

IQ in the Production Function: Evidence from Immigrant Earnings

Garett Jones^{*}

Department of Economics
and
Center for Study of Public Choice
George Mason University

W. Joel Schneider
Department of Psychology
Illinois State University

June 2008

Forthcoming, *Economic Inquiry*

Abstract

We show that a country's average IQ score is a useful predictor of the wages that immigrants from that country earn in the U.S., whether or not one adjusts for immigrant education. Just as in numerous microeconomic studies, 1 IQ point predicts 1% higher wages, suggesting that IQ tests capture an important difference in cross-country worker productivity. In a cross-country development accounting exercise, about one-sixth of the global inequality in log income can be explained by the effect of large, persistent differences in national average IQ on the private marginal product of labor. Taken together with the results of Jones and Schneider (2006) and Hanushek and Kimko (2000), this suggests that cognitive skills matter more for groups than for individuals.

^{*}Corresponding Author. E-mail: gjonesb@gmu.edu. We would like to thank participants at the Missouri Economics Conference, the Southern Economic Association meetings, the Society for Economic Dynamics, DEGIT XI, McGill University, and George Mason University for helpful comments. We especially thank Francesco Caselli, Michael Davis, William Smith, Petia Stoytcheva, Bryan Caplan, editor Vincenzo Quadrini, and an anonymous referee for particularly helpful recommendations, and the Graduate School of Southern Illinois University Edwardsville for financial support. An earlier version of this paper circulated under the title "IQ in the Ramsey Model." The usual disclaimer applies with particular force.

The cross-country growth literature (esp. Sala-i-Martin 1997; Sala-i-Martin, Doppelhofer, & Miller 2004) has found that traditional education measures rarely have a robust relationship with growth and productivity—elementary education being a rare exception. By contrast, a new empirical growth literature (Lynn & Vanhanen 2002; Weede & Kampf 2002; Weede 2004; Jones & Schneider 2006; Ram 2007) has shown that a nation’s average IQ has a remarkably robust relationship with its productivity (Figure 1). That a test designed by *psychologists* should have such a robust relationship with *economic* variables is a puzzle that demands explanation.

For instance, Jones & Schneider showed that national average IQ was more robust than other human capital variables, and was statistically significant at the 1% level in all 455 growth regressions that controlled for all 18 factors that passed Sala-i-Martin *et al.*’s (2004) robustness tests. Of course, as with all of the growth regression literature, a key difficulty is disentangling cause and effect. Thus, in this paper, we run no growth regressions whatsoever. Instead, we perform a simple calibration of the IQ-productivity relationship based on widely-agreed-upon microeconomic parameters. That means we can directly estimate one causal channel running from cognitive ability to productivity. In the process, we learn the following:

1. If one knows the average IQ of a nation’s citizens as estimated by Lynn and Vanhanen (2002, 2006), one can predict the average wages that immigrants from that country will earn upon their arrival in the U.S.—whether or not one controls for immigrant education, and even if the test is completely visual rather than verbal. In other words, national average IQ predicts part of what Hendricks (2002) calls “unmeasured worker skill.”
2. We find that a one point increase in national average IQ predicts one percent higher immigrant wages—*precisely the value found repeatedly in microeconomic studies* (note that by construction, 1 IQ point $\approx 1/15^{\text{th}}$ of a standard deviation within any large national population). Together, points 1 and 2 provide further evidence that cross country IQ tests are valid predictors of worker productivity.
3. When IQ is added to the production function in the form implied by traditional, externality-free human capital theory, differences in national average IQ are quantitatively significant in explaining cross-country income differences. That said, our productivity accounting exercise does not resolve the puzzle of why high-IQ countries are 15 times richer than low-IQ countries.

In a related vein, Hanushek and Kimko (2000) use national math and science test scores to verify that cognitive skills appear to matter more for groups than for individuals. Like

us, they use immigrants to the U.S. as a way to test whether immigrants bring their home-country productivity levels along with them when they immigrate to the U.S.. When they interpret their results within a Solow-type framework, they conclude that “the [cross-country] growth equation results are much larger than the corresponding results for individual earnings.”

In sum, our paper rigorously explores the quantitative magnitude of the puzzle uncovered by Hanushek and Kimko (2000) and Hanushek and Woessman (2007). But by using IQ tests rather than other widely-used math and science test scores, we can often double our sample size while simultaneously using the most widely-analyzed, best-understood form of cognitive test.

We begin with an overview of the recent psychological literature on the validity of IQ tests, and then proceed to our discussion of the link between IQ and immigrant wages. The discussion of IQ and immigrant wages yields a key parameter, γ , the IQ semi-elasticity of wages, which we use in our development accounting exercise. We then discuss the questions of reverse causality and trends in the IQ-productivity relationship over the last 40 years, and conclude by discussing how our results fit into the growth literature.

II. IQ: A psychologist’s perspective

It is not possible to have confidence in these or any other IQ-related findings without an adequate understanding of how IQ is measured and why psychologists believe that well-constructed IQ tests are legitimate tools for the study of cognitive abilities. Unfortunately, space does not permit a comprehensive review of the large research literature that adequately addresses the many reasonable doubts and concerns a properly skeptical reader might have about the validity of IQ tests. Readers wishing for scholarly, balanced, and accessible introductions to intelligence research are advised to consult Bartholomew (2004), Cianciolo & Sternberg (2004), Deary (2001), or Seligman (1992).

Considerable effort has gone into producing nonverbal IQ tests that can be used in any culture. These “culture-fair” or “culture-reduced” IQ tests have been shown to predict important life-outcomes with validity coefficients comparable to traditional IQ tests designed for specific populations (Court 1991; Kendall, Verster, & Von Mollendorf

1988; Rushton, Skuy, & Bons 2004). As we note below, the correlation between national average IQ and GDP per worker is essentially unchanged if we only use data from such “culture-reduced” tests. Unlike traditional IQ tests that measure a very diverse set of cognitive abilities, culture-reduced IQ tests necessarily measure a much smaller number of abilities, focusing on nonverbal reasoning and novel problem-solving. Fortunately, the types of tests that lend themselves to cross-cultural research correlate very highly with the overall scores from traditional IQ tests (Jensen 1998). For our purposes, it does not matter if one believes that IQ tests are valid measures of whatever “real intelligence” is (if there is such a thing as “intelligence”). The tests measure a set of skills that appear to be very advantageous in societies with modern economies. Unlike other measures of human capital such as reading comprehension and mathematical reasoning tests, culture-fair IQ tests have no literacy prerequisites. Because the tests are nonverbal, the test items are the same for everyone and thus results are more comparable across language groups and cultures.

