PUTTING THINGS IN ORDER: PATTERNS OF TRADE DYNAMICS AND GROWTH

ROBERT C. FEENSTRA AND ANDREW K. ROSE

Working Paper Series #97-14

Department of Economics
University of California
Davis, California 95616-8578

Putting Things in Order: Patterns of Trade Dynamics and Growth

Robert C. Feenstra and Andrew K. Rose*

Working Paper Series No. 97-14
April, 1997

*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 Professor and Chair of 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 Michael Burda, Jeff Frankel, Gene Grossman, Ben Hermalin, Rich Lyons and Lars Svensson for comments and encouragement.

Abstract

We develop a procedure to rank-order countries and commodities using dis-aggregated American imports data. We find strong evidence that both countries and commodities can be ranked, consistent with the "product cycle" hypothesis. 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 semi-parametric 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. For instance, our country rankings derived from dis-aggregated trade data, turn out to be highly correlated with macroeconomic phenomena such as national productivity levels and growth rates.

I: Introduction

This paper has two goals. The first is methodological; we develop techniques to estimate rank-orderings from large dis-aggregated panel data sets. We apply these techniques to 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. Country-rankings can be estimated in a comparable fashion. Our methodology accounts for the fact that observations may be missing in a non-random fashion.

The second objective of this paper is to ask whether countries can be ranked in a meaningful way using trade data. Since some theories of international trade suggest such rankings, we are interested in whether countries can actually be ordered in a systematic and sensible way. We find that they can, providing evidence consistent with the product cycle hypothesis. We also investigate the relationships between our country rankings and macroeconomic phenomena such as national growth-rates and productivity levels. Our rankings (estimated solely with dis-aggregated bilateral trade data) 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 empirical methods are motivated by the trade and growth models in Grossman and Helpman (1991), which we briefly outline below in section II. After a discussion of our data set, in section IV we develop a statistical methodology to estimate rankings. We then apply our techniques in section V, which contains a discussion of empirical results. We estimate and analyze rankings for both commodities and countries, and link these to per-capital productivity levels and growth rates. We conclude with a brief discussion of ways in which our analysis can be extended. Proofs are provided in the Appendix.

II: Economic Framework

Our work lies within the framework of the "product cycle," due to Vernon (1966). 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.

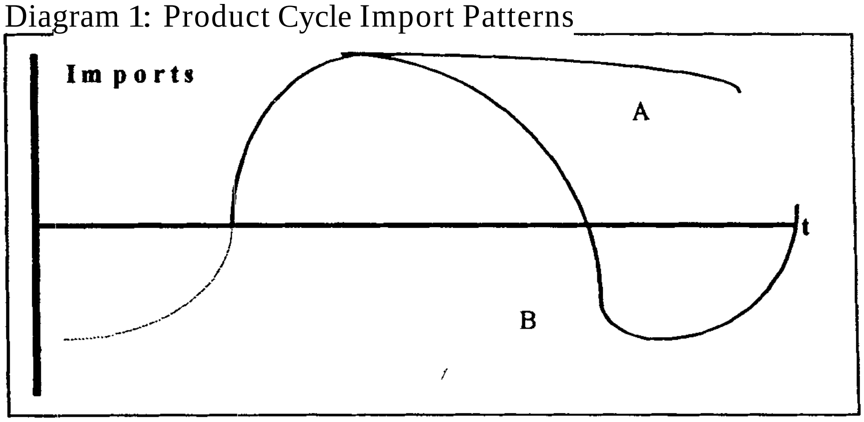
While the theoretical work often predicts that trade will have dynamic effects over and above the static gains from specialization, the empirical evidence points to only a limited role for trade in influencing growth. For instance, Frankel and Romer (1996) have introduced "openness" (measured by trade relative to GDP) into the extended Solow growth model, and found that while its estimated impact on growth is positive and significantly different from zero, it is sometimes just barely so. In this paper we will provide stronger evidence of the link between trade and growth, using a new measure of export orientation. Rather than looking at aggregate measures of openness, such as those considered by Frankel and Romer (1996) or Sachs and Warner (1995), we instead consider the dis-aggregate trade patterns of countries, and how they evolve over time.

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. In this model, once Southern countries begin to export a good to the North, Northern production ceases. This is illustrated as case "A" in Diagram 1.

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 recumng 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.

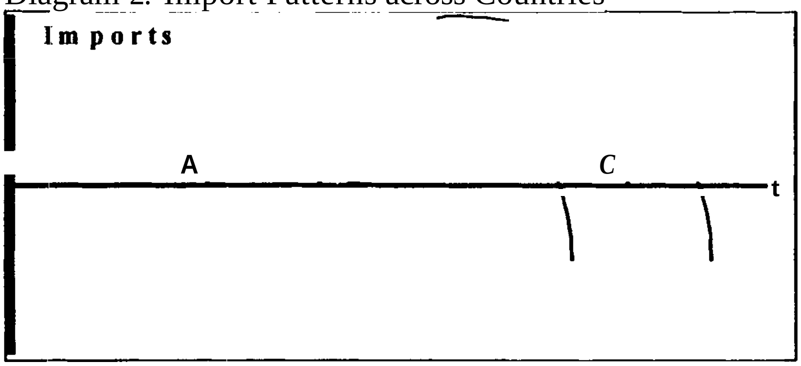
Diagram 1: Product Cycle Import Patterns



[1] In future work, we plan to check the sensitivity of the use of "the year of first import." A number of perturbations are natural. First, one could use the first year that imports reach a given size either in terms of dollars, or as a fraction of the (partner-country) export base. Second, one could use the first time cumulative imports reach a given size.

The intuition behind our technique for rank-ordering both commodities and countries is simple. We assume that 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 provide 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.

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 fiom 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 fiom, 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).

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 all possible values for x_{ik}.[2]

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}. After briefly describing our data in sedtion 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 dataset. 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

Much of our methodology is driven by features of typical panel data sets. 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 162 countries and other geographical jurisdictions (which we refer to as "countries" for simplicity); and 1,434 commodities ("goods").[3] For each good and each country, we initially use only the *first year* of *export* to the United States.[4] 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 the case for data sets at higher levels of aggregation (such as United Nations data for country's world-wide 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

---

[2] 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.

[3] 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).

need to be modified to account from these "missing" observations. But the presence of non-random censoring in many large panel data sets makes our techniques more generally applicable.

IV: A Methodology to Rank Countries and Commodities

IVa: Motivation

Initially suppose that we have a full sample of observations without any "missing" observations, so that each good was exported by each country during the sample. 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, and 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).

---

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

The difficulty is that there are no known results for determining an optimal ranking when tht: sample is non-balanced, **i.e.** when there are "missing" observations. To see this **difficulty, suppose** that Mexico exports only the first and last goods:

Example 2

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

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, C, E, D. We believe this result is nonsensical, because E has a higher rank good D for Canada, and no comparative information is provided for Mexico, so it should not be the case that the ranking of E and D 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 $I=\{1,\ldots,N\}$, 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.[5]

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:

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

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

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

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

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:

[5] **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.**

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

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 $x_{ik}(I_k)=X_i(I_k)=i$):[6].

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

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:

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

To establish this result, note that the expectation in (5) is taken over all possible sets A, of which there are $N_A = \binom{N}{\rho_k N}$. The summation in (5) contains $\rho_k N_k$ 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

---

[6] **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).**

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 expected value consists of $\rho_k N_A$ summations identical to (5), divided by N_A (which is the probability of each set A occuring), so that:

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

Substituting this into (5), we obtain the result $E(r_k) = \rho_k$. That is, ***the Spearman rank correlation is an unbiased estimate of the fraction of observations for which the country and overall ranks are equal.***[7]

IVd: Numerical Estimation of 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:

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

[7] 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,...,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 in 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 possibility is to numerically maximize this objective function. To do so, first simplify the objective function in (6) as:

$$\sum_{k=1}^{M}\frac{r_k}{M} = \sum_{k=1}^{M} \frac{12}{M(N_k^3 - N_k)}\left[\sum_{i\in I_k} X_i(I_k)x_{ik}(I_k) - \frac{1}{4}N_k^2(N_k+1)^2\right], \tag{7}$$

where this line follows from (6) because $X_i(I_k)$ and $x_{ik}(I_k)$ both sum to $N_k(N_k+1)/2$.

