Testimony of Charles H. Fay
GAO Personnel Reform: Does it Meet Expectations
House Subcommittee on the Federal Workforce, Postal Service, and the District of Columbia
Senate Subcommittee on the Oversight of Government Management, the Federal Workforce, and the District of Columbia
Tuesday, May 22, 2007

Witness Background

I am Charles Fay. I am professor of human resource management and Chair of the Human Resource Management Department at the Rutgers University School of Management and Labor Relations, where I have specialized in the fields of compensation and performance management. These areas draw on the pure disciplines of economics, psychology, business strategy and human resource management, and courses covering both topics are offered in most business schools and all schools focusing on management and labor relations.

I have taught undergraduate, masters' level and doctoral classes in compensation and performance management since 1979. Most of my research since 1979 has focused on compensation (particularly performance driven pay) and the results have been published in a variety of scholarly and professional journals. One area of compensation that is my specialty is incentive pay, which is the intersection of performance management and compensation. I co-authored a leading text in compensation, titled *Compensation: Theory and Practice*, which has been widely used by colleges and universities as well as human resource managers in business and government. I have co-edited, and written major chapters for *The Executive Handbook on Compensation* and *New Strategies for Public Pay*. I have chaired the Research Committee of the American Compensation Association (now WorldatWork), and served as a member of that organization's Certification Program, where I taught several courses on compensation, HRIS and performance management. I was a member of the first Federal Salary Council and chaired the technical working group of the Council. I have also served as a consultant to the Bureau of Labor Statistics on several projects concerning the National Compensation Survey.

I have also served as a consultant to private and public sector organizations on the creation, evaluation and revision of compensation programs and in that capacity have conducted and critiqued job evaluation processes and labor market surveys. I have also consulted on the creation, implementation and evaluation of performance management systems for private and public sector organizations.

Introduction and Outline of Testimony

I have been asked to testify on "the use of external market data to establish pay ranges in the federal government, and, specifically, the 2004 market based compensation study conducted by the Watson Wyatt consulting firm for the U.S. Government Accountability Office." Before speaking to those two specific issues I think it will be helpful to provide context in terms of three related areas. The first of these areas is an overview of some common misunderstandings about compensation practice that may lead to a misplaced confidence in the results of any compensation study and its results. The second area is a discussion of the shortcomings of most compensation and benefit surveys, and particularly those conducted by most consulting firms. The third area, based on a research project I am doing for WorldatWork (the leading professional association in the area of compensation and benefits), is a discussion of market pricing practices in the private sector, and the difficulties perceived by compensation professionals who do this work on a daily basis.

Having completed that, I will then speak to the market pricing of the General Schedule conducted by OPM for the Federal Salary Council, which is based on Bureau of Labor Statistics data collected for that purpose. (Market pricing of federal government jobs is nothing new; both the General Schedule and the Federal Wage System have used variations of market pricing techniques for many years.) Then, I will provide my opinion of the market pricing study done for GAO, and how the market data were used by GAO in setting wages. I will close with a short summation.

Common Misunderstandings about Compensation

People who do not work in the area of compensation (and even many who do) share some common misunderstandings about compensation that lead to belief in the results of specific practices that are unwarranted. Correcting these misunderstandings allows for much better judgments about compensation practices and outcomes.

> **Misunderstanding One**: Compensation is a science. Because compensation uses the language of numbers, many people assume it to be a science. In fact, compensation is an art. In the modern corporation compensation is driven by organizational strategy, and innumerable judgment calls are made on practices and outcomes to arrive at a rewards system that is aligned with strategy. When practices produce some outcomes that are undesirable, those outcomes are changed to those thought to be desirable. For example, any job evaluation system provides results that are at odds with market data and those results may be set aside in favor of the market data. Conversely, when surveys report wage

levels for jobs that clash with the importance of the job to the organization, the survey data may be ignored.

Misunderstanding Two: Every job has an inherent value that is knowable with the right evaluation and survey tools. In fact, one of the difficulties faced by labor economists and compensation scholars is determining what makes any job valuable. A large body of empirical work has tried to tease out the source of wage differentials and any number of theories have been developed to explain job value. While many of these theories provide insights into job value (e.g., human capital theory, time span of discretion, marginal revenue product, tournament theory) none provide a comprehensive view of what lends value to jobs, and neither the empirical work on wage differentials nor the range of compensation theory can tell us what to pay for any specific job.

Indeed, the whole notion of market pricing relies on assuming that other organizations know more about job value than we do, so if we find out what they pay and pay about the same, we'll have captured the value of the job by that process. Market pricing may allow us to be competitive in some segment of the labor market but does not mean we have any real idea of the value of the jobs we have market priced.

