Discuss this article at Jt: http://journaltalk.net/articles/5466

Econ Journal Watch,
Volume 1, Number 2,
August 2004, pp 222-234.

Kuznets Curveball: Missing the Regional Strike Zone

JEFF EDWARDS[*] AND ANYA MCGUIRK[**]

A COMMENT ON: JIH Y. CHANG AND RATI RAM. 2000. LEVEL OF DEVELOPMENT, RATE OF ECONOMIC GROWTH, AND INCOME INEQUALITY. *ECONOMIC DEVELOPMENT AND CULTURAL CHANGE* 48(4): 787-799.

ABSTRACT, KEYWORDS, JEL CODES

THE OBJECTIVE OF THIS PAPER IS TO EMPHASIZE THE NEED FOR adhering to the standard normal and identically distributed (NID) assumptions that are the basis of proper statistical inference in Ordinary Least Squares econometric models. The testing of and adherence to these assumptions are particularly important with small samples. In this paper we offer a statistical critique of a study by Chang and Ram (2000). The inferences they draw from their models are incorrect because of a violation of the assumption of identically distributed errors. Using graphical analysis and descriptive statistics, we were able to determine that significant regional fixed effects resulted in the invalidation of their conclusions that the Kuznets hypothesis holds, and that high growth countries maintain a lower level of income inequality over all income levels than do low growth countries. We add to this critique by commenting on the data used, as well as, the way Chang and Ram force the data through a somewhat arbitrary theoretical form.

[*] Department of Economics, Texas Tech University.
[**] Department of Statistics and Department of Agricultural and Applied Economics, Virginia Tech.

According to Simon Kuznets (1955) there should exist an inverted-U correlation between income inequality and the level of economic development (usually measured as per capita GDP). The inverted-U correlation suggests that for low levels of economic development there will exist low levels of inequality in per capita GDP, and inequality will increase as income increases. However, as incomes reach a particular level, i.e., as the economy develops past a certain point, the relationship reverses and income inequality decreases. Kuznets empirically documented this relationship for nations transitioning from agricultural to industrial economies.

The most common measure of income inequality is the Gini coefficient developed by Corrado Gini (1936). If the area between the line of perfect equality and the Lorenz curve is A, and the area beneath the Lorenz curve is B, then the Gini coefficient is A/(A+B).[1] The Gini coefficient ranges from 0% (perfectly even distribution) to 100% (perfectly concentrated distribution). In terms of income inequality, a relatively low Gini coefficient reflects low levels of income inequality between the rich and the poor, while a relatively high Gini coefficient reflects high levels of inequality.

THE MODELS, THE DATA, AND THE DUPLICATION OF CHANG AND RAM

In their paper *Level of Development, Rate of Economic Growth, and Income Inequality*, Chang and Ram evaluate the Kuznets hypothesis using a common functional form that includes an additive growth term. They estimate the following regression:

$$Ineq_i = a_1 + b_1(LY_i) + c_1 (LY_i)^2 + d(GY_i) + u_{1i} \qquad (1)$$

where $Ineq_i$ is the Gini coefficient for country i, LY_i is the natural log of income per capita, GY_i is the rate of growth of income per capita, and u_{1i} is a disturbance term that is assumed to be normal and identically distributed (NID). Multi-country cross-sectional studies prior to Chang and Ram lead to ambiguous evidence regarding the relationship between income inequality and the additive growth term in a pooled sample. To investigate this issue, Chang and Ram stratify their multi-country data by growth rate:

[1] See Appendix A.

high (2% and above), medium (between zero and 2%), and low growth countries (zero or negative growth). In their final regression they actually throw out the medium growth countries and estimate the following:

$$Ineq_i = a_2 + b_2(LY_i) + c_2(LY_i)^2 + a_{22}D + b_{22}(D^*LY_i) + c_{22}(D^*(LY_i)^2) + u_{2i} \quad (2)$$

where D takes the value of 1 for high growth countries and 0 for low growth countries.

The data that Chang and Ram use come from a variety of sources. The Gini coefficient measuring income inequality consists of the "high quality" observations from the Deininger and Squire (1996) data set ". . . for or around the year 1985" (790). The per capita GDP variable comes from the Penn World Tables (PWT), version 6.1 constructed by Heston, Summers, and Aten (2002). On the other hand, the growth variable comes from the United Nations Human Development Report (1993) and is averaged over the years 1980-90. The choice of this data set brings forth several interesting issues.

