# LITIGATION ECONOMICS DIGEST

Volume III Summer 1998 Number 1

## NOTES

The *Litigation Economics Digest (LED)* specializes in presenting applied articles in the discipline of forensic economics, and useful consulting information to the forensic practitioner.

The *LED* is published bi-annually at a subscription price of $50 a year for NAFE members, and $60 for non-members. Subscriptions should be sent to the National Association of Forensic Economics, P.O. Box 30067, Kansas City, Missouri 64112.

We are now soliciting articles in forensic economics for publication in future issues. The next issue will be Fall 1998. The *Digest* contains articles in the following arenas: Applied Forensic Economics, Law and Economics, Data Sources, Computer Programs and Ethics in Economics. Those interested should forward two copies of their article, along with any other pertinent information, and a reviewer fee of $25 (payable to LED), to the Editorial Office at the address shown below:

Litigation Economics Digest
7311 Borla Place
Carlsbad, CA 92009

(760) 944-9721 FAX: (760) 944-4551 email: litdigest@aol.com



While the *LED* does not accept advertisements of consulting services, we will accept advertisements concerning computer software, books and consulting support services. Advertising rates are available upon request.

The National Association of Forensic Economics expresses its thanks to the University of Missouri - Kansas City, for technical support in producing the *Litigation Economics Digest*.

# TABLE OF CONTENTS

Volume III Summer 1998 Number 1

## *Feature Articles*



## *Comment & Reply*

# THE USE OF STATISTICAL METHODS IN DISPARATE IMPACT CASES: THE NORTHERN MARIANA ISLANDS CASE

by

Raymond S. Hartman*

## I. Introduction and Overview

Regardless of whether the protected class is distinguished on the basis of gender, race, age and/or natural origin, in most litigation alleging discrimination against a protected class, statistical studies are performed to assess whether discrimination was indeed present. Not infrequently, the statistical approach involves comparing an average measure of economic performance or welfare for the protected class with an average measure of the same economic variable for the unprotected class. If, for example, average salaries or average rates of promotion are found to be statistically lower for a protected class relative to all other employees, that difference is often taken as prima facie evidence of a pattern and practice of discrimination.

Such simple statistical comparisons are appropriate only if the protected and unprotected classes are identical in all respects but for the attribute used to distinguish the protected class. If average compensation for a protected group of employees (say distinguished by race) is found to be less than the average compensation for all other employees, and that difference is statistically significant, this finding supports a claim of discrimination (disparate impact) only if the measured difference can be attributed to race alone. If the groups differ also in terms of education, training and job performance, observed differences in compensation will be attributable in part to these other differences. Before any weight can be given to the observed difference in average salaries across racial groups as evidence of discrimination, the differences attributable to education and performance must be taken into account.

While this may seem reasonable and obvious, it is surprising how frequently these other factors are ignored. To make this point more clearly and to explicate the methods required when the protected and unprotected classes are not identical (that is, in almost all real world cases), let me discuss arguments made in a particular case. Hence, while the discussion of this paper is anecdotal, it's focus ultimately is methodological.

In 1992, the United States Department of Justice (DOJ) conducted a study of the salaries of employees of the Public School System (PSS) of the Commonwealth of the Northern Mariana Islands (CNMI). Using 1990 data, their study purported to find that the PSS discriminated among three groups of employees distinguished by national origin: indigenous islanders (local Chamorros and Carolinians), Filipinos and United States citizens ("State-Siders"). They found the average salary of State-Side employees to be $3,600 more per year than the average salary of Filipino employees and $3,000 more than the average salary of the indigenous Chamorro/Carolinian employees. This difference, which was

* The author, who is Director, Cambridge Economics Inc., 4 Cambridge Center, Cambridge, MA 02142, gratefully acknowledges the comments of Jerry Hausman and Gregory Leonard.

statistically significant, was argued to be a violation of Title VII of the Civil Rights Act of 1964, as amended, 42, U.S.C. Section 2000e, et seq, and was the basis for a civil complaint against the Commonwealth and a variety of individuals.[1] The complaint alleged that the discrimination occurred over the period 1987 through 1993. Based upon the wage and salary discrimination complaint alone, the DOJ estimated compensatory damages of $2.8 million were due the group of Filipino teachers and $7.6 million were due the indigenous Chamorro/Carolinian teachers.

As part of the litigation, I was retained by the Commonwealth to examine whether there existed evidence of such wage and salary discrimination. My analysis relied upon employment data jointly gathered by plaintiff and defendants under court order (hereafter referred to as the Joint Data Base (JDB)[2]). I subjected these data to a more focused comparison of means and hedonic regression analysis and concluded the following:

- The initial DOJ analysis was flawed and did not support a finding of wage and salary discrimination. The analysis aggregated all employees with salaries between $14,000 and $40,000 in 1990. It compared the average salaries of all State-Side employees, all Filipino employees and all indigenous Chamorro/Carolinian employees and found the salary differentials mentioned above.

  This comparison ignored the composition of employees within each group. For example, the group of State-Siders included relatively more teachers with advanced degrees who occupied supervisory positions. These teachers should have been paid more, everything else equal. As discussed below, a comparison within comparable job categories is required to avoid such compositional differences.

- Analysis of the means of teacher salaries within comparable job categories reveals no pattern and practice of discrimination.

  A finding of a "consistent pattern and practice" of wage and salary discrimination is required for a violation of Title VII of the Civil Rights Act of 1964. In order to assess whether such a consistent pattern was in evidence, I analyzed average salaries for specific job categories in the Public School System in 1987, 1989, 1991 and 1993. I found no consistent pattern of discrimination in favor of State-

[1] *United States of America, Plaintiff, v. Commonwealth of the Northern Mariana Islands (CNMI), a commonwealth of the United States of America; Board of Education (BOE), an agency of the CNMI; Public School System (PSS), an educational system of the CNMI; and Commissioner of Education of the PSS, William S. Torres, Civ. No. 92-0016.*

[2] ***Ibid.***, Order For Creation of a Joint Data Base, May 14, 1993. The data include approximately 130 variables summarizing the qualifications, experience and performance of some 1100 teachers within the PSS over the 1987-1993 period.

Siders in any of the teacher classifications over the four years.

- Hedonic regression analysis of teacher salaries within comparable job categories reveals no pattern and practice of discrimination.

  Regression analysis allows us to estimate the salary that any teacher should have received, based upon her/his qualifications, experience and performance. I compared that "estimated salary" with the salary that each teacher did indeed receive. If teachers in a particular group distinguished by national origin consistently received salaries above or below the salaries they should have received, that would constitute a pattern and practice of wage and salary discrimination.

  Corroborating the analysis of mean salaries, I found no consistent pattern of wage and salary discrimination in favor of State-Siders in any of the job classifications over the four years. There were some years when Filipino and local teachers in certain job classifications were paid less than they should have been paid, relative to State-Side teachers in the same classification. However, there were other years when Filipino and local teachers were paid as much or more than they should have been paid, relative to State-Side teachers in the same job classification. These differences reflected competitive market conditions, not a pattern and practice of wage and salary discrimination.

- In more general terms of methodology, the method of analysis employed by the DOJ produced very misleading results and grossly exaggerated measures of damages. The comparison of simple means across employee groups with very different compositions of skills and training provides no information about disparate impact or treatment. More refined econometric techniques, that correct for differences in qualifications and performance, must be employed before conclusions concerning disparate impact or treatment can be drawn.

The discussion of my underlying analysis and conclusions proceeds as follows. In Section 2, I provide background information concerning the CNMI, the Public School System and the DOJ's complaint of discrimination. I also formulate several hypotheses concerning the determinants of teachers' wages and salaries. In Section 3, I formulate the methodological approaches required to test the hypotheses concerning the alleged wage and salary discrimination. I implement those methods in Section 4 using the Joint Data Base. Section 5 summarizes the paper.

## II. Background - Summary of Events and the Determination of TeacherCompensation

### A. Events

From 1986 through April 1989, all educational matters in the Commonwealth were under the jurisdiction of the Department of Education (DOE or CNMI). Hiring of teachers was implemented through graded civil service contracts for "locals" (Chamorros and Carolinians) and ungraded contracts for "aliens" (Filipinos, State-Siders and "Others"). The civil service contracts were regularized with Grades and Steps, and employees hired under these contracts encountered standard procedures regarding hiring, promotion, compensation and raises.[3] The terms of the ungraded contracts were quite idiosyncratic. Procedures regarding hiring, promotion, compensation and raises for personnel hired under these contracts were not standardized and the DOE could negotiate very aggressively with these employees on salary and terms of employment.[4] Given the more aggressive negotiating approach by DOE in recruitment and the lack of formalized salary increases, all ungraded teachers would have lower salaries than graded-contract teachers, everything else equal. These differences were independent of national origin, except for the fact that all Chamorros and Carolinians were graded contract holders. Among the ungraded contract holders, those recruits who were more aggressive negotiators did better, independent of national origin.[5]

The Public School System (PSS) was implemented in April, 1989, in part to consolidate the graded and ungraded systems so that all teachers would be graded consistently. A set of Grades/Steps and salary levels were proposed for all teachers (and revised in later plans). The civil service teachers (locals) were simply moved over to a similar Grade/Step and salary. The ungraded contract holders were placed on a Grade and Step that was determined by their then current salary, in order to maintain budget discipline. Hence, their initial Grade/Step was not determined by their educational background and teaching experience, as was the case with locals. To the extent that an ungraded teacher had historically accepted an initial salary offer below the Grade/Step implied by his/her educational background and teaching experience, this transition placed that teacher on a Grade/Step below that required of his/her educational background and teaching experience. The PSS transition also attempted to standardize procedures regarding hiring, promotions and raises.

[3] Over the alleged damage period, the Grades (job categories) included Teacher Aide I, II, III and IV; Developmental Teacher; Classroom Teacher; Classroom Teacher I, II, III, IV, V. Within each Grade, salary Steps were used to compensate differences in training, skills and performance.

[4] These teachers were not placed within a given Grade and/or Step; the salary offered was usually determined by competitive market forces and the negotiating skills of the recruit.

[5] In years when the overall budget constraints were tighter (e.g., FY 1989), the negotiating tactics of the recruiters were even more aggressive and the impacts on the starting salary offers to the ungraded teachers were more severe.

In 1990, a reclassification audit was undertaken to assess the progress of the transition under the PSS, and a 5% raise was given to all teachers. The reclassification audit made explicit many of the implicit differences in the system and contributed to dissatisfaction and the complaints of discrimination that initially brought the attention of the DOJ. Further changes occurred with the Salary Act of 1991, which raised each teacher's salary 3 Steps (5% per Step) while she/he remained at the same Grade.[6]

In 1992 and 1993, more serious budget pressures arose. Salaries and raises (within-Grade-increases or WGIs) were paid by continuing resolution (i.e., by off-budget agreements) and the WGIs were paid late and retroactively. Salary increases for promotions generated by reclassification (i.e., a teacher adding a degree and/or additional credits) were not allowed for the first time.

B. The Determinants of Teacher Salaries

Formal policies and procedures of the DOE/PSS during this period identify the following factors as most important in determining the salary of a given teacher.

- Educational background in the form of specific degrees, which determine Grade (job classification) and Step.
- Teaching experience, either within CNMI/PSS or outside.
- Seniority as measured by years within the CNMI system, whether as a teacher or not.
- Additional educational credits beyond the most recent degree.
- Performance evaluation, which led to merit increases before 1989 and which allow for a standard WGI post-1989.

In addition to these formal criteria, a number of other factors were cited as important in determining compensation, including the subjects and level of school (elementary or high) taught; whether the teacher participated in the Head-Start Program; the competitive conditions in the labor market and the extent PSS budget constraints were binding at recruitment; the number of DOE courses taken; whether the teacher had special education credentials; the negotiating skills of the recruit; and the type of contract originally binding the recruit (graded or ungraded).

## III. Formulation of Methodologies to Test for the Presence of Wage and Salary Discrimination

A. Comparison of Means

As discussed above, the analysis performed by the DOJ in this matter used the most

[6] As part of the law, raises for performance (within-grade-increases or WGIs) would be processed upon availability of the funds but would be paid retroactively, while raises for reclassification would not be retroactive.

simple comparison of means. DOJ focused upon all PSS employees with salaries between $14,000 and $40,000 in 1990 and compared the average salaries of all State-Side, Filipino and indigenous Chamorro/Carolinian employees. Because this comparison ignores the composition of employees within each group, it is flawed and without merit. Specifically, if the composition of employees within one group is predominated by Classroom Teachers and PSS administrators while the composition of employees within a second group is predominated by Teacher Aides and maintenance personnel, the average salary for the two groups must be different to account for the different average level of skills, training, background and human capital included in each group.

A more valid approach is to compare the differences in average salaries across national origin for all employees within a common job classification. Employees within a common job classification will possess comparable qualifications and human capital. Hence, any remaining differences in average salaries will be a more accurate measure of patterns of differences related to national origin. Therefore, I analyze the average salaries for the following specific classifications of teachers in the PSS for the years 1987, 1989, 1991 and 1993:

| | |
|---|---|
| Teacher Aide I | Developmental Teacher |
| Teacher Aide II | Classroom Teacher |
| Teacher Aide III | Classroom Teacher I |
| Teacher Aide IV | Classroom Teacher II |
| | Classroom Teacher III |
| | Classroom Teacher IV |
| | Classroom Teacher V |

B. Regression Analysis

Regression analysis can be used to measure the relationship between the salary of any particular teacher and his/her human capital attributes, which include qualifications, experience and performance. Several attributes were identified in Section 2.B as important to the DOE/PSS. Denote the value of each of these attributes for teacher i as $A_{ij}$, where i indexes each of the approximately 1100 teachers in the Joint Data Base and j indexes each of the N attributes above. The salary of teacher i can be explained by his/her human capital using the following equation:

$$(1) \qquad Salary_i = b_0 + b_1A_{i,1} + b_2A_{i,2} + \ldots + b_NA_{i,N} + e_i,$$

where $e_i$ is a random error, and the effect of any attribute j on salary is measured by the regression coefficient $b_j$. $b_j$ will measure either the dollar increase induced (say $750 per year of teaching experience) or the percentage increase induced by a given attribute (say

3.46% per year of teaching experience).[7]

Having estimated $<b_0, b_1, ..., b_N>$, we can use Equation (1) to predict the salary any teacher should receive, given his/her background and human capital attributes. If teacher background and attributes completely determine teacher salary, then the difference between the actual salary and the salary predicted by the regression model will be small and random across all teachers. If these residuals are randomly distributed across all teachers, we shall have demonstrated that there exists no pattern of preferential treatment for any specific group of teachers. To the extent that the residuals are not random, preferential treatment will be indicated. For example, if the measured residuals are uniformly negative for one group of teachers identified by national origin (that is, actual salary < predicted salary), then these teachers were systematically underpaid, everything else being equal. If on the other hand, the residuals are uniformly positive for another group of teachers (i.e., actual salary > predicted salary), then this group of teachers were systematically overpaid, everything else being equal.[8]

C. Hypotheses to be Tested

While I investigate the following hypotheses using both methodologies, let me discuss the hypotheses using the regression methods.

Hypothesis 1 — There exists a pattern of wage and salary discrimination against Filipino and indigenous (Chamorro/Carolinian) teachers and in favor of State-Side teachers.

This hypothesis is the basis for the Justice Department's complaint. DOJ is essentially claiming that the residuals for all Filipinos and Chamorros/Carolinians are consistently negative (i.e., actual salary < predicted salary, for most years and most teacher categories) and that the residuals for all State-Siders are consistently positive (i.e., actual salary >

[7] Whether the regression coefficients measure the dollar effect or the proportional effect depends upon the specification of the salary variable. The most standard form is to measure salary as Log (Salary) and the $b_j$ in proportional terms. Both specifications led to equivalent results with this data base.

[8] A third alternative, a selectivity model, could have been specified and estimated to analyze the **joint** hiring and compensation practices of the CNMI. Such a model, described in Appendix A, could be used to analyze whether a pattern and practice of discrimination in hiring existed across groups differentiated by national origin. The hedonic wage and salary regression in the text can be thought of as conditional upon those hiring practices. Since hiring practices were not subject to the complaint, data were not gathered with which such selectivity models could be estimated. If a pattern and practice of discrimination in hiring had existed, it would have revealed itself in the hedonic wage and salary regressions. As discussed in Appendix B and the text, I found no evidence of such a pattern and practice.

predicted salary) for most years and most teacher categories.

However, there exist several alternative hypotheses arising from the historical record (Section 2).

Hypothesis 2 Teachers hired during competitive hiring years or in competitive labor markets (geographic) during periods of serious DOE/PSS budget constraints reveal a non-random pattern of negative residuals (actual salary < predicted salary).

Anecdotal information suggests that State-Siders, Filipinos and persons of "Other" national origin fit into this group. To the extent that teachers in this group were not as aggressive in bargaining as were the DOE/PSS recruiters, this non-random pattern of negative residuals will be compounded. Anecdotal information suggests that Filipinos and a subset of State-Siders (e.g., secondary wage earners in a two-person, State-Side household) fit into this group.

Hypothesis 3 Teachers grand-fathered into the PSS system from ungraded contracts under DOE reveal non-random negative residuals (actual salary < predicted salary).

This should be particularly true in the early years of PSS. The non-random pattern may dissipate with the reclassification efforts that occurred from 1990 onward. However, those reclassification efforts were slowed down with the budget problems since FY 1992.

Hypothesis 4 Teachers who always had graded contracts (with standard raises and seniority increases) and who have not had to confront the aggressive negotiating tactics of DOE/PSS recruiters in competitive markets (by year and by geographic area) reveal a non-random pattern of positive residuals (actual salary > predicted salary).

Anecdotal information suggests that native Chamorros and Carolinians fit into this group.

Hypothesis 5 All teachers promoted since the severe budget constraints of 1992 reveal a non-random pattern of negative residuals (actual salary < predicted salary).

Anecdotal information suggests that this group contains teachers of all national origins.

## IV. Results of the Analysis

### A. Focused Comparison of Means

Table 1 summarizes, by national origin, the number of teachers in each job category and their average salaries in 1987 and 1993. Similar measures were developed for

1989 and 1991. We find that indigenous peoples represented the majority of teachers in 1987. 213 out of 321 (66%) teachers were CNMI natives. The CNMI personnel were distributed throughout the job classifications. In 1987, ungraded contract "alien" teachers were fairly equally divided among the 3 other groups distinguished by national origin.

Over the next six years, the number of PSS teachers approximately doubled to 604. The local population was not sufficient to meet this growth; CNMI staff increased by only 20 (to 233). PSS turned to off-island sources of labor, i.e., the US, Philippines and "Other" areas. By 1989, the size of the Filipino group and "Other" group rose to approximately 60 persons. By 1993, Filipinos accounted for 150 positions, State-Siders accounted for 120 and Other national origins accounted for 101. In 1993, CNMI persons accounted for 233/604 (39%) of the teaching staff.

In Table 1, the highest average salary is underlined for each job classification in 1987 and 1993. The group with the consistently highest salaries is not the State-Siders, as maintained by DOJ. Instead, it is the CNMI group. For comparable job classifications, CNMI teachers receive the highest average salary 26 out of 39 times for which comparable data exist for all four years. Filipinos and State-Siders have the highest salary in approximately the same number of cases (5 for Filipinos and 6 for State-Siders). Viewed another way, State-Siders have the lowest salary in 9 of the comparable cases; Filipinos have the lowest salary in 11 comparisons; and CNMI teachers have the lowest salary in only 6 cases.

I find therefore no evidence of a pattern of wage and salary discrimination in favor of State-Siders and against Filipinos and the indigenous peoples of CNMI. I find instead a pattern of wage and salary preference in favor of the indigenous peoples of CNMI. This pattern reflects normal competitive market forces. The indigenous peoples of CNMI are the long-term residents, who have worked within the school system for a longer period and will continue to work within the school system after many of the more transient "alien" teachers have returned to their home countries. The CNMI teachers have been protected by graded contracts for a longer period of time (Hypotheses 3 & 4). They have not been subjected to the same competitive market forces in recruitment (Hypothesis 2). I would expect this group to reveal higher wages and salaries, everything else equal.

I would also expect that in competitive labor markets, State-Siders and Filipinos would reveal higher wages and salaries in some teacher classifications and in some years, reflecting the specific hiring needs of the PSS in those years, the availability of teachers with specific qualifications and the negotiating skills of the relevant applicants.

