
A COMMENT ON THE RELATIONSHIP BETWEEN JUDICIAL SALARY AND JUDICIAL QUALITY

STEPHEN MARKS*

Professor Scott Baker was kind enough to present his empirical research on the relationship between judicial salary and judicial quality¹ to the Law and Economics Workshop run by Professor Keith Hylton² and me last fall and I am honored to be able to comment on it today. It is part of a growing body of literature in law that tries to shed light on important issues through statistical analysis. Baker's paper, even before its publication, generated a significant amount of buzz.

As Baker points out and as is well-known, the Chief Justice of the Supreme Court has opined that low judicial pay threatens to undermine significantly the quality of the judicial system.³ This view, as Baker notes, has been "endorsed by prominent law school deans, the American Bar Association, and leading members of the corporate bar."⁴ Baker's statistical analysis casts doubt upon this thesis.

In this Reply I will first review the economic arguments underlying the thesis that low salaries reduce the quality of the judiciary, as well as the countervailing arguments. I will then look at Scott Baker's statistical analysis and try to determine if this analysis provides evidence as to whether salary affects the quality of the judiciary. I was impressed by the sophistication and creativity of the statistical analysis in the face of formidable evidentiary problems, e.g. how does one measure the effect of salary when all judges earn the same amount and how does one measure judicial quality. I will note my reservations about some of the assumptions underlying the statistical analysis and highlight certain things that I might have done differently. Nevertheless, Baker's work is a heroic effort carried out competently and creatively. Whether it ultimately succeeds and overcomes the inherent evidentiary

* Professor of Law, Boston University School of Law.

¹ The paper was ultimately published in the Boston University Law Review. Scott Baker, *Should We Pay Federal Circuit Judges More?*, 88 B.U. L. REV. 63 (2008).

² Paul J. Liacos Scholar in Law, Professor of Law, Boston University School of Law.

³ Chief Justice John G. Roberts, *2006 Year-End Report on the Federal Judiciary*, 39 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> (predicting that if the difference between judges' and lawyers' pay remains too large, "the judiciary will over time cease to be made up of a diverse group of the Nation's very best lawyers.").

⁴ Baker, *supra* note 1, at 65.

problems is a matter for the reader to decide. I will, however, try to shed some light on the issue.

Using a traditional statistical approach, Baker takes as his null hypothesis the proposition that current variations in judicial pay do *not* affect judicial quality. He then asks: is there sufficient evidence to reject this hypothesis? Baker does not find sufficient evidence. This result, as statisticians know, could mean two things. It could mean that there is indeed no relationship between salary and judicial quality, or it could mean that there is a relationship, but the data is so poor that it simply cannot reveal the relationship. I will discuss these issues below in greater detail.

My bottom line is that Baker's study provides mild support for his thesis that increasing judicial salaries will not improve the judiciary.

THE THEORY AND SOME PRELIMINARY EVIDENCE

The evidence for the pernicious effect of low judicial salary is twofold. First, distinguished lawyers receive significantly higher monetary compensation than do distinguished judges.⁵ This fact, coupled with economic theory, suggests that the best and brightest legal minds may go into lawyering rather than judging. Second, those involved in the recruitment of judges report that there are instances in which distinguished lawyers have declined offers of judgeships due to the pay differential.⁶

Economic theory suggests that higher pay will attract more candidates. Higher pay will produce a candidate pool that is a superset of the pool that would exist with lower pay. Put another way, higher pay attracts additional candidates without losing those candidates who would be attracted to the job even with lower pay.

Economic theory also suggests that candidates attracted by a higher salary are better candidates. Suppose, for example, (1) that candidates are motivated only by money and (2) that higher quality judicial candidates can command higher lawyer salaries as well. Under these assumptions, high quality judicial candidates will consider being judges only if the salary is commensurate with the high salaries they would receive as lawyers. In this model, higher judicial salaries will bring higher-qualified judges into the pool of judicial candidates. If the selection from this pool is also based on quality, or even if judges are picked randomly, then higher salaries should result in better judges. This simple economic model underlies the claims that low judicial salaries, relative to lawyer salaries, are threatening the quality of the judiciary. Under the

⁵ See, e.g., SALARIES OF FEDERAL JUDGES, ASSOCIATE JUSTICES, AND CHIEF JUSTICE SINCE 1968 1 (2007) (stating that in 2005, federal circuit judges received a salary of \$171,800); *The AmLaw 100, 2006*, AM. LAW., May 2006, at 165 (stating that in 2005, average partner salaries at a top Chicago law firm exceeded \$2 million).

