NBER WORKING PAPER SERIES

WHAT YOU EXPORT MATTERS

Ricardo Hausmann
Jason Hwang
Dani Rodrik

Working Paper 11905
http://www.nber.org/papers/w11905

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2005

What You Export Matters
Ricardo Hausmann, Jason Hwang, and Dani Rodrik
NBER Working Paper No. 11905
December 2005
JEL No. F1, O4

ABSTRACT

When local cost discovery generates knowledge spillovers, specialization patterns become partly indeterminate and the mix of goods that a country produces may have important implications for economic growth. We demonstrate this proposition formally and adduce some empirical support for it. We construct an index of the "income level of a country's exports," document its properties, and show that it predicts subsequent economic growth.

Ricardo Hausmann
John F. Kennedy School of Government
Harvard University
ricardo_hausmann@harvard.edu

Jason Hwang
Department of Economics
Harvard University
jjhwang@fas.harvard.edu

Dani Rodrik
John F. Kennedy School of Government
Harvard University
79 JFK Street
Cambridge, MA 02138
and NBER
dani_rodrik@harvard.edu

What You Export Matters*

Ricardo Hausmann
John F. Kennedy School of Government
Harvard University

Jason Hwang
Department of Economics
Harvard University

Dani Rodrik
John F. Kennedy School of Government
Harvard University

December 2005

Abstract

When local cost discovery generates knowledge spillovers, specialization patterns become partly indeterminate and the mix of goods that a country produces may have important implications for economic growth. We demonstrate this proposition formally and adduce some empirical support for it. We construct an index of the "income level of a country's exports," document its properties, and show that it predicts subsequent economic growth.

1 Introduction

Why do countries produce what they do, and does it matter? The conventional approach to these questions is driven by what we might call the "fundamentals" view of the world. In this view, a country's fundamentals–namely its endowments of physical and human capital, labor, and natural resources along with the overall quality of its institutions–determine relative costs and the patterns of specialization that go with them. Attempts to reshape the production structure beyond the boundaries set by these fundamentals are likely to fail and hamper economic performance.

We argue in this paper that specialization patterns are partly indeterminate and may be shaped by idiosyncratic elements. While fundamentals play an

*Hausmann and Rodrik thank the Center for International Development for financial support. Oeindrila Dube and Bailey Klinger provided excellent research assistance.

important role, they do not uniquely pin down what a country will produce and export. Furthermore, not all good are alike in terms of their consequences for economic performance. Specializing in some products will bring higher growth than specializing in others. In this setting, government policy has a potentially important positive role to play in shaping the production structure.

Our analysis is driven by what Hausmann and Rodrik (2003) have earlier called "cost discovery." An entrepreneur who attempts to produce a good for the first time in a developing economy necessarily faces considerable cost uncertainty. Even if the good comes with a standard technology ("blueprint"), domestic factor endowments and institutional realities will require tinkering and local adaptation (see Evenson and Westphal 1995, Lall 2000). What the entrepreneur effectively does is to explore the underlying cost structure of the economy. This process is one with considerable positive externalities for other entrepreneurs. If the project is successful, other entrepreneurs learn that the product in question can be profitably produced and emulate the incumbent. In this way, the returns to the pioneer investor's cost discovery become socialized. If the incumbent ends up with failure, on the other hand, the losses remain private. This knowledge externality implies that investment levels in cost discovery are sub-optimal unless the industry or the government find some way in which the externality can be internalized.

In such a setting, the range of goods that an economy ends up producing and exporting is determined not just by the usual fundamentals, but also by the number of entrepreneurs that can be stimulated to engage in cost discovery in the modern sectors of the economy. The larger this number, the closer that the economy can get to its productivity frontier. When there is more cost discovery, the productivity of the resulting set of activities is higher in expectational terms and the jackpot in world markets bigger.

In this paper we provide a simple formal model of this process. We also supply some evidence that we think is suggestive of the importance of the forces that our formal framework identifies. We are interested in showing that some traded goods are associated with higher productivity levels than others and that countries that latch on to higher productivity goods (through the cost discovery process just described) will perform better. Therefore, the key novelty is a quantitative index that ranks traded goods in terms of their implied productivity. We construct this measure by taking a weighted average of the per-capita GDPs of the countries exporting a product, where the weights reflect the revealed comparative advantage of each country in that product. So for each good, we generate an associated income/productivity level (which we call $PRODY$). We then construct the income/productivity level that corresponds to a country's export basket (which we call $EXPY$), by calculating the export-weighted average of the $PRODY$ for that country. $EXPY$ is our measure of the productivity level associated with a country's specialization pattern.

While $EXPY$ is highly correlated with per-capita GDPs, we show that there are interesting discrepancies. Some high-growth countries such as China and India have $EXPY$ levels that are much higher than what would be predicted based on their income levels. China's $EXPY$, for example, exceeds those of countries in Latin America with per-capita GDP levels that are a multiple of that of China. More generally, we find that $EXPY$ is a strong and robust predictor of subsequent economic growth, controlling for standard covariates. We show this result for a recent cross-section as well as for panels that go back to the early 1960s. The results hold both in instrumental variables specifications (to control for endogeneity of $EXPY$) and with country fixed effects (to control for unobserved heterogeneity).

Our approach relates to a number of different strands in the literature. Recent work in trade theory has emphasized cost uncertainty and heterogeneity at the level of firms so as to provide a better account of global trade (Bernard et al. 2003, Melitz and Ottaviano 2005). In contrast to this literature, we focus on the spillovers in cost information and are interested in the economic growth implications. There is also an empirical literature on the so-called natural resource curse, which examines the relationship between specialization in primary products and economic growth (Sachs and Warner 1995). The rationale for the natural resource curse is based either on the Dutch disease or on an institutional explanation (Subramanian and Sala-i-Martin 2003). Our approach has different micro-foundations than either of these, and yields an empirical examination that is much more fine-grained. We work with two datasets consisting of more than 5,000 and 700 individual commodities each and eschew a simple primary-manufactured distinction.

The outline of the paper is as follows. We begin in section 2 with a simple model that develops the key ideas. We then present the empirical analysis in section 3. We conclude in section 4.

2 A simple model

We are concerned with the determination of the production structure of an economy in which the standard forces of comparative advantage play some role, but not the exclusive role. The process of discovering the underlying cost structure of the economy, which is intrinsically uncertain, contributes a stochastic dimension to what a county will produce and therefore how rich it will be.

