STATISTICAL APPENDIX TO REPORT ON ACHIEVING GREATER CONSISTENCY IN SOCIAL SECURITY DISABILITY ADJUDICATION: AN EMPIRICAL STUDY AND SUGGESTED REFORMS

ANALYSIS OF ADMINISTRATIVE LAW JUDGE DISPOSITION AND ALLOWANCE RATES (FISCAL YEARS 2009 – 2011)

HAROLD J. KRENT

IIT CHICAGO-KENT COLLEGE OF LAW

SCOTT MORRIS

IIT COLLEGE OF PSYCHOLOGY

Final Report: Apr. 3, 2013

This report was prepared by was prepared for the consideration of the Administrative Conference of the United States. The views expressed are those of the authors and do not necessarily reflect those of the members of the Conference or its committees.

TABLE OF CONTENTS

I. Methodology for ALJ Adjudication Analysis 1
A. Sample 2
B. Variables 3
C. Analyses 5
II. Descriptive Statistics on ALJ Dispositions 5
A. Number of Dispositions 5
1. Distribution of Disposition Frequency 5
2. Yearly and Monthly Trends 9
3. "Hurry Up" Dispositions 10
B. Allowance Rates 13
1. Distribution of Allowance Rates 13
2. Changes in Allowance Rates Over Time 14
C. Time to Disposition 17
D. Variance Decomposition Analysis 17
E. Identifying Outliers 19
III. Analysis of Outlier ALJs 23
A. Outliers Based on Disposition Frequency 23
B. Outliers Based on Allowance Rates 24
C. Outliers based on Time-to-Disposition 25
IV. Correlates of Disposition Outcomes 25
A. Analysis 25
B. Correlates of Monthly Disposition Outcomes 26
C. Correlates of Yearly Disposition Outcomes 29
D. Multiple Regression Analysis 33
1. Analysis 33
2. Predictors of Number of Dispositions. 35
3. Predictors of Allowance Rates. 37
E. Effects of Individual Case Characteristics 39

1. Analysis 39
2. Video Hearings 40
3. Claim Type 42
4 New Evidence. 43
5. Medical Expert 45
6. Vocational Expert 46
7. Claimant Representative. 48
8. Bench Decisions, OTR Decisions, and Use of Decisionwriters 50
9. Prototype States 50
10. Type of Impairment. 52
V. Analysis of Remands 54
A. Analysis of Federal Court Remands 54
1. Data 54
2. Analysis of Court Remands 54
3. Remands by Region 55
4. Other Case Characteristics 55
5. Reason for Court Remand 56
6. Remands Related to New Evidence. 58
B. Appeals Council Remands 58
1. Frequency of Appeals Council Remands 58
2. Reasons for Appeals Council Remand 59
3. Remands Related to New Evidence. 59

LIST OF TABLES

Table A- 1: Descriptive Statistics on Monthly ALJ Data 6
Table A- 2: Descriptive Statistics on Yearly ALJ Data 8
Table A- 3: Yearly Allowance Rate (Fully + Partially Favorable) Including Senior Attorney Decisions 15
Table A- 4: Change in Allowance Rates (FY 2009 - FY 2011) 16
Table A- 5: Percent of ALJs with Extreme Allowance Rates (Fully Favorable + Partially Favorable) 16
Table A- 6: Intraclass Correlation Coefficients (ICC) for Within- and Between-ALJ Variance 19
Table A- 7: Extreme Values for Number of Dispositions and Allowance Rates (Fully + Partially Favorable), FY 2011 22
Table A- 8: Number of ALJs with Multiple Years in Top and Bottom 1% of the Disposition Frequency and Allowance Rate Distributions (FY 2009 - FY 2011) 23
Table A- 9: Within- and Between-ALJ Correlations 27
Table A- 10: Descriptive Statistics on Yearly Disposition Outcomes and Case Characteristics 30
Table A- 11: Within- and Between-ALJ Correlations Among Yearly Disposition Outcomes and Case Characteristics 32
Table A- 12: Within-ALJ Predictors of the Number of Dispositions 36
Table A- 13: Between-ALJ Predictors of the Number of Dispositions 36
Table A- 14: Within-ALJ Predictors of Total Allowance Rate (Fully + Partially Favorable) 38
Table A- 15: Between-ALJ Predictors of Total Allowance Rate (Fully + Partially Favorable) 38
Table A- 16: Effect of Video Hearing on Allowance Decision Rate 41
Table A- 17: Effect for Title II and Title XVI Claims on Allowance Decision Rate 43
Table A- 18: Effect of New Evidence on Allowance Decision Rate 44
Table A- 19: Effect of Medical Expert Presence on Allowance Decision Rate 46
Table A- 20: Effect of Presence of Vocational Expert on Allowance Decision Rate 48
Table A- 21: Effect of Presence of Claimant Representative on Allowance Decision Rate 49
Table A- 22: Disposition Outcomes for Prototype and Non-Prototype States 51
Table A- 23: Trends in Number of Dispositions by Impairment (FY 2009 - FY 2011) 53
Table A- 24: Frequency of Federal Court Filings and Remands 54
Table A- 25: Number of Remands by Region 55
Table A- 26: Relative Frequency of Reasons for Remand Categories 57
Table A- 27: Appeals Council Remand Rates 58
Table A- 28: Frequency of Appeals Council Remand Reason Categories (FY 2009 - FY 2011) 59

