Putting Things in Order:
Trade Dynamics and Product Cycles

Robert C. Feenstra and Andrew K. Rose*

November 12, 1999

Abstract

We develop a procedure to rank-order objects using censored panel data sets. We illustrate this by ranking countries and commodities using dis-aggregated American import data. We find evidence that countries and commodities can be ranked. Countries habitually begin to export goods to the United States according to an ordering; goods are also exported in order. We estimate these orderings using a methodology, which takes account of the fact that most goods are not exported by most countries in our sample. Our orderings seem sensible, robust and intuitive. They are correlated with macroeconomic phenomena such as productivity and growth rates.

JEL Classification Number: F10

Keywords: empirical; data; SITC; dis-aggregate; bilateral; semi-parametric; rank; missing; country; commodity; product cycle.

Robert C. Feenstra
Department of Economics
University of California
Davis, CA 95616
Tel: (916) 752-7022
Fax: (916) 752-9382
E-mail: rcfeenstra@ucdavis.edu

Andrew K. Rose
Haas School of Business
University of California
Berkeley, CA 94720-1900
Tel: (510) 642-6609
Fax: (510) 642-4700
E-mail: arose@haas.berkeley.edu

* Feenstra is Professor of Economics at U.C. Davis, Visiting Professor at the Haas School of Business, and Director of the International Trade and Investment Program of the NBER. Rose is B.T. Rocca Professor of International Business and Economic Analysis and Policy in the Haas School of Business at the U.C. Berkeley, Acting Director of the International Finance and Macroeconomics Program of the NBER, and Research Fellow of the CEPR. This research was performed in part while Rose was visiting the European University Institute; the Bank of Israel; and the US Department of Treasury. We thank two anonymous referees, George Akerlof, Giuseppe Bertola, Michael Burda, Jeff Frankel, Gene Grossman, Ben Hermalin, Olivier Jeanne, Rich Lyons, Robert Pindyck, John Rogers, Roger Studley, Lars Svensson, Bharat Trehan, Shang-Jin Wei and seminar participants at European University Institute, Federal Reserve Board, Federal Reserve Bank of San Francisco, the Graduate Institute for International Studies, Humboldt University, U.C. Berkeley and U.C. Davis for comments. This is a shortened version of a working paper with the same name, NBER WP #5975, CEPR DP #1629, available as a PDF file at http://haas.berkeley.edu/~arose, where the data set is also available.

I: Introduction

This paper is concerned with ranking objects. Ranking is a pervasive phenomenon in economic life. Consumers rank alternatives. Firms rank investment projects. Policymakers rank objectives. Faculty rank students, students rank faculty, and faculty rank faculty. In this paper we develop techniques to estimate rank-orderings from large dis-aggregated panel data sets. Our methodology accounts for the fact that observations may be missing, possibly in a non-random fashion.

The particular application we are interested in here is a ranking of goods and countries within the framework of the "product cycle" theory of international trade due to Vernon (1966). Traditionally, this hypothesis has been tested using detailed dis-aggregated data on preferences, technologies, factor endowments, or some combination; Deardorff (1984) provides a survey. Here, we pursue a more tractable approach that relies simply on easily observed trade patterns. We rank commodities using the order in which they are exported to the United States. The "product cycle" hypothesis of international trade suggests that there is an ordering of commodities that a country develops and begins to export. We also use our methodology to rank countries, and find sensible country rankings, which are correlated with interesting macroeconomic phenomena.

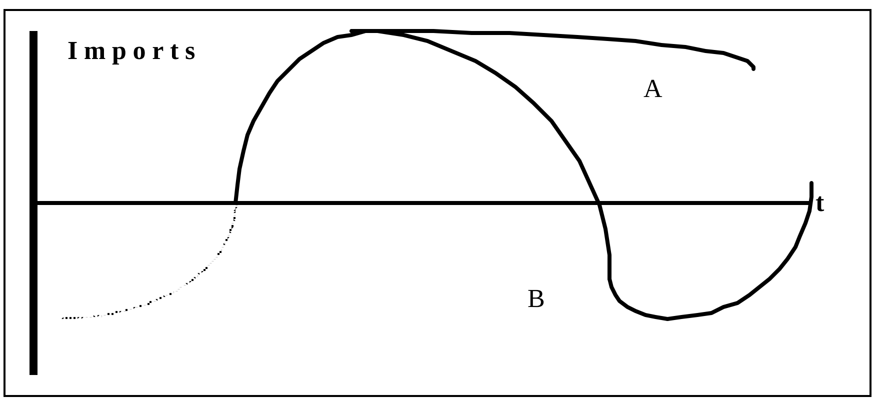
Our empirical methods are motivated by the trade and growth models in Grossman and Helpman (1991), which we briefly outline below in section II. This is followed by a brief discussion of our data set. Section IV constitutes the core of the paper; in it we develop a statistical methodology to estimate rankings. Our technique is applied in section V, which contains a discussion of empirical results. Section VI provides a summary, conclusion, and suggestions for future research. Proofs are provided in the working paper version of the paper.

II: Economic Framework

Building on the framework of endogenous growth, much work has been done recently on models of international trade with dynamic product markets. A comprehensive treatment of these models is provided by Grossman and Helpman (1991), who have linked growth to models of international trade with dynamic product markets.

Grossman and Helpman set out two formalized models of the product cycle. The first relies on the familiar Krugman model of intra-industry trade with imperfect substitutes sold by monopolistic competitors. Northern countries innovate by producing new varieties of horizontally differentiated goods. Southern countries eventually imitate these new goods and begin to export them to the North, taking advantage of lower costs. Once Southern countries begin to export a good, Northern production ceases. This is case "A" in Diagram 1.

Diagram 1: Product Cycle Import Patterns



The second model considered by Grossman and Helpman relies on their "quality ladder" model of continued innovation in the same industry. As an example, suppose the Northern country sells and exports personal computers. Eventually the technology is cloned and Southern clones drive the more expensive Northern PCs out of the market. But as North innovates by moving to superior machines based on the next generation of computer chip, the clone manufacturers lose their export base and the North begins to export again. Here, exports by the South are recurring and cyclic; case "B" in Diagram 1.

We are not certain which model of the product cycle best characterizes the data, if indeed there is any evidence of a product cycle at all. Therefore, we rely initially only on a single datum for each country-commodity observation. In particular, we exploit "the year of first export"; the year in which the country in question first exported the commodity in question to the US.[1] This datum does not depend on whether the good is subject to continued quality changes (as would be true of e.g., the year of last export).

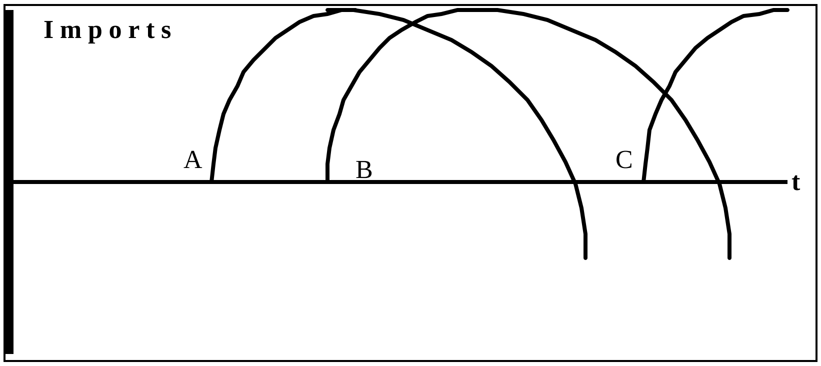
The intuition behind our technique for rank-ordering both commodities and countries is simple. Goods that are exported earlier are *less "advanced"* than goods exported later. In Diagram 2, product "A" is exported before "B", which in turn precedes "C". Thus, the ranking of goods in the order they are exported provides a measure of their "sophistication"; we would rank "A" the least advanced good, followed by "B", then "C". Alternatively, for each commodity, we consider the order at which *countries* first begin exporting that good (simply consider "A", "B", and "C" to be countries exporting a given good in Diagram 2). Countries that begin exporting earlier are considered to be *more "advanced"* than those exporting later.[2]

---

[1] One could check the sensitivity of the use of "the year of first import" by using the first year that imports (or cumulative imports) reach a given size.