Table 1: Summary of Average Salaries (in $): By Teacher Classification and National Origin for Selected Years

1987

| Teacher Classification (Total = 321) | State-Siders (34) | Filipinos (33) | CNMI (213) | Other (41) |
|---|---|---|---|---|
| Developmental Teacher | -- | --- | --- | --- |
| Classroom Teacher (38) | 18,041 (12) | 15,536 (16) | 6888 (9) | 21,224 (1) |
| Classroom Teacher 1 (38) | 9,980 (2) | 10,223 (1) | 12,365 (26) | 10,585 (9) |
| Classroom Teacher 2 (80) | 13,519 (3) | 11,269 (3) | 15,042 (64) | 14,063 (10) |
| Classroom Teacher 3 (52) | 14,285 (6) | 15,856 (4) | 17,629 (34) | 16,781 (8) |
| Classroom Teacher 4 (21) | 23,861 (3) | 22,011 (4) | 20,225 (11) | 23,116 (3) |
| Classroom Teacher 5 (15) | 25,452 (5) | 24,916 (5) | 22,895 (2) | 24,255 (3) |
| Teacher Aide 1 (29) | 6,597 (1) | --- | 6,783 (27) | 6,283 (1) |
| Teacher Aide 2 (6) | --- | --- | 8,448 (6) | --- |
| Teacher Aide 3 (6) | --- | --- | 9,215 (6) | --- |
| Teacher Aide 4 (36) | 10,233 (2) | --- | 11,171 (28) | 12,530 (6) |

1993

| Teacher Classification (Total = 604) | State-Siders (120) | Filipinos (150) | CNMI (233) | Other (101) |
|---|---|---|---|---|
| Developmental Teacher (62) | 20,234 (2) | 21,243 (1) | 23,380 (46) | 21,249 (13) |
| Classroom Teacher (4) | 21,506 (1) | 24,889 (1) | 16,607 (2) | --- |
| Classroom Teacher 1 (25) | 19,983 (13) | 24,909 (3) | 22,200 (2) | 21,478 (7) |
| Classroom Teacher 2 (134) | 25,427 (41) | 25,678 (37) | 29,098 (21) | 26,378 (35) |
| Classroom Teacher 3 (90) | 27,790 (20) | 26,849 (52) | 35,857 (13) | 32,634 (5) |
| Classroom Teacher 4 (98) | 31,072 (38) | 28,666 (49) | 33,704 (2) | 33,262 (9) |
| Classroom Teacher 5 | --- | --- | --- | --- |
| Teacher Aide 1 (111) | 12,462 (3) | 13,778 (4) | 14,161 (87) | 13,837 (17) |
| Teacher Aide 2 (45) | --- | 17,482 (1) | 16,635 (38) | 16,154 (6) |
| Teacher Aide 3 (35) | 16,671 (2) | 15,520 (2) | 16,960 (22) | 16,117 (9) |
| Teacher Aide 4 | --- | --- | --- | --- |

Notes:
Number of Teachers in Parentheses.

B. Hedonic Regression Analysis

The comparison of means focuses upon actual wages and salaries. This is precisely the comparison performed, incorrectly, by Justice in initiating this litigation. In Table 1, the measure has been correctly developed for comparable job classifications. While these comparisons indicate that Filipinos and the indigenous peoples of the CNMI (Chamorros/Carolinians) received higher salaries in as many cases or more than did State-Siders, the DOJ could claim that the salaries for the Filipinos and CNMI teachers should have been even higher than they were, and that the fact that these salaries were not higher reveals a pattern and practice of wage and salary discrimination.

This interpretation can be critically examined with hedonic regression analysis.[9] If the PSS determined the salary of each teacher using a common criteria summarizing the qualifications, experience and performance of that teacher, and if those criteria were implemented independently of national origin, then Equation 1 should predict the salary expected for that teacher, given that teacher's attributes. If the predicted (expected) salary is consistently more or less than the actual salary paid to the members of a specific group of teachers identified by national origin, then those wages and salaries will reveal the pattern of wage and salary discrimination identified by Hypothesis 1 in Section 3.C.

I examine the validity of this hypothesis by estimating regression Equation 1 for all teachers within each job category identified in Table 2.[10] Because many

[9] Hedonic analysis has come to be used in a variety of applications, including demand analysis, cost analysis and analyses supporting litigation. The References provide an introduction to this literature.

[10] This particular pooling is supported at the 99% confidence level. To test the validity of pooling the data across the 11 job classifications in Table 1, I estimate Equation (1) for all teachers for all years and all job categories, yielding $b_{pool}$. For the unpooled model, I estimate Equation (1) for all teachers within each of the 11 job classifications, yielding ($b_{tc}$, tc = 1, 11). I reject the pooled model at well above the 99% level with $F_{230,1026} = 9.644$. Furthermore, closer scrutiny of the classification "Classroom Teacher" reveals a miscellaneous catch-all category which includes data on classroom teachers, teacher aides, and Headstart teachers. I tested the appropriateness of this pooled category versus disaggregation into "Classroom Teachers--Other" and "Teacher Aides/Headstart Teachers". I rejected the pooled miscellaneous category at well above 99% ($F_{9,99} = 8.173$). As a result, the data argue for differentiating the 12 job categories in Table 2. These results and tests are available from the author upon request.

A complete description of the regression variables is available upon request. Briefly, the data include the following. The dependent variable is log (teacher $Salary_t$, deflated by the CNMI CPI). The attributes hypothesized to explain salary include: **number** of degrees earned prior to t, by type {DEG_AS (associate), DEG_BA (bachelors), DEG_MA (masters), DEG_PHD (doctorate), DEG_OT (other)}; **whether** a degree by type had been earned prior to t {DEG_AS1, DEG_BA1, DEG_MA1, DEG_PHD1, DEG_OT1}; teaching experience {EXPER_1 (years within DOE/PSS), EXPER_2 (years within CNMI private schools), EXPER_3 (years outside CNMI), EXPER_4 (equivalent

of the attributes are similar for teachers within each job classification, I expect those attributes will add little explanatory power and consequently will drop out of the estimated regression equation.[11] Table 2 identifies those attributes that do prove to be statistically important in determining teacher salary for each job classification. For example, the number of associate degrees (DEG_AS), regular teaching experience within the CNMI (DOE/PSS) public school system (EXPER_1), regular teaching experience outside CNMI (EXPER_3), and years spent in full-time education (YRS_ED) were the determining factors of salaries for all Developmental Teachers over the 1987-1993 period.[12] Alternatively, the factors determining level IV Classroom Teacher salaries were the following: whether the teacher had a masters and/or PHD degree (DEG_MA1, DEG_PHD1); teaching experience within the CNMI public school and private school system (EXPER_1, EXPER_2); experience as a substitute teacher (EXPER_5); years of seniority (YRSSEN); age (AGE); special education credentials (SP_ED); and whether the teacher works in a high school (HIGH).[13]

I expect that the statistically important attributes will differ by teacher classification. For example, I do not expect to find that the presence of a BA or MA degree will affect the salaries of Teacher Aides, because those personnel do not have such degrees. Likewise, I expect that the impact of the same attribute

years of summer or part-time teaching), EXPER_5 (experience as a substitute teacher)}; years of DOE/PSS seniority (YRSSEN); semester credit hours of continuing education (SEM_CRED); general non-specific human capital proxied by AGE; number of BOE courses completed (BOE_COR); years spent in full-time higher education (YRS_ED); special education experience and credentials (SP_ED); experience in CNMI high schools (HIGH); participation in CNMI Headstart program (HEAD); most recent performance evaluation (PER_EVAL); and fixed effects for snapshot date {YEAR87, YEAR89, YEAR91, YEAR93}, for national origin {NO_STAT (State Side), NO_FILI (Filipino), NO_CNMI (CNMI, Chamorro/Carolinian), NO_OTH (Other)}, and for national origin/year {(FILI87, FILI89, FILI91, FILI93 -- designating Filipino teachers by year), (CNMI87, CNMI89, CNMI91, CNMI93 -- designating CNMI teachers by year), (OTHE87, OTHE89, OTHE91, OTHE93 -- designating teachers of "Other" national origin by year)}.

[11] I eliminate those variables which fail standard t and F tests (available upon request).

[12] The regression results indicate that an Associate Degree adds 4.96% to a Developmental Teacher's salary and each year of CNMI experience adds 1.68%. See Appendix B.

[13] These factors have the following effects upon expected salaries: masters degree (+2.71%); PHD (+9.67%); an additional year of teaching experience within either the CNMI public (+0.72%) or private school systems (+1.42%); experience as a substitute teacher (+3.54%); a year of seniority (+1.14%); special education credential (+7.58%); and high school position (+2.01%). Notice further that the important degree variables are **whether** the teacher had a degree (DEG_MA1), rather than the **number** of degrees (DEG_MA). Both degree variables were tested. See Appendix B.

Table 2: Summary of Regression Results[a]

| Teacher Classification | Statistically Significant Teacher Attributes | National Origin Effects[b] Filipino | CNMI | Other |
|---|---|---|---|---|
| Developmental Teacher | DEG_AS, EXPER_1, EXPER_3, YRS_ED | - | + | + |
| Classroom Teacher (Catch-all) | | | | |
| ■ Classroom Teacher | DEG_MA, AGE, HIGH | - | + | - |
| ■ Teacher Aide/ Head Start | HEAD | n.a. | - | n a. |
| Classroom I | EXPER_1 | +7.11% | + | +3.93% |
| Classroom II | DEG_AS1, EXPER_1, HIGH | -4.47% | -5.33% | -3.21% |
| Classroom III | EXPER_1 | -7 53% | + | - |
| Classroom IV | DEG_MA1, DEG_PHD1, EXPER_1, EXPER_2, EXPER_5, YRSSEN, AGE, SP_ED, HIGH | -7 59% | - | + |
| Classroom V | DEG_MA, EXPER_2, EXPER_5, YRSSEN, SEM_CRED, BOE_COR, SP_ED, HIGH | - | + | + |
| Teacher Aide I | EXPER_1, BOE_COR, HEAD | + | + | + |
| Teacher Aide II | EXPER_1, EXPER_3, HEAD | + | - | - |
| Teacher Aide III | EXPER_1, YRSSEN | - | + | - |
| Teacher Aide IV | EXPER_1, EXPER_4, AGE | + | - | - |

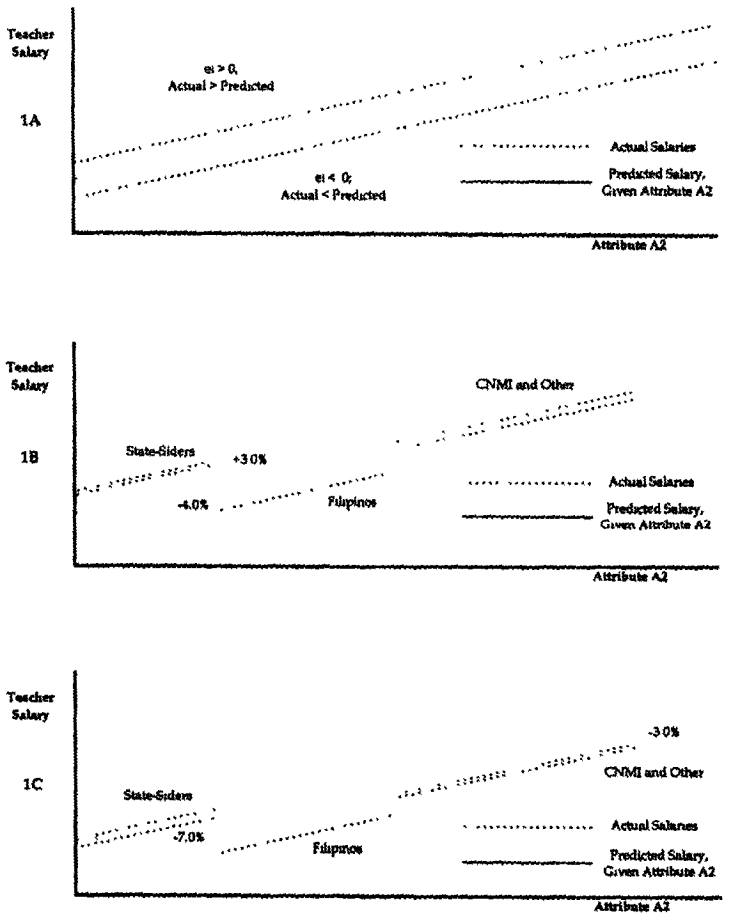
Notes.

a) Results (available from the author on request) summarize effects over the entire 1987-1993 period.

b) Deviation of actual salary from expected salary, relative to State-Siders. For example, Filipino Classroom I teachers are paid 7.11% more than expected, given their attributes, relative to State-Siders. While the sign of each national origin effect is presented, the size of the effect is presented only when statistically significant at the 95% level

c) n.a. signifies "not appropriate", meaning that there were no employees of this national origin within this job classification.

## Figures 1A - 1C



will differ for different teacher classifications.[14]

The estimated regressions indicate how the PSS implements its wage and salary criteria for each job category. Those attributes that are identified as statistically significant in Table 2 are the ones used by PSS in determining wage and salary offers within that job classification. These criteria will be administered without discrimination if two conditions obtain: 1) actual salaries are approximated by the salaries predicted by the regression for the teachers, given their attributes; and 2) differences between actual and predicted salaries reveal no consistent pattern for any particular group of teachers. Figure 1A indicates the pattern of results that will obtain if there is no wage and salary discrimination in a job classification. In Figure 1A, a teacher's expected salary is predicted by the regression line, given the value of the teacher's attribute $A_2$.[15] The regression line predicts a teacher's salary with some error ($e_i$), which is the regression residual in Equation 1. If all teachers are distributed randomly around their expected salary (the regression line), there exists no observable pattern in the $e_i$s for identifiable groups of teachers. In this case, there is no discernable pattern of wage and salary discrimination.

Figure 1B indicates the results that will obtain if there is a pattern of wage and salary discrimination. If the distribution of actual teacher salaries around their expected salaries is conditioned by national origin, there will exist an observable pattern in the $e_i$s for identifiable groups of teachers. In this case, some groups will be distributed above the regression line (actual salary > expected salary; $e_i > 0$; teachers are paid more than expected given their attributes) while some groups will be distributed below the regression line (actual salary < expected salary; $e_i < 0$; teachers are paid less than expected given their attributes). In the hypothetical example in Figure 1B, the actual salaries of State-Side teachers are shown as being consistently higher than those predicted by the regression line, given the level of attribute $A_2$. Assume that the actual salaries are approximately 3% greater than they should be for this group. The actual salaries of Filipino teachers are shown as being consistently less than those predicted by the regression line, given the level of attribute $A_2$. Assume that the actual salaries are approximately 4% less than they should be for this group. Finally, assume that the actual salaries of CNMI and "Other" teachers are distributed randomly around the regression line. Hence, the actual salaries of these two groups are good approximations of the salaries that they should receive, given their attributes. Notice that if the hypothetical results in Figure 1B were to obtain for the teachers in a particular job classification, then Filipino teachers would receive approximately 7% less than State-Side teachers were receiving, given their attributes, and CNMI and "Other" teachers would receive 3% less than the State-Side teachers were receiving, given their attributes.

[14] For examples, the value of a year of teaching experience (EXPER_1) has the following respective salary impacts for Developmental Teachers, Teacher Aides I, Teacher Aides III, Classroom Teachers I and Classroom Teachers IV: 1.68%, 2.72%, 0.49%, 2.12%, 0.72%. See Appendix B.

[15] For example, $A_2$ could summarize either DEG_AS, EXPER_1, EXPER_3 or YRS_ED for Developmental Teachers.

We can test for this pattern of wage and salary discrimination across groups distinguished by national origin by including a variable in Equation (1) that identifies the national origin of each teacher. The test is usually conducted relative to one of the groups. Given the Department of Justice's complaint, I use State-Siders as my basis of comparison, which leads to the same insights developed with the hypothetical in Figure 1B. Specifically, the regression line is run through the actual salaries of the group used as the basis for comparison (State-Siders), and average treatment of the other groups relative to the comparison group is then made explicit (Figure 1C).

Table 2 summarizes the regression results from these tests over the entire 1987-1993 period. This test examines whether any national origin effect occurred on average over the period. The test should be thought of as preliminary because it does not pinpoint the years in which the effect actually occurred. To obtain the results in Table 2, I added to the original hedonic regressions binary variables summarizing national origin for Filipinos, CNMI teachers and Other teachers.[16] The results in Table 2 describe the sign of all national origin effects (whether the effect was positive or negative) and present the size of the effect only in those cases that were statistically significant.

Notice that the national origin effects are not uniform. For each national origin, some effects are positive and some are negative. Likewise, those effects which are statistically significant are not uniformly positive or negative. For example, the salaries of Filipino and Other Teachers were higher than they should have been (by 7.11% and 3.93%), relative to State-Side teachers in the Classroom Teacher I category. Within the Classroom Teacher II category, however, the teachers of all three national origin groups were less than they should have been, relative to State-Siders. Furthermore, in all cases that are not statistically significant (27 out of 34), there exists no pattern by national origin. In all of these cases, all four groups are treated equivalently. Based upon these findings, I conclude that, overall, there is no evidence of a consistent pattern of wage and salary discrimination against any of these groups distinguished by national origin over the 1987-1993 period on average.

I disaggregate these results further by year in Table 3,[17] where statistically significant incidents of overpayment or underpayment by national origin occur in only 10 out of 144 possible cases, or about 7% of the cases. In all other cases, the wages and salaries of the teachers of the specific group designated by national origin were equivalent to those paid to State-Siders, given their attributes. Furthermore, the 10 cases that are statistically significantly different from 0 are themselves not uniform. Some are positive; some are negative. Based upon these results, I conclude that there exists no consistent pattern of wage and salary discrimination in favor of State-Side teachers in any of the snapshot years from 1987-1993. While there do exist a few cases (8 out of 144) where teachers of national origin other than State-Side were paid less than they should have been, relative to State-Siders, these findings do not represent a pattern of wage and salary

[16] The binary variable summarizing State-Side national origin is included in the intercept, as pictured in Figure 1C. These regression results are available upon request.

[17] This measurement is accomplished by interacting the binary dummy variables for national origin with the binary dummy variables for years. See footnote 10 and Appendix B.

discrimination.

## V. Summary and Conclusions

While I have not documented the alarming extent to which I have generally found simple-minded comparisons of group averages (protected v. unprotected classes) used incorrectly to support complaints of discrimination, I have described in some detail a particular case. Given the suspected inadequacy of the Justice Department's original comparison of average wages and salaries for all Filipino, all CNMI and all State-Side teachers, I conducted two sets of more specific analyses to better assess whether there existed any evidence of a pattern of wage and salary discrimination in the Public School System of the Commonwealth of the Northern Mariana Islands over the period 1987-1993.

In the first set of analyses, I compared average wage and salary for the teachers in the protected and unprotected classes within distinct job categories. The focus upon distinct job categories insures that the teachers being compared are, if not identical, at least much more similar in all ways but for national origin. If differences are found in these average wages across groups distinguished by national origin, such differences can more reliably be attributed to disparate treatment based upon national origin alone. In the second set of analyses, I made use of hedonic regression methods to analyze and estimate the wages and salaries that all teachers within comparable job categories should have received, given their qualifications, experience and performance. I then compared the expected compensation with the wages and salaries that the teachers did indeed receive and examined whether the differences between the expected and the actual wages and salaries showed any pattern by national origin. Table 1 summarizes the results of my comparison of average wages and salaries for 11 job categories from 1987 to 1993. Table 2 and 3 summarize my hedonic regression analysis.

In Table 1, I find absolutely no evidence of a pattern of wage and salary discrimination in favor of State-Siders and against Filipinos and the indigenous peoples of CNMI. I find that CNMI teachers receive the highest average salary 26 out of 39 times for which data exist to support such a comparison; Filipinos and State-Siders have the highest salary in approximately the same number of cases (5 for Filipinos and 6 for State-Siders). These results demonstrate that the Justice Department's original aggregate comparison of wages and salaries was flawed and without merit for the purposes of the complaint in this matter.

Indeed, the patterns in Table 1 reflect normal competitive market forces. The indigenous peoples of the CNMI are the long-term residents, who have worked within the school system for a longer period and will continue to work within the school system after many of the more transient "alien" teachers have returned to their home countries. The CNMI teachers have been protected by graded contracts for a longer period of time and have not been subjected to the same competitive market forces in recruitment. As a result, I expected this group to reveal higher wages and salaries, everything else equal. I also expected that State-Siders and Filipinos would reveal higher wages and salaries in certain teacher classifications and in certain years due to the following: competitive conditions in the labor markets in those years, in those countries and for those job categories; the hiring

needs of the PSS in those years; and the negotiating skills of the relevant applicants.

Based upon statistical evidence, I find that one of the original 11 job categories, "Classroom Teacher (Other)", is too broad and includes teachers too diverse to be grouped together. Therefore, I divide that category into two sub-categories: Classroom Teacher and Teacher Aide/Head Start Participant. For these 12 job categories over the four snapshot years of my analysis, I find in Table 3 that there exist only 10 cases (out of 144) for which there exists statistically measurable differences by national origin between expected and actual wages and salaries. In all other cases (134 out of 144, or 93%), teachers of all national origin are paid commensurately, given their qualifications, experience and performance. Based upon these findings, I conclude that there exists absolutely no evidence of a pattern of wage and salary discrimination among teachers in the Public School System in favor of State-Siders and against indigenous and Filipino peoples.

The evidence of a lack of a consistent pattern is even stronger because the measured national origin effects are not consistently negative. Of the six cases for which the national origin effect for Filipinos is statistically significant, 1 (or 16% of the total cases) is positive. Specifically, Filipino teachers in the Classroom Teacher I job category were paid 16.42% more than they should have been paid, relative to State-Siders, given their qualifications, experience and performance. Likewise, of the three cases for which the national origin effect for CNMI teachers is statistically significant, 1 (or 33% of the total cases) is positive.

Table 3· National Origin Effects.
by Job Category and Year

| Teacher Classification | National Origin Effects | | |
|---|---|---|---|
| | Filipino | CNMI | Other |
| Developmental Teacher | | | |
| 1987, 1989, 1991, 1993 | -- | -- | -- |
| Classroom Teacher (Catch All) | | | |
| Classroom Teacher | | | |
| 1987, 1989, 1991, 1993 | -- | -- | -- |
| Teacher Aide/Head Start | | | |
| 1987, 1989, 1991, 1993 | -- | -- | -- |
| Classroom I | | | |
| 1987, 1989, 1991 | -- | -- | -- |
| 1993 | +16.42% | -- | -- |
| Classroom II | | | |
| 1987, 1989, 1993 | -- | -- | -- |
| 1991 | -6.30% | -9.80% | -5 59% |
| Classroom III | | | |
| 1987, 1989 | -- | -- | -- |
| 1991 | -9.51% | -- | -- |
| 1993 | -6.92% | -- | -- |
| Classroom IV | | | |
| 1987 | -- | -9.57% | -- |
| 1989 | -- | -- | -- |
| 1991 | -9.77% | -- | -- |
| 1993 | -5.88% | -- | -- |
| Classroom V | | | |
| 1987, 1989, 1991, 1993 | -- | -- | -- |
| Teacher Aide I, III, IV | | | |
| 1987, 1989, 1991, 1993 | -- | -- | -- |
| Teacher Aide II | | | |
| 1987, 1991, 1993 | -- | -- | -- |
| 1989 | -- | +13.44% | -- |

Notes:

All numerical entries are statistically significant at the 95% level. The results for all other years and national origins (--) are not statistically different from zero

Specifically, CNMI teacher aides in the Teacher Aide II job category were paid 13.44% more than they should have been paid, relative to State-Siders, given their qualifications, experience and performance. These finding only strengthen my conclusion that there exists absolutely no evidence of a pattern of wage and salary discrimination among teachers in the Public School System in favor of State-Siders and against indigenous and Filipino peoples.