LIST OF FIGURES

Figure A- 1: Distribution of ALJs' Dispositions by Month 6
Figure A- 2: Distribution of ALJs' Disposition by Year 8
Figure A- 3: Monthly Distribution of Average Disposition Frequencies 9
Figure A- 4: Disposition Outcome Rates for ALJs Close to Yearly Disposition Goal 10
Figure A- 5: Median Monthly Disposition Frequency for ALJs Close to Yearly Goal 11
Figure A- 6: Standard Deviation of Monthly Disposition Frequency for ALJs Close to Yearly Goal 12
Figure A- 7: Disposition Outcomes for ALJs Just Above the Yearly Goal with High Year-End Activity 13
Figure A- 8: Distribution of Allowance Rates (Fully or Partially Favorable) 14
Figure A- 9: Change in Allowance Rates Over Time 16
Figure A- 10: Predicted Allowance Rate for Video Hearings (FY 2011) 41
Figure A- 11: Predicted Allowance Rate by Claim Type (FY 2011) 43
Figure A- 12: Predicted Allowance Rates by Introduction of New Evidence (FY 2011) 45
Figure A- 13: Predicted Allowance Rate by Presence of Medical Expert 46
Figure A- 14: Predicted Allowance Rates by Presence of Vocational Expert (FY 2011) 48
Figure A- 15: Predicted Allowance Rate by Presence of Claimant Representative (FY 2011) ... 49
Figure A- 16: Comparison of Rulings for Non-Prototype and Prototype States 52
Figure A- 17: Comparison of Rulings Based on Impairment Type 53
Figure A- 18: Frequencies of Reason for Remand Categories 57

This appendix describes the results of statistical analyses conducted to understand variability in adjudication decisions issued by Social Security Administration administrative law judges (ALJs). The report is divided into five sections. The first section describes the data used for this analysis. The second section provides descriptive statistics on the data and describes the degree of variability in disposition frequencies. The third section provides a description of outlier ALJs (*i.e.*, those with atypical disposition frequencies or allowance rates). The fourth section examines correlates of allowance rates, and provides information on case characteristics that may explain some of the variability in adjudication decisions. The fifth section analyzes remands from the Appeals Council and the federal courts to identify common reasons for remand.

I. METHODOLOGY FOR ALJ ADJUDICATION ANALYSIS

Data on ALJ dispositions and related hearing information were obtained from the Social Security Administration's (SSA's) case processing management system (CPMS) management information data tables. Some variables that were not directly available in these tables were computed by staff of the Office of Electronic Services and Strategic Information, Division of Management Information and Analysis (OESSI/DMIA) from information in CPMS (*i.e.*, average number of claims, number of cases with mental impairment, number of cases with decisionwriter, number of cases with new evidence). Whether the case was processed in a prototype state[1] was determined by OESSI/DMIA staff based on the address of the ALJ.

Because data were compiled from multiple sources, there were a few instances in which the matching of data to individual ALJs was imperfect. For example, there were a few cases where the number of dispositions involving a decisionwriter slightly exceeded the total number

[1] In non-prototype states, a dissatisfied claimant can ask for reconsideration of the claim by a second team from the state disability determination service. *See* HAROLD KRENT & SCOTT MORRIS, ACHIEVING GREATER CONSISTENCY IN SOCIAL SECURITY DISABILITY ADJUDICATION: AN EMPIRICAL STUDY AND SUGGESTED REFORMS 6 (2013) [hereinafter ACHIEVING GREATER CONSISTENCY IN SOCIAL SECURITY ADJUDICATION]. Prototype states do not have a reconsideration step as part of the disability determination process. *See id.*

of dispositions. This might occur, for example, due to a change in an individual's job title affecting which cases were included in different reports. These situations were rare and the discrepancies are not expected to substantially alter the results. This issue only applies to variables that were compiled by OESSI/DMIA from multiple summary tables: average number of claims, number of cases with mental impairment, number of cases with decisionwriter, and number of cases with new evidence.

The data used for this analysis contained no personally identifiable information. Cases were identified only by a randomly generated pseudo-name. Further, data on dispositions contained no information about region or hearing office.

Monthly data on disposition frequency and allowance rates were available for each ALJ across fiscal years (FY) 2009 - 2011. Data on other case characteristics were provided on a yearly basis.

A. Sample

Data were obtained on 1,661 ALJs, who provided between 1 and 3,620 dispositions in a FY. Following the procedure used in a congressional response report by SSA's Office of the Inspector General,[2] we excluded from our analyses ALJs who had an unusually low number of dispositions (less than 200 disposition in a FY). These low frequencies might be due to new hires, retirement, part-time work or ALJs with other duties in addition to adjudicating cases. Because low activity ALJs were excluded on a yearly basis, an ALJ excluded in one year might be present in other years. One hundred fifty-two (9%) yearly data points were removed due to low activity.

[2] SOC. SEC. ADMIN. OFFICE OF THE INSPECTOR GENERAL, CONGRESSIONAL RESPONSE REPORT: OVERSIGHT OF ADMIN. LAW JUDGE WORKLOAD TRENDS, A-12-11-01138 (Feb. 2012).

The final data set used for the analysis consisted of 1,509 ALJs. Data were available from 1,129 ALJs in FY 2009, 1,256 ALJs in FY 2010 and 1,360 ALJs in FY 2011. Three years of data were available on 981 ALJs (65%), two years on 274 ALJs (18%) and one year on 254 ALJs (17%). Complete data for all 36 months was available for 895 ALJs (59%). ALJs provided on average 29.4 months of data, and only 7% had less than 12 months of data (minimum = 5 months). The data included a total of 44,396 monthly data points.

B. Variables

Monthly data contained frequency counts on the number of dispositions issued and the number of disposition that were Fully Favorable, Partially Favorable, Unfavorable and Dismissed. In order to avoid confounding of disposition outcomes with the number of dispositions in a month, outcomes were converted to a proportion of the total number of dispositions for that month. Because data on some of the characteristics examined were not recorded for dismissed cases, proportions were also examined for these characteristics as a fraction of the number of decisions (excluding dismissals).