Suppose that the goods have been re-numbered by increasing rank, so that $X_i(I)=i$, and consider reversing the rank of goods i-1 and i within the overall ranking $X_i(I)$. This will have an impact on the restricted ranking $X_i(I_k)$ if and only if both these goods are supplied by country k. Def ne the indicator variable,

$$\delta_{ik} = \begin{cases} 1 & \text{if } i-1 \in I_k \text{ and } i \in I_k \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

Then it is immediate that the change in the objective function (7) from reversing the rank of goods i-1 and i within the overall ranking $X_i(I)$ is simply equal to:

$$\Delta_i = \sum_{k=1}^{M} \frac{12}{M(N_k^3 - N_k)} \delta_{ik}[x_{i-1,k}(I_k) - x_{ik}(I_k)]. \tag{9}$$

If $\Delta_i>0$ then the objective function is increased by reversing the rank of i-1 and i. Suppose we do so, and then re-number all the goods by that new ranking so that $X_i(I)=i$. Then for each adjoining pair of goods, the change in the objective can again be computed as in (9), rind whenever $\Delta_i>0$ then the position of goods i-1 and i can be reversed and the set of goods re-numbered. When it is no longer the case that $\Delta_i>0$ for any adjoining pair of goods, the algorithm has converged to a ranking $X_i(I)$.

Thus, from some initial value for the overall ranking $X_i(I)$, it is easy to compute the (discrete) change in the objective function from swapping the position of two adjoining goods in the overall ranking: whenever this change is positive, the swap should be made. We call this the "numerical approach" to maximizing the objective function (6), and illustrate some results from it in section VI.

One difficultywith the numerical approach is that it may not enough to just check whether the position of all adjoining goods in the ranking should be swapped; it also seems necessary to check whether the position of ***any*** two goods should be reversed. With about 1,400 goods, this would mean that one would need to check about $1{,}400^2$ possibilities on each iteration. It is not computationally feasible to perform all these comparisons, and our program to implement the numerical maximization is limited to comparing the ***ten*** adjoining goods for each product on each iteration. For these reasons, we cannot establish that our numerical approach necessarily reaches a *global* maximum of the objective function. Accordingly, in the remainder of the paper we will pursue an alternative approach to determining the overall ranking, suggest# by econometric analogies.

IVe: Analogy to a 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:

$$x_{ik}(I_k)-(N_k+1)/2 = \rho_k[X_i(I_k)-(N_k+1)/2] + \varepsilon_{ik}, \quad i \in I_k, \tag{10}$$

where,

$$\varepsilon_{ik} = (1-\rho_k)Xi(I_k) \qquad \text{for } \rho_k N_k \text{ observations, and,} \tag{11a}$$

$$= x_{ik}(I_k)-\rho_k X_{ik}(I_k) \qquad \text{for the remaining } (1-\rho_k)N_k \text{ observations,}$$

$$\text{with } E[x_{ik}-(N_k+1)/2][X_i-(N_k+1)/2]=0. \tag{11b}$$

The regression in (10) is identical to the model in (1'), given our definitions of the error terms in (11). 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.[8]

[8] This result is obtained despite that fact that the error terms in (11) 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.

Thus, minimizing the sum of squared residuals for (10) yields the rank correlation coefficient as the estimate for ρ_k. The question is whether this minimization problem also be used to solve for the overall ranking $X_i(I)$. It turns out that this is indeed the case:

Proposition 1 ***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 ρ, will minimize the following weighted sum of squared residuals:***

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

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 (12). This SSR is obtained by *pooling* over all goods i and countries k in (10), while imposing a common slope coefficient for p. The weights used in (12) when adding up across countries reflect the differing number of observations within each country. This result does not necessarily provide an easier way to obtain the optimal overall ranking, since the numerical difficulties that we noted in maximizing (3) apply equally well to minimizing (8). Rather, 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, which 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.[9]

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

Ng: 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 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

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

are too sophisticated for the country to produce them in the sample, where x_k^{max} will be (implicitly) estimated.

3. Third, trade is driven by other considerations (e.g., factor abundance); we ignore this possibility throughout.

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:

$$x_{ik}(I) = x_{ik}(I_k) + x_k^{min} \quad \text{for } i \in I_k . \tag{13}$$

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 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:

$$[x_{ik}(I)-(N+1)/2] = \rho[X_i(I)-(N+1)/2] + \varepsilon_{ik} , \quad \text{for } i \in I_k, k=1,\ldots,M. \tag{14}$$

Note that in (14), 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 (13), we can rewrite (14) as:

$$[x_{ik}(I_k)-(N+1)/2] = -x_k^{min} + \rho[X_i(I)-(N+1)/2] + \varepsilon_{ik} , \quad \text{for } i \in I_k, k=1,\ldots,M. \tag{14'}$$

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 (14'). 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 (14') by the ***inverse*** of M_i so that 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 **dataset**, 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 ρ, minimizes the weighted sum of squared residuals:**

$$\min_{X_i(I),\rho} \sum_{i=1}^{N} \sum_{k \in K_i} \frac{1}{M_i}\left[x_{ik}(I_k) - \frac{N+1}{2} + x_k^{min} - \rho\left(X_i(I) - \frac{N+1}{2}\right)\right]^2 . \qquad (15)$$

If the optimal choice for ρ 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 the 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 (14), where the left-hand side of (14) 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 (14') 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 Canada exported all five goods in consecutive order, but Mexico exported only goods A and E). Given that both countries exported A before E, it is plausible to specify an initial ranking of the five X(I)=(1,2,3,4,5). When regression (14') is run over the observations for Canada and Mexico, applying the appropriate weights, we obtain the values $x_{mex}^{min}=1.5$ and $\rho=0.7$. Then according to step 2, we add this value for x_{mex}^{min} to the initial rankings $x_{mex}(I_{mex})=(1,2)$ for Mexico, and calculate the new average ranking as:

Example 3

	Goods:	A	B	C	D	E
Canada:	New goods ranking:	1	2	3	4	5
Mexico:	New goods ranking:	2.5	.			3.5
	Average of new ranks:	1.75	2	3	4	4.25

Ranking the averages in the last line, we obtain the new estimate of the overall ranking, X(I)=(1,2,3,4,5). This is the same as its initial value, so the procedure has converged, and this is the optimal ranking.[10]

IVh: Three Observations

We conclude this section with three observations about the "regression-based method in steps 1-4.

First, it is immediate fiom the proof of Proposition 2 that the values of $X_i(I)$ chosen to minimize (15) also maximize (when $\rho>0$) the weighted correlations,

$$\frac{12}{(N^3-N)}\sum_{i=1}^{N}\sum_{k\in K_i}^{N}\frac{1}{M_i}[x_{ik}(I_k)+x_k^{min}-\tfrac{1}{2}(N+1)][X_i(I)-\tfrac{1}{2}(N+1)] \; . \qquad (16)$$

[10] It can be shown that this overall ranking is also obtained for other starting values.

This objective function can be compared to (6), which we attempt to maximize with our "numerical method." 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 procedure can be viewed as an application of the "EM" algorithm. The two equations are:

$$x_{ik}(I^j) = \alpha_k + \beta X_i^j(I) + \varepsilon_{ik}, \qquad \text{for } i \in I_k \qquad \text{at iteration j}$$

$$X_i^{j+1}(I) = f(x_{ik}(I^j)) + \text{error}$$

One first takes the ***expectation*** when filling in the missing values (in step 2 above); then ***maximizes*** (in step *3*).

Third, we have outlined this methodology as a way to estimate an 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. But it should be clear that an identical methodology can be used to estimate country rankings as primitive (with appropriate changes to subscripts), using *cross-country* variation in the year of first export. In our empirical work, we pursue both schemes and compare estimates derived using the different techniques.

V: Empirical Results

Va: Estimates of Country Rankings

We have estimated both commodity and country rankings using the regression-based method outlined in section IV.

Table 1 presents a number of different sets of *country* **rankings**; these are easier to interpret than comparable commodity **rankings**. We derive these estimates by first estimating primitive *goods* **rankings** using the methodology outlined above.[11] We then average these goods-**rankings** ***over the good!actually exported*** on a country by country basis, and rank the resulting **averages**. We refer to this as our "baseline goods-based **methodology**.[12]

Our baseline methodology yields quite sensible **results**.[13] The top countries 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 (ignoring implicit US leadership), followed by the UK, Germany, Japan and France. Mexico is ranked higher (at position 6) 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.[14] Overall, there is strong evidence of intuitively reasonable orderings of both countries and commodities, consistent with the product cycle hypothesis.[15]

---

[11] 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 **(10).** This generalization results in some computational economies, but **insignificantly different** results; the overall ranking derived **from** the pooled **regression** setup of (14)-(15) has a .999 correlation with that derived from the country-specific regression **framework** of equation **(10).**

[12] The list of goods at the "early" end of the list includes: special mail transactions **(SITC 93 100);** 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 band **paintings etc. (89601).** At the other end of the spectrum **are:** vinyl chloride **(51131);** mechanically propelled **cars (79130);** wine lees **(8 194); Linseed (22340); methacrylic** acid **(51373);** slag etc. from iron **(27861); natural** sodium nitrate **(27120);** paper pulp filter-blocks (641%); tin **tubes (68724);** uranium **(68800);** and **oxy-fnct** alddhyde derivatives **(5 1622).**

[13] Our iterative technique **seems** to converge 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 wer **95%),** and significantly for most.