We do not conceptualize culture-fair IQ tests as measures of some immutable quantity that is solely determined by genes. Although it is quite clear that genes play an important role in the development of cognitive abilities, it is equally clear that cognitive abilities are quite sensitive to environmental inputs and can change considerably over the lifespan (Shaie 2005). It is relatively easy to disrupt the delicate processes of the brain with disease, malnutrition, parental abuse and neglect, environmental toxins, and brain injury. With considerable effort, it is also possible to raise IQ somewhat with high-quality personal health care, sound public health policies, adequate nutrition, reasonable parental involvement, and excellent education (Armor 2003). The fact that IQ scores have been rising 0.2 standard deviations per decade in most countries ever since mass IQ testing started in the 1920’s (Dickens & Flynn 2001; Flynn 1987) suggests that in many societies people have increased access to some of these things.

III. IQ as a measure of “unmeasured worker skill.”

In this section, we investigate whether the average IQ in the immigrant’s home country is a useful predictor of the wages of immigrants from that country. Our estimates of immigrant wages come from Hendricks (2002), who used data on earnings, education,

and age from 106,263 immigrants from the 1990 Census of Population and Housing.¹ These immigrants were between the ages of 20 and 69 and worked full-time in the U.S. and had immigrated as adults. For further information on the immigrant data, see Section II and especially Table B1 of Hendricks (2002).

Hendricks extracted systematic wage differences due to education and age by comparing weighted averages of the earnings of native-born and immigrant workers. He did this by creating 10 age categories and 6 education categories for each country's immigrants as well as for U.S. natives. The average immigrant wage per source-country was weighted according to the U.S. distribution of education and age levels. Thus, countries whose emigrants have a low (high) average education level would have the wages of their highly-educated emigrants overweighted (underweighted). For example, immigrants from Taiwan have an average of 15.9 years of education (above the average of native U.S. workers), so the adjustment process would downweight the earnings of Taiwan's highest-educated immigrants, putting more weight on the earnings of those with less than a high-school education.

After thus controlling for age and education, Hendricks concludes that the only remaining explanation for wage differences between workers from different countries is what he calls differences in "unmeasured worker skill."² Hendricks created estimates of unmeasured worker skill for 76 countries.

Perhaps surprisingly, this "unmeasured worker skill" estimate varies widely for immigrants from different countries. The standard deviation of log unadjusted immigrant wages is 0.29 across Hendricks's sample of 76 countries, while the standard deviation of log unmeasured worker skill across these countries is still a sizable 0.19. Henceforth we refer to uws_i , the log of "unmeasured worker skill" in country i .

Our goal in this section is to show that national average IQ is a useful predictor of Hendricks's "unmeasured worker skill." We use Lynn and Vanhanen's (2006) database of national average IQ estimates. We should briefly review how Lynn and Vanhanen (henceforth LV) (2006) created their dataset: They used hundreds of IQ tests from 113

¹ Hendricks's census data on "earnings" combine all forms of income, but we will follow Hendricks's practice and treat them as useful proxies for wages.

² Hendricks addresses the question of immigrant self-selection in detail, and finds little evidence that this is quantitatively important. We refer interested readers to his valuable analysis.

countries across the 20th and 21st centuries in LV (2006). They aggregated these results using best-practice methods to create estimates of “national average IQ” for these countries.³ LV show that the IQ gaps between regions of the world have not appreciably changed during the 20th century.

Lynn and Vanhanen’s 2006 dataset overlaps with 59 of Hendricks’s observations. The mean and median IQ across these 59 countries are both 91 and the standard deviation of IQ across these countries is 9. This is a slightly more intelligent, less varied sample than the full 113 countries: LV’s full-sample mean and median are both 87 and the standard deviation is 12. For comparison, we note that within the United Kingdom, mean IQ is defined as equal to 100, and the standard deviation of IQ within the U.K. population is defined as equal to 15.

Data in hand, we regress $\log uws_i$ onto the level of national average IQ and a constant. The goal is to see whether the estimated relationship between immigrant wages and national average IQ is close to conventional microeconomic estimates of the IQ-wage relationship. In a variety of previous studies, the semi-elasticity of wages (denoted γ) has been close to 1%: Thus, 1 IQ point is associated with 1% higher wages, and a one-standard-deviation rise in IQ is associated with about a 1% rise in wages.⁴ The semi-elasticity γ has a similar magnitude whether one measures in developing countries or in the U.S.. Perhaps surprisingly, Zax and Rees (2002) find that γ appears to rise later in life—so childhood IQ predicts one’s wage better as one gets older—while the coefficient

³ LV made one noteworthy change between their 2002 and 2006 IQ estimates: In cases where they had more than two IQ estimates for a country, they chose the median as their national average IQ estimate rather than their mean.

⁴ For instance, the widely-cited work of Zax and Rees (2002) uses data from Wisconsin to estimate the impact of teenage IQ on lifetime earnings. They find that for men in their 50’s, $\gamma = 0.7\%$ higher earnings when they control for education, and $\gamma = 1.4\%$ when they do not. Since some education is surely caused by prior IQ, and since that education causes higher wages, Zax and Rees note that we should place some weight on the estimates that do not control for education when trying to determine the impact of IQ on wages. They find that IQ—which was measured when these men were teenagers—does a better job predicting wages in a worker’s 50’s than in his 20’s. Neal and Johnson (1996) find that one IQ point is associated with $\gamma = 1.3\%$, while Bishop (1989) finds $\gamma = 1.1\%$.

Cawley et al. (1997) find U.S. estimates in a similar range, even when they break the estimates down by ethnic group and gender—and their estimates drop by about one-third when they control for education. Behrman et al. (2004) survey some developing country studies, and find that the mean and median estimates both imply $\gamma = 0.8\%$. We take $\gamma = 1\%$ as reasonable estimate of best-practice labor econometric work; U.S. estimates often run a bit larger, while developing country estimates and estimates that control for education often run a bit smaller.

risers only by about 1/3 when one controls for education in a typical Mincer wage regression.

Now, let us return to our main question. Do our 59 observations roughly replicate these intra-country estimates of the IQ-wage relationship, where 1 IQ point predicts about 1 percent higher wages? Yes, they do, as seen in Figure 2 and Table 1. When we run a simple bivariate correlation between uws_i and national average IQ, we find a correlation of +0.47, and OLS yields a regression coefficient of $\gamma=0.95$ (White std. error=0.31). This is remarkably close to the coefficient estimates seen elsewhere.

Our estimate, which we round to unity, provides a number of insights. First, it shows that LV's national average IQ measures are useful for predicting more than just cross-country productivity differences, cross-country growth rates (both positive correlations), cross-country suicide rates (also a positive correlation (*sic*): Voracek 2004, 2005), and other cross-country factors. We have now shown that they are also useful for predicting the age-and-education adjusted wages of the average immigrant coming from her home country to the United States.⁵ This is surely evidence that national average IQ is an important predictor of what Hanushek and Kimko (2000) call "labor quality."

Further, we have shown that the estimate is quite close to conventional microeconomic estimates of the IQ-wage relationship.⁶ Whatever an IQ test can tell us about worker wages, it appears to be measuring the same thing across countries as within countries. This is confirmatory evidence that cross-country IQ comparisons are indeed possible, despite the claims of many (e.g., Ehrlich, 2000; Diamond, 1999) to the contrary.