First, even though the Deininger and Squire data set is carefully constructed, with special care taken to obtain 'high quality' observations, the data on the Gini coefficient comes from many different sources, possibly leading to systematic measurement error in this series, e.g., a less democratic government source may want to report the coefficient as low as possible while a more liberal source may want to report the coefficient as high as possible. Second, it is unclear why growth in per capita GDP is averaged over a 10 year period when we are trying to measure the contemporaneous effects of per capita GDP on the Gini coefficient. One reason for the averaging may be to smooth the business cycle effects in the series, however, the same argument can then be used for the contemporaneous Gini data as well as per capita GDP, even though it is well-known that the time rate of any measured time series is 'noisier' than the levels variable measured over time. Third, it seems that using consistent data sources would be preferable to using many sources. For instance, the growth rate in per capita GDP could have easily been calculated from the PWT, but was not. It can be assumed that most data will have particular flaws with the way it was collected or calculated. Given these flaws, data should be taken from the same source if at all possible. Doing this will help control for the possibility of different biases resulting from data attained from different sources.

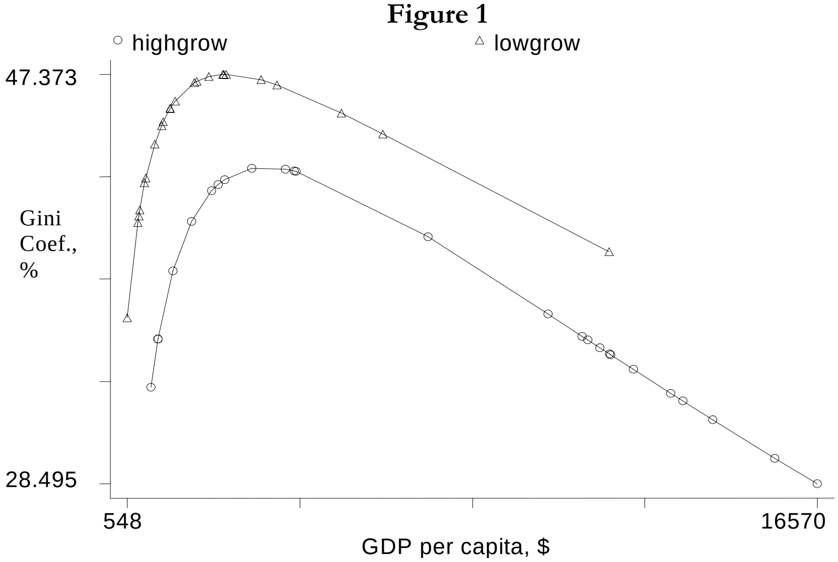
Despite the fact that there are these potential problems with the data used in Chang and Ram's paper, we now replicate their results for model (2) as shown below:

Table 1
Chang and Ram's Results of Model (2)

	Constant	LY	LY^2	D	D*LY	D^*LY^2
Estimate (t-statistics)	-212.774 (-1.62)	65.366* (1.92)	-4.106* (-1.87)	-193.792 (-0.89)	43.875 (0.82)	-2.529 (-0.77)
The number of observations is 48 The unadjusted R^2 is 0.40 The adjusted R^2 is 0.33 $^*P \leq 0.10$						

When graphed, the estimates of income inequality for the high and low growth countries take the following form:[2]

Figure 1



[2] Chang and Ram actually use predetermined values of real GDP per capita starting at 750 and increasing to 17,000, while we use the *observed* values of per capita GDP. The resulting graphs are almost identical.

The conclusion that is drawn by Chang and Ram is that high growth countries have lower levels of income inequality over all levels of income than do low growth countries. However, as we will demonstrate below, the model above is incorrectly specified because of differences in mean levels of income inequality across regions. Because relevant regional fixed effects have been omitted, Chang and Ram's results are biased and inconsistent.