[14] Mexico's ranking may also **reflect** the **"806/807"** program or **reexports.** China **also** is ranked higher than some of the newly-industrialized Asian countries, which we plan to investigate more carefully in future work. We do not yet have a convenient **method** for estimating the statistical sensitivity of **country** rankings, though presumably some **simulation** technique can **be used.**

[15] For instance, all the country-specific correlations between the overall ranking and the **country-specific rankings** in **(10)** are positive.

To check the sensitivity of our results, we also tabulate in Table 1 four perturbations to our basic methodology. First, we restrict the sample of goods ranked to those with SITC codes between 60,000 and 80,000, which can be thought of as manufactured commodities. Second, we repeat our analysis but weigh each country (in the Kendall estimation procedure) by the number of individual goods it exported in the sample. This ensures that countries with a large number of exports are given more weight in determining the overall ranking; without weighting, countries which exported few goods to the US will be treated identically with countries which exported many goods. Finally, we estimate country rankings in the first and last halves of the sample. We do this by weighting the goods-rankings for each country by: 1) only the goods the country first exported before 1985; and 2) only those goods first exported by the country after 1984. To ease the comparison of the five different perturbations of the methodology, we also provide cross-scatterplots in Figure 1

Our results appear to be quite robust for the countries at the top df the rankings, but somewhat sensitive 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.[16] Still, the different rankings are quite highly correlated overall. Spearman rank correlations between the rankings are high (>.9) and statistically significant, and the rankings share essentially one common factor.[17]

---

[16] 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.

[17] 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 sane year. But manifestly dispersion can be found at finer levels of dis-aggregation; this dispersion also appears to in systematic and meaningful.

We have also compared the results in Table 1 (for manufacturing goods) to those obtained use the numerical method for maximizing (6). Using the overall goods ranking obtained from steps 1-4 as starting values, the average rank correlation in (6) was 0.4404. We then ran the numerical method for over 100 iterations until it converged, yielding a value for (6) of 0.4480. Unsurprisingly, the ordering of individual goods was quite similar for the two techniques. The correlation between the rankings of the "numerical method and the "regression method" was an extremely high 0.999.

Table 2 provides four different estimates of our country rankings, derived fiom the regression-based method. For these results, we estimate the country ranking as the primitive overall ranking in equation (14), rather than deriving it fiom some underlying estimate of a goods ranking. We also provide three perturbations to our basic methodology: a) estimates using only manufacturing data; b) weighted estimates; and c) an estimate derived with imputed data (when we actually "fill in" missing data using (13) and (14), and use this imputed data in our estimation). The cross-scatterplots are provided in Figure 2. Again, the results seem sensible and insensitive.

The results in Tables 1 and 2 are similar. That is, when ordering countries, it does not matter much whether we treat goods-rankings or c ountry-rankings as primitive. Table 3 provides a direct comparison of the baseline country rankings estimated both directly (treating the country ranking as primitive, as in Table 2) and indirectly (i.e., from a country-specific weighted average of goods-rankings, as in Table 1). It is comforting to note that the two rankings are closely related.[18] This can be seen more easily from the graphical analysis in figures 3-5. The latter

[18] We can see 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.

present scatterplots of the country- and goods-based rankings derived fiom our baseline methodology and two perturbations: a) baseline; b) using only manufacturing data; and c) weighted.

Figure 6 compares the "early" and "late" country rankings graphically. Few countries changed places dramatically, though the decline in some of the European rankings is interesting (bearing in mind that a country with a low numerical ranking is interpreted as "advanced). Figure 7 plots the country rankings (derived fiom goods rankings by simply averaging the latter over the set of goods exported on any given year) on a year by year basis for sixteen countries. Each of the "small multiple" graphics portrays a time-series plot of country ranking fiom 1972 through 1994. In future work, we plan to analyze the dynamics associated with changes in country rankings over time more closely.

Vb: Linking Country Rankings with Aggregate Variables

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

Figure 8 presents a simple bivariate scatterplot of country rankings (derived treating country rankings as primitive, as in Table 2) with the growth rate of real GDP per capita (taken fiom 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 "high" 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 4, our (ordinal) country ranking appears to be significantly negatively related to the growth rate of real GDP per capita.[19] 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 9 is a simply scatterplot of our country rankings (again, treating countries as primitive) and the level of 1985 real GDP per capita. 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 the *level* of total *factor* productivity and country ranking. 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.

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

Figure 10 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 5 contains the corresponding regression evidence.

VI: Topics for Future Research

Our country rankings are derived from a complicated semi-parametric estimation procedure using only dis-aggregated international trade data. We find it both reassuring and promising that they turn out to be related to important economic phenomena such as growth rate and level of real GDP per capita and the level of total factor productivity. Yet much remains to be done.

There is no explicit alternative hypothesis to our product cycle theory. An alternative explanation of these correlations stemming from a factor-endowments theory of international trade is a natural candidate, since factor proportions change slowly over time.

A closer examination of both the determinants and effects of country rankings is warranted. 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? America has been the richest country in the world for our sample, and (according to the product cycle theory) should be the first country to develop new goods. But the rankings should be similar when derived from the imports of any country; even though trade volumes differ systematically by country (the "gravity" effect).[20]

We could also combine our data with export data, and search for cycles in the data. Product cycles occur when a period of net imports follows one of net exports; this may, in turn, lead to another period of net export. Checking for recurrent cycles could allow us to compare the empirical import of "quality ladder" models of international trade with models which rely on an ever-increasing number of goods.

[20] It is interesting to note the negative relationship between trade volume and country ranking.

References

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

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.

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.

Appendix

Proof of Proposition 1

The value of p that minimizes (12) is given by:

$$\hat{\rho} = \frac{1}{M}\sum_{k=1}^{M}\frac{12}{(N_k^3 - N_k)}\sum_{i\in I_k}[x_{ik}(I_k) - \tfrac{1}{2}(N_k + 1)][X_i(I_k) - \tfrac{1}{2}(N_k + 1)], \qquad \text{(A1)}$$

which equals the average rank correlation in (6). Substituting this back into (12), it is straightforward to show that the objective function equals $M(1-\hat{\rho}^2)$. Thus, choosing the overall ranking $X_i(I)$ to maximize (6) is equivalent, when $\hat{\rho} > 0$, to minimizing the weighted sum of squared residuals in (12).

Proof of Proposition 2

The value of p that minimizes (15) is given by:

$$\hat{\rho} = \frac{\sum_{i=1}^{N}\sum_{k\in K_i}\frac{1}{M_i}[x_{ik}(I_k) + x_k^{min} - \tfrac{1}{2}(N+1)][X_i(I) - \tfrac{1}{2}(N+1)]}{\sum_{i=1}^{N}\sum_{k\in K_i}\frac{1}{M_i}[X_i(I) - \tfrac{1}{2}(N+1)]^2} \qquad \text{(A2)}$$

Substituting this back into (15), the objective function equals:

$$\sum_{i=1}^{N}\sum_{k\in K_i}\frac{1}{M_i}\left[x_{ik}(I_k) + x_k^{min} - \frac{N+1}{2}\right]^2 - \hat{\rho}^2\sum_{i=1}^{N}\sum_{k\in K_i}\frac{1}{M_i}\left[X_i(I) - \frac{N+1}{2}\right]^2. \qquad \text{(A3)}$$

The first double-summation that appears in (A3) does not depend on the choice for $X_i(I)$. The second double-summation can be simplified by noting that there are M_i elements in each set K_i, and that the terms being summed do not depend on the index k. Thus, thesecond double-summation simplifies to,

$$\sum_{i=1}^{N} \sum_{k \in K_i} \frac{1}{M_i} \left[X_i(I) - \frac{N+1}{2}\right]^2 = \frac{(N^3 - N)}{12}, \qquad \text{(A4)}$$

which is similar to the summation given in (4). Since this term is constant, it follows that choosing $X_i(I)$ to minimize (15) is equivalent to choosing $X_i(I)$ to maximize $\hat{\rho}^2$ in (A3). Provided that $\hat{\rho} > 0$, this is equivalent to maximizing the numerator of (A3), since the denominator is constant by (A4).

The numerator of (A4) can be rewritten as:

$$\sum_{i=1}^{N} \sum_{k \in K_i} \frac{1}{M_i} [x_{ik}(I_k) + x_k^{min}][X_i(I) - \tfrac{1}{2}(N+1)], \qquad \text{(A5)}$$

where we have used the fact that $X_i(I)$ has the average value of (N+1)/2 over i=1,...,N. It is evident that in order to maximize (A5), we shculd let the highest rank $X_i(I)$ multiply the highest value for $\sum_{k \in K_i} \frac{1}{M_i} [x_{ik}(I_k) + x_k^{min}]$, the second-highest rank multiply the second-highest value, and so on. In other words, $X_i(I)$ equals the ranks of the averages $\sum_{k \in K_i} \frac{1}{M_i} [x_{ik}(I_k) + x_k^{min}]$.

