpen* represent greater openness to international trade. μ_i and v_i are country-specific error terms.

Table 2 provides ordinary least squares (OLS) estimates of Eqs. 20 and 21, using alternative fractionalization indexes to proxy the effects of social barriers to communication. For each equation, a battery of diagnostic tests was performed to test for normality of the errors, heteroskedasticity and functional form misspecification. There is evidence of heteroskedasticity for the models in columns (5)–(7), so we report heteroskedasticity-consistent standard errors, although conventional standard errors give similar results.

Consistent with proposition 1, the estimated coefficients for *Language*, *Culture* and *ELF* are negative and statistically significant at the 5-percent level or better in both the human capital and physical capital equations, regardless of the human capital measure used and whether the capital stocks are averaged over a period or are for a selected recent year. In addition, in all models, *GADP* has a positive coefficient that is statistically significant at the 5-percent level or better; *YrsOpen* also has a positive coefficient but this is not statistically significant in all models. Although measurement of human and physical capital stocks is problematical, our results provide support for the hypothesis that the larger the economy-wide social barriers to communication, the lower are levels of physical and human capital.

3.2 Proposition 2

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. To test whether higher social barriers to communication (lower β) have a negative effect on TFP, we estimate variants of:

$$\ln TFP_i = \pi_0 + \pi_1 Ethnic + \pi_2 Language + \pi_3 Religion_i + \psi\, Control_i + \xi_i. \tag{22}$$

Our main proxy for ln*TFP* is Hall and Jones' (1999) estimate, which is solved as a labor-augmenting measure of productivity from a Cobb–Douglas production function,

[3] Following a referee's suggestion, we report results for *levels* regressions for the capital stock variables, for consistency with our TFP levels results. However, in an earlier, working-paper version (Grafton et al. 2004), we report cross-sectional results for capital accumulation equations, in which the dependent variables are long-period differences in the capital stocks (for both human and physical capital over, respectively, 39 and 25 years). Base-period fractionalization measures have statistically significant negative effects on both types of capital accumulation.

Table 1 Summary statistics for key variables

	N	Mean	Standard deviation	Minimum	Maximum
ln*TFP*	110	7.9570	0.7195	6.2845	9.0154
Ethnic	110	0.4424	0.2763	0.0000	0.9302
Language	110	0.3771	0.3028	0.0021	0.9227
Religion	110	0.4217	0.2500	0.0028	0.8603
Culture	106	0.2951	0.2156	0.0000	0.7330
GADP	110	0.6167	0.1958	0.3080	1.0000
YrsOpen	110	0.3581	0.3453	0.0000	1.0000
Telephones	110	128.19	176.86	0.6224	663.94
Popn Density	110	189.00	680.06	1.5527	5683.4
Radios	110	379.25	344.95	0.2517	2119.3
Road Density	110	0.5450	0.9323	0.0043	4.7438
ELF	91	0.3311	0.2968	0.0000	0.8902
*YS*99	97	6.3631	2.8220	0.8390	12.049
AYS	91	5.1088	2.6187	0.5239	10.876
ln*HL*	118	0.5938	0.2884	0.0724	1.2147
ln*K*90	62	9.2039	1.3329	5.4072	11.204
ln*KAV*	57	8.9853	1.2928	5.0956	10.866

Notes: *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* = 91 to Table 2, column (3), *N* = 97 to Table 2, column (1), *N* = 118 to Table 2, column (4), *N* = 62 to Table 2, column (5), and *N* = 57 to Table 2, column (7)

given estimates of output per worker, physical capital stock, labor input and years of schooling.[4] As a check on the sensitivity of our results we also examine Islam's (1995) estimates of TFP.