The evidence in Table 3 indicate that in the overwhelming majority of cases, the PSS consistently and successfully implemented personnel procedures and wage and salary criteria that fairly and commensurately compensated teachers of all national origins based upon each teacher's qualifications, experience and performance. In the few cases where actual compensation measurably deviated from expected compensation, the pattern of those deviations are not consistently for or against Filipino and indigenous teachers.

Again, I find that some of the measurable deviations in Table 3 reflect normal competitive forces. As indicated in Sections 2 and 3, PSS budget constraints became progressively more severe starting in 1989. As a result, PSS recruitment attempted to be aggressively cost effective with all new hires, the majority of which were off-island. If they did not behave in this way, they would be derelict in their duty to their taxpayers. I find evidence that aggressive salary negotiations were conducted regardless of national origin. However, competitive market forces and the negotiating posture/behavior of Filipino teachers made them more likely to accept actual salaries which were less than expected, given their attributes. Indeed, based upon normal competitive market forces, I expected to have found more cases of actual salaries being less than expected for Filipino teachers. I conclude from the fact that there exist so few measurably negative (and one positive) national origin effects for Filipinos, that the PSS went out of its way to aggressively implement fair and unbiased wage and salary criteria, in spite of the competitive market forces and budget constraints that they faced.[18]

## Appendix A: Issues in Selectivity Model Specification and Estimation

Specified most generally, the statistical problem to be analyzed here is a regression switching model with endogenous switching induced by an ordered response model.[19] For each job category k (k = 1 to 12) identified in the text and in Appendix B, the salary of each teacher i in that job category ($S_i$) is explained (Equation (1) in the text) as

$$S_i = A_i\beta_k + \epsilon_{ki}, \tag{A1}$$

where $A_i$ is the vector describing all attributes of teacher i, and $\beta_k$ is the vector of average valuations of each attribute for job category k. Note that if all teachers in job category k do not have a given attribute (e.g., Developmental Teachers will not have a PhD), the

[18] None of these findings were refuted by the DOJ during litigation. The case settled before trial.

[19] In order to minimize technical discussion and given the fact that these models are well developed in the literature, I discuss these issues at a fairly heuristic level while providing appropriate citations.

corresponding attribute and ß weight will be zero.

The job category (ordered by grade) into which a teacher is hired is determined (probabilistically) by that teacher's attributes and the stated job criteria of the school system; or

(A2)
$$\begin{aligned}
&\Pr(\text{teacher i is not hired}) = &&\Pr(A_i\tau + \epsilon_i < c_1)\\
&\Pr(\text{teacher i is hired into category 1}) = &&\Pr(c_1 < A_i\tau + \epsilon_i < c_2)\\
&\Pr(\text{teacher i is hired into category 2}) = &&\Pr(c_2 < A_i\tau + \epsilon_i < c_3)\\
&\qquad\vdots &&\\
&\Pr(\text{teacher i is hired into category k}) = &&\Pr(c_k < A_i\tau + \epsilon_i < c_{k+1})\\
&\qquad\vdots &&\\
&\Pr(\text{teacher i is hired into category 12}) = &&\Pr(c_{12} < A_i\tau + \epsilon_i).
\end{aligned}$$

If $\epsilon_i$ is distributed N(0,1) and if $Z_{ik} = 1$ if individual i is hired into job category k (0 otherwise), then $\Pr(Z_{ik} = 1) = \Phi(c_{k+1}-A_i\tau)-\Phi(c_k-A_i\tau)$, where $c_{13} = \infty$ and $\Phi$ is cumulative standard normal. The likelihood function for a model of the hiring of all teachers is therefore,

(A3) $$L = \pi_i\pi_k\,[\Phi(c_{k+1}-A_i\tau)-\Phi(c_k-A_i\tau)]^{Z_{ik}}.$$[20]

We are interested in the likelihood of observing the joint hiring and compensation decisions of the PSS. To that end, if job selectivity is determined by (A2); if salary is determined by (A1) conditional on job category; and if ($\epsilon_{1i}$, $\epsilon_{2i}$, .. $\epsilon_{12i}$, $\epsilon$) are jointly normal and appropriately normalized (and denoted as $\Sigma$)[21], the likelihood function for the full model is

(A4) $$L(ß_1, ß_2, .. ß_k, .. ß_{12}, \tau, \Sigma) = \pi_i \{\int_{\phi k}^{\phi k+1} (g_k(S_i-A_iß_k, \epsilon))\, d\epsilon\}^{Z_{ik}},$$

where $g_k$ is the bivariate normal density function for ($\epsilon_{ki}$, $\epsilon$) for all k; and $\phi k = c_k-A_i\tau$ for all k.

Estimation of (A4) can be accomplished through maximum likelihood methods and a variety of two-stage methods.[22] However, convergence of the maximum likelihood estimators for this likelihood function is not always assured. Furthermore, the two stage estimators are known to be extremely inefficient. Indeed, monte carlo analysis suggests that OLS with dummy variables for each job category dominates two-stage estimators.[23] For these reasons and the fact that data concerning the hiring decision were not gathered under

[20] See Maddala [1983, pp. 46-48].

[21] See Maddala [1983, Section 8.3].

[22] See Maddala [1983, Section 8.3] and Hartman [1991].

[23] See Hartman [1991].

the DOJ complaint, the OLS/GLS estimation in the text and in Appendix B is appropriate for this analysis.

## Appendix B: Econometric Issues and Selected Results

Based upon standard pooling tests reported in the text, I estimated regression Equation 1 in the text for all teachers within each of the following 12 job categories:

| | |
|---|---|
| Teacher Aide I | Developmental Teacher |
| Teacher Aide II | Classroom Teacher |
| Teacher Aide III | Classroom Teacher I |
| Teacher Aide IV | Classroom Teacher II |
| Teacher Aide/Head Start | Classroom Teacher III |
| | Classroom Teacher IV |
| | Classroom Teacher V |

I also estimated the regression equation in the standard "semi-log" form (see Brown and Rosen [1982], Griliches [1971], Hartman [1987], Hartman and Doane [1987], Ohta and Griliches [1976]), where

$$(1a) \quad \text{Log}(\text{Salary}_i) = b_0 + b_1 A_{i,1} + b_2 A_{i,2} + \ldots + b_{13} A_{i,13} + b_N A_{i,N} + e_i,$$

and the variables are defined in the text. I found the results equivalent and report the semi-log results.

In estimation, I first examined the correlation of all variables to be included in the regression, to assess for the presence of multicollinearity. I included all variables in the regression that did not reveal severe multicollinearity. In the early runs, I included all variables listed in footnote 10, in addition to period dummies for 1989, 1991 and 1993. These variables summarize all teacher attributes hypothesized to effect salary. The time dummies capture across-the-board salary increases unrelated to specific teacher attributes. The time dummies measure changes in real wages and salaries relative to 1987. I eliminated those variables from subsequent regressions that fail standard t and F tests for statistical significance (95% confidence level).

Having estimated the effects of teacher attributes and year effects (1989, 1991, 1993), I added binary variables for the national origin of the teacher. These variables measure the additional explanatory effect (if any) of the national origin of the teacher. The national origin dummies measure national origin effects relative to State-Siders (see Figure 1C in the text). Once I added these binary variables, I perform the same set of t and F tests mentioned above, to eliminate statistically insignificant national origin effects. These regression results are summarized in Table 2.

Finally, in order to assess the national origin effects by year, I interacted the binary dummy variables for national origin with the binary dummy variables for years 1987, 1989, 1991 and 1993 and added them to each final equation in attributes. I then performed the same set of statistical tests (t and F) to eliminate statistically insignificant national origin variables as described above. These results are summarized in Table 3.

While all regression results are available from the author upon request, for brevity

I present in this Appendix (Table B.1) results summarized in Table 3 for selective job categories discussed in the text.

TABLE B.1 SELECTED FINAL REGRESSIONS IN ATTRIBUTES AND NATIONAL ORIGINS DESIGNATIONS BY YEAR

TEACHER AIDE I

Log (SALARY) = 8.7630 + 0.0272 EXPER_1 + 0.0216 BOE_COR - 0.0364 HEAD
(555.57) (11.93) (3.38) (-2.49)

- 0.1192 YEAR89 + 0.2904 YEAR91 + 0.3437 YEAR93
(-5.58) (16.44) (19.47)

N = 229; F = 232.54; $R^2$ = .86

TEACHER AIDE III

Log (SALARY) = 9.0669 + 0.0049 EXPER_1 + 0.0076 YRSSEN - 0.0953 YEAR89
(218.87) (1 94) (2 23) (-1.90)

+ 0.2179 YEAR91 + 0.2308 YEAR93
(4.67) (5 52)

N = 63; F = 25.87; $R^2$ = .69

DEVELOPMENTAL TEACHER

Log (SALARY) = 9.3949 + 0.0496 DEG_AS + 0 0168 EXPER_1 + 0 0046 EXPER_3
(267.91) (2.52) (9.87) (1.79)

+ 0.0217 YRS_ED - 0 0868 YEAR91
(2 48) (-4.61)

N = 133; F = 31 63; $R^2$ = .55

CLASSROOM TEACHER I

Log (SALARY) = 9.2255 + 0.0212 EXPER_1 - 0 0750 YEAR89 + 0.4022 YEAR91
(536.90) (12.96) (-4.10) (11.84)

+ 0.3123 YEAR93 + 0.1642 FILI93
(12.98) (3.18)

N = 118; F = 95.85; $R^2$ = .81

CLASSROOM TEACHER II

Log (SALARY) = 9 4283 + 0 0203 DEG_AS1 + 0.0153 EXPER_1 + 0 0426 HIGH
(641.07) (1.95) (16.33) (3.64)

- 0 0966 YEAR89 + 0.2800 YEAR91 + 0 2866 YEAR93
(-6.14) (14.10) (20.52)

- 0 0630 FILI91 - 0.0980 CNMI91 - 0 0559 OTHE91
(-3 08) (-3.98) (-2.43)

N = 373; F = 98 40; $R^2$ = .71

CLASSROOM TEACHER IV

Log (SALARY) = 9 4315 + 0 0271 DEG_MA1 + 0.0967 DEG_PHD1 + 0 0072 EXPER_1
(64 91) (1.71) (2.56) (1.84)

+ 0 0142 EXPER_2 + 0.0354 EXPER_5 + 0.0114 YRSSEN
(3 40) (1.65) (3.31)

+ 0 0174 AGE - 0.0002 $AGE^2$ + 0.0758 SP_ED
(2.60) (-2.44) (1 88)

+ 0 0201 HIGH - 0.1683 YEAR89 - 0.0977 FILI91
(1.48) (-7 19) (-5.13)

- 0 0588 FILI93 - 0.0957 CNMI87
(-3 40) (-2.61)

N = 203, F = 31 12; $R^2$ = .70

Notes. t statistics for $H_0$: $\beta = 0$ in parentheses.

## References

Anderson, P.A., A. de Palma and J-F. Thisse [1992], *Discrete Choice Theory of Product Differentiation*, MIT Press.

Baldus, D.C. and J.W.L. Cole [1980], *Statistical Proof of Discrimination*, McGraw-Hill.

Brown, J.N. and Rosen, H.S. [1982], "On the Estimation of Structural Hedonic Price Models," *Econometrica*, 50(3).

Epple, D. [1987], "Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products," *Journal of Political Economy*, 95(1).

Griliches, Z., ed. [1971], *Price Indexes and Quality Changes*, Harvard University Press, Cambridge.

Hartman, R.S. [1987], "Product Quality and Market Efficiency: The Effect of Product Recalls on Resale Prices and Firm Valuation," *Review of Economics and Statistics*, 69(2).

Hartman, R.S. [1989], "An Empirical Model of Product Design and Pricing Strategy," *International Journal of Industrial Organization*, 7(4), December.

Hartman, R.S. [1989], "Hedonic Methods for Evaluating Product Design and Pricing Strategies," *Journal of Economics and Business*, 41(3), August.

Hartman, R.S. [1991], "A Monte Carlo Analysis of Alternative Estimators in Models Involving Selectivity," *Journal of Business and Economic Statistics*, 9(1).

Hartman, R.S. and M.J. Doane [1987], "The Use of Hedonic Analysis for Certification and Damage Calculations in Class Action Complaints," *The Journal of Law, Economics and Organization*, 3(2).

Heckman, J.J. [1978], "Dummy Endogenous Variables in a Simultaneous Equation System," *Econometrica*, 46.

Heckman, J.J. [1979], "Sample Selection Bias as a Specification Error," *Econometrica*, 47.

Maddala, G.S. [1983], *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge University Press, Cambridge

Ohta, M. and Z. Griliches [1976], "Automobile Prices Revisited: Extensions of the Hedonic Hypothesis," in N.E. Terleckyj, editor, *Household Production and Consumption*, National Bureau of Economic Research Conference on Research in Income and Wealth: Studies in Income and Wealth, 40.

Spady, R.H., and A.F. Friedlaender [1978], "Hedonic Cost Functions for the Regulated Trucking Industry," *Bell Journal of Economics*, 9(1).

# STATIONARITY OF THE NET DISCOUNT RATE: ADDITIONAL EVIDENCE

by

James E. Payne, Bradley T. Ewing and Michael J. Piette*

## I. Introduction

Within the forensic economics literature a debate has emerged with respect to the use of the net discount rate approach to calculate the present value of future earnings loss. The empirical relationship between interest rates and earnings growth has been examined by researchers (see Hosek, 1982; Harris, 1984; Parks, 1985; Schilling, 1985; Lambrinos, 1985; Carpenter, et.al., 1986; Pelaez, 1989, 1991; Anderson and Roberts, 1989; Nowak, 1989; Lewis, 1991; Benich, 1992; Bonham and La Croix, 1992; Gamber and Sorensen, 1993, 1994; Lynch and Stauffer, 1993; Lawlis and Male, 1994). The task of this empirical note is to examine whether or not the net discount rate is characterized as a stationary process or a nonstationary process. If the net discount rate is stationary around its mean, then the use of historical averages of the net discount rate may be used for forecasting. On the other hand, if the net discount rate is nonstationary around its mean, due to deterministic trend or unit roots, then the use of historical averages serves no useful purpose in the forecasting of the net discount rate.

Section II discusses the net discount rate, data, and unit root tests while Section III provides concluding remarks.

## II. Net Discount Rate and Unit Root Tests

The net discount method, or the offset method, recognizes that earnings growth and the interest rate do not have to be equal, as in the case of the total offset method, but that the relationship between earnings growth and the interest rate is stable. Paralleling the work by Nowak (1991) and Benich (1992), we specify the net discount rate in nominal terms as follows.

(1) Net Discount Rate (Nominal) = $(1 + g)/(1 + i)$

where g denotes nominal earnings growth and i the nominal interest rate. Following the research by Romans and Floss (1992) we use the 3-month U.S. Treasury bill interest rate (taken from the Economic Report of the President. 1996). Earnings growth is examined by sector using the one-digit industry code. These hourly earnings converted to natural logarithms are obtained from Employment and Earnings. For the period 1964-1995, annual earnings growth from nine sectors will be examined: Total Private Sector; Transportation

* James E. Payne, Ph.D. is an Associate Professor with the Department of Economics and Finance at Eastern Kentucky University, Richmond, KY 40475-3176. Bradley T. Ewing, Ph.D. is an Assistant Professor with the Economics Department at Texas Tech. University. Michael J. Piette, Ph.D. is an Economist with Analytical Economics, Inc. At 1535 Killearn Center Blvd., Suite B-1, Tallahassee, FL 32308.

and Public Utilities; Retail Trade; Finance; Insurance and Real Estate; Services; Manufacturing; Mining; Construction; and Wholesale Trade.[1]

In order to differentiate whether or not the net discount rate is stationary or nonstationary, we apply the popular Augmented Dickey-Fuller (ADF) unit root tests (Dickey and Fuller, 1979). Unlike the unit root tests undertaken by Pelaez (1991), Bonham and La Croix (1992), and Gamber and Sorensen (1993), we test several hypotheses concerning the unit root hypothesis. The ADF unit root test is constructed from the ordinary least squares estimation of the following:[2]

$$(2) \qquad \Delta X_t = \alpha + \beta t + \gamma X_{t-1} + \delta_1 \Delta X_{t-1} + \epsilon_t$$

where $\Delta$ is the first difference operator; t is a linear time trend; and $\epsilon_t$ is a covariance stationary random error. The null hypothesis is that $\gamma=0$, making a nonstationary series containing a unit root. The various forms of "random walk" processes have unit roots, are nonstationary, and are said to contain a "stochastic trend" component. The critical values for $\gamma=0$ are based on the $\tau_\tau$ test statistic as reported in Fuller (1976). The $\tau_\tau$ statistic fails to reject the null hypothesis of a unit root for each of the sectors as reported in Table 1.

In addition to tests of the null hypothesis that $\gamma=0$, Dickey and Fuller (1981) provide tests of hypotheses concerning the significance of the drift term $\alpha$ and time trend coefficient $\beta$ in equation (2). To test the significance of the drift parameter $\alpha$, the null hypothesis that $\alpha=0$ given $\gamma=0$ is examined using the $\tau_{\alpha\tau}$ statistic. With the exception of the net discount rate for transportation and public utilities as well as finance, insurance and real estate, Table 1 shows that the $\tau_{\alpha\tau}$ statistic fails to reject the null hypothesis. In the cases of the net discount rates for transportation and public utilities and for finance, insurance and real estate, one can reject the null hypothesis of a unit root with drift at the 10% level of significance. This is consistent with a "random walk with drift" process. To test the significance of the time trend $\beta$, the null hypothesis that $\beta=0$ given $\gamma=0$ is examined using the $\tau_{\beta\tau}$ statistic. The $\tau_{\beta\tau}$ statistic fails to reject the null hypothesis, as reported in Table 1, for each of the sectors.

In addition to the individual and conditional tests of parameters, we test two joint hypotheses on the coefficients. From equation (2) we test the null hypothesis $\alpha=\gamma=\beta=0$ using the $\Phi_2$ statistic. In this case the null hypothesis is that the data are generated by the restricted version of equation (2), with $\alpha=\gamma=\beta=0$, against the alternative hypothesis that the data are generated by equation (2). The second joint hypothesis with respect to equation (2)

[1] As pointed out by Bonham and La Croix (1992), nominal interest rate and nominal earnings growth rates are used since state and local authorities tax nominal interest and labor income. Moreover, forecasting real earnings growth and real interest rates does not eliminate the need to forecast inflation (see footnote 4, p. 222 of Bonham and La Croix, 1992).

[2] Phillips and Perron (1988) use a nonparametric adjustment to the Dickey-Fuller test statistics which allows for weak dependence and heterogeneity in the error term. However, Kim and Schmidt (1990) indicate that the Phillips-Perron tests do not perform well in finite samples. The Augmented Dickey-Fuller tests include one lag.

specifies the null hypothesis $\gamma=\beta=0$ using the $\Phi_3$ statistic. In this case the null hypothesis is that the data are generated by the restricted version of equation (2), with $\gamma=\beta=0$, against the alternative hypothesis that the data are generated by equation (2). Both the $\Phi_2$ and $\Phi_3$ statistics fail to reject the respective null hypotheses. Thus, the restricted versions are not binding. The results suggest that the net discount rate for each of the sectors follows a random walk without drift or deterministic trend, i.e. difference stationary. Thus, as recommended by Bonham and La Croix (1992), the forensic economist may use the random walk rule for forecasting a nonstationary series. The random walk rule uses the current observation of the net discount rate as an estimate of future values of the net discount rate.

## III. Concluding Remarks

The simple approach undertaken in this note of directly testing the stationarity of the net discount rate supports the cointegration analysis by Bonham and La Croix (1992) and Gamber and Sorensen (1993) in examining the stationarity of the long-run trend between interest rates and earnings growth. Given that the net discount rate for each of the nine sectors is not stationary about its mean (DSP), forecasting based on the use of historical averages of the net discount is questionable. These findings suggest that the forensic economist should use caution when applying the net discount rate to earnings growth with respect to projections of lost future earning capacity. Based on the random walk behavior of the net discount rate by sectors, the forensic economist should perhaps use the current value of the net discount rate as an estimate of future values.

Table 1
ADF Unit Root Tests
Net Discount Rate (NDR)

| NDR | $\tau_\tau$ | $\tau_{\alpha\tau}$ | $\tau_{\beta\tau}$ | $\Phi_2$ | $\Phi_3$ |
|---|---|---|---|---|---|
| NDR1 | -2.6806 | 2.6822 | -1.7996 | 0.0349 | 0.0231 |
| NDR2 | -3.1587 | 3.1733[c] | -2.3111 | 0.0487 | 0.0325 |
| NDR3 | -2.1126 | 2.0974 | -1.1815 | 0.0220 | 0.0144 |
| NDR4 | -2.9566 | 2.9565[c] | -0.9060 | 0.0420 | 0.0280 |
| NDR5 | -2.5435 | 2.5326 | -1.5980 | 0.0319 | 0.0207 |
| NDR6 | -2.2960 | 2.3007 | -1.5938 | 0.0258 | 0.0171 |
| NDR7 | -2.2603 | 2.2595 | -1.5355 | 0.0249 | 0.0164 |
| NDR8 | -2.4576 | 2.4428 | -1.7329 | 0.0298 | 0.0194 |
| NDR9 | -2.4416 | 2.4312 | -1.3928 | 0.0292 | 0.0191 |

Critical values defined below are based on a sample size n=25. Significance levels: a (1%), b (5%), and c (10%).