Misunderstanding Three: A job's value in one organization is equivalent to its value in another organization. Market pricing that benchmarks all jobs at the same percentile (e.g., median, 65th percentile, 70th percentile) of the market relies on this misunderstanding. Every organization has some jobs that provide competitive advantage to it and others that, while necessary, do not provide any particular advantage. Organizations with more sophisticated rewards strategies price the "A" jobs high in market, and pay lower in the market for other jobs. Executives frequently exercise judgment to overpay certain job categories in the belief that those jobs are critical to the success of the organization.

Misunderstanding Four: One compensation consulting firm is better than another; if you only choose the correct consulting firm your project is assured of success. In fact, all the major consulting firms (and many of the smaller ones) can provide excellent service. The correct unit of analysis is the consultant, not the firm. As I noted above, compensation is an art, and some consultants are better artists than others. Similarly, some organizations know how to make use of consultants and consulting products better than others. Even the best consultant can have the hiring organization provide inaccurate or incomplete information in the course of the study and misuse his/her study results.

Compensation and Benefits Surveys[1]

While most seasoned compensation professionals often take survey results with a grain of salt, they frequently have no other data on which to base compensation decisions. I will speak to misgivings of compensation professionals about surveys and survey results in the next section of this testimony. In this section I will note the weaknesses shared by most compensation and benefits surveys. The use of survey data by organizations can also create poor results; I will speak to these issues in this section, too.

Survey problems.
Anyone developing a wage or benefit survey has to decide the job(s) for which data will be collected, and the appropriate market from which data will be collected. Even broad national surveys conducted by the major consulting firms face this problem, since one of their services to clients is providing different survey "cuts," which focus on specific labor or product markets, or even a selected list of organizations.

1. Job definition. The first problem arising from this decision is how the job will be defined on the survey. Job title alone would clearly be insufficient, since very different jobs might have the same title in different organizations. A detailed job description listing all tasks and outcomes associated with the job, the job specifications (knowledge, skills and abilities), and job criticality to the organization would assure a perfect job match, but such details are rarely used because the survey would be too cumbersome to respondents and minor differences would rule out a match. Some surveys with more detailed job descriptions have tried to get around the minor differences problem by allowing respondents to note the job for which they are entering data is a "smaller" or "larger" job than the survey job. However, most surveys provide only a short job description. Sometimes the job description is modified by a job family level definition.

   An example of this is one of the surveys used by Watson Wyatt in their GAO study that is produced by WTPF.[2] The 2007 version of this survey form (2007 Compensation Survey Guide), on the WTPF web site,[3] provides the following job description of the match used for Analyst (PE-347) in Bands I, II and III:

---

[1] Much of the argument in this section is based on Rynes, S. L. and Milkovich, G. T. (1986) "Wage Surveys: Dispelling Some Myths About the "Market Wage" **Personnel Psychology** (39) 71-90. While this article was written in the context of comparable worth and other wage discrimination issues, their points about the difficulties associated with salary determination based on wage surveys are relevant to the current issues with the surveys used in the Government Accountability Office study.

[2] Letter from Gary L. Kepplinger, General Counsel, GAO to The Honorable Danny K. Davis, dated March 30, 2007, pp. 4-5.

[3] http://209.200.109.246/home/documents/07WTPFGuidewithcoverfinal.pdf , accessed 13 May 2007.

"OPOR – Operations Research/Analysis: Conducts analytical studies of military, commercial or civil operations. Projects and/or programs may involve engineering, scientific, information systems, logistics, administrative, or strategic planning expertise, and lead to recommendations to improve operational effectiveness in the client organization." (p.15)

The analysts in Bands I, II and III are differentiated by a "Career Level" marker. Thus, the Band I analyst is a P2, or Intermediate Professional level, defined as "Exhibits technical and operational proficiency in the primary duties of the job family. Plays a key role in implementing projects and programs in the function. Acts as a resource to managers and employees in the organization. Typically requires: Bachelor's; 2 – 4 years of related professional experience."

The Band II analyst is matched at a P4, or Career Level. I have not included the definition because the 2007 Compensation Survey Guide lists "P4" as "Advanced Professional" and "P3" as "Career Level Professional." It is possible that the career levels have changed since the 2004 survey or that the General Counsel's letter contains a typographical error. It may also be that data from both levels were blended.

The Band III analyst appears to be a blend, since it is listed in the General Counsel's letter both as an "Advanced Level Professional (P4)" and as a "2nd Level Manager (M2)."

Under any circumstance, the generic sort of description, while better than a job title, may not provide sufficient detail to assure that all the jobs entered under that category are a good match.