THE STATISTICAL CRITIQUE, RESPECIFIED RESULTS, AND THE CONCLUSION

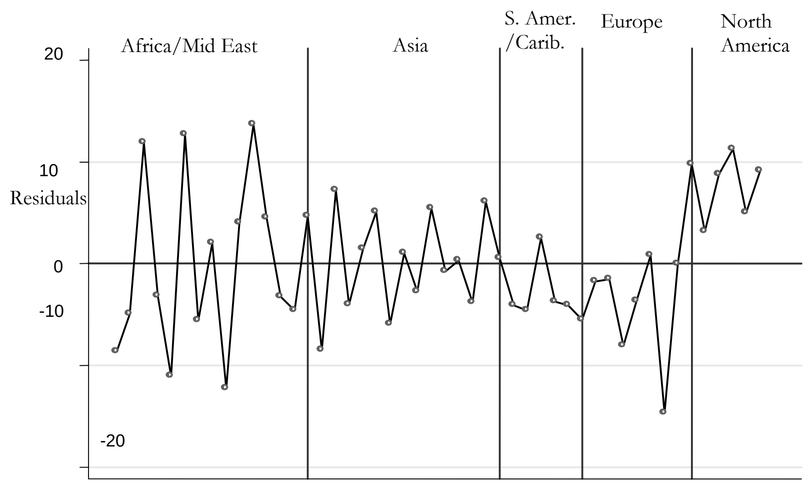
The correction of mean heterogeneity through the inclusion of fixed effects is critical to drawing accurate statistical inference from the estimated parameters. The errors of any Ordinary Least Squares regression, such as the ones used here, are assumed to be normal and identically distributed. Mean heterogeneity causes the assumption of identical distribution to break down resulting in biased and inconsistent estimators—hence, no inference can be drawn.

A close analysis of the Chang and Ram results led us to suspect the need for regional fixed effects. Figure 2 best illustrates this statistical evidence.[3] This figure, which shows a plot of the residuals from the Chang and Ram model ordered by geographical region, indicates a downward structural shift in the residuals for Europe, an upward structural shift for North America, while the other residuals are fairly evenly distributed around the mean of zero.[4] The inclusion of fixed regional effects, rather than some other sort of fixed effect, can also be justified from a theoretical point of view—one might expect that for political and socio-economic reasons, inequality in incomes may spill over from one nation to another within a particular geographical region.

[3] These regions are delineated according to the Penn World Tables version 6.1, which separates countries into 15 regions, however, because of the limited number of observations, testing all of them would greatly decrease our degrees of freedom and can be akin to data mining. Furthermore, as Figure 2 depicts, this detailed delineation is not needed.

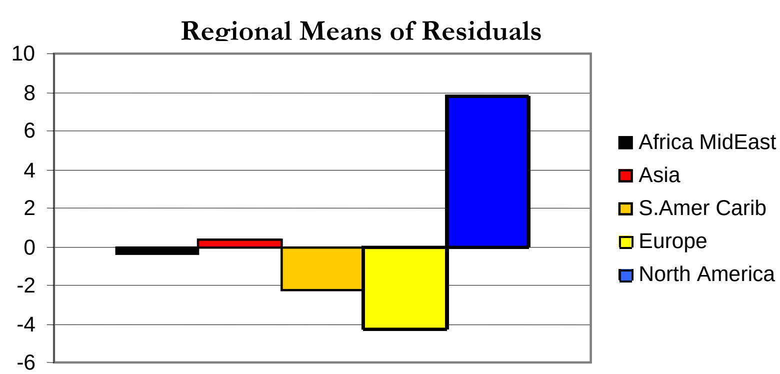
[4] It could be argued that there exists slight heterogeneity in the South America/Caribbean region, however, with South America as the control group, the p-value for the significance of the Caribbean was 0.909, and an f-test of the joint significance of Africa and Middle East, Asia, and the Caribbean returned a p-value of 0.998. These statistics indicate no statistically significant difference in the South America/Caribbean region from the regions of Africa and Middle East, or Asia.

Figure 2



In order to emphasize the differences in the means of the residuals across regions, Figure 3 charts these means as they deviate from zero. We see that estimation of model (2) results in residual means that vary from -0.3379 for the Africa and Middle East region to 7.8344 for North America.