We normalize units of goods such that all goods have an exogenously given world price p. Each good is identified by a certain productivity level θ, which represents the units of output generated by an investment of given size. We align these goods on a continuum such that higher-ranked goods entail higher productivity. The range of goods that an economy is capable of producing is given by a continuous interval between 0 and h, i.e., $\theta \in [0, h]$ (see Figure 1) We capture the role of comparative advantage by assuming that the upper boundary of this range, h, is an index of the skill- or human-capital level of the economy. Hence a country with higher h can produce goods of higher productivity ("sophistication").

Projects are of fixed size and entail the investment of b units of labor. When investors make their investment decisions, they do not know whether they will end up with a low-productivity good or a low-productivity good. The θ associated with an investment project is discovered only after the investment is sunk. All that investors know ex ante is that θ is distributed uniformly over the range $[0, h]$.

However, once the θ associated with a project/good is discovered, this becomes common knowledge. Others are free to produce that same good without incurring additional "discovery" costs (but at a somewhat lower productivity than the incumbent). Emulators operate at a fraction α of the incumbent's productivity, with $0 < \alpha < 1$. Each investor can run only one project, so having discovered the productivity of his own project, the investor has the choice of sticking with that project or emulating another investor's project.

An investor contemplating this choice will compare his productivity θ_i to that of the most productive good that has been discovered, $\theta^{\max}$, since emulating any other project will yield less profit. Therefore, the decision will hinge on whether θ_i is smaller or bigger than $\alpha\theta^{\max}$. If $\theta_i \geq \alpha\theta^{\max}$, investor i will stick with his own project; otherwise he will emulate the $\theta^{\max}$-project. Therefore the productivity range within which firms will operate is given by the thick part of the spectrum shown in Figure 1.

Now let's move to the investment stage and consider the expected profits from investing in the modern sector. These expected profits depend on expectations regarding both the investor's own productivity draw and the maximum of everybody else's draws. As we shall see, the latter plays a particularly important role. Obviously, $E(\theta^{\max})$ will be an increasing function of the number of investors who start projects. Let m denote the number of investors who choose to make investments in the modern sector. Given our distributional assumptions, we have a particularly simple expression for $E(\theta^{\max})$:

$$E\left(\theta^{\max}\right) = \frac{hm}{m+1}$$

Note that $E(\theta^{\max})$ equals 0 when $m = 0$, and converges to h as $m \to \infty$.

Since productivity is distributed uniformly, the probability that investor i will stick with his own project is

$$prob(\theta_i \geq \alpha\theta^{\max}) = 1 - \frac{E(\theta^{\max})}{h} = 1 - \frac{\alpha m}{m+1}.$$

This eventuality yields the following expected profits

$$E(\pi \,|\theta_i \geq \alpha\theta^{\max}\,) = \frac{1}{2}p[h + \alpha E(\theta^{\max})] = \frac{1}{2}ph[1 + \frac{\alpha m}{m+1}]$$

since $\frac{1}{2}[h + \alpha\theta^{\max}]$ is the expected productivity of such a project. We can similarly work out the probability and expected profits for the case of emulation:

$$prob(\theta_i < \alpha\theta^{\max}) = \frac{E(\theta^{\max})}{h} = \frac{\alpha m}{m+1}.$$

$$E(\pi \,|\theta_i < \alpha\theta^{\max}\,) = \frac{1}{2}p[h + \alpha E(\theta^{\max})] = ph\left(\frac{\alpha m}{m+1}\right)$$

Putting these together, we have

$$E(\pi) = ph\left[\left(1-\frac{\alpha m}{m+1}\right)\frac{1}{2}\left(1+\frac{\alpha m}{m+1}\right)+\left(\frac{\alpha m}{m+1}\right)^2\right]$$
$$= \frac{1}{2}ph\left[1+\left(\frac{\alpha m}{m+1}\right)^2\right] \tag{1}$$

Note that expected productivity in the modern sector is

$$E(\theta) = \overline{\theta} = \frac{1}{2}h\left[1+\left(\frac{\alpha m}{m+1}\right)^2\right] \tag{2}$$

Expected profits shown in (1) are simply the product of price and expected productivity. Expected productivity, and in turn profitability are determined both by "skills" (h) and by the number of investors engaged in cost discovery (m). The larger m, the higher the productivity in the modern sector. Hence we have increasing returns to scale in the modern sector, but this arises from cost information spillovers rather than technological externalities. If α were zero, productivity and profits would not depend on m.

2.1 Long-run equilibrium

In long-run equilibrium, the number of entrants in the modern sector (m) is endogenous and is determined by the requirement that excess profits are driven to zero. Let us express the flow (expected) profits in this sector as

$$r(p,h,m^*) = E(\pi)_{LR} = \frac{1}{2}ph\left[1+\left(\frac{\alpha m^*}{m^*+1}\right)^2\right]$$

where m^* denotes the long-run level of m. Remember that each modern sector investment requires b units of labor upfront, resulting in a sunk investment of bw, where w is the economy's wage rate. Long-run equilibrium requires equality between the present discounted value of $r(p,h,m^*)$ and the sunk cost of investment:

$$\int_0^\infty r(p,h,m^*)e^{-\rho t}dt = \frac{r(p,h,m^*)}{\rho} = bw^* \tag{ZP}$$

where ρ is the discount rate.

Wages are determined in turn by setting the economy's total labor demand equal to the fixed labor supply L. The modern sector's labor demand equals m^*b. Let the traditional sector's labor demand be given by the decreasing function $g(w)$, $g'(w) < 0$. Labor market equilibrium is then given by

$$m^*b + g(w^*) = L \tag{LL}$$

Equations (ZP) and (LL) determine the long-run values of the endogenous variables m and w. The equilibrium is shown in Figure 2, which plots these two equations in (m, w)-space. Note that (ZP) and (LL) are both positively sloped. We have drawn (ZP) as less steep than (LL), because otherwise scale economies would be so strong that the dynamic behavior of the model would be unstable under reasonable specifications. This amounts to assuming that α is not too large.

2.2 Short-run equilibrium

In short-run equilibrium we require labor markets to clear but take m as fixed. This means we are always on the (LL) schedule, with the wage rate determined by equation (LL) for a given m.

2.3 Dynamics

Given our assumptions so far, if m were allowed to adjust instantaneously we would jump immediately to the long-run equilibrium given by the intersection of the (ZP) and (LL) schedules. In fact, forward-looking behavior on the part of investors in the modern sector provides an additional mechanism for immediate convergence to the long-run equilibrium. Suppose, for example, that we start at a level of m which falls short of m^*. On the transition to the long-run equilibrium, we know that m and w will both rise. Consider how these dynamics influence the decision to enter. The rise in m implies that productivity will be higher in the future than it is today, and is a force that will induce delay in the decision to invest in the modern sector ceteris paribus. The rise in w, on the other hand, implies that investment will be more costly in the future than it is today, and is a factor that will precipitate investment. Given the relative slopes we have assumed, the second factor outweighs the first–i.e., wages increase faster than the rate at which productivity benefits come in–and investors would rather invest today than wait.

