NBER WORKING PAPER SERIES

# GLOBALIZATION AND THE MARKET FOR TEAMMATES

Edward P. Lazear

Working Paper 6579
http://www.nber.org/papers/w6579

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 1998

This research was supported in part by the National Science Foundation through a grant to the National Bureau of Economic Research. Able research assistance by Joseph Guzman and Muhamet Yildiz is gratefully acknowledged. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

Globalization and the Market for Teammates
Edward P. Lazear
NBER Working Paper No. 6579
May 1998


## ABSTRACT

The globalization of firms is explored at theoretical and empirical levels. The idea is that a global firm is a multi-cultural team. The existence of a global firm is somewhat puzzling. Combining workers who have different cultures, legal systems, and languages imposes costs on the firm that would not be present were all workers to conform to one standard. In order to offset the costs of cross-cultural dealing, there must be complementarities between the workers that are sufficiently important to overcome the costs. Disjoint and relevant skills create an environment where the gains from complementarities can be significant. It is also necessary that teammates be able to communicate with one another. The search for the "best practice" is analyzed and empirical support from an examination of trading patterns is provided.



Edward P. Lazear
Graduate School of Business
Stanford University
Stanford, CA 94305-5015
and NBER
lazear_edward@gsb.stanford.edu

It is impossible to pick up a business publication these days without reading about the wonders of teamwork. In the same publication, the reader is likely to come across a discussion of globalization. The two topics are often discussed in the same article: Once teamwork is accepted as a basic business principle, it is not much of a stretch to think about teams that are comprised of diverse individuals, coming from different countries and cultures.

This essay focuses on the apparent growth in globalization that has occurred over recent years. Why do firms become global? When workers from different countries are employed in the same organization, difficulties arise that would not be present in the absence of international mixing. Different laws, different languages and different cultures must be integrated into the same firm. Of course, the ability to take advantage of cheap sources of labor might motivate a firm to produce internationally. But there is no necessity to have the foreign labor part of the same firm. Nothing prevents a firm in one country from using an impersonal market to buy factors and intermediate goods from a cheap source of supply. The various parts of the supply chain need not be linked formally or even informally through some long term implicit relationship.

The global firm can be thought of as a team where members come from different cultures or countries. Global firms face costs of translation, which include costs of transacting across borders. Teams must select members and compensate them accordingly. What kinds of members do they

select? How does the market establish pay for members with different skills, particularly skills that allow them to act as liaisons? Is it better for firms to hire employees who are similar or employees who are different and how does the choice vary with the circumstance?

The existence of a global firm is somewhat puzzling.[1] Combining workers who have different cultures, legal systems, and languages imposes costs on the firm that would not be present were all workers to conform to one standard. In order to offset the costs of cross-cultural dealing, there must be complementarities between the workers that are sufficiently important to overcome the costs. Disjoint and relevant skills create an environment where the gains from complementarities can be significant.

There are gains from having a firm that is comprised of diverse individuals because skills and knowledge sets may be culture-specific. Three factors determine the gains from putting together diverse teams. The gains from diversity are greatest when groups have information sets that are disjoint,[2] that are relevant to one another, and that can be learned by the other group at low cost.[3]

---

[1]No firm is truly global, so the model is one of multinational firms that have a presence in more than one country, rather than universal firms.

[2]Hong and Page (1997) focus on the gains from diversity that come about when different agents, each of whom possess limited ability, work collectively. Lazear (1998a) applies this to the choice of an immigrant group. Kole and MacDonald (1997) consider interactions between customers and suppliers to argue for a business reason for diversity. But O'Reilly, Williams, Barsade (1997) find that the gains from diversity are in fact negative. Because diversity creates conflict, any creativity gains are swamped by those associated with the conflict itself.

[3]An informal presentation of these ideas is in Lazear(1998b), pp. 310-15

A more formal model will be presented below, but the intuition can be stated verbally.

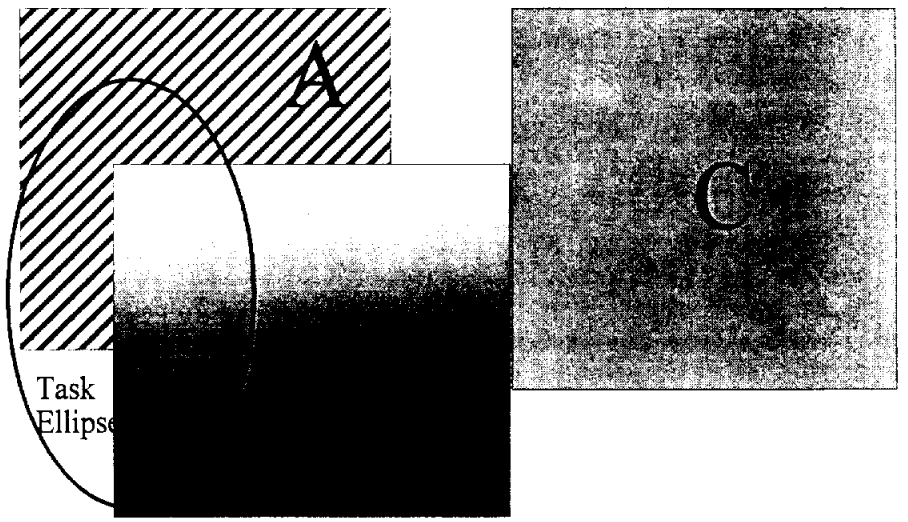
First, the diversity gains are greatest when individuals have different information. If information or skill sets are completely disjoint, then group A can benefit from working with group B, and vice versa. If information and skill sets are completely overlapping, then the two groups do not contribute much to each other's knowledge. The issue here involves complementarity. Disjointness is important when complementarities are important, which leads to the second point.

Skills or information possessed by the other group must be relevant. For example, the knowledge that an auto mechanic has is quite different from that held by an economist. The information sets are quite distinct and thereby meet the disjointness criterion. But they are not relevant to one another. Knowing how to repair the differential[4] on a 1963 Buick in unlikely to help an economist analyze wage differentials.

Figure 1 illustrates the point. Consider three individuals, A, B, and C. The information sets of each of the three are represented by the rectangles, A, B and C, respectively. Each point on the diagram maps into a specific piece of information. Suppose that a firm would like to put together a two person team. On the basis of disjointness alone, it would seem best to put A and C or B and C together, because C has information that is disjoint from that of either A or B. A team of A and B has considerable overlap of information so there is more potential from gain in an AC or a BC team than in an AB team. But relevance matters. Suppose that the skills required to perform the tasks needed by the firm are represented by the oval, marked "task ellipse." Then, the best team consists of A and B, and excludes C. Although C knows many things that neither A nor B know,

[4]A differential is a device that allows two wheels on the same axis to spin at different rates.

Figure 1




C's information is irrelevant for performing the tasks. The team consisting of A and B has almost all the knowledge necessary to perform the tasks needed by the firm.

In addition to knowledge, communication is necessary. It is important that A and B be able to communicate with each other in order to perform the relevant tasks. Thus, the third point is that even if skill and information sets are disjoint and relevant, they are useless unless they can be understood by the other group. For example, it might be better to express a particular thought in French than it is in English, but in order for English speakers to get the benefit of this improvement, they must be able to understand French themselves. If it were prohibitively costly to learn the language or obtain the information possessed by the other group, then disjointness and relevance would have no value. Similarly, an economist who tries to communicate with a sociologist might

encounter difficulties because the jargon in the two fields and level of mathematical expertise is quite different.

Even if there are gains to having a diverse team, the market need not price all types of labor similarly. Most important in this context is that *A*s may not have the same wage as *B*s, and bilingual *A*s may earn more than monolingual *A*s. In designing the optimal team, it is necessary to take the different prices into account.[5]

The first part of the paper discusses the general theory of choosing and rewarding teammates. The last part extends the theory to look at the global corporation and at returns to being able to act as liaison between cultures. The approach is to view the analysis of the global organization as an application of the theory of choosing teammates. "Going global" means choosing teammates who are from a different business culture or language group than the members of the initial organization. The main points are:

1. The primary reason for employing inter-cultural teams is that the value from putting together disjoint and relevant information swamps the communication costs incurred.

2. The search for "best practices" is an example of how firms gain by constructing multi-cultural teams. Such teams create the most added value when cultures are less positively correlated with one another.

3. There is a bias toward lopsided firms, where one culture dominates the firm. This results even in a world with complete symmetry and may have nothing to do with social preferences or chauvinism.

---

[5]The problem of sorting and marriage (see, e.g., Becker (1991) and Pollak (1986 and 1987) is related to the one considered here.

4. Firms in and from rich countries employ more bilingual members of poor countries than firms in and from poor countries employ bilinguals from rich countries.