The monthly data included the average time in days from hearing to disposition, the number of dispositions that were drafted by the ALJ versus non-ALJ staff (e.g., decisionwriters). The monthly data also included two variables expected to reflect the complexity of the cases: the number with a mental impairment and the average number of claims per claimant. All frequency counts were converted to proportion of the number of dispositions for the month.

Also included in the monthly data was the number of dispositions on cases from prototype states. An initial analysis confirmed that the majority of data points had either 0 or 100 percent of cases from prototype states, with a small number of instances with values close to

these extremes. This variable was treated as dichotomous in the analyses, indicating whether the majority of dispositions from a particular month were from prototype states.

Additional case characteristics were only available on a yearly basis. These characteristics included use of bench decision, use of on-the-record decision, presence of medical expert at hearing, presence of vocational expert at hearing, presence of claimant representative at hearing, use of video hearing, impairment type, case type, introduction of new evidence after the hearing, and use of a non-ALJ decisionwriter.

Several limitations of the use of CPMS data should be noted. Several variables (presence of medical or vocational expert, presence of claimant representative, use of video hearing) reflected the most recent hearing conducted. If a case involved multiple hearings, the presence of these features at earlier hearings is not represented, even though the testimony from these experts would have been part of the record reviewed by the ALJ. Whether a case involved a bench or on-the-record decision was entered manually by administrative staff and may be less accurate than other case processing data. The indicator for new evidence only tracks if the new evidence changed the ALJ decision, and does not reflect introduction of new evidence that did not change the decisions.

Missing data was found on some case characteristics because the values were not recorded in the CPMS system. This typically occurred because the case was dismissed without a full hearing. Data were missing on 20% of the dispositions for the variables presence of medical expert, presence of vocational expert, and use of video hearing. Data were missing from 29% of the dispositions for use of a decisionwriter.

Data were analyzed at three levels of aggregation: monthly, yearly, and total for each ALJ. Yearly statistics were computed by summing the monthly frequencies, and then computing

outcome rates as a proportion of the total number of dispositions. For non-categorical variables (average claims per claimant and time from hearing to disposition), monthly data were averaged to obtain yearly values. For dispositions in prototype or non-prototype states, yearly data were classified as prototype or non-prototype based on the majority of cases.

Similar procedures were used to compute summary statistics for each ALJ. Disposition outcome rates were computed by summing the monthly data for each ALJ and then dividing by the total number of dispositions for that ALJ. In order to adjust for differing amounts of data available, and to make the ALJ-level data comparable to the yearly data, the number of dispositions was expressed as the average dispositions per year, computed by dividing the total number of dispositions by the number of years of data provided for the ALJ.

C. Analyses

All statistical analyses were conducted using SAS/STAT software.[3] Because the study was largely exploratory, the alpha level was set as a moderately conservative level (.01) for determining statistical significance.

II. DESCRIPTIVE STATISTICS ON ALJ DISPOSITIONS

A. Number of Dispositions

1. Distribution of Disposition Frequency

On average, ALJs issued 45.5 dispositions per month; Median=44, standard deviation (SD)=20.1. 95% of the dispositions fell in the range 13-85. Only 1% of ALJs were above 104 dispositions per month and only 1% were below 6. There were a number of very high disposition frequencies (max=475). Excluding these outliers, the distribution of was approximately normal (*see* Figure A-1). Descriptive statistics on the monthly data are presented in Table A-1.

[3] Version 9.3, SAS System for Windows, SAS Institute Inc. (2011).