$\tau_\tau$ from Table 8.5.2. of Fuller (1976): -4.38 for 1% level; -3.60 for 5% level; and -3.24 for 10% level.

$\tau_{\alpha\tau}$ from Table II of Dickey and Fuller (1981): 4.50 for 1% level; 3.20 for 5% level; and

2.77 for 10% level.

$\tau_{\beta\tau}$ from Table III of Dickey and Fuller (1981): 3.74 for 1% level; 2.85 for 5% level; and 2.39 for 10% level.

$\Phi_2$ from Table V of Dickey and Fuller (1981): 0.61 for 1% level; 0.89 for 5% level; and 1.10 for 10% level.

$\Phi_3$ from Table VI of Dickey and Fuller (1981): 0.74 for 1% level; 1.08 for 5% level; and 1.33 for 10% level.

Sector definitions for net discount rate:

NDR1 = Total Private Sector hourly earnings
NDR2 = Transportation and Public Utilities hourly earnings
NDR3 = Retail Trade hourly earnings
NDR4 = Finance Insurance and Real Estate hourly earnings
NDR5 = Services hourly earnings
NDR6 = Manufacturing hourly earnings (excluding overtime)
NDR7 = Mining hourly earnings
NDR8 = Construction hourly earnings
NDR9 = Wholesale Trade hourly earnings

## References

Anderson, G.A. and D.L. Roberts (1989), "Stability in the Present Value Assessment of Lost Earnings", *Journal of Risk and Insurance*, March, 50-66.

Benich, J. (1992), "Comment: On the Stability and Predictability of the Interest Rate and Earnings Growth Rate", *Journal of Forensic Economics*, 6(1), 51-52.

Bonham, C. and S.J. La Croix (1992), "Forecasting Earnings Growth and Discount Rates: New Evidence from Time Series Analysis", *Journal of Forensic Economics*, 5(3), 221-231.

Carpenter, M.D., D.R. Lange, D.S. Shannon, and W.T. Stevens (1986), "Methodologies of Valuing Lost Earnings: A Review, A Criticism, and A Recommendation", *Journal of Risk and Insurance*, 53(1), 104-118.

Dickey, D.A. and W.A. Fuller (1979), "Distribution of Estimates for Autoregressive Time Series with Unit Root", *Journal of the American Statistical Association*, 74, 427-431.

Dickey, D.A. and W.A. Fuller (1981), "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root", *Econometrica,* 49(4), 1057-1072.

Fuller, W.A. Introduction to Statistical Time Series, (New York: John Wiley), 373.

Gamber, E.N. and R.L. Sorensen (1993), "On the Testing for the Stability of the Net Discount Rate", *Journal of Forensic Economics*, 7(1), 69-79.

Gamber, E.N. and R.L. Sorensen (1994), "Are Net Discount Rates Stationary? The Implications for Present Value Calculations: Comment", *Journal of Risk and Insurance*.

Harris, W.G. (1984), "Comment", *Journal of Risk and Insurance*, March, 122-126.

Hosek, W.R. (1982), "Problems in the Use of Historical Data in Estimating Economic Loss in Wrongful Death and Injury Cases", *Journal of Risk and Insurance*, June, 300-308.

Kim, K. and P. Schmidt (1990), "Some Evidence on the Accuracy of Phillips-Perron Tests Using Alternative Estimates of Nuisance Parameters", *Economics Letters*, December, 34, 345-350.

Lambrinos, J. (1985), "On the Use of Historical Data in Estimating Economic Losses", *Journal of Risk and Insurance*, September, 464-476.

Lawlis, F. and R. Male (1994), "Methodological Issues: Interest Rate and Wage Growth Forecasting", *Journal of Legal Economics*, Winter, 55-61.

Lewis, W.C. (1991), "On the Relative Stability and Predictability of the Interest Rate and Earnings Growth Rate", *Journal of Forensic Economics*, 5(1), 9-25.

Lynch, L.A. and R.F. Stauffer (1993), "Single Equation Models for Valuing Economic Loss", *Journal of Legal Economics*, 3(2), 53-60.

Nowak, L.S. (1991), "Empirical Evidence on the Relationship Between Earnings Growth and Interest Rates", *Journal of Forensic Economics*, 4(2), 187-202.

Parks, R. (1985), "The Evaluation of Earnings Loss in Alaska Courts: The Implications of Beaulieu and Guinn", *Alaska Law Review*, December, 311-337.

Pelaez, R.F. (1989), "The Total Offset Method", *Journal of Forensic Economics*, 2(2), 45-60.

Pelaez, R.F. (1991), "Valuation of Earnings Using Historical Growth-Discount Rates", *Journal of Forensic Economics*, 5(1), 27-44.

Phillips, P.C.B. and P. Perron (1988), "Testing for a Unit Root in Time Series Regression", *Biometrika*, June, 75, 335-346.

Romans, T.J. and Floss, F.G. (1992), "Four Guidleines for Selecting a Discount Rate", *Journal of Forensic Economics*, 5(3), 265-266.

Schilling, D. (1985), "Estimating the Present Value of Future Income Losses: An Historical Simulation, 1990-1982", *Journal of Risk and Insurance*, March, 100-116.

# THE MEASURE OF ECONOMIC DAMAGES IN INSURANCE SUBROGATION: A CASE STUDY

by

Melvin A. Hughes and Paul E. Radloff*

## I. Introduction

This paper will address the measure of economic losses in insurance subrogation litigation[1] in the context of a case study. In particular, we examine various economic issues in *Travelers Indemnity Company of Illinois v. U.S.A.* No. CV 90-5237-MRP(Ex). In this case the court ruled that the proper measure of property loss to a storm damaged hotel was "depreciated replacement cost,"[2] as opposed to "reproduction cost new."[3] Although an insurer paid $14.1 million under a replacement cost insurance policy, the insurer was only entitled to recover $6.45 million from a liable third party.

After summarizing key facts of the case, we will examine the legal and economic arguments that were advanced regarding property and business interruption damages. Then we will analyze additional economic issues in insurance subrogation from a "Law and Economics" perspective.

## II. Facts of the Case

In 1965 a seaside hotel was constructed in Southern California. By 1986 the property had fallen into a state of disrepair and occupancy rates were low. In December 1986 the property was purchased for $16 million. By January 1988 the new owner had

* Melvin A. Hughes is with Valuation & Consulting, 28687 Triton Street, Hayward, CA 94544. Paul E. Radloff is with the Udinsky Group, Inc., 2941 Telegraph Avenue, Berkely, CA 94705.

[1] *Subrogation* refers to the substitution of one person for another. An insurance company that has paid a claim has the right to sue a negligent third party in place of the insured if the policy contains a *subrogation clause*. The insured is generally barred from suing the third party in order to prevent double recovery for the same damage. According to Shavell (1987), optimal risk bearing by insureds requires that the insured does not recover more than the actual economic loss from the negligent act.

[2] *Depreciated replacement costs* refers to the replacement value of property minus physical depreciation and obsolescence; insurance adjusters estimate the *actual cash value* of property based on its depreciated replacement cost (Rubin, 1991). Similarly, real estate appraisers use depreciated replacement cost to estimate *fair market value* using the *cost approach* to value (Friedman et al, 191).

[3] *Replacement cost new* refers to the cost to replace damaged property with like property of the same functional utility without regard to depreciation (physical wear and tear) and obsolescence.

almost completed renovations. That month, during an unusually high tide, a storm caused severe damage to the hotel, which had been set for a grand reopening under new management. The loss was insured and a claim was submitted for property and business interruption loss.

The insurer hired an independent adjuster and a CPA firm to assist in adjusting the claim. By the middle of 1989, repairs were completed and the hotel was reopened for business. The insurer paid $14.1 million for the loss, consisting of $8.8 million for property damage and $5.3 million for business interruption loss. The insurer became subrogated to the hotel owners in an action against the United States Government, and sought to recover its payments to the insured pursuant to a property and casualty policy. The court found that the damage to the hotel resulted from the negligence of the United States Government and damages were argued by the parties. The court ruled that the defendant's damage calculations were correct, and awarded that amount to the insurer. A summary of the economic damage theories follows.

## III. Insurer's Economic Damages Arguments

Legal counsel for the insurer argued that a subrogee insurance company is *entitled to recover any amount paid to its insured* in good faith based on a reasonable belief that the moneys were due and owing under the policy. The insurer acted in good faith for the following reasons: (1) All bills submitted as evidence of loss had been audited and verified for actual payment by independent experts; (2) The insurer's independent adjuster had closely monitored reconstruction and clean-up costs with the assistance of consultants; and (3) A nationally known CPA firm had reviewed all the bills. The accountants projected a business interruption loss based on the occupancy levels in the market area of the hotel during the period of restoration, rather than using its historical occupancy rates prior to restoration.

Based on the above facts, the insurer asked for the entire $14.1 million that it paid. In support of its arguments, the insurer's counsel cited the following authorities, *inter alia*: *Agricultural Insurance Company vs Smith* (1968) 262 Ca.App.2d 772; *Employers Mutual Liability Insurance Company vs Pacific Indemnity Company* (1959) 156 Ca.App.2d 369; *and, State Farm vs Cooperative of American Physicians, Inc.* (1984) 163 Cal.App.3d 199.

## IV. Defendant's Economic Damages Arguments

Legal counsel for the United States Government argued that the insurer is only entitled to recover the amount of economic loss actually suffered by the hotel owners. In subrogation, the insurer "stands in the shoes" of the insured and *cannot recover payments in excess of the actual economic loss sustained by the insured. The actual economic loss must be proved by the evidence.* The mere fact that the insured paid claims in good faith is not sufficient to prove economic loss to the insured.

### A. Property Loss

In a subrogation action, the insurer succeeds to the legal rights of the insured. Thus, the correct measure of economic damages to the property should be the diminution in

value caused by the acts and omissions of the liable third party.[4]

A key fact in this case was that the policy provided payment for "replacement cost without deduction for depreciation". Reimbursement under such a policy is likely to lead to an economic betterment of the insured because it means that payment will be made to replace old, depreciated property with new property. For example, termite infested boards in the hotel were replaced with new boards. Thus, the insured may "profit" by incurring a loss when covered by this type of property insurance. *The defendant should be obligated to pay the amount necessary to place the insured in the same financial and economic position after a loss as prior to the loss.* The negligent party should not be required to reimburse the insurer for payments which unjustly enrich the insured.

Defendant's economic expert estimated the dollar value of damage to the insured's property at $4.6 million (compared to $8.8 million paid under the policy). Defendant's expert first estimated the pre-storm value of the hotel based on an appraisal of the depreciated improvement value of the hotel. This value was obtained from an appraisal made shortly before the storm, for the purpose of obtaining a loan to finance the renovation of the hotel. Amounts spent on renovating the property between the time of the appraisal and the date of loss were added to the appraised depreciated replacement cost.

Next, the property damage was estimated at 35 percent of the pre-storm value, based on the report of the independent adjuster regarding the extent of the storm's destruction. The final step was to add the cost of required code upgrades, extra expenses, and supervisory fees. The required code upgrades could be viewed as a betterment. However, but for the storm, the insured could have continued to defer code upgrades.

B. Loss of Business Income

Defendant's counsel contended that the insurer had overpaid for the business interruption loss. The loss of net income was projected by the insurer's accountants before the hotel was reopened. Actual operating results were substantially less than the accountant's projections. In addition, the insured's accountant had overestimated the actual continuing expenses during the period of restoration.

Defendant's economic expert estimated the business interruption loss at $2.85 million (compared to $5.3 million paid under the policy). The defense argued that the best possible estimate for loss of income is the subsequent operating experience of the insured after reopening, provided that economic conditions in the market were similar during the loss period and the proxy period. This would allow a determination of the net income that would have been realized had the hotel reopened as scheduled in January, 1988.

According to the business interruption loss coverage of the policy, the continuing expenses incurred by the insured during the restoration period is added to projected loss of net income to obtain the total amount of business interruption loss. This formula can also be stated as lost revenue minus avoided variable cost (See Foster and Trout, 1989), which

[4] BAJI 14.20 D (*California Jury Instructions*, 1986) equates reasonable compensation for damages to property as the lesser of the difference in the fair market value of the property immediately before and after the accident or the cost to repair the property so as to restore the fair market value of the property as it existed immediately before the accident.

yields an amount called "contribution margin" (Garrison 1991). If projections are accurate, such a formula generally measures actual economic loss. The accountant for the insured made his estimates of continuing expenses after only one-half of the period of restoration had elapsed. Thus, the accountant was forced to predict "estimated actual" continuing expenses for the second half of the period of restoration. This estimate turned out to be larger than the actual continuing expenses. Therefore, the accountants overestimated the business interruption loss, and the finder of fact accepted the defendant's calculation of loss, which was base on actual continuing expenses over the loss period.

## V. Economic Issues in Liability and Subrogation

Up to now, we have addressed the *existence* of economic betterment in the context of insurance subrogation, without considering the implications for economic efficiency in society. In this section we will discuss the broader economic implications of the case under study. To set the foundation for this discussion, we will first review selected topics in Law and Economics, from the vast literature that exists. Interested readers are referred to books such as Landes and Posner (1987), or especially Shavell (1987) for a more comprehensive treatment. Also see Stewart (1994) for a review of the moral hazard literature. The present discussion will be focused on how economic theory relates to insurance subrogation situations.

In *California Jury Instructions* (1986) we find a succinct statement of the public policy issue in question with respect to insurance subrogation in an inverse condemnation action. BAJI 11.71 states: "The term `just compensation' means just not only to the owner but just to the public which must pay just compensation. You must be fair and just to all parties."[5] From the perspective of an economist, "fair and just to all parties" can be discussed in terms of welfare economics.

In an "optimal" economic system, society will achieve the greatest possible social welfare. In the neoclassical view of welfare economics, as summarized by Bator (1957), we can achieve Pareto optimality when marginal utility equals the marginal rates of substitution of goods in production and exchange. Supposing we could determine the map of social welfare functions, we could achieve the maximum social welfare: constrained bliss. Unfortunately, the analytics are not so simple; this elegant theory was dealt two devastating blows. First, Arrow (1951) shows that it is impossible to derive the maximum social welfare. Second, Lipsey and Lancaster (1956) show that under the neoclassical assumptions, we cannot reach welfare maximization from another point; we can only reach a "second-best" optimum in our journey towards the first-best optimum. It seems that bliss is unattainable. Therefore, economists can only speak of Pareto optimality, as opposed to some overall level of social welfare.

Efficient liability rules will achieve Pareto optimality. For simplicity, most (but not all) liability models assume linear utility for money, so that we can discuss costs and benefits in terms of dollars as opposed to utils. Thus, our working criteria for efficiency are Pareto optimality and cost minimization, the two criteria suggested by Shavell (1987) and others. See Landes and Posner (1987) for a further defense of the linear utility and risk neutrality

[5] See Trout and Wade (1995) for a further discussion of just compensation.

assumptions.[6]

Judge Learned Hand provides an early analysis of the economics of liability rules. Judge Hand (*United States vs. Carroll Towing, 1947*) suggests that a negligent party should be liable for damages if the "burden" to prevent the loss is less than the extent of loss times the probability that a loss occurs. In algebraic terms, a party is liable if $B < PL$, where B is the burden (cost to prevent the loss), P is the probability of loss, and L is the value of the loss. If the cost to prevent the loss (B) is greater than the expected loss (PL), then the overall cost to society is less if the loss is allowed to occur.

Economists have recast the Learned Hand Rule in terms of marginal analysis. When economic agents are deciding whether to spend an additional dollar on loss prevention, they should do so if $B < PL$. Agents should continue to spend money on loss prevention until $B = PL$ for the last dollar spent. More formally, we should use the notation of calculus to address the marginal benefits and marginal costs in question. See DeSerpa (1985, Chapter 19), or Landes and Posner (1987) for a more complete discussion.

The Learned Hand Rule, per se, does not address the bilateral nature of many accident situations. A "bilateral precaution" is called for when efficiency requires that the injurer and victim both take precautions to prevent an accident (Cooter 1991). For example, pedestrians should take the bilateral precaution of due care when crossing the street. In a property and casualty insurance setting, bilateral precautions would include loss prevention measures such as fire sprinklers, burglar alarms and smoke detectors on the part of the insured.

The existence of insurance can affect people's incentives to take precautions against loss. Consider the problem of moral hazard. With respect to the economics of insurance, "moral hazard here refers to the tendency of insurance protection to alter an individual's motive to prevent loss." (Shavell 1979, p. 541) For example, a person with automobile insurance might not be as careful locking their car to protect against theft, as compared to a person who self-insures against theft.

The literature on moral hazards has neglected two key factors related to insurance and incentives. First, insurance policies may lead to "economic betterment" in the event of loss, as discussed above. Second, the prospects for subrogation in the event of loss can affect incentives. Hughes (1997) develops a classification system of moral hazard, reflecting the effects of subrogation and economic betterment. Moral hazards are classified as first-degree through forth-degree moral hazard. The definitions are:

[6] When we add insurance to liability models, the assumption of risk neutrality becomes problematic conceptually. Risk neutral agents will not purchase insurance unless the premiums are actuarially unfair in favor of the insured (If premiums are actuarially fair, then agents would be indifferent to purchasing insurance). The purchase of insurance is based on diminishing marginal utility beyond some level of wealth (Arrow, 1971). Can we on the one hand derive the benefit of insurance assuming risk aversion, and on the other hand assume away risk aversion? So, we are faced with a common problem in forensic economics: balancing economic validity and adequate simplicity for use in a courtroom setting.

1. First-degree moral hazard is the situation where there are not good subrogation prospects, and there is no economic betterment.

2. Second-degree moral hazard is the situation where there are good subrogation prospects, but there is no economic betterment.

3. Third-degree moral hazard is the situation where are not good subrogation prospects, but there is economic betterment.

4. Fourth-degree moral hazard is the situation where there are both good subrogation prospects and economic betterment.

The present case study falls into the category of fourth-degree moral hazard.

Let us consider how economic betterment affects economic incentives. An insurance policyholder has *no* incentive to spend resources on loss prevention, because if the loss occurs, the insured is better off. Indeed, the policyholder has the incentive to *cause* the loss to occur, either through inaction (e.g. negligence) or action (e.g. arson). Consider home owners on a flood plain who have replacement cost insurance policies. Every few years they get brand new houses, so there is little incentive to move, to build their houses on stilts, or to pay to build levees.

Now consider subrogation. When an insurer has good subrogation prospects, the insurer has less incentive to research the risk of loss from a specific policyholder. Also, they have less incentive to monitor the loss prevention activities of the policyholder. In the event of loss, the insurer has reduced incentives to diligently audit loss claims submitted by the policyholder. *A fortiori*, if there is also a threat of a "bad faith" lawsuit by the insured, incentives for rigorous auditing of claims are reduced. Consider the following statement from a post-trial brief in the case under study: "The motivation and interest of the accounting firm and of TRAVELERS was to minimize such projected [income] losses." The discussion above casts doubt on this assertion. At a minimum, the assertion cannot be considered true, strictly *a priori*.

We saw above that moral hazards reduce an insured's incentive to spend money on loss prevention. When insurance policies provide for "economic betterment" in the event of loss, incentives for prevention are further reduced. When there are good subrogation prospects, the insurer has reduced incentives to research risks *ex ante*, and reduced incentives to audit claims, *ex post*. If damage awards in subrogation lead to economic betterment, then insureds may be spending too little on loss prevention, and insurers may be requiring too small a premium for the risks they have underwritten. In such cases, it could be inefficient to award loss compensation that provides economic betterment. *Ceteris paribus*, "just compensation" should not subsidize lack of loss prevention on the parts of the insured and inadequate premiums charged by the insurer.

The above caveat "*ceteris paribus*" is important. We must also consider loss prevention efforts on the part of the liable third party. Did the third party

spend enough on sea walls, levees, maintenance, and the like? In certain cases, it may be efficient to award betterment, to increase the third party's incentives to engage in loss prevention activities. One can think of this as a "tax" or "penalty" to the third party, to encourage economically efficient behavior.

Notice that there is a paradox, in that when the third party's incentives for precautions are increased, the insurer and insured's incentives are reduced, and vice versa. When an insured can be made better off in the event of an insured loss, incentives for loss prevention are reduced. Conversely, potentially liable third parties also should have the economic incentives to take efficient levels of precautions against loss. Forensic economists could help the courts to understand this paradox, and thereby could take an active role in reducing the cost of accidents to society. For a more rigorous discussion of these incentive issues, see Hughes (1997).

## VI. Conclusions

In light of the court ruling under study, insurers cannot count on recovering all moneys paid out under property and casualty policies in successful subrogation actions. Although this is not an appealed case, and thus may have no standing as a precedent, the economists did persuade the court not to award economic betterment to the subrogated insurance company. So, the defendant may not be liable for betterment that occurs when an insurer pays full reproduction cost to replace depreciated property. Also, when insurers settle claims quickly with incomplete information, they risk less than full recovery in subrogation. Insurers must balance the cost of bad faith lawsuits against the cost of paying out sums which might not be recoverable in subrogation.