5. The costs of communication show up as a wage premium that bilingual individuals receive over those who can speak only one language. In countries with low levels of education, the sectoral wage differences reflect the premium to being bicultural, which tends to be higher than in countries with high levels of education.

6. Somewhat limited evidence suggests that globalization has increased over time. It is also true that countries trade with those that are more relevant and that can communicate with. Over time, as the world has gone to English as the language of business, these trading patterns have become less pronounced.

## Theory

Why have a global firm at all? There are clear communication costs so one can only argue in favor of the global firm if the benefits from complementarity of selecting teammates across countries outweigh the costs. It is useful to start with a disclaimer. Since we do not have a good theory of the firm, it is unlikely that this paper will present a good theory of the global firm. Instead, the discussion will focus more on relationships. The boundaries of a firm have posed problems for economists for a long time. Thus, the term "global firm" refers to personal interaction rather than impersonal market interaction through a few translators.

What is a global firm? There is some ambiguity in the definition, but I want to think of a global firm as having employees in more than one country, and even better, in many countries. A wheat farmer in Kansas whose wheat is used to bake bread in Moscow is not considered a global firm. The intermediary agency who arranges for transportation of the wheat from Kansas to Moscow is a global firm, but the farmer is not.

It is not the legal limits of the firm that are important here, but rather the personal interaction that takes place. For example, students of haute couture who live in New York might benefit from having direct personal contact with a French designer. Whether the designer is actually an employee of the school or a sub-contractor is inessential. What is crucial is that direct interaction occurs. And this means that the designer must speak English or the students must speak French. With this in mind, the formal model is presented.

**Production**

Let there be two types, *A* and *B*. The production function is

(1) Output = $f(x_1, x_2)$

where $x_1$ and $x_2$ are skills. Suppose that *A*s are identical and each has one unit of $x_1$. *B*s are also identical and each B has one unit of skill $x_i$, i=1 or 2, but not both. Thus, either all *B*s have skill 1 or they all have skill 2, to be considered below.

In order for *A* s to work with *B* s effectively they must be able to communicate. Suppose that *A* s and *B* s speak different languages. As mentioned earlier, "language" is defined loosely to refer to jargon, rules or other aspects of the business culture that are shared by all *A* s and by all *B* s but are different between the groups *A* and *B*. In order for *A*s and *B*s to work together, either *B*s must speak *A* or *A*s must speak *B*.

The first part of table 1 lists output under a variety of circumstances. The number of monolingual *A* s and *B* s employed is given by A and B, respectively. The number of bilingual *A* s and *B* s employed is given by A* and B*, respectively. Rows 1 through 4 consider the case where *B* s have the same skill as *A* s. Put differently, the skills of *A* s and *B* s are completely overlapping or non-disjoint. This is the perfect substitutes situation. When *B* possesses skill $x_1$, it never pays to use both *A* s and *B* s. The firm will specialize in the type of labor that is cheapest. If *A* s and *B* s had identical wage structures, the firm would still specialize in one type or the other as long as bilingual labor earned a premium over monolingual labor. Since there is no value to having two types of labor, there is no reason for the firm to choose to pay the bilingual premium. In row 4, the firm must choose to use one type or the other because monolingual *A* s cannot work with monolingual *B*s.

Next consider the case where the *B*s' skills are disjoint from the *A*s'. Rows 4 through 8 of

the table depict this situation. Now there is value to combining $A$ s and $B$ s as long as $f_{12} > 0$. To the extent that $A$ is complementary with $B$, it pays to have some of each type of worker in the firm. But unlike the standard production problem, there is an added wrinkle here. In order to use $A$ s and $B$s

Table 1

| Row # | *B*'s skill | Communication | Output | First-Order Condition |
|---|---|---|---|---|
| *A* s and *B* s have same skill | | | | |
| 1 | $x_1$ | *A* s monolingual; *B* s bilingual | f(A+B*,0) | $f_1 = w_i$ where i is the cheapest type of labor. The other type of labor is set equal to zero. |
| 2 | " | *A* s bilingual; *B* monolingual | f(A*+B,0) | " |
| 3 | " | Both bilingual | f(A*+B*,0) | " |
| 4 | " | Both monolingual | f( A,0 ) or f(B,0) | $f_1(A,0)=w_A$ or $f_1(B,0)=w_B$ |
| *A* s and *B* s have different skill | | | | |
| 5 | $x_2$ | *A* s monolingual; *B* s bilingual. Firm speaks *A* | f(A,B*) | $f_1(A,B^*)=w_A$ ; $f_2(A,B^*)=w_B^*$ |
| 6 | " | *A* s bilingual; *B* monolingual. Firm speaks *B* | f(A*,B) | $f_1(A^*,B)=w_A^*$; $f_2(A^*,B)=w_B$ |
| 7 | " | Both bilingual. Firm speaks both | f(A*,B*) | $f_1(A^*,B^*)=w_A^*$; $f_2(A^*,B^*)=w_B^*$ |
| 8 | " | Both monolingual. | f(A,0) or f(0,B) | $f_1(A,0) = w_A$ or $f_2(0, B) = w_B$ |

together, they must be able to communicate with one another. Communication costs, which may be interpreted as transaction costs across borders, play a key role in the model.

When *A* s and *B* s speak different languages, the firm must hire either all bilingual *A* s who can communicate with each other and the *B* s, bilingual *B* s, who can communicate with each other and the *A* s, or hire exclusively bilingual workers, which introduces some redundancy. Each of these choices results in identical output levels, namely f(A,B). If neither group of workers is bilingual, as shown in row 8, then none of the complementarities can be enjoyed and the firm must choose to use one type or the other. Since they cannot work together, the labor type that is not chosen is wasted. This results in output f(A,0) or f(0.B), depending on the choice.

It is now possible to present the first result:

**Output is non-decreasing in the disjointness of the workforce.**

To see this, note that the value of having Y members, all of whom are *A*s is f(Y,0). The output of Y members split between *A* s and *B* s when *B* s have skill $x_1$ and (at least) one group is bilingual is also f(Y.0) because A+B=Y. The value of having a diverse workforce is the difference between the two, which is obviously zero. There is no value to having a diverse workforce when skills are completely overlapping.

When *B* s have skill $x_2$, then, output in the diverse firm is f(A, Y-A), assuming that *A* s and *B* s can communicate with one another. Output in a homogeneous firm with the same number of employees is f(Y,0) or f(0,Y). If f(A, Y-A) > max { f(Y,0), f(0,Y) }, then the gains to diversity are

larger skills are non-overlapping. If $f(A, Y-A) < \max \{ f(Y,0), f(0,Y) \}$, the firm has the option of using only one type of labor, in which case the gains to diversity are zero. But the gains were zero when types had overlapping skills. Thus, disjointness increases or at worst, leaves the same, the value of having a diverse workforce.

The second result follows immediately from this discussion:

**Relevance increases the gains from diversity where relevance is interpreted as gains from complementarity.**

The difference between the output in row 1 and row 5 (or 2 and 6 or 3 and 7) equals $f(A^*, B) - f(A+B,0)$, which is increasing in $f_{12}$ .

The third result follows directly from the assumption about production:

**Without communication, there can be no gains from diversity.**

If *A*s and *B*s have the same skills, then output in the diverse firm without communication is given by row 4. Thus, if Y workers are hired, the highest level of output is

$\max \{ f(A,0) , f(Y-A, 0) \}$.

This is maximized by setting A=Y, i.e., hiring only *A*s. There is no gain to diversity in this case.

Additionally, if *A*s and *B*s had different skills, then output in the diverse firm is given by row 8, again by

$\max \{ f(A,0) , f(0, Y-A) \}$ .

This is maximized by setting A equal to Y or to zero. The firm specializes either in *A*s or *B*s.

The production technology assumed in table 1 emphasizes the importance of three factors: disjointness (*B* s and *A* s have different skills or knowledge), relevance (their skills are complementary), and communication (some members of the firm are "bilingual").

**The Practical Value of Multi-Cultural Teams**

Table 1 makes assumptions about the nature of production. Once it is assumed that $x_1$ and $x_2$ are complementary skills, that *A*s know $x_1$ and that *B*s know $x_2$, it is quite obvious that there may be gains to creating teams of *A*s and *B*s, even when this involves costly communication between *As* and *Bs*. It is helpful, however, to make the discussion concrete. Are there real world circumstances where different cultures do have disjoint and complementary skills that can be combined to create a whole that is greater than the sum of its parts?