[2] To illustrate the technique by way of an example, consider unexposed photographic roll film, (SITC code 88222). A total of 61 countries exported this to the United States during the sample period. In 1972, when the sample began, 13

Diagram 2: Import Patterns across Countries



To formalize this idea somewhat, let i=1,N denote the set of commodities, and let the (unobserved) rank order of their sophistication be X_i. That is, X_i is a set of integers running from 1 to N, indicating the order that we expect goods to be developed and exported. We do not observe X_i, but instead observe the actual rank-order by year of export, denoted by x_{ik} for countries k=1,M. We would not expect these orders to be identical to X_i: even in the models of Grossman and Helpman, a Southern country that adapts a technology from the North will generally have a range of possible goods that it can choose from, and it does not necessarily adapt in the same order that goods were developed in the North. The similarity between these rankings in theory will depend on characteristics of the goods (whether they are vertically or horizontally differentiated) and of the countries in question (such as the difference in their factor prices, as in Grossman and Helpman's "wide gap" and "narrow gap" cases).

countries were exporting film to the US, including Japan, Canada, Germany and Switzerland. In 1974, Korea began to export; Taiwan in 1976; Malaysia and China in 1989; Bangladesh in 1991; and Senegal in 1994. A total of 79 countries exported color TV receivers (SITC 76110) to the US in the sample. Japan was exporting in 1972; Canada began in 1973; Singapore in 1978; Brazil in 1981; Indonesia in 1983; Morocco in 1987; Niger in 1989; and Jamaica in 1993.

We model the imperfect correlation between the ranks x_{ik} and X_i by supposing that there is an integer-value $\rho_k N$ of the observations for which they are equal, while for the remaining observations the ranks are uncorrelated:

$$x_{ik} = X_i \qquad \text{for } \rho_k N \text{ observations and,} \tag{1a}$$

$$E[x_{ik}-(N+1)/2][X_i-(N+1)/2]=0 \quad \text{for the remaining } (1-\rho_k)N \text{ observations.} \tag{1b}$$

Note that in (1b) we measure both ranks relative to their mean values, which are (N+1)/2. We consider all possible sets of the $(1-\rho_k)N$ observations, of which there are $\binom{N}{(1-\rho_k)N}$. For each of these sets, the ranks X_i are *randomly reassigned* to the country ranks x_{ik}. Then the expectation in (1b) is taken over all possible sets of the $(1-\rho_k)N$ observations and values for x_{ik}.[3]

With this specification, the "product cycle" is measured by the rank ordering of commodities Xi, which we shall refer to as the "overall" ranking. Our goal in this paper is to obtain a meaningful measure of this overall ranking, using data on the country rankings x_{ik}. Implicitly we are assuming that all international trade is driven by the product cycle, rather than more standard considerations such as differing factor proportions or increasing returns. Of course, this is unlikely to be true in practice. If other types of international trade were important in practice for most goods and countries, we would expect our rankings to be meaningless and/or not well determined. What we find intriguing is that in fact, our rankings are both robust and meaningful in the sense that they are closely related to other

[3] In order for this expectation to be zero, it must be that the set of $(1-\rho_k)N$ observations contains more than one element, since otherwise we would have to assign $x_{ik}=X_k$ for that element.

(macroeconomic) phenomena. This suggests that product cycle considerations are more important than commonly considered.[4]

After briefly describing our data in section III, we then review methods suggested by Kendall and Dickinson (1990) to obtain an overall ranking. These methods do not depend on the particular specification in (1), but we will argue that they are inadequate to deal with the unbalanced nature of our data set. Accordingly, we develop alternative methods to estimate the underlying ranking, that allows for an unbalanced panel and also uses the specification in (1).

III: The Data Set

Features of typical panel data sets drive much of our methodology. We exploit a data set of American imports by source country, extracted from the CD-ROM data set of Feenstra (1996). In particular, we examine imports at the five-digit level of Standard International Trade Classification (SITC), revision 2, between 1972 and 1994. These span 160 countries and other geographical jurisdictions (which we refer to as "countries" for simplicity); and 1,434 commodities ("goods"). Our data set is the most dis-aggregated comprehensive data set available (to our knowledge).[5] The more dis-aggregated the data set, the better; we would like to be able to distinguish e.g., modern luxury cars

[4] For example, Leamer and Levinsohn (1995) spend two pages on product cycle evidence in their fifty-page survey of international trade empirics.

[5] So as to preclude data problems associated with revisions to the SITC, we only consider commodities which were exported by at least one country in 1972. Examples of such commodities include: "Human Hair" (29191); "Varnish Solvents" (59897); "Cotton Yarn 14-40 KM/KG" (65132), "High Carbon Steel Coils" (67272), and "Piston Aircraft Engines" (71311).

with air bags and anti-lock brakes from older cars. Our data set is unable to do that, but has the enormous advantage of covering all traded goods.

For each good and each country, we initially use only the *first year of export* to the United States.[6] There are 88,292 non-zero entries in the data set. One important feature of this data set is that there are *many* goods that are not exported by countries initially, but become exported during the sample period. That is, there are a great many zero values for imports by source country that become positive later in the sample period. This feature is essential for our empirical methodology, and would not be as prevalent in data sets at higher levels of aggregation (such as United Nations data for country's worldwide exports).

There are also *many* instances of "missing" observations, by which we mean that a given commodity is *never* exported by a given country in the sample. If each country had exported each commodity at least once during the sample period, there would be 232,308 entries in our data set. Since we actually have only 88,292 non-zero entries, over 60% of the potential country-commodity observations were censored. This means that even our simple framework in (1) will need to be modified to account from these "missing" observations. The presence of non-random censoring in many large panel data sets makes our techniques more generally applicable.

IV: A Ranking Methodology

IVa: Motivation

[6] As a weighting variable, we use below the presence and/or value of exports subsequent to the year of first export.

Initially suppose that we have a full sample of observations without any "missing" observations, so that each country during the sample exported each good. An example is provided below, with just two countries (Canada and Mexico) and five goods:

Example 1

		Goods:	A	B	C	D	E
Canada:	Exports goods in the order:		1	2	3	4	5
Mexico:	Exports goods in the order:		1	3	4	5	2
	Average of rank orders:		1	2.5	3.5	4.5	3.5

Consider ranking goods. For each country, we have observations on the year of first export of each good. Each country then provides a relative ranking of goods (by their year of first export), as shown in Example 1. With this balanced and complete panel, Kendall and Dickinson (1990, chaps. 6-7) establish the following procedure for determining the best "overall" ranking: average the ranks for each good across countries, and then rank these averages. In the above example, we would therefore assign the goods the ranking A, B, C tied with E, then D. According to this ranking, A would be the least sophisticated, and D is the most sophisticated. Kendall and Dickinson show that this method for determining the overall ranking is optimal in the sense of maximizing a certain objective function (described below).

The difficulty is that there are no known results for determining an optimal ranking when the sample is non-balanced, i.e. when there are "missing" observations. To see this difficulty, suppose that each country exports only a subset of the goods, in the following orders:

Example 2

		Goods:	A	B	C	D	E

Canada:	Exports goods in the order:	.	.	.	1	2
Mexico:	Exports goods in the order:	1	2	3	4	.
	Average of rank orders:	1	2	3	2.5	2

In this case, if we applied the method of averaging the rank-orders over the observations in the sample, then we arrive at the ranking indicated the last line of Example 2: the goods would be ranked A, B tied with E, then D and C. We believe this result is nonsensical, because C has a lower rank than good D for Mexico, while D has a lower rank than E for Canada, so it should not be the case that the ranking of C, D and E is *reversed* in the overall ranking. We conclude from this example that the simple average-ranking method is *not* appropriate when there are missing observations. Since this is a pervasive feature of our data set, we need to develop the statistical techniques to deal with this case.

IVb: Notation

To make all this more formal, we begin with some notation. We tackle the problem of ranking goods, although the logic will be identical for ranking countries.

Selecting from the entire list of goods I={1,N}, let $I_k \subseteq I$ denote the set of goods supplied at any point in the sample by country k. The number of elements in I_k is denoted $N_k \leq N$, where N is the total number of goods (just over 1,400 for the second revision of the 5-digit Standard International Trade Classification). We denote the *rank* of "first year of export to the US" by $x_{ik}(I_k)$ where i denotes the good and k denotes the country. This ranking is done over the goods i, for each country k.[7]

[7] We handle ties in the following way. Arrange the N_k goods (exported by the country at some point in the sample), in order. For the j goods exported in the first year, assign the rank of (j/2). Assign the next j' goods (exported in the second year) (j+j'/2). And so on.