⁶ See, e.g., Abner Mikva, *Judicial Pay: Attracting the Best*, NAT'L L.J. (June 4, 2007) (cited in Frank Cross, *Perhaps We Should Pay Federal Circuit Judges More*, 88 B.U. L. REV. 815 (2008)).

assumptions of this model, if judges are offered only 60% of what top lawyers make, then the inevitable result is an inferior judiciary. Of course, even with this model, the magnitude of the effect is not clear. It could be that lawyers who earn 60% of what top lawyers make would still make excellent judges. In this model, the degree of the effect on quality depends on how rapidly quality falls off as salary goes down.

Of course, people, including judges and lawyers, are not motivated by money alone. It is possible, as Baker points out, that judging may provide non-pecuniary benefits relative to lawyering.⁷ It may also be possible that the qualities that make a good judge are not the same as those that make a highly-paid lawyer. Academia provides a good illustration of both these effects. All of us know excellent law teachers and researchers who would not make it as highly-paid lawyers because, although they possess skills that make them excellent professors, they lack the skills that would make them highly-paid lawyers. The opportunity costs of taking a professorship are low for these people. We also know of excellent law teachers and researchers who could be highly-paid lawyers, but prefer the life of a professor, even at a greatly reduced salary. For these individuals, we say that the non-pecuniary benefits of being a professor, plus the professor salary, provide greater utility than the combined salary and non-pecuniary benefits of being a highly-paid lawyer. Thus, there are two reasons to suspect that, even at a lower salary, we still end up with an excellent set of professors. The same reasons apply to judges. Some excellent judicial candidates may have relatively low-paying alternatives because, in spite of having excellent judicial skills, they lack skills that would make them highly-paid lawyers. Other excellent judicial candidates might have very attractive alternatives as lawyers, but the non-pecuniary benefits of judging outweigh the salary differences.

Scott Baker focuses on the non-pecuniary benefits of judging. He posits that for some, judging confers large non-pecuniary benefits, inducing them to accept lower-paying judicial appointments rather than higher-paying lawyer positions.⁸ But do people who receive large non-pecuniary benefits make good judges? It could be that these non-pecuniary benefits are antithetical to good judging. For example, non-pecuniary benefits might include the ability to favor one's friends in important cases or to take large amounts of leisure time. If so, then lower judicial salaries could result in worse judges. It could also be that non-pecuniary benefits are unrelated to good judging. Perhaps some people just enjoy the process of judging more than the process of lawyering. This would mean that even with lower salaries, there could be excellent judicial candidates. A third possibility is that non-pecuniary benefits are actually positively related to good judging. Such benefits may include the ability to serve the public, to be engaged intellectually, to preserve the rule of

⁷ See Baker, *supra* note 1, at 66.

⁸ See *id.* at 73 (explaining that the non-pecuniary benefits of judging include "status, prestige, leisure, power to affect policy, and public service").

law, and so forth. In such a case, raising judicial salaries will increase the proportion of those in the pool who are interested in judging only for the money but who are unlikely to be better judges than those already in the pool.

The presence of non-pecuniary benefits thus complicates the picture and makes it less obvious that lower judicial pay will result in worse judicial candidates or that higher pay will produce better judicial candidates. It is possible that a salary of \$179,500, lifetime tenure, generous retirement benefits, the attendant prestige, and a work day that is intellectually stimulating are sufficient to draw into the applicant pool more than enough highly qualified judges to fill all positions. Raising salaries may just increase this surplus of qualified candidates, without increasing the quality of the judiciary at all.

In short, it is an empirical question. Other professions have similar issues. One might suspect, for example, that increasing professional major league baseball salaries would not increase the quality of play. On the other hand, one might suspect that an increase in the salaries of high school teachers would increase teacher quality. Of course, it is not surprising that judges feel their salaries should be higher. Nor is it surprising that there are stories about potential judges declining appointments due to the salary differential. Whether this will lead to worse judges, however, is unclear.