Two cases come to mind. The first is what might be called "knowing the ropes." This is probably the most common and straightforward use of an individual from a different culture. For example, when a European oil company wants to do business in Kazakhstan, it is useful to have as a team member someone who knows the customs, laws, formal and informal, and people in Kazakhstan to grease the skids. It is very different for, say, a British expatriate to establish the same ties and acquire the same information as that held by a native of the country in which the oil extraction is occurring. The local knows the ropes.

A somewhat less obvious case involves the quest for "best practices." It is common to hear managers talk about the search for the "best practice." British Petroleum, for example, uses the

slogan "Somewhere in the world, someone is doing it better," with the implication that it is up to managers to discover the practice and put it to use. At Ford, outside London, a manager relayed a story of using Brazilian engineers to design seals, because Brazilians were accustomed to dealing with dusty environments where seals were particularly important. The idea that individuals from other cultures may know something that locals do not is a statistical proposition. It relies on the notion that different cultures provide different distributions of skills and knowledge and that the correlation the skills of two individuals drawn from the same country is likely to be larger than the correlation between two individuals drawn from different countries.

To see this, consider choosing two random variables, $\epsilon$ and $\eta$. The goal is to get the best practice, which can be represented as finding $\max[\epsilon, \eta]$. The firm is hiring two workers, each of whom has knowledge or skill drawn from a particular distribution. The value of the first worker's skill is represented by a scalar (thought of as profit) which is equated with $\epsilon$ and that of the second worker is $\eta$. The firm can either hire two *A* s, two *B*s, or one *A* and one *B*.

To get the intuition of the diversity point, consider an extreme case. Suppose that all *As* are identical and all *B*s are identical, but *A* s may differ from *B* s. Since the information that *A* s have is perfectly correlated across *A* s, it is better to choose one *A* and one *B* if the goal is to maximize the expected value of the highest order statistic, i.e., the value of $\max[\epsilon,\eta]$. Choosing a second *A* has no value since the value of $\eta$ will be identical to that of $\epsilon$. By choosing a *B*, there is some chance that $\eta$ will exceed $\epsilon$, thereby improving the best practice. This is true even if *Bs* are on average much less skilled than *A*s so that $E(\eta) < E(\epsilon)$, where $\eta$ now reflects the draw on a *B*. As long as the A distribution does not lie everywhere to the right of the B distribution, there is some chance that $\eta$ will exceed $\epsilon$ where $\eta$ represents the value of a *B*.

The point can be proved more generally. Let $\epsilon$ and $\nu$ be i.i.d. random variables. Further, let

$$\eta=\rho\epsilon + (1-\rho)\nu ,$$

where $0\leq\rho\leq1$ .

The general proposition to be shown is that the expected value of the highest order statistic, $\max[\epsilon,\eta]$, decreases in the correlation between $\epsilon$ and $\eta$. That is, disjointness increases the gains to diversity. In this context, firms are most likely to seek best practices by hiring workers from pools where knowledge is less correlated. Given that $\eta$ and $\epsilon$ are i.i.d., $\rho$ is the correlation coefficient between $\eta$ and $\epsilon$. Thus, showing that the expected value of the highest order statistic decreases in the correlation between $\eta$ and $\epsilon$ boils down to showing that $E\{\max[\epsilon,\eta]\}$ decreases in $\rho$.

Now,

$$E\{\max[\epsilon,\eta]\} = \text{prob}(\eta>\epsilon)\, E\{\eta \mid \eta > \epsilon\} \quad + \quad \text{prob}(\eta<\epsilon)\, E\{\epsilon \mid \eta < \epsilon\} .$$

From the definition of $\eta$, $\eta>\epsilon$ holds iff $\nu>\epsilon$. Thus, $\text{prob}(\eta>\epsilon)= \text{prob}(\nu>\epsilon)$. Further, since $\nu$ and $\epsilon$ are i.i.d. $\text{prob}(\nu>\epsilon) = \frac{1}{2}$ . Similarly, $\text{prob}(\eta<\epsilon) = \frac{1}{2}$ so that we can rewrite

$$E\{\max[\epsilon,\eta]\}= \tfrac{1}{2} [\, E\{\eta \mid \eta > \epsilon\} \quad + \quad E\{\epsilon \mid \eta < \epsilon\}] .$$

Using the same logic, this means that

$$E\{\max[\epsilon,\eta]\}= \tfrac{1}{2} [\, E\{\eta \mid \nu > \epsilon\} \quad + \quad E\{\epsilon \mid \nu < \epsilon\}]$$

so

$$E\{\max[\epsilon,\eta]\} = \tfrac{1}{2}\, E\{\, [\rho\epsilon + (1-\rho)\nu] \mid \nu > \epsilon \,\} + \tfrac{1}{2}\, E\{\epsilon \mid \nu < \epsilon\}$$

Thus,

$$d\, E\{\max[\epsilon,\eta]\} / d\, \rho = \tfrac{1}{2} [\, E\{\epsilon \mid \nu > \epsilon\} - \; E\{\nu \mid \nu>\epsilon \,\} \,]$$

which is negative because the expression inside the brackets must be negative.[6] The value of

---

[6] This follows because $E(\epsilon \mid \nu>\epsilon) - E(\nu\mid\nu>\epsilon) = E(\epsilon-\nu\mid\nu>\epsilon) < 0$.

drawing from different distributions increases with the independence of those distributions.

The discussion above fits the problem that businesses face when choosing a new technology. Suppose, that a new product is about to emerge, but that there are competing technologies for producing the product. Previous examples from television are the choice of color TV system to be used, which arose during the 1950s, and the more recent choice between British and American high definition TV systems. Consider a British manufacturer who believes it most likely that the British system will dominate, but acknowledges that there is some chance that the American system will be better. Specifically, assume that p of the time the British design will dominate and 1-p of the time the American design will dominate with $p > 1/2$. Suppose further that British engineers' knowledge is specialized to the British system and that American engineers' knowledge is specialized to the American system. Within each group, there is a distribution of talent and the actual technology that is use will be selected from the best designs of all the engineers. The profit (ignoring development costs) associated with a design by an British engineer i is given by $x_i$, and the profit associated with design by an American engineer j is given by $y_j$. Underlying talent generates a density of profit possibilities given by $f_x(x)$ for British engineers and $f_y(y)$ for American engineers. Suppose that engineers cost w. The firm's problem is to hire n engineers, chosen endogenously, where m are British and n-m are American.

It is the "best practice" aspect of the problem that leads to the gains from diversity. If output were additive across all workers, only one type of worker would be hired. The firm would specialize in hiring the worker with the highest expected output. In this case, the firm would hire only *B* s (for British) if $pE(x) > (1-p)E(y)$. If this condition were not met, it would hire only *A* s. This is the

solution to the standard problem of worker choice.[7]

When only one design is used, the problem becomes akin to selecting a best practice. In its search for a best practice, the firm cares about the highest order statistic only. In cases where it pays to hire more than one type of labor, some interesting results obtain.

Denote by $X_m$ the expected value of the $m^{th}$ order statistic in a sample of size m drawn from the density $f_x(x)$. Denote by $Y_{n-m}$ the expected value of the $n\text{-}m^{th}$ order statistic in a sample of size n-m drawn from the density $f_y(y)$. Assume x and y are non-negative. Then, the firm maximizes profit by choosing m and n to maximize

(2) Expected Profit = $p X_m + (1-p) Y_{n-m} - n w$

The expectation of the highest order statistic is a concave function of the sample size and this is all that is necessary to derive the main result.[8] For simplicity and concreteness, we assume that

---

[7]See, for example, Lazear (1998b), ch. 2 for a discussion of this point.

[8]It pays to hire some of *A* and some of *B* when

$$\frac{w}{X_m} \le p \le 1 - \frac{w}{Y_{n-m}}.$$

That this is true can be shown as follows: Rewrite (3) as

$$\text{Expected Profit} = pX_m - wm + (1-p)\,Y_{n-m} - (w)(n-m)$$

or as

$$EP_A(m) - EP_B(n-m)$$

where

$$EP_A(m) \equiv pXm - wm$$

and

$$EP_B(n-m) \equiv (1-p)(Y_{n-m}) - (w)(n-m).$$

$f_x$ and $g_y$ are from the class of normal, gamma, logistic, or exponential densities. In those cases, the expectation of the kth order statistic from a sample of k (k>1) is given by $\theta \ln(k)$ where $\theta$ is a parameter derivable from the parameters of the underlying density function. Then, (1) can be rewritten as

(3) Expected Profit = $p\,\theta_x \ln(m) + (1-p)\,\theta_y \ln(n-m) - n\,w$

which has first-order conditions

$$\frac{p\theta_x}{m} - \frac{(1-p)\theta_y}{n-m} = 0$$

$$\frac{(1-p)\theta_y}{n-m} = w$$

---

An interior solution, where both types of labor are used, exists if $EP_A(1) > 0$ and $EP_B(1) > 0$. Since $EP_A$ is concave, $EP_a(1) \leq 0$ implies that $A$ is not used. Now,

$$EP_A(1) = pX_m - w$$

so $EP_A(1) > 0$ if

$$\frac{w}{X_m} < p.$$

Analogously, $EP_B(1) > 0$ if $(1-p)Y_{n-m} > w$ or if $p < 1 - \frac{w}{Y_{n-m}}$. Thus, an interior solution exists if

$$\frac{w}{X_m} \leq p \leq 1 - \frac{w}{Y_m}.$$

As labor cots goes to zero, an interior solution is sure to exist.