⁵ Vinogradov and Kolvereid (2006) show that Lynn and Vanhanen's national average IQ estimates are good predictors of the self-employment rates of immigrants coming to Norway.

⁶ At first glance, this is surprising: If one thought that workers from low-IQ countries faced enormous hardships, hardships that would impact their level of human capital in ways that wouldn't show up on a so-called "pencil and paper IQ test," then one would expect immigrants from those countries to have much lower earnings upon their arrival in the U.S than an IQ test would predict. In other words, an IQ of 81 for an American citizen would mean something much less serious than an IQ of 81 for a person from Ecuador. The Ecuadorean 81 would likely come bundled with a history of poor nutrition and education, weak public health services, and other adverse factors. Can a mere "pencil and paper IQ test" capture the impact of all of these various insults on a person's wage-earning ability? The answer appears to be yes, on average. So while one might have expected $\gamma \gg 1$ in this cross-country regression, that was not the case. At the same time, one might have expected the OLS estimate of γ to be smaller than 1: If IQ tests in general were a *Mismeasure of Man* (Gould, 1981), then one would expect cross-country IQ tests that were aggregated to the national level and then imputed to the average immigrant to have multiple levels of errors-in-variables problems. This would likely bias the IQ coefficient downward, yielding $\gamma < 1$. But neither turned out to be the case: Our estimated coefficient is quite close to conventional microeconomic estimates.

IV. Robustness Tests: Endogenous Education and Outliers

As mentioned above, Zax and Rees (2002) note that controlling for education may bias the γ coefficient downward. After all, IQ is quite likely to have an impact on the quantity of future education a student acquires, so some of the estimated effect of education on earnings is likely to represent IQ's indirect impact on earnings. As a practical solution, they recommend a simple regression of earnings on IQ alone.

In our case, the equivalent regression would involve regressing Hendricks's "log unadjusted earnings" on IQ. This regression is then the average wage of all Mexicans or all Canadians or all Italians working in the U.S., regressed on the average IQ in that country. This will provide us with an upper bound for IQ's impact on immigrant earnings. In such a regression (Table 2), the correlation coefficient is +0.42, with an OLS regression coefficient of $\gamma=1.3$ (White s.e., 0.44). This is quite close to the upper bound of current estimates found in micro-level panel and cross-sectional studies, and is only 30% larger than our baseline estimate of $\gamma=1$.

Further, our original uws_i results do not appear to be sensitive to obvious outliers. There are three obvious outliers, and all three tend to push γ downward: high-wage South African immigrants, (IQ=72) and low-wage Chinese (IQ=105) and South Korean immigrants (IQ=106); they are the only three with regression residuals more than 2.5 standard deviations away from zero, and all three are in fact over 4 standard deviations away from zero. Thus, they are not small outliers. But are they driving our results from the previous section? It would appear not. One-at-a-time omission of these outliers has a negligible impact on the γ estimate, and eliminating all three raises the coefficient to just 1.4, at the high end of microeconomic estimates.

Another way to check for outliers would be to include dummies for regions of the world that appear to be econometrically "special." Sala-i-Martin (1997) and Sala-i-Martin et al. (2004) found that geographic dummies for East Asia, Latin America, and Sub-Saharan Africa were robust across millions of growth regressions. At the present state of knowledge, it is difficult to know just what these dummies are proxying for; it could be geography, culture, genetics, natural resource availability, persistent political institutions, or many other factors. When we include dummies for these regions (Table

2), we find that our result actually becomes slightly more robust. Thus, these data provide little evidence that a few “special” regions of the world are driving this result.

Overall, our results appear to be robust to endogenous education and to outliers. In our development accounting exercises below, we investigate the implications of imposing various γ values. We tentatively conclude that cross-country IQ measures, as aggregated by LV, are a useful indicator of the private marginal productivity of workers. Cross-country IQ scores pass this “market test” with little difficulty, a result that strengthens our confidence in the validity of cross-country IQ tests as indices of one form of labor quality.

V. IQ in the Production Function

We now turn to the question of whether IQ’s impact on the private marginal product of labor can explain the massive differences in living standards we see across countries. We begin by assuming an IQ-augmented Cobb-Douglas production function,

$$Y_i = K_i^\alpha (e^{\gamma IQ_i} A_i L_i)^{1-\alpha} \quad (1)$$

The subscript i is the country subscript, Y , K , A , and L are output, the capital stock, disembodied technology, and the labor supply respectively, and γ is the semi-elasticity of wages with respect to IQ. In other words, γ is the impact of IQ on human capital. Since our concern is with cross-country comparisons, we suppress time subscripts. We reorganize the production function to make it amenable to development accounting:

$$\left(\frac{Y}{L}\right)_i = A_i e^{\gamma IQ_i} \left(\frac{K}{Y}\right)_i^{\frac{\alpha}{1-\alpha}} \quad (2)$$

This is the equation we use (sometimes in log form) to evaluate the impact of IQ differences on steady-state living standards. Writing in terms of a capital-output ratio is useful since in a Solow or Ramsey growth model the economy heads to a steady-state capital-output ratio that is independent of the level of technology (or by extension, the level of IQ). IQ appears in the production function just as any other form of human capital would. As such, we can estimate IQ’s impact on output in the same way that economists estimate education’s impact on output: By looking at microeconomic

estimates of the link between wages and this form of human capital. Thus, we will repeatedly reuse our $\gamma=1$ estimate from Section 3, but will also consider $\gamma=1.25$ as an upper bound and $\gamma=0.5$ as a lower bound (Bowles et al. 2001), in a meta-study of the labor literature, find a median estimate of $\gamma=0.5$; their meta-study includes all possible studies, without regard to econometric technique).

Before we do so, let us briefly review the power of national average IQ to predict national productivity. Lynn and Vanhanen (2002) found a correlation of 0.7 between national average IQ and the *level* of GDP per worker in 81 countries. Jones and Schneider (2006) found a correlation of 0.82 between national average IQ and log GDP per worker, and also found that national average IQ was statistically significant at the 1% level in 455 cross-country growth regressions that used all of Sala-i-Martin, Doppelhofer, and Miller's (2004) robust growth variables as controls.⁷

In the results below, GDP per worker estimates are from the Penn World Tables. In total, we have complete data for 87 countries that are broadly representative of the world's economies. Data and software are available upon request, and the raw data underlying Lynn and Vanhanen's IQ estimates are readily available in table form on the web (Sailer 2004). The Sailer website's charts are especially useful for demonstrating that these IQ differences have been persistent and do not turn on the type of IQ test employed.