Figure A- 1: Distribution of ALJs' Dispositions by Month

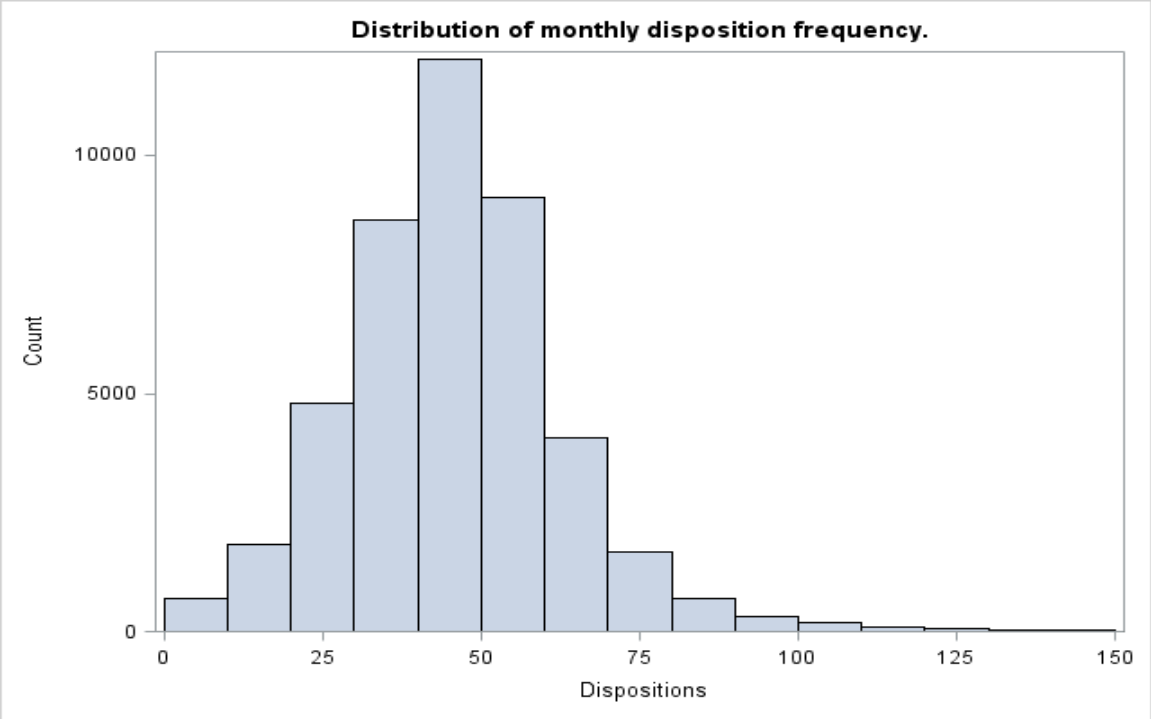


Table A- 1: Descriptive Statistics on Monthly ALJ Data

Variable	N	Mean	SD	Median	Minimum	Maximum
Number of Dispositions	44396	45.5	20.1	44	1	457
Days From Hearing to Disposition	44396	54.1	34.9	47	0	515
Average Number of Claims	44396	1.6	0.2	1.6	1.0	6.0
Percent Dismissal	44396	15%	11%	14%	0%	100%
Percent Fully Favorable	44396	51%	19%	50%	0%	100%
Percent Partially Favorable	44396	5%	6%	4%	0%	100%
Percent Unfavorable	44396	29%	17%	27%	0%	100%
Percent Drafted by ALJ	44396	8%	17%	0%	0%	100%
Percent Mental Claim	44396	42%	13%	42%	0%	100%
Prototype State	44396	0.31	0.46	0.0	0.0	1.0

Note: The data consist of 44,396 monthly datapoints obtained from 1509 ALJs across three years (FY 2009 -2011).

When aggregated yearly, ALJs issued an average 538.9 dispositions per year (Median=530, SD=180.5). 95% of the dispositions fell in the range 255-878. The majority of

ALJs (67%) issued 500 or more dispositions per year. Only 1% of ALJs were above 1,079 dispositions per year, and only 1% were below 221. There were a number of very high disposition frequencies (max=3,620). The distribution of yearly disposition frequencies (*see* Figure A-2) displays three notable characteristics. First, the distribution has a slight positive skew, resulting from (a) the exclusion of data with fewer than 200 dispositions, and (b) the long positive tail due to a small number of extremely high disposition frequencies. Second, the distribution is asymmetric in that there are more ALJs moderately below the mean (300-400 dispositions) than moderately above the mean (700-800 dispositions). Third, the distribution displays a large jump between 450 and 500 dispositions. This suggests that individuals in this range likely made an extra effort to meet the SSA target of 500 dispositions per year.

Descriptive statistics for all other variables calculated using yearly statistics were similar to those reported for monthly data (*see* Table A-2).

Figure A- 2: Distribution of ALJs' Disposition by Year

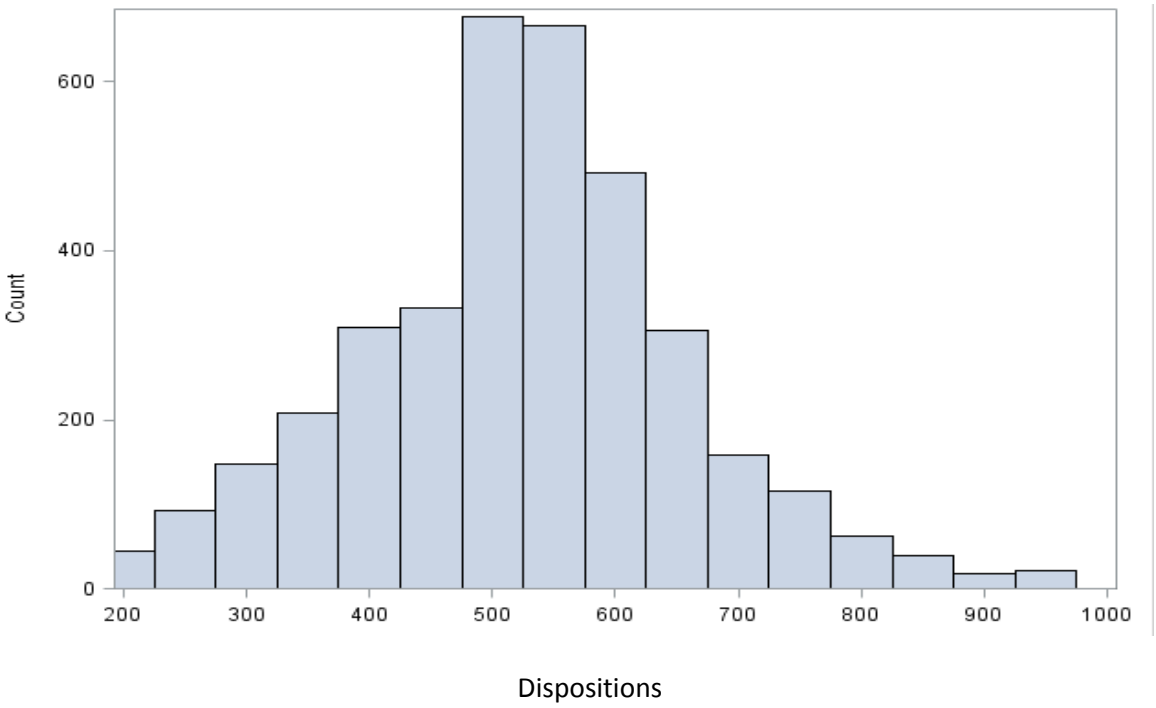


Table A- 2: Descriptive Statistics on Yearly ALJ Data

Variable	N	Mean	SD	Median	Minimum	Maximum
Number of Dispositions	3745	538.9	180.5	530.0	200.0	3620.0
Days From Hearing to Disposition	3745	54.0	28.6	49.1	4.3	397.1
Avg Number of Claims	3745	1.6	0.1	1.6	1.0	2.1
Percent Dismissal	3745	15%	7%	14%	2%	72%
Percent Fully Favorable	3745	51%	16%	50%	4%	97%
Percent Partially Favorable	3745	5%	4%	4%	0%	51%
Percent Unfavorable	3745	29%	13%	29%	0%	80%
Percent Drafted by ALJ	3745	10%	19%	0%	0%	100%
Percent Mental	3745	42%	9%	42%	11%	79%
Prototype State	3745	0.30	0.46	0.0	0.0	1.0