The solution is

(4) $m = p\,\theta_x / w$ ; $n = [\theta_y - p(\theta_y - \theta_x)] / w$ .

Further, when x and y are i.i.d., then $\theta_x = \theta_y$ and (3) can be rewritten as

(5) $m = np$ ; $n = \theta_y / w$ .

The firm hires workers in proportion to the probability that the worker's type succeeds. In this case, the firm hires p of the *B*s and (1-p) of the *A* s. Further, the demand for labor in general has unitary elasticity.

The conclusion of this section is this. Because correlation is important in choosing teammates, a firm that seeks to achieve the best practice gives up mean for negative correlation. It pays to locate some of the workforce in a country or field that has a lower expected value in order to obtain variety in outcomes that maximize the value of the highest order statistic.

The fact that a firm is willing to trade off mean to obtain lower correlation between inputs provides a way to think about the number and identity of the countries in which has teammates. Consider a firm that has a British engineer and a Brazilian engineer and is considering the addition of a third engineer. Should he come from one of the two countries that are already represented, or should he be, say, a Frenchman? This depends on the cost of the Frenchman relative to other labor types and on the correlation between the Frenchman's skills and those of the current workforce. If, for example, the French engineer's output were perfectly correlated with the British engineer's output, there would be no gain to adding a Frenchman, but there would be a cost. The firm would either have to have engineers who could speak French, or find a French engineer who was an English speaker (or a Spanish speaker, if that were the chosen language of the firm). The cost of using an English speaking French engineer shows up as a premium that must be paid to bilingual workers.

Even if the firm decides to hire the Frenchman, eventually it will not pay to add new nationalities to the firm. Either the mean of the incremental country will be too low relative to the wage, or the amount of new variation will not be high enough. This is analogous to the finance notion of redundant securities. Unless an additional country can diversify the firm's "portfolio" of skills sufficiently, there is no additional value to adding new countries, especially when this requires the addition of new languages to the set of those spoken at the firm.

**Choosing to act as a Liaison**

The examples in the last section illustrate that firms may want to construct multi-cultural teams. A multi-cultural team can only add value when some of the individuals are bilingual. In this section, the individual's decision to become bilingual is considered.

Individuals who come from different groups may have different backgrounds, cultures or languages, but they can invest in learning the culture of language of another group. Within any population, some persons are better at acquiring language skills than others. There is evidence that much of this relates to age, but even within a age group, there are differences. Let us suppose that the cost to individual i of learning the other group's language is $k_i$ and that $k_i$ has density function $g_a(k)$ and distribution function $G_a(k)$ for the *A*s, and similarly $g_B(k)$ and $G_B(k)$ for the *B*s. Further, assume that $G(0) = 0$; no individuals have negative costs of learning a language.

An individual will invest in learning the other language if the cost of doing so is less than the benefit. Suppose that a monolingual *A* earns $w_A$ and a bilingual *A* earns $w_a^*$, and similarly for *B* s. Then if individual i is an *A*, he learns to speak *B* if

(6)

$$k_i < w_A^* - w_A$$

so that

(7) Proportion of *A* s that is bilingual = $G_A(w_A^* - w_A)$ .

Analogously,

(8) Proportion of *B* s that is bilingual = $G_B(w_B^* - w_B)$ .

The bilingual individuals are choosing to act as liaisons because in equilibrium, described in the next section, bilingual *A* s work in *B* firms where *B* is spoken and bilingual *B* s work in *A* firms where *A* is spoken.

**Wage Determination and Market Equilibrium**

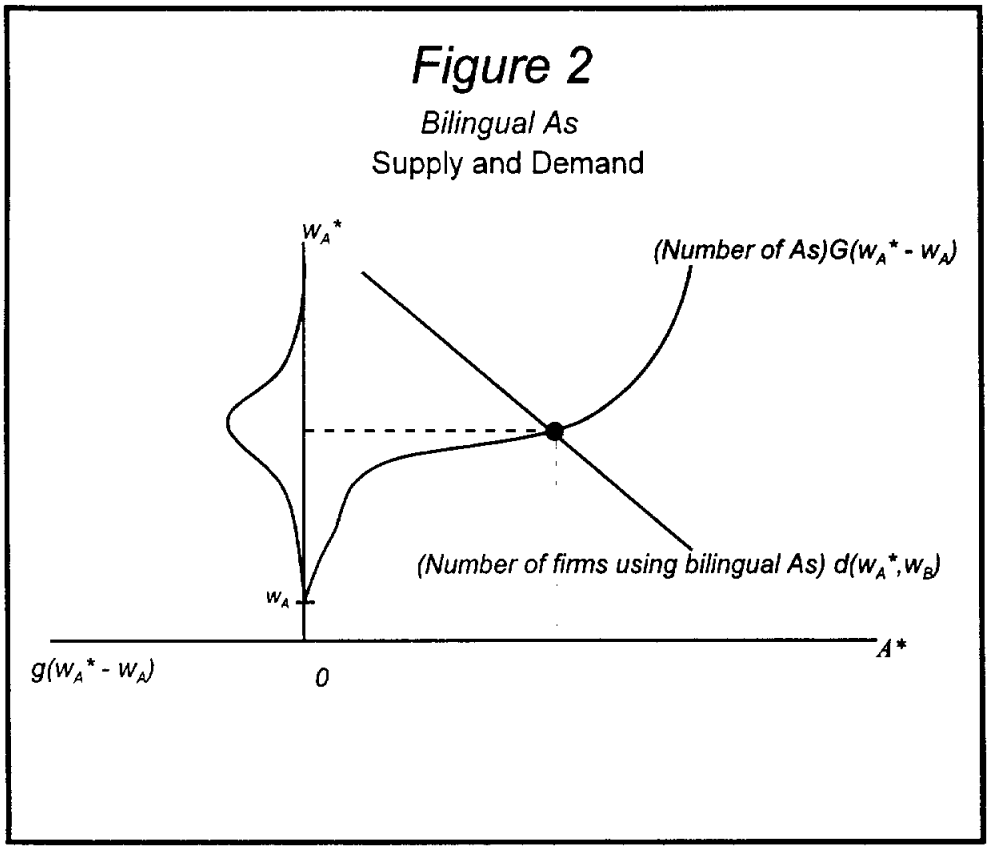
Consider first country A. Suppose that country A has two types of firms. Most have technologies that cannot take advantage of the gains from multicultural interaction. These include local services, retailing, and mining of natural resources that are sold in domestic markets only. Suppose that the domestic sector is large relative to the global sector. Wages of monolingual *A* s are determined by their marginal product in the domestic sector. Since the number of bilingual individuals is small, assume for simplicity that the wage of monolingual *A*s is given.[9] Also, because monolingual *A* workers are indifferent between working in domestic firms and in firms that have foreigners, but speak *A*, wages of monolingual *A* s, irrespective of workplace, is $w_A$ . A similar

[9]It is possible to make endogenous the wages of those employed in domestic firms, but this simply adds algebra without yielding any insight. What is important is the difference between wages of monolingual and bilingual workers, not the level of either type's wages.

argument determines $w_B$ in country *B*.

Labor supply of bilingual *A* s can be derived from (7), which depends on $w_A$, given, and on $w_A^*$, to be determined. Demand for bilingual *A* s comes from the first order condition in row 6 of table 1. Since $w_B$ is determined, it is possible to solve the two first order equations to derive the demand by each firm for bilingual *A* s as a function of $w_A$. After aggregating over the entire economy, the market demand for A* is determined. Coupled with labor supply, $w_A^*$ is derived. Analogous logic lies behind the derivation of B* and $w_B^*$.

Figure 2 shows the equilibrium. The density of k is shown, sideways, on the left hand side of the diagram. It is shifted so that the density has its lower support at $w_A^* = w_A$. Multiply the proportion that is the integral under g() up to $w_A^*$ by the number of *A* s in the population to derive the supply of A*, shown as the upward sloping curve. The demand for A* is simply the solution to the firms' first order conditions, aggregated and inverted. The equilibrium wage and quantity of bilinguals is shown by the intersection of the supply and demand curves.




