
REPLIES & RESPONSE

TO SCOTT BAKER,
SHOULD WE PAY FEDERAL CIRCUIT JUDGES MORE?,
88 B.U. L. REV. 63 (2008).

PERHAPS WE SHOULD PAY FEDERAL CIRCUIT JUDGES MORE

FRANK B. CROSS*

Scott Baker's article takes a creative approach to the question of whether the federal judiciary should receive a pay raise and concludes that the data counsel otherwise.¹ While I am a great advocate of empirical analysis to inform questions such as this, and Professor Baker may have done the best job possible with the available measures, I fear that this study contributes little to the debate and cannot support its author's conclusions. The available tools to measure the effects of lower pay and judicial performance are so extremely crude they cannot tell us much. Perhaps more significant, Baker's failure to prove that judicial pay *does* matter, given the limitations of the available measures, provides no evidence that it *does not* matter. I think it makes far more sense, given the lack of reliable measures, to rely on basic economic intuition and more direct anecdotal experience.

I. WAGE ECONOMICS AND JUDGES

In ordinary economic circumstances, better pay obviously results in higher quality workers. Professor Baker notes some of this research in his article.² Basic economic intuition tells us that an employer offering less money for a job will get fewer and less qualified applicants than an employer offering more. The best candidates won't even want the job, which will be filled by someone of lesser qualifications.

* Herbert D. Kelleher Centennial Professor of Business Law, McCombs School of Business, University of Texas at Austin; Professor of Law, University of Texas Law School; Professor of Government, University of Texas at Austin. The author would like to thank his colleagues, Dain Donelson and Robert Prentice, for their helpful insights into this response.

¹ Scott Baker, *Should We Pay Federal Circuit Judges More?*, 88 B.U. L. REV. 63, 66 (2008).

² *Id.* at 73 n.42.

This line of thinking is behind current efforts to increase judicial pay. Chief Justice Roberts has said that the failure to raise judicial pay has created a “constitutional crisis.”³ Justice Scalia declared that because of insufficient judicial pay, “we cannot attract the really bright lawyers” because it is “too much of a sacrifice”⁴ for them to give up their private practice salaries. Abner Mikva, writing about his White House experience, has discussed the difficulty in getting desired candidates for the bench due to low pay.⁵ Federal judicial salaries, adjusted for inflation, have declined significantly in recent years. Since 1969, district court judge salaries have declined 21.5%, while law partner profits have grown 74.1% and top law school senior professor salaries have increased 114.6%.⁶ Congress is currently considering raising the pay of federal judges in response to these concerns.⁷

There are some legitimate reasons to question whether this standard wage economics analysis properly applies to the federal judiciary. The first difference is the non-monetary compensation offered by such an appointment. Federal judges can gain considerable non-monetary utility from, among other factors, their ability to project (political or legal) authority, opportunities for more leisure, increased social status, and guaranteed life tenure.⁸ Of course, individuals value these forms of utility differently. Lower pay will result in more federal judicial candidates who place relatively more importance on projecting power. This consideration means that our decisions on pay must consider what type of judges we want on the federal bench. The answer to this question is indeterminate, however.⁹

³ Chief Justice John G. Roberts, Jr., *2006 Year-End Report on the Federal Judiciary*, THE THIRD BRANCH: NEWSLETTER OF THE FEDERAL COURTS (Admin. Office of the U.S. Courts, Wash., D.C.), Jan. 2007, at 1, available at <http://www.uscourts.gov/ttb/jan06ttb/yearend/index.html>.

⁴ Posting of Peter Lattman to Wall Street Journal Law Blog, <http://blogs.wsj.com/law/2006/12/14/justice-scalia-bemoans-judicial-pay/> (Dec. 14, 2006, 10:53 A.M.) (quoting a Dec. 13, 2006 speech by Justice Scalia).

⁵ Abner Mikva, *Attracting the Best*, NAT'L L.J., June 4, 2007, at 23.

⁶ CONG. RESEARCH SERV., JUDICIAL SALARY: CURRENT ISSUES AND OPTIONS FOR CONGRESS, at CRS-25 (2007), available at http://assets.opencrs.com/rpts/RL34281_20071212.pdf.