Control is a vector of regressors to control for variables such as institutional quality, trade openness, population density, and measures of mass communication that may influence TFP, ψ is its associated vector of parameters, and ξ_i is a country-specific error term. If social barriers to communication inhibit the transmission of productivity-enhancing ideas, then we would expect the estimated coefficients for at least some of the fractionalization regressors, especially linguistic fractionalization, to be negative and statistically significant.[5]

Table 3 provides OLS estimates for variants of Eq. 22. 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. Column (2) gives the results of a model that includes *GADP* and *YrsOpen*. The coefficient on *Ethnic* is no longer statistically significant, but the results for *Language* are robust to the addition of these controls. Diagnostic tests suggest the presence of heteroskedasticity for the models in columns (2), (3) and (4). Heteroskedastic-consistent standard errors are therefore reported for these models, although these give qualitatively similar results to the conventional standard errors.

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

The results in columns (4)–(6) provide some evidence on the robustness of the initial results. Column (4) presents the results from re-estimating the model in column (2), but removing observations identified by studentized residuals and leverage statistics as outliers and/or influential observations.[6] The overall goodness of fit improves and the coefficient on *Language* increases in absolute size, but 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; the

[4] Hall and Jones (1999) assume that the relative efficiency of labor is a piecewise linear function of years of schooling, with decreasing rates of return to additional education, and that the capital share is equal to one third. They 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).

[5] Given the timing of the TFP estimates, we use the more recent Alesina et al. and Fearon fractionalization measures in the TFP equations rather than the base-period fractionalization measure.

[6] The cut off values used were 2 for the studentized residuals and $2k/N$ for the leverage statistics (Belsley et al. 1980).

Table 2 Capital stocks and social barriers to communication

Dependent variable	(1) *YS99*	(2) *YS99*	(3) *AYS*	(4) ln*HL*	(5) ln*K90*	(6) ln*K90*	(7) ln*KAV*
Constant	2.072 (0.876)	1.239 (0.710)	0.230 (0.666)	0.094 (0.082)	6.842 [0.453]	6.796 [0.450]	7.563 [0.426]
Ethnic	–0.958 (0.995)			–0.006 (0.089)	0.353 [0.505]		
Language	–1.775 (0.803)			–0.246 (0.072)	–1.283 [0.489]		
Religion	0.686 (0.754)			0.074 (0.069)	-0.876 [0.447]		
Culture		–2.178 (0.828)				–1.259 [0.608]	
ELF			–1.711 (0.584)				–2.024 [0.429]
GADP	7.226 (1.290)	8.248 (1.176)	7.934 (1.052)	0.862 (0.124)	3.878 [0.615]	3.126 [0.689]	2.206 [0.614]
YrsOpen	1.382 (0.661)	1.520 (0.692)	0.935 (0.607)	0.110 (0.065)	0.651 [0.361]	1.186 [0.493]	0.851 [0.419]
R^2	0.684	0.671	0.696	0.679	0.696	0.628	0.700
Regression SE	1.629	1.646	1.468	0.167	0.768	0.842	0.727
N	97	96	91	118	62	59	57

Notes: Conventional standard errors are in parentheses and heteroskedastic-consistent standard errors in square brackets

results, in column (5), are similar to those in columns (2) and (4). Column (6) provides estimates using an alternative measure of ln*TFP* obtained from Islam (1995). The coefficient for *Language* remains negative and statistically significant at the 5-percent level throughout.

4 Total factor productivity: robustness results and economic significance

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.[7] Table 4, column (1) reports results for the model specified in Eq. 22 with, in addition to *GADP* and *YrsOpen*, measures of mass communication, population density and interaction effects included as controls. 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, it is not surprising that few of the individual coefficients are statistically significant at conventional levels.[8] Nevertheless, we can use this as a starting point for the application of a Gets simplification process.

Estimates for the final specific model obtained using the Gets model selection algorithm are reported in column (2) of Table 4. Two measures of social barriers to communication, *Language* and *Religion*, and a measure of mass communication, the number of telephones per capita (*Telephones*), are selected and have coefficients that are statistically significant at the 1-percent level, with the expected signs. Also selected is the interaction term *Language**Radios*. Its coefficient is positive, implying that the negative effects of linguistic fractionalization are reduced with improvements in mass communication.