It is now possible to make some statements about equilibrium wages and in particular, about the premium to being bilingual. The rent that a bilingual *A* receives is

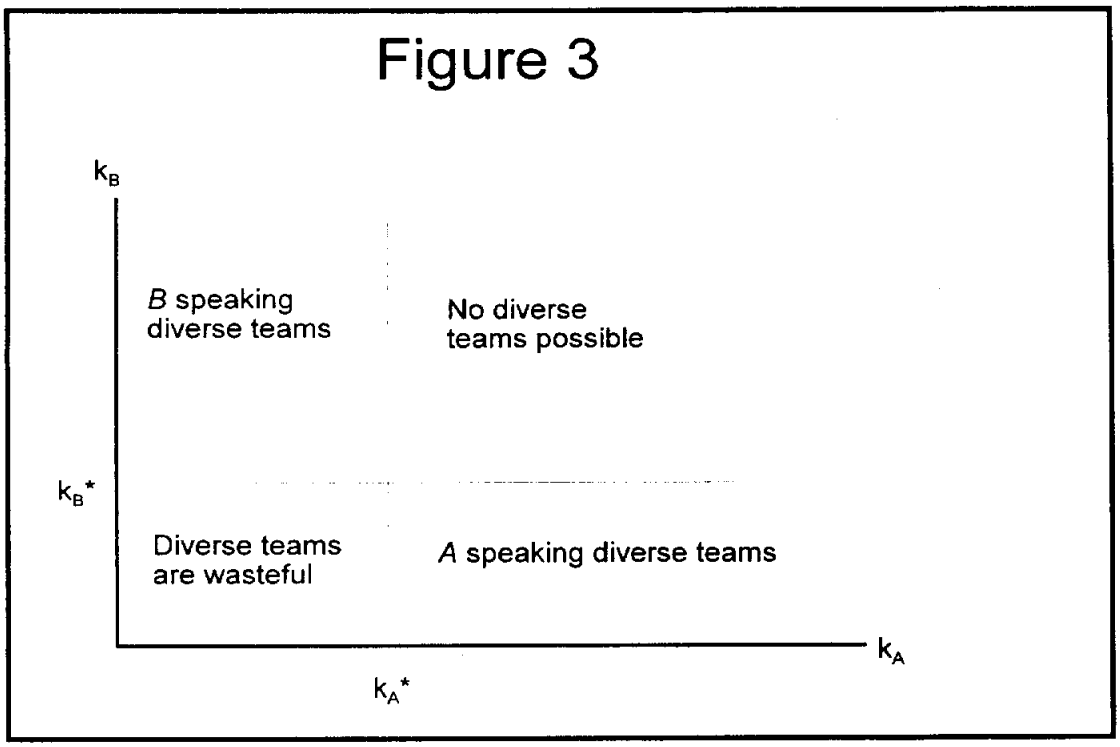
(9) Rent to a bilingual A with cost k = $w_A^* - w_A - k$

or

$$= f_1(A^*, B) - w_A - k .$$

The marginal bilingual, who is just indifferent between becoming bilingual, earns higher wages than monolingual individuals, but no rent. The difference in wages between bilinguals and

monolinguals just covers the cost of learning the other language. Since the marginal individual must face positive costs of learning, $w_A^* > w_A$. All inframarginal bilinguals earn rent by virtue of their




scarce rapid language learning skill.

Figure 3 shows the equilibrium situation. The cutoff cost levels are $k_A^* \equiv w_A^* - w_A$ and $k_B^* \equiv w_B^* - w_B$, for *A* s and *B* s respectively. All *B* s whose costs of learning *A* are below $k_B^*$ opt to learn *A* and all *A* s whose costs of learning *B* are below $k_A^*$ opt to learn *B*. The lower right quadrant shows the kinds of individuals, by costs, that form *AB* teams and speak *A*. These teams are comprised of *A* s who have high costs of learning *B* and *B* s who have low costs of learning *A*. The upper left quadrant shows the kinds of individuals, by costs, that form *AB* teams and speak *B*. These teams are comprised of *A* s who have low costs of learning *B* and *B* s who have high costs of learning *A*. High cost *A*s and high cost *B*s never work together in the same firm. This is shown in the upper right

quadrant. Furthermore, it is wasteful to put low cost , bilingual *A* s together with low cost, bilingual *B*s, since there is no gain to having all individuals bilingual, given the production technology of table 1.

Some comparative statics are informative. First, the bilingual premium increases when complementarities are strong so that $f_{12}(A^*,B)$ is high for any given A*, B pair. Loosely speaking, the larger is the value of diversity, the larger will be the demand for bilinguals and the more individuals who will be induced to invest in language skills. In terms of figure 2, more complementarity raises the value of $d(w_A^*, w_B)$. The marginal individual continues to m*2Uearn no rent, but this happens at higher levels of k and therefore higher numbers of *A*s who opt to become bilingual.

More formally, consider two production functions, f(A*,B) and h(A*,B) such that f(A*,0) = h(A*,0) and such that $f_{12} > h_{12} \; \forall$ A, B. The complementarities between A* and B in the f function exceed those in the h function. Now, since $f_{11} < 0$ and since $h_{11} < 0$, $A^*_0$, which solves the first order condition

$$f_1(A^*_0, B) = w_A^*$$

must exceed $A^*_1$, which solves the first order condition

$$h_1(A^*_1, B) = w_A^*$$

for any B>0. Thus, the demand for A* is higher with production function f() than it is with production function h().

Second, the premium also increases when the supply of potential bilinguals is low. The supply of bilinguals is derived from the distribution of costs. Thus, if most individuals have high

costs of learning a language, then the supply will be low in the relevant range and the equilibrium wage differential will be high. Thus, for example, a rightward shift in the density function consistent with first-order stochastic dominance will cause the equilibrium wage of bilinguals to rise. Formally, if J(k) is a distribution function such that $J(k) < G(k)$ for all k, then the equilibrium bilingual premium is higher when costs are distributed according to J(k) than when they are distributed according to G(k). This is clear because $G(w_A^* - w_A)$ lies everywhere above $J(w_A^* - w_A)$.

In countries where general levels of education are low and language skills are difficult to acquire, individuals who can act as liaisons earn high premiums for their services. The difference between the wages of those in the multinational sector and those in the rest of the economy is large. This premium associated with working in the "primary sector" results from a return to a scarce skill, i.e., the ability to travel in others' circles. Connections may be particularly important in this context. Companies often talk of the necessity of "employing" locals who have government and other connections which facilitate doing business. If there are few individuals who both have such connections and can act as a liaison to a foreign firm, the premium associated with this activity will be high.

Another interpretation describes the way in which teams form between minority and non-minority business people. When minority businesses are favored as a result of an affirmative action program, it pays for majority member to team up with a minority member who has learned to "speak" the majority's business language. The minority member has connections, whereas the majority member may have other expertise or access to capital markets. The minority liaisons earn a premium over the wages earned by minority members who work only with other minority members.

The costs of communication show up as a wage premium to bilingual individuals. The more to the left is the distribution of learning costs, the larger is the supply of bilinguals and the lower is the premium. When many individuals can become bilingual cheaply, the cost of intercultural communication is low and multicultural teams are more likely to form. Advances in communication technology, such as television, are likely to reduce the costs of learning other languages and reduce the premium earned by bilingual individuals.

Third, when alternatives are poor, more individuals work in foreign firms as liaisons. The proportion who opt to learn the foreign language or culture is seen, from differentiating (2), to decrease in $w_A$. As local alternatives become poorer, more individuals opt to bear the costs of learning the other language so that they can be liaisons in foreign firms. As a result, if country *A* is richer than country *B*, more *B* s will learn to speak *A* and work in *A* firms than will *A* s learn *B* and work in *B* firms. This implies that fewer British work in Saudi run firms than do Saudis in British run firms.

This model provides predictions about the wages of white collar workers who are employed abroad, often called "expats," and of immigrants. It is quite obvious that the wages of expats and immigrants are higher than those of monolingual workers who work in domestic firms in their native countries. Specifically,

$$w_A^* = w_A + k_A^*$$

and $$w_B^* = w_B + k_B^* .$$

Further, if *A* is the high wage country, $w_A > w_B$. Coupled with the previous condition, the ranking of wages is as follows:

$$w_A^* > w_A = w_B^* > w_B$$

Expats from country *A* and other *A*s who in *B* dominated firms earn the most. Monolinguals in country *B* earn the least. Without additional information, the relation of wages of monolinguals in country *A* and expats from country *B* cannot be determined.