To provide the model with some non-trivial dynamics, we can simply assume that there is a limit to how much investment is feasible per unit of time. To be concrete, let the rate at which m increases be restricted by the exogenous parameter μ. That is

$$\left|\dot{m}(t)\right| \leq \mu$$

Given the considerations discussed in the previous paragraph, there will be maximal adjustment in m whenever net returns at time t are non-zero. Hence,

$$\begin{array}{lll} \dot{m}(t) = \mu & if & \frac{r(p,h,m(t))}{\rho} > bw(t) \\ \dot{m}(t) = -\mu & if & \frac{r(p,h,m(t))}{\rho} < bw(t) \\ \dot{m}(t) = 0 & otherwise & \end{array}$$

2.4 Comparative dynamics

We are now ready to analyze the behavior of the economy. Starting from an initial equilibrium given by (m_0, w_0), consider an increase in the economy's labor endowment. This shifts the LL schedule down since, at a given m, labor-market equilibrium requires lower wages. Hence the impact effect of larger L is a lower w. However, the lower wage induces more firms to enter the modern sector and engage in cost discovery, which in turn pulls wages up. How high do wages eventually go? As Figure 2 shows, the new equilibrium is one where wages are higher than in the initial equilibrium. A larger labor endowment ends up boosting wages! What is key for this result is the presence of information spillovers in the modern sector. Once the modern sector expands, productivity rises, and zero profits can be restored only if wages go up.

Increases in p and h operate by shifting the ZP schedule up. They both result in higher m and w eventually. These results are less surprising.

3 Empirics

The model shows that productivity in the modern sector is driven by $\theta^{\max}$, which depends on m, which in turn is driven by country size (L), human capital (h), and other parameters. In our empirical work, we shall proxy $\theta^{\max}$ with a measure calculated from export statistics which we call $EXPY$. This measure aims to capture the productivity level associated with a country's exports. Focusing on exports is a sensible strategy since $\theta^{\max}$ refers to the most productive goods that a country produces and we can expect a country to export those goods in which it is the most productive. Besides, we have much more detailed data on exports across countries than we do on production.

In order to calculate $EXPY$ we rank commodities according to the income levels of the countries that export them. Commodities that are exported by rich countries (controlling for overall economic size) get ranked more highly than commodities that are exported by poorer countries. With these commodity-specific calculations, we then construct country-wide indices.

3.1 Construction of $EXPY$

First, we construct an index called $PRODY$. This index is a weighted average of the per capita GDPs of countries exporting a given product, and thus represents the income level associated with that product. Let countries be indexed by j and goods be indexed by l. Total exports of country j equals

$$X_j = \sum_l x_{jl}$$

Let the per-capita GDP of country j be denoted Y_j. Then the productivity level associated with product k, $PRODY_k$, equals

$$PRODY_k = \sum_j \frac{(x_{jk}/X_j)}{\sum_j (x_{jk}/X_j)} Y_j$$

The numerator of the weight, x_{jk}/X_j, is the value-share of the commodity in the country's overall export basket. The denominator of the weight, $\sum_j (x_{jk}/X_j)$, aggregates the value-shares across all countries exporting the good. Hence the index represents a weighted average of per-capita GDPs, where the weights correspond to the revealed comparative advantage of each country in good k.

The rationale for using revealed comparative advantage as a weight is to ensure that country size does not distort our ranking of goods. Consider an example involving Bangladesh and US garments, specifically, the 6-digit product category 620333, "men's jackets and blazers, synthetic fiber, not knit." In 1995, the US export value for this category was $28,800,000, exceeding Bangladesh's export value of $19,400,000. However, this commodity constituted only 0.005 percent of total US exports, compared to 0.6 percent for Bangladesh. As defined above, the $PRODY$ index allows us to weight Bangladesh's income more heavily than the U.S. income in calculating the productivity level associated with garments, even though the U.S. exports a larger volume than Bangladesh.

The productivity level associated with country i's export basket, $EXPY_i$, is in turn defined by

$$EXPY_i = \sum_l \left(\frac{x_{il}}{X_i}\right) PRODY_l$$

This is a weighted average of the $PRODY$ for that country, where the weights are simply the value shares of the products in the country's total exports.

3.2 Data and methods

Our trade data come from two sources. The first is the United Nations Commodity Trade Statistics Database (COMTRADE) covering over 5000 products at the Harmonized System 6-digit level for the years 1992 to 2003. The value of exports is measured in current US dollars. The number of countries that report the trade data vary considerably from year to year. However, we constructed the $PRODY$ measure for a consistent sample of countries that reported trade data in each of the years 1999, 2000 and 2001. It is essential to use a consistent sample since non-reporting is likely to be correlated with income, and thus, constructing $PRODY$ for different countries during different years could introduce serious bias into the index. While trade data were actually available for 124 countries over 1999-2001, the real per capita GDP data from the World Development Indicators (WDI) database was only available for 113 of these countries. Thus, with the COMTRADE data, we calculate $PRODY$ for a sample of 113 countries. We calculate $PRODY$ using both PPP-adjusted GDP and GDP at market exchange rates. In what follows we shall present most of our results only with the PPP-adjusted measures of $PRODY$; we have found no instance in which using one instead of the other makes a substantive difference.

The average $PRODY$ from 1999-2001 is then used to construct an $EXPY$ measure for all countries reporting trade data during the period from 1992 to 2003. Since the number of countries reporting COMTRADE data varies from year to year, and the coverage is especially patchy for earlier years, the total number of countries for which we could calculate $EXPY$ ranges from a low of 48 in 1992 to a high of 133 in 2000. Table 1 shows the country coverage for each of the years between 1992 and 2003.

Some limitations of COMTRADE data are its relatively short time-span and limited coverage of countries earlier in the period. To check the robustness of our findings against these concerns, we have also constructed our measures with the World Trade Flows dataset which has recently been updated to extend coverage back to 1962 (Feenstra et al. 2005). Trade flows are based on 4-digit standard international trade classifications (SITC rev. 2) comprising over 700 commodities. Our $PRODY$ and $EXPY$ indices are calculated by combining the World Trade Flows data on export volumes with PPP-adjusted GDP from the Penn World Tables, yielding a sample of 97 countries for the period 1962-2000.

We prefer to work with the indices based on more disaggregated data, and the basic patterns in the data are very much consistent between the two datasets. Hence we limit our discussion of descriptive statistics below to the COMTRADE data. We return to the 4-digit data when we turn to growth regressions.