We wish to determine an *"overall" ranking* of the goods $X_i(I)$. We will sometimes want to restrict $X_i(I)$ to be defined only over those goods supplied by country k. This restricted ranking is defined by:

$$X_i(I_k) \equiv \{\text{the ranking of values } X_i(I) \text{ over the set } I_k\}. \qquad (2)$$

With these definitions, we modify (1) slightly to account for "missing" observations:

$$x_{ik}(I_k) = X_i(I_k) \qquad \text{for } \rho_k N_k \text{ observations, and,} \qquad (1a')$$

$$E[x_{ik}(I_k)-(N_k+1)/2][X_i(I_k)-(N_k+1)/2]=0 \quad \text{for the remaining } (1-\rho_k)N_k \text{ observations.} \quad (1b')$$

We will sometimes want to extend $x_{ik}(I_k)$ to cover the entire set of goods, even those not supplied by country k, by imputing where these "missing" goods appear in the ordering for that country. This extended ranking will be denoted by $x_{ik}(I)$.

IVc: Rank Correlation

For any country k, the (Spearman) *rank correlation* between its own ranking $x_{ik}(I_k)$ and the overall ranking $X_i(I)$ is defined as:

$$r_k \equiv \frac{12}{(N_k^3 - N_k)} \sum_{i \in I_k} [X_i(I_k) - \tfrac{1}{2}(N_k+1)][x_{ik}(I_k) - \tfrac{1}{2}(N_k+1)]. \qquad (3)$$

The term $(N_k^3 - N_k)/12$ is the highest possible value for the summation in (3), which is obtained when $x_{ik}(I_k)=X_i(I_k)$ for all observations (and re-ordering the observation so that $x_{ik}(I_k)=X_i(I_k)=i$):[8]

$$\sum_{i=1}^{N_k}[i - \tfrac{1}{2}(N_k + 1)]^2 = \frac{(N_k^3 - N_k)}{12}. \tag{4}$$

Dividing (3) by this term, it can be seen that that the rank correlation lies between -1 and 1.

Let A denote the $\rho_k N_k$ observations for which (1a') holds. Using (1b') and evaluating the expected value of (3), we find that:

$$E(r_k) = \frac{12}{(N_k^3 - N_k)} E\left\{\sum_{i \in A}[X_i(I_k) - \tfrac{1}{2}(N_k + 1)]^2\right\} = \rho_k . \tag{5}$$

To establish this result, note that the expectation in (5) is taken over all possible sets A, of which there are $N_A \equiv \binom{N_k}{\rho_k N_k}$. The summation in (5) contains of $\rho_k N_k$ terms, so writing the expectation in full over all sets A, there will be total of $\rho_k N_k N_A$ terms. Each of these terms will be of the form $[i - (N_k + 1)/2]^2$, where i is an integer within the set A. But by choosing the sets A randomly, it must be that each of the integers $i=1,\ldots,N_k$ appears an *equal number of times*. Thus, each of these integers will appear $\rho_k N_k N_A / N_k = \rho_k N_A$ times within the expected value summation. It follows that the

[8] The equality in (4) can be obtained using the formula $\sum_{i=1}^{N_k} i^2 = \frac{1}{6} N_k (N_k + 1)(2N_k + 1)$, which is reported in elementary mathematics textbooks (and can be proved by induction).

expected value consists of $\rho_k N_A$ summations identical to (4), multiplied by $1/N_A$ (which is the probability of each set A occurring), so that:

$$E\left\{\sum_{i\in A}[X_i(I_k)-\tfrac{1}{2}(N_k+1)]^2\right\}=\left(\frac{\rho_k N_A}{N_A}\right)\frac{(N_k^3-N_k)}{12} \ .$$

Substituting this into (5), we obtain the result $E(r_k)=\rho_k$. *The Spearman rank correlation is an unbiased estimate of the fraction of observations for which country and overall ranks are equal*.[9]

IVd: The Overall Ranking

Kendall and Dickinson (1990) consider the problem of optimal ranking when the number of goods supplied by each country is the same. The objective function that they propose is the *average* of the rank correlations between each country's ranking and the overall ranking. Adopting this same objective function even when the set of goods supplied by each country differs, we can consider choosing the overall ranking $X_i(I)$ to maximize:

$$\sum_{k=1}^{M}\frac{r_k}{M}=\sum_{k=1}^{M}\ \frac{12}{M(N_k^3-N_k)}\sum_{i\in I_k}[X_i(I_k)-\tfrac{1}{2}(N_k+1)][x_{ik}(I_k)-\tfrac{1}{2}(N_k+1)], \qquad (6)$$

[9] A different result is established in Kendall and Dickinson (1990, chaps. 4-5), where the sample rank correlation is shown to be a *biased* (but asymptotically consistent) estimate of the population rank correlation. In our notation, let ρ denote the rank correlation computed as in (3) over the entire population $I=\{1,\ldots,N\}$. Consider taking a random sample of size N_k from that population, and computing the sample rank correlation r_k as in (3). Then taking the expected value over all possible samples, it turns out that $E(r_k)\neq\rho$.

where M is the number of countries. For any choice of $X_i(I)$ the restricted rankings $X_i(I_k)$ are readily computed as in (2), so this is a well-defined optimization problem.

In the case without "missing" observations, so that N_k=N for all k, then Kendall and Dickinson (section 7.10, p. 151) show that (6) is maximized by choosing the overall ranks $X_i(I)$ as the rank of the averages $\frac{1}{M}\sum_{k=1}^{M} x_{ik}(I)$. However, when there are "missing" observations so that N_k<N for some k, then there is no known analytical solution to maximize (6); our objective in this paper is to provide such a solution.

One way to proceed is to maximize this objective function numerically, as was described briefly in our working paper (Feenstra and Rose, 1997). This approach is computationally burdensome, and it is difficult to find the global maximum. Accordingly, in the remainder of the paper we will pursue an alternative approach to determining the overall ranking, suggested by econometric analogies.

IVe: Analogy to Regression

We begin by expressing the country and overall rankings in (1') as a difference from their means of $(N_k+1)/2$, and re-writing the model as:

$$x_{ik}(I_k)-(N_k+1)/2 = \rho_k[X_i(I_k)-(N_k+1)/2] + \varepsilon_{ik} \quad , \ i\in I_k, \tag{7}$$

where,

$$\varepsilon_{ik} \equiv (1-\rho_k)[X_i(I_k)-(N_k+1)/2] \ \text{ for } \rho_k N_k \text{ observations, and,} \tag{8a}$$

$$\equiv [x_{ik}(I_k)-(N_k+1)/2]-\rho_k[X_{ik}(I_k)-(N_k+1)/2] \ \text{ for the remaining } (1-\rho_k)N_k$$
$$\text{observations, with } E[x_{ik}(I_k)-(N_k+1)/2][X_i(I_k)-(N_k+1)/2]=0. \tag{8b}$$

The regression in (7) is identical to the model in (1'), given our definitions of the error terms in (8). Using the standard formula for the least squares estimate of ρ_k, it is immediate that this estimate is identical to the rank correlation coefficient r_k in (3). Since $E(r_k)=\rho_k$ from (5), least squares therefore provides an unbiased estimate of the slope coefficient ρ_k.[10]

Thus, minimizing the sum of squared residuals for (7) yields the rank correlation coefficient as the estimate for ρ_k. The question is whether this minimization problem can also be used to solve for the overall ranking $X_i(I)$. It turns out that this is indeed the case:

<u>Proposition 1</u> *Suppose that when $X_i(I)$ is chosen to maximize (6), the value of (6) is positive. Then the identical values of $X_i(I)$, when chosen along with the coefficient **r**, will minimize the following weighted sum of squared residuals:*

$$\min_{X_i(I),\rho} \sum_{k=1}^{M} \frac{12}{(N_k^3 - N_k)} \sum_{i \in I_k} \left[x_{ik} - \frac{(N_k+1)}{2} - \rho\left(X_i(I_k) - \frac{(N_k+1)}{2} \right) \right]^2 . \tag{9}$$

In other words, there is a very close connection between the objective function in (6) and that obtained by minimizing the weighted sum of squared residuals (SSR) in (9). This SSR is obtained by *pooling* over all goods i and countries k in (7), while imposing a common slope coefficient for ρ. The

[10] This result is obtained despite that fact that the error terms in (8) are clearly correlated with the regressor $X_i(I_k)$ in each observation. However, summing across the observations, it can be shown that $E\left(\sum_{i\in I_k} \varepsilon_{ik} X_i(I_k)\right) = 0$, by using arguments similar to those used in establishing (5). Thus, the regression satisfies the requirement of least squares that the errors are orthogonal to the regressor in expected value.

weights used in (9) when adding up across countries reflect the differing number of observations within each country. The advantage of using the regression-based framework is that it enables us to think about imputing the ranks for goods not supplied by a country, the task that we turn to next.