The Optimal Size of Teams:

There are three considerations in determining the optimal size of a team. They are incentives, communication, and knowledge. Incentives in teams have been discussed in many different contexts[10] and are not discussed here. Instead, the focus is on communication and knowledge, which fit more closely with the earlier part of the analysis.

Two points to can made, first verbally, and then more formally. Most readers have experienced the difficulty associated with getting something accomplished in a committee setting. Communication is problematic. As the group size gets large, everyone speaks at once (or no one speaks at all). Often, the group divides itself into smaller, more workable units. Yet, there remains a tendency in business to call meetings and operate in a committee setting. Indeed, an academic seminar is a committee of sorts that is regarded as beneficial to scholarship. This is because of the second factor. The amount of information in the group tends to increase as the size of the group increases. The skill and knowledge set represented by a room of thirty individuals is usually considerably greater than that of just one of those thirty. It is for this reason that scholars benefit from airing early versions of their work in seminars. Of course, the value of the seminar depends on the quality of the participants and on the relevance of their knowledge to the topic at hand. It also depends on disjointness; it is better to present to a group who has information that is different from

[10]See, for example, Holmstrom (1982) and Kandel and Lazear (1992).

that of the author and relevant to him than to a group whose information and skills are virtually identical. These points can be made more precise.

The tradeoff between disjointness and communication are captured in the technology of table 1. Let us reinterpret the inability of *A*s and *B*s to communicate costlessly as the difficulty in a committee of making one's thoughts known to other members of the committee. Two *A*s, as previously assumed, can communicate without cost, however, because they have more common background, experiences, or frames of reference.[11] When do communication costs make a team too large?

First, if *A*s and *B*s have perfectly overlapping information and skills, so that *A* s and *B*s know $x_1$, then any positive communication costs make teams of *A* s and *B*s worse than homogeneous firms: Specifically,

$$2f(1,0)-w_A-w_B \quad > \quad f(2,0)-w_A-w_B^*$$

since $w_B^* > w_B$ and $f_{11} < 0$. There is no gain to creating a larger team because there is only additional communication cost without enlarging the skill set. It is better to have two one-member teams than one *A*,*B* team.

More interesting is the case where *A*s and *B*s have separate knowledge. The gain from having one team of two rather than two teams of one is shown as

$$Gain \quad = \quad f(1,1)-w_A-w_B^* \quad - \quad f(1,0)+f(0,1)-w_A-w_B$$

which can be positive or negative, depending on the importance of the complementarity and the cost

---

[11]This is extreme. All that is required is that communication between *A*s and *B*s is more expensive than communication between *A*s.

of communication. Here, the cost of communication is represented as the difference between $w_B^*$ and $w_B$. The premium paid for a bilingual *B* can be reinterpreted as the cost of communication between one member of the team and the other.

The gain represented above relates to a team where members have disjoint information or skill sets, but this is not enough. Relevance matters. When individuals have skills that are not particularly relevant to one another, $f_{12}$ is close to zero and the optimal team size is small. In this case, when $f_{12}$ is close to zero, the gain is likely to be negative. Even small communication costs lead to a preference of two one member teams over one two member team. In order for

$$f(1,1) - w_A - w_B^*$$

to exceed

$$f(1,0)+f(0,1) - w_A - w_B \quad ,$$

it is necessary that $f_{12}$ be sufficiently large to swamp the additional communication costs. The conclusion, then, is that small teams are better unless the following conditions hold:

1. Information is disjoint (*A* s know $x_1$ ; *B* s know $x_2$)
2. Information is relevant ($f_{12}$ positive and large)
3. Communication costs are small ($w_B^* - w_B$ is small)

## The Structure of Global Firms

The term "global" implies diversity, but most global firms are lopsided, having many more members from one country than from other countries. In this section, it is shown that asymmetry is a natural outcome in a symmetric world. The asymmetry, which may make the parent company appear chauvinistic or even imperialistic, is the optimal organizational form under very general

circumstances.

Let us continue to suppose that there are only two countries, *A* and *B*, and that the countries are identical in every respect. This means that $G_A(k) = G_B(k)$ for all k so that the distribution of costs for *A* s learning *B* is the same as that for *B* s learning *A*. Also, assume that the number of workers, the number of firms, and production technologies are identical in the two countries. As a result,

$$w_A = w_B .$$

*A* s know $x_1$ and *B*s know$x_2$. Further, assume that the production function in those firms that can take advantage of diversity (as shown in table 1) has the property that

$$f(x,y) = f(y,x)$$

so that a firm with x *A*s and y *B*s produces exactly the same amount as a firm with y *A*s and x *B*s. This is a description of a world with complete symmetry. Despite the symmetric nature of the world, firms are asymetric. In particular, firms tend to be lopsided, having either large numbers of *A*s and small numbers of *B*s or vice versa. Indeed, because translation is costly, symmetric firms can only come about if the underlying world in which they operate is asymmetric.

First, it is shown that firms tend to be asymmetric, given an initially symmetric world. If there were no need to use bilingual labor, then firms would choose to use monolingual labor because, as shown above, the wages of monolinguals are lower than the wages of bilinguals. The first order condition, under these circumstances, is

$$f_1(x,y) / w_A = f_2(x,y) / w_B .$$

A solution to this first order condition is that x=y, because $f(x,y) = f(y,x)\ \forall\ x,y$ implies that

$$f_1(x,y) = f_2(y,x) .$$

Since $w_A = w_B$, x=y solves the first order condition because $f_1(x,y)=f_1(x,x)=f_2(x,x)=f_2(x,y)$.

By extension, if bilingual workers of (at least) one type must be used in order to get output described by f(x,y), then the wage that is paid to *A* s by a firm that employs bilingual *A* s is $w_A^*$, which exceeds $w_A$. This means that $f_1(x,y)<f_2(x,y)$ when the first order condition is satisfied. But since $f_{ii}<0$ and $f_{12}>0$, fewer bilingual *A* s than monolingual *B* s will be used. This firm is lopsided. More than half of the workers are monolingual *B* s; fewer than half are bilingual *A* s. The language spoken and the culture used in the firm is *B*.

Of course, the argument could have been reversed. Since the world is symmetric, just as some firms will have many monolingual *B*s and few bilingual *A* s, so too are there firms that will be the mirror image. They will have many monolingual *A* s and few bilingual *B* s. These firms will speak *A* and have an *A* culture. Thus, a symmetric world produces asymmetric firms. Some are dominated by *A*s and have culture *A*; some are dominated by *B* and have culture *B*.

The firms may look chauvinistic in that they do not integrate. Rather, they impose the plurality language and culture on the rest of the firm. But this may have nothing to do with chauvinism. It is a direct result of communication costs which provide incentives for the firm to select a common standard - in this case a language and business culture.

These firms are not parochial. In fact, one could argue the reverse. Unlike strictly domestic firms, these global firms bring in talent from other countries, but the cost of doing this is imposing a common standard, which appears takes a form that some might describe as corporate imperialism.

It is also true that in order for firms to have equal numbers of *A* s and *B* s, some underlying asymmetry is required. In a symmetric world, the wage of all bilinguals are equal and the wages of all monolinguals are equal, irrespective of type. For a firm that employs monolingual *A* s and bilingual *B* s to have the same number of each type, it is necessary that $w_A = w_B^*$, by the previous

argument. But in order for $w_A = w_B^*$, it must be that $w_B < w_A$ because $w_B^* > w_B$. However, this requires asymmetry; *B* would be have to be a lower wage country, which requires some differences and immobility in factors, technology and products. Thus, if the number of types is the same in firms that speak *A*, *B* must be a lower wage country.

Now, if it is necessary for country *B* to have lower wages than country *A* in order to have the same number of workers of each type in firms that speak *A*, what does this imply about firms that speak *B*? Since *B* is the low wage country, $w_A > w_B$ which implies that $w_A^*$ is even higher, relative to $w_B$ than it would be in a symmetric world. Bilingual *A* s are very expensive in this asymmetric world, which means that many fewer are used in *B* speaking firms than are bilingual *B* s in *A* speaking firms.

This analysis provides an implication: Firms in and from rich countries employ more bilingual members of poor countries than firms in and from poor countries employ bilinguals from rich countries. Thus, we expect fewer British workers in Romanian firms than Romanian workers in British run firms.

Federalist Structures:

In some rare circumstances, firms have small numbers of employees in many different countries, without having a large plurality in any one. When this occurs, the firm may adopt a "federalist" structure. Each country's branch office or offices may use their own language and communication between them may occur in some third language. Most often the language is English because English has become the language of business. This phenomenon is explained by the model.

Consider a company, like an airline, that must have offices throughout a large part of the world. Swissair is a good example. One subsidiary, Gate Gourmet, runs very independently with a decentralized structure. There is no large center; food services are provided in each of the airport countries. The offices are loosely confederated with Swissair, but each sets it human resources practices somewhat independently.