BAKER'S METHOD

In testing his hypothesis that salaries of circuit judges do not affect quality, Baker faces two problems. First, the independent variable does not vary because all circuit judges make the same amount of money. Second, the quality of judging is difficult to ascertain, making the dependent variable unobservable. Yet things may not be as hopeless as they appear. As for the problem of uniform judicial salaries, Scott Baker observes that the judges' salaries *relative to lawyers' salaries* have varied over time and across geography.⁹ Both judges' salaries and lawyers' salaries have changed over time, but at different rates, causing the differential between them to change. Additionally, while federal judges' salaries do not vary across geography, the salaries of lawyers do, which means the differential varies as well. These variations can be used to test whether the difference between lawyer and judge salaries affects the quality of the judiciary.

Consider a hypothetical. Suppose that the country were divided into two regions: the regions are identical except that in Region A, all prices and salaries are double those in Region B. Now let us introduce the market for judges. Suppose that judges are offered the same nominal salary in both regions. The following table gives a numerical example with these assumptions:

⁹ See *id.* at 76-77.

	Region A	Region B
(A) Cost of Living Index ¹⁰	2	1
(B) Lawyer Salaries	400,000	200,000
(C) Lawyer Salaries Adjusted for Cost of Living	200,000	200,000
(D) Judge Salaries	150,000	150,000
(E) Judge Salaries Adjusted for Cost of Living	75,000	150,000
(F) Net cost of becoming a judge (C) – (E)	125,000	50,000

Judges in Region A are effectively making half of what judges in Region B are making. In such a case, we would expect the applicant pool in Region B to be larger than the applicant pool in Region A. Provided that we had good measures of quality, we could test whether the quality of appointed judges was better in Region B.

Note that if judges are not mobile then judges are drawn from regional pools. The pool for Region B is bigger than that for Region A. If judges are completely mobile, then both regions would draw from a national pool of judges. The national pool available to Region A, however, would be a subset of the pool available to Region B. Thus, in this story, judicial mobility does not impede the analysis.¹¹ However, mobility will affect the analysis, as we will see below.

SOME RESERVATIONS ABOUT THE METHODOLOGY

Calculation of Real Lawyer Salaries

In order for the analysis to work, real judicial salaries must be compared to real lawyer salaries. The notion is to compare what a judge makes to what the judge could have made as a lawyer. What a judge earns is simply the nominal salary that the judge actually receives, discounted for cost-of-living and inflation. But what about the foregone lawyer salary? Which geographical legal market serves as a benchmark for potential lawyer salary depends on the assumptions about judicial mobility. Consider the following modified example:

¹⁰ This measure represents the cost of living in nominal dollars. In our example, it takes twice as many nominal dollars in Region A to purchase the same basket of goods as in Region B.

¹¹ I point this out because some critics of Baker's approach seem to believe Baker's analysis assumes judicial immobility. For example, Frank Cross, in a reply to Baker's article, states that many judges are in fact mobile, and that even within the regions that Baker uses, there exist large salary disparities. Frank Cross, *Perhaps We Should Pay Federal Circuit Judges More*, 88 B.U. L. REV. 815 (2008) (pointing out that "a good number of circuit court judges have relocated" and that the South Atlantic region used in Baker's Judge Sprouse example encompasses sub-regions with greatly differing salary norms).

	Region A	Region B
(A) Cost of Living Index	2	1
(B) Lawyer Salaries	400,000	350,000
(C) Lawyer Salaries Adjusted for Cost of Living	200,000	350,000
(D) Judge Salaries	150,000	150,000
(E) Judge Salaries Adjusted for Cost of Living	75,000	150,000
(F) Net cost of becoming a judge (no mobility)	125,000	200,000
(G) Net cost of becoming a judge (mobility)	275,000	200,000