3.3 Descriptive statistics

Some descriptive statistics on $PRODY$ are shown in Table 2. The first row shows $PRODY$ calculated using GDP at market exchange rates and the second row shows $PRODY$ with PPP-adjusted GDP levels. As the table reveals, the income level associated with individual traded commodities varies greatly, from numbers in hundreds (of 2000 US dollars) to tens of thousands. This reflects the fact that specialization patterns are highly dependent on per-capital incomes.

The five commodities with the smallest and largest $PRODY$ values are shown in Table 3. As we would expect, items with low $PRODY$ tend to be primary commodities. Consider for example product 10120, "live asses mules and hinnies." The main reason this product has the lowest income level is that it constitutes a relatively important part of the exports of Niger, a country with one of the lowest per capita GDPs in our sample. Similarly, sisal. cloves, and vanilla beans have low $PRODY$ values because they tend to be significant exports for poor sub-Saharan African countries. On the other hand, product 7211060, flat rolled iron or non-alloy steel, has the highest $PRODY$ value because it holds a substantial share of Luxembourg's exports, and this country has the highest per capita GDP in our sample.

Table 4 and Figure 3 summarize some basic descriptive statistics for $EXPY$. We note that the mean $EXPY$ for the sample of countries included exhibits a downward trend over time. Mean $EXPY$ has fallen from \$12,994 in 1992 to \$10,664 in 2003. Since the income levels associated with individual products are held constant over time (as explained above), this is due partly to the changing composition of the sample of countries (with more low-$EXPY$ countries being included over time) and partly to the reduction in $EXPY$ levels in many of the countries. Indeed, Table 5 shows that a majority of countries (among those that have $EXPY$ values throughout our sample period) have experienced a reduction in $EXPY$ over time. This downward trend may be specific to the recent period, since we do not see a similar trend since the 1960s when we use 4-digit trade data.

How does $EXPY$ vary across countries? Figure 4 shows a scatterplot of $EXPY$ against per-capita GDP. Unsurprisingly, there is a very strong correlation between these two variables. The correlation coefficient between the two is in the range 0.80-0.83 depending on the year. Rich (poor) countries export products that tend to be exported by other rich (poor) countries. Although in our framework this relationship has a different interpretation, it can also be explained with the Heckscher-Ohlin framework if rich country goods are more intensive in human capital or physical capital. The relationship between $EXPY$ and per capita GDP exists partly by construction, since a commodity's $PRODY$ is determined by the per capita GDPs of the countries that are important exporters of that commodity. However, the relationship is not just a mechanical one. Calculating country specific $PRODY$s by excluding own exports from the calculation of these measures does not change the results much. Note also that the variation in $EXPY$ across countries is much lower than the variation in per-capita GDPs. This is a direct consequence of the fact that $PRODY$ (and therefore $EXPY$) is a weighted average of national income levels.

Table 6 shows the countries with smallest and largest $EXPY$ values for 2001 (the year with the largest possible sample size). Note that French Polynesia (PYF) ranks in the top 5 among those with the largest $EXPY$. This surprising outcome arises in part because cultured pearl exports contribute heavily to a French Polynesia's export basket and this product has a relatively large $PRODY$ value of $22,888. A few other cases where countries appear to have very large $EXPY$ values relative to per capita GDP are Mozambique (MOZ), Swaziland (SWZ), Armenia (ARM), India (IND), and China (CHN). In a couple of these instances, the culprit is once again a specific commodity with a high $PRODY$ value: unwrought, alloyed aluminum for Mozambique and "mixed odoriferous substances in the food and drink industries" for Swaziland. But in the remaining cases (China, India, and Armenia), this is the result of a portfolio of a high $PRODY$ exports, and not one or two specific items. At first sight, diamonds seem to play a large role in India and Armenia, but both countries retain their high $EXPY$s even with diamonds removed from the calculation. And China has a very diversified set of exports, with no single product category standing out in terms of high export shares. It is worth remembering at this juncture that China and India have both been experiencing very rapid economic growth (as has Armenia more recently) .

Figure 5 shows the time trend for $EXPY$ for China, India, and a sample of other Asian and Latin American countries. Among the Latin American countries included (Argentina, Brazil, Chile, and Mexico), only Mexico has a level of $EXPY$ that is comparable to those in East Asia. This probably reflects the fact that the exports of the other three are heavily based on primary products and natural resources, which tend to have lower $EXPY$s. Chile has the lowest $EXPY$ by far, and its $EXPY$ has been steadily drifting downwards. At the other end, South Korea and Hong Kong have the highest $EXPY$ s. Note how China has significantly closed the gap with these countries over time. China's $EXPY$ has converged with that of Hong Kong, even though Hong Kong's per capita GDP remains five times larger (in PPP-adjusted terms). And China's $EXPY$ now exceeds those of Brazil, Argentina, and Chile by a wide margin, even though China's per-capita GDP is roughly half as large as those of the Latin American countries. India's $EXPY$ is not as spectacular as China's, but that is in large part because our measure is based on commodity exports and does not capture the explosion in India's software exports. Nonetheless, by 2003 India had a higher $EXPY$ than not only Chile, but also Argentina, a country that is roughly four times richer.

Do all natural-resource exporting countries have low $EXPY$s? Figure 6 shows a similar chart for five primary-product exporting countries: Canada, Norway, New Zealand, Australia, and Chile. The variation in $EXPY$ among these countries turns out to be quite large. Once again, Chile is at the bottom of the scale. But even among the remaining four advanced countries, the range is quite wide. Canada's $EXPY$ is between 20-25 percent larger than Norway's or Australia's. Therefore, our measure seems to capture important differences among primary product exporting countries as well.

3.4 Determinants of $EXPY$

What might be some of the fundamental determinants of the variation across countries in levels of $EXPY$? We have shown above that $EXPY$ is highly correlated with per-capita GDP. The model laid out in the early part of the paper suggests that specialization patterns will be determined both by fundamentals and by idiosyncratic elements. Among fundamentals, the model pointed to human capital and the size of the labor force as two key determinants. The first extends the range of "discoverable" goods, and the second promotes cost discovery through (initially) lower wages. We find support for both of these implications in the cross-national data. Human capital and country size (proxied by population) are both associated positively with $EXPY$, even when we control for per capita GDP separately (Table 7). It may be difficult to give the relationship with human capital a direct causal interpretation, since the causal effect may go from $EXPY$ to human capital rather than vice versa. But it is easier to think of the relationship with country size in causal terms: it is hard to believe that there would be reverse causality from $EXPY$ to population size. Interestingly, institutional quality (proxied by the Rule of Law index of the World Bank, a commonly used measure of institutional quality) does not seem to be associated with $EXPY$ once we control for per capita GDP (Table 7, column 3).