IVf: Estimation with Censoring

To avoid the difficulties of dealing with an unbalanced panel, there are at least two approaches that can be taken.

Conceptually, we could imagine "shrinking" the panel down to a *balanced* but *incomplete* "Youden" panel. This would be a panel where each country contributed the same number of commodity-observations and each good was observed the same number of times. However, there are two problems with this strategy. First, there is no guarantee that *each* country has exported enough goods to ensure that *all* commodities are covered and could be ranked. Second, much information would be lost, and with it, the benefits of our large data set.[11]

Alternatively, we can "stretch" the panel up to a complete balanced panel by imputing missing observations. We now proceed to that issue.

IVg: Accounting for "Missing" Observations

There are three economic reasons why a country might not have exported a good during the sample.

1. First, the country may have been "too advanced" to export the good during the sample; it had exported the good before the start of the sample and ceased exporting before the start

[11] On a more technical level, it is hard to figure out a scheme for dropping observations randomly that would satisfy the requirements of an incomplete balanced panel.

of the sample. For each country, we denote by $(1,2,\ldots,x_k^{min})$ the ranks of all goods (relative to the entire set I) that are too "unsophisticated" for the country to have produced them in the sample, where x_k^{min} will be estimated.

2. Second, the country may not have been "advanced enough" to export the good during the sample, but will export it at some point in the future. For each country, we will denote by $(x_k^{max}, x_k^{max}+1,\ldots,N)$ the ranks of all good (relative to the entire set I) that are too sophisticated for the country to produce them in the sample, where x_k^{max} will be (implicitly) estimated.[12]

3. Third, trade is driven by other considerations (e.g., factor abundance, increasing returns, trade restrictions, or simply because new goods need not be sophisticated goods); we ignore this possibility throughout, treating it as the alternate hypothesis.

Denote the "filled-in" ranking by $x_{ik}(I)$, which is defined over the entire set of goods. For those goods actually supplied by country k, $x_{ik}(I)$ is related to $x_{ik}(I_k)$ by:

$$x_{ik}(I) = x_{ik}(I_k) + x_k^{min} \quad \text{for } i \in I_k . \tag{10}$$

That is, we take the ranking $x_{ik}(I_k)$, which runs from 1 up to N_k, and increase each of these by the number of goods that we estimate have already been dropped by country k. Since we are supposing

[12] In fact, this possibility turns out to account for many of the missing data in our sample.

that there are no omitted goods "in the middle" of this ranking, given any estimate for x_k^{min}, the corresponding estimate for x_k^{max} would be $x_k^{max} = x_k^{min} + N_k + 1$.

With this preliminary specification of $x_{ik}(I)$, consider choosing x_k^{min} and the overall ranking $X_i(I)$ to minimize the (weighted) SSR of the following pooled regression:

$$[x_{ik}(I)-(N+1)/2] = \rho[X_i(I)-(N+1)/2] + \varepsilon_{ik}, \quad \text{for } i \in I_k,\ k=1,\ldots,M. \tag{11}$$

Note that in (11), the right and left-hand side variables are both defined over the entire set I, so they are expressed relative to their mean values (N+1)/2. Making use of (10), we can rewrite (11) as:

$$[x_{ik}(I_k)-(N+1)/2] = - x_k^{min} + \rho[X_i(I)-(N+1)/2] + \varepsilon_{ik}, \quad \text{for } i \in I_k,\ k=1,\ldots,M. \tag{11'}$$

This is a regression equation in which the left-hand side is data, and the right-hand side variable is simply the overall ranking $X_i(I)$ at some iteration. It follows that $-x_k^{min}$ can be estimated from the various country fixed-effects in this regression.

If the overall ranks $X_i(I)$ were not constrained to be the integers 1,…,N, then it would be possible to estimate them as commodity fixed-effects in (14'). Indeed, these commodity fixed-effects would be chosen to given an average residual of zero for each commodity, so the fixed-effects would equal $\frac{1}{M_i}\sum_{k \in K_i}[x_{ik}(I_k) + x_k^{min}]/\rho$. Then when estimating these as ranks, it seems very plausible that we should simply rank the values of $\frac{1}{M_i}\sum_{k \in K_i}[x_{ik}(I_k) + x_k^{min}]$, provided that the estimate of ρ is positive.

In order to demonstrate the optimality of this procedure, we need to apply certain weights to the observations in (11'). For each good i, let $K_i \subseteq \{1,\ldots,M\}$ denote the set of countries that supply that good sometime during the sample period. We will denote the number of countries within K_i as $M_i \leq M$. Then we will consider weighting the observations in (11') by the *inverse* of M_i so those goods supplied by only a small number of countries receive the largest weight. By this weighting scheme, we achieve a kind of artificial balance in the data set, and obtain the result:

<u>*Proposition 2*</u> *Let $X_i(I)$ denote the overall ranking that, when chosen together with $x_k^{\min}$ and $\boldsymbol{r}$, minimizes the weighted sum of squared residuals:*

$$\min_{X_i(I),\rho} \sum_{i=1}^{N} \sum_{k \in K_i} \frac{1}{M_i} \left[x_{ik}(I_k) - \frac{N+1}{2} + x_k^{\min} - \rho\left(X_i(I) - \frac{N+1}{2} \right) \right]^2 . \tag{12}$$

If the optimal choice for $\boldsymbol{r}$ is positive, then $X_i(I)$ equals the rank of $\frac{1}{M_i} \sum_{k \in K_i} [x_{ik}(I_k) + x_k^{\min}]$.

That is, the optimal overall ranking is simply obtained as the rank of the average country ranking for each good, computed over those countries that actually supply the good. This is a generalization of the Kendall and Dickinson recommendation, derived in the context of an unbalanced panel. It is obtained in the present framework because we have weighted the observations in the unbalanced panel by the inverse of the numbers of times each good appears, which creates a kind of artificial balancing.

In order to compute the averages, however, we must have an estimate of $x_k^{\min}$ for each country. These coefficients can be obtained as the country-fixed effects from the pooled regression

(11), where the left-hand side of (11) is data, and the right-hand side uses the overall ranking $X_i(I)$ at some iteration. To obtain the solution values in Proposition 2, we use the following iterative estimation strategy:

1. Start with a guess for the overall ranking $X_i(I)$.
2. Run (11') over $i \in I_k$ and $k=1,\ldots,M$ to estimate x_k^{min}, applying weights of $1/M_i$ to each observation.
3. Calculate a new optimal overall ranking $X_i(I)$ by averaging values of $(x_{ik}(I_k) + x_k^{min})$ for each commodity over all exporting countries $k \in K_i$, and ranking the results.
4. Return to step 2, until convergence is reached.

This procedure can be illustrated on Example 2 (where Mexico exported the first four goods and Canada the last two). Suppose that we start with the average of the rank orders shown at the bottom of Example 2, which is expressed in rank order as X(I)=(1,2.5,5,4,2.5) since goods B and E are tied. As we argued earlier, this ranking is implausible because goods C, D, and E are each produced before each other by one country, whereas the average ranking has reversed their ordering, with C having the highest rank. Nevertheless, using this as an initial guess for X(I), we can apply the regression in (11') to obtain $x_{can}^{min} = 1.55$ and $x_{mex}^{min} = 0.525$. Then according to step 2, we add these values to the initial rankings for each country, and calculate the new average ranking as:

Example 2 (cont'd)

		A	B	C	D	E
	Goods:	A	B	C	D	E
Canada:	New goods ranking:	.	.	.	2.55	3.55
Mexico:	New goods ranking:	1.525	2.525	3.525	4.525	.
	Average of new ranks:	1.525	2.525	3.525	3.538	3.55

Ranking the averages in the last line, we obtain the new estimate of the overall ranking, $X(I)=(1,2,3,4,5)$. If this is applied again in regression (11'), we obtain the country fixed-effects x^{min}_{can} = 1.80 and x^{min}_{mex} = 0.35, and exactly the same overall ranking $X(I)=(1,2,3,4,5)$. Thus, the procedure has converged and this is the *optimal* ranking.