2. Market definition. A second major difficulty facing a survey manager is the market from which to collect data. Most national surveys try to collect data from a sufficient number of organizations so that a variety of geographic, industry, size and other cuts can be provided to a client. In this case it appears that Watson Wyatt got data covering non-profit and other research firms in the Washington DC area.

   One of the rationales for collecting data in a specific market (whether geographic or industry) is that this is the market from which the organization attracts labor and to which they lose labor. Yet, organizations have some choices as the markets in which they will compete. They choose where to recruit labor and they choose those employees who are leaving who will receive a counteroffer. That is, organizations are not merely passive players in labor markets.

3. Sampling the Market. The largest threat to the validity and usefulness of compensation data from surveys is the lack of sampling sophistication in the large commercial survey firms. In most cases the surveys rely on what is called a "convenience" sample – a list of current and former clients and purchasers of surveys. Purchased mailing lists are also used to find respondents. I receive invitations to participate in salary surveys from several major consulting firms, and the name and address on the mailing label is usually the same as that on materials I receive from one or more of the professional human resource management organizations to which I belong.

   In some cases surveys are aimed at a specific group of organizations specified by the organization for which the survey is done, so that sampling is not an issue. Even in this case, the organization may not have specified the respondents in a way that meets its data needs.

   As a result, when multiple surveys are compared, results may be very different. Differences may be attributable to differences in job descriptions or the different set of respondents providing data for each survey. An additional source of differences is the different data editing rules used by different survey organizations and the different estimating techniques that may be used. Some survey organizations, for example, drop outliers from their calculations, while others contact respondents and check for accuracy, dropping those that aren't supported.

   Only the Bureau of Labor Statistics uses rigorous sampling methodology in its wage surveys. They are also the only wage survey organization that uses rigorous statistical methodology to evaluate survey data, and they do not publish data that do not meet their criteria.

4. Differences in Data Collected. While all wage surveys collect data on base salaries, there are differences in which other parts of the rewards package are included in the survey. Rewards include not only base salaries but also short term incentives, long term incentives, recognition awards, perquisites, benefits, work/family accommodations and other things of value to employees. Labor economists have long argued that employers seek a reasonable level of labor costs, and then allocate labor costs across various parts of the reward system. For example, an employer with high cost levels of benefits would pay less in wages than would another similarly situated employer with lower cost levels of benefits.

   Most surveys include questions on the cash value of short term incentives so that a "total cash compensation" figure is provided. Without more inclusive data on the value of other parts of the rewards package surveys do not provide data that allows the user to make a meaningful comparison with benchmark organizations.

Other data collected (or not) makes it easier (more difficult) to compare the results of two surveys. Labor economists have done extensive work on sources of wage differentials. I have provided a list of these (Appendix A). I know of no survey that provides all these data (especially on job and individual differences), but without them survey users have less assurance of the comparability of benchmarks.

5. Impact of Discounts to Survey Participants. Most commercial surveys have two prices: the normal price and a discounted price for survey participants. Watson Wyatt, for example, lists charges of $1200 to a non-participant for its survey of human resource personnel compensation, but only $500 to a participant. If discounts are driving participation, they do not necessarily drive conscientious participation.

6. User Problems. Even with the best surveys available, a user wanting to benchmark a specific job has significant choices to make that call into question the accuracy of the results:
    - How many different surveys are needed to make a match?
    - How does the user reconcile differences between the different wage levels reported by different surveys for the same job?
    - How does the user adjust for different job level breakouts? If one survey has research associate levels one through six, a second has levels one through three, and the user has levels one through four, how can adjustments be made to assure equivalence?
    - What impact do relatively small differences in job descriptions have on wage data, and what adjustments might be made to data to account for these differences?
    - What adjustments should be made if the job in the user's organization is critical to the success of that organization but might not be to those of many survey respondents?
    - What can be done if a satisfactory survey match can't be found?
    - What can be done if the user organization's job is a "hybrid" of two or more jobs for which wage survey data are readily available?
    - What should be matched? Choices include base pay, total cash compensation, total rewards (including long term incentives and benefits).

At this point a reader may wonder whether there is any value to wage and benefit surveys at all. It is my contention that they are much better than no data at all, but that compensation professionals have to be careful to understand survey data shortcomings and that the collection and use of wage data is an art rather than a science. Most importantly, the process of collecting and using wage data has so many points at which judgment is used that certainty of the goodness of results is rarely warranted, especially at the individual job level.

Market Pricing in the Private Sector

In the private sector market pricing of individual jobs and entire pay structures is common. Some organizations conduct some sort of job evaluation to create an "internal value hierarchy" of jobs. Jobs within a given spread of job evaluation points are assigned to the same salary grade. This set of grades forms the basis of a salary structure, which is then priced using market data for as many of the jobs in the structure as can be obtained. Market data is used to develop a midpoint for each salary grade, and minimum and maximum salaries for the grade are calculated. Jobs for which data can not be obtained are paid the same rate as other jobs in the same salary grade.