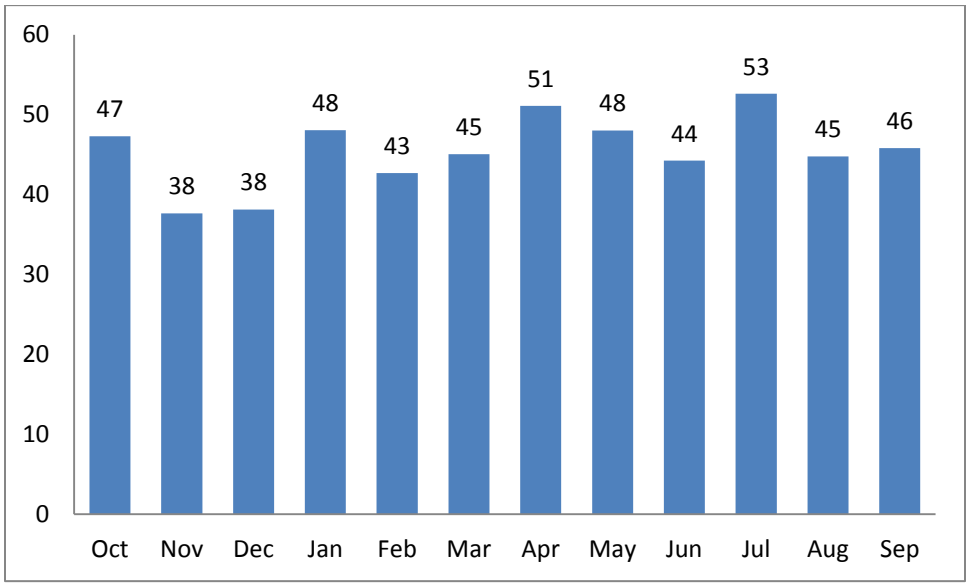
Note: The data consist of 3745 yearly datapoints obtained from 1509 ALJs across three years (FY 2009 – FY 2011).

2. Yearly and Monthly Trends

The average number of dispositions has not changed substantially across years, although the difference is statistically significant, $F(2,2234)=4.77$, $p<.01$. On average ALJs issued an average of 543.8 dispositions in FY 2009, 535.9 in FY 2010, and 537.5 in FY 2011. Variability in the number of dispositions decreased over the years SD=194.1 in 2009, SD=178.8 in 2010, and SD=170.1 in 2011. The SD in FY 2009 was significantly greater than in FY 2010, $F(1128, 1255)=1.18$, $p=.002$) and greater than in FY 2011, $F(1128, 1359)=1.30$, $p<.001$. The difference between FY 2010 and FY 2011 was not statistically significant, $F(1255, 1359)=1.11$, $p=.04$.

Monthly data suggest the presence of quarterly cycles in disposition frequency. The number of dispositions tended to be higher in January, April, July, and October. This pattern was less distinct during the spring months, with March and May disposition frequencies only slightly lower than April. February, June, August, November, and December tended to have low disposition frequencies (*see* Figure A-3).

Figure A- 3: Monthly Distribution of Average Disposition Frequencies

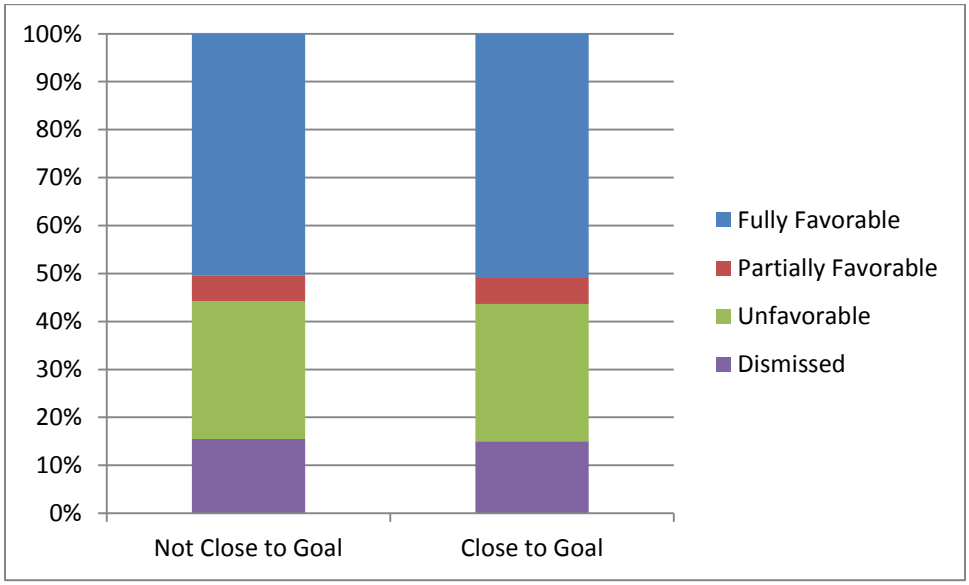


3. "Hurry Up" Dispositions

The distribution of yearly disposition frequency suggested a tendency of individuals close to the goal of 500 dispositions per year to put in extra effort to reach that goal. That is, the distribution displayed a lower number of dispositions in the range of 450-449 than would be expected from a normal distribution, and an excess of dispositions in the range of 500-549. This suggests that ALJs who were close to the goal may have made an extra effort to increase their productivity.