An economically efficient legal system should create incentives such that the insured, the insurance company, and the liable third party are motivated to take adequate precautions against losses. However, recall that we have a paradox with respect to the parties' incentives for precautions. So at this stage, the public policy prescriptions are ambiguous. It would seem that each case is unique, and each case requires an individually tailored analysis. A key issue for forensic economists to address may be which party can prevent losses most cost efficiently. Demsetz (1972) suggests that liability fall on the "least cost avoider" of the accident.[7] Brown's (1973) relative negligence model is a bilateral version of the Learned Hand Rule framework which formalizes the least cost avoider notion. Further theoretical research may lead to less ambiguous public policy prescriptions (from an economic perspective). Theoretical research in this area is likely to prove challenging; economic models of insurance markets can quickly

[7] For a discussion of the present case in the context of the least cost avoider, see Hughes (1997).

become intractable.[8]

References

*Agricultural Insurance Company vs Smith* (1968) 262 Cal.App.2d 772.

Arnott, R. J. and J. E. Stiglitz. 1988. "The Basic Analytics of Moral Hazard." *Scand. Journal of Economics*. 90(3):383-413.

Arrow, K. J. 1951. *Social Choices and Individual Values*. Wiley: New York.

Arrow, K.J. 1971. "The Theory of Risk Aversion.' in *Essays in the Theory of Risk Bearing*. Markham: Chicago.

Bator, F. M. 1957. "The Simple Analytics of Welfare Maximization." *American Economic Review*. 47(2):22-59.

Brown, J. P. 1973. "Toward an Economic Theory of Liability." *Journal of Legal Studies*. 2(Jan):11-30.

*California Jury Instructions. Civil.* West Publishing: St. Paul. 1986.

Cooter, R. D. 1991. "Economic Theories of Legal Liability."*Journal of Economic Perspectives*. 5(3):11-30.

Demsetz, H. 1972. "When Does the Rule of Liability Matter?" *Journal of Legal Studies*. 1(Jan):13-28.

DeSerpa, A.C. 1985. *Microeconomic Theory: Issues and Applications.* Allyn and Bacon: Boston.

*Employers Mutual Liability Insurance Company vs Pacific Indemnity Company* (1959) Cal.App.2d 369.

Foster, C. B. and R. R. Trout. 1989. "Computing Losses in Business Interruption Cases." *Journal of Forensic Economics*. 3(1):9-22.

Friedman, J. P. et al. 1991. *Dictionary of Real Estate Terms.* Barron's

[8] Consider the following statement from the literature (Arnott and Stiglitz, 1988): "[U]sing the simplest possible models of the insurance market . . . the indifference curves and feasibility set . . . are not [mathematically well behaved' - indifference curves need not be convex and feasibility sets never are; price and income consumption lines may be discontinuous; and effort is not in general a monotonic or continuous function of the parameters of the insurance policies provided."

Educational Series: Hauppauggee, NY.

Garrison, R. H. 1991. *Managerial Accounting.* Irwin: Boston.

Hughes, M. A. 1997. "Moral Hazard, Subrogation, and Economic Efficiency: Definitions and Issues." Paper presented to the Southern Economic Association, Atlanta, November, 1997.

Landes, W. M. and R. A. Posner. 1987. *The Economic Structure of Tort Law.* Harvard University Press: Cambridge.

Lipsey, R. G. and K. Lancaster. 1956. "The General Theory of Second Best." *Review of Economic Studies.* 24;11-32.

Rubin, H. W. 1991. *Barron's Dictionary of Insurance Terms.* Barron's educational Series: Hauppauggee, NY.

Shavell, S. 1979. "On Moral Hazard and Insurance." *Quarterly Journal of Economics.* 93(4):541-62.

Shavell, S. 1987. *Economic Analysis of Accident Law.* Harvard University Press: Cambridge.

*State Farm vs Cooperative of American Physicians, Inc.* (1984) 163 Cal.App.3d 199.

Stewart, J. 1994. "The Welfare Implications of Moral Hazard and Adverse Selection in Competitive Insurance Markets." *Economic Inquiry.* 32(April):193-208.

Trout, R. R. and W. Wade. 1995. "The Role of Economics in Regulatory Takings Cases." *Litigation Economics Digest.* 1(1):63-74.

*Travelers Indemnity Company of Illinois v U S.A.* No. CV 90-5237-MRP(Ex).

*United States v. Carroll Towing Co* 159 F.2d 169, 2d Cir. 1947.

# DEFENDING AGAINST A DAUBERT CHALLENGE: AN APPLICATION IN PROJECTING THE LOST EARNINGS OF A MINOR CHILD

by

Lawrence M Spizman and John Kane*

## I Introduction

Expert economic testimony can be excluded from evidence by a trial judge if the testimony does not meet the elements of an expert as outlined in Rule 702 of the Federal Rules of Evidence Since the U.S. Supreme Court's *Daubert v Merrell Dow Pharmaceutical (1993) (1995)* ruling, the standard for admissibility of scientific testimony has become more restrictive. The purpose of this article is to demonstrate how an economist might respond to a Daubert challenge. The specific example considered in this article is a response to a hypothetical challenge of expert testimony about the lost earnings of an injured child that relies on the methodology developed by Spizman and Kane (SK) (1992), and replicated by Gill and Foley (GF) (1995).

District courts under Daubert are charged to act as "gatekeepers," ensuring "that any and all scientific testimony or evidence admitted is not only relevant but reliable"[1] This "gatekeeper" option is intended to prevent misleading or wrong testimony from being admitted as evidence The Daubert standard for admissibility of scientific testimony, consists of (but is not limited to) the following five standards.

1. Whether the theory or technique employed is generally accepted in the scientific community;
2. Whether the theory has been subjected to peer review and publication;
3 Whether the theory can be, and has been, tested,
4. Whether the known or potential rate of error is acceptable,[2] and,
5 Whether experts are testifying about matters growing directly out of research or have simply developed opinions expressly for the purpose of testifying [3]

These standards no longer require trial courts to defer to scientific experts for acceptance of expert testimony as has been the case since Frye (1923) The courts under Daubert presumably will be able to distinguish between real and "junk science" by questioning the scientific validity of the expert testimony. An important question for forensic economists is whether the courts, under the Daubert standard, will allow

* Lawrence M. Spizman is a Professor of Economics at the State University of New York at Oswego, Oswego, New York, 13126 John Kane is an Associate Professor of Economics at the State University of New York at Oswego, Oswego, New York, 13126.

[1] Daubert (1993) at 2798-99.

[2] These four standards come from Daubert 1993 at 2796-97

[3] This fifth standard comes from Daubert 1995 at 1316-17.

economists to testify when new economic methodologies are used

Economists traditionally use average income figures to create a "statistical person" for determining what a minor child with no work history would have earned in the absence of the injury. There is no clear consensus among economists concerning the proper methodology for creating this statistical person. In fact, the broad average income data used to create the "statistical person" is not based on direct statistical evidence of the individual child. Under this traditional approach both the plaintiff and defendant would argue that the injured child would have attained a certain level of education. Then they would claim that the child would have earnings commensurate with the average income for that educational level. Under this approach the plaintiff has an incentive to overestimate the educational attainment, while the defendant has an incentive to underestimate educational attainment. A typical example by the plaintiff would be to claim that a particular child will receive an MBA degree even though there is no evidence that anyone from the child's family ever graduated from high school A typical example by the defendant would be to argue that the child of two parents with high school degrees would also acquire a high school degree (despite the fact that many children acquire more education than their parents). The dichotomy between the plaintiff and defendant may lead the court under the *Daubert* standard to reject economic testimony on this important matter

This problem can be addressed by estimating the economic losses of an injured minor child using the procedure developed by SK and refined by GF This methodological approach is an improvement over existing techniques since it relies on information concerning the individual child's characteristics rather than relying solely on general statistical information[4] The SK educational attainment model is useful to plaintiffs, defendants and the courts in estimating the educational attainment of the individual child This technique provides a less subjective method of forecasting educational attainment In some situations the results would be advantageous to the plaintiff while in other situations it would be advantageous to the defendant It depends on what each side originally believed the final educational attainment of the child to be. The neutrality of the SK and GF procedure allows the judicial system to settle discrepancies between plaintiffs and defendants with respect to educational attainment of minor children

The SK procedure uses an ordered probit model to forecast the probabilities of alternative levels of educational attainment for an injured minor child with no work history In this model, it is assumed that this probability is affected by a vector of observable demographic and family background characteristics. Once these estimated probabilities are computed, the expected potential lifetime earnings of a child with these characteristics can be predicted A major contribution of the GF study is that it validates the SK model by testing it with another sample, thus showing that the model and results are robust The GF data set has the advantage of including people who did not graduate from high school (all individuals in the SK study were high school graduates), and considering an expanded set of family characteristics that influence educational attainment Both studies show what most members of the legal profession intuitively believe family background characteristics

[4] Valuing children in wrongful death cases has a more accepted methodology, called the investment approach For a discussion of this approach, see Ireland (1990), and Ward (1989)

**do** contribute to the probability of a child attaining a certain educational level and, hence, have an impact on expected lifetime income

To date, the authors know of no court decisions allowing or disallowing testimony of economic damages to an injured minor child based on the SK model. Thus whether the Daubert standards will allow the admissibility of expert testimony on this matter is for the courts to decide The following sections in this paper will present an argument that the SK methodology meets the Daubert Standards of admissibility of expert testimony.

## II. Application of Daubert to Personal Injury of a Child

Daubert has opened the door for trial courts to become more assertive in reviewing scientific theories and, if necessary, to discard them as being too speculative. Miller and Rein (1997) argue that Daubert should be used by judges in "disposing of cases in which causality is premised on scientific analysis rather than observation" Under Daubert, the judge, rather than the jury, is deciding whether expert testimony is admissible. While the Daubert case involved the "hard sciences" its application has been used in different types of economic testimony.[5] Given the proper facts, the Daubert standard may well be applied in a personal injury of a child case using the SK model. Will this model meet those standards?

Although Daubert appears to be the final word on admissibility of expert testimony two recent district courts rulings point out the ambiguity of Daubert U S. District Court Judge Robert E Jones in *Hall v Baxter Healthcare Corp (1996),* agreed with the "disposal of cases" concept and ruled that attorneys cannot introduce evidence, when such evidence, according to Jones, is not scientifically valid This is important because Judge Jones, under the guidance of Daubert, believes that lower courts can restrict the admissibility of expert witness testimony that is not uniformly acceptable to the scientific community A different position, however, was taken in a joint decision by Judges Jack B Weinstein and Harold Baer Jr when they ruled *In re Breast Implant Cases 1996,* that cases should not be disposed of by the courts because expert testimony standards were not satisfied They ruled that new scientific theories should not be rejected because they lack adequate support, but instead should be given time to develop in order to support plaintiff's claims Judge Weinstein believes that "at times it will be appropriate to delay a decision or provide for intermediate relief while studies go forward" The importance of these two different interpretations of Daubert is reflected by the Supreme Courts grant of certiorari in *General Electric Co. v Joiner, 96-188.* In this case the trial judge dismissed the plaintiff's experts while an appeal court allowed the experts testimony The Supreme Court ruled on December 15, 1997 that appellate courts must ordinarily defer to a trial judge's ruling on admitting expert testimony. While the above cases encompass the "hard sciences" they do provide precedents for future courts to use in other types of expert testimony, such as economics. As cases pertaining to economic damages appear before lower courts, guidance from past decision (even if in the "hard sciences") will help them rule on the admissibility and reliability of

[5] See Ghosh (1997) for a thorough review of all the types of economic cases in which the Daubert standard has been applied Also see Slesinger (1996), and Johnson and Ireland (1996), on how the Daubert decision may pertain to forensic economists

economic testimony especially when that testimony is based on unfamiliar econometric techniques, such as the SK model.

Since economists' testimony may rely on econometric techniques that will be unfamiliar to judges, it is important that forensic economists be able to provide a credible defense to a Daubert challenge. In Section III, an example of such an argument is provided in the case of a Daubert challenge to the methodology developed in the SK model

## III. Economic Loss to a Injured Child and the Daubert Standard

The Daubert standard of scientific acceptability creates the danger that evidence on economic damages of a child may not be admissible at trial. Since damages to a minor child is not as well grounded in the judicial system as the economic damages to an adult, courts may use Daubert as a justification for not allowing expert testimony on this matter

In the absence of the GF and SK models defense attorneys could claim that other methods of estimating losses to a injured minor child are speculative, and, in light of Daubert, inadmissible. Knowing that courts may reject plaintiff's economic testimony may make the defense pretrial negotiation position more intransigent. However, the defense's new position may be counterbalanced if the SK model meets the Daubert standard of acceptability because both sides recognize that the economist may now testify under Daubert The defense will be less willing to risk a trial, while the plaintiff would be more willing to go to trial, thus the chance of settlement is increased.

The first Daubert standard must show that the methodology used has wide acceptance in the economic community. Since Daubert, the courts have allowed testimony by an economist based on valid and commonly accepted statistical and multivariate analysis which have been subjected to peer review.[6] The use of econometric techniques, such as the ordered probit model used by SK and GF, does have wide acceptance in the economic profession

The second Daubert standard recognizes that peer-reviewed research is a relevant criterion for determining the reliability, and hence the admissibility, of evidence [7] The fact that both studies were published in a refereed journal, therefore, becomes important [8] Economist have been dismissed because of their publication records.[9] It is also important to note that the ordered probit model has been used in other more generalized economic journals. Thus even though the SK and GF studies were published in a specialized journal the techniques used have appeared in more general journals.

The third standard of admissibility of scientific testimony developed in Daubert

[6] Flavel (1994)

[7] Bernstein (1994) p 2150.

[8] Courts have rejected the testimony of experts because the studies which their testimony was based on was not peer-reviewed and untestable. Hall v Baxter (1996 at 1406), and Lynch v. Merrell-National Laboratories (1987).

[9] In re Aluminum Phosphide (1995)

stressed the importance of testing in deciding the relevancy of testimony If a model can be tested a judge will be better able to determine the relevancy of evidence as well as its reliability Daubert requires that the reasoning or methodology underlying the testimony be scientifically valid by having a reasonable basis in fact Ghosh (1997) discusses how courts, in applying Daubert to economic testimony, are more inclined to accept such testimony if it was based on empirical results rather than pure theory because the speculative element of pure theory is reduced.

The empirical evidence of SK and GF support the theory that family background characteristics influence educational attainment. The results of the ordered probit model are based on empirical observations The probabilities of reaching alternative levels of educational attainment are calculated using data from a large heterogenous random sample of the national population This provides consistent estimates of the probability that a child with a given set of demographic characteristics will achieve each alternative level of educational attainment. These probabilities will be helpful to the trier of the facts

A judge may rule that the Daubert standard of reliability and validity were not satisfied by the SK study by itself. The judge may rule this way because of unfamiliarity with the ordered probit methodology used in estimating damages to a minor child or concern with the reliability of a study based on a single sample set Courts have rejected testimony, even though the methodology used was correct, because the hypothesis has not been subjected to more rigorous statistical testing[10] or experts did not use the appropriate sample[11] or the data was inadequate [12] The GF study becomes an important criterion for accepting the SK model under Daubert because the results were almost identical despite the use of a totally different data set. The consistency and robustness of the SK model should be compelling statistical evidence that the SK model meets the third Daubert Standard

The fourth standard created by Daubert is more problematic for the social sciences in general and economics specifically Conceptually, the fourth standard is relevant to the hard sciences but does not apply to the social sciences unless the question is asked differently The only way an error rate can be known is if we can prove that the model is specified correctly. It cannot be proved that any model is specified correctly. Economists can only demonstrate statistically that a given model is better specified than any particular alternative. While the fourth Daubert standard is not literally met (actually the five standards are only suggestions for the trial court to follow), it is the best that can reasonably be expected given statistical limitations.

The fifth and final standard created by Daubert attempts to prevent experts from initiating research for the sole purpose of testifying in a specific case The court in City of Tuscaloosa et al (1995) rejected economic testimony in part because the method used by the economist had only been cited and published by other economists involved in the case The SK empirical model was developed out of economic inquiry with no agenda for litigation, thus meeting the fifth Daubert standard of admissibility of testimony. Arguably the SK and

[10] City of Tuscaloosa (1995)

[11] Contractors v City of Philadelphia (1995)

[12] Scales v. George Washington University (1993)

GF studies satisfy the Daubert standards thus allowing the courts to admit evidence based on the use of this methodology. While some trivial differences between the two studies occur, both studies support the methodological approach developed by SK As a scientific model that estimates the economic damages of a injured minor child, the SK approach seems to meet Daubert's standard of scientific acceptability

## VI Conclusion

The Daubert standard has been applied to scientific testimony with respect to the "hard sciences". Some economic testimony has also been scrutinized by the courts under Daubert As forensic economics evolves new methods of projecting economic losses will be developed. These methods will at some point be required to meet the Daubert standard for admissibility of evidence One new method developed in SK uses an ordered probit technique to estimate the economic damages to a injured minor child. That study was replicated with another set of data by GF arguably showing that the Daubert Standard can be met. The SK educational attainment model is useful to the legal profession because of its neutrality For example, the court can now consider lost income for a child for different educational levels All levels of educational attainment can be shown weighing each level by the probability (calculated by the SK model) of the child attaining that level. The jury then can decide which level of education the facts of the case support Since the lost earnings are adjusted by the probability of attaining that level a more precise estimate is attained. These two studies increases the probability of convergence toward settlement between both sides of the dispute The trial court as "gatekeeper" under Daubert and the jury will base their decision's on a neutral method of estimating the probability of a child attaining a certain educational level

## References

Bernstein, David, "The Admissibility of Scientific Evidence After Daubert v Merrell Dow Pharmaceuticals, Inc.", *Cardozo Law Review*, 1994, Vol 15:2139, pp 2139-2181.

City of Tuscaloosa et al v Harcros Chemicals, Inc. et al, 877 F. Supp. 1504 (N Dist. Ala.1995)

Contractors v City of Philadelphia 893 F Supp 419 (E D Penn 1995)

Daubert v Merrell Dow Pharmaceuticals, Inc , 509 U.S. 579, 113 S Ct 2786, L.Ed.2d 469 (1993).

Daubert v Merrell Dow Pharmaceuticals, Inc , 43 F 3d 1311,1316 (9th Cir. 1995)

Flavel v Svedala Industries, 875 F. Supp. 550 (E.D Wis. 1994)

Frye v United States, 293 F 1013 (D C. Cir 1923)

Ghosh, Shubha, "Economics in Law Through an Evidentiary Lens· Appraising the Paradox

of Daubert as Applied to Social Science Testimony" Unpublished paper, Oklahoma City University School of Law, Available from Author, 1997

Gill, Andrew and Foley, Jack, "Predicting Educational Attainment for A Minor Child· some Further Evidence", *Journal of Forensic Economics*, Vol IX, No 2, Spring/Summer 1996, pp. 101-112

Hall v. Baxter Healthcare Corp., 947 F Supp.1387(D Or 1996)

In re Aluminum Phosphide Antitrust Litigation, 893 F Supp 1497 (D Ct. Kan. 1995)

In re Breast Implant Cases, 942 F Supp 958, 961 (E.D.N.Y. & S D N Y. 1996)

Ireland, Thomas, "Loss in the Death of a child. An Extension of the Ward Analysis", *Journal of Forensic Economics,* Vol. 3, No 2, Spring/Summer 1990, pp 73-77.

Lynch v Merrell-National Laboratories, 830 F 2d 1190.1195 (1st Cir 1987)

Johnson, Walter and Ireland, Thomas , "Qualifications and Admissibility. Applying the Daubert Mandate to Economic Testimony", Presented at the Allied Social Science Association Meeting, San Francisco, California, January 1996

Miller, Paul and Rein, Bert, "Science and Courts Still Struggle to Cooperate", *The National Law Journal,* Monday May 12, 1997, p a22.

Scales v. George Washington University, 1993 U.S dist. Lexis 10692 (D.C 1993)

Slesinger, Reuben, "The Demise of Hedonic Damages Claims in Tort Litigation", *Journal Of Legal Economics,* Vol 6, No 2, Fall 1996, pp 17-28.

Spizman, Lawrence and Kane, John, "Loss of Future Income in the Case of Personal Injury of a Child. Parental Influence on a Child's Future Earnings", *Journal of Forensic Economics*, Vol V, No 2, Spring/Summer, 1992, pp 159-168

Ward, John, "Estimating Economic Loss in the Case of A Death of a Child", *Journal of Forensic Economics*, Vol 2, No 3, August 1989, pp 19-29.