VA. IQ differences: Magnitude

In this section, we combine the IQ-augmented production function (2) with conventional parameter values for γ to illustrate how IQ differences can impact steady-state living standards. Consider two countries that differ only in average IQ—i.e., their levels of technology and their capital-output ratios are equal across countries. The ratio of living standards in these two countries would then be:

⁷ Does this strong IQ-productivity correlation depend on the type of IQ test used? Apparently not, if we look at the IQ tests underlying Lynn and Vanhanen's (2002) estimates. For example, looking only at the 25 scores (out of the 163 total in their 2002 book) that used Cattell's Culture-Fair test, the correlation with 1998 PPP-adjusted log GDP per capita was 0.74, slightly below the 0.82 in the aggregated sample. For one form of Raven's Progressive Matrices (a non-verbal, visual pattern-finding IQ test), the correlations were 0.92 (35 tests), and for the other form of the Raven, the correlation was 0.69 (53 tests). These were the only three tests appearing more than 25 times in the LV (2002) database. Clearly, regardless of the type of test used, national average IQ can still predict about half or more of a nation's productivity.

$$\frac{(Y/L)_{hi}}{(Y/L)_{lo}} = e^{\gamma \Delta IQ} \quad (3)$$

where ΔIQ is the difference in IQ between the two countries. Lynn and Vanhanen (2006) show that if countries are ranked according to IQ, then the country in the 5th percentile has an estimated average IQ of 66, while the country in the 95th percentile has a median IQ of 104. This yields an IQ gap of 38 points—a bit more than two standard deviations if one were looking within the U.S. population. As noted above, we take $\gamma = 1$ as our preferred estimate; under this assumption a rise of 1 IQ point raises wages (and hence the marginal product of labor) by a modest 1%.

Therefore, as Figure 3 illustrates, if a country moved from the middle of the bottom IQ decile to the middle of the top IQ decile (a rise of 38 points), steady state living standards would be about 1.5 times greater in the higher-IQ country ($e^{0.38} \approx 1.46$). This compares to the factor of 2 commonly cited for the impact of cross-country differences in education on productivity—some of which may in fact reflect differences in IQ endogenously driving education choices. If the true γ were equal to 1.25, toward the high end of current estimates, a 38-IQ-point gap would raise living standards by a multiple of 1.64. And if γ were half our preferred estimate, as denoted in the lowest of the three lines, a 38-point IQ gap would cause living standards to diverge by a factor of 1.23.

But perhaps the 5th and 95th percentiles are outliers, driven by test error or idiosyncratic environmental factors. Therefore we look at the 90/10 and 80/20 ratios. The gap between the 90th and 10th percentiles is 31 IQ points (102 and 71 points), while the gap between the 80th and the 20th percentiles is 21 IQ points (99 and 78 points). In these cases, productivity levels between these countries in the $\gamma = 1$ case would differ by a bit more than 30% and a bit more than 20%, respectively.

Since living standards across countries differ by perhaps a factor of 30, and since the natural log of 30 is about 3.4, then if $\gamma = 1$, the channel running from national average IQ to the private marginal product of labor explains perhaps $0.46/3.4$, or a bit less than $1/6^{\text{th}}$ of the log difference in living standards across countries.

We should note these development accounting results do not depend on IQ being exogenous. We suggest below that simple reverse causality (running from productivity

to IQ) is unlikely to be the main explanation for the strong empirical IQ-productivity relationship. However, even if reverse causality *were* important, the development accounting results would still hold, since microeconomic studies demonstrate convincingly that IQ has an independent impact on the marginal product of labor.

VB: Calibration Results

The calibration exercise is quite straightforward, and is similar to that of Dhont and Heylen (2008). For each country in the dataset, we predict the level of GDP per worker using equation (2), assuming that technology and capital per worker are identical across countries. Thus, the only source of cross-country productivity differences is IQ working through the narrow channel of the private marginal product of labor, γ . This gives us predicted values for 87 countries, which we then compare to the actual values of GDP per worker. We compare this prediction again the actual level of GDP per worker for each country.

As a goodness-of-fit measure, we use R^2 . This R^2 is the percentage of the global income distribution that can be explained through a single channel: the steady-state impact of differences in national average IQ on labor productivity by way of the private marginal product of labor, γ . For reference, note that the R^2 between log GDP per worker in 2000 and Lynn and Vanhanen's (2006) national average IQ estimate is 58%, and in an cross-country OLS regression, 1 IQ point is associated with 6.7% higher GDP per worker (Figure 1).

Results are reported in Table 2.^{8,9} For the preferred parameter value of $\gamma=1.0$, IQ can explain 16% percent of log cross-country income variation. Therefore, IQ's impact

⁸ Results were substantially unchanged if 2000 log GDP per person was used instead of log GDP per worker. They were also substantially unchanged if national average IQ was windsorized at values of 70, 80 or even 90 IQ points (first recommended by McDaniel and Whetzel (2004)). For example, IQ scores less than 70 were set equal to 70, and the estimates were substantially unchanged when rerun. This windsorizing addresses the concern than IQ scores from the poorest countries are "too low to believe": Even if we bump the lowest scores up a few (dozen) points, the results still hold.

⁹ Results were likewise substantially unchanged if we omitted the 8 observations that Jones and Schneider (2006) also omitted. They omitted observations from LV's dataset that were based on fewer than 100 test subjects per country or that relied exclusively on immigrant data. They also omitted two observations (Peru and Columbia) that partially relied on imputing IQ scores based on the average IQ's of residents of nearby countries. Omitting these possibly weaker data points had no substantial effect on the results.

on wages would explain 29% (i.e., 16%/58%) of the relationship between IQ and log productivity.

If, instead, IQ has a 25% larger impact on wages ($\gamma=1.25$) then IQ's effect on wages can explain 20% of the variance in log productivity and 34% ($=20\%/58\%$) of the IQ/log productivity relationship. And even if $\gamma=0.5$ --half of our preferred estimate--IQ's impact on wages explains 8% of the log global income distribution. So even under unusually conservative assumptions, IQ's impact on the private marginal product of labor appears to belong on any top 20 list of explanations for cross-country income differences.

VI: Addressing Reverse Causality

The quantitative results of the last two sections imply that differences in national average IQ are substantial drivers of global income inequality. Can simple reverse causality explain this relationship? In other words, does a dramatic rise in GDP per worker cause a dramatic rise in national average IQ?

The region of the world that has witnessed the most rapid increases in living standards the world has ever known is unambiguously East Asia. Surely, this region would be an ideal testing ground for the productivity-causes-IQ hypothesis. If most of the IQ-productivity relationship were reverse causality, then we would expect to see the East Asian economies starting off with low IQ's in the middle of the 20th century, IQ's that would rapidly rise in later decades, perhaps even converging to European IQ levels. In short, one would expect to see Solow-type convergence in national average IQ.

However, this is not the case. Lynn and Vanhanen's (2006) country-level IQ data shows that average East Asian IQ's were *never* estimated below 100 before the 1980's (Figure 4). These IQ scores come from South Korea, Japan, Hong Kong, China, and an East Asian offshoot, Singapore. In all cases, IQ scores are above 100—even in a poor country like China. Thus, East Asians both started and ended the period with high IQ scores.