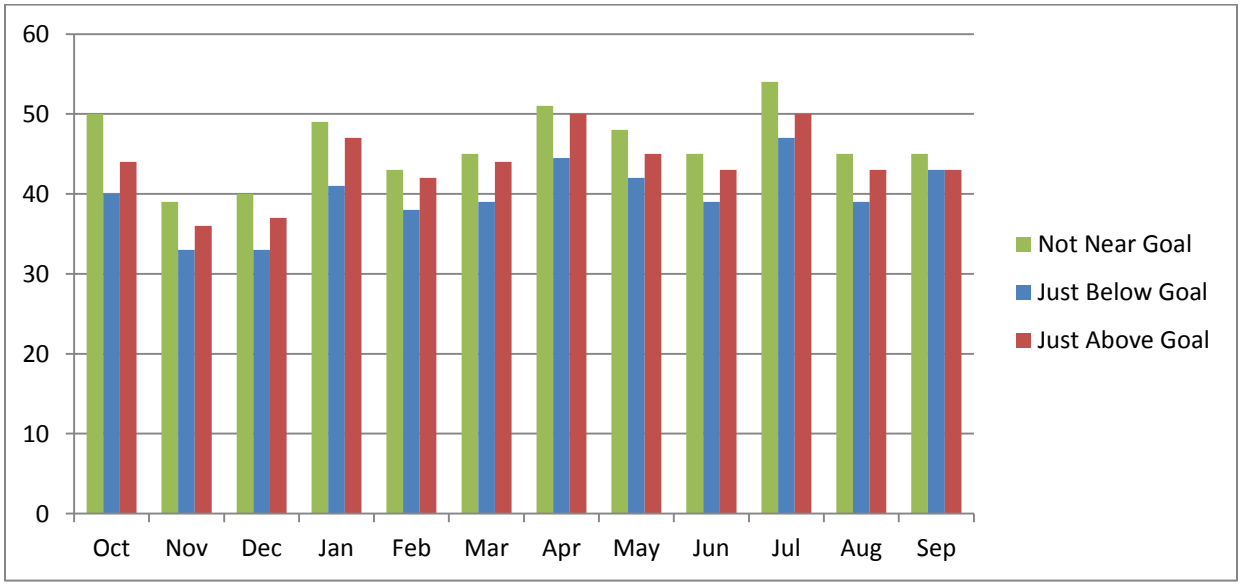
An analysis was conducted to determine whether this push to meet the goal led to changes in disposition outcomes, that is, whether ALJs issue more allowance dispositions if they were near the goal. Those near the goal (450-549 dispositions per year) did not show a different pattern of allowance rates, when compared to other ALJs (*see* Figure A-4).

Figure A- 4: Disposition Outcome Rates for ALJs Close to Yearly Disposition Goal



If ALJs were increasing their activity to reach the goal, one would expect the disposition rate to increase toward the end of the FY. An examination of monthly disposition frequencies provides some support for this hypothesis. ALJs just above the goal show a pattern of monthly disposition frequencies similar to those not near the goal. ALJs just below the goal generally show lower number of dispositions than the other groups, except for the last month of the FY (September). Thus, for the ALJs who ended the year just below the goal, there is evidence of an increase in activity in the final month (*see* Figure A-5).

Figure A- 5: Median Monthly Disposition Frequency for ALJs Close to Yearly Goal

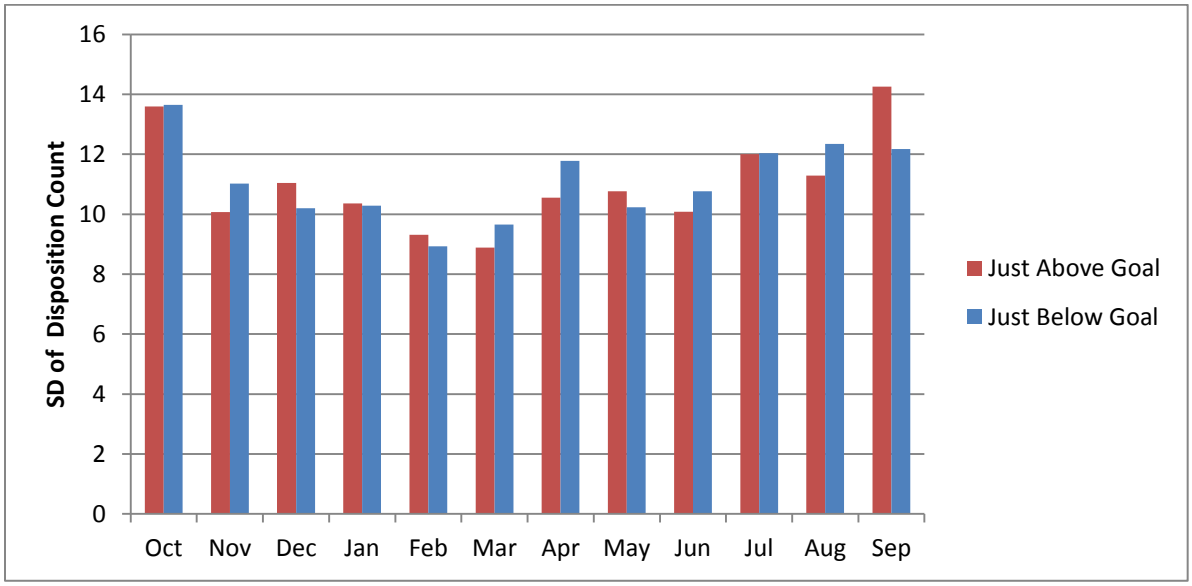


Note: Just Below Goal = 400-499 dispositions; Just Above Goal = 500-549 dispositions: Not Near Goal = less than 400 or greater than 549 dispositions.