Wilt v Bruacker, 191 W. Va 39, 1993, 443 S E 2d 196, 203-207, W Va (1993)

# CHOOSING THE APPROPRIATE SUPPLY OF HOUSEWORK EQUATION

by

Shmuel Sharir*

## I Introduction

The economic value of housework lost by a plaintiff may play a role in wrongful death and injury cases Such a calculation mandates separation of housework hours from leisure hours In the absence of direct personal data, forensic economists may want to use an econometrically-estimated supply of housework hours, along with data on the relevant wage rate [1]

When considering the use of estimated equations of supply of housework hours, forensic economists have to decide which of the available equations, if any, should be chosen. This is what Greenwood (1996) tries to do, but her discussion does not provide general guidelines on how to deal with the issue. The purpose of this paper is to use economic theory to provide guidance regarding the circumstances under which certain specifications of the supply of housework hours are appropriate. As a by-product, it will also provide commentary on Greenwood's specific claims that market-work hours should be included in the equation, while a race variable should not

The next section uses Greenwood[2] as an example of what the issue is; it presents the estimated supply of housework which she discusses, summarizes her criticisms of it and highlights their weaknesses. The third section uses economic theory to determine the roles of (economic) variables such as wages, hours and incomes in the equation The roles of other (demographic) variables such as race are discussed in the fourth section The final section contains concluding remarks

## II. An Example. An Equation and Greenwood's Criticism

Greenwood (1996) focuses on the following equation estimated by Bryant, Zick and Kim (1993) for married women using the Cornell sample

* Shmuel Sharir is a Professor of Economics at the University of Alberta, Edmonton, Alberta, T6G 2H4, Canada and Visiting Fellow in Economics, School of Economics and Financial Studies, Macquarie University, Sydney, NSW 2190, Australia

[1] See Romans and Floss (1989) for a discussion about the relevant wage for tort liability cases within the economic model of the family which is used in this paper

[2] See Greenwood (1996)

N = 3770 59 + .8676 UNEARN + 521.69 UND3 - 94 2772 AGEF + 1.23 AGESQ
(911 55) (5 26) (98.17) (42.56) (0.48)

-46.93 EDF + 199.83 FAMGT3 - 235 87 OWN - 7.22 URBAN - 651.78 BLACK
(20.29) (40.94) (187 82) (99 76) (265.78)

where standard errors are in parentheses, and the variable definitions are.

| | |
|---|---|
| N | amount of housework by a married woman (in minutes per week) |
| UNEARN | after tax annual nonlabor family income (in thousands of dollars) |
| UND3 | number of children up to three years old in the family |
| AGEF | age of the married woman (in years) |
| AGESQ | AGEF squared |
| EDF | educational attainment of the married woman (in years) |
| FAMGT3 | number of family members older than 3 years |
| OWN | dummy variable equaling 1 if family owns home |
| URBAN | dummy variable equaling 1 if lives in city with 50,000 people or more |
| BLACK | dummy variable equaling 1 if married woman is black |

Greenwood counsels forensic economists against using this equation. She claims that "The most glaring deficiency is that the reported equation . contains no variable measuring time spent in the labor force! The economic theory of the household tells us that hours worked in the labor market is a critical determinant of hours spent in household production .". However, Greenwood does not provide the model that will prove her right. As well, Greenwood faults the estimated regression for the inclusion of a race dummy variable Although acknowledging that the variable is statistically significant, she discounts this result on two accounts: "First, there were a very small number of blacks in the sample – only 14 in 1981. Second, the use of a readily identifiable characteristic which has no theoretical basis is always highly questionable " However, Greenwood neither explains why the exact number of blacks is an issue, nor how many blacks would have made the result acceptable to her. As well, her only point regarding the second issue is to mock it by raising the possibility of including a dummy variable of the height of the woman Without providing any rationale, she takes it as self-evident that height is an irrelevant variable in the present context

## III The Role of Market-Work Hours, Wages and Income

### A. The Traditional Model. Individuals are Wage Takers

Consider a traditional (two-person) family model of maximizing utility

$$1) \qquad U(L_1, L_2, N_1, N_2, X ; V^0)$$

subject to the, say, weekly (real) income and time constraints

$$2) \quad X = w_1^0 H_1 + w_2^0 H_2 + A^0$$

$$3) \quad L_i + H_i + N_i = 168 \qquad i = 1,2$$

$L_i$, $H_i$, $N_i$ are, respectively, the weekly hours of leisure, market work and housework; 1 and 2 identify the person; X denotes (the family) weekly real income or expenditures on market goods;[3] A denotes the weekly real nonlabor income and $w_i$ is the real hourly wage of the i-th person; V stands for a vector of "tastes" and "productivity-at-home" variables. The superscript 0 implies that the variable's value is pre-determined.

The endogenous variables determined within the model are, $L_1$, $H_1$, $N_1$, $L_2$, $H_2$, $N_2$, and X. The values of these variables are affected by the values of the exogenous variables of the model, namely, $w_1$, $w_2$, A and V. From the first order conditions and the comparative statics analysis of the model, one can deduce that the supply of housework hours of the i-th person is

$$4) \quad N_i = f_i(w_1, w_2, A, V) \qquad i = 1,2.$$

The traditional model assumes that people face a given wage rate, but are free to choose how to allocate their time among market work, housework and leisure. For most people this model is probably quite appropriate since they may adjust their market-work hours by working overtime, switching to jobs with more acceptable hours, moonlighting, be absent from the job (with or without pay), leave the labor market temporarily, becoming self-employed, etc. The wage rate is usually determined by the market, i.e. supply and demand for labor, or by collective agreements, minimum wage legislation, professional association price lists and social customs. Thus, the analysis suggests that in most cases the forensic economist is likely to conclude that the present model is applicable, and that housework hours of a plaintiff should be estimated on the basis of an equation like (4)

The main implication of the analysis is that in most cases forensic economists should prefer using an estimated equation which includes among its arguments the (actual or last) real wage rates of **both** (or **all**) family members,[4] as well as the real nonlabor income of the family Estimated equations of housework hours of married women which include such variables are reported, for example, by Cochrane and Logan (1975) and Gronau (1977)

[3] Saving can be viewed as a market good within this one-period model Its utility equals the present value of the utilities of future purchases of consumer goods financed by the savings.

[4] If the family is made up of only one (adult) member, only one wage rate should appear in the equation, and it can be shown that if a family member is not a market participant, her/his wage should not appear in the equation. Note that in such cases, forensic economists should prefer an estimated equation which uses relevant data. For example, in the former case, it may be data of single people, and if the data include single and married people, it can be shown that in order to be useful in the present context, the estimated equation should include among its arguments one wage variable as well as an interaction variable of the wage and of the marital status dummy variable

and for married women and married men by Kooreman and Kapteyn (1987) Kiker and Heath (1987), who also report such equations for married women and men, use the appropriate estimation technique for seemingly unrelated regressions, and report them according to whether there are one or two earners in the family, moreover, they report an equation for females who head households. Note that even these equations do not include as variables wages of family members other than the husband and wife, but this shortcoming may be unimportant in most cases dealt with by forensic economists

The Cornell equation for married women reported in the previous section does include among its variables the family's nonlabor income (UNEARN) It also includes age (AGEF and AGEFSQ), and education (EDF) variables of the married woman, according to the human capital approach they can be viewed as representing the woman's wage rate. However, the equation does not include (variables representing) the husband's wage rate. This shortcoming is likely to render the use of the equation by forensic economists inappropriate.

The conclusion that the Cornell equation for married women should not be used by forensic economists, is the same as Greenwood's, but the rationales are totally different. As a matter of fact, her main objection to the equation is actually determined by the current analysis to be unfounded; contrary to Greenwood's claim, the traditional economic model of the family suggests that market-work hours should **not** appear in this supply equation.[5] The explanation for this result is that within the traditional model, housework hours (N), market-work hours (H) and leisure hours (L) of both (or all) family members, and consumption of market goods (X), are determined at the same time and by the same exogenous variables of the model, i.e. the arguments of equation (4) [6]

It should be emphasized that the "family" in the model is an **economic** decision-making unit. As a result, the conclusions above are applicable to traditional as well as non-traditional families, i e. common-law families, or "families" of the same sex Thus, in the absence of estimated equations for non-traditional families, forensic economists may well use equations which are based on data for traditional families

B. A Modified Model: Individuals Are Also Hours Takers

There may be cases in which the forensic economist concludes that the individual is not only a wage taker but also an hours taker in the labor market· A working individual

[5] This argument suggests that even the equation estimated by Gramm (1974) is not useful The equation uses the wife's and husband's **earnings** instead of their wage rates As a result, the equation can be viewed as including the wage **and** market-work hours of each of them when their separate effects are constrained to be identical

[6] The equations solving for the endogenous variables of the model ($N_i$, $H_i$, $L_i$ and X) are the reduced form, not a simultaneous equations system This is another way of proving that Greenwood's claim that market-work hours (H) should appear in the supply of housework hours (N) is wrong when the present model is appropriate

may not have a choice regarding the amount of market work, due to a standard work period, mandatory overtime (at management's discretion), regulations regarding hours of work and the large depreciation of earnings capabilities resulting from a temporary withdrawal from the labor market Assuming that all the family members are (also) hours takers, one can write $H_i = H_i^0$ (i = 1,2). These additional constraints on the behavior of the family are incorporated into the model by substituting them into equations (2) and (3) One gets,

$$2') \qquad X = w_1^0 H_1^0 + w_2^0 H_2^0 + A^0$$
$$3') \qquad L_i + N_i = T_i^0 - H_i^0 \qquad I = 1,2$$

It is clear from equation (2') that within this version of the model X is no longer an endogenous variable. The first order conditions and the comparative statics analysis of this version of the model – where equation (1) [with a constant X] is maximized subject to equations (2') and (3') – imply that the supply of housework hours is now,

$$4') \qquad N_i = g_i(w_1 H_1 + w_2 H_2 + A, V) \qquad I = 1,2$$

When circumstances of a case lead the forensic economist to believe that the present version of the model is applicable, equation (4') becomes the appropriate supply function of housework hours of the plaintiff In this case forensic economists should prefer an estimated supply of housework hours which includes among its arguments the real family (total) income, instead of separate variables of all the real wages, market-work hours and nonlabor incomes of the family members. It should be noted that within the present (non-traditional) version of the model, Greenwood's (1993) claim regarding the importance of market-work hours as a determinant of housework hours becomes relevant. However, they play a role within an income variable, not as separate variables as Greenwood suggested; moreover, market-work hours of **all** family members, not only the wife's, now become relevant.

It is easy to see that if some family members can choose their market-work hours and others cannot, the supply of housework hours of a plaintiff will be a hybrid of equations (4) and (4') In addition to the family's real nonlabor income, the equation includes among its arguments the real **wage** rates of the former and the sum of real **earnings** of the latter.

C. Other Modified Versions of the Model

Two other modified versions of the model can be formulated: In one, people are hours takers but not wage takers, while in the other they are neither wage nor hours takers. The likelihood that forensic economists will conclude that one of these versions is appropriate is small, therefore, they will not be carefully analyzed However, the case where people are hours, but not wage takers, is of particular interest and further comments are warranted Since market-work hours are exogenous variables and wage rates are endogenous variables in this case, the former replace the latter as arguments of the supply of housework hours. And it can be shown from the first order conditions and comparative statics of this version of the model, that under certain conditions Greenwood's suggestion that only the wife's market-work hours should be introduced into the her housework supply equation is correct, the conditions are that the wife and other family members are not

substitutes in housework activities,[7] and that the age and education variables are included in the equation as, say, tastes variables rather than to represent her own wage effect

## IV The Role of Race and Other (Demographic) Variables

Up to this point, the analysis has not dealt with the exogenous variable V The introduction of this variable to the model is an admission on the part of economists that variables other than wage rates and nonlabor income may have important effects on the family's behavior As was already suggested, variable V stands for variables which reflect the role of subjective "tastes" in choosing allocations of time and goods consumption of family members, and for variables which reflect family members' (relative) productivity in housework activities

The economic theory of family behavior is not yet developed enough to provide more than few clues regarding what "other variables" should appear in the supply of housework hours, or what their predicted effects are This is due a certain degree of ignorance on the part of economists about how "productivity in housework" and mainly "tastes" are developed and changed Thus, when it comes to choosing tastes and housework-productivity variables, economists may also have to rely on production theory, and on theories in other sciences, like sociology and psychology

Since economic theory and evidence seem to suggest that educational attainment affects productivity of market work, it is reasonable to expect it to affect productivity of some housework activities too, and the idea that the existence of young children raises productivity of housework time seems logical too.[8] Sociological theory and evidence regarding the importance of reference groups in determining values and norms of families seem to provide a basis for expecting that the level of education, race, place of residence etc affect "tastes" Thus, there are good theoretical reasons for including these (and other) variables in an equation estimating the supply of housework hours

Notwithstanding the above discussion, economists are still not able to determine when a particular demographic variable **must** be introduced into the equation and when it should not, nor can they always predict what its effect will be when it is introduced From this point of view, there is a difference in our understanding of the role of economic variables, like wage rates and nonlabor income, and the role of demographic variables

Contrary to Greenwood's (1993) claim, the above discussion suggests that there are sound **theoretical** reasons for including race in the housework equation. Moreover, since the discussion applies to race as well as other demographic variables, such as place of residence and its size, one wonders why Greenwood objects to the use of the former but not the latter.

Greenwood's (1993) attempt to show the absurdity of the race issue by referring

[7] In this case, the family utility function is said to be separable in housework hours. In the case of a two-person family it can be written as
$U = k(L_1, L_2, X; V^0) + j(N_1, V^0) + l(N_2, V^0)$

[8] For an economic model relating education, the existence of children and housework see Becker (1985)

to a hypothetical finding that tall housewives spend less time on housework than short ones, deserves a response As a matter of fact, there may well be reasonable explanations to such a result within the model For example, it may reflect a higher productivity of tall women in housework, because, say, their reach is higher As a result, a given amount of home production may require less of the tall woman's time.

Turning to statistical issues, Greenwood (1996) realizes that the race variable is statistically significant in the Cornell equation, but she is not impressed because she thinks that there are too few blacks in the sample. Such a view is unjustified. Other things being equal, a decline in the number of blacks in a sample decreases the variance of the race dummy variable This decreased variance appears in the denominator of the formula of the standard error of the estimated regression coefficient of race and increases it, as a result, the t-value and statistical significance of the race variable decrease As well, other things being equal, a decline in the number of blacks in a sample increases the degree of collinearity between the race variable and the constant of the regression, this raises the degree of multicollinearity in the sample, which, in turn, also leads to a smaller t-value and lower statistical significance of the race variable.[9] Thus, the test of significance of the race variable takes into account the small number of blacks in the sample If in spite of a small number of blacks in the sample the race variable is still statistically significant, one is bound by such a result [10]

There is another statistical argument for including the race variable in the housework equation It was argued above that the age and education variables represent the wage in that equation It is not unreasonable to add race to the list of variables representing the wage, as Kooreman and Kapteyn (1987) do.

A finding that black women spend less time on housework than white women with the same other characteristics may not be to Greenwood's liking, but it may have a reasonable explanation within the model For example, if black women are less exposed than black men to labor-market discrimination, it will be optimal for the family that they will work more in the market and less at home Other family members may pick up the slack in housework, and/or less housework will be done within a black family. If the above interpretation is correct, a race dummy variable will have a positive effect on the supply equation of women's market-work hours And while a black-woman plaintiff will be compensated less for her housework than a white woman with identical other characteristics, she will be compensated more for her market work.

The implication of the above discussion for forensic economists is that the preferred estimated supply of housework hours should include race and other demographic and socioeconomic variables However, the evidence for every variable should be examined carefully Various interpretations, with possible opposing implications for plaintiffs, may be consistent with, or explain, the same fact It is up to forensic economists to look for

[9] The former argument is a statement about a diagonal element of the variance-covariance matrix, while the latter is about an off-diagonal element of it Both affect the standard error of the estimated coefficient of the race variable

[10] Discussions on the statistical issue with my colleagues Adolf Buse and Suart London proved useful.

supporting evidence, explain to the courts what the most logical interpretation of the evidence is, and what the implications for the damage award are.

## V. Conclusions

Economic theory suggests that the specification of the supply of housework hours depends on the circumstances of the case facing the forensic economist. Since most people are wage takers, but have control of their market-work hours, the traditional time allocation model suggests that equation (4) is appropriate; forensic economists should use in this case an estimated equation of supply of housework hours which includes among its explanatory variables real wages but not market-work hours of each family member. In other circumstances, the specification of the supply of housework hours will be different, e g equation (4') may become relevant.

It should be emphasized that the above conclusions are based on a traditional **static** model. A model which has dynamic elements, with multi-period horizon and a certain sequence of the timing of decisions regarding the allocation of time may yield different specifications of the supply of housework hours than those discussed in the paper Such a framework seems to be used by Hersch and Stratton (1994) to estimate supplies of housework hours of married women and married males in which both their market-work hours and the family's total labor income are among the variables However, in the absence of a formal model, it is unclear whether their specification is correct. And if it is, they should have used a simultaneous equations estimation procedure, since both housework and market-work hours are endogenous variables, but there is no indication that they do so. Due to these shortcomings, it is doubtful that forensic economists can use their equations. The above comments suggest that future research effort should be devoted to estimating a supply of housework hours within a dynamic model.

The use of demographic and socioeconomic variables such as race, religion, and even height, as proxies for tastes and productivity at home, is well grounded in economic and sociological theories However, forensic economists should be careful when they use and interpret empirical evidence regarding such variables, since our understanding of the role that each of these variables play within the model leaves much to be desired.

## References

Becker, Gary S "Human Capital, Effort, and the Sexual Division of Labor," *Journal of Labor Economics,* 3(1, part 2) January 1985, 533-58

Bryant W Keith, Cathleen D Zick, and Hoshin Kim. *The Dollar Value of Household Work,* Revised edition Ithaca New York State College of Human Ecology, Cornell University, 1993

Cochrane, Susan H and Sandra P Logan. "The Demand for Wife's Nonmarket Time: A Comparison of Rules From Surveys of Chicago School Teachers and South Carolina College Graduates," *Southern Economic Journal,* 42(2), October 1975, 285-93

Gramm, Wendy Lee "The Demand for the Wife's Non-Market Time," *Southern Economic Journal*, 41(1), July 1974, 124-33.

Greenwood, Daphne T. "Estimating Hours of Lost Household Production Using Time Use Data: A Caution," *Litigation Economics Digest*, 2(1), Fall 1996, 89-91

Gronau, Reuben "Leisure, Home Production, and Work – the Theory of Allocation of Time Revisited," *Journal of Political Economy*, 85(6), December 1977, 1099-1123.

Hersch Joni and Leslie S Stratton "Housework, Wages, and the Division of Housework Time for Employed Spouses," *American Economic Review, Papers and Proceedings*, 84(2), May 1994, 120-6.

Kiker, B F and Julia A Heath "Evaluating Household Services Additional Evidence," *Journal of Forensic Economics*, 1(1), September 1987, 19-26.

Kooreman, Peter and Arie Kapteyn. "A Asaggregated Analysis of the Allocation of Time Within the Household," *Journal of Political Economy*, 95(2), April 1987, 223-49

Romans, J. Thomas and Frederick G Floss "Evaluation of Homemaker services. Replacement Cost, Opportunity Cost or Something Else?," *Journal of Forensic Economics*, 2(2), April 1989, 61-5

# USING REVEALED PREFERENCE TO EVALUATE THE RESERVATION WAGE AND THE VALUE OF LEISURE

by

Phyllis W. Isley and Robert Rosenman*

## I. Introduction

One element of economic damages, often considered in cases involving an injury, is the value of the lost pleasure of life. This is intimately connected to the value of leisure. Therefore, it is reasonable to ask what estimate of the value of leisure should be used to calculate the lost value of leisure. This note explores how the reservation wage affects the estimation of the value of leisure. We draw two conclusions from this analysis. First, we conclude that the market wage at which work is accepted or at which a well documented amount of time is devoted to market production is the lower bound of the value of leisure and is, therefore, a conservative estimate of the value of leisure. We draw this conclusion because the reservation wage represents an individual's assessment of the wage which is consistent with the individual's job market qualifications. It can be viewed as a minimum rate of return for work which the individual deems to be the compensation necessary before a job is acceptable. Our second conclusion is that an expert needs compelling reasons to deviate from the value of leisure revealed by the pre-injury allocation of time.

A priori, the reservation wage at which a job is considered acceptable is known only to the individual. However, after the search process and after a job is accepted, the individual is assumed to have revealed the reservation wage. This means that the norm is to assume that the reservation wage is at most equal to the market wage at which a job is accepted. In a similar manner, the value of household production must be greater than the wage rate earned by the individual. Otherwise, more market work would supplant the household production, and others hired to do the household work.

This note assumes that both the reservation wage and the value of leisure are unknown prices in a world of three activities. Members of a household have one basic endowment - time - which enables the consumption of three types of goods:

1. purchased goods and services funded by money income from the production of goods and services for the market;
2. goods and services which are produced by the household for consumption by the household; and,
3. leisure.

The pre-injury allocation of time between the alternatives reveals the household's

* Phyllis Isley is an Adjunct Professor of Law and Economics at the Vermont Law School, South Royalton, Vermont. Robert Rosenman is a Professor of Economics at Washington State, Pullman, Washington.

consumption preferences.[1] This allocation represents the utility maximizing combination of the three goods[2] given the budget constraint:[3]

$$1. \qquad w_r * T = w_o * W + w_h * H + w_l * L.$$

Here T is the total amount of time per day or per year which may be allocated to the three production activities. The amount of time allocated to work for money income is W. The amount of time allocated to household production is H. The amount of time allocated to leisure is L. The values for $w_o$ and $w_h$ are market determined. The values for $w_r$ (the reservation wage) and $w_l$ (the value of leisure) are not directly observable. Without additional information about $w_r$, one cannot determine the value of leisure. Or alternatively, without additional information about the value of leisure time $w_l$, one cannot determine the reservation wage.

One possible way to determine $w_r$ and $w_l$ is to examine all of the combinations of goods which the household could purchase given the household's endowment of time and various rates of return for $w_r$ and $w_l$, where $w_o$ and $w_h$ are known. This amounts to using the budget frontier with different relative values to approximate the household's indifference curve. To illustrate the determination of $w_r$ and $w_l$, a numerical example based on an actual case will be used.[4] The facts of the case are representative of the type of data commonly

[1] The model briefly described here has been more fully developed by Rosenman and Fort (1995) and Isley and Rosenman (1997).

[2] The household's utility function is $U = f(q1, q2, q3)$ where q1, q2 and q3 represent the quantity of market goods, household production and leisure respectively. Along any indifference curve, the total differential $dU = f_1dq1 + f_2dq2 + f_3dq3$ is equal to zero. It is assumed that $f_1>0$, $f_2>0$ and $f_3>0$.

[3] The members of the household have one basic endowment, time denoted T. The total amount of time available can be allocated to hours devoted to producing three good or services: hours devoted to market work W; hours devoted to home production H; and, hours devoted to leisure production L. Each of these production activities has a rate of return such as the market wage $w_o$ or a value such as household production $w_h$. Hence, the household's total income in each period is the sum of the value generated by each activity.