Other large organizations rely entirely on market data for their pay systems. In these organizations jobs are paid their market rate. Statistical models are developed to estimate market rates where none exist. The most common statistical method used is regression. Job attributes (e.g., experience required, education required,) are collected for each job and market rates are regressed on this set of attributes. In some cases the attributes come from job descriptions and specifications; in other cases they are drawn from the employee data in the human resource information system of the organization. The estimated market rate becomes the pay base for these jobs.

Although private sector organizations rely much more heavily on market pricing than do public sector organizations, this does not mean that private sector organizations are necessarily satisfied with the results or do not have problems with the processes used to develop and use market rates. The results of stage one of a research project on market pricing I have conducted for WorldatWork[4] speaks to the concerns private sector professionals have about market pricing, and forms the basis for the following discussion. Three areas are considered: how organizations strategy should affect market pricing processes and outcomes, the value of various market pricing sources, and the analysis of data from surveys to determine the appropriate rate for a job.

Compensation Strategy and Market Rates. An organization's compensation strategy typically reflects its business strategy. It provides the guideline for various compensation decisions, including how to balance market pricing data and internal equity.

The most frequently raised issue was how to decide whether internal or external equity should take precedence. Many respondents would like guidelines helping them determine what a market pricing policy should look like, and a means of resolving conflicts between market rate and job evaluation (or other) indicators of job value.

A second critical issue for many respondents is determining which market should be priced against. There seems to be widespread recognition of the existence of

[4] Fay, C.H. and Tare, M. (2007) Market Pricing Concerns. WorldatWork Journal, 16(2), 61-69.

different definitions of markets (e.g., geographic breakouts, from local to global, product market competitor or industry breakouts, size breakouts) but much concern about when each might be used. One respondent noted that while the "textbook" answer was available it didn't seem to match the reality reflected in discussions with colleagues and other professionals.

The third compensation strategy issue raised by respondents is the desire for guidelines on how competitive to be in a variety of different situations, and what a competitiveness strategy that delineated these situations would look like. Different breakout groups (e.g., hot jobs, critical jobs, typical jobs, executive jobs) were the focus of different respondents, but the common thread running through comments in this area is how one should best determine the competitive level of rewards for a set of jobs.

The fourth issue raised with some frequency focuses on whether competitiveness in a labor market should be based solely on wages, or whether a broader rewards (total cash compensation, total compensation, compensation plus work/life balance, etc.) measure should be used. There appears an unsatisfied need for strategy guidelines on individual and joint reward segment competitiveness.

There is clear recognition that the proper use of market data is a critical issue for organizations that are trying to stay competitive in attracting and retaining human capital while staying competitive in product and service markets. Most respondents note that no "one best strategy" exists while at the same time there is a perceived need for best practices in market pricing strategies taking into account industry, organizational and business strategy characteristics.

Market Pricing Sources. Compensation professionals typically obtain market pricing data from several sources, e.g., standard and custom surveys done by consultants, industry association surveys and formal surveys they have done themselves. While the federal government conducts the largest survey (BLS' National Compensation Survey), few of the respondents mention it. Respondents have to make decisions on a variety of source issues: the number of surveys to use, type of surveys (commercial or non-commercial, standard or custom), selection of an appropriate survey and number of data points to be used for benchmarking a job. Market pricing sources elicited the largest set of concerns from respondents.

The reliability, quality and breadth of data from market pricing surveys are all major areas of concern for compensation professionals. Added to this is the perception that survey costs are very high. There are concerns that as organizations are participating in fewer surveys, the number of useful surveys is smaller and the quality of data is dropping. Since the survey costs are high, some organizations may participate only to get a lower price for the survey but not put in the effort to provide quality data for the survey.

Respondents feel that although there are many surveys available, most are general rather than industry-specific and finding one that covers some industries is difficult. Getting two comparable data sources is hard when there are different industry participant groups for each survey. Most respondents raised issues of the number of organizations/positions required before wage data for a specific job would be useful. There is considerable desire for guidelines on this point. Similarly, many respondents expressed a desire for guidelines on the number of surveys that should be sought for each market-priced job.

There is concern whether an informal telephone survey will provide valid data not available in a more formal survey. On a related issue there is concern that few surveys have made an attempt to incorporate "team" and "hybrid" jobs. In part this is inevitable as the nature of work changes. Different organizations have reconstructed work in different ways so there are many variants on any team or hybrid job, even though the original jobs from which these new jobs were reconstructed are similar across organizations. It is unfortunate that survey organizations have not yet developed techniques to capture the value of job parts that can then be combined.