Figure 3



When dummy variables are included for the regions of North America, Asia, Africa and the Middle East, Europe, and the Caribbean (South America is the control group), an F-test of the joint significance of the regional intercept shifts yields a p-value of 0.0188 (the F-statistic is 3.13).[5] To save degrees of freedom, we drop the insignificant regional dummy variables and keep Europe and North America. The resulting estimates are summarized below:

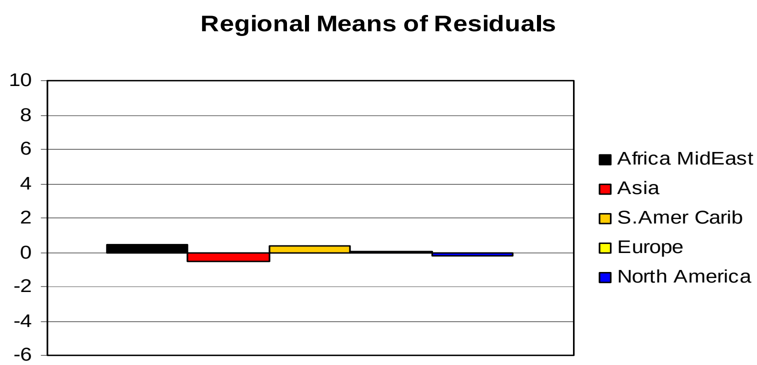
Table 2:
Respecified Results

	Constant	LY	LY^2	D
Estimate (t-statistic)	-143.350 (-1.242)	47.678 (1.595)	-3.018 (-1.571)	-349.76* (-1.775)
	D*LY	$D*LY^2$	Europe	N. Amer.
Estimate (t-statistic)	81.692* (1.672)	-4.757 (-1.581)	-6.147* (-1.804)	8.793* (2.837)
The number of observations is 48 The unadjusted R^2 is 0.58 The adjusted R^2 is 0.51 * $P \leq 0.10$				

As we can see above, LY and LY^2 are no longer statistically significant at the 10% level, whereas they were under the specification in Table 1 (the t-statistic cutoff is 1.684 for a 90% confidence interval with 40 degrees of freedom). Furthermore, our mean heterogeneity results suggest that the relationship between the Gini coefficient versus incomes (comparable to those of Chang and Ram) will vary by region—i.e., the intercepts will differ. In addition, the residuals of this new model show no un-modeled systematic differences by regions.

[5] We could have used a Chow-type test here to test for structural change, however, to get a reliable test statistic, we would need a much larger data set due to the design of the Chow test.

Figure Four



However, there is a slight problem with the estimates in Table 2.[6] The region of North America includes the countries of Canada, Mexico, the United States, Honduras, Guatemala, and Panama. This is problematic because the average gini coefficient of the United States and Canada is 35.03, while the average gini coefficient of Mexico, Honduras, Guatemala and Panama is 55.06. To correct for this, we separated the United States and Canada from the rest—redefining the North American dummy variable into two separate dummy variables. The results are that the dummy variable controlling for the United States and Canada had a t-statistic of -0.213, while the dummy variable controlling for Mexico, Honduras, Guatemala, and Panama was 3.394. We dropped the former dummy variable and kept the latter and then re-estimated the model.

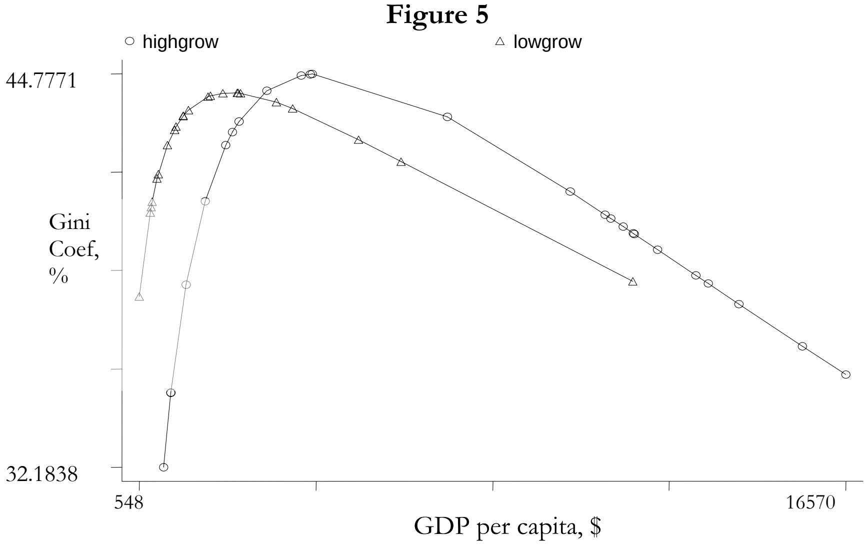
[6] This observation was graciously contributed by the authors of the original paper in correspondence prior to the submission of this paper to *Econ Journal Watch*.