We think that the optimal ranking makes a good deal of sense in this example, since good C was produced before D for Mexico, and D before E for Canada, so that this ordering should be preserved. We could imagine that Canada has already progressed beyond goods A, B, and C in this example, whereas Mexico has not yet produced good E, and our estimation procedure makes a suitable adjustment for these "missing" goods. Notice that this example meets the criterion outlined in the beginning of this section, whereby the "missing" goods are either too simple or too sophisticated for a country to produce, but are not "in the middle" of the overall ranking. While our methods relies on this hypothesis, other examples can be constructed whereby reasonable overall rankings are obtained even when goods are missing "in the middle".[13]

IVh: Three Observations

We conclude this section with three comments.

First, it is immediate from the proof of Proposition 2 that the values of $X_i(I)$ chosen to minimize (12) also maximize (when $\rho>0$) the weighted correlations,

[13] In our working paper we considered an example where Canada exports the five goods in consecutive order (1,2,3,4,5), whereas Mexico exports only the first and last, in that order. The simple average of the ranks for the two countries give nonsensical results, whereas our methods yields that optimal ranking (1,2,3,4,5) in one iteration.

$$\frac{12}{(N^3 - N)} \sum_{i=1}^{N} \sum_{k \in K_i}^{N} \frac{1}{M_i} [x_{ik}(I_k) + x_k^{min} - \tfrac{1}{2}(N+1)][X_i(I) - \tfrac{1}{2}(N+1)] \ . \tag{13}$$

This objective function can be compared to (6). While the objective functions obviously differ in the weights used across observations, we might expect that the overall ranking that maximizes one also does quite well on the other. We find that this is indeed the case below.

Second, our application allows us to rank goods through the order of export by different countries, with missing data. Re-labeling would allow us to rank investment projects via the ranking of different project attributes (cost of capital, payoff structure, technological flexibility, etc.), or consumption choices via the ranking of various characteristics of goods (a car's legroom, mileage, space, styling, etc.). Thus we think of our methodology as being rather general

Third, we have exposited our methodology as a way to estimate the overall ranking of goods, using *cross-good* variation in the year of first export. We refer to this technique below as one in which we consider goods-rankings to be "primitive." From these goods rankings, countries can be ordered according to the ranks of their exports; countries with more "advanced" exports are more "sophisticated". In Example 2 above, the final ranking X(I)=(1,2,3,4,5) means that Canada is exporting goods with rank (4,5), whereas Mexico is exporting goods with rank (1,2,3,4). The average (unweighted) ranking for these countries is therefore 4.5 and 2.5, respectively, indicating that Canada has a higher "goods based" country ranking.

But the identical methodology can also be used to estimate country rankings as primitive (with appropriate changes to subscripts), using *cross-country* variation in the year of first export. In Example 2 above, only good D provides any information on the cross-country year of first export. If Canada exported this good first, then it would be ranked as "more advanced" than Mexico, treating country rankings as primitive. To continue the example, in this case the "goods based" country ranking would agree with the direct country ranking. (We adopt the convention of assigning a *lower number* to the countries exporting earliest, or with the highest "goods based" ranking.) There does not seem to be any mathematical reason why these two rankings need necessarily cohere, and a counterexample can easily be constructed, as follows.

In example 3, we introduce a third country called China, which exports goods in the order shown. We now interpret the values in the table as the *cardinal year* in which each good is first exported, so that China first exports good A in year 2, good B in year 3, etc. We assume that goods continue to be exported after their first year. If we convert these cardinal years to ranks, and apply the methodology above, it is immediate that the final goods ranking is (1,2,3,4,5). Taking an average of the goods ranking for each country, we see that in year two Canada has an average of 4.5 (since it is exporting goods D and E), Mexico has 1.5 (exporting goods A and B), and China has average rank 1 (exporting only good A). Over time, Canada maintains the lead of 4.5, but China eventually overtakes Mexico, so that by year six China has average rank of 3 (exporting all the goods), while Mexico has 2.5 (exporting only the first four goods). Thus, the "goods-based country ranking" can easily change over time.

Example 3

		A	B	C	D	E
	Goods:	A	B	C	D	E
Canada:	First exports goods in the year:	.	.	.	1	2
Mexico:	First exports goods in the year:	1	2	3	4	.
China:	First exports goods in the year:	2	3	4	5	6
	Goods ranking:	1	2	3	4	5

In contrast, the direct country ranking using our methodology above (where we run regression (11') using fixed effects for each good) will lead to a country ranking that is *fixed* over time. In example 3, it turns out (not surprisingly) that the ranking is Canada as most advanced, followed by Mexico, followed by China. The fact that this ranking is fixed can be interpreted as a limitation of the direct country-ranking approach. Accordingly, in our empirical work we will stress the goods-based country ranking, instead.

V: Empirical Illustration

Va: Estimates of Country Rankings

We estimated both commodity and country rankings using the method outlined in section IV. Table 1 presents three different sets of *country* rankings (these are easier to interpret than comparable *commodity* rankings). The "goods-based" estimates are derived by first estimating primitive *goods* rankings, then averaging these goods-rankings *over the goods actually exported* on a country by country basis, and ranking the resulting averages.[14,15] Table 1 also includes "country-based" estimates

---

[14] We actually use a slightly more general version, allowing the slope of the relationship between the country-specific ranking and the overall ranking to vary by country, as in (7). This generalization results in some computational economies, but insignificantly different results; the overall ranking derived from the pooled regression setup of (8)-(9) has a .999 correlation with that derived from the country-specific regression framework of equation (7).

[15] The list of goods at the "early" end of the list includes: special mail transactions (SITC 93100); coins (89605); antiques nes (89606); furniture (82100); women's outerwear (84300); other wood article manufactures (63599); imitation jewelry (89720); printed books (89211); wood manufactures (63549); and hand paintings etc. (89601). At the

when we treat the country ranking as the primitive overall ranking in equation (11), rather than deriving it from some underlying estimate of a goods ranking. Finally, there is a goods-based set of rankings which only exploits data for *manufacturing* goods (about 60% of our goods are manufactured goods).

Our ranking techniques yield quite sensible results.[16] The top countries (i.e., those with low numbers) are for the most part advanced rich OECD countries; poor countries tend to be clustered at the bottom. Unsurprisingly, Canada is ranked the most sophisticated country overall (ignoring implicit US leadership), followed by the UK, Germany, Japan and France. When manufacturing goods are considered by themselves, the Canadian ranking follows somewhat, and Japan plausibly captures first place. Mexico is ranked higher than one might expect; this may well have to do with either Mexico's proximity to the US or special trade arrangements, and is a topic worthy of further investigation.[17] There is strong evidence of intuitive orderings of countries and commodities, consistent with the product

---

other end of the spectrum are: vinyl chloride (51131); mechanically propelled cars (79130); wine lees (8194); linseed (22340); methacrylic acid (51373); slag etc. from iron (27861); natural sodium nitrate (27120); paper pulp filter-blocks (64196); tin tubes (68724); uranium (68800); and oxy-fnct aldehyde derivatives (51622).

[16] Our iterative technique converges quite quickly. Our default specification converges after three iterations. We have also experimented with random starting values, and our procedure still converges to the same final estimates quickly. Also, the rank correlation coefficients between this overall ranking and the individual country rankings turn out to be positive for essentially all the goods in our sample (well over 95%), and significantly for most.

[17] Mexico's ranking may also reflect the "806/807" program or re-exports. Generally, our results are affected by FDI activity of U.S. firms (such as the *maquiladora* firms in Mexico, exporting back to the U.S.), or of firms from other countries (such as an Asian firm setting up in Mexico, and exporting to the U.S.). Trade flows motivated by FDI activity may well lead a country to be ranked as more technologically advanced than would be expected. This may also account for the high rank received by China, above some of the newly industrialized Asian countries. We do not yet have a convenient method for estimating the statistical significance of country rankings.

cycle hypothesis.[18] This is true despite the fact that we model *all* trade as being driven by the product-cycle, implicitly ignoring alternative theories such as those which rely on economies of scale, distance, or factor endowments.