It is easy to see how this happens. Suppose that a firm has offices in N countries. One possibility is to integrate fully, by hiring people who can speak all N languages. Another is to hire people who can speak their own language and one neutral language. It is intuitive that hiring someone who is multilingual, defined as being able to speak N languages, is more costly than hiring someone who is bilingual. It is also a direct implication of the formal analysis.

From (6) and (7), it is possible to derive the equilibrium wages of multilinguals and compare them to bilinguals. Consider multilinguals from country A who have wage $w_A'$. Bilinguals receive wage $w_A^*$ and monolinguals receive wage $w_A$. The cost of becoming multilingual for individual i is $N k_i$ so the condition for becoming multilingual is

$$Nk_i < w_A' - w_A .$$

Using the logic in (7), the proportion of *A*s who become multilingual is then

$$G_A( (w_A' - w_A)/N )$$

which is necessarily less than $G_A (w_A^* - w_A)$ for $w_A^* = w_A'$ and for N>1. Thus, if the value of multilinguals were no lower than that of bilinguals, the wage of multilinguals would be higher than that of bilinguals because the supply is lower.

Beyond cost saving associated with selecting one common language rather than many,

additional saving is achieved in a federalist structure by using a limited interface. Instead of insisting that all individuals in each branch office be multilingual or even bilingual, a firm may choose to use only a few liaisons. The local office operates primarily in its native tongue and communicates with the center and other offices through the common language.[12]

Indeed, an SAirGroup officer reports that "our official company language is English; i.e., all written communications to/from abroad are in English. This also means that people abroad, communicating with people in other countries or with those at the head office in Zurich use English."[13]

This is the structure that applies to Gate Gourmet. Language can be interpreted more broadly to connote business or general culture. The human resources system, wages, and industrial relations tend to conform to the country in which the airport is located.

The Business Case for Diversity:

Some have suggested that there may be business reasons for having employees of different types.[14] The argument relies on a correlation between the type of the salesperson and the type of the customer. As such, it focuses on communication among teammates. One can think of the customer and salesperson as comprising a team, and communication between similar types is easier than communication between different types. A specific example of which the author is aware involves a large American bank. The bank hires bank tellers who are from the same ethnic background as the

---

[12]The production technology described in table 1 does not permit this kind of leverage.

[13]Claudio Corti in correspondence, Dec. 23, 1997.

[14]See, for example, O'Reilly, Williams and Barsade (1997), Kole and MacDonald (1997).

majority of customers in the area served by that teller, e.g., Hmong tellers in Hmong communities in the North-Central United States.

The focus on communication makes the argument most appropriate for services, which often emphasize one-on-one interaction with little room for economies of scale. But in other industries, the argument is unpersuasive. It would be more difficult to argue, for example, that an American would be unlikely to buy Saudi oil because he is unable to speak Arabic. Cheap translation between a few individuals at the level of extraction, refining and exporting make widespread bilingualism unnecessary.

The Quintessential Team:

The most prevalent team in society consists of the pairing of a male and female in marriage. No theory of teams that is inconsistent with this fact deserves serious consideration. But the desiderata of this model are consonant with the characteristics of marriage between a man and woman. First, men and women have disjoint skills, particularly when it comes to procreation. Second, their skills are relevant to one another, again, best exemplified by reproduction. Third, they can communicate with one another, although this area seems to be subject to dispute.[15] Becker, Landes, and Michael (1977) provide evidence on divorce that reveals that marital dissolution (not to mention failure to marry in the first place) is most likely when the married parties are from different cultures, language groups or religions. All of those factors are interpreted in this model as reflecting costs of communication between types.

---

[15]See, for example, Tannen, Debra, *You Just Don't Understand.*(July 1991)

## Empirical Analysis

### Trading Patterns

The theory presented suggests that teams should be more likely to form between cultures that have easy communication, and complementary knowledge or skills. Complementary skills are those that are different from those possessed by one group of individuals, but are also relevant to that group. At the empirical level, it is difficult, if not impossible, to obtain direct measures of who works with whom. Even if this could be done, it would then be necessary to obtain information on the characteristics, skills, and knowledge of the individuals who are engaged in team production.

A much less accurate, but also less ambitious approach is to examine trading patterns by country. This provides only indirect evidence of the ways in which teams are formed, but it is suggestive. The assumption is that in order to trade with another country, it is necessary to have at least some workers who are from or stationed in that other country. "Translators" are needed. What will be shown in this section is that to the extent that trade patterns are a proxy for team formations, the predictions of the model hold up. Specifically, countries are more likely than expected, as compared with similarly situated countries, to trade with other countries that speak their language.

Table 2

Imports, Selected European Countries

1996

| | Imports from: | | | | | |
|---|---|---|---|---|---|---|
| Country | Total | US | Canada | Australia, NZ, SA,HK,Sing, India | West Hemi- | Brazil |
| UK | 285831 | 36002 | 3873 | 12830 | 5438 | 1533 |
| France | 455683 | 32551 | 3295 | 11702 | 9975 | 1840 |
| Spain | 121784 | 7720 | 644 | 2006 | 4912 | 1160 |
| Germany | 455683 | 32551 | 3295 | 11702 | 9775 | 3240 |
| Neth | 180639 | 14689 | 896 | 3503 | 4567 | 1721 |
| Italy | 206883 | 10175 | 1858 | 10716 | 5050 | 1891 |
| Portugal | 33932 | 1083 | 117 | 262 | 973 | 472 |
| Percentages | | | | | | |
| UK | 1 | **0.126** | **0.014** | **0.045** | 0.019 | 0.005 |
| France | 1 | 0.071 | 0.007 | 0.026 | 0.022 | 0.004 |
| Spain | 1 | 0.063 | 0.005 | 0.016 | **0.040** | 0.010 |
| Germany | 1 | 0.071 | 0.007 | 0.026 | 0.021 | 0.007 |
| Neth | 1 | 0.081 | 0.005 | 0.019 | 0.025 | 0.010 |
| Italy | 1 | 0.049 | 0.009 | 0.052 | 0.024 | 0.009 |
| Portugal | 1 | 0.032 | 0.003 | 0.008 | 0.029 | **0.014** |
| Average | 1 | 0.071 | 0.007 | 0.027 | 0.026 | 0.008 |

Source: *Direction of Trade Statistics Yearbook, 1997*, Real Sector Division, IMF Statistics Department. Washington, DC, 1997.

Table 2 examines imports from selected European countries. The reason for using European countries is that the travel distances to each of these countries from the countries considered are similar. Furthermore, all countries considered have ports and ready access to shipping.

The table reveals that countries are more likely to import from countries that speak their own language. The bottom half of the table calculates the percentage of total imports by a given country that come from a country listed in that column. For example, 12.6% of the UK's imports comes from the US. The bold numbers in the bottom half of the table reflect imports from countries with a common language.[16] All of the bold numbers are greater than the averages for all countries in the sample. For example, on average, the selected European countries import 7.1% of their goods from the United States. The UK, on the other hand, imports 12.6% of its goods from the US. Countries listed as Western Hemisphere include all of those in the Western Hemisphere excepting Canada and the US. They are primarily Spanish speaking. Note that Spain is more likely than average to trade with Western Hemisphere countries. Finally, Brazil is the Western Hemisphere country that speaks Portugese and Portugal is more likely than average to import from Brazil.

Are there other explanations? One obvious possibility is that countries that share language were once colonies of the countries with which they now trade and the historic patterns of colonial trade are continued, through inertia. Although true, this raises two more basic questions. First, why should a country necessarily trade with its colonies? After all, trading with colonies is endogenous. Second, why should there be persistence? Most of the countries have severed formal ties with their parent countries.

---

[16]Canada is counted as English, rather than French speaking because the majority speak English.

The answer to both of these questions is the same. Countries trade with their colonies and vice versa because culture and language is similar. The parent country makes is the most natural trading partner for the colony, even when distances are great. These relations persist over time because of the value of communication and culture. In Lazear (1997), it is shown that immigrants to the United States tend to cluster, locating in areas where others speak the same language and come from the same native country. This is particularly true for immigrants who are not fluent in English. Because of the need to communicate to trade, they locate near others with whom they can most easily converse. A colony or country trading with its parent is analogous. Communication costs are saved by this pattern of trade.

A comparison of table 3 to table 2 reveals that inertia is not a major force. An examination of trade patterns in 1965 reveals that they are very different from those in 1996. For example, the UK imported from Canada about 3/4 of what it imported from the US in 1960. In 1996, that fraction had fallen to around 11%. Imports from Germany accounted for about 5% of the UK's total imports in 1965. In 1996, Germany accounts for over 13% of the UK's imports.