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). Column (4) presents the results of the final model selected from a Gets search applied to a model of the form in column (1) of Table 4, except that Fearon's (2003) *Culture* index replaces

[7] 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). A detailed discussion of the steps in the PcGets algorithm is available in Hendry and Krolzig (2001, 2005). Monte Carlo evidence to date (e.g., Krolzig and Hendry 2001; Hendry and Krolzig 1999, 2001, 2005; Hoover and Perez 2004) suggests that the algorithm has impressive properties: model selection is consistent, the size of the overall model selection process is close to the nominal size of the tests used in the search, and power approaches that obtained if commencing from the data generating process.

[8] 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-pecent level.

Table 3 Determinants of TFP: OLS results

Dependent variable: ln*TFP*	(1)	(2)	(3)	(4)	(5)	(6)
Constant	8.533 (0.138)	7.206 [0.300]	7.237 [0.183]	7.297 [0.243]	7.342 (0.369)	5.905 (0.300)
Ethnic	–0.755 (0.301)	0.182 [0.290]		0.311 [0.283]	0.148 (0.335)	–0.233 (0.311)
Language	–0.567 (0.278)	–0.532 [0.211]		–0.763 [0.259]	–0.560 (0.260)	–0.652 (0.251)
Religion	–0.087 (0.254)	–0.417 [0.223]		–0.465 [0.229]	–0.502 (0.272)	–0.070 (0.270)
Culture			–0.618 [0.213]			
GADP		1.310 [0.364]	0.952 [0.280]	1.273 [0.367]	1.190 (0.603)	2.293 (0.463)
YrsOpen		0.644 [0.189]	0.853 [0.180]	0.655 [0.200]	0.588 (0.235)	0.672 (0.231)
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

Notes: Conventional standard errors are in parentheses and heteroskedastic-consistent standard errors in square brackets. 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*

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.

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 control variables may be endogenous, so that OLS estimates could be inconsistent. To address this issue, 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), ln *FraRom*, 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 Sachs (2003) argues that this is a poor proxy for geographical factors such as climate, so 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 100 km of the coast, *LT100 km*, and total land area, *LandArea*. In addition, we include a measure of 'state antiquity', *StateHist*, constructed by Bockstette et al. (2002), which their results suggest is a significant predictor of Hall and Jones' (1999) social infrastructure measure.[9] We also include 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*.

The values of the partial R^2 for the first-stage regressions reported in Table 5 suggest that the instrument sets are reasonably strongly associated with the endogenous right-hand-side variables.[10] Also, the hypothesis that the over-identifying instruments are independent of the error terms is not rejected for any of the models using Sargan's (1964) general misspecification test for instrumental variables estimation of over-identified models.[11] 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).[12] 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).

[9] 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 et al. (2002) corresponding to a discount rate of five percent.

[10] These represent the correlations between the dependent variable and the additional instruments after partialling out the correlations with the exogenous regressors, which are also included in the instrument sets but essentially act as instruments for themselves. If most of the explanatory power of the first-stage regression is due to the exogenous regressors in the instrument set, then the partial R^2 will be low even though the overall R^2 may be high.

[11] 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.

[12] Under the null that OLS estimates are consistent, the test statistic is asymptotically distributed as a central chi-square with degrees of freedom equal to the number of potentially endogenous right-hand-side variables.

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)
Ethnic	0.122 (0.305)			
Language	–1.331 (0.908)	–0.755 (0.219)	–0.981 (0.311)	
Religion	–0.501 (0.258)	–0.507 (0.217)	–0.705 (0.328)	
Culture				–0.570 (0.245)
GADP	–0.171 (0.922)			
YrsOpen	0.206 (0.393)			0.722 (0.203)
Telephones	0.002 (0.001)	0.002 (0.0004)	0.001 (0.0006)	0.001 (0.0004)
Popn Density	0.00001 (0.00003)			
Radios	–0.0001 (0.0005)			
Road Density	0.017 (0.117)			
*Language*Telephones*	–0.002 (0.003)			
*Language*Radios*	0.002 (0.002)	0.002 (0.0006)	0.002 (0.0007)	
*Language*Popn Density*	0.0003 (0.001)			
*Language*Road Density*	–0.025 (0.407)			
*Language*GADP*	1.032 (1.859)			
*Language*YrsOpen*	0.103 (0.887)			
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