Although ALJs just over the goal did not show an increased mean number of dispositions at the end of the FY, they did show an increase in the variability of activity at year end. Figure A-6 shows the SD of disposition frequency for those just above the yearly goal (500-549 dispositions) and just below the yearly goal (450-499 dispositions). For ALJs just over the goal, the SD was 14.3 in September. This value is substantially greater than the SD in other months,

with the exception of October, which showed high variability for both groups. The higher variability in September may be due to some of the individuals increasing their activity, although this number may not have been sufficient to raise the overall mean.

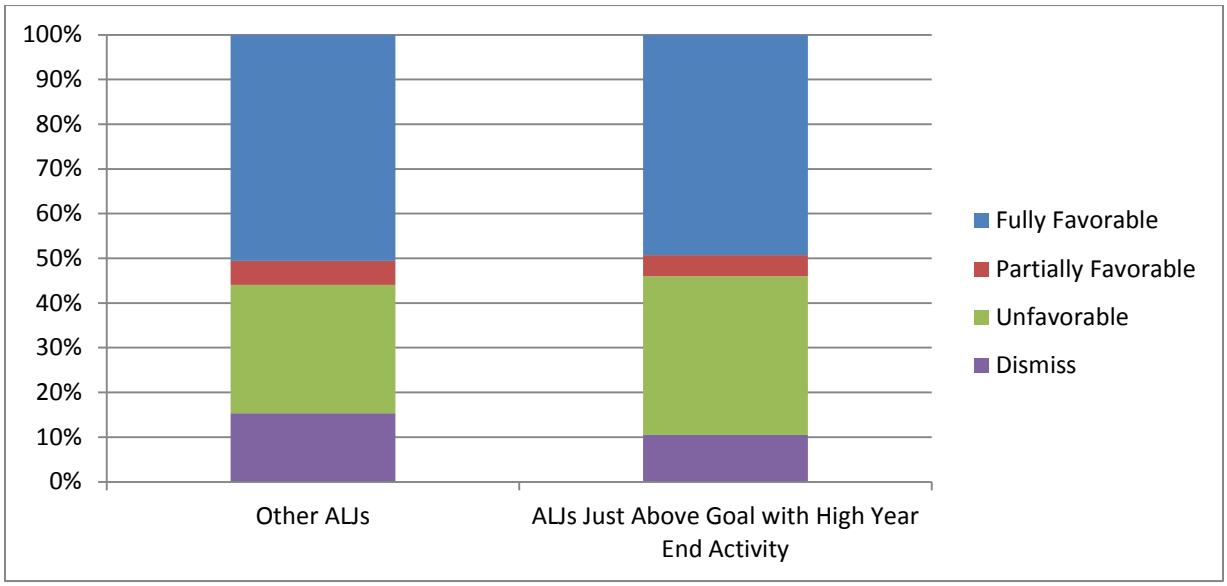
Figure A- 6: Standard Deviation of Monthly Disposition Frequency for ALJs Close to Yearly Goal



To investigate whether the year-end increase in activity affected disposition outcomes, we examined ALJs who were just over the goal (500-549 dispositions), and who were greater than 1 SD above the mean number of dispositions for September. These individuals had a similar Fully Favorable rate (49% vs. 51%) to other ALJs, but had higher Unfavorable rates (35% vs. 29%, $t[44394]=2.94$, $p<.01$), and lower dismissal rates (11% vs. 15%, $t[44394]=3.32$, $p<.01$), as shown in Figure A-7. These results do not support the hypothesis that ALJs near the goal are issuing more allowance dispositions in order to increase the number of dispositions near the end of the year.

To further explore the year-end rush, we examined ALJs who were just over the yearly goal (500-549 dispositions per year) and who had high activity in September (greater than 1 SD above the September average number of dispositions). These ALJs had September allowance rates that were slightly lower than in other months (54% in September, 58% in other months). However, a similar differential was observed among other ALJs (55% in September, 57% in other months). Thus, the data provide no evidence that allowance rates were increasing at year-end among ALJs who were hurrying to meet the yearly goal.

Figure A- 7: Disposition Outcomes for ALJs Just Above the Yearly Goal with High Year-End Activity