Table 3:
Respecified Results

	Constant	LY	LY^2	D
Estimate t-statistic	-143.350 (-1.242)	42.097 (1.453)	-2.675 (-1.438)	-260.863 (-1.436)
	D*LY	$D*LY^2$	Europe	N. Amer.
Estimate t-statistic	59.027 (1.314)	-3.310 (-1.200)	-9.483* (-3.079)	11.568* (3.435)
The number of observations is 48 The unadjusted R^2 is 0.61 The adjusted R^2 is 0.54 * $P \leq 0.10$				

After re-estimation only the dummies for Europe and the four Central American countries are significant and none of the coefficients for GDP are significant. Therefore, the Kuznets hypothesis is invalid. However, if we assume the hypothesis does hold, and force the data through Chang and Ram's quadratic functional form from equation (2), we can graph the relationship as in Figure 1 assuming the countries are all in the control group (we subtracted out the estimated effects of the two dummy variables from the estimates of the gini coefficient).

Figure 5



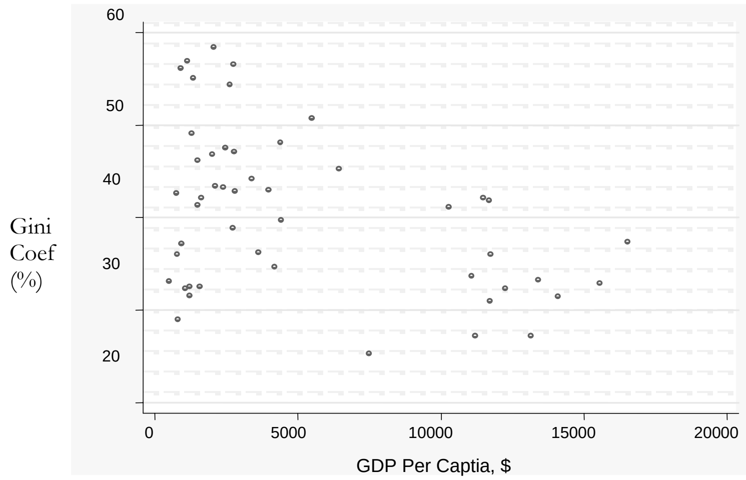
What we find in this new graph is that for low levels of GDP per capita, high growth countries have less income inequality than low growth countries, and for medium to high levels of GDP per capita, low growth countries have less income inequality than high growth countries. When important regional heterogeneity is modeled, the conclusion of Chang and Ram that higher growth countries have lower levels of inequality no longer holds for all income levels. The economic explanation of this crossing pattern would be beyond the scope of this paper, but is moot given that the GDP terms are both statistically insignificant and, given the multiple problems of the entire exercise, economically insignificant.

What has just been depicted is a complete structural breakdown of Chang and Ram's results. This breakdown results from the violation of two critical assumptions—the assumption that the Kuznets hypothesis holds in the first place, and the assumption that the errors corresponding to the regression of model (2) are normal and identically distributed. In order to address the former, we must first address the latter.

We have seen that the identically distributed assumption of the errors breaks down with regard to regionally-specific fixed effects. Including the dummy variables of Europe and (what we define as) Central America, results in the statistical insignificance of all per capita GDP variables in Table 3. In fact, only the two dummy variables are statistically significant. This is in contrast to Chang and Ram's results where at least the Kuznets hypothesis holds for the low growth countries, but not for the high growth countries. One might argue that the regional dummy variables are simply proxies for the high and low growth countries, however, this line of reasoning is flawed. The flaw lies in the simple fact that the complete set of high or low growth countries are not embodied in the regional dummies, and furthermore, the intercept dummy for the high growth countries in Table 1 was highly insignificant. This indicates that there never was a statistically significant structural difference between high and low growth countries.

The first assumption—that the Kuznets hypothesis holds at all—is much easier to address. Given the fact that the structural dummy for the high growth countries is statistically insignificant in both the Chang and Ram regression (Table 1) and in our final regression (Table 3), a simple plot of the Gini coefficient on per capita GDP would tell us if there is any inverted-U pattern to estimate.