Another place to look for massive IQ increases would be in a region of the world that experienced a dramatic increase in the price of its exports: The oil-rich countries of the Middle East. But a glance at that data, likewise, shows little evidence that being richer, *per se*, increases IQ within ten or twenty years:

Year	IQ	Country
1957	77	Egypt
1957	82	Lebanon
1959	84	Iran*
1972	81	Egypt
1972	83	Iran*
1972	87	Iraq*
1972	87	Iraq*
1987	80	Iran*
1987	84	Jordan
1987	78	Qatar*
1989	83	Egypt
1992	89	Iran*
1997	85	Yemen
2005	86	Kuwait*

(*: OPEC countries)

If one uses 1973 as a breakpoint—since real oil prices increased fourfold between 1973 and 1986, before declining—then one would expect IQ scores to be higher in oil-rich countries if simple reverse causality drove IQ scores. Casual inspection of the evidence doesn’t show such a relationship—indeed, Qatar and Kuwait, two low-population, high-GDP-per-capita countries, fail to stand out along the IQ dimension.

Further, after 1973, there is no clear difference between OPEC and non-OPEC countries, contrary to what one would expect if income caused IQ in an important way. Finally, a simple difference-in-difference test shows that OPEC countries have a median IQ score falling 5.5 points *lower* compared to non-OPEC countries after 1973 (Given the small sample size, we will refrain from calculating standard errors—consider these results as suggestive). All told, if one wants to use a reverse causation argument to explain the IQ-productivity relationship, it will have to be more subtle than the simple tests of East and Southwest Asian IQ’s presented here.

VII. Conclusion

Hendricks (2002) showed convincingly that workers from different countries have different average levels of what he calls “unmeasured worker skill.” We have provided evidence that conventional, out-of-the-box IQ tests can measure an important part of that heretofore unmeasured skill. This supports the claims of Lynn and Vanhanen (2002,

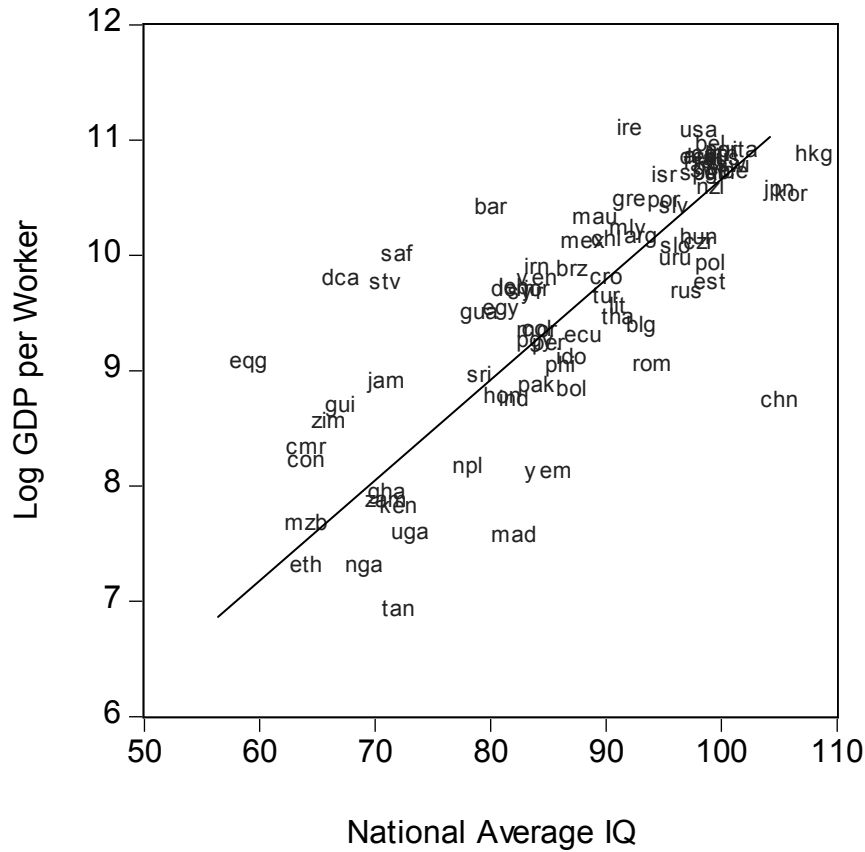
2006) that national average IQ is an important determinant of economic outcomes across countries.

We have further shown that the between-country coefficient of the IQ semi-elasticity of wages, γ , is essentially identical to the within-country coefficient, and we have used that fact to conduct a conventional, externality-free development accounting exercise. In such an exercise, we found IQ's impact on productivity to be quantitatively modest: It explains about $1/6^{\text{th}}$ of the variance in log productivity between countries, and about $1/6^{\text{th}}$ of the predicted steady-state relationship between IQ and log productivity.

To put this in perspective, note that if a nation moved from the 5^{th} to the 95^{th} percentile of national average IQ, our development accounting exercise predicts that its output per worker would rise by perhaps 50%. But in reality, these countries have living standards that differ by a factor of 15, not 1.5. We hope that future research investigates why these relatively modest IQ differences between countries are correlated with such massive differences in national living standards.

We also hope that economists can bring their powerful econometric and theoretical tools to bear on the question of why IQ gaps across poor countries are so large. If economists can find ways to narrow these persistent IQ gaps, the world's poorest citizens may be able to make full use of their productive potential.

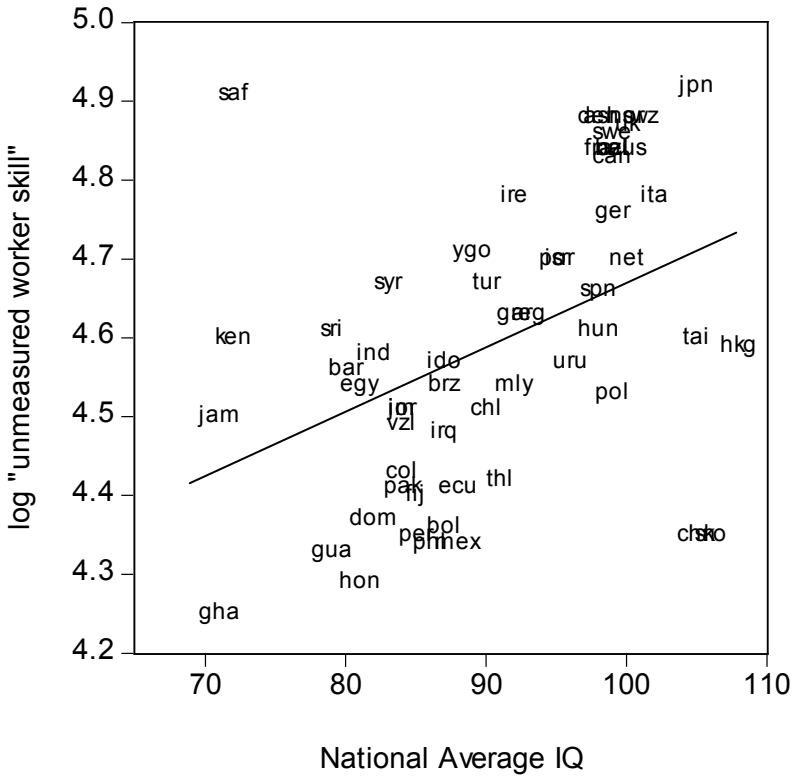
Figure 1
National Average IQ
and Year 2000 GDP Per Worker



Notes: Y-axis shows GDP per worker in logarithmic scale. In this bivariate regression, the coefficient on national average IQ is 0.067, and the R^2 is 58%. Thus, a one-point rise in IQ is associated with 6.7% higher output per worker. The sample covers 87 countries. The outlier in the lower-right corner is China (IQ=105). High outliers, from left to right, are Equatorial Guinea, Dominica, St. Vincent, South Africa, and Barbados.