Even if we ascribe a causal role to per-capita income and human capital, there is a lot that remains unexplained in the determination of $EXPY$. Figure 7 shows a scatter plot of deviations from the cross-country norms established in column 4 of Table 7 against per capita GDP. There are big outliers in either direction, especially among low-income countries. Mozambique (+88 percent), Swaziland (+55 percent), and Senegal (+29 percent) have $EXPY$ levels that are much higher than would be predicted on the basis of the right-hand side variables in Table 7, while Guinea (-66 percent), Niger (-55 percent), and Burundi (-57 percent) have much lower $EXPY$s. If indeed such differences matter to subsequent economic performance (and we claim that they do), it is important to understand where they arise from. Moreover, to the extent that $EXPY$ levels exert an independent influence on per capita income levels and human capital stocks, the "unexplained" component of the cross-national variation in $EXPY$ is naturally much larger. Hausmann and Rodrik (2003) provide some anecdotal evidence which suggests that successful new industries often arise for idiosyncratic reasons. Fundamentals are only part of the story.

3.5 $EXPY$ and growth

We finally turn to the relationship between $EXPY$ and economic growth. Table 8 shows a set of cross-national regressions in which growth is regressed on initial values of $EXPY$ and other regressors.

The maximum time span that we can use for these regressions based on COMTRADE data is a time horizon of 11 years (1992-2003). However, this leaves us with a sample of only some 40 odd countries. By focusing on a somewhat shorter time horizon–between 1994 and 2003–we can nearly double the sample of countries included in the regression. The table shows results with both samples. All regressions include initial per-capita GDP as a covariate. Human capital is also included in some of the specifications. Finally, we show both OLS and IV results. We appeal to the theory developed previously and the empirical results above in using country size (population and land area) as instruments in the IV specification. Country size is plausible exogenous with respect to $EXPY$ levels, and it is rarely included as an explanatory variable in growth regressions. There is no established theory as to why it should exert an independent effect on economic growth (other than through the $EXPY$ channel which is our focus).

$EXPY$ enters with a large and positive coefficient that is statistically significant in all of these specifications (but only at the 10% level of significance in one instance). The estimated coefficient varies from 0.034 to 0.082, with IV estimates being larger than OLS estimates. Taking the midpoint of this range, the results imply that a 10 percent increase in $EXPY$ boosts growth by half a percentage points, which is quite large. Figure 8 shows a representative scatter plot.

A shortcoming of these regressions is that the time horizon is short, and that they suffer, as with all cross-national specifications, from possible omitted variables bias. While 6-digit aggregation based on COMTRADE does not allow us to examine pre-1992 data, 4-digit calculations based on World Trade

Flows allows us to construct a panel going back to 1962. Table 9 shows results from panel regressions. Data are grouped into 5- and 10-year intervals and four different estimators are used: pooled OLS, IV, OLS with fixed effects (for countries and years), and GMM. (See notes at the bottom of the table for more details.) The estimated coefficient on $EXPY$ is significant in all cases, with a magnitude that is comparable to that in the cross-section results reported above. The fixed effects results are particularly telling, since these explicitly control for time-invariant country characteristics and identify the impact of $EXPY$ off the variation *within* countries. They are significant in both the 5- and 10-year panels. These fixed effects estimates suggest that a 10 percent increase in $EXPY$ raises growth by 0.14 to 0.19 percentage points. This is a smaller effect than what we found in the cross-national specifications, but it is still noteworthy.

4 Concluding remarks

What we have shown in this paper is that there are economically meaningful differences in the specialization patterns of otherwise similar countries. We have captured these differences by developing an index that measures the "quality" of countries' export baskets. We provided evidence that shows that countries that latch on to a set of goods that are placed higher on this quality spectrum tend to perform better. The clear implication is that the gains from globalization depend on the ability of countries to appropriately position themselves along this spectrum.[1]

[1] Some of the policy implications of this is discussed in Rodrik (2004).

5 References

Bernard, A. B., J. Eaton, J. B. Jensen, and S. Kortum, "Plants and Productivity in International Trade," *American Economic Review* 93, 2003, 1268-1290.

Evenson, Robert, and Larry E. Westphal, "Technological Change and Technology Strategy," ch. 37 in Jere Behrman and T.N. Srinivasan, eds., *Handbook of Development Economics,* vol. 3A, Amsterdam, North-Holland, 1995, pp. 2209-2229.

Feenstra, Robert C., et al., "World Trade Flows, 1962-2000," NBER, January 2005.

Hausmann, R. and D. Rodrik, "Economic Development as Self Discovery," *Journal of Development Economics*, December 2003.

Melitz, Marc J., and Gianmarco I.P. Ottaviano, "Market Size, Trade, and Productivity," Harvard University, October 10, 2005.

Rodrik, Dani, "Industrial Policies for the Twenty-First Century," Harvard University, October 2004.

Sachs, J., and A.Warner, "Natural Resource Abundance and Economic Growth," in G. Meier and J. Rauch (eds.), *Leading Issues in Economic Development*, New York: Oxford University Press, 1995.

Sala-i-Martin, Xavier, and Arvind Subramanian, "Addressing the Natural Resource Curse: An illustration from Nigeria," IMF Working Paper, July 2003.

Table 1: Sample size of *EXPY*

Year	*No. reporting countries*
1992	48
1993	65
1994	87
1995	99
1996	111
1997	119
1998	119
1999	126
2000	133
2001	133
2002	127
2003	122

Table 2: Descriptive statistics for *PRODY* (2000 US$)

Variable	*No. obs.*	*Mean*	*Std. Dev.*	*Min*	*Max*
mean *PRODY*, 1999-2001, at market exchange rates	5023	11,316	6,419	153	38,573
mean *PRODY*, 1999-2001, PPP-adjusted	5023	14,172	6,110	748	46,860

Table 3: Largest and smallest *PRODY* values (2000 US$)

	product	*product name*	*mean* PRODY, *1999-2001*
smallest	140490	Vegetable products nes	748
	530410	Sisal and Agave, raw	809
	10120	Asses, mules and hinnies, live	823
	90700	Cloves (whole fruit, cloves and stems)	870
	90500	Vanilla beans	979
largest	721060	Flat rolled iron or non-alloy steel, coated with aluminium, width>600mm	46,860
	730110	Sheet piling of iron or steel	46,703
	721633	Sections, H, iron or non-alloy steel, nfw hot-roll/drawn/extruded > 80m	44,688
	590290	Tyre cord fabric of viscose rayon	42,846
	741011	Foil of refined copper, not backed, t < 0.15mm	42,659