Notes: Standard errors are given in parentheses. Results in columns (1), (2) and (4) are obtained using OLS. Results in column (3) are median regression estimates

Columns (1) and (2) in Table 5 report the IV estimation results for the models corresponding to the OLS estimates in columns (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 Gets selection process. Apart from a reduction in the statistical significance of the coefficient on *Telephones*, the IV results are 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 Gets simplification, but based throughout on IV estimation, using the specified instrument set, rather than OLS. *YrsOpen*, and *Road Density* are selected, in place of *Telephones*, but linguistic and religious fractionalization continue to have a significant negative effect.[13] 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 Gets 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' Gets 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, i.e., higher economy-wide social barriers to communication have a negative impact on productivity.

To assess the economic significance of the effect of social barriers to communication, we carried out a simple simulation. Taking the representative results from Table 4, column (2), 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

[13] 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).

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)
Ethnic	–0.001 (0.373)				
Language	–0.540 (0.268)		–1.056 (0.272)	–0.838 (0.267)	–1.142 (0.219)
Religion	–0.299 (0.333)		–0.558 (0.237)	–0.525 (0.268)	–0.750 (0.246)
Culture		–0.800 (0.326)			
GADP	0.112 (1.136)	–0.468 (1.078)			
YrsOpen	1.479 (0.657)	1.999 (0.687)		0.871 (0.458)	
Telephones			0.001 (0.0006)		
Road Density				0.147 (0.158)	0.362 (0.113)
*Language*Radios*			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 [*P*-value]	1.887 [0.596]	2.216 [0.529]	5.374 [0.146]	2.556 [0.923]	6.084 [0.638]
Hausman χ^2 [*P*-value]	3.386 [0.184]	5.819 [0.055]	1.990 [0.370]	6.849 [0.077]	6.612 [0.037]
Partial R^2 for first-stage regressions					
GADP	0.579	0.648			
YrsOpen	0.401	0.516		0.515	
Telephones			0.634		
Road Density				0.437	
*Language*Radios*			0.593	0.673	

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*, ln*FraRom*; Column (2): *Culture*, *MeanTemp*, *LT100km*, *StatHist*, *EurFrac*, ln*FraRom*; 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*, ln*FraRom*, *MeanTemp*, *LT100km*, *LandArea* and the interaction of each of the last three variables with *Language*

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. At face value, and assuming 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.

5 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, 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. The empirical results obtained from OLS and IV estimation, which include 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 proxied by higher linguistic fractionalization, reduce total factor productivity and stocks of human capital and physical capital. Some evidence is found that the effects of social barriers to communication may be mitigated by improvements in mass communications.

Our results provide a potentially important explanation for the large cross-country differences in total factor productivity, and also 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. 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 within and also across countries.

Acknowledgments 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 two anonymous referees, Kevin Fox, Tue Gorgens, Jack Pezzey, Rhema Vaithianathan, and participants at various seminars and conferences. 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 and the research assistance of Tuong Nhu Che.

Appendix Data sources and definitions

ln*TFP*: 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 100 km 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)

ln*FraRom*: 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 et al. (2002)

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

ln*HL*: Human capital per worker (in natural logs). Source: Hall and Jones (1999)

ln*KAV*: Real non-residential capital stock per worker (1985 international prices) (in natural logs) averaged over the period 1965–1990. Source: Penn World Tables 5.6

ln*K*90: Real non-residential capital stock per worker (1985 international prices) (in natural logs) in 1990. Source: Penn World Tables 5.6

AYS: Average schooling years in the total population (aged 15 years and over) averaged over the period 1960–1999. Source: Barro and Lee (2001)

*YS*99: Average schooling years in the total population (aged 15 years and over) in 1999. Source: Barro and Lee (2001)

References

Acemoglu D, Johnson S, Robinson JA (2001) The colonial origins of comparative development: an empirical investigation. Am Econ Rev 91:1369–1401

Alesina A, Devleeschauwer A, Easterly W, Kurlat S, Wacziarg R (2003) Fractionalization. J Econ Growth 8:155–194