Rows (F) and (G) represent the cost of becoming a judge in the two regions under the assumptions of *no mobility* and *mobility*, respectively. If there is no mobility between regions – which assumes that judges are appointed in the regions in which they reside – then we should calculate the lawyer salary within the region. In our example, the net cost of becoming a judge would then be higher in Region B. Based on Baker’s non-pecuniary analysis, a judge in Region A would need the equivalent of \$125,000 in non-pecuniary benefits to become a judge, while a judge in Region B would need \$200,000 in non-pecuniary benefits. On the other hand, if there is complete mobility, a candidate considering an appointment in Region A would compare the real lawyer salary in Region B versus the judicial salary in Region A. Thus, a candidate considering a judicial appointment in Region A would require \$275,000 in non-pecuniary benefits. A judge considering an appointment in Region B would require \$200,000 in non-pecuniary benefits.

Note that the mobility assumption completely affects the net cost of becoming a judge. Under the “salary matters” hypothesis¹² and the assumption of immobility, we would expect better judges in Region A. Under the “salary matters” hypothesis and the assumption of mobility, we would expect better judges in Region B. Thus, the mobility assumption has a critical effect on the analysis. Which is the better assumption? It would seem that if a judge has moved from her place of origin to a new place to judge, then the mobility assumption would be reasonable. In that case, the sacrificed lawyer salary should be that of the highest paying area, in real terms. Baker uses real lawyer salaries from the place of origin. This assumes no mobility, even when the judge demonstrated mobility by moving. The effect of getting this assumption wrong is to throw randomness into the calculation of the opportunity costs of judging. This, in turn, increases the likelihood that the coefficients will show

¹² Baker, *supra* note 1, at 74 (“Under the ‘salary matters’ theory, increased competition affects the kind of person eventually selected for the bench.”).

statistical insignificance, even if there is in fact a relationship between opportunity costs and judicial quality.

Cumulation Under the Direct Comparison Approach

Scott Baker utilizes two different methodologies to examine the relationship between judicial compensation and judicial quality. The first methodology, which he calls the “direct comparison approach,” looks at the amount each judge gives up over his or her career to be a judge.¹³ It assumes that those who become judges earlier in life give up more than those who become judges later, simply because the salary differential operates for a greater number of years. Unfortunately, cumulating the salary differential in this way is not, in my opinion, justified theoretically and has the effect of introducing unnecessary randomness into the independent variable. This randomness could easily mask any relationship between salary and quality that may exist under the “salary matters” hypothesis.

To see why this is the case, consider two hypothetical judges. Judge Younger serves for two periods, while Judge Older serves for one. Each judge receives some non-pecuniary benefit for being a judge as opposed to being a lawyer. Let us speculate that this non-pecuniary benefit comes from not having to work as hard. Specifically, each judge works five hours fewer per week than he or she would as a lawyer. The result of working fewer hours is that opinions take longer to write. In this model, taking longer to write an opinion is the sign of a less hardworking and thus a lower quality judge. Suppose that the value of this extra leisure is worth \$150,000 a year. Now suppose that judicial pay for both judges is \$175,000 and that each could earn \$300,000 a year as a lawyer:

Judge Younger

Potential Lawyer Pay: \$300,000

Judge Pay: \$175,000

Amount of Leisure if Judge: 5 hours a week

Value of Leisure: \$150,000 per year

Judge Older

Potential Lawyer Pay: \$300,000

Judge Pay: \$175,000

Amount of Leisure if Judge: 5 hours a week

Value of Leisure: \$150,000 per year

Both prefer to be judges rather than lawyers despite the decrease in pay because of the non-pecuniary benefits of judging relative to lawyering. Each takes the same amount of leisure and each produces opinions at the same speed. Because these judges are identical, except for length of service, we

¹³ See *id.* at 77-83.

really cannot use them to test the “salary matters” hypothesis. Yet under Baker’s methodology, finding identical results for these two judges is used as evidence against the “salary matters” hypothesis. This is because Baker cumulates the pecuniary losses over time. Since Judge Younger serves for two periods, her losses are \$250,000. Judge Older’s losses are only \$125,000. Thus, Baker’s methodology treats Judge Younger as if she had a salary significantly less than that of Judge Older. If the time to produce an opinion is identical for both judges, then, in Baker’s methodology, it is evidence against the “salary matters” thesis. This does not seem right – after all, the salaries per period are identical and one would not expect the data on quality to tell us anything. Consider another example:

Judge Younger

Potential Lawyer Pay: \$270,000

Judge Pay: \$175,000

Amount of Leisure as Judge: 4 hours a week

Value of Leisure: \$120,000 per year

Judge Older

Potential Lawyer Pay: \$300,000

Judge Pay: \$175,000

Amount of Leisure as Judge: 5 hours a week

Value of Leisure: \$150,000 per year

In this case, Judge Younger is getting a higher salary relative to lawyering than Judge Older. The “salary matters” thesis would predict that, because Judge Younger is making more money relative to being a lawyer, she should be a harder-working judge. And indeed she is, as evidenced by the fact she has less leisure time than Judge Older. However, Baker’s methodology treats Judge Younger as if she is making less money than Judge Older – she loses \$240,000 while Judge Older only loses \$150,000 – because Baker cumulates the salary differential over time. The fact that Judge Younger is harder working is treated as evidence against the “salary matters” thesis.

One can come up with many different examples. Sometimes cumulation amplifies the differences in single-period salary and sometimes it diminishes or even reverses the differences. In general, cumulating losses over time creates meaningless variation in the NETCOST variable – the measure of the pecuniary loss in taking a job as a judge relative to taking a job as a lawyer – and will likely mask any causal effects between NETCOST and measures of quality. That is, even if the “salary matters” thesis were true, the unrelated variation in the NETCOST variable from cumulation would likely mask the relationship between salary and judicial quality.

All is not lost, however. Baker also applies a pool approach to test the “salary matters” hypothesis. The “pool approach” forms a measure of the strength of the pool of judicial candidates at the time of a judge’s

appointment.¹⁴ The pool measure is the salary differential of a typical 49-year-old candidate at the time of the appointment and in the geographical area of the appointment.¹⁵ As with the direct comparison approach, pecuniary losses from judging are cumulated over time. However, the cumulation period is uniform (sixteen years) and, as a result, there is a 0.98 correlation between the cumulated net cost and the single period net cost.¹⁶ Thus, the NETCOSTPOOL variable can be taken as a near perfect proxy for single period differentials between lawyer salary and judge salary. Because this measure does not suffer from the problem of randomness generated by nonuniform cumulation, one might expect that it would better test the “salary matters” hypothesis.

Measures of Quality

Even if we had a well-defined and theoretically justified measure of judicial salary, such a study still requires measures of judicial quality. The thesis that “salary matters” is really a thesis that low judicial pay is compromising judicial quality. Baker uses the following proxies for judicial quality:

1. *Ideological voting.* Better judges are less likely to vote ideologically in controversial cases, where voting ideologically is voting in line with the ideology of the President and Senators involved in the judge’s nomination.¹⁷
2. *Ideological citing.* Better judges are more likely to cite the opinions of judges with contrary ideologies.¹⁸
3. *Dissents.* Better judges are likely to dissent more.¹⁹
4. *Opinion writing.* Better judges are likely to write their opinions more quickly.²⁰

¹⁴ *Id.* at 83-84 (addressing the “salary matters” argument that higher salaries would create a deeper – and better – pool of judgeship candidates).

¹⁵ *Id.*

¹⁶ This correlation was calculated from the data that Baker was kind enough to provide to his commenters.

¹⁷ *See id.* at 85 (“The operative assumption is that a more ideological judiciary will engage in more partisan voting patterns in [controversial] cases.”).

¹⁸ *Id.* at 85-86, 95-96 (“Under this measure, a more ideological judiciary consists of judges who seldom, if ever, recognize the opinions of judges from the other political party as persuasive authority.”); *see also* Stephen J. Choi & G. Mitu Gulati, *Ranking Judges According to Citation Bias (as a Means To Reduce Bias)*, 82 NOTRE DAME L. REV. 1279, 1281-82 (2007) (hypothesizing that ideologically motivated judges tend not to cite authority from other circuits when it was written by judges of a different political party).

¹⁹ *See Baker, supra* note 1, at 98-99 (explaining how dissenting has both temporal and social costs, and thus can be a measure of how hard a judge is willing to work).