Even when matching jobs are found, respondents are concerned that the brief job descriptions provided make job matches problematic. Scope data are not always provided and are rarely sufficient for most respondents to feel certain of their matches. This experience leads several respondents to question whether wage survey participants can make valid matches when responding to surveys. Uncertainty at both ends of the survey process has led some survey participants to question the value of any data from some surveys.

Respondents strongly feel the need for industry standards for survey companies. This is because the methodology, reporting procedures and data of survey companies vary considerably. In such a scenario, it is difficult for respondents to accept inconsistent data. There is a perception that survey samples may not truly reflect the marketplace, as the participant lists seem skewed in many surveys. The reliability of data is also questioned, when data fluctuate considerably from year to year for the same survey. Consulting firms providing wage surveys were perceived to make a better effort to scrub and validate data than trade and professional associations.

Comparing jobs is difficult, as some benchmarks are based on incumbents and others on positions being offered. This is a problem with hot jobs especially. Moreover, data for hot jobs doesn't reflect what companies must offer to get a candidate interested. Jobs that seem similar (e.g. marketing and sales), are hard to compare, as they vary from industry to industry and even across companies within the same industry.

Data on rewards components (base, variable, and total compensation and job levels) is perceived to be inadequate for matching purposes. Respondents would like more information on variable pay programs; including both actual and target data. Short term incentive programs and their interaction with wages and benefits is of particular concern to respondents.

Many respondents did not stop with simply complaining about survey vendors, and provided input as to how some of the problems noted could be lessened, if not solved. Respondents feel that to begin with, companies should actively participate with vendors (consultants or industry association) to assure that better data comes out of the survey. Another way of ensuring this is that survey companies should make an effort to reach the right person, not just the right sounding title, while collecting data. Organizations who want survey data have an obligation to take part in surveys, and use the resources to provide the most accurate and complete data they can. Some respondents note they pay the most attention to surveys that have active participant groups.

More comments were made on survey sources than any other area. Respondents are suspicious of the goodness of survey results and indicate many vendors do not provide value for money, given the perceived flaws in survey results.

Analysis of Survey Data. Comments on analyzing survey data parallel concerns about market pricing sources. Some of the issues respondents raised concerning survey data analysis include which survey data can be trusted, how many matches can be made, how close must the survey job description be to the organizational job description for it to be considered a match, how to account for differences in job levels, how to price hybrid and cross-functional jobs, and how to deal with outlier jobs.

Respondents note the difficulty of judging the reliability of data from small sample surveys. Making geographic adjustments and adjusting rates for jobs that have specialized requirements (e.g., heavy travel and bad working conditions), is a challenge.

Determining what characteristics make for a strong or weak match is difficult. So is adjusting rates for a combination of strong and weak matches. When survey data are available by job level (e.g., 3 job levels), it is difficult to reconcile this to a greater (e.g., 5 level job) level job in the organization.

Cross-functional jobs and those having rapidly evolving roles present special problems when the analyst must blend data for two or more survey jobs to estimate the value of an organizational job. Determining the criteria to be used for weighting and the actual weights to be used present a challenge.

Even when incentive, benefit and/or option awards data are available in the survey, determining equivalency and balancing trade-offs is not easy. Several respondents noted that the lack of starting salaries in many surveys makes it difficult to determine what starting salaries should be offered, especially for hot jobs where new entrants to the field may make more than survey means or medians. The format of many surveys is such that some analyses are not possible, even when the data are delivered electronically. Respondents recognize that confidentiality issues place limitations on what data can be delivered to survey purchasers, but would like more freedom in analysis.

Respondents noted several approaches to deal with inconsistent and insufficient survey data. Some respondents document the job pricing process, so that it is consistent from year to year and across analysts. They also document any adjustments made to survey data and have set percentage limits on how much data can be adjusted. Other respondents have developed differentials between peer company wage levels and the rest in survey where possible. This is used to estimate appropriate wage level for jobs where peer companies do not report.

Benchmarking sources and process are also evaluated by respondents. Some collect market data for multiple years and do five year trends for each survey for composite market rate. They analyze changes in survey participants, survey price rises, etc., to understand why market rates reported may have changed. They also often look at rates offered for jobs at 'Careerbuilder' and other websites as a check on benchmark rates developed internally.

It is clear that the concerns raised by compensation professionals mirror the concerns of many of the people questioning the results of the Watson Wyatt study done for GAO and GAO's use of survey data in pricing jobs.

Market Pricing in the Federal Government

The federal government has used market rates in setting wages for government employees for many years. Two systems cover most federal workers: the General Schedule and the Federal Wage System. Both systems use market data to price wage structures rather than individual jobs.