⁷ *Id.* at CRS-36 to -37.

⁸ See, e.g., Russell Smyth, *Do Judges Behave as Homo Economicus, and If So, Can We Measure Their Performance? An Antipodean Perspective on a Tournament of Judges*, 32 FLA. ST. U. L. REV. 1299, 1302-09 (2005) (reviewing evidence that judges are concerned with incentives other than financial ones).

⁹ We might very well want judges who put less importance on projecting political power, because they would be less inclined to reach political or ideological outcomes (which Baker attempts to measure, see Baker, *supra* note 1 at 85-97). However, we might prefer judges who put greater importance on projecting legal power, which presumably would induce them to produce legally better decisions and opinions.

The second difference involves the screening involved in the selection process. Higher pay will produce a larger and better pool of candidates.¹⁰ In private industry, applicants are screened for the relevant qualifications. People will be hired and paid more according to their economic value, as best measured by economic productivity. If higher pay is to increase judicial quality, candidates must be effectively screened for quality, which is clearly an uncertain matter.

Federal judges are appointed by the President and confirmed by the Senate. Clearly, both use criteria other than pure judicial quality in their decisions. Indeed, it is well established that ideological considerations play a role in the selection of federal judges. If judicial quality is not used in screening selections, one can have little confidence that a higher quality pool of candidates will yield a higher quality judiciary. This question has been studied at the Supreme Court level, though, and the empirical research suggests that a judicial candidate's professional merit *is* a significant determinant in the confirmation process.¹¹ Thus, the quality of the pool of candidates may influence the quality of the judiciary. While it is too facile to assume that higher pay means higher quality judges, it is likewise too facile to presume the contrary. Hence the value of empirical analysis.

II. BAKER'S EMPIRICAL ANALYSIS

Professor Baker certainly deserves credit for thought and creativity in designing an empirical test of judicial pay levels. However, his analysis, though it may be the best job realistically possible, does not add much to the judicial pay debate. His measures, both for the effects of judicial pay and for judicial quality, are extremely crude and reveal little or nothing. The analysis of those measures suffers from problems with multicollinearity and omitted third variables. Finally, he places too much significance on a failure to reject the null hypothesis. This is especially true because the flaws in his analysis conspire to produce results that would fail to reject a null hypothesis, even when the null hypothesis was plainly false.

A. *Measures*

To conduct an empirical study on the question of the effect of judicial pay, a researcher needs measures to capture the relevant variables. Because federal judges at the same hierarchical level are paid equally, it is difficult to distinguish the effect of compensation.¹² Professor Baker employs a measure of foregone wages to study the effect of judicial pay, which could capture the

¹⁰ In addition, many workers in private industry are motivated to do a better job by the promise of incentive compensation, but this is inapplicable to the federal judiciary.

¹¹ See Lee Epstein et al., *The Changing Dynamics of Senate Voting on Supreme Court Nominees*, 68 J. POL. 296, 296 (2006).

¹² See Baker, *supra* note 1, at 76.

effect of low salaries on individuals' willingness to become federal judges.¹³ Foregone salary is based on a comparison of judicial pay with the average salary of a law firm partner in the region of the judge's service, adjusted for the age of joining the judiciary and the number of years of private practice compensation lost, called "NETCOST."¹⁴ Unfortunately, this is a somewhat flawed measure to assess the effect of judicial pay.

The underlying notion of Professor Baker's analysis is that increases in judicial pay will reduce the NETCOST of joining the judiciary.¹⁵ Thus, a finding that his NETCOST measure has a significant association with undesirable consequences would be evidence of a judicial pay effect. Thus, if judges with lower NETCOST are less ideological, that could counsel for raising pay to reduce ideological decision making. The individual's tradeoff between income and his or her "taste for being a judge" may be the best available quantitative tool for measuring the consequences of judicial pay, but it is extremely rough.

The measure itself is of uncertain accuracy because it assumes the appropriate salary comparison for judges is that of the region of their judicial service.¹⁶ There is no reason to assume such immobility. While most circuit court judges have come from the same region in which they serve, a good number have relocated to become judges.¹⁷ More might have moved, absent the judicial appointment.¹⁸ Even for those who would not move to become a judge, the measure is extremely rough. Professor Baker used Fourth Circuit Judge Sprouse as an example.¹⁹ Judge Sprouse spent nearly all his life in West

¹³ *Id.* at 78.