²⁰ *See id.* at 101-03 (measuring the speed of case disposition in controversial cases that have advanced past the oral argument stage).

5. *Influence*. Better judges do not try to be influential.²¹

The notion is as follows: lower pay results in judges who receive higher non-pecuniary benefits for judging. These non-pecuniary benefits could lead to negative results in the judiciary: more ideological judging, fewer hardworking judges, or judges who value being influential. More ideological judging is evidenced by ideologically consistent voting and citing. Fewer hardworking judges is evidenced by a longer time to write opinions and fewer dissents. Influence mongering is evidenced by the production of opinions that are cited more often.

Of course, non-pecuniary benefits could also result in better judging. Positive non-pecuniary benefits would perhaps include the enjoyment of intellectual challenges, a desire to engage in public service, and so forth. This casts some doubt upon the proxies used by Baker. For example, a more diligent, careful, and intelligent judge may take longer to write opinions and may also produce better opinions, which are likely to be cited more often. So arguably, two of the proxies used by Baker as evidence of bad judging could also signify good judging. It also is not clear that better judges dissent more often. Good judges may seek consensus and be more willing to compromise.

Of course, since Baker finds the relationship between salary and any of these variables to be either statistically or economically insignificant, it is perhaps not as important to interpret whether they mean good judging or bad judging. Still, if we are trying to test a hypothesis about good judging, it is important that the proxies are strongly correlated with the quality of judging. If they are not strongly correlated, then the results of the statistical analysis, whether significant or not, will not say much about the relationship between salary and the quality of judging.

And here is a scary thought: suppose what we are really testing for is the mere presence of non-pecuniary benefit. That is, suppose that the null hypothesis is that judges do not receive any non-pecuniary benefit from judging relative to lawyering. In that case, statistically or economically insignificant results would support the null hypothesis and we might conclude that judges are not motivated by any non-pecuniary benefits. The only remaining explanation is that those who become judges do so only because they have no other attractive alternatives. If judging skills and lawyering skills are at all correlated, then this would support the “salary matters” thesis.

Other Quibbles

Scott Baker has ably anticipated other objections. Some of these objections are more serious than others. For example, it does not bother me that the NETCOST data are not judge-specific, especially since I believe that the pool approach is the better one in any case.²² I do worry about the multicollinearity

²¹ See *id.* at 105-07 (addressing the “salary matters” argument that lower judicial salaries result in more judges who crave influence).

²² See *id.* at 109-10.

that is generated by including circuit dummies. Baker points out that multicollinearity does not bias the results, and this is correct.²³ However, since a number of circuit dummies were statistically significant, I wonder if the NETCOST coefficient would be statistically significant if these dummies were removed. Multicollinearity caused by including the age variable should not be much of an issue in the pooled approach.

CONCLUDING REMARKS

As I indicated above, there are a number of things I might have done differently. I suspect, however, that the methodological changes that I would implement would not drastically change the results. First of all, I suspect that Baker's null hypothesis is true – that a salary of \$179,500, lifetime tenure, generous retirement benefits, the attendant prestige, and a workday that is intellectually stimulating is sufficient to draw into the applicant pool enough very highly qualified judges to fill all the positions and more. However, I also suspect that even if it were not true, and salary did matter, it would be very difficult to prove. As Baker points out, the data contain certain inherent limitations. The very imperfect measures of judge quality in particular make the analysis difficult. Nevertheless, there is enough here, in my mind, to at least nudge one in the direction of the camp that holds increasing judicial salary will not likely bring forth a better judiciary.²⁴ Whether it is enough to move anyone from one camp to another I cannot say. What is clear is that this study is creative, thoughtful, and meticulous. It squeezes more from the data than I would originally have thought possible. It was a pleasure to read and contemplate.

²³ *Id.*

²⁴ Bayesian analysis suggests that if there is even a small probability that the analysis could uncover a relationship between salary and judicial quality, the failure to find such a relationship should cause some revising of prior beliefs. For a brief description of Bayes' Theorem, see Stanford Encyclopedia of Philosophy, Bayes' Theorem, <http://plato.stanford.edu/entries/bayes-theorem/> (last visited Apr. 27, 2008).