Table 1: Goods-Based Rankings

	Base	Mnfg	Wghtd	Early	Late
CANADA	1	3	1	1	1
UKINGDOM	2	4	2	2	11
GERMAN	3	2	3	3	15
JAPAN	4	1	4	4	6
FRANCE	5	5	5	5	2
MEXICO	6	9	6	6	3
NETHLDS	7	8	7	7	7
ITALY	8	6	8	8	8
BEL_LUX	9	11	9	9	14
SWITZLD	10	7	10	11	18
CHINA	11	17	13	21	17
SWEDEN	12	13	11	10	19
TAIWAN	13	12	12	12	5
SPAIN	14	14	16	13	4
BRAZIL	15	15	14	14	9
AUSTRAL	16	20	15	15	10
HONGKONG	17	19	17	17	13
KOREA_S	18	18	18	20	16
DENMARK	19	21	19	16	20
AUSTRIA	20	16	20	18	24
S_AFRICA	21	25	22	23	21
ISRAEL	22	22	23	25	22
NORWAY	23	23	21	22	27
INDIA	24	24	24	24	12
IRELAND	25	26	25	27	23
FINLAND	26	27	28	26	29
ARGENT	27	31	26	28	26
SINGAPR	28	28	29	32	25
USSR	29	34	30	38	39
VENEZ	30	29	32	33	38
UNKNOWN	31	10	27	19	160
NEW_ZEAL	32	33	31	30	36
THAILAND	33	30	33	37	31
PHIL	34	32	34	29	28
PORTUGAL	35	35	35	31	33
CHILE	36	41	38	41	43
COLOMBIA	37	40	36	39	30
POLAND	38	39	37	36	35
DOM_REP	39	44	39	35	41
MALAYSIA	40	38	40	42	34
YUGOSLAV	41	37	41	40	32
CZECHO	42	36	42	46	45
GERMAN_E	43	43	43	34	50
GREECE	44	42	44	45	37
PERU	45	46	45	44	40
HUNGARY	46	45	46	47	44
INDONES	47	48	47	54	42
TURKEY	48	49	48	57	49
ST_K_NEV	49	47	51	76	55
COS_RICA	50	50	49	52	51
ROMANIA	51	51	50	43	52
JAMAICA	52	56	52	49	48
GUATMALA	53	54	53	48	46
PANAMA	54	52	54	53	53
ECUADOR	55	57	55	59	54
SD_ARAB	56	53	60	70	68
EGYPT	57	62	62	69	60
PAKISTAN	58	55	56	61	57
NIGER	59	59	67	113	77
TRINIDAD	60	71	59	63	62
HONDURA	61	69	58	56	67
HAITI	62	61	57	50	47
MOROCCO	63	67	61	75	65

N_ANTIL	64	65	64	60	58
KENYA	65	60	65	74	70
BAHAMAS	66	76	63	55	61
BULGARIA	67	68	69	87	71
SALVADR	68	66	70	62	56
ICELAND	69	63	68	64	72
URUGUAY	70	73	66	66	59
MRITIUS	71	58	75	142	79
MACAU	72	64	76	78	63
IVY_CST	73	80	74	105	85
SRI_LKA	74	79	72	72	69
ARAB_EM	75	72	78	67	84
JORDON	76	74	84	132	83
IRAN	77	81	71	73	64
GABON	78	93	77	128	109
LEBANON	79	75	81	81	66
GILBRALT	80	70	73	58	73
KIRIBATI	81	95	89	118	101
YEMEN_S	82	107	79	51	150
GUYANA	83	108	83	65	103
NIGERIA	84	91	85	82	78
BARBADO	85	78	86	80	75
MOZAMBQ	86	89	82	71	122
NICARAGA	87	101	80	68	80
CYPRUS	88	84	90	77	82
BOLIVIA	89	85	87	88	74
MONGOLA	90	145	88	131	125
SURINAM	91	82	94	79	104
ZIMBABWE	92	90	91	91	102
GUINEA	93	88	106	147	110
NEW_CALE	94	106	93	127	96
BAHRAIN	95	96	105	93	114
BELIZE	96	104	96	83	107
BERMUDA	97	83	100	84	100
GHANA	98	100	92	111	86
MALI	99	86	107	145	105
SEYCHEL	100	77	111	102	113
CAMEROON	101	109	101	133	97
ALGERIA	102	116	97	99	93
TUNISIA	103	94	104	119	94
SYRIA	104	110	99	94	81
GUADLPE	105	98	110	92	76
FIJI	106	105	108	129	89
ZAIRE	107	134	98	86	106
BNGLDSH	108	87	103	95	92
ALBANIA	109	112	95	124	87
LIBERIA	110	115	109	85	115
NEPAL	111	92	112	107	88
AFGHAN	112	99	113	97	99
MALTA	113	97	114	104	95
SENEGAL	114	102	123	143	117
PARAGUA	115	124	102	89	123
BURMA	116	103	116	112	120
SIER_LN	117	136	121	110	118
G_BISAU	118	120	122	90	138
MADAGAS	119	125	115	109	124
OMAN	120	117	124	122	112
KUWAIT	121	111	130	116	116
CONGO	122	127	118	141	108
QATAR	123	137	127	115	128
FR_GUIAN	124	118	134	101	121
KOREA_N	125	133	117	96	146
FR_IND_O	126	114	142	148	126
TANZANIA	127	147	119	103	139
LIBYA	128	146	120	100	144

SUDAN	129	126	135	117	148
GREENLD	130	121	125	106	136
US_NES	131	131	139	144	129
NEW_GUIN	132	142	126	135	98
LAO	133	144	131	140	90
ZAMBIA	134	129	129	108	151
YEMEN_N	135	119	136	121	149
ANGOLA	136	143	128	114	133
S_HELNA	137	113	140	123	145
SP_MQEL	138	132	145	120	143
MALAWI	139	151	138	125	141
VIETNAM	140	140	133	134	111
UGANDA	141	139	132	137	127
ASIA_NES	142	128	146	150	140
SAMOA	143	152	144	136	130
SOMALIA	144	123	149	138	154
IRAQ	145	141	141	130	91
GAMBIA	146	122	151	151	132
MAURITN	147	135	143	126	134
BURUNDI	148	158	137	98	156
C_AFRICA	149	153	148	139	142
TOGO	150	149	153	153	137
BURKINA	151	130	152	156	147
RWANDA	152	138	150	146	131
ETHIOPIA	153	150	147	152	119
BENIN	154	154	155	154	135
CHAD	155	156	154	149	153
CAMBOD	156	148	157	157	152
FALK_IS	157	155	159	160	157
DJIBOUTI	158	157	158	159	155
CUBA	159	160	156	155	158
EQ_GNEA	160	159	160	158	159

Table 2: Country-Based Rankings

	Base	Mnfg	Wghtd	Imputed
CANADA	1	1	1	1
UKINGDOM	2	2	2	2
GERMAN	3	3	3	3
JAPAN	4	4	4	4
FRANCE	5	5	5	5
ITALY	6	6	7	6
MEXICO	7	7	6	7
NETHLDS	8	8	8	8
BEL_LUX	9	9	9	9
SWITZLD	10	10	10	10
SWEDEN	11	11	11	11
SPAIN	12	13	13	12
TAIWAN	13	12	12	13
HONGKONG	14	14	14	14
DENMARK	15	15	17	15
BRAZIL	16	16	15	17
AUSTRIA	17	18	18	16
AUSTRAL	18	17	16	19
INDIA	19	19	20	18
KOREA_S	20	20	19	20
ISRAEL	21	21	21	21
NORWAY	22	22	23	22
IRELAND	23	23	24	24
PORTUGAL	24	25	28	23
CHINA	25	24	22	34
FINLAND	26	28	27	27
S_AFRICA	27	26	25	30
ARGENT	28	27	26	31
COLOMBIA	29	29	31	29
SINGAPR	30	31	29	33
PHIL	31	30	30	32
POLAND	32	32	35	37
YUGOSLAV	33	33	34	36
UNKNOWN	34	45	88	28
THAILAND	35	35	32	42
GREECE	36	34	37	38
NEW_ZEAL	37	36	33	43
DOM_REP	38	37	40	46
VENEZ	39	38	36	53
HAITI	40	39	52	35
GERMAN_E	41	43	42	49
MALAYSIA	42	40	38	54
USSR	43	42	39	62
CZECHO	44	48	43	58
JAMAICA	45	41	48	44
ROMANIA	46	47	51	47
PERU	47	44	41	55
GUATMALA	48	46	50	52
PAKISTAN	49	49	54	48
HUNGARY	50	51	44	61
TURKEY	51	50	46	60
IRAN	52	53	57	45
INDONES	53	54	47	71
PANAMA	54	57	53	66
COS_RICA	55	58	49	70
SALVADR	56	59	58	56
ECUADOR	57	55	55	64
CHILE	58	56	45	83
NICARAGA	59	52	78	41
LEBANON	60	61	65	57
GILBRALT	61	65	72	59
BAHAMAS	62	60	68	63
MOROCCO	63	62	59	68