Figure 6

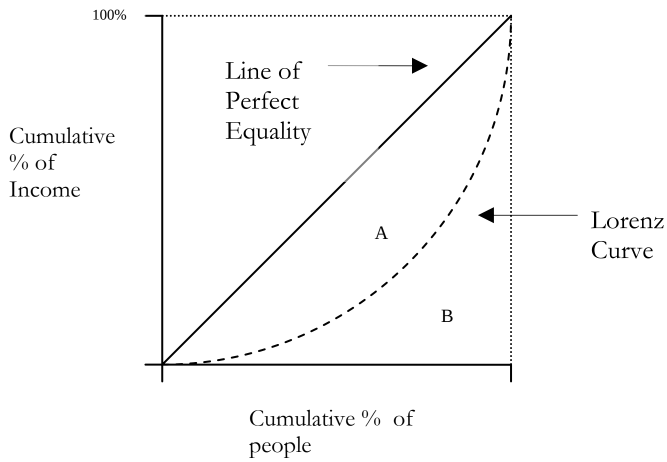


The appearance of any quadratic correlation between the Gini coefficient and per capita GDP is not apparent in Figure 6. In fact, if pressed, there would almost seem to be a somewhat linear, negative correlation between the two. We must remember that the reason why Figures 1 and 5 depict the typical inverted-U pattern is simply because the coefficients to the squared GDP variables happen to be negative (even though not statistically significant) and because *the data was forced through an equation that produces a quadratic-type curve*. It is also unclear why this type of quadratic structure was chosen among many others that produce the same shape. Randolph and Lott (1993) address seven possible forms, with supporting results, for the inverted-U pattern of the Kuznets hypothesis.

In conclusion, in Chang and Ram's paper, not only is the data set troublesome and the functional form ad hoc, but the forcing of the data through the chosen functional form when there are apparent mean heterogeneity problems is statistically improper.

APPENDIX A

Graph of the Gini coefficient and the Lorenz curve with regard to income inequality:



REFERECNES

Chang, Jih Y., and Rati Ram. 2000. Level of Development, Rate of Economic Growth, and Income Inequality. *Economic Development and Cultural Change* 48(4): 787-799.

Gini, Corrado. 1936. On the Measure of Concentration with Special Reference to Income and Wealth. Abstracts of Papers Presented at the Cowles Commission Research Conference on Economics and Statistics, Colorado College Publications.

Klaus Deininger and Lyn Squire. 1996. A New Data Set Measuring Income Inequality. *World Bank Economic Review* 10(3): 565-591.

Kuznets, Simon. 1955. Economic Growth and Income Inequality. *American Economic Review* 45: 1-28. need issue number.

Randolph, Susan, and William F. Lott. 1993. Can the Kuznets Effect Be Relied on to Induce Equalizing Growth? *World Development* 21(5): 829-840.

Heston, Alan, Robert Summers, and Bettina Aten. 2002. The Penn World Table Version 6.1. Center for International Comparisons at the University of Pennsylvania. October.

United Nations Development Program. 1993. *Human Development Report*, 188-89.

ABOUT THE AUTHORS

Jeffrey Edwards is an assistant professor of economics at Texas Tech University. He holds a Ph.D. in economics from Virginia Tech with a major concentration in econometrics. His research interests are empirically broad from testing probabilistic hypotheses of empirical models, to growth and development, to water resource management. He is also an assistant editor for the Journal of Economics, and currently sits on the advisory panel of the first international water resources conference sponsored by Texas Tech. His email address is jeffrey.edwards@ttu.edu.

Anya McGuirk is a full professor at Virginia Tech. She holds a joint appointment with the Departments of Statistics, and Agricultural and Applied Economics. Her primary areas of research are applied econometrics and microeconomics. Her most recent works include studies of the restrictions imposed when modeling error autocorrelation and the nature of the implications of multicollinearity in regression analysis. McGuirk also teaches econometrics and microeconomic theory at the graduate and undergraduate levels. Anya obtained her MSc. as a Marshall Scholar at the University of Reading, England. She received her Ph.D. from Cornell University and her dissertation, "Technological Change and Agricultural Growth in Punjab, India," received the AAEA Outstanding Dissertation Award (1989). While McGuirk enjoys her research and teaching, her real loves are fly-fishing and family.

GO TO RESPONSE BY JIH CHANG AND RATI RAM

Discuss this article at Jt: http://journaltalk.net/articles/5466