It is striking that the country- and goods-based rankings are so similar. That is, when ordering countries, it does not matter much whether we treat goods-rankings or country-rankings as primitive. It is comforting to note that the two rankings are closely related; there is no reason why the country- and goods-based rankings need necessarily deliver similar results for any *statistical* reason. Further, the two different applications rely on different economic assumptions, namely whether countries or goods can be ranked in terms of sophistication.

Figure 1 plots the country rankings (derived from goods rankings by averaging the latter over the set of goods actually exported in any given year) on a year by year basis for sixteen countries. Each of the graphs is a time-series plot of country ranking from 1972 through 1994.

Vb: Sensitivity Analysis

To check the sensitivity of our results, we have also estimated rankings for a number of perturbations to our basic methodology. First, we repeated our analysis but weighed each country (in the Kendall estimation procedure) by the number of individual goods it exported in the sample. Thus countries with a large number of exports were given more weight in determining the overall ranking. Second, we estimated separate country rankings for the first and last halves of the sample. We did this by weighting the goods-rankings for each country by: 1) only the goods the country first exported

---

[18] For instance, all the country-specific correlations between the overall ranking and the country-specific rankings in (7) are positive, and significantly so (using conventional or bootstrapped standard errors).

before 1985; and 2) only goods first exported by the country after 1984. Next, we adjusted our x_k^{min} estimates (in step 2 of our procedure) by regressing our initial x_k^{min} estimates on four standard gravity regressors (the log of distance, the log of real GDP per capita and dummy variables for common land borders and common language). Fourth, we dropped goods-country observations where the good was exported by the country for only a single year. Finally, we repeated our analysis using bilateral American *export* data (dis-aggregated to the same 5-digit SITC level) instead of American import data.

Our orderings appear to be quite robust to all these perturbations in our methodology, at least for the countries at the top of the rankings. The rankings are somewhat sensitive for countries towards the bottom of the rankings. This comes as no surprise to us; the poor countries that tend to be ranked towards the bottom provide relatively few exports to the United States, and are accordingly difficult to rank precisely.[19] Still, the different rankings are quite highly correlated overall. Spearman rank correlations between the rankings are quite high and statistically significant.[20] The working paper version contains more evidence.

[19] Indeed, there is a strong negative correlation between the number of goods a country exports and its ranking. This comes as no real surprise to us; rich countries tend to be open and diversified exporters, while poor countries tend to be closed and specialized exporters. Sachs and Warner provide evidence on the linkages between openness and growth; Hall and Jones provide evidence on the linkage from openness to productivity.

[20] In passing, we note that the dis-aggregated nature of the data seems critical for the actual estimation of these rankings. When we aggregated our data to the 2-digit SITC level, over a quarter of our countries showed literally no dispersion in "year of first export" across commodities; all commodities exported were exported first in the same year. But manifestly dispersion can be found at finer levels of dis-aggregation; this dispersion also appears to be systematic and meaningful.

Vc: Linking Country Rankings with Aggregate Variables

Our country rankings appear to be robustly estimated, stable and sensible. Derived as they are from dis-aggregated bilateral trade flows, there is no obvious reason why they need necessarily be linked to macroeconomic phenomena. Are they?

Figure 2 presents a simple bivariate scatterplot of country rankings (derived treating country rankings as primitive) with the growth rate of real GDP per capita (taken from the Penn World Table). A non-parametric data smoother has been included to "connect the dots". An economically and statistically significant negative correlation appears. Sophisticated countries (which export first and consequently have low numerical rankings) tend to have high growth rates of real GDP per capita. Of course, the causal interpretation of this finding is unclear.

To pursue this matter further, we have merged our data with the Barro-Lee data on economic growth and added our country-rankings to standard cross-country growth equations. As is apparent from Table 2, our (ordinal) country ranking appears to be significantly negatively related to the growth rate of real GDP per capita.[21] We have conditioned growth rates on the share of GDP devoted to investment (one of the few variables consistently associated with growth) as well as the initial level of GDP; we have also added other regressors, including measures of human capital, political stability, and other proxies for openness. Partial correlations between growth rates and country rankings, like simple correlations, are significant and negative. Countries which export sooner tend to grow faster.

Our rankings are not simply highly correlated with the *growth rates* of output; it turns out that they are also correlated with the *levels* of economic activity. Figure 3 is a scatterplot of our country

[21] The same is true of our baseline orderings, treating goods-rankings as primitive.

rankings and the level of 1985 real GDP per capita; the latter is a standard datum used by economists to rank countries. A strongly negative correlation emerges clearly in the graph. High-income countries tend to have low ("advanced") rankings.

The same negative correlation characterizes the relationship between our country rankings and the *level of total factor productivity* (another standard metric of country sophistication). We have added our rankings to the Hall and Jones (1996) productivity data set, and found that our country ranking is significantly negatively related to productivity. This is true both unconditionally, and when the effects of the Hall-Jones factors have been taken into account. The latter include such measures as the fraction of the populace speaking an international language, the country's latitude, government intervention in the economy, and other measures (including the Sachs-Warner openness indicator) that Hall and Jones found important in determining total factor productivity differentials across countries. Figure 4 contains the graphical evidence. It contains four scatterplots, corresponding to two measures of country rankings (derived from primitive orderings of both goods and countries) graphed against two measures of productivity (raw, and after the effects of the Hall-Jones variables have been "partialled out"). Table 3 contains the corresponding regression evidence.[22]

VI: Summary and Topics for Future Research

Ranking objects is a pervasive feature of everyday life. In this paper, we have developed a general methodology for ranking objects. Our techniques takes advantage of the large panel data sets

[22] The distance between Chicago and the capital of the other countries is slightly positively correlated with our country rankings. The relationships between our rankings and macroeconomic phenomena like growth and TFP are insignificantly affected if we condition them on distance.

that are frequently available to produce such rankings. Our technique also takes account of non-randomly missing data.

We applied our technique to a large panel data set of international trade data, and rank both countries and commodities. Consistent with the product cycle theory of international trade, we find sensible, insensitive rankings. We also investigate the relationships between our country rankings and macroeconomic phenomena such as national growth-rates and productivity levels. Our rankings turn out to be closely linked with both productivity levels and growth rates. Countries which are "advanced" in the sense that they export commodities early, also tend to have both high productivity levels and fast growth rates.

Our country rankings are derived from a complicated semi-parametric estimation procedure using only dis-aggregated international trade data. It is both reassuring and promising that they turn out to be related to important macroeconomic phenomena. Still, while this evidence is suggestive and consistent with the product cycle, it clearly does not constitute a test of the product cycle theory against an explicit alternative hypothesis.

Future research could follow a number of different directions. Does government policy (e.g., industrial policy) affect rankings? Is there causality in the reverse direction? Do our rankings depend on the fact that our data covers *American* imports? Our rankings should be similar when derived from the imports of *any* country (or indeed exports from one country to *any other* country), even if trade volumes differ systematically by country (the "gravity" effect).[23]

---

[23] It is interesting to note the negative relationship between trade volume and country ranking. It is also interesting to note that the missing observations are disproportionately goods which are classified as "sophisticated goods" from "unsophisticated countries"; "sophisticated countries" tend to export many goods, while poorer countries tend

only to have exported simple goods during the sample. Our country rankings do not appear to be due to country size; when we split the Canadian data into two parts randomly, each part was ranked with approximately the same ranking as Canada.

References

Deardorff, Alan V. (1984) "Testing Trade Theories and Predicting Trade Flows" in *The Handbook of International Economics vol. III* (R. Jones and P. Kenen, eds.), Amsterdam: North-Holland

Feenstra, Robert C. (1996) "NBER Trade Database. Disk 1: U.S. Imports, 1972-1994," NBER Working Paper no. 5515.

Feenstra, Robert C. and Andrew K. Rose (1997) "Putting Things in Order: Patterns of Trade Dynamics and Growth," NBER Working Paper no. 5975, March.

Frankel, Jeffrey A. and David Romer, 1996, "Trade and Growth: An Empirical Investigation," NBER Working Paper no. 5476.

Grossman, Gene M. and Elhanan Helpman (1991) *Innovation and Growth in the Global Economy*, Cambridge: MIT Press.