Source: Lynn and Vanhanen (2006) and Penn World Tables 6.1 for IQ and GDP data, respectively.

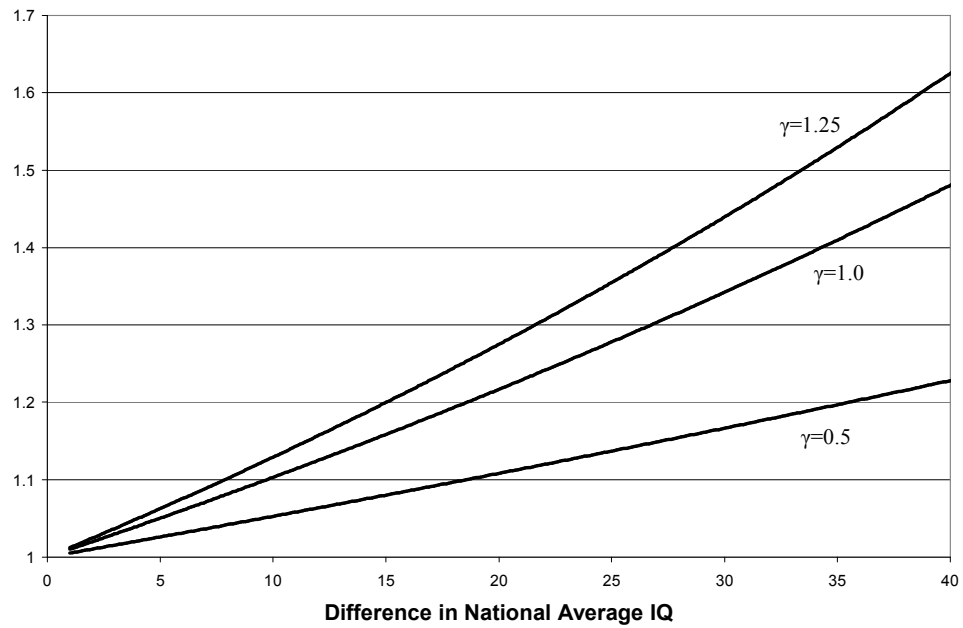
Figure 2
IQ and Immigrant Skill



Notes: The x-axis reports estimates for national average IQ for country i from Lynn and Vanhanen (2006). The y-axis reports values for uws_i , the unmeasured worker skill estimate for immigrants from country i , as estimated in Hendricks (2002). uws_i is the log average wage of immigrants for country i , adjusting for age and education. The trendline reflects the OLS coefficient of 0.95 reported in the text, and the R^2 is 22%.

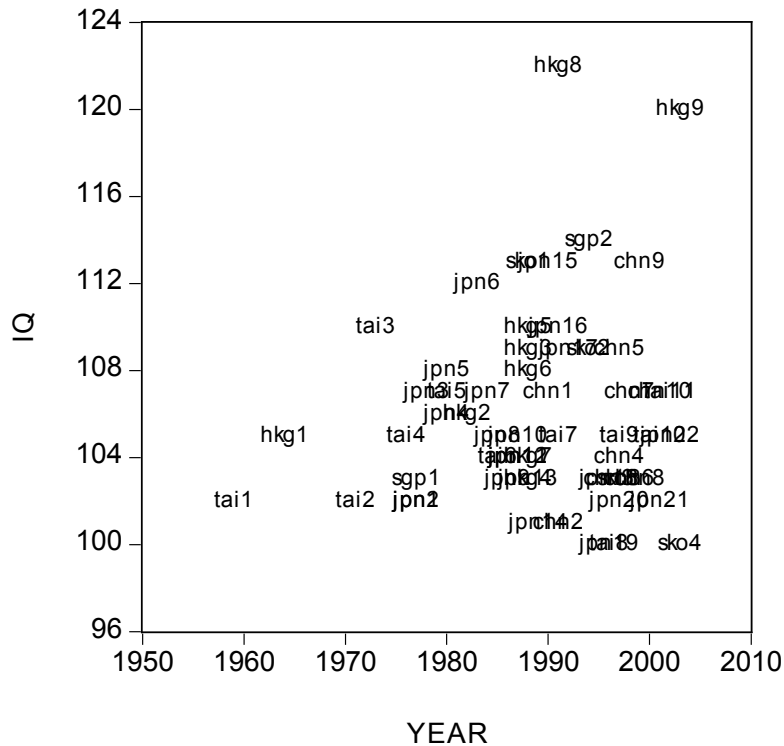
Figure 3

IQ's impact on Steady-State Living Standards



Notes: The value on the y-axis is $(Y/L)_{hi}/(Y/L)_{lo}$, the ratio of living standards in two countries who differ only in national average IQ. This chart is based on equation (3).

Figure 4
IQ in East Asia: 1959-2003



Source: Lynn and Vanhanen (2006). Data labels indicate country and sequence order. Thus, hkg9 is the 9th IQ test given to a Hong Kong sample as reported in LV (2006).

Table 1: National IQ as a predictor of immigrant earnings

	Coeff	S.E.	P-value	R²
<i>Dependent Variable: Log Unmeasured Worker Skill (i.e., education- and age-adjusted earnings)</i>				
IQ	0.95%	0.31%	0.3%	23%
Controls: None				
IQ	1.16%	0.35%	0.2%	41%
Controls: East Asia/Sub-Saharan Africa/Latin America dummies (from Sala-i-Martin 1997, Sala-i-Martin et al. 2004)				

	Coeff	S.E.	P-value	R²
<i>Dependent Variable: Log unadjusted earnings</i>				
IQ	1.30%	0.44%	0.5%	18%
Controls: None				

Note: 59 Observations; White standard errors.

Table 2: Log productivity variance explained by IQ's impact on wages

$\gamma=0.5$	$\gamma=1.0$	$\gamma=1.25$
8%	16%	20%

Notes: γ is the semi-elasticity of skill/wages with respect to IQ. The percentages indicate the variance in year 2000 log GDP per worker that can be explained by IQ's steady-state impact on the private marginal product of labor, as set forth in equation (3). These calibrations are based on data from 87 countries.

For reference, the R^2 from a simple regression of year 2000 log GDP per worker on national average IQ is 58%. IQ and GDP data are from Figure 1.