Alesina A, La Ferrara E (2005) Ethnic diversity and economic performance. J Econ Lit 43:762–800

Barro RJ, Lee J-W (2001) International data on educational attainment: updates and implications. Oxford Econ Papers 53:541–563

Becker GS, Murphy KM (1992) The division of labor, coordination costs, and knowledge. Quart J Econ 107:1137–1160
Belsley DA, Kuh E, Welsch RE (1980) Regression diagnostics: identifying influential data and sources of collinearity. John Wiley, New York
Bénabou R (1996) Heterogeneity, stratification, and growth: macroeconomic implications of community structure and school finance. Am Econ Rev 86:584–609
Bertrand M, Luttmer EFP, Mullainathan S (2000) Network effects and welfare cultures. Quart J Econ 115:1019–1056
Bockstette V, Chanda A, Putterman L (2002) States and markets: the advantage of an early start. J Econ Growth 7:347–369
Borjas GJ (1992) Ethnic capital and intergenerational mobility. Quart J Econ 107:123–150
Borjas GJ (1995) Ethnicity, neighborhoods, and human capital externalities. Am Econ Rev 85:365–390
Brown JS, Duguid P (2000) The social life of information. Harvard Business School Press, Boston MA
Buchinsky M (1998) Recent advances in quantile regression models: a practical guideline for empirical research. J Hum Res 33:88–126
Burt RS (2002) Bridge decay. Soc Networks 24:333–363
Calvó-Armengol A, Jackson MO (2004) The effects of social networks on employment and inequality. Am Econ Rev 94:426–454
Canning D (1998) A database of world infrastructure stocks, 1950–95. World Bank Policy Research Working Paper No. 1929
Che TN, Kompas T, Vousden N (2001) Incentives and static and dynamic gains from market reform: rice production in Vietnam. Aust J Agr Res Econ 45:547–572
Coleman JS, Katz E, Menzel H (1966) Medical innovation: a diffusion study. Bobbs-Merrill, Indianapolis IN
Davis JA (1967) Clustering and structural balance in graphs. Human Rels 20:181–187
Easterly W, Levine R (1997) Africa's growth tragedy: policies and ethnic divisions. Quart J Econ 111:1203–1250
Fearon JD (2003) Ethnic and cultural diversity by country. J Econ Growth 8:195–222
Frankel JA, Romer D (1999) Does trade cause growth? Am Econ Rev 89:379–399
Gradstein M, Justman M (2002) Education, social cohesion, and economic growth. Am Econ Rev 92:1192–1204
Grafton RQ, Knowles S, Owen PD (2002) Social divergence and productivity: making a connection. In: Sharpe A, St-Hilaire F, Banting K (eds) The Review of Economic Performance and Social Progress 2002: towards a social understanding of productivity. Institute for Research on Public Policy, and Centre for the Study of Living Standards. Montreal, Ottawa, pp 203–223
Grafton RQ, Knowles S, Owen, PD (2004) Total factor productivity, per capita income and social divergence. Econ Record 80:302–313
Grafton RQ, Kompas T, Owen PD (2004) Bridging the barriers: knowledge connections, productivity, and capital accumulation. Australian National University, Asia Pacific School of Economics and Government, International and Development Economics Working Papers IDEC04–5
Granovetter MS (1973) The strength of weak ties. Am J Sociol 78:1360–1380
Hall RE, Jones CI (1996) The productivity of nations. NBER Working Paper No. 5812
Hall RE, Jones CI (1999) Why do some countries produce so much more output per worker than others? Quart J Econ 114:83–116
Hausman JA (1978) Specification tests in econometrics. Econometrica 46:1251–1271
Helliwell JF, Putnam RD (1995) Economic growth and social capital in Italy. Eastern Econ J 21:295–307
Hendry DF (1995) Dynamic econometrics. Oxford University Press, Oxford
Hendry DF, Krolzig H-M (1999) Improving on 'data mining reconsidered' by K. D. Hoover and S. J. Perez. Economet J 2:202–219