¹⁴ *Id.*

¹⁵ *See id.* at 89.

¹⁶ *Id.* at 91.

¹⁷ Many D.C. Circuit judges come from outside of Washington, D.C. and Professor Baker has appropriately adjusted for this by using their preexisting locale. Baker, *supra* note 1, at 83-84 n.75. However, the effect is not entirely limited to the D.C. Circuit. Michael Boudin of the First Circuit had been in private practice in Washington, D.C. prior to taking the bench in Boston, although he was first nominated for the D.C. District Court. J. Michael Luttig of the Fourth Circuit had spent a decade in government in Washington D.C. This list is incomplete but shows that mobility is not uncommon. Federal judicial biographies can be found at Judges of the United States Courts, <http://www.fjc.gov/public/home.nsf/hisj> (search by judge name; then click on judge name when results appear). Most circuit court judges are drawn from within the circuit, so the measure has some accuracy, but it is rough.

¹⁸ For example, when Judge Luttig left the bench, he moved from Richmond, Virginia to Chicago. See Jerry Markon, *Appeals Court Judge Leaves Life Appointment for Boeing*, WASH. POST, May 11, 2006, at A11. This move likely would have occurred earlier had he not served as a federal judge.

¹⁹ Baker, *supra* note 1, at 79-81.

Virginia,²⁰ but his average salary calculation in Baker's model is for the entire South Atlantic region, within which there are significant salary disparities.²¹ Likewise, First Circuit judges may come from Maine or Boston, two legal markets with significant salary differential. Professor Baker attempts to adjust for this effect, but only with the rough tool of "top-five legal market," not considering any effect outside these markets.²²

An additional problem with the NETCOST measure is its use of average partner salaries in the relevant markets. There is no particular reason to assume circuit court judges would receive an "average" salary. Some might receive more, others less. This is a problem, because it is theoretically possible that the true NETCOST for every judge in the sample is identical. Suppose those judges with high NETCOST scores based on the average would actually have received lower than average partner salaries, while those with low NETCOST scores based on the average would in fact have received higher than average salaries.²³ In this case, the measure is obviously distorted and there might in fact be no difference in the actual income foregone regardless of the measure.

Baker's measures for judicial quality are even cruder. His first hypothesis measures ideological voting, on the theory that higher quality judges are less ideological.²⁴ The use of ideological voting is one plausible proxy for judicial quality but a very weak one. The direct ideological outcome variable is binary – a decision is classified as either liberal or conservative, with no distinction regarding *how* liberal or conservative the decision is.²⁵ The NETCOST variable could have moderated the extent of conservatism or liberalism, but Baker's model would not have identified such an effect. Moreover, a liberal or conservative opinion may be of higher or lower quality. An ideological result may be reached through sound legal reasoning based on ample authority or it may be reached arbitrarily. The voting measure makes no attempt to capture that.

Professor Baker's second ideological variable, citation bias, attempts to capture something of opinion quality.²⁶ This proxy is also extraordinarily rough, however, as it treats all citations equally.²⁷ The R-squared for this

²⁰ See Judges of the United States Courts, *supra* note 18.

²¹ Baker, *supra* note 1, at 80.

²² *Id.* at 91.

²³ At a general level, this effect is not implausible. It is harder to be an average quality lawyer in very high-paying areas which require more aptitude than it is to be an average quality lawyer in lower paying areas.

²⁴ Baker, *supra* note 1, at 85.

²⁵ *Id.* at 86.

²⁶ *Id.* at 95-97.

²⁷ *Id.* at 95-96. Some of the shortcomings of Baker's citation analysis are summarized briefly in Frank B. Cross et al., *Warren Court Precedents in the Rehnquist Court*, 24 CONST.

equation is quite low, indicating that none of Professor Baker's independent variables really addresses this end. While NETCOST once again fails to load as significant, the judicial ideology independent variable of "selfpref" also fails to show significance on the ideological dependent variable, which should call the results into some question.²⁸ The analysis implies that ideology does not affect ideology.