Data Appendix 1: IQ and Earnings Data

Country	National Average IQ	Log Adjusted Earnings
Argentina	93	4.63
Australia	98	4.88
Austria	100	4.84
Barbados	80	4.56
Belgium	99	4.84
Bolivia	87	4.36
Brazil	87	4.54
Canada	99	4.83
Chile	90	4.51
China	105	4.35
Colombia	84	4.43
Denmark	98	4.88
Dominican Republic	82	4.37
Ecuador	88	4.41
Egypt	81	4.54
Fiji	85	4.40
France	98	4.84
Germany	99	4.76
Ghana	71	4.25
Greece	92	4.63
Guatemala	79	4.33
Honduras	81	4.29
Hong Kong	108	4.59
Hungary	98	4.61
India	82	4.58
Indonesia	87	4.57
Iran	84	4.51
Iraq	87	4.48
Ireland	92	4.78
Israel	95	4.70
Italy	102	4.78
Jamaica	71	4.50
Japan	105	4.92
Jordan	84	4.51
Kenya	72	4.60
Malaysia	92	4.54
Mexico	88	4.34
Netherlands	100	4.70
New Zealand	99	4.84
Norway	100	4.88
Pakistan	84	4.41
Peru	85	4.35
Philippines	86	4.34
Poland	99	4.53
Portugal	95	4.70
South Africa	72	4.91
South Korea	106	4.35
Spain	98	4.66
Sri Lanka	79	4.61
Sweden	99	4.86
Switzerland	101	4.88
Syria	83	4.67
Taiwan	105	4.60
Thailand	91	4.42
Turkey	90	4.67
United Kingdom	100	4.87
Uruguay	96	4.57
Venezuela	84	4.49
Yugoslavia	89	4.71

Notes: National Average IQ data are from Lynn and Vanhanen (2006). Adjusted earnings data are from Table B1 of Hendricks (2002), and draw on the 1990 U.S. census.

Data Appendix 2: Reliability of IQ scores

Since the Lynn and Vanhanen (2002, 2006, henceforth LV) IQ scores have been used in only a few papers in the economics literature, some effort to measure the reliability of their database is warranted. In LV (2006), over 300 IQ tests from 113 nations are used. Their database combines many types of IQ tests from the purely visual, multiple-choice Raven's Progressive Matrices to the three-hour-long Wechsler, which is always given one-on-one by a professional psychometrician. When LV have multiple IQ estimates for a country, they choose the median score. LV's data come from a variety of sources, but the two most important categories are "standardization samples" and individual published studies. Standardization samples are typically created by publishers of IQ tests to learn about the first, second, and higher moments of the distribution of IQ scores within a particular national population. By doing so, they can convert a raw test score into a percentile ranking within a national population.

For example, Angelini *et al.*, (1998) gave over 3000 Brazilian children the Raven's Colored Progressive Matrices test (a simple multiple-choice visual IQ test, but a powerful predictor of overall IQ). In creating a standardization sample, psychometricians attempt to create a genuinely representative sample, so that their product—purchased by school districts around the world—will build a good reputation and find many customers. In LV (2006), 25 countries have at least one score from a standardization sample. In LV (2006) most standardization samples are from Ravens-type tests.

The other individual published studies tend to come from "opportunity samples," perhaps a classroom or a school district near the researchers. Some such studies are simply an attempt to document how typical children perform on one type of IQ test, while other studies look into how nutrition, level of schooling, environmental lead, or other forces impact an individual child's IQ. An important question is how the "best" studies—the standardization samples—compare to the "rest." Within a country, are the standardization scores similar to the individual study scores? And are the standardization scores similar to LV's country-level estimated national average IQ?

The answer to both questions appears to be "yes." We assembled data from countries that had at least one standardization sample IQ estimate plus at least one individual published study IQ estimate. In the "standardization" category, we also include five Latin American estimates from a UNESCO (1998) IQ test of verbal reasoning; each country had a sample of 4000 students. Omitting these observations had no noticeable impact on the results below. When a country had more than one standardization sample (common in rich countries plus India), we took the median standardization sample IQ score and compared it pairwise against the other lower-quality IQ scores.

We arrive at a total sample of 23 countries and 63 comparisons. The mean absolute deviation between a country's median standardization score and that country's other lower-quality scores is 3.2 IQ points, 1/5 of a standard deviation within the U.S. (the standard deviation is 4.4 IQ points). Therefore, it will be rare for a lower-quality IQ

score to be off by 15 points, a full U.S. standard deviation. The mean deviation is -0.2 IQ points, so the lower-quality scores display negligible bias, with standardization scores ever-so-slightly lower than other scores. The correlation between high- and lower-quality IQ scores is 0.9, but since this sample is weighted toward the higher-IQ countries where there is little variation in IQ scores, a correction for restriction of range would raise this correlation even higher.

Since both high-quality and lower-quality samples appear to tell roughly the same story about a country's IQ, there is little to be gained from painstakingly creating a standardization sample for every country: "The best" differ little from "the rest." This finding shows up in the LV's estimated national average IQ: The mean absolute deviation between the median standardization score and LV's national average IQ is a negligible 1.1 IQ points (standard deviation 1.9 IQ points) while the mean deviation is 0.1 IQ points. So for countries where we have standardization scores for comparison, LV's method of choosing the median IQ score is quite similar to choosing the highest-quality score.

Another important question is how the Flynn effect impacts these scores: might the Flynn correction introduce some bias? All scores used thus far in this paper employ only Flynn-adjusted IQ scores. LV's Flynn adjustment uses 1979 as a base year, and following the best-practice in the literature, assumes that scores on the Raven's increase by 3 IQ points per decade and increase by 2 points per decade on all other IQ tests. LV (2002) report both Flynn-adjusted IQ scores and raw IQ scores: using that data we find that the correlation between LV's national average IQ and year 1998 log GDP per worker (Penn World Tables) is 0.83 with unadjusted scores and 0.85 with Flynn-adjusted scores, a minor difference.

Indeed part of the reason Flynn adjustments cannot matter much is because both poor and rich countries have IQ scores going back many decades—so on average, the Flynn adjustment impacts all types of countries about equally. Therefore even if the Flynn adjustments are incorrect, they combine an irrelevant shift in intercept with a shift in slope. A mere glance at the datasets in LV (2002) will be enough to convince many readers that Flynn corrections are unlikely to be relevant; Sailer (2004) has put the LV (2002) database into a convenient online tabular format.

Finally, there is the question of whether the quality of the IQ scores impacts the immigrant wage results reported in this paper. Apparently, the answer is no: When we use only standardization sample and UNESCO (1998) scores, we have observations for a mere 21 countries, but an OLS regression finds that 1 IQ point predicts 1.2% higher immigrant wages ($p=0.02$, $\text{corr}=0.5$), similar to the results reported using the full data sample. Similarly, at the cross-country level, one "high-quality" IQ point predicts 8% higher national GDP per worker ($p=0.0001$, $\text{corr}=0.7$, $n=27$). So even using high-quality national average IQ estimates, IQ predicts small within-country productivity differences but large cross-country productivity differences. This replicates the central finding of this paper.