Until the Federal Employees Pay Comparability Act of 1990 (FEPCA) the General Schedule was priced using national pay rates for selected benchmark jobs. Recognizing geographic differentials, FEPCA mandated the use of survey data from major labor markets to price the GS differently in each of those markets to achieve "locality pay" comparability. Currently there are thirty one pay localities plus "Rest of US" to make a total of 32 different possible General Schedules. Benchmark jobs are no longer used; instead data from the National Compensation Survey (conducted by BLS) is weighted by GS employment in each area to provide comparable pay benchmarks for each locality.

The Federal Wage System (FWS) covers appropriated fund and nonappropriated fund blue collar workers who are paid by the hour in 131 appropriated fund areas and 125 non-appropriated fund local wage areas. Each local wage area consists of one or more survey counties and one or more "areas of application" counties. Survey data from the survey counties are used to price wage structures in the area of application counties. Again, the wage structure is priced rather than individual jobs.

Pricing the structure means that differentials between different federal pay grades (whether GS of FWS) will remain constant across areas even though that may not mirror the relationship of different jobs in different areas. (Survey data show that it is not unusual for job "A" to be priced much higher than job "B" in one area while just the opposite is the case in another area.)

The kind of market pricing done with respect to the GS structure and the FWS structure can not be compared to the market pricing study done by Watson Wyatt and the application of the data developed in that study by GAO. In both cases the structure is priced rather than individual jobs. BLS uses impeccable methodology in gathering reliable and valid data to price the GS, and applies sophisticated statistical methods to evaluate survey data and to apply it to the GS for the Federal Salary Council. FWS methodology is methodologically less sophisticated than that used by BLS, but is done by government employees and employee union representatives, so that the results have a high level of acceptance by federal managers and employees.

Market Pricing and the General Accountability Office

The various materials I have received on which I base my discussion of the market pricing study done by Watson Wyatt for GAO include:

- Career Stream Published Survey Job Links and Position Descriptions. Dated September 1, 2004
- Executive Committee Briefing: Compensation Design Task 2. Dated October 18, 2004
- Executive Committee Briefing: Compensation Design Options. Task 2. Dated October 29, 2004
- Letter dated March 30, 2007 from Gary L. Kepplinger to The Honorable Danny K. Davis
- Letter dated April 3, 2007 from Gary L. Kepplinger to The Honorable Danny K. Davis
- Letter dated April 3, 2007 from David M. Walker to Daniel P. Mullholland (with enclosures)
- GAO Compensation Design. Presented by Watson Wyatt. April 12, 2007
- GAO Custom Data (undated)
- Watson Wyatt Contract and Related Costs (undated)

Miscellaneous data from web sites of Abbot Langer, Cardom Associates, prm Consulting, and WTPF. These organizations produced surveys used in the market pricing project.

I have also had conversations with Allan Hearne, of OPM, who studied additional materials concerning the surveys and the process used to set GAO pay bands.

I focus in this section only on the market pricing of the analyst jobs.

I will begin by noting that it is not entirely clear what Watson Wyatt and GAO did in their study and in the application of study results to GAO pay bands. It is possible that documentation exists that would explain in greater detail and with more clarity exactly what was done and how and why it was done. If that documentation does not exist, that constitutes an additional flaw in the study. Pay is one of the most visible links between an employee and an organization, and it is critical that each employee understand how his or her salary is determined. One of the hallmarks of the general schedule and the FWS is that every employee in those systems can understand how his or her pay was determined. It is certainly possible to do this in a market pricing system, and organizations who rely on market pricing, such as Johnson & Johnson and Motorola, almost always devote the necessary resources to make sure employees understand the system.

Problem One

My first concern with this market pricing project is the lack of a coherent competitive strategy. Excerpts from the GAO web page "Why work at GAO" include such phrases as "epicenter of government decision making," "our recommendations result in hundreds of actions-including landmark legislation-that lead to meaningful improvements in government operations and billions of dollars in direct financial benefits on behalf of the American people," "employees are at the front line of congressional oversight, and our work depends on their knowledge, analyses, and specialized skills," and "attract some of the brightest, most dedicated people in government."

The unique nature of the work performed by the Government Accountability Office, its scope and the level of impact on the nation, suggest that pay levels at GAO should be set to attract and retain the best employees with the requisite skills, abilities, and experience. Selecting a competitive pay strategy that only matches the market means that the pay system is benchmarked only against the typical worker, not the best and brightest. Most organizations who market price differentiate between critical and less critical jobs when developing a competitive pay strategy. More critical jobs might be benchmarked at anywhere from the 60th to the 75th percentile, rather than at the 50th percentile (median). I know of large corporations who have, at times, benchmarked some of their most critical jobs at the 90th percentile, to be certain that they attract and retain the "A" players for the "A" jobs.