TRINIDAD	64	63	62	74
HONDURA	65	64	61	78
URUGUAY	66	66	63	77
N_ANTIL	67	67	64	80
KENYA	68	69	60	88
SYRIA	69	68	81	72
ST_K_NEV	70	70	56	98
MACAU	71	74	71	85
MALTA	72	83	83	73
BARBADO	73	76	73	81
SRI_LKA	74	72	70	89
ICELAND	75	75	66	90
EGYPT	76	73	67	94
AFGHAN	77	80	90	69
BOLIVIA	78	78	79	87
GUYANA	79	71	97	84
YEMEN_S	80	77	116	67
GHANA	81	81	91	93
NIGERIA	82	79	85	97
PARAGUA	83	84	103	82
CYPRUS	84	85	82	99
SD_ARAB	85	91	69	111
VIETNAM	86	86	108	76
BULGARIA	87	93	74	109
GUADLPE	88	92	87	95
TANZANIA	89	82	111	79
NEPAL	90	105	95	96
BNGLDSH	91	100	80	101
GREENLD	92	99	115	75
BELIZE	93	89	93	102
BERMUDA	94	103	92	103
LIBERIA	95	95	105	91
MOZAMBQ	96	88	101	114
IRAQ	97	90	127	50
ETHIOPIA	98	96	120	65
ARAB_EM	99	106	76	117
TUNISIA	100	110	84	105
LIBYA	101	94	150	26
NEW_CALE	102	107	94	115
IVY_CST	103	104	77	121
ZAIRE	104	97	113	110
ALGERIA	105	98	123	112
NEW_GUIN	106	102	121	100
MADAGAS	107	101	117	113
LAO	108	113	112	104
SURINAM	109	114	96	122
ANGOLA	110	87	146	39
MRITIUS	111	122	75	126
FIJI	112	111	100	119
NIGER	113	120	89	136
KOREA_N	114	128	155	86
JORDON	115	124	86	132
CAMEROON	116	118	99	128
KUWAIT	117	131	106	116
UGANDA	118	108	145	108
SIER_LN	119	119	118	125
MALI	120	129	98	130
BURMA	121	121	114	127
CONGO	122	109	143	120
ZAMBIA	123	123	130	107
BENIN	124	117	147	40
ZIMBABWE	125	132	104	141
ALBANIA	126	112	153	148
GABON	127	126	109	153
S_HELNA	128	127	129	118

KIRIBATI	129	130	102	142
OMAN	130	136	122	135
SENEGAL	131	137	107	134
MAURITN	132	133	134	106
BAHRAIN	133	140	119	145
G_BISAU	134	135	126	140
BURUNDI	135	115	154	51
FR_GUIAN	136	134	141	131
SAMOA	137	116	139	124
SEYCHEL	138	144	110	146
BURKINA	139	148	133	137
MALAWI	140	125	137	123
QATAR	141	143	144	149
YEMEN_N	142	139	128	151
SUDAN	143	138	131	150
RWANDA	144	151	142	143
ASIA_NES	145	149	125	144
GUINEA	146	146	124	155
US_NES	147	153	132	152
SP_MQEL	148	141	152	129
CUBA	149	142	160	25
TOGO	150	150	136	133
CAMBOD	151	156	138	139
MONGOLA	152	147	149	160
GAMBIA	153	157	140	154
C_AFRICA	154	154	151	156
SOMALIA	155	155	148	157
FR_IND_O	156	158	135	158
CHAD	157	152	157	147
DJIBOUTI	158	145	158	138
FALK_IS	159	160	156	159
EQ_GNEA	160	159	159	92

Table 3: Comparison of Goods- and Country-Based Rankings

	Country-Based	Goods-Based
CANADA	1	1
UKINGDOM	2	2
GERMAN	3	3
JAPAN	4	4
FRANCE	5	5
ITALY	6	8
MEXICO	7	6
NETHLDS	8	7
BEL_LUX	9	9
SWITZLD	10	10
SWEDEN	11	12
SPAIN	12	14
TAIWAN	13	13
HONGKONG	14	17
DENMARK	15	19
BRAZIL	16	15
AUSTRIA	17	20
AUSTRAL	18	16
INDIA	19	24
KOREA_S	20	18
ISRAEL	21	22
NORWAY	22	23
IRELAND	23	25
PORTUGAL	24	35
CHINA	25	11
FINLAND	26	26
S_AFRICA	27	21
ARGENT	28	27
COLOMBIA	29	37
SINGAPR	30	28
PHIL	31	34
POLAND	32	38
YUGOSLAV	33	41
UNKNOWN	34	31
THAILAND	35	33
GREECE	36	44
NEW_ZEAL	37	32
DOM_REP	38	39
VENEZ	39	30
HAITI	40	62
GERMAN_E	41	43
MALAYSIA	42	40
USSR	43	29
CZECHO	44	42
JAMAICA	45	52
ROMANIA	46	51
PERU	47	45
GUATMALA	48	53
PAKISTAN	49	58
HUNGARY	50	46
TURKEY	51	48
IRAN	52	77
INDONES	53	47
PANAMA	54	54
COS_RICA	55	50
SALVADR	56	68
ECUADOR	57	55
CHILE	58	36
NICARAGA	59	87
LEBANON	60	79
GILBRALT	61	80
BAHAMAS	62	66

MOROCCO	63	63
TRINIDAD	64	60
HONDURA	65	61
URUGUAY	66	70
N_ANTIL	67	64
KENYA	68	65
SYRIA	69	104
ST_K_NEV	70	49
MACAU	71	72
MALTA	72	113
BARBADO	73	85
SRI_LKA	74	74
ICELAND	75	69
EGYPT	76	57
AFGHAN	77	112
BOLIVIA	78	89
GUYANA	79	83
YEMEN_S	80	82
GHANA	81	98
NIGERIA	82	84
PARAGUA	83	115
CYPRUS	84	88
SD_ARAB	85	56
VIETNAM	86	140
BULGARIA	87	67
GUADLPE	88	105
TANZANIA	89	127
NEPAL	90	111
BNGLDSH	91	108
GREENLD	92	130
BELIZE	93	96
BERMUDA	94	97
LIBERIA	95	110
MOZAMBQ	96	86
IRAQ	97	145
ETHIOPIA	98	153
ARAB_EM	99	75
TUNISIA	100	103
LIBYA	101	128
NEW_CALE	102	94
IVY_CST	103	73
ZAIRE	104	107
ALGERIA	105	102
NEW_GUIN	106	132
MADAGAS	107	119
LAO	108	133
SURINAM	109	91
ANGOLA	110	136
MRITIUS	111	71
FIJI	112	106
NIGER	113	59
KOREA_N	114	125
JORDON	115	76
CAMEROON	116	101
KUWAIT	117	121
UGANDA	118	141
SIER_LN	119	117
MALI	120	99
BURMA	121	116
CONGO	122	122
ZAMBIA	123	134
BENIN	124	154
ZIMBABWE	125	92
ALBANIA	126	109
GABON	127	78

S_HELNA	128	137
KIRIBATI	129	81
OMAN	130	120
SENEGAL	131	114
MAURITN	132	147
BAHRAIN	133	95
G_BISAU	134	118
BURUNDI	135	148
FR_GUIAN	136	124
SAMOA	137	143
SEYCHEL	138	100
BURKINA	139	151
MALAWI	140	139
QATAR	141	123
YEMEN_N	142	135
SUDAN	143	129
RWANDA	144	152
ASIA_NES	145	142
GUINEA	146	93
US_NES	147	131
SP_MQEL	148	138
CUBA	149	159
TOGO	150	150
CAMBOD	151	156
MONGOLA	152	90
GAMBIA	153	146
C_AFRICA	154	149
SOMALIA	155	144
FR_IND_O	156	126
CHAD	157	155
DJIBOUTI	158	158
FALK_IS	159	157
EQ_GNEA	160	160

Table 4: 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.