Professor Baker's second hypothesis is that more pay would produce a harder working judiciary.²⁹ In this case he *confirms* the hypothesis about the effect of NETCOST.³⁰ He correctly notes, though, that dissent is explained by factors other than simple effort, so the significance of the finding is uncertain.³¹ The second measure for effort, time to render a published opinion, is an extremely poor proxy. Taking more time to issue an opinion might be considered a sign of more effort, not less. One can easily imagine a lazy judge telling clerks to promptly produce an opinion, as well as a hard working judge who may have clerks produce the first draft but then carefully scrutinizes their work and takes longer to produce a final opinion. Justice William O. Douglas was renowned for quickly producing opinions, but this was not a sign of his hard work – some were scribbled out on the tray-table during a plane flight.³²

Baker's final hypothesis, about influence maximization and the effect of NETCOST on outside citations,³³ is probably the most interesting of the three. Again, the citation indicator is very rough – a judge's outside circuit citations will depend to some degree, perhaps substantially, on the cases that the judge is randomly assigned. And yet, the associations in Baker's analysis were almost statistically significant (the figure for average outside citations was only about six percent likely random chance).³⁴ This is close to a confirmation of the hypothesis, even by statistical significance testing. The coefficient was small, but Professor Baker concedes that his measure was "an imprecise and messy measure of judicial influence."³⁵

The proxy variables in Professor Baker's study are extremely poor tools for capturing what we hope to measure to evaluate the effect of judicial pay. No measures used in statistical analyses are precise – certainly not in the case of

COMM. 3, 5-6 (2007). While these limitations do not invalidate the measure, they counsel for caution in placing heavy reliance on results of the measure.

²⁸ *Id.* at 97.

²⁹ *Id.* at 98.

³⁰ *Id.* at 101 (finding that NETCOST has a significant impact on the rate of dissent in controversial cases).

³¹ *Id.*

³² JEFFREY ROSEN, *THE SUPREME COURT: THE PERSONALITIES AND RIVALRIES THAT DEFINED AMERICA* 129-30 (2007).

³³ Baker, *supra* note 1, at 105.

³⁴ *See id.* at 109.

³⁵ *Id.* at 105 n.128.

studying judicial opinions – and it is important to understand the significance of their limitations. The effect of very crude proxies for both independent and dependent variables is to introduce a great deal of random noise into the statistical equation. Insofar as this noise is indeed random, it will be uncorrelated with the dependent variable and therefore serve to obscure even a true significant association.

Quantitative studies typically must deploy a numerical proxy for some matter of concern. If the proxy is very closely correlated to the theoretical matter being studied, the association from the regression should be informative. If the proxy has no correlation whatsoever with the matter of concern, however, the association should be random. The degree to which the proxy fails to correlate with the matter of concern may be greater or lesser, and this failure creates something known as statistical noise. The general effect of such noise is to confound a finding of a true association. Thus, a low signal-to-noise ratio will run “the risk of failing to confirm a valid research hypothesis (i.e., Type II error).”³⁶ A design with “too much noise” is more likely to produce such a Type II false negative error.³⁷ Thus, the limitations of Professor Baker’s measures were likely to yield no statistically significant results, even in the presence of a true association. The lack of identified statistical significance, therefore, is not very meaningful. Professor Baker suggests that his analysis is “the best that can be done with the available data.”³⁸ This may well be true, but it does not mean much if the “best” is still a very poor measure. The variables simply don’t capture the true concerns about judicial performance.

B. *Statistical Analysis*

Professor Baker’s statistical analysis is appropriate to his variables and objectives, but he observes that the multicollinearity between his NETCOST measure and the age and circuit dummies “increases the standard errors, which might then generate the insignificant results.”³⁹ Multicollinearity, like the noise resulting from mismeasured variables, conspires against finding statistical significance, even when two variables are in fact associated. Here,

³⁶ See Amos Tversky & Daniel Kahneman, *Belief in the Law of Small Numbers*, in A HANDBOOK FOR DATA ANALYSIS IN THE BEHAVIORAL SCIENCES: METHODOLOGICAL ISSUES 341, 343 (Gideon Keren & Charles Lewis eds., 1993).