Problem Two
My second concern is the lack of involvement of a wider variety of employees in defining GAO jobs and in determining whether the survey jobs chosen were good matches. The first rule of job matching is that subject matter experts be involved. It is possible that senior executives are sufficiently expert in the jobs that were being benchmarked that they could provide accurate data. However, work today is changing rapidly, and even though a senior executive may have been a job incumbent at one time it is unlikely they are expert in the evolved job.

Even if the senior executives do know the jobs sufficiently it is appropriate to involve incumbents in the process. Employee involvement helps assure buy-in for rewards system development and implementation. Most organizations who value their employees also value their employee's input into critical human resource processes such as rewards. If employees lose faith in the reward system the likely outcome is reduced effort, increased turnover, and employees seeking a required governance role through unionization drives.

Problem Three
Given the uniqueness of the jobs involved I was surprised that Watson Wyatt used off the shelf commercial surveys rather than developing a specific survey to cover the job set. Watson Wyatt and GAO would then have had much more control over the sample and could have assured that the organizations as well as the jobs were equivalent. The surveys employed in this market pricing project use convenience samples and the number of participating organizations is low enough that arguing that any one or combination of these surveys represents the "market" is a stretch.

Most of the surveyed organizations listed are much smaller than GAO and may have only one or two analysts. These analysts are likely to be fairly narrow in terms of the analyses they must do, whereas GAO analysts must undertake a much broader range of projects. Breadth and depth of knowledge required are likely to be greater among GAO analysts.

The work of GAO analysts is likely to be more critical than that of analysts in many non-profits: there is a lot of difference in an analysis done to support a particular point of view for an organization urging a policy decision and an impartial analysis done evaluating the strengths and weaknesses of programs or policy proposals.

Problem Four
The off the shelf commercial surveys used are not flawless. As an example, one of the surveys used to price the analyst job (Band 1) is produced by Cordom Associates. The two Cordom jobs used (out of a total of five jobs used to price the analyst job Band I) are #79, Public Policy Analyst and # 81, Research Associate. For #79, 19 companies report a total of 108 jobs. Fifty-one of these

jobs were in one company. The lowest salary reported for this job was $35,000 per year; the maximum was $125,000 (3.57 times as much). One would not expect to see such a range for a single job in a single labor market. For job #81, 21 companies report a total of 90 incumbents. Fifty-eight incumbents are in one company. (The companies having the largest number were not the same; the company with 51 incumbents in job #79 had only 1 incumbent in job #81.) The lowest salary reported for job #81 was $34,000 per year; the maximum was $106,621 (3.14 times as much).

The stability of survey data reported also calls the data's validity as an accurate benchmark into question. The weighted average of job #79 rose by 11.5% between 2003 and 2004 while the weighted average of job #81 dropped by 8.8%. If both of these jobs are representative of the analyst job one would expect similar changes from year to year. The difference in changes is either because the jobs are not both good matches for the GAO job or because volatility in survey participation masks true changes in wages from year to year. I expect both problems play a role.

Similar problems exist with Cordom's job #82, Research Fellow, which was one of five jobs used to price the analyst job Band II. Nine survey respondents provided job data on 40 incumbents. Two of the companies provided 26 (65%) of the 40 data points used. Salaries for job #82 ranged from $28,000 to $125,058 (4.47 times as much). The weighted average increased by 13.6% between 2003 and 2004.

Problem Five

Watson Wyatt's selection of cuts of survey data is hard to understand. WTPF's (Washington Technical Professional Forum 2004 Compensation Survey provides several cuts of data of its OPOR (Operations Research/Analysis) job at levels P2 (which is matched to analyst), P3 (which is matched to senior analyst) and P4 (which is matched to supervisory analyst). Three industry cuts are provided: Government Contractors, R&D, and Technical/Professional Services. It appears from working backwards from the results to the surveys that Watson Wyatt matched the analyst to the Government Contractors' cut, but matched the senior analyst and the supervisory analyst to the Technical/Professional Services cut. This inconsistency in application is not explained or documented in the materials I have seen.

Data from a PRM Consulting wage survey of research organizations was used to benchmark Band I, II, and III analysts. Unlike other surveys used in the Watson Wyatt study the PRM survey provides data on total cash compensation rather than base pay. It is not clear what impact this has on the market benchmark. The other anomaly in this data is that Watson Wyatt evidently used not only data for the Washington area but also data from New York and "Other." For example, the PRM survey job Social Policy Researcher III was used to benchmark the Band II analyst. To get to the 393 incumbents reported by Watson Wyatt it is necessary

to add the 294 incumbents for Washington, the 27 incumbents listed for New York City, and the 72 "Other" incumbents.