Table 5: 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: Perturbations of Goods-Based Rankings

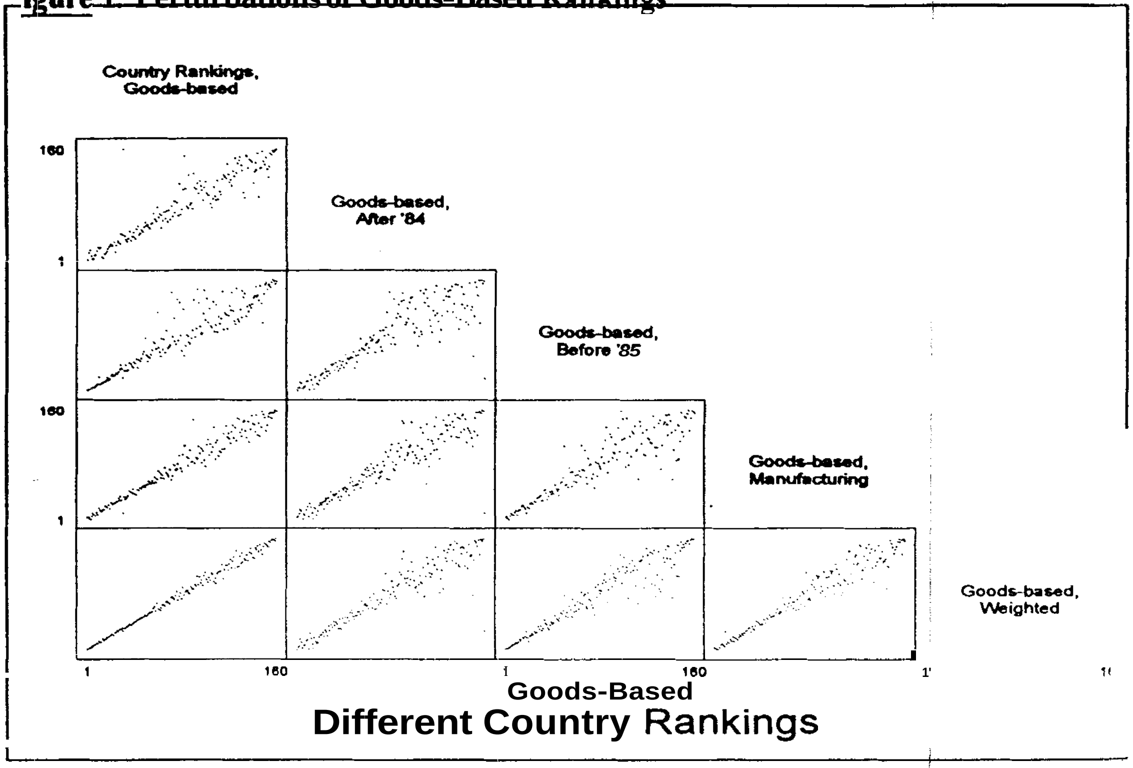


Figure 2: Perturbations of Country-Based Rankings

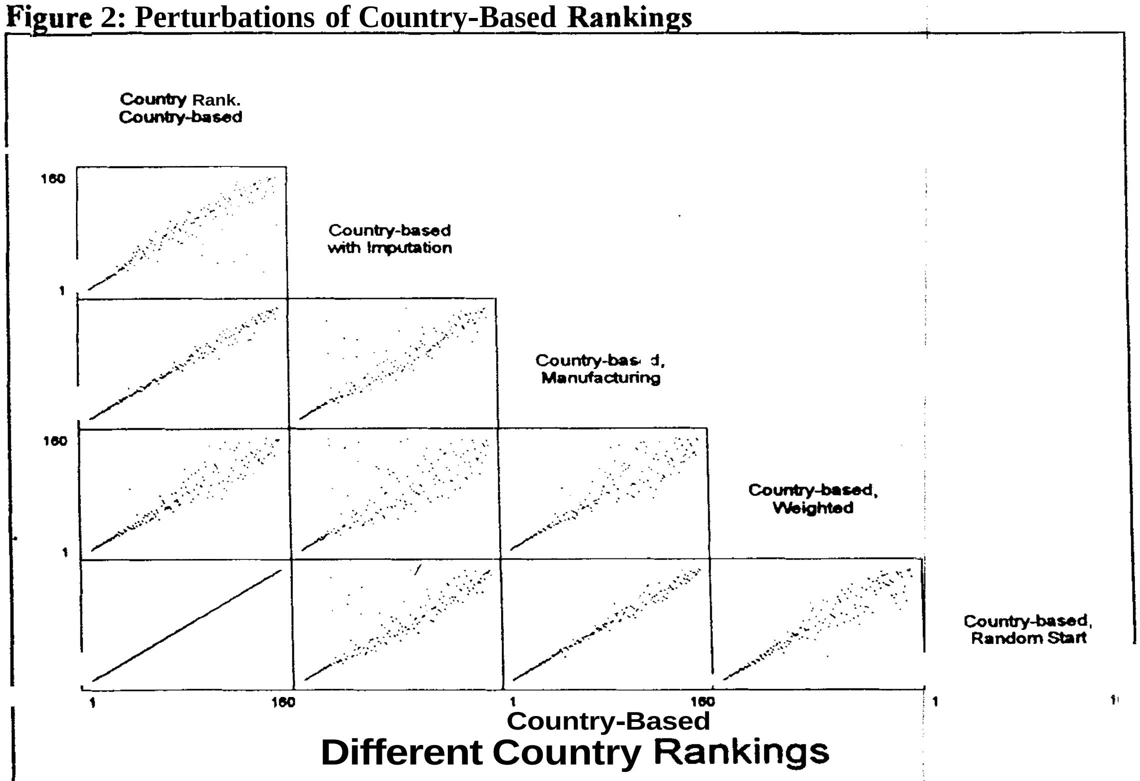


Figure 3: Comparison of Goods- and Country-Based Rankings

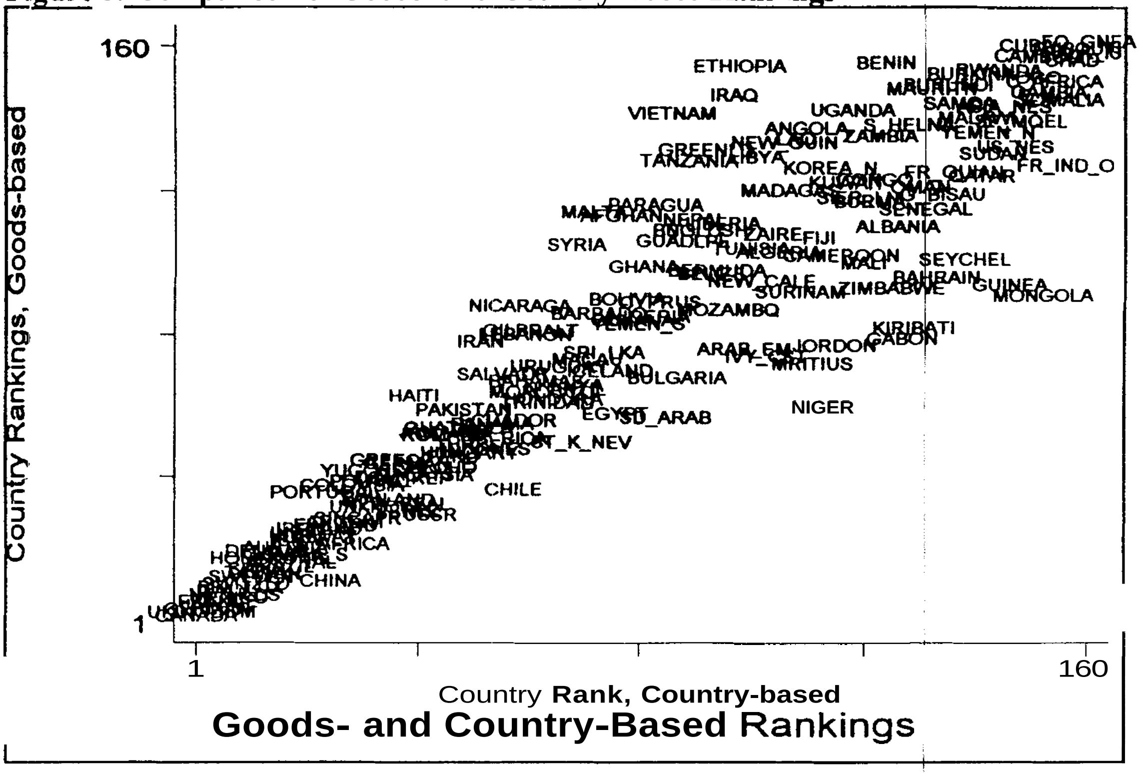