Bibliography

- Angelini, A.L., Alves, I.C., Custodio, E.M., and Duarte, W.F. (1988). *Manual Matrices Progressivas Coloridas*. Sao Paulo: Casa do Psicologo.
- Armor, D. J. (2003). *Maximizing intelligence*. New Brunswick: Transaction Publishers.
- Bartholomew, D.J. (2004). *Measuring intelligence: Facts and fallacies*. Cambridge: Cambridge University Press.
- Behrman, J., Alderman, H. & Hoddinott, J. (2004). *Copenhagen Consensus – Challenges and Opportunities: Hunger and Malnutrition*.
<http://www.copenhagenconsensus.com/Default.asp?ID=223>
- Bishop, J. H. (1989). “Is the Test Score Decline Responsible for the Productivity Growth Decline?” *American Economic Review*, 79, March, 178-197.
- Bowles, S., Gintis, H. & Osborne, M. (2001). “The Determinants of Earnings: Skills, Preferences, and Schooling,” *Journal of Economic Literature*, 39, 1137–1176.
- Cawley, J., Conneely, K., Heckman, J. & Vytlačil, E. (1997). “Cognitive Ability, Wages, and Meritocracy,” in B. Devlin, S. E. Fienber, D. P. Resnick, & K. Roeder (Eds), *Intelligence, Genes, and Success: Scientists Respond to The Bell Curve*. New York: Springer-Verlag.
- Cianciolo, A. T. & Sternberg, R. J. (2004). *Intelligence: A brief history*. Malden, MA. Blackwell Publishing.
- Court, J. H. (1991). “Asian applications of Raven's Progressive Matrices,” *Psychologia*, 34, 75-85.
- Deary, I. J. (2001). *Intelligence: A very short introduction*. Oxford: Oxford University Press.
- Dhont, Tine and Freddy Heylen, (2008). “Why do Europeans work (much) less? It is taxes and government spending,” *Economic Inquiry*, 46(2), April, 197-207.
- Diamond, J. (1999). *Guns, Germs, and Steel: The Fates of Human Societies*. New York: W.W. Norton.
- Dickens, W. T., & Flynn, J. R. (2001). “Heritability estimates versus large environmental effects: The IQ paradox resolved,” *Psychological Review*, 108, 346–369.
- Ehrlich, P. (2000). *Human Natures: Genes, Culture, and the Human Prospect*. Washington, DC: Island Press.
- Flynn, J. R. (1987). “Massive IQ gains in 14 nations,” *Psychological Bulletin*, 101, 171-191.
- Gould, S. J. (1981). *The mismeasure of man*. New York: W.W. Norton/Harmondsworth, U.K.: Penguin Books.
- Hanushek, E. & Kimko, D. (2000). “Schooling, Labor Force Quality, and the Growth of Nations,” *American Economic Review*, 90, 1184-1208.
- Hanushek, E. & Woessman, L. (2007). “The Role of School Improvement in Economic Development,” NBER Working Paper 12832, January.
- Hendricks, L (2002). “How Important is Human Capital for Economic Development? Evidence from Immigrant Earnings.” *American Economic Review*, Vol. 92:1, pp. 198-219.

- Heston, A., Summers, R. & Aten, B., (2002). *Penn World Table Version 6.1*, Center for International Comparisons at the University of Pennsylvania (CICUP), October 2002.
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.
- Jones, G. & Schneider, W. J. (2006). "Intelligence, Human Capital, and Economic Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach," *Journal of Economic Growth*, 11:71-93.
- Kendall, I. M., Verster, M. A. & Von Mollendorf, J. W. (1988) Test performance of blacks in Southern Africa. In S. H. Irvine and J. W. Berry (Eds.), *Human abilities in cultural context* (pp. 299-339). Cambridge, UK: Cambridge University Press.
- Lynn, R. & Vanhanen, T. (2002). *IQ and the Wealth of Nations*. Westport, CT: Praeger Publishers.
- Lynn, R. & Vanhanen, T. (2006). *IQ and Global Inequality*. Augusta, GA: Washington Summit Publishers.
- McDaniel, M.A. & Whetzel, D.L. (2004). *IQ and the Wealth of Nations: Prediction of GDP Not Dependent on Precise g Estimates*. Working paper, Virginia Commonwealth University.
- Neal, D. A. & Johnson, W. R. (1996). "The Role of Premarket Factors in Black-White Wage Differences," *Journal of Political Economy*, 104, 5, 869-895.
- Neisser, U. (1998). *The Rising Curve*. Washington, DC: American Psychological Association.
- Ram, R. (2007). "IQ and economic growth: Further augmentation of Mankiw-Romer-Weil model." *Economics Letters*, 94:1, pages 7-11.
- Rushton, J. P., Skuy, M. & Bons, T. A. (2004). "Construct Validity of Raven's Advanced Progressive Matrices for African and Non-African Engineering Students in South Africa." *International Journal of Selection and Assessment*, 12, 220-229.
- Sala-i-Martin, X. (1997). "I just ran two million regressions," *American Economic Review*, 87(2), 178-183. Data available at www.columbia.edu/~xs23.
- Sala-i-Martin, X., Doppelhofer, G. & Miller, R. (2004). "Determinants of long-run growth: A Bayesian averaging of classical estimates (BACE) approach," *American Economic Review*, 94(4), 813-835.
- Sailer, S. (2004). "IQ and the Wealth of Nations, Lynn and Vanhanen: data table of national mean IQ studies." http://www.isteve.com/IQ_Table.htm
- Schaie, K. W. (2005). *Developmental influences on adult intelligence: The Seattle Longitudinal Study*. London: Oxford University Press.
- Seligman, D. (1992). *A question of intelligence: The IQ debate in America*. New York: Birch Lane Press.
- UNESCO (1998). *Statistical Yearbook 1998*. Paris: UNESCO Publishing and Bernan Press.
- Vinogradov, E., and Kolvereid, L. (2006). "Cultural Background, Home Country National Intelligence, and Self-Employment Rates Among Immigrants in Norway," Working Paper, Bodo Graduate School of Business.
- Voracek, M. (2004). "National intelligence and suicide rate: an ecological study of 85 countries." *Personality and Individual Differences*, 37: 543-53.

- Voracek, M. (2005). "National intelligence, suicide rate in the elderly, and a threshold intelligence for suicidality: an ecological study of 48 Eurasian countries." *Journal of Biosocial Science*, 37: 721-40.
- Weede, E. (2004). "Does Human Capital Strongly Affect Economic Growth Rates? Yes, but only If Assessed Properly." *Comparative Sociology*, 3, 2, 115-134.
- Weede, E. & Kampf, S. (2002). "The Impact of Intelligence and Institutional Improvements on Economic Growth," *Kyklos*, 55, 361-380.
- Zax, J. S. & Rees, D. I. (2002). "IQ, Academic Performance, Environment, and Earnings," *Review of Economics and Statistics*, 84, 4, 600-616.