[4] The estimation of this numerical technique is based on the assumption that the household is at equilibrium. The importance of this condition is elaborated in "Measuring Consumer Surplus with Unknown Hicksian Demand" by Irvine and Sims (1998). The importance of this equilibrium condition is also related to a question posed by one referee. The referee asked if a forensic economist, charging the going rate for her services, was not so fully employed that she was willing to take on one more hour of work would this not imply that the market wage exceeded the opportunity of leisure. We assert that the answer to this is no and goes to the very heart of why one must consider the 'unknown' reservation wage when determining the value of leisure.

Suppose that the forensic economist believe that her skills are such that she can charge the

available to the forensic economist.

## II. Facts of the Case

The household in question consisted of two persons. The injured member of the household, IM, was 49 at the time of the accident. She was employed part-time as a school bus driver. At the time of the accident her husband, hereafter the spouse, was age 65 and fully retired.

IM had quit her job as an assembly technician in a manufacturing firm in 1987 so that she and her spouse could return to Vermont. This was precipitated by her spouse's heart condition. With the exception of occasional work filling in as a temporary bus driver, the spouse's income was all non-wage income from a pension and social security.

Together the couple developed the site for their home and maintained a large lawn and garden. They also spent many hours sailing, snowmobiling, traveling and dancing.

IM took the part-time job as a school driver because the schedule permitted her to be home for breakfast with her spouse. She was able to return home for lunch, shopping and other household tasks before she went back to work in the afternoon. IM worked 35 weeks per year. Her regular schedule required 20 hours of work per week and she was paid $6.00 per hour. IM also averaged 18 additional hours per week driving for field trips and sporting events.[5] For the extra hours she was paid $7.00 per hour, and like the other hours of work, IM found that the schedule fit the household's needs. If necessary she could prepare the evening meal before leaving and on many occasions her spouse would travel with her to sporting events. The job had the further advantage that IM's summers were entirely free and school holidays for IM tended to coincide with school holidays for the couple's grandchildren.

going market rate per hour for her forensic services and that this rate is equal to or exceeds her true reservation wage. However, potential clients may not believe that she has sufficient experience as a forensic economist and she finds herself under-employed. To get to equilibrium she will either adjust her market wage down, the case where the market wage is greater than the reservation wage, or she will accept work which pays her true reservation wage, the case where the reservation wage is equal to the market wage. In either case, the value of leisure is still approximately the market wage at equilibrium.

[5] A referee questioned the extent to which the allocation of time worked, especially the allocation of hours to overtime, could be considered as a continuous choice. We recognize that there is not perfect continuity in the choice about how many hours are work. However, we would argue that for an increasing portion of work force there is an increasing amount of control over the choices about the allocation of time. Some examples of the way in which labor markets have adjusted to accommodate workers desire for a greater correspondence between the hours they are willing to offer and the hours which employers accept include: 1) job sharing program; 2) the increasing portion of all types of workers who supply services through temporary agencies; 3) the increasing portion of the labor force which is self employed; and, 4) the type of partial commitment to work as observed in this case for those who are retired or semi-retired.

Pre-injury IM and her spouse each distributed 16 hours per day between each of the three goods: work, home production and leisure. IM devoted 1,330 hours to work per year at an average wage rate of $6.71. She also devoted an estimated 670 hours per year to home production and 3,840 hours per year to leisure.

Based on the facts of the case, one may pose several questions about IM's possible reservation wage. First, did IM's acceptance of the job at $6.00 per hour mean that IM's reservation wage was $6.00 per hour? Alternatively, based on other characteristics of the job such as the convenient schedule, does the acceptance of this job at $6.00 per hour mean that IM accepted a position where the wage represents a discount of her actual reservation wage and lacking certain characteristics, would IM reveal that her true reservation wage was higher than $6.00 per hour? Finally, one might ask was IM's reservation wage really $6.71 per hour instead of $6.00 per hour?

It is true that any wage accepted by IM probably represented a discount from her real reservation wage. IM had left a high paying manufacturing job to return to a rural area which offered few comparable opportunities. Additionally, because of her spouse's retirement and because of his health, the relative time demands of a position were a very important part of her preferences.

Using the facts of the case and the options available to IM, i.e. the option to work only the required 20 hours for $6.00 per hour or work additional hours for $7.00 per hour, it is possible to illustrate that both the reservation wage and the value of leisure may be simultaneously determined, but the solution is dependent on the distribution of time spent on each activity.

## III. Analysis: Estimating the Value of Leisure and the Reservation Wage

From the facts of the case, Equation 1 can be written as:

$$2(a).\quad w_r * 5840 = \$6.71 * 1330 + \$9.11 * 670 + w_l * 3840.$$

This represents IM's observed or revealed preference for each activity.

However, suppose that IM had never worked any overtime. Then she would have had an additional 630 hours to allocate to household production H, or leisure L, or both. If one assumes that all of the additional hours would have been allocated to leisure, Equation 1 can be written as:

$$2(b).\quad w_r * 5840 = \$6.00 * 700 + \$9.11 * 670 + w_l * 4470.$$

The two equations can be solved for the value of $w_r$ and the value of $w_l$. In this case, the value of both $w_l$ and $w_r$ is $7.50. The value of $w_r$ and $w_l$ represents values of the reservation wage and leisure which were clearly within IM's set of choices.

On the other extreme, it is possible that IM could have devoted nearly all 630 hours

to home production.[6] In that case equation 2(b) would be shown as equation 2(c) where *h* stands for the number of hours, up to 630, which could be re-allocated from work to home production or leisure production:

$$2(c). \quad w_r * 5840 = \$6.00 * 700 + \$9.11 * (670 + h) + w_l * (3840 + (630 - h)).$$

Using equations 2(a) and 2(c) to solve for $w_l$, $w_l = (\$4724.3 - (\$9.11 * h))/ (630 - h)$, which is negative if *h* exceeds 518.58. While such values are feasible and within IM's choice set, we believe that under her pre-injury preferences (as reveled by the actual distribution of time), such a distribution of time would never have been selected by IM.

Alternatively, IM might have divided the additional 630 hours evenly between household production and leisure. In that case equation 2(b) would be shown as equation 2(d):

$$2(d). \quad w_r * 5840 = \$6.00 * 700 + \$9.11 * 985 + w_l * 4155.$$

Using equation 2(a) and 2(d) to solve for $w_r$ and $w_l$, $w_r$ and $w_l$ are \$6.45 and \$5.89 respectively. Again, these values for $w_r$ and $w_l$ were clearly within IM's set of choices.

## IV. Findings

The purpose of establishing a reservation wage is to increase the searcher's efficiency in locating a job which provides compensation consistent with the searcher's human capital. If the reservation wage is set too high relative to the actual market wage for those jobs which the searcher has the requisite skills, then the individual will find no job alternatives. Similarly, if the reservation wage is too low relative to the actual expected market wage, few jobs are eliminated from the search. This reduces the efficiency of the job search. Hence, in order maximize the chance of a desirable position when searching for a job, the individual is likely to choose a reservation wage which is marginally lower, but not significantly lower, than the actual expected market wage for jobs within the individual's range of skills.

The solution in which $w_r$ tends to a negative value based on Equation 2(a) and 2(c) means that IM would accept any job with a positive wage. Such a reservation wage may indicate that only the non-wage characteristics of a job were relevant to the searcher. In IM's case, this may in fact have been true. However, the solution also results in a negative value for $w_l$ which implies leisure is consumed beyond the point of saturation. But, if this were the case, it is unlikely that IM would have selected the observed distribution of time shown in Equation 2(a), since if IM was saturated with leisure at 3,840 hours per year, she could have easily spent more time on home production without altering her choice of time spent on work.

The solution for $w_r$ using Equation 2(a) and 2(b) results in a reservation wage of \$7.50 which is higher than the market wage. One interpretation of this is that in money

[6] Note, if all 630 hours are re-allocated to home production instead of leisure, i.e. leisure is not a substitute for work, then there is no solution to the two equations.

terms, IM values the characteristics of the job at a rate of at least approximately $0.79. Hence, the wage at which IM accepted a job may have, in fact, represented a discount of the real reservation wage, since it has been established that non-wage characteristics were important to IM. The finding that $w_r$ is $7.50 may also be interpreted as indicating that IM's true reservation wage was more nearly the average wage of $6.71 or her marginal wage of $7.00 was consistent with her work effort. Without the overtime hours she may well have found the job unacceptable.

Additionally, in the case of Equations 2(a) and 2(b), the estimated value of leisure is $7.50. This would suggest that both the reservation wage and market wage represent lower bound, or conservative, estimates of the value of leisure time. This finding is also consistent with the fact that IM required an additional payment of $1.00 in order to forgo leisure for extra hours of work.

Using Equations 2(a) and 2(d), the reservation wage of $6.45 is less than the wage at which a job is accepted. In this case, the $5.89 value of leisure is less than both the reservation wage and the market wage. This estimate of the value of leisure is not substantially less than the market wage, and it should be noted that these findings are sensitive to the assumption that the additional time from reduced work was divided evenly between home production and leisure. As relatively more leisure is substituted, the value of leisure increases relative to the reservation wage and the market wage.

It should also be noted that each of the above sets of results are sensitive to the relative values of $w_o$ and $w_h$. For example, if the value of $w_h$ is lower than $w_o$, then for Equations 2(a) and 2(c) the value of $w_r$ falls to $7.03 while the value of $w_l$ remains the same at $7.50. For Equations 2(a) and 2(d), the values of $w_r$ and $w_l$ are $8.68 and $10.00 respectively. Hence, with $w_h$ at $5.00 in Equations 2(a) and 2(d), $w_l$ is greater, rather than less than the value of $w_r$. In this case, the market wage again is a reasonable lower bound estimate for the value of leisure. Hence, even with different assumptions about the market value of home production, the market wage remains a conservative estimate of the value of leisure.

## V. Conclusions

What the above analysis shows is that if it is assumed that $w_r$ and $w_l$ are unknown or not approximated by the market wage, then their values are dependent on the degree of substitution between work and household production and work and leisure. The degree of substitution may only be determined if the expert has more than one observation on the distribution of time among the three activities.

While it is clear that the market wage may be an imprecise estimate of the value time spent on leisure, the above demonstrates that the market wage provides a reasonable lower bound estimate of the value of leisure. Further, to the extent that leisure is an increasingly preferred substitute over other uses of time, the market wage increasingly underestimates the true value of leisure. Hence, an expert needs compelling reasons to deviate from the usual assumption that the market wage represents the value of leisure.

In general, an expert is unlikely to have numerous observations on a household's revealed preferences about the allocation of time among work, household production and leisure. However, in cases where at least two sets of observations about the distribution of time are available, estimates of the value of leisure may be calculated and may differ from

the market wage.

References

Fort, Rodney and Robert Rosenman. 1995. "Rethinking the Value of Lost Health." *Journal of Legal Economics* 5:63-73

Irvine, Ian J. and William A. Sims. 1998. "Measuring Consumer Surplus with Unknown Hicksian Demand." *The American Economic Review* 8:314-322.

Isley, Phyllis and Robert Rosenman. 1997. "An Economic Analysis of the Value of Lost Health." *Journal of Legal Economics* 6.

# THE CALCULATION OF LOST PENSION BENEFITS - A COMMENT

by

Eric Frye and David Hatcher*

## I. Summary

The computation of damages arising from a personal injury claim should take account of lost pension benefits to which, absent injury, the plaintiff would otherwise have been entitled. In the context of estimating such damages under the Federal Employer's Liability Act (FELA), Ciecka and Donley [1997] have made a useful contribution to an understanding of the way in which retirement benefits are actually computed for railroad workers covered under Railroad Retirement Act. Their article provides a good summary of the various components of pension benefits under the railroad retirement system, and it generally delineates the appropriate calculations necessary to compute specific benefit amounts. In this otherwise valuable article, however, Ciecka and Donley offer precisely the wrong conclusion about the inadvisability of following the dicta of a recent court case; dicta that ironically make the substance of their article that much more useful. Specifically, Ciecka and Donley conclude -

> "In regard to pension benefits in FELA actions, or non-FELA cases as well, the most straightforward manner to compute the present value of future pension benefits entails calculating the present value of employee-paid and employer-paid pension contributions." [Ciecka and Donley, p. 136]

The authors contrast their conclusion with the decision reached in *Rachel v. Consolidated Rail Corp.* [1995] in which the court ruled that lost pension benefits should not be valued in the manner recommended by Ciecka and Donley; rather the calculation requires the precise computations specified under the terms of the railroad pension system. The present comment is intended to demonstrate why the recommendation of Ciecka and Donley regarding their seemingly straightforward approach to lost benefits is flawed and why the court's ruling is on point.

## II. The "Straightforward" Albeit Misleading Approach to Estimating Lost Pension Benefits

In an effort to explain the logic behind their observation that contributions are a satisfactory, indeed, the preferred basis for establishing the value of lost pension benefits, Ciecka and Donley offer the following "economically intuitive"[1] argument:

* Mr. Frye is Vice President of Spectrum Economics in Overland Park, KS. Dr. Hatcher is a Principal with that firm in Mercer Island, WA.

[1] "Economically intuitive" is a claim made by Ciecka & Donley for their method of estimating lost pension benefits. Specifically, "[E]conomic intuition implies that the present value of contributions should approximate the present value of benefits..." [Ciecka & Donley, 145]

> "Such a procedure seems appropriate because pension contributions are a part of the value of a person's labor. In addition, for defined contribution plans, an individual specific pension account exists in the name of each employee; and the accumulated value of the pension account directly depends on employee and employer contributions. The same procedure makes sense for defined benefit plans because the benefits that are ultimately paid to all pensioners must, in an overall actuarial present value sense, equal the value of all contributions made to fund benefits."

The problems with this rationale can be demonstrated in several ways. A simple test might entail substitution of the phrase "social security" for "pension" in the quoted paragraph. The logic would appear just as intuitive; indeed, every attribute described by Ciecka and Donley as pertaining to the railroad retirement program is just as applicable to social security, but it is hardly the case in either that there is anything other than a crude positive correlation between the "contributions" made on behalf of the employee and the actual retirement benefits received on his or her behalf. Nevertheless, Ciecka and Donley interpret that crude correlation as a direct dependency relationship. Quite simply, it is difficult to sustain an argument that social security is a contribution-based retirement program in which an employee's future retirement benefits are a direct function of the payroll taxes declared to be social security contributions.[2] As is more fully discussed below, it is equally difficult to sustain such an argument with regard to the railroad retirement system.

Alternatively, one can see the flaw in the intuitive logic offered by Ciecka and Donley by examining, for instance, their description of the Tier-I annuity portion of the railroad retirement program:

> "This annuity is approximately equivalent to benefits paid under Social Security. Employee and employer contributions are the same as under Social Security; however, eligibility rules are somewhat different. Most notably, a railroad worker qualifies for Tier 1 benefits at age 60 if he or she has 30 years of railroad service,

[2] Social security exhibits the same characteristics that Ciecka and Donley have claimed for the railroad retirement system. That is, contributions are just as much "a part of the value of the person's labor" as any other employee-benefit program funded from that employee's nominal wage. Similarly, there is a comparable dependency relationship between contributions and value of the account, (but the relationship exhibits only a crude correlation with a huge variance, just as with the railroad retirement system). And finally, the social security benefits that are ultimately paid to the aggregate of beneficiaries are supposed to "in an overall actuarial present value sense equal the value of all contributions made to fund benefits." (That, at any given time, the social security "fund" is or is not going bankrupt is not particularly relevant. Nor is it particularly relevant that one-fifth of social security payroll taxes fund medicare, the government sponsored healthcare plan, available regardless of contribution and which is widely forecast to be bankrupt within a few years. It is clear that occasionally, and as the federal government sees fit, actuarial-based estimates of the contribution levels required to keep the system from going bankrupt – whatever that may mean – will be undertaken and payroll taxes will be adjusted accordingly. As we note in the body of this comment, similar circumstances and similar adjustments describe parts of the railroad retirement system.)

> but benefits are reduced by 20 percent. In addition, a railroad worker receives unreduced Tier 1 benefits if he or she has 30 years of railroad service and is 62 years old at the date of retirement. Spousal benefits are equal to 50 percent of an employee's benefits while the employee is alive. If the employee dies, the spouse receives the full employee benefit, and the previously paid spousal benefit ceases. In addition, eligible children receive 75 percent of the employee's benefits and a dependent receives 82.5 percent in survivor's benefits." [Ciecka and Donley mention in a footnote that payments to dependents must not exceed 150 percent of the employee's Tier 1 payments and that there is a maximum survivor's benefit cap which typically becomes binding with three or more eligible family members.]

There are two observations that can be made regarding this summarization of the benefits attributable to the Tier-I portion of the railroad retirement system. First, when one recognizes the nature of the disconnect between contributions and benefits which are universally acknowledged in the social security system, then it is apparent that the railroad retirement system may be more "disconnected." This is the case because Tier 1 benefits are "calculated using social security formulas, but with railroad retirement age and service requirements."[3] Social Security, to account for its own disconnect problem, has extended beyond 65 the retirement age at which recipients born after a certain date qualify for full benefits. Likewise, the RRB has recently modified its eligibility criteria to copy the social security schedule for full retirement eligibility, but the railroad retirement system continues the "accelerated" feature that qualifies an employee for full pension benefits at age 62 with 30 years of service. In a similar vein, the spouse's benefit is not reduced if the 30-year railroad employee retires at 62 even if the spouse retires at age 60 rather than 62. The social security system's spousal benefits reflect a more strict criterion that reduces the surviving spouse's benefit if any payments are received before age 65.[4]

But most fatal to the logic offered by Ciecka and Donley is that family circumstances, which control critical portions of the benefit calculations, will have major quantitative effects on any lost-benefit calculation. It has to be recognized that family circumstances are not readily generalized across beneficiaries, yet Ciecka and Donley's estimation requires just such generalization. In reality, the pension benefit received by a railroad retiree who is either unmarried, or a widow(er) with no dependent children, will be less, by 33 percent, than an otherwise similar retiree whose marriage partner is qualified to receive spousal benefits. If the married pensioner happens to have dependent children, the difference in pension benefits is even greater. But in no case will the sum of the employee- and employer-paid contribution toward the pension benefit be different among these individuals, and thus, Ciecka and Donley's preferred method of estimation of benefits will not differ among these individuals.

Quite simply, the "straightforward" estimate of a lost benefits premised on the contributions made toward that pension will produce inconsistent and factually irrelevant results because such an estimate is independent of, at minimum, the marital and dependents status associated with the pension recipient. The only way to accurately estimate the pension benefit to which the particular employee will be entitled is to perform the detailed benefit calculations that the court has required. Ciecka and

[3] Railroad Retirement and Unemployment Insurance Systems Handbook [1997] p. 32.

[4] Social Security Handbook [1997] §407, Part B.

Donley outline those calculations in their article but reject them as a matter of recommended practice.

## III. A Comparison of the "Straightforward" Computation and the Court-Directed Computation of Lost Benefits

Ciecka and Donley attempt to demonstrate the validity of their premise regarding the "contribution-based" pension benefit by doing two computations – one based on their claim that future benefit valuations are adequately represented by the present value of actual and prospective employee- and employer-paid contributions, and the other based on the admittedly complex formulae specified by the railroad pension benefit system. Ciecka and Donley's analysis shows a huge variation between results for the two computational methods when applied to their prototypical case. Their "straightforward" approach produces a lost pension benefit value of $255,477. This contrasts with $140,245, if the actual benefit formulae applicable to the railroad pension program are used. Thus, Ciecka and Donley demonstrate that there are substantial differences between estimating retirement benefits premised on employee/employer contributions as contrasted with calculating those benefits based upon the applicable benefit formulae.

Logic suggests that future retirement benefits computed from actual formulae might be used to test whether a proxy such as contribution percentages — or any other proxy for that matter — can reasonably approximate the present value of future pension benefits. Ciecka and Donley, however, invert this logic. Rather than analyzing the root causes of the computed difference, Ciecka and Donley attack the benefit formulae as the culprit because the results violate their "economic intuition," leaving unexamined the question of whether computations based solely on railroad contributions are reliable.

In the section of their paper that presents this analysis, they do not emphasize the fact that such comparisons are very much influenced by the assumptions that describe the railroad employee under consideration — such things as the employee's work history, family structure, and disability status. Instead, they demonstrate that variation in tax status attributable to the pensioner will have a negligible effect on calculations.[5] They next assert that "scope of recipiency" will have a substantial effect on benefit calculations. By scope of recipiency, they mean the extent to which a benefit calculation encompasses the railroad employee and his/her dependents (to include spouse). For some unspecified reason Ciecka and Donley believe there is a defensible case to be made that the courts would accept a lost-benefits analysis that ignored the railroad employee's dependents and their entitlement to pension benefits under that employee's retirement program. There is no precedent for such an interpretation to our knowledge, and Ciecka and Donley have not suggested one.

Only at the end of their paper do they return to the possibility that there may be wide variation in circumstances that cause computational variances. [Ciecka and Donley, pp. 147-8] This is apparently suggested to them because their comparison of the "average" benefits paid out to families covered under the railroad retirement system differs dramatically from the result they computed for their hypothetical example. We present some observations dealing with this difference in the section below titled "Why Contributions Do Not Predict Even Aggregate Benefits."

[5] Reasonable people may disagree about what adequately describes a "prototypical" case; that is not the basis for our commenting on the comparison made by Ciecka and Donley, (although it must be pointed out that such differences in circumstances are precisely what can only be adequately revealed in the estimation procedure required by the Court and not by the procedure recommended by Ciecka and Donley).

In their concluding comments, Ciecka and Donley proffer that aggregate RRB benefit payments confirm their methodology. After noting employee retirement and supplemental pensions were 62 percent of all benefits paid in 1994, Ciecka and Donley assume the remaining 38 percent are family benefits. This supposition is offered as a demonstration of their hypothesis that family benefits are the difference between benefits estimated using contribution percentages and benefits specified by the Court. That is hardly likely to be the case.