Table 4: Descriptive statistics for *EXPY* (2000 US$)

Year	*Obs.*	*Mean*	*Std. Dev.*	*Min*	*Max*
1992	48	12,994	4,021	5,344	20,757
1993	65	12,407	4,179	3,330	20,361
1994	87	11,965	4,222	2,876	20,385
1995	99	11,138	4,513	2,356	19,823
1996	111	10,950	4,320	2,742	20,413
1997	119	10,861	4,340	2,178	19,981
1998	119	11,113	4,621	2,274	20,356
1999	126	11,203	4,778	2,261	26,218
2000	133	10,714	4,375	1,996	25,248
2001	133	10,618	4,281	2,398	24,552
2002	127	10,927	4,326	2,849	24,579
2003	122	10,664	3,889	2,684	23,189

Table 5: Number of countries that show an increase/decrease in *EXPY*, 1992-2003

	EXPY, ppp	*EXPY, market XRs*
Increase	8	13
Decrease	37	32

Table 6: Countries with smallest and largest *EXPY*s

	Reporter	*EXPY*
Smallest	Niger	2,398
	Ethiopia	2,715
	Burundi	2,726
	Benin	3,027
	Guinea	3,058
Largest	Luxembourg	24,552
	Ireland	19,232
	Switzerland	19,170
	Iceland	18,705
	French Polynesia	18,550

Table 7: Correlates of *EXPY*

	(1)	(2)	(3)	(4)
	Dependent variable: log *EXPY* in 2001			
log GDP per capita	0.354	0.298	0.288	0.282
	(14.75)**	(9.37)**	(6.96)**	(7.47)**
log human capital		0.281	0.268	0.157
		(2.08)*	(1.79)	(1.16)
rule of law index			0.019	0.065
			(0.41)	(1.58)
log population				0.089
				(5.01)**
log land area				-0.032
				(2.30)*
constant	6.090	6.405	6.497	5.523
	(27.39)**	(26.45)**	(18.03)**	(14.66)**
Observations	131	102	101	100
R-squared	0.72	0.75	0.74	0.79

Robust t-statistics in parentheses
* significant at 5% level; ** significant at 1% level

Table 8: Cross-national growth regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dependent variable: Growth of GDP per capita over 1992-2003				Dependent variable: Growth of GDP per capita over 1994-2003			
	OLS	OLS	IV	IV	OLS	OLS	IV	IV
log initial GDP/cap	-0.015	-0.019	-0.017	-0.025	-0.008	-0.013	-0.012	-0.018
	(2.37)*	(2.89)**	(2.47)*	(3.79)**	(1.90)	(2.78)**	(1.41)	(2.54)*
log initial *EXPY*	0.060	0.056	0.072	0.082	0.035	0.034	0.046	0.053
	(3.96)**	(3.83)**	(3.42)**	(3.93)**	(3.05)**	(2.74)**	(1.95)	(2.47)*
log human capital		0.028		0.024		0.021		0.015
		(2.02)*		(1.83)		(2.20)*		(1.41)
constant	-0.419	-0.357	-0.501	-0.550	-0.242	-0.201	-0.305	-0.323
	(4.32)**	(3.68)**	(3.49)**	(3.59)**	(3.15)**	(2.36)*	(2.08)*	(2.35)*
observations	46	43	44	42	85	69	76	68
R-squared	0.35	0.40	0.36	0.36	0.20	0.26	0.20	0.23

Robust t-statistics in parentheses
Instruments for IV regressions: log population, log land area.
* significant at 5% level; ** significant at 1% level

Table 9: Panel growth regressions, 1962-2000

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		5-year panels				10-year panels		
	OLS	IV	FE	GMM	OLS	IV	FE	GMM
log initial GDP/cap	-0.0117	-0.0299	-0.0272	-0.0143	-0.0128	-0.0384	-0.0318	-0.0177
	(4.39)**	(4.78)**	(4.24)**	(2.65)**	(4.42)**	(4.37)**	(5.69)**	(2.37)*
log initial EXPY	0.0287	0.0739	0.0185	0.0446	0.0286	0.0919	0.0141	0.0444
	(5.38)**	(5.06)**	(2.26)*	(4.10)**	(5.22)**	(4.54)**	(1.97)*	(2.29)*
log human capital	0.0068	0.0041	0.0049	0.0035	0.0077	0.0045	0.0038	0.0085
	(3.27)**	(1.76)	(1.08)	(0.92)	(3.75)**	(1.75)	(0.81)	(1.23)
constant	-0.1146	-0.3372	0.0937	-0.2301	-0.1076	-0.4197	0.1640	-0.2023
	(4.08)**	(4.68)**	(1.35)	(3.91)**	(3.68)**	(4.25)**	(2.53)*	(1.75)
observations	604	604	604	604	299	299	299	299
R-squared	0.16	0.05	0.13		0.24		0.28	

Robust t-statistics in parentheses

All equations include period dummies. IV regressions use log population and log area as instruments. Fixed effects(FE) include dummies for countries. GMM is the Blundell-Bond System-GMM estimator using lagged growth rates and levels as instruments. The GMM estimation also use log population and log area as additional instruments.