Figure 4: Comparison of Goods- and Country-Based Rankings, Manufacturing Data

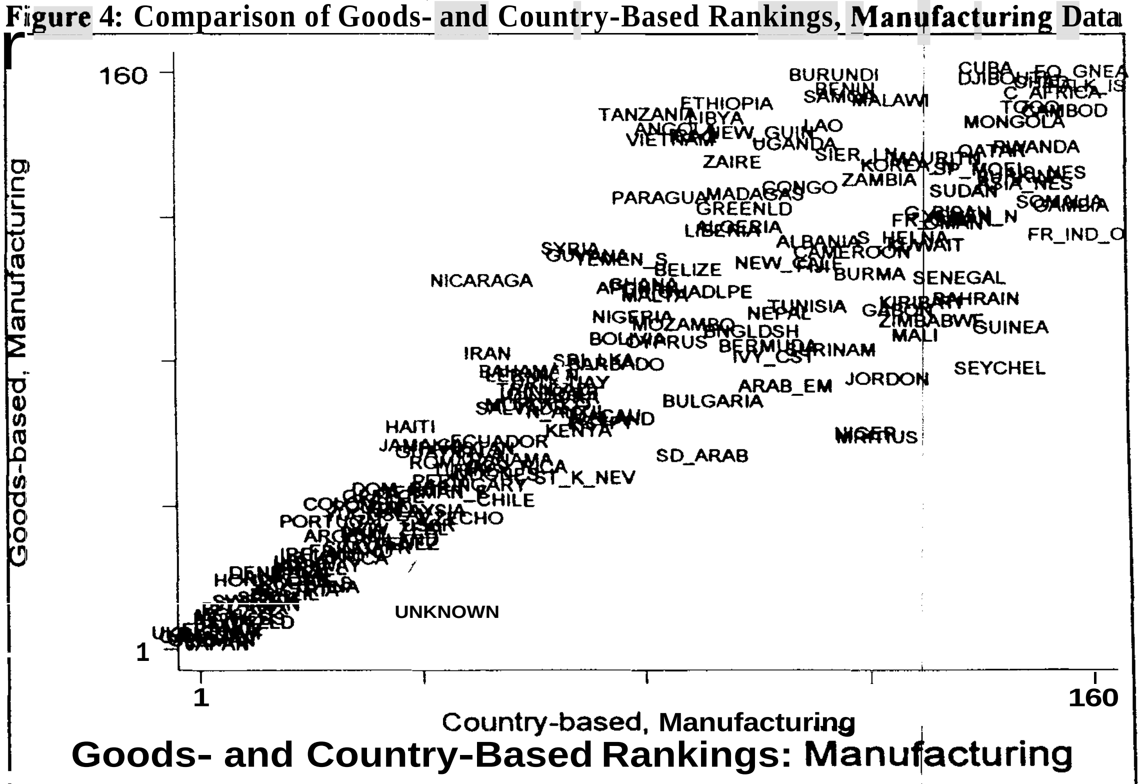


Figure 5: Comparison of Goods- and Country-Based Rankings, Weighted Data

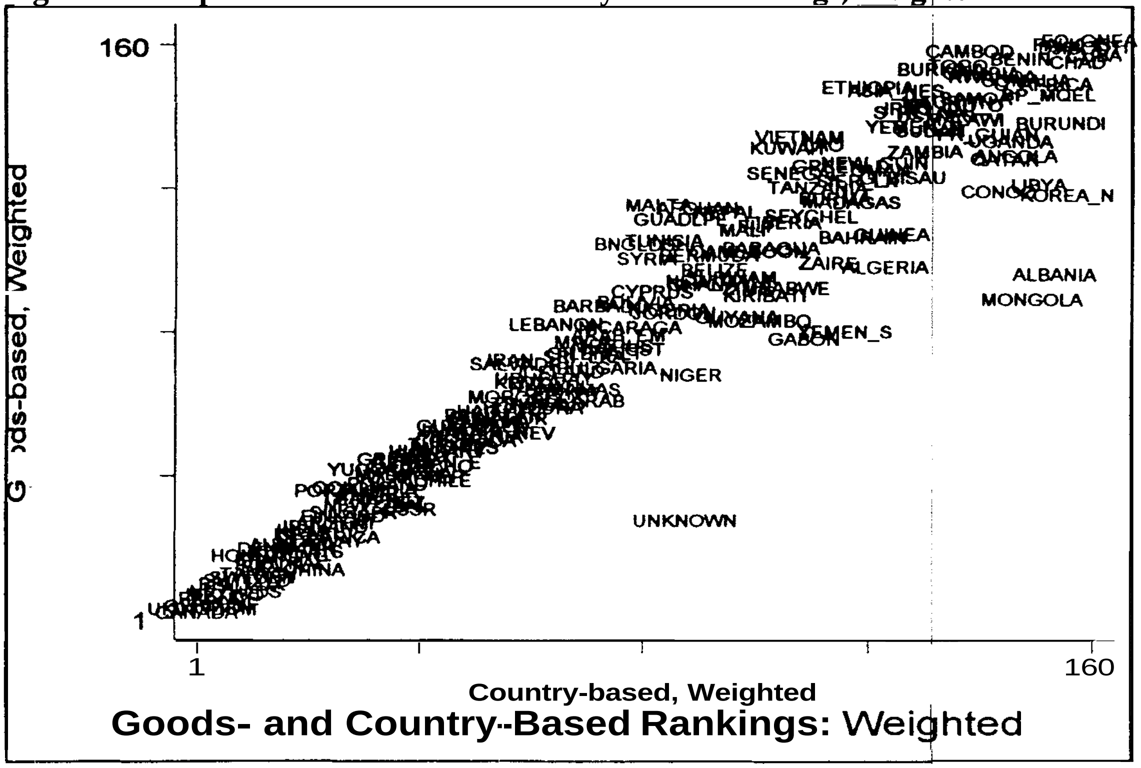


Figure 6: Comparison of Early and Late Goods-Based Rankings

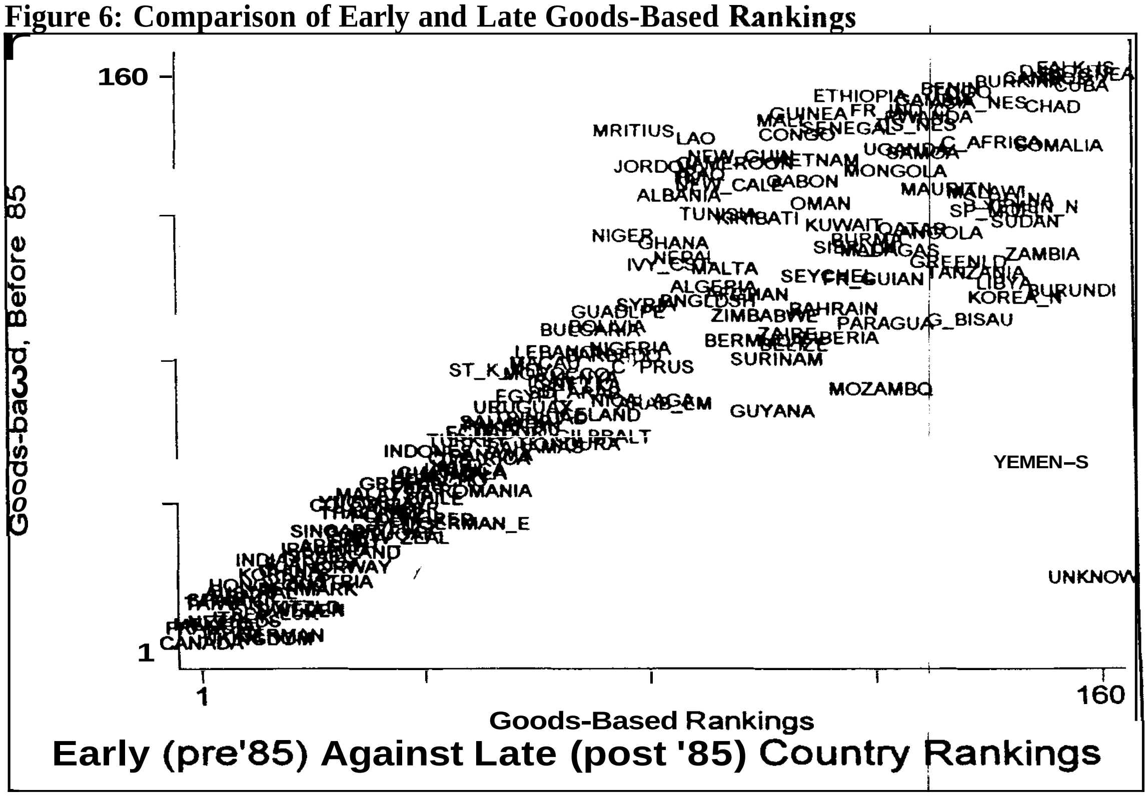