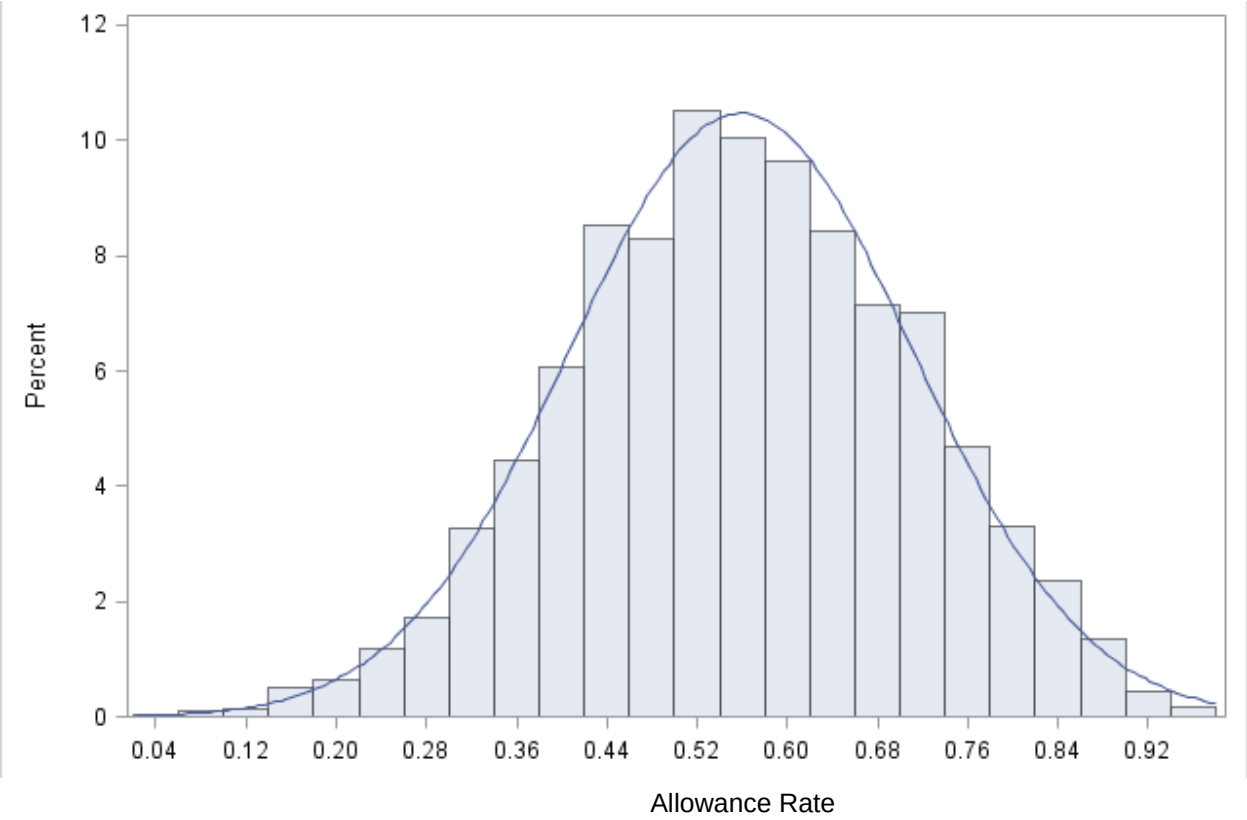
B. Allowance Rates

1. Distribution of Allowance Rates

The distribution of yearly allowance disposition rates (Fully Favorable plus Partially Favorable) was approximately normal, with a mean of 56% (standard error (SE)=0.25%) and a SD of 15%. As shown in Figure A-8, the distribution covered a wide range of allowance rates, with 95% of the rates falling between 26% and 85%. The lowest allowance rate was 4% and the

highest allowance rate was 98%. A similar pattern was found when considering only Fully Favorable dispositions, where 95% of the rates were between 20% and 82%.

Figure A- 8: Distribution of Allowance Rates (Fully or Partially Favorable)



2. Changes in Allowance Rates Over Time

Allowance rates have been declining over time. The mean ALJ allowance rate dropped from 59% in FY 2009 to 58% in FY 2010 and 53% in FY 2011. The variability of allowance rates has been stable, with a SD of 16% in FY 2009 and 15% in FY 2010 and FY 2011. The variability of allowance rates has been stable, with a SD of 16% in FY 2009 and 15% in FY 2010 and FY 2011.

Table A- 3: Yearly Allowance Rate (Fully + Partially Favorable) Including Senior Attorney Decisions

FY	Senior Attorney Decisions[4]	ALJ Dispositions			Combined ALJ + Senior Attorney		
		Favorable	Total	%	Favorable	Total	%
2009	36,366	361278	613921	59%	397,644	650,287	61%
2010	54,186	387667	673050	58%	441,853	727,236	61%
2011	53,253	388011	731051	53%	441,264	784,304	56%
2012[5]	37,423	373343	783061	48%	410,766	820,484	50%

The decline in allowance rates occurred across the entire distribution of allowance rates, although there tended to be less change at the lower end of the distribution. Table A-4 and Figure A-9 show the yearly trend for the 10th percentile, median, and 90th percentile of the allowance rate for ALJ-issued decisions (not including senior attorneys). The percentile statistics are defined such that 10% of ALJs fall below the 10th percentile, and 10% fall above the 90th percentile. For the median and 90th percentile, allowance rates dropped by 2-3% FY 2009 - FY 2010 and by 5% FY 2010 - FY 2011. At the lower end of the distribution, there was no change in the 10th percentile FY 2009 - FY 2010, and a change of 4% FY 2010 - FY 2011.

Corresponding to the drop in the mean allowance rate, the number of ALJ with extreme rates has been declining. As shown in Table A-5, the percentage of ALJs with allowance rates over 80% has dropped by half, from 9% in 2009 to 4% in 2011. At the other end of the distribution, the percentage of ALJs with extremely low allowance rates has been more stable (*i.e.*, the percentage with allowance rates under 30% ranged from 5% in FY 2009 to 6% in FY 2011).

[4] Data on senior attorney dispositions were provided by the Soc. Sec. Admin, Office of Appellate Operations. *See* E-mail from Rainbow Forbes, Appeals Officer, Soc. Sec. Admin. Office of Appellate Operations, to Harold J. Krent, Dean and Professor of Law, IIT Chicago-Kent College of Law (February 11, 2013) (on file with author).
[5] FY 2012 ALJ disposition statistics were obtained from Data.gov, ALJ Disposition Data, https://explore.data.gov/Social-Insurance-and-Human-Services/ALJ-Disposition-Data/zdyy-hq2m (last visited Feb. 21, 2013).

Table A- 4: Change in Allowance Rates (FY 2009 - FY 2011)

Percentile	FY 2009	FY 2010	FY 2011
10th %tile	38%	38%	34%
Median	59%	57%	52%
90th %tile	79%	76%	72%

Figure A- 9: Change in Allowance Rates Over Time