Specifically, such a "theory" of the railroad retirement system suffers from fatal flaws. Most alarming is their failure to account for the changing demographics of railroad employment that have had predictable effects on the contribution/benefit calculus. For this particular population stratum, female spouses significantly outlive their male partners. Yet the overwhelmingly male railroad workforce, having fallen approximately 85 percent since World War II, is not a stable population base or benefit group against which to compare survivor and spousal benefits.[6] Past benefit distribution patterns associated with a contracting workforce are not representative of the benefits expected for current employees, unless someone can demonstrate how the deficiency is to be made up outside the employer/employee contribution system. It should be apparent that aggregate benefit payments that reflect historical biases are simply unsuitable for extrapolating the type of relationship Ciecka and Donley advocate.

## IV. Why Contributions Do Not Predict Even Aggregate Benefits

To adopt Ciecka and Donley's argument that a specific defined-benefit plan (i.e. the railroad retirement system) should behave like any defined contribution plan is to disregard the considerable history that surrounds the plan. In constructing their argument, Ciecka and Donley never address the fact that contribution percentages that fund the railroad retirement system have been altered repeatedly to rescue the system from insolvency.

Since October 1981, the federal government has instituted three major revisions to the railroad retirement system in an effort to restore its financial underpinnings. These revisions mainly invoked substantial increases in contribution rates without any significant increase in benefits. Starting with the Omnibus Budget Reconciliation Act of 1981, and ending with federal legislation passed in 1987, total employee- and employer-Tier 2-contribution rates have increased by more than 120 percent (from 9.5 percent of earnings to the present combined contribution rate of 21.0 percent). Over this same time period, Tier 1 contribution rates have also increased in step with increased social security contribution rates necessary to maintain the solvency of that system. Using the "economic intuition" put forth by Ciecka and Donley, however, one would have to wrongly conclude that these increased contribution rates resulted in substantially enhanced retirement benefits.

To put this last point in perspective, had Ciecka and Donley used the railroad contribution rates in effect prior to the tax changes that rescued the railroad retirement system, their analysis would have computed a retirement-benefit loss for their prototypical disabled employee of approximately $160,000. Using today's contribution rates, however, they instead compute a figure that is $95,000 larger, or some 57 percent greater than the retirement benefit under the old contribution rates. Yet the reality is that these increases in the contributions to the retirement system produced no significant change in average benefit entitlements!

[6] Railroad Retirement Board, Statistical Tables, Table D1 – Number of employees and their compensation, for all employers and class I railroads, 1937-95.

The Courts are fully cognizant of this characteristic of the railroad retirement system, and they have made known that knowledge in the two court decisions cited by Ciecka and Donley. In *Adams v. BNR*, the Appellate Court declared "[A]ny link between taxes paid and the benefits is too tenuous to provide a true measure of the Plaintiff's loss." In *Rachel v. Conrail*, the Court quoted the prior passage from *Adams* and went on to remark that "Congress determines the size of the tax contributions and the size of Plaintiff's annuity, and it has no obligation to balance the two." This latter statement by the Court makes clear that it has recognized the "disconnect" between Tier 1/Tier 2 payroll taxes and the associated pension benefits, and thus it has correctly rejected the fundamental supposition posited by Ciecka and Donley, which is that "the accumulated value of the pension account directly depends on employee and employer contributions." [Ciecka and Donley, p.136]

Subsequent to submission of Ciecka and Donley's article, the Court issued its ruling in *Edwards v. Atchinson, Topeka and Santa Fe Railway.* In that case the Court cited both *Adams* and *Rachel* as follows:

> Rachel v. Consolidated R. Corp., 891 F. Supp. 428, 429-30 (N.D. Ohio 1995), is instructive, and we quote its discussion of the issue.
>
> "Pursuant to the Railroad Retirement Act of 1974, 45 U.S.C. § 231 et seg., the Railroad Retirement Board administers disability and retirement annuities for eligible railroad workers, paid from a fund maintained by the United States Treasury. The internal revenue code requires employees and employers alike to contribute tax payments to the annuity fund. (U.S.C. §§ 3201, 3221. ) Both employees and employers presently pay an amount equal to 7.65% of the employee's gross wage in 'Tier I' taxes, which taxes sustain the Railroad Retirement Board Disability and Retirement Annuities that supplant social security benefits. The employee pays an additional 4.9% of his total compensation as a 'Tier II' tax toward the retirement fund's pension component, and the employer adds an amount equal to 16.1% of the employee's compensation in Tier II taxes. *** Defendant expects [plaintiff's expert economist] to offer a projected loss of earnings figure that includes as 'fringe benefits' the Tier I and Tier II taxes that Plaintiff and Defendant would have paid in the future but for Plaintiff's disability. Defendant argues that these sums must be excluded ***.
>
> Had Plaintiff continued in Defendant's employ until his natural retirement, he would have been eligible for a larger retirement annuity. Defendant concedes Plaintiff's right to seek damages that reflect the loss of that bigger annuity. *** But Defendant insists, and the Court must agree, that the total tax contributions by the parties do not fairly approximate the value of Plaintiff's loss. Defendant aptly quotes the Missouri court in Adams v. Burlington Northern R.R. Co., 865 S.W.2d 748, 750 (Mo. Ct. App. 1993): 'Any link between the taxes paid and the benefits is too tenuous to provide a true measure of plaintiff's loss.' Congress determines the size of the tax contributions and the size of Plaintiff's annuity, and it has no obligation to balance the two. *** As for the Tier II taxes, the Railroad Retirement Act provides the proper method for determining Plaintiff's expected benefits at § 3(b), 45 U.S.C. § 231b(b). Adams, id. at 751. Accordingly, the Court will allow Plaintiff to present evidence of the value of his lost pension benefits calculated in a manner

> consistent with 45 U.S.C. § 231b(b). It will not, however, permit [plaintiff's expert economist] to offer his opinion that the Tier I and Tier II taxes that would have been paid by the parties save Plaintiff's injury represent lost fringe benefits for which Plaintiff should be compensated in kind."
>
> Thus, Rachel held that the formula set out in Adams and 45 U.S.C. § 231b(b) is the correct method for computing retirement benefits. That formula requires a calculation of the amount of benefits plaintiff would have received had he continued working to retirement age and the amount plaintiff will actually receive. Adams, 865 S.W.2d at 751. The difference between the two amounts, discounted to present value, represents plaintiff's lost retirement benefits. Adams, 865 S.W.2d at 751.

## V. A Note Regarding "RRAMAX"

While the present authors disagree with the conclusions of Ciecka and Donley regarding estimations of lost pension benefits, we believe that they have done an admirable job of summarizing much of the basis for computing railroad retirement pensions. In particular, their explanation of the Railroad Retirement Annuity Maximum (RRAMAX) computation is a model of clarity. However, there is a subtle assumption that Ciecka and Donley have made in their prototypical example that casts the RRAMAX in a particularly dubious light. More importantly, that assumption makes it erroneous to apply RRAMAX to their hypothetical example.

RRAMAX, adopted by the Railroad Retirement Board in 1974, establishes a mathematical ceiling to a family's railroad retirement benefits. As stated by the RRB in *Railroad Retirement Maximum*, [1997] "[m]ost retirees and spouses are not affected by this maximum on benefit payments, which was intended as a reasonable cap on family retirement benefits in relation to an employee's earnings." Ciecka and Donley, however, endeavor to show that a potential, and unfortunate consequence of RRAMAX is to virtually eliminate Tier II retirement benefits for injured railroaders who have less than 20 years of service and are more than 10 years from early retirement. They assert in their prototypical case that such an individual would lose all Tier II retirement benefits due to having no railroad or other earnings in the ten-year period preceding retirement.

Ciecka and Donley then calculate RRAMAX using zero earnings, making their hypothetical Mr. Jones subject to the RRAMAX-specified minimum and limiting his pension to accrued Tier I benefits. Lost to Ciecka and Donley, however, is the fact that their hypothetical Mr. Jones is by their own assumption "totally disabled" [Ciecka and Donley, p. 143] — that means, under the provisions of the RRA, he qualifies for a full disability pension (which later converts to a retirement pension) shortly following his injury. This makes their RRAMAX calculation inappropriate.

It is conceivable that applying RRAMAX to post-injury earnings could cause some loss of Tier II benefits. A total loss due to RRAMAX, however, would be very rare. First, the railroader must be old enough to have worked for the railroad at least 10 years (thus becoming vested under Tier II), but have less than 20 years of service. Next, that same individual must also be young enough to be more than 10 years away from early retirement. Finally, that person's skills and ability to learn new trades must be sufficiently below average as to keep him from earning normal wages throughout his remaining worklife. But in no case can the injured worker be totally disabled.

In order to achieve the results postulated by Ciecka and Donley, their hypothetical Mr. Jones would had to have been partially disabled, and over his remaining worklife he would have had to be incapable of earning as little as half his prior annual railroad wages for RRAMAX to have any impact.

(RRAMAX is calculated using all "...railroad retirement and social security covered earnings...", not just wages from railroad employment.) In addition, inflation or productivity induced wage increases also must be assumed to be minimal throughout Mr. Jones' remaining working years. The existence of individuals that meet these criteria is not impossible, but they will be so rare as to provide little indication of how the railroad retirement system works other than in a pathological sense.

## VI. Ciecka and Donley's Criticism of Pension Formulae Used to Compute Railroad Retirement Benefits

In critiquing the court-ordered method of computing lost pension benefits, Ciecka and Donley argue that the computationally intensive approach required by the court lends itself to sloppiness by economists doing the computation because they may, for instance, ignore the railroad pension rights of the injured worker's family.[Ciecka and Donley, p. 147] This is a perplexing criticism. That complex calculations encourage sloppy analysis is not something we are aware has been adequately studied, and in any event, it is a worrisome professional code of conduct that suggests an analyst declare that a complex calculation need not be performed because the temptation to perform it incorrectly is too great.

Ciecka and Donley further argue that the computationally intensive approach requires much speculation regarding future family circumstances of the injured party.[Ciecka and Donley, pp. 146-7] On the one hand, this would seem to be an acknowledgment of the point we have made here — railroad pension benefits reflect the formulae that are driven by the retiree's circumstances. But more importantly, there does not seem to be anything particularly noteworthy about the speculative nature of the estimation process. The entire procedure of estimating damages — not just lost pension benefits — hinges on reasonable assumptions about unknowable future events, things such as life expectancy, real wage-escalation rates, probability of employment, etc. If an analyst is troubled by the requirement that defensible assumptions must be made in doing these estimations, it would not seem a reassuring statement regarding that analyst's suitability for the task.

## VII. Conclusion

> "The well trained economist is first a theorist, second a social scientist and last a number cruncher." [Havrilesky, 1990]

It is perhaps regrettable, but probably inevitable, that complicated systems give rise to complicated analysis. If economists are going to say useful things regarding the behavior and results that obtain in complicated systems, they are likely going to confront the third activity to which the late Professor Havrilesky referred. The railroad retirement system is a complicated system. Ciecka and Donley have made a useful contribution to the toolbox of the economist responsible for analyzing that system and have thereby eased somewhat the difficulty of crunching the numbers; but there should be no doubt, the numbers to be crunched are those defined by specific pension benefit formulae prescribed under the Railroad Retirement Act, not some proxy derived from payroll taxes.

References

Ciecka, James, and Thomas Donley, "The Calculation of Lost Pension Benefits for Railroad Workers," *Litigation Economics Digest*, 2(2) Summer 1997, 136-150

Havrilesky, Thomas, "Those Who Only Remember the Past May Be Doomed to Repeat Its Mistakes," *Journal of Forensic Economics* 3(1) 1990, 23-8.

*Rachel v. Consolidated Rail Corporation*, US District Court for the Northern District of Ohio, Eastern Division, Case No. 5:93, CV 2748,891 F. Sup. 428, April 4, 1995.

*Social Security Handbook* [1997] Washington, DC.

*US Railroad Retirement Board, Railroad Retirement and Unemployment Insurance Systems* [1997] Chicago, IL, 17-65.

*US Railroad Retirement Board, Twentieth Actuarial Valuation of the Assets and Liabilities Under the Railroad Retirement Acts as of December 31, 1995*, [1997] Chicago, IL.

*US Railroad Retirement Board, Railroad Retirement Maximum,* [1997] Chicago, IL.

*Vital Statistics of the United States, 1992*, US Department of Health and Human Services

# REGARDING THE CALCULATION OF LOST PENSION BENEFITS FOR RAILROAD WORKERS: A REPLY

by

James Ciecka and Thomas Donley

In their comment, Eric Frye and David Hatcher indicate that we have made "a useful contribution to an understanding of the way in which retirement benefits are actually calculated for railroad workers" and that we have provided "a good summary of the various components of pension benefits under the railroad retirement system." They also say that we have done "an admirable job of summarizing much of the basis for computing retirement pensions" and that our explanation of the RRAMAX provision is a "model of clarity." However, the foregoing compliments notwithstanding, Frye and Hatcher take us to task in various ways.

In this note, the term Payout Method refers to calculating the present value of lost pension benefits as the difference between the present value of pension benefits that would have been received if a person were not injured, and the present value of pension benefits that will be received given that an injury occurred. The term Contribution Method refers to the calculation of the pension loss as the present value of lost employee and employer pension contributions arising from an injury.

We feel that most of the comments made by Frye and Hatcher are due to their misinterpretation of our position regarding the complexity of the Payout Method. A basic conclusion of our paper was that the Payout Method does not completely capture the value of lost pension benefits if family and social insurance benefits are not included in the valuation. In brief, we argued that the Payout Method provides an accurate estimate of lost pension benefits as far as it goes -- but that in many instances it may not go far enough. In what follows, we briefly indicate Frye's and Hatcher's disagreements with our paper and our reactions to their criticisms.

> *1. Frye and Hatcher say that we "demonstrate that there are substantial differences" between the outcomes from the Payout Method and the Contribution Method, but they criticize us because they feel that we reject the Payout Method and recommend the Contribution Method. Their characterization of our paper is incorrect.*

Most of our paper is devoted to delineating and clarifying the complex set of calculations that comprise the Payout Method. We did this because a US District Court in *Rachel v. Conrail* required the use of the Payout Method. However, even though we did describe the Contribution Method as straightforward -- which it certainly is relative to the Payout Method, we did not suggest the Contribution Method was preferable.

Moreover, we emphasized the difference in estimates of lost benefits that arises between the two methodologies. As we noted in our paper "This amount differs so dramatically from the estimate derived by calculating the present value of pension contributions that it becomes clear that the decision in *Rachel v. Conrail* raises a non-trivial issue in estimating economic loss." It was our intention to demonstrate this difference in order to ensure a more rigorous approach to estimating the value of the lost benefits. Our conclusions were based on the recognition that the present value of premiums paid does not

provide an accurate assessment of economic loss. In fact, we stated that "In practical terms, we suggest that estimates of loss (in order to be consistent with new case law), be set equal to the value of direct benefits plus an estimate of additional family benefits ...." Nowhere in the paper did we suggest that the practitioner ignore *Rachel v. Conrail* and estimate the value of lost benefits as equal to the present value of premiums paid.

> *2. Frye and Hatcher argue that the most fatal flaw to calculating the value of lost pension benefits as equal to the present value of premiums is that it ignores the differences that family structure play in actual benefits received. We find their criticism mystifying as we argued that it is often exactly this factor that accounts for a substantial portion of the difference in estimates of pension benefits.*

In particular, we noted that the railroad retirement system provides a comprehensive package of benefits to retirees, spouses, dependents, survivors, and disabled workers. Both employee and employer contributions finance these benefits, but there is a danger that the Payout Method may not pickup the value of some of these benefits. Although the court required the use of the Payout Method in *Rachel v. Conrial*, the court did not consider, to our knowledge, several questions which immediately arose from its own decision. For example, suppose an injured railroad worker attempts to recover lost pension benefits in a law suit. It seems very possible that the injured employee's spouse may not have legal standing in that law suit, and reduced spousal benefits (usually 50 percent of Tier I and 45 percent of Tier II) may not be recoverable.

Consider an example of a plaintiff who is 20 years old, unmarried, and has no children. Also assume that it is very likely the plaintiff will marry and have children in the future, or would have married and had children but for the injury. Will a court requiring the use of the Payout Method allow recovery of lost pension benefits for a presently nonexistent spouse and presently nonexistent children? To our understanding, courts that have required the use of the Payout Method have not clearly specified the property rights that a plaintiff has in the comprehensive set of family benefits and other benefits (including the value of social insurance) which collectively comprise the railroad retirement system. We think this a significant problem with the Payout Method at the present time; Frye and Hatcher do not see this as a problem.

> *3. Frye and Hatcher argue that we incorrectly applied the RRAMAX in our example and that the RRAMAX occurs so rarely as to "provide little indication of how the railroad system works other than in a pathological sense" We disagree on both counts.*

In the detailed example in our paper, we dealt with a male who suffers an injury at age 45 and who had more than ten years, but fewer than 20 years, of railroad service. The example assumes that the injured person cannot perform his regular railroad work because of this injury. Such a person would not qualify for an occupational disability because he lacks 20 years of service and is under age 60. Furthermore, we assumed that the injured person has no earnings after being injured. Under these assumptions, future pension payments would be subject to the RRAMAX provision, and the maximum pension payment

is $1,200 per month when payments commence at age 60. Frye and Hatcher maintain that the RRAMAX would not apply because we use the words "totally disabled" in our example. However, our example clearly refers to a person who is occupationally disabled, and the use of the words "totally disabled" was only meant to convey the point that there were no earnings after the injury occurred. We believe that the occupationally disabled person described in our example would be subject to the RRAMAX provision.

We disagree with Frye's and Hatcher's interpretation that the RRAMAX can only occur in "pathological" situations. For example, in the 1980s some relatively young employees took buyouts from their railroad employers even though these employees were more than ten years from receiving their railroad pensions. When their railroad annuities began in the 1990s, these people discovered, much to their dismay, that their pensions were subject to the RRAMAX; and their pensions were much lower than they anticipated because they had no railroad earnings in the ten years prior to the commencement of benefits. In addition, starting in approximately 1983, US railroad employees working in Canada were no longer given credit for railroad service; and their future pension benefits could easily be affected by the RRAMAX provision.

> *4. Frye and Hatcher note that the connection between premiums paid and benefits received is tenuous, a contention that we agree with; however, they take our statement out of context when they state the court, ". . . has correctly rejected the fundamental supposition posited by Ciecka and Donley, which is that the accumulated value of the pension account directly depends on employee and employer contributions."*

We clearly stated in our paper that "the benefit structure does not return dollar for dollar benefits for premiums paid". We went into some detail about the possible role of social insurance in explaining this discrepancy. In our view, social insurance (*i.e.*, a system that provides a safety net for low-income earners by granting benefits in excess of their contributions by redistributing contributions from high-income earners) is a valuable commodity. Our paper recognized that courts may rule that plaintiffs do not have property rights in the social insurance aspect of their pensions. However, lost pension benefits are not only lost benefits accruing directly to plaintiffs. An accurate assessment of benefits should include family benefits as well.

We suggested that for younger workers (whose family structure at normal retirement is quite unclear) family benefits be set equal to 60% of the worker's direct benefits as calculated by the Payout Method. For older workers with a more clearly determined family structure, we suggested an individual-specific calculation based on the Payout Method using the current family structure because the inputs required in order to use this method are more clearly known. In defense of the 60% figure for younger workers, we offer the following facts which are consistent with the actuarial record of actual family benefits: spouses receive 50% of employee Tier I benefits and 45% of Tier II benefits, survivors receive 100% of Tier I benefits and 50% of Tier II benefits, spouses are typically women who are younger than their railroad employee husbands and women have longer life expectancies, and there are other family benefits (*e.g.*, for dependent children and parents). Therefore, 60% seems to be a conservative figure; and it was proffered in that sense. Frye and Hatcher contend that our 60% figure is demographically sensitive. Okay, perhaps it is;

suppose half the railroad labor force are women, suppose men start marrying women who are older than them, or suppose female life expectancies decline. It is still the case that spousal benefits are 50% of Tier I and 45% of Tier II, survival benefits are 100% of Tier I and 50% of Tier II, and there are still a whole host of other family benefits. Therefore, the 60% figure for family benefits for a young worker is still reasonable and conservative.

Finally, our paper never says, contrary to Frye's and Hatcher's interpretation, that "the computationally intensive approach required by the court lends itself to sloppiness by economists doing the computation because they may, for instance, ignore the railroad pension rights of the injured worker's family." Quite the contrary, most of our paper is devoted to an explanation of the Payout Method in the context of the railroad industry because we have every confidence that economists will get it right once they become familiar with railroad pensions. We are not worried that economists will ignore family benefits; but rather that considerations regarding the legal status of family members, or as yet nonexistent family members, may prevent economists from including family benefits in their calculations. We do not wish to avoid the "computationally intensive approach" or "crunching the numbers" if that is the most economically correct thing to do. We are reminded of an aphorism that has been attributed to E. J. Mishan; we paraphrase it as follows: *It is far better to have an approximate estimate of the precisely correct economic concept than to have a precise measure of a wrong concept*

## References

Ciecka, James and Thomas Donley, "The Calculation of Lost Pension Benefits for Railroad Workers," *Litigation Economics Digest*, 1997, 2(2), 136-150.

Frye, Eric and David Hatcher, "The Calculation of Lost Pension Benefits - A Comment," *Litigation Economics Digest*, 1998,Vol. III, No. 1, 69-77.

Rachel v. Consolidated Rail Corporation, US District Court for the Northern District of Ohio, Eastern Division, Case No. 5:93,CV 2748, 891 F. Sup. 428, April 4, 1995.