Hendry DF, Krolzig H-M (2001) Automatic econometric model selection using PcGets. Timberlake Consultants Ltd, London
Hendry DF, Krolzig H-M (2005) The properties of automatic Gets modelling. Econ J 115:C32-C61
Hoover KD, Perez SJ (2004) Truth and robustness in cross-country growth regressions. Oxford Bull Econ Stat 66:765–798
Islam N (1995) Growth empirics: a panel data approach. Quart J Econ 110:1127–1170
Knack S, Keefer P (1997) Does social capital have an economic payoff? A cross-country investigation. Quart J Econ 112:1251–1288
Krolzig H-M, Hendry DF (2001) Computer automation of general-to-specific model selection procedures. J Econ Dyn Control 25:831–866
La Porta R, Lopez-de-Silanes F, Shleifer A, Vishny R (1999) The quality of government. J Law Econ Organ 15:222–282
Lazarsfeld PF, Merton RK (1954) Friendship as social process: a substantive and methodological analysis. In: Berger M, Abel T, Page CH (eds) Freedom and control in modern society. D. Van Nostrand, New York, NY, pp 18–66
Lazear EP (1999) Culture and language. J Pol Econ 107: S95-S126
Lucas RE Jr (1988) On the mechanics of economic development. J Monetary Econ 22:3–42
Marshall A ([1890], 1916) Principles of economics: an introductory volume. Macmillan, London
McArthur JW, Sachs JD (2001) Institutions and geography: comment on Acemoglu, Johnson and Robinson (2000). NBER Working Paper No. 8114
Meyer DR (1998) Formation of advanced technology districts: New England textile machinery and firearms, 1790–1820. Econ Geog 74:31–45
Nettle D (2000) Linguistic fragmentation and the wealth of nations: the Fishman-Pool hypothesis reexamined. Econ Dev Cult Change 49:335–348
Penn World Tables. Center for International Comparisons, University of Pennsylvania, http://pwt.econ.upenn.edu/php_site/pwt_index.php
Powell WW (1990) Neither market nor hierarchy: network forms of organization. Res Organ Behav 12:295–336
Rauch JE (2001) Business and social networks in international trade. J Econ Lit 39:1177–1203
Rodrik D, Subramanian A, Trebbi F (2004) Institutions rule: the primacy of institutions over geography and integration in economic development. J Econ Growth 9:131–165
Rivera-Batiz LA, Romer PM (1991) Economic integration and endogenous growth. Quart J Econ 106:531–555
Rogers EM (1995) Diffusion of innovations, 4th edn. The Free Press, New York, NY
Ryan B, Gross NC (1943) The diffusion of hybrid seed corn in two Iowa communities. Rural Sociol 8:15–24
Sachs JD (2003) Institutions don't rule: direct effects of geography on per capita income. NBER Working Paper No. 9490
Sachs JD, Warner AM (1995) Economic reform and the process of global integration. Brookings Pap Econ Act 1:1–95
Sargan JD (1964) Wages and prices in the United Kingdom: a study in econometric methodology. In: Hart PE, Mills G, Whitaker JK (eds) Econometric analysis for national economic planning. Butterworths, London, pp 25–63

Saxenian A (1994) Regional advantage: culture and competition in Silicon Valley and Route 128. Harvard University Press, Cambridge, MA

Schelling TC (1978) Micromotives and macrobehavior. W.W. Norton, New York, NY

Sherif M, Harvey OJ, White BJ, Hood WR, Sherif CW (1961) Intergroup conflict and cooperation: the Robbers Cave experiment. Institute of Group Relations, University of Oklahoma, Norman, OK

Solo RA (1967) Economic organizations and social systems. Bobs-Merrill, Indianapolis, IN

Tarde G (1895) Les lois de l'imitation, 2nd edn. Félix Alan, Paris

Temple J, Johnson PA (1998) Social capability and economic growth. Quart J Econ 113:965–990

Valente TW (1995) Network models of the diffusion of innovations. Hampton Press, Cresskill, NJ

World Bank (2000) World development indicators on CD-ROM. The World Bank, Washington, DC

Zak PJ, Knack S (2001) Trust and growth. Econ J 111:295–321