Hall, Robert and Charles Jones (1996) "The Productivity of Nations" NBER Working Paper no. 5812.

Kendall, Maurice and Jean Dickinson (1990) *Rank Correlation Methods*, fifth edition, London: Arnold.

Leamer, Edward E. and James Levinsohn (1995) “International Trade Theory: The Evidence” in *The Handbook of International Economics vol. III* (G. Grossman and K. Rogoff, eds.), Amsterdam: North-Holland.

Sachs, Jeffrey and Andrew Warner (1995) “Economic Reform and the Process of Global Integration” *Brookings Papers on Economic Activity.*

Vernon, Raymond (1966) “International Investment and International Trade in the Product Cycle” *The Quarterly Journal of Economics* LXXX, 190-207.

Table 1: Country Rankings

	Goods-Based	Country-Based	Man'g Goods
CANADA	1	1	3
UK	2	2	4
GERMANY	3	3	2
JAPAN	4	4	1
FRANCE	5	5	5
MEXICO	6	7	9
NETHLDS	7	8	8
ITALY	8	6	6
BELG/LUX	9	9	11
SWITZLD	10	10	7
CHINA	11	25	17
SWEDEN	12	11	13
TAIWAN	13	13	12
SPAIN	14	12	14
BRAZIL	15	16	15
AUSTRAL	16	18	20
HONG KNG	17	14	19
S KOREA	18	20	18
DENMARK	19	15	21
AUSTRIA	20	17	16
S AFRICA	21	27	25
ISRAEL	22	21	22
NORWAY	23	22	23
INDIA	24	19	24
IRELAND	25	23	26
FINLAND	26	26	27
ARGENT	27	28	31
SINGAPR	28	30	28
USSR	29	43	34
VENEZ	30	39	29
UNKNOWN	31	34	10
NEW ZEAL	32	37	33
THAILAND	33	35	30
PHIL	34	31	32
PORTUGAL	35	24	35
CHILE	36	58	41
COLOMBIA	37	29	40
POLAND	38	32	39
DOM REP	39	38	44
MALAYSIA	40	42	38
YUGOSLAV	41	33	37
CZECHO	42	44	36
E GERMAN	43	41	43
GREECE	44	36	42
PERU	45	47	46
HUNGARY	46	50	45
INDONES	47	53	48
TURKEY	48	51	49
ST K NEV	49	70	47
COST RICA	50	55	50
ROMANIA	51	46	51
JAMAICA	52	45	56
GUATMALA	53	48	54
PANAMA	54	54	52
ECUADOR	55	57	57
SD ARAB	56	85	53
EGYPT	57	76	62
PAKISTAN	58	49	55
NIGER	59	113	59
TRINIDAD	60	64	71
HONDURA	61	65	69
HAITI	62	40	61
MOROCCO	63	63	67
N ANTIL	64	67	65
KENYA	65	68	60
BAHAMAS	66	62	76
BULGARIA	67	87	68
SALVADR	68	56	66
ICELAND	69	75	63
URUGUAY	70	66	73
MRITIUS	71	111	58
MACAU	72	71	64
IVY CST	73	103	80
SRI LKA	74	74	79
UAE	75	99	72
JORDON	76	115	74
IRAN	77	52	81
GABON	78	127	93
LEBANON	79	60	75
GILBRALT	80	61	70
KIRIBATI	81	129	95
S YEMEN	82	80	107
GUYANA	83	79	108
NIGERIA	84	82	91
BARBADO	85	73	78
MOZAMBQ	86	96	89
NICARAGA	87	59	101
CYPRUS	88	84	84
BOLIVIA	89	78	85
MONGOLA	90	152	145
SURINAM	91	109	82
ZIMBABWE	92	125	90
GUINEA	93	146	88
NEW CAL	94	102	106
BAHRAIN	95	133	96
BELIZE	96	93	104
BERMUDA	97	94	83
GHANA	98	81	100
MALI	99	120	86
SEYCHEL	100	138	77
CAMERN	101	116	109
ALGERIA	102	105	116
TUNISIA	103	100	94
SYRIA	104	69	110
GUADLPE	105	88	98
FIJI	106	112	105
ZAIRE	107	104	134
BNGLDSH	108	91	87
ALBANIA	109	126	112
LIBERIA	110	95	115
NEPAL	111	90	92
AFGHAN	112	77	99
MALTA	113	72	97
SENEGAL	114	131	102
PARAGUA	115	83	124
BURMA	116	121	103
SIER LN	117	119	136
G BISAU	118	134	120
MADAGAS	119	107	125
OMAN	120	130	117
KUWAIT	121	117	111
CONGO	122	122	127
QATAR	123	141	137
FR GUIAN	124	136	118
N KOREA	125	114	133
FR IND O	126	156	114
TANZANIA	127	89	147
LIBYA	128	101	146
SUDAN	129	143	126
GREENLD	130	92	121
US NES	131	147	131
NEW GUIN	132	106	142
LAO	133	108	144
ZAMBIA	134	123	129
N YEMEN	135	142	119
ANGOLA	136	110	143
S HELNA	137	128	113
SP MQEL	138	148	132
MALAWI	139	140	151
VIETNAM	140	86	140
UGANDA	141	118	139
ASIA NES	142	145	128
SAMOA	143	137	152
SOMALIA	144	155	123
IRAQ	145	97	141
GAMBIA	146	153	122
MAURITN	147	132	135
BURUNDI	148	135	158
CAR	149	154	153
TOGO	150	150	149
BURKINA	151	139	130
RWANDA	152	144	138
ETHIOPIA	153	98	150
BENIN	154	124	154
CHAD	155	157	156
CAMBOD	156	151	148
FALK IS	157	159	155
DJIBOUTI	158	158	157
CUBA	159	149	160
EQ GNEA	160	160	159

Table 2: Cross-Country Growth Equations

Ranking(x100)	-.02 (3.4)	-.02 (3.8)
Log of 1960 GDP (x100)	-1.0(3.8)	-1.2 (3.4)
Investment/GDP	.15 (5.9)	.13 (4.5)
Average Years of School in the Population over 25 (1985)		.00 (.3)
Percentage of the Population without Schooling (1985)		.00 (.3)
Assassinations per million population (1985)		-.01 (1.4)
Average Annual Number of Revolutions and Coups		.002 (.3)
Exports/GDP		.02 (1.4)
Own Import-Weighted Tariffs on Intermediate Inputs and Capital Goods		-.07 (3.9)
Measure of Tariff Restriction		.66 (4.4)
Black Market Premium (1985)		-.001 (1.6)
Observations	82	62
R^2	.50	.61

Country Rankings estimated treating countries as primitive.
Absolute value of t-statistics in parentheses.
OLS with an unreported constant.
A lower numerical country ranking corresponds to a more "sophisticated" country.

Table 3: Hall-Jones Cross-Country Productivity Equations

Economic Organization	.02 (.03)	.02 (.03)	.03 (.03)
Openness	.55 (.15)	.53 (.15)	.50 (.14)
GADP	.88 (.27)	.21 (.30)	.31 (.28)
International Language	.55 (.09)	.43 (.10)	.46 (.09)
Latitude	.003 (.002)	.002 (.002)	.002 (.002)
Country-Ranking: Goods-Based		-.005 (.001)	
Country-Ranking: Country-Based			-.005 (.001)
Observations	122	122	122
R^2	.57	.62	.62
RMSE	.432	.404	.407

Regressand is log-level of total factor productivity.
Huber-consistent standard errors in parentheses.

OLS with an unreported constant.

Figure 1: Country Rankings over Time

A lower numerical country ranking corresponds to a more "sophisticated" country.