³⁷ See John Connolly, *Estimation of Sample Size Required for Experiment* (Apr. 3, 2008) (unpublished manuscript, available at <http://www.ucd.ie/statdept/jconnolly/pgm/sample%20size/samplesize.doc>). The noise would only produce a false positive Type I error if the noise were somehow systematically associated with the independent variables. There is no plausible explanation for this in Baker’s study.

³⁸ Baker, *supra* note 1, at 112.

³⁹ *Id.* at 110.

the sixty-percent overlap is not disqualifying, but this multicollinearity problem may be more significant than Professor Baker acknowledged.

A study of structural equation models found that when multicollinearity was between 0.6 and 0.8, there was greater than fifty-percent risk of Type II errors (false negatives).⁴⁰ While Professor Baker did not use a structural equation model in his study, such models are typically considered *more robust* to multicollinearity problems, so this study does not overstate the risk of error in Baker's analysis. Another analysis of traditional multiple regression models showed that multicollinearity could cause an extremely high Type II error rate, at least when the model's R-squared was relatively low (i.e., ≤ 0.25).⁴¹ Thus, the multicollinearity of variables in Baker's analysis created a high risk of a false negative.⁴²

Any valid statistical analysis also requires consideration of "third variables" that may help explain the results. The most obvious third variable that Professor Baker's analysis omits in the ideological voting hypothesis involves panel effects. His study uses data from the Chicago Judge's Project.⁴³ Other research with this data has demonstrated that the ideology of other panel members has a substantial effect on the votes of individual judges.⁴⁴ According to the Sunstein study, although the panel effect does not totally eliminate individual ideological voting, it substantially affects it.⁴⁵ Using a larger database of circuit court opinions, I found that the ideology of other panel judges was clearly the most significant determinant of outcomes, and the ideology of the voting judge was of uncertain significance.⁴⁶ The panel effect means that if Professor Baker was looking for evidence of ideological effect, he was looking in the wrong place. Moreover, the panel composition variable is a demonstrably relevant third variable and its omission, added to the noise, could obscure an authentic relationship between pay and ideological voting. Indeed, the measure of judicial ideology (selfpref) had no statistically

⁴⁰ Rajdeep Grewal et al., *Multicollinearity and Measurement Error in Structural Equation Models: Implications for Theory Testing*, 23 *MARKETING SCI.* 519, 526-27 (2004).

⁴¹ Charlotte H. Mason & William D. Perreault, Jr., *Collinearity, Power, and Interpretation of Multiple Regression Analysis*, 28 *J. MARKETING RES.* 268, 277 (1991).

⁴² Professor Baker suggests the multicollinearity should not bias the size of the coefficients, which are small. Baker, *supra* note 1, at 110. This may or may not be true and is difficult to establish. Even if true, a small coefficient could still justify some increase in judicial pay, as I explain in my conclusion. See *infra* notes 57-68 and accompanying text.

⁴³ Baker, *supra* note 1, at 86.

⁴⁴ See CASS R. SUNSTEIN ET AL., *ARE JUDGES POLITICAL?* 22 (2006) (concluding that "votes of judges are significantly influenced by the party affiliation of the president who appointed the other two judges on the same panel").

⁴⁵ See *id.*

⁴⁶ FRANK B. CROSS, *DECISION MAKING IN THE U.S. COURTS OF APPEALS* 165 (2007). Similar results were found in Richard L. Revesz, *Environmental Regulation, Ideology, and the D.C. Circuit*, 83 *VA. L. REV.* 1717, 1719 (1997).

significant effect on ideological outcomes in Tables 4 through 6 of Baker's study.⁴⁷ While this might be taken as evidence that judicial ideology has no effect on ideological voting, it seems more likely that a true association was obscured by statistical noise.

Another potential third variable is the intervening effect of clerks. Circuit court judges have an extremely heavy caseload and must delegate some of their work to clerks. We presumably want judges to delegate as little as is reasonably possible. Baker's study cannot measure the relative delegation among judges, however. Between the problems of multicollinearity and unmeasured third variables in Professor Baker's analysis, it is hard to place much weight on the failure of the study to find profound statistical significance.