Table 3 also yields another finding. Language based trading patters seem to have been more important in 1965 than they are today. The UK traded more with English speaking countries in 1965 than it does today, and Spain traded more with Spanish speaking countries in 1965 than it does today. One exception is Portugal, which traded relatively less with Brazil in 1965 than today.

That language was more important in 1965 than in 1990 is not surprising. Since English has become the neutral language of business, the costs of trading across different non-native speaking borders has been reduced. This effect should be particularly observable with respect to UK imports. If the rest of the world has become English speaking, then the need to trade with other English

speaking countries has been reduced. In fact, that is what is observed. In 1996, 18.5% of the UK's imports came from English speaking countries. This is down considerably from the 33.5% that prevailed in 1965.

Others have provided evidence on these points as well. Rauch (1997) tests a "gravity" model of trade and finds that the amount of trade is greater between counties that have common language or colonial histories together. This is more true for differentiated products than it is for homogeneous goods. Further, Rauch and Trindade (1997) find that Chinese networks, as proxied by ethnic Chinese population shares, are positively associated with bilateral trade between the two countries. Rauch's model is not quite the same as the one emphasized in this paper, but he does push the importance of network effects.

Table 3

Imports, Selected European Countries

1965

| | Imports from: | | | | | |
|---|---|---|---|---|---|---|
| Country | Total | US | Canada | Australia, NZ, SA,HK,Sing, India | West Hemi- | Brazil |
| UK | 16138 | 1886 | 1284 | 2227 | 1084 | 79 |
| France | 10341 | 1088 | 116 | 312 | 468 | 82 |
| Spain | 3004 | 70 | 7 | 45 | 235 | 28 |
| Germany | 17482 | 2299 | 228 | 456 | 1186 | 204 |
| Neth | 7460 | 765 | 65 | 85 | 305 | 59 |
| Italy | 7347 | 989 | 91 | 237 | 553 | 89 |
| Portugal | 896 | 71.6 | 5.5 | 11.4 | 31.6 | 5.3 |
| Percentages | | | | | | |
| UK | 1 | **0.117** | **0.080** | **0.138** | 0.067 | 0.005 |
| France | 1 | 0.105 | 0.011 | 0.030 | 0.045 | 0.008 |
| Spain | 1 | 0.023 | 0.002 | 0.015 | **0.078** | 0.009 |
| Germany | 1 | 0.132 | 0.013 | 0.026 | 0.068 | 0.012 |
| Neth | 1 | 0.103 | 0.009 | 0.011 | 0.041 | 0.008 |
| Italy | 1 | 0.135 | 0.012 | 0.032 | 0.075 | 0.012 |
| Portugal | 1 | 0.080 | 0.006 | 0.013 | 0.035 | **0.006** |
| Average | 1 | 0.099 | 0.019 | 0.038 | 0.059 | 0.009 |

Source: *Direction of Trade*, International Monetary Fund and The International Bank for Reconstruction and Development. Washington, D.C., (1966?)

**Firm-Based Statistics**

As mentioned earlier, it is difficult, if not impossible, to get a systematic sample of the nationalities of employees on the payroll of even large international corporations. As a result of direct contact with five global firms, I have been able to obtain some statistics on the proportion of workers who are currently employed in the home country and elsewhere. I have also obtained some time series evidence for these firms. Obviously, this sample is too small to be of any statistical value. Furthermore, it is not a random sample in any sense. But the patterns that are observed in all three cases are interesting. Table 4 reports the results for these five large firms.[17]

Table 4

Averages for Five Firm Sample

| Variable | 1960 | 1980 | 1997 |
|---|---|---|---|
| Percent of Total Employees working Outside Home Country (s) | 38% | 48% | 68% |
| Percent of Production that is Outsourced | <10% | <10% | 24% |
| Total Number of Employees in Firm (in thousands) | 152 | 194 | 193 |

There are two patterns that are shown in table 4. First, the percent of individuals employed outside the home country has risen from 38% in 1960 to 68% in 1997. Second, the percent of production that is outsourced has gone up significantly over time. To the extent that the numbers

[17]The firms are ABB (Asea Brown Boveri), British Petroleum, Ford Motor Company, Swissair, and Unilever.

in the table reflect globalization, these five firms have become more global over time.

## Conclusion

Economics can be used to understand cultural relations in general and the predictions about the mixing of cultures in particular. Elsewhere, (Lazear 1995), I have used this general idea to explore language acquisition and cultural assimilation. Here, the idea is used to analyze the globalization of firms. The building block of this theory of global firms is the notion that the global firm is a multi-cultural team. Teams that span cultures carry costs, but also benefits. Why would a firm bother to combine cultures when doing so may be very expensive relative to using like-cultured individuals? The answers relies on gains from using complementary factors that are more easily or cheaply obtained by hiring from a different culture.

The gains derive from two features of teammates: disjointness and relevance. Costs are incurred because some team members must be bilingual or bicultural. The cost shows up as a premium that bilingual and bicultural workers earn over their monolingual and monocultural counterparts.

The search for best practices provides one example of the gains from factor complementarity. The best practice can be thought of as the expected value of the highest order statistic, which has been shown to be decreasing in correlation between skills and knowledge of the underlying factors. The best practice is expected to be better when teams of unlike individuals are formed. As a result, firms may hire from an inferior labor pool to reduce the correlation between factors of production. At equal wages, under quite general conditions, firms should hire individuals in proportion to the probability that their technology or market will dominate. Thus, if there is a 90% chance that the British approach will dominate, then at comparable wages, 90% of a firm's engineers should be

British. Market wages adjust, however, to take into account supply and demand differences across cultures.

Because bilingual labor earns a premium, firms tend to be lopsided, using mostly home labor in management positions. This makes a firm appear chauvinistic or even imperialistic, where employees outside the home country are forced to use the home country's language or business culture. But this behavior is a natural outcome of profit maximization that may have neither social nor nationalistic causes whatsoever. Lopsided firms result even when there is no preference for one's own type.

The results of the analysis hinge costly communication. The empirical evidence on trade among countries suggests that communication costs are significant determinants of trade patterns. Countries are more likely than expected to trade with countries that share their language. It is not inappropriate, then, to infer that communication costs are important within firms as well.

To sum up, economic models can be used to understand cultural issues. They give us predictions that seem to be borne out in a large range of cultural experiences.

References

Becker, Gary S. *A Treatise on the Family.* Enlarged Edition. Cambridge: Harvard University Press, 1991.

Becker, Gary S.; Landes, Elisabeth; and Michael, Robert. ``An Economic Analysis of Marital Instability," *Journal of Political Economy* **85** (December 1977): pp. 1141-87.

Hölmstrom, Bengt. ``Moral Hazard in Teams," *Bell Journal of Economics* **13** (Autumn 1982): pp.324-40.

Hong, Lu and Scott Page, "Probem Solving by Heterogeneous Agents," Draft, October 1997, Department of Economics, Syracuse University and Division of Humanities and Social Science, California Institute of Technology.

Kandel, Eugene and Edward P. Lazear. "Peer Pressure and Partnerships," *Journal of Political Economy* 100(4) (August 1992): pp.801-17.

Kole, Stacey and Glenn M. MacDonald, "Economics, Demography and Communication," Draft, October 1997, William E. Simon School, University of Rochester.

Lazear, Edward. "Culture and Language," NBER Working Paper #5249 (September 1995)

Lazear, Edward. "Diversity and Immigration" NBER Working Paper, forthcoming1998. Also forthcoming in George Borjas, ed., NBER Conference Volume, Chicago: University of Chicago Press, expected 1999.(a)

Lazear, Edward *Personnel Economics for Managers.* New York: John Wiley and Sons, 1998.(b)

O'Reilly, Charles A. III, Katherine Y. Williams and Sigal Barsade, "Demography and Group Performance: Does Diversity Help?" Research Paper No. 1426 (January 1997), Graduate School of Business, Stanford University

Pollak, Robert A. "A Reformulation of the two-Sex Problem," *Demography* 23 (May 1986): pp. 247-59.

Pollak, Robert A., " The Two-Sex Problem with Persistent Unions: a Generalization of the Birth Matrix-Mating Rule Model," *Theoretical Population Biology* 32 (October 1987) pp. 176-87 (1987b).

Rauch, James "Networks Versus Markets in International Trade," UCSD August, 1997. Forthcoming, *Journal of International Economics.*

Rauch, James and Victor Trindade. "Ethnic Chinese Networks in International Trade." UCSD,

draft, September, 1997.

Tannen, Debra, *You Just Don't Understand.* Ballantine Books: New York (July 1991).