* significant at 5% level; ** significant at 1% level

Figure 1: The production space

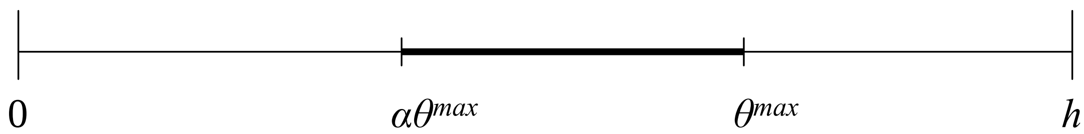

Figure 2: Equilibrium and comparative dynamics

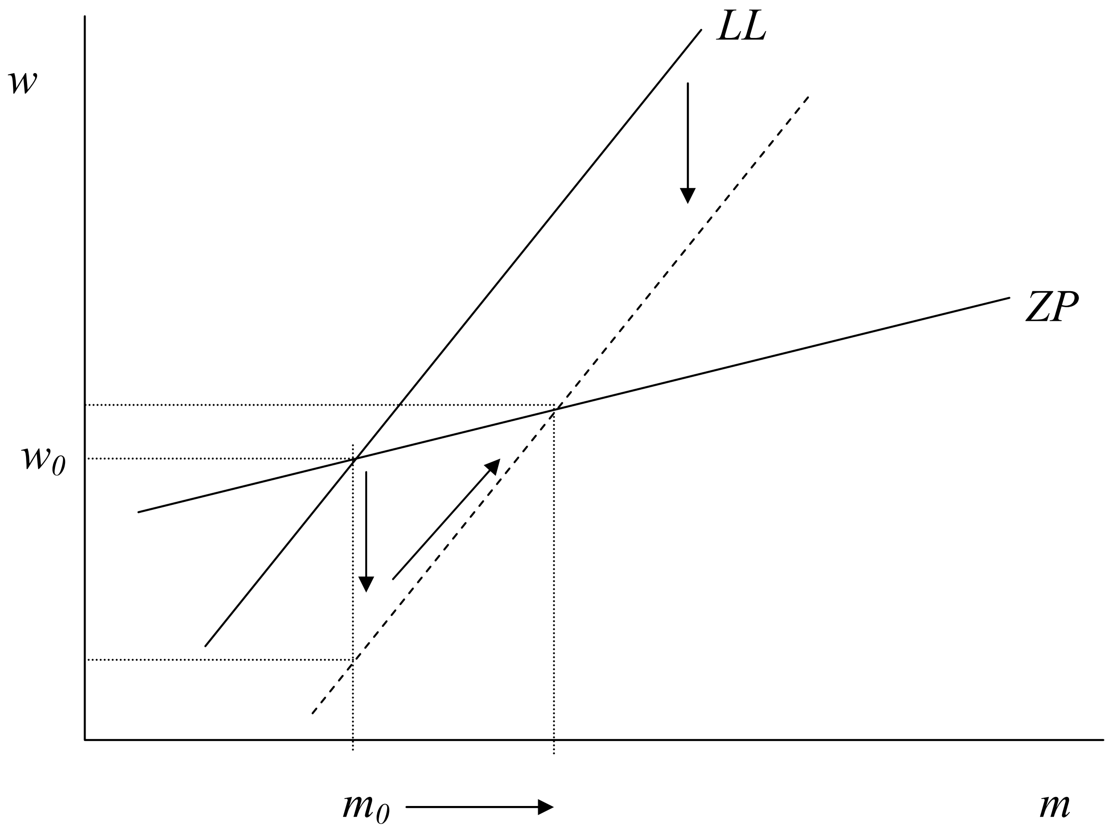

Figure 3: How *EXPY* varies over time

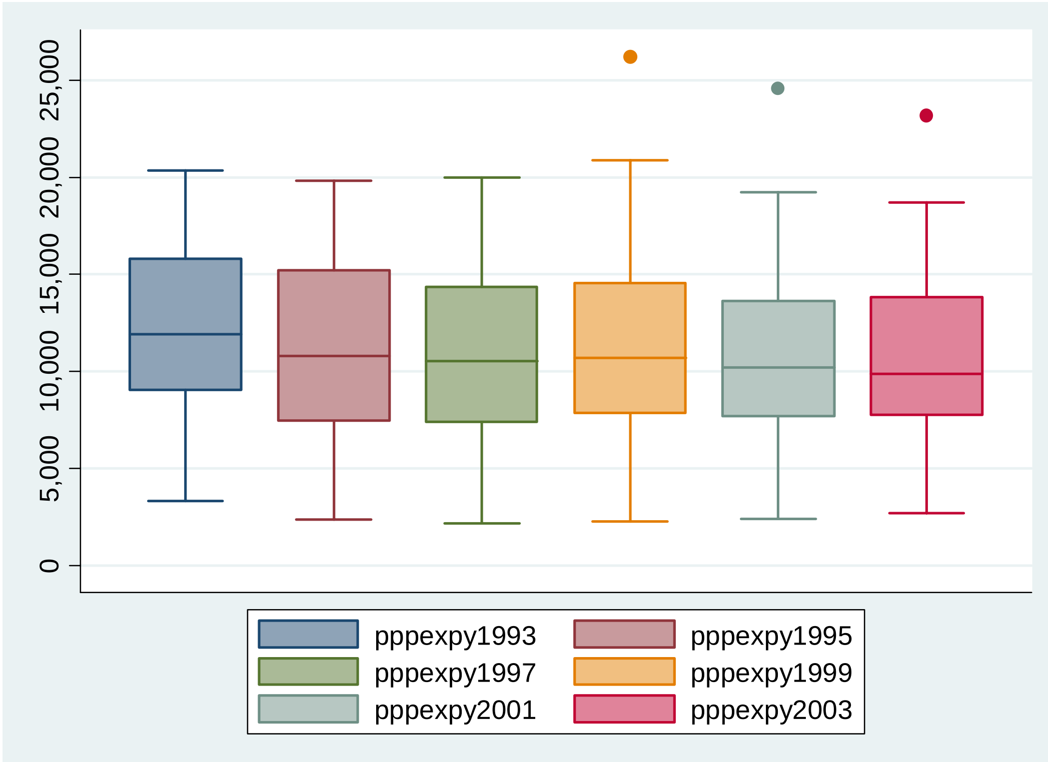

Figure 4: Relationship between per-capita GDP and *EXPY*, 2003

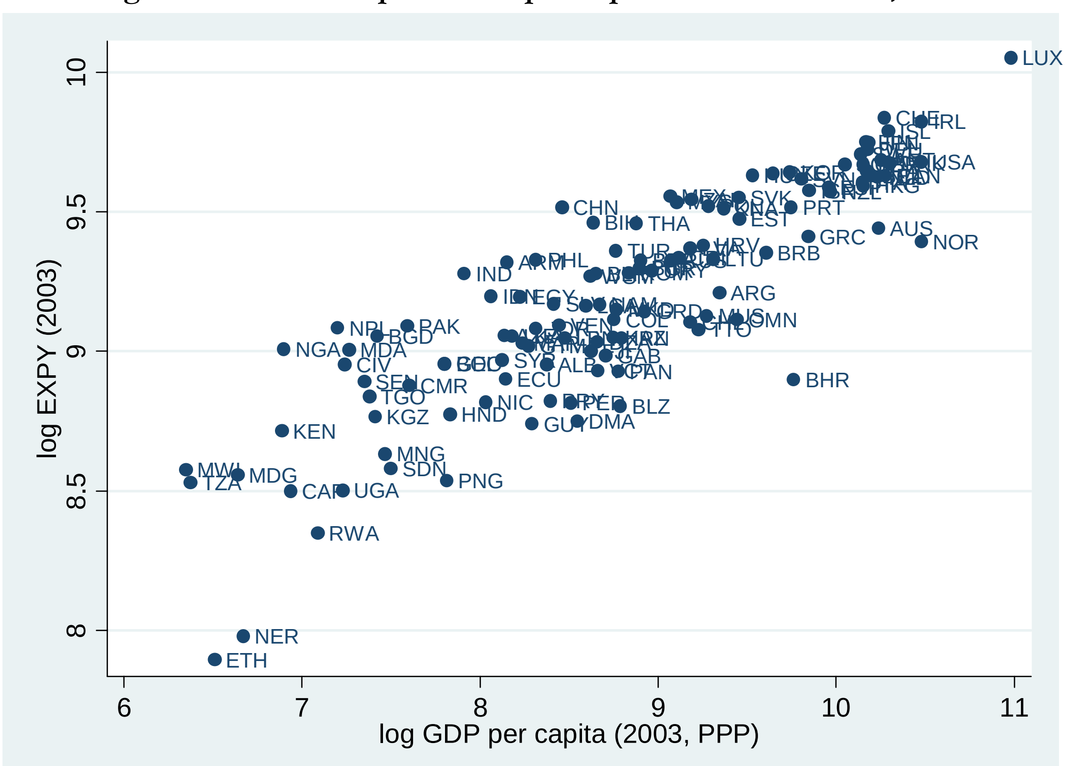