C. *Results and Implications*

The most serious flaw in Professor Baker's article lies in the conclusion regarding its implications. A statistical analysis tests a null hypothesis. If the variables in a study reveal a statistically significant difference, the null hypothesis is rejected because there is a less than one-in-twenty probability that random chance would produce such an outcome. Thus, positive results of such a study provide substantial evidence that the null hypothesis is incorrect. Negative results, however, say little or nothing about the truth of the null hypothesis.

Professor Baker tests the null hypothesis that "judicial pay does *not* affect" his measures of quality.⁴⁸ Because three of the four regressions do not produce statistically significant results, he concludes that "the effect of low judicial pay is non-existent" and "does not impact voting patterns, citation practices, the speed of controversial case disposition, or opinion quality."⁴⁹ This is a misapplication of statistical methods; the lack of statistically significant results does not warrant such strong conclusions. It only demonstrates that we cannot confidently reject the null hypothesis. Thus, the failure to reject the null hypothesis "offers no evidence for choosing between two possibilities: there is no signal to detect or noise overwhelms the signal."⁵⁰

Suppose instead that Professor Baker were testing the null hypothesis as "judicial pay *does* influence" his measures of quality. The results of this alternative study would also have meant that we could not reject this null hypothesis. Thus, he might have said that low judicial pay "does impact voting

⁴⁷ Baker, *supra* note 1, at 92-94.

⁴⁸ *Id.* at 67 (emphasis added).

⁴⁹ *Id.* at 112.

⁵⁰ See Kevin D. Hoover & Mark V. Siegler, *Sound and Fury: McCloskey and Significance Testing in Economics*, J. ECON. METHODOLOGY (forthcoming) (manuscript at 19), available at <http://www.econ.duke.edu/~kdh9/Source%20Materials/McCloskey/Sound%20and%20Fury%206%20March%202007.pdf>.

patterns, citation practices, the speed of controversial case disposition or opinion quality.” That too would be a misapplication of statistics. In fact, the lack of statistically significant associations simply means that the test left the proper conclusion uncertain, which is unsurprising given the extreme crudeness of the measures for both the independent and dependent variables. Professor Baker can fairly say only that his study does not prove the need for judicial pay increases, but he cannot claim that it disproves the need.

III. BEHAVIORAL CONSIDERATIONS

The analysis thus far has relied on traditional economics, but recent years have seen increased recognition that individuals are not purely *homo economicus*. We make decisions not exclusively on expected return but also on considerations such as fairness. This field of study is known as behavioral economics. Professor Baker openly acknowledges a failure to consider this sort of effect and its potential significance.⁵¹

A behavioral economist would suggest that the NETCOST measure could not possibly capture all the effects of a pay increase, because it takes no account of behavioral responses. Under the behavioralist theory, a judge who believed that he or she was not being fairly compensated for his or her work would put correspondingly less effort into that work. Because this effect could operate across the board, its appearance would not be limited to those judges who sacrificed the greatest alternative earning opportunities.

Research indicates that higher pay does in fact motivate greater work effort: “gift exchange” models suggest that worker productivity is enhanced by generosity in wage benefits.⁵² Many employers make generous wage offers to employees in anticipation of greater effort, which the employers often receive.⁵³ When workers perceive that they are treated unfairly, they often react by reducing their work effort.⁵⁴

This is not to suggest that judges would consciously and blatantly shirk their responsibility in a sort of “blue flu” strike. However, in the presence of pay unfairness, behavioral economists might expect judges to work a little less, perhaps only subconsciously. This might take the form of judges writing less or perhaps exerting less effort to get the law “right” in particular opinions. It could also mean delegating more workload to clerks. This sort of tradeoff between pay and effort could be independent of NETCOST and hence would not be captured by Professor Baker’s empirical study.

⁵¹ Baker, *supra* note 1, at 112.

⁵² See George A. Akerlof, *Labor Contracts as Partial Gift Exchange*, 97 Q.J. ECON. 543, 567 (1982); Gillian Lester, *Careers and Contingency*, 51 STAN. L. REV. 73, 135 (1998).

⁵³ Ernst Fehr & Simon Gächter, *Fairness and Retaliation: The Economics of Reciprocity*, 14 J. ECON. PERSP. 159, 169 (2000).