Thus there is an inconsistent choice of benchmark cuts chosen by Watson Wyatt for at least three of the surveys used to benchmark the analyst series of jobs. There may be documentation of the rationale for these inconsistencies but I have not seen it. This sort of inconsistency casts doubt on the entire benchmarking process and its results.

Problem Six

In its presentation titled GAO Compensation Design (dated April 12, 2007) Watson Wyatt notes (page 17) that since "GAO competes for talent against general industry, including for profit, not for profit, federal government, and general industry," for "pay competitiveness assessment and design each market was weighted equally." Yet, the weights reported are not equal – for the analyst job five jobs from four surveys were used as benchmarks and the weights used in combining these were Abbott Langer Consultant weight 33%, PRM Social Policy II weight 11%, WTPF Operations Researcher P2 weight 33%, Cordom Not For Profit Research Associate 81 weight 11%, and Cordom Not For Profit Public Policy Analyst 79 weight 11%. It would be interesting to know how the Watson Wyatt consultant arrived at these weights and how these weights approximate equal weights for each of the four sectors.

Problem Seven

The pay range options developed by Watson Wyatt and chosen by GAO are not well explained in the documentation I have received. The "General Methodology" described (page 7) in Executive Committee Briefing: Compensation Design Options (October 29, 2004) is typical of organizational practice but depends on the quality of the market data developed, which in this case is not high. The statement that each market is weighted equally appears not to be the case.

Building ranges within bands "derived by clustering and grouping median market data" (Executive Briefing, p. 14) has no rationale provided. The market medians for jobs by band (Executive Briefing, p. 4) do not indicate any natural clusters or groupings and could not be expected to since the exhibit notes "not to scale." There are not enough data points to do any mathematical clustering and the fact that Watson Wyatt could suggest two different clusterings (one of three ranges, one of six) for band 1 indicates this was merely some preferred breakout rather than a logical one. Given the quality of the data driving the process it would be difficult to maintain that the resulting salary ranges were much more than arbitrary and artificial constructs.

Summary and Conclusions

Ordinarily, when a market pricing process indicates some significant number of incumbents are either overpaid or underpaid most compensation professionals would pause and look for explanations before assuming the process and resulting data were correct. That appears not to have happened at GAO. Yet the problems in this study are significant.

1. There is a disconnect between the rhetoric of GAO being the home of the best and the brightest and the competitive compensation strategy.
2. Only executives appear to have been involved in the study, which could result in poor job matches and lack of buy-in from employees whose pay was impacted by the study.
3. Off the shelf surveys were used that are unlikely to have captured appropriate market data. A custom survey would have provided a better basis for benchmarking these jobs.
4. The data from the surveys used is problematic. Too much of the data comes from too few organizations, the range of data for each job is very broad, and the data are not stable from one year to the next.
5. Watson Wyatt used inconsistent data cuts in developing benchmark medians.
6. The process used by Watson Wyatt to blend data is at odds with the process they claim to have used.
7. The pay ranges developed within bands are problematic, both because of the data input and because of the "clustering/grouping" technique.
8. Documentation of the study process and the resulting pay structure are ambiguous and confusing. Employees should understand how their pay structure was established and nothing I have seen is likely to lead to that.

Compensation is an art, not a science. That does not mean that it is, or should be, free of any standards. GAO is noted for the quality of its analyses. It is unfortunate that the same care was not taken with the analysis of its own pay system.

APPENDIX A
Sources of Wage Differentials

1. Geographic Factors
    a. Region
    b. Urban/Suburban/Rural
    c. Ease of Commuting (local)
    d. Desirability of Location (local)
    e. Supply/Demand Imbalances
    f. Labor Market Demographics
    g. State/Local Legal Requirements

2. Industry Factors
    a. Longevity
    b. Profitability
    c. Technology
    d. Capital/Labor Ratio

3. Organizational Characteristics
    a. Size
    b. Structure
    c. Stage in Growth Cycle
    d. Degree of Unionization
    e. Economic Success
    f. Competitive Pay Policy
    g. Other HR Policies (e.g., internal vs. external labor market strategy)

4. Job Differences
    a. Requirements
    b. Contribution
    c. Incumbent Characteristics (e.g., occupational segregation)
    d. Job Characteristics (e.g., degree of supervision)
    e. Setting Characteristics (e.g., safety, health)

5. Individual Differences
    a. Ability
    b. Performance
    c. Potential
    d. Demographic Characteristics (e.g., gender bias)

Note: There is an interaction between many of these factors. Industry is not distributed proportionally across or within regions. Different industries are characterized by different sets of jobs. Different jobs are likely to exhibit different incumbent demographic patterns.