Figure 7: Country Rankings over Time

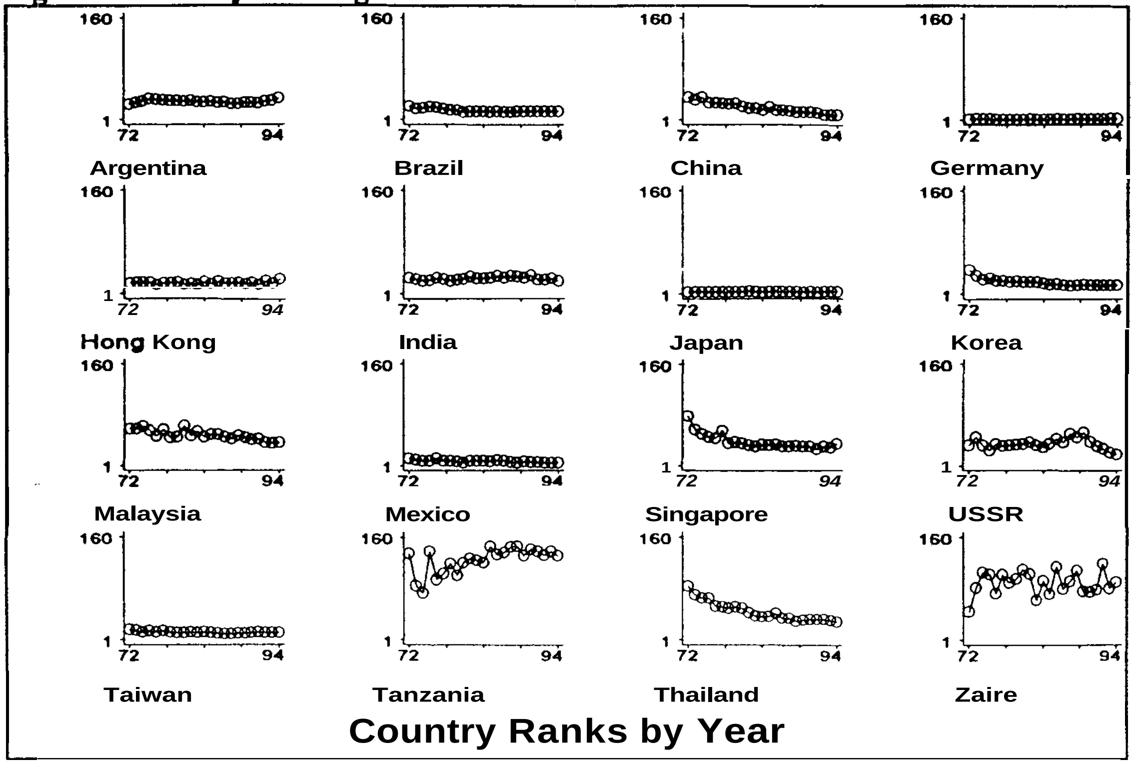


Figure 8: Country Rankings and Real GDP per capita Growth

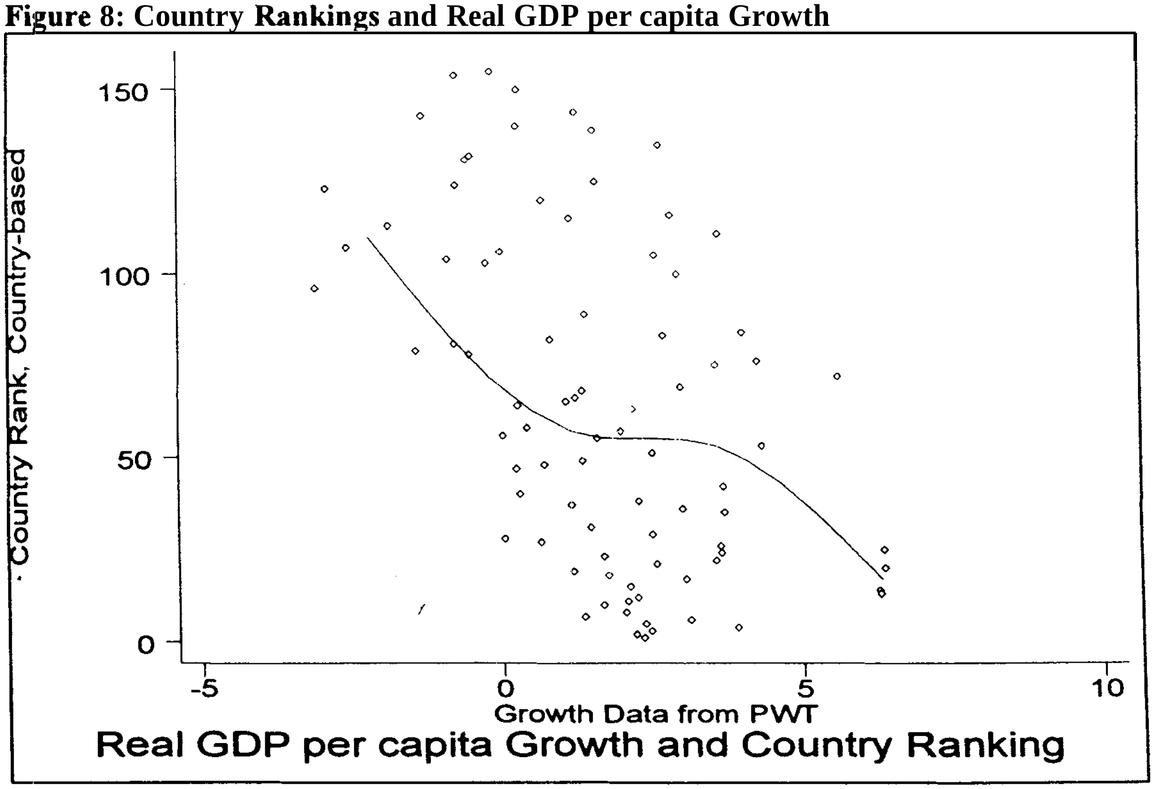


Figure 9: Country Rankings and Real GDP per capita in 1985

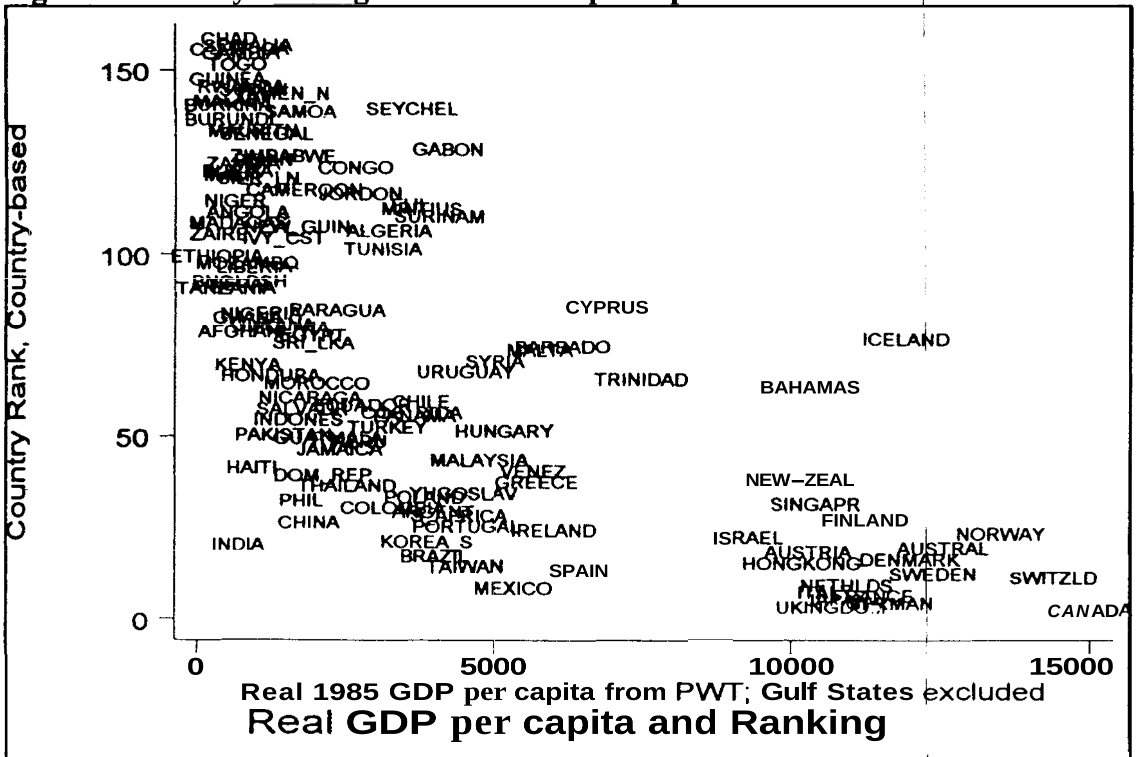


Figure 10: Country Rankings and the Log-Level of Total Factor Productivity