⁵⁴ George A. Akerlof & Janet L. Yellen, *The Fair Wage-Effort Hypothesis and Unemployment*, 105 Q.J. ECON. 255, 255 (1990).

The “fairness” of prevailing judicial pay levels is debatable; however, the justices of the Supreme Court have made it quite clear they consider current levels to be far too low.⁵⁵ One might think the “first-year associate” pay comparison might be a factor in fairness. The salaries of newly minted associates are approaching the salaries of federal judges.⁵⁶ One can readily imagine a federal judiciary, consisting of some of this country’s best legal minds, finding it fundamentally unfair if their pay drops below the level of those with no actual legal experience.

CONCLUSION

While I admire the great effort undertaken by Professor Baker to empirically assess judicial salaries, I do not think he has made much of a case against pay raises. The tools available for his empirical analysis are extraordinarily crude and unlikely to reveal even a true association.⁵⁷ Moreover, the proposition that increased pay will not increase judicial quality is contrary to the simplest of economic intuitions.

Obviously, there *must* be some tradeoff of pay and quality in the federal judiciary. No one would recommend paying federal judges zero salary, which is what the absence of any tradeoff would suggest is efficient. If a salary of \$170,000 yields better judges than a salary of zero, it seems logical to conclude that a salary of \$200,000 would yield still better judges. The difference in quality, of course, would not be nearly so great, but there should be some marginal effect. The effect of small pay changes may be so marginal that they would not appear to be statistically significant even in an empirical study with better variable measurements than Professor Baker’s. Nevertheless, they still exist.

In defense of Professor Baker’s research and choice of null hypothesis, one might argue there should be a presumption against increasing judicial pay which must be overcome by data disproving that presumption. Given the

⁵⁵ See, e.g., Roberts, *supra* note 3.

⁵⁶ Baker, *supra* note 1, at 82.

⁵⁷ It should be noted that Professor Baker is not the only person to have attempted to study the effect of judicial pay levels. Others have analyzed state judicial pay, which has the advantage of providing differential cross-sectional pay levels, but the disadvantage of being only a rough parallel to the federal judiciary. One study used an approach similar to Professor Baker’s to examine state judicial salaries and found little positive effect on their measures of judicial output (which included total opinions, outside citations, and dissents). See Stephen J. Choi et al., *Are Judges Overpaid?: A Skeptical Response to the Judicial Salary Debate* 59 (Univ. of Chicago Law School John M. Olin Law & Economics Research Paper Series, Working Paper No. 376, 2008), available at <http://ssrn.com/abstract=1077295>. Another study found that higher levels of state judicial pay were significantly associated with lower levels of general governmental corruption. See Adriana S. Cordis, *Judicial Checks on Corruption in the United States* 4 (February 2008) (unpublished manuscript, available at <http://ssrn.com/abstract=1019897>).

extreme difficulty of finding reliable measures, however, this unfairly stacks the deck, asking advocates of greater pay to prove the impossible. Increased judicial pay does add budgetary costs, so one might nevertheless put some burden on those seeking pay increases. But any such burden should not be a heavy one.

The law is extremely important to the economy and societal welfare. Douglass North, Nobel Prize-winning economist and leader of the new institutional economics, has stressed that the law is the key to economic growth.⁵⁸ Empirical research has shown that the maintenance of a reliable rule of law is responsible for significantly greater national economic growth.⁵⁹ It is not so much the content of substantive laws as it is the reliable implementation of those laws that enables growth.

Judges are central to maintaining the quality of the national legal system. “[T]he judiciary is a vital factor in the rule of law and more broadly in economic development.”⁶⁰ Studies in less-developed foreign nations have demonstrated that low judicial salaries result in a low-quality judiciary.⁶¹ While these findings may not be directly applicable to the United States, they illustrate a fundamental economic point: “[y]ou pay peanuts, you get monkeys.”⁶² The fact that relatively low judicial salaries in the United States have not seriously undermined the overall rule of law does not mean they have not had some marginal negative effect.

Choi, Gulati, and Posner suggest that the appropriate test is whether an “incremental increase in pay will improve the social value of judicial performance more than the social cost of the higher pay.”⁶³ This cannot be directly measured, but given the value of a quality legal system (perhaps more than a trillion dollars), it would take only a very slight improvement in quality to justify an increase in judicial pay.