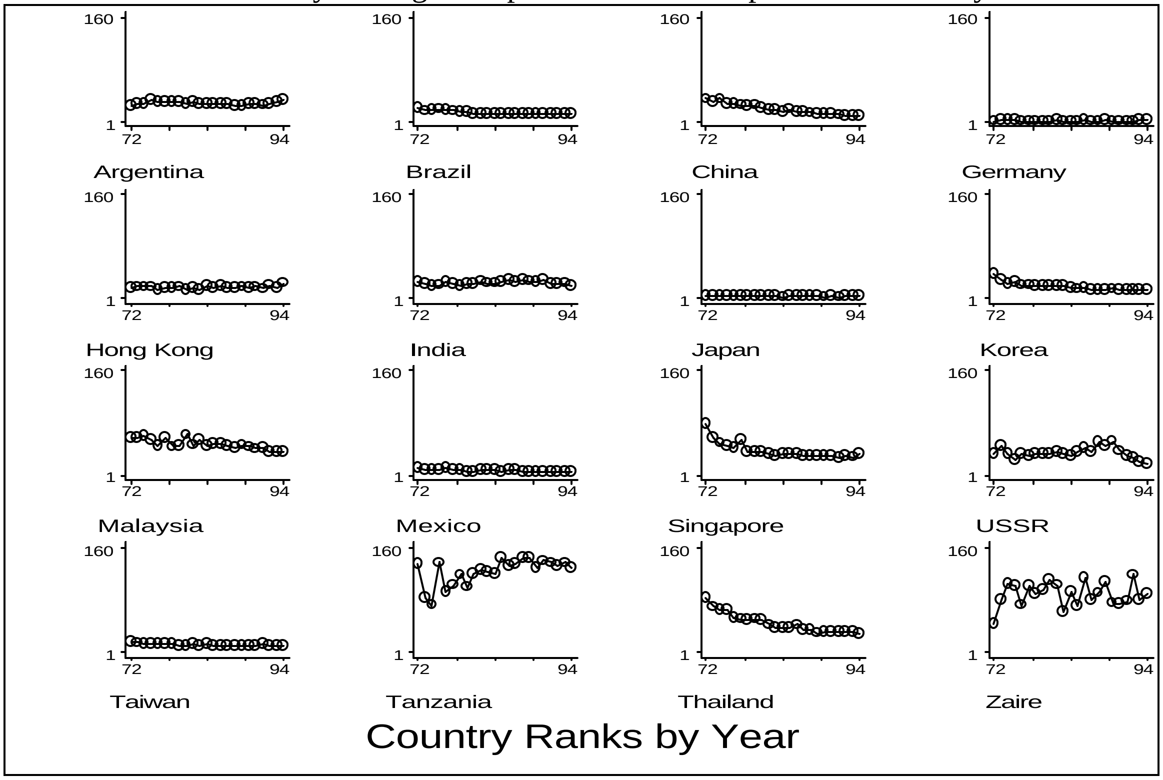


Figure 2: Country Rankings and Real GDP per capita Growth

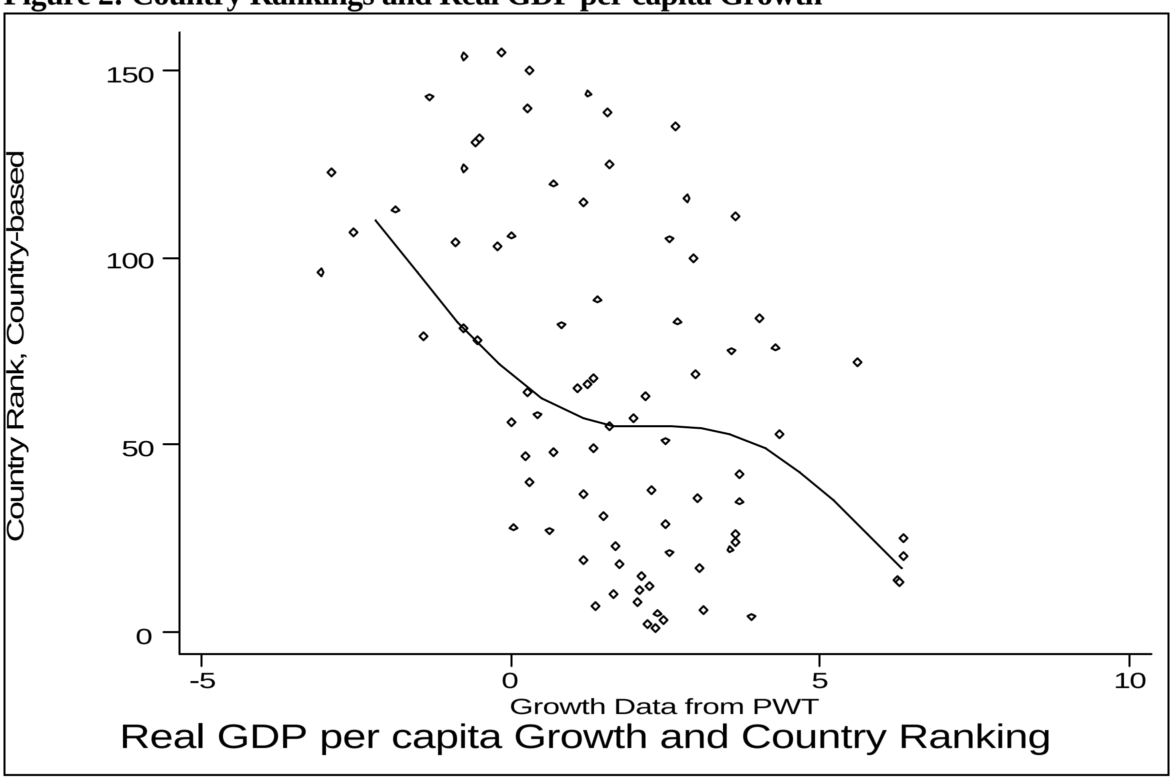


Figure 3: Country Rankings and Real GDP per capita in 1985

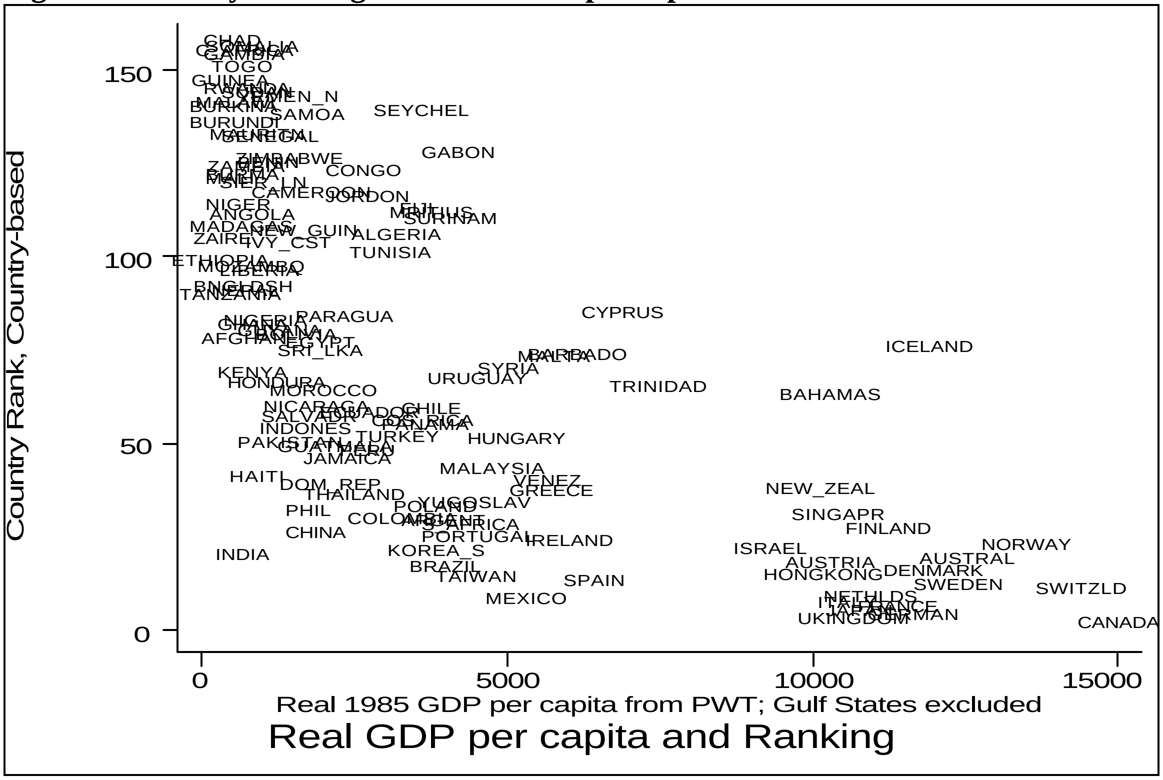


Figure 4: Country Rankings and the Log-Level of Total Factor Productivity