Figure 5: *EXPY* over time for selected countries

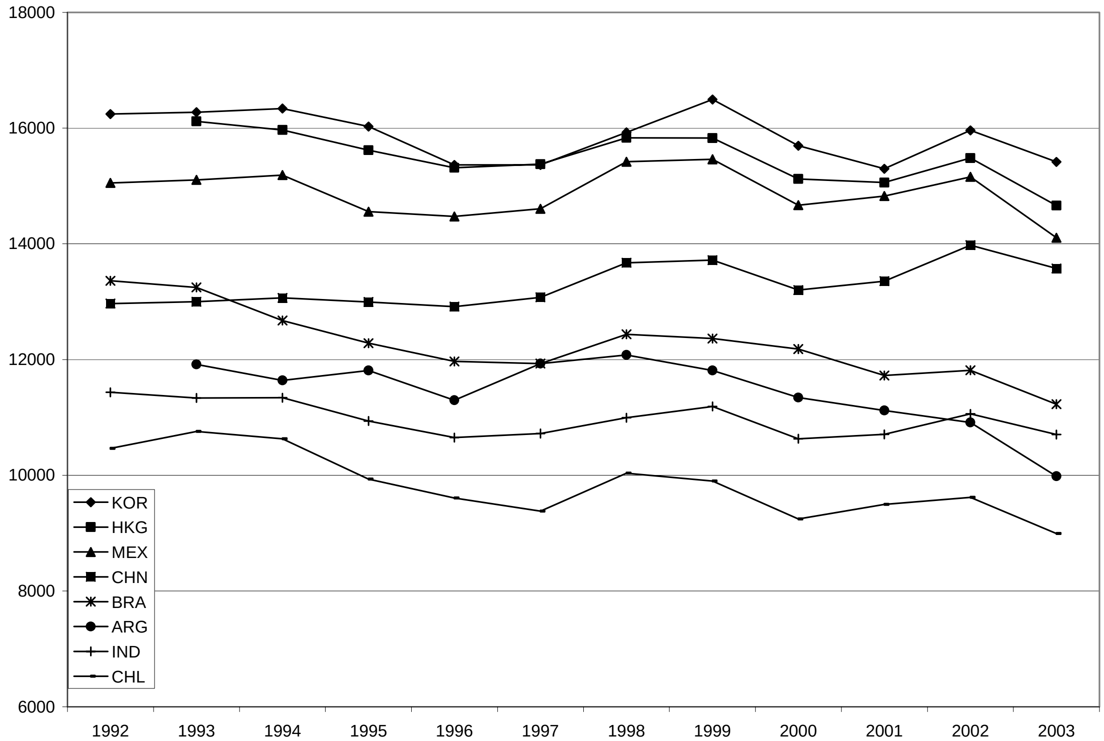

Figure 6: *EXPY* over time for natural-resource exporting countries

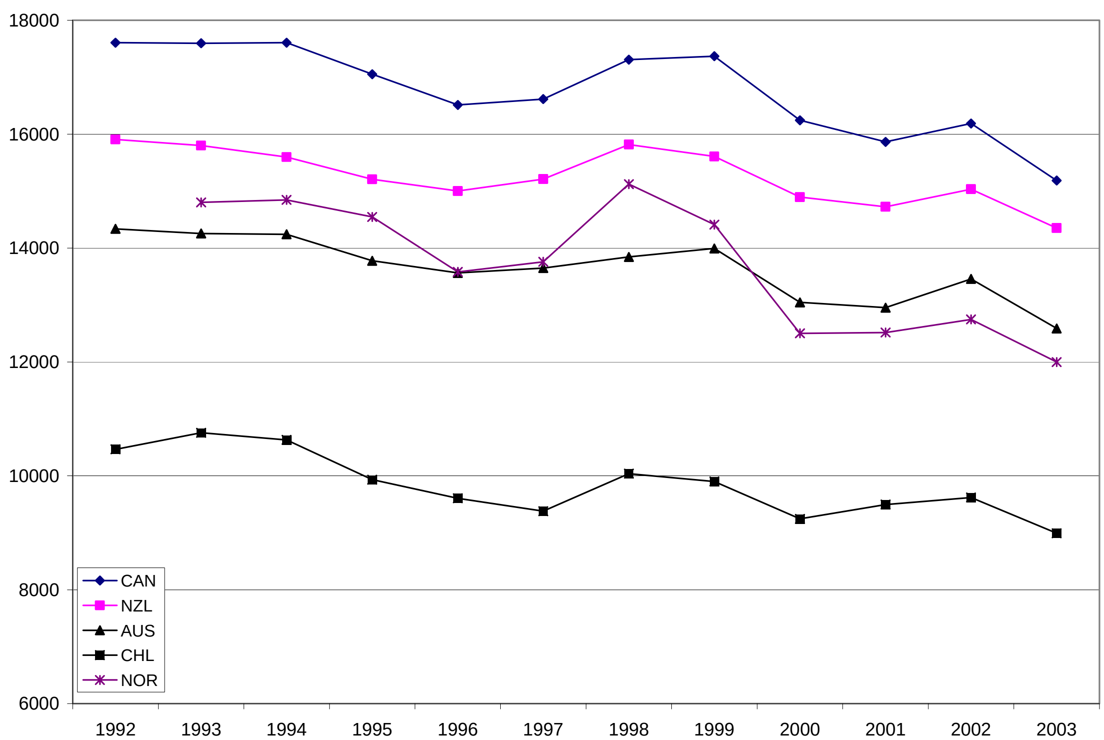

Figure 7: Deviations from cross-national norm for *EXPY*
(Percent differences from the regression specification in Table 7, column 4)

Figure 8: Partial relationship between *EXPY* and subsequent growth (Table 8, col. 5)