Earlier empirical research also informs the debate over judicial salaries. A study of circuit court resignations between 1893 and 1991 found that lower

⁵⁸ See DOUGLASS C. NORTH, INSTITUTIONS, INSTITUTIONAL CHANGE AND ECONOMIC PERFORMANCE 54 (1990) (“[T]he inability of societies to develop effective, low-cost enforcement of contracts is the most important source of both historical stagnation and contemporary underdevelopment in the Third World.”).

⁵⁹ ROBERT J. BARRO, DETERMINANTS OF ECONOMIC GROWTH: A CROSS-COUNTRY EMPIRICAL STUDY 26-28 (1997).

⁶⁰ KENNETH W. DAM, THE LAW-GROWTH NEXUS: THE RULE OF LAW AND ECONOMIC DEVELOPMENT 93 (2006). Dam notes that “[b]etter-performing courts have been shown to lead to more developed credit markets” and that a “stronger judiciary is associated with more rapid growth of small firms as well as with larger firms in the economy.” *Id.* (citation omitted).

⁶¹ *Id.* at 116 (discussing Ukraine and Mexico as examples of countries where low salaries have diverted the best law students into private practice rather than the judiciary).

⁶² *Id.* (quoting Singaporean Prime Minister Lee Kwan Yew).

⁶³ Choi, *supra* note 57, at 13.

judicial salary was associated with increased retirement.⁶⁴ A study of district court resignations found similar results.⁶⁵ While resignation rates do not test the effect of salaries on judicial selection or the pool of available candidates, the information suggests there is an association between pay and who is willing to serve on the bench.

Recent years have seen a number of very prominent federal judicial resignations, with the retiring judge often citing low pay as the reason for leaving the bench. In 2006, Fourth Circuit Judge Luttig resigned to become an officer of Boeing, citing family financial concerns.⁶⁶ In early 2007, Judges Hector Laffitte and David Levi left the district court bench, both expressing concerns about salary.⁶⁷ Later that year, district Judge Paul Cassell tendered his resignation, expressing similar concerns.⁶⁸ These anecdotal examples certainly offer evidence that higher pay is important to the retention of good judges.

Professor Baker can correctly say that his model produced no convincing evidence in support of a judicial pay raise. It is not accurate for him to say, however, that it produced convincing evidence to reject a judicial pay raise. Failure to reject a null hypothesis never provides proof of the null hypothesis. And a failure in the presence of as much noise, collinearity, and omitted variable bias as existed in Professor Baker's study, means very little. Given the anecdotal evidence, economic logic, and the great importance of a quality judiciary, the case for a judicial pay raise is reasonably strong.

In closing, I feel compelled to observe that Professor Baker's article also implicates the salaries of law professors. The choice to teach at a law school and take non-monetary compensation, rather than enter private practice, is not unlike the choice to become a judge. Perhaps someone should undertake a study of whether greater pay for law professors produces better professorial output. The independent variables for comparative research productivity are certainly more reliable than the measures for judicial quality. Speaking as a law professor, however, I cannot say that I look forward to the results of that study.

⁶⁴ James F. Spriggs, II & Paul J. Wahlbeck, *Calling It Quits: Strategic Retirement on the Federal Courts of Appeals, 1893-1991*, 48 POL. RES. Q. 573, 589-90 (1995).

⁶⁵ Deborah J. Barrow & Gary Zuk, *An Institutional Analysis of Turnover in the Lower Federal Courts, 1900-1987*, 52 J. POL. 457, 473-74 (1990).

⁶⁶ Markon, *supra* note 18.

⁶⁷ *Leaving So Soon: Will Pay Issues Mean Only the Wealthy Can Afford To Serve?*, THE THIRD BRANCH: NEWSLETTER OF THE FEDERAL COURTS (Admin. Office of the U.S. Courts, Wash., D.C.), Feb. 2007, at 1.

⁶⁸ See Posting of Peter Lattman to Wall Street Journal Law Blog, <http://blogs.wsj.com/law/2007/09/21/judge-paul-cassell-resigning/> (Sep. 21, 2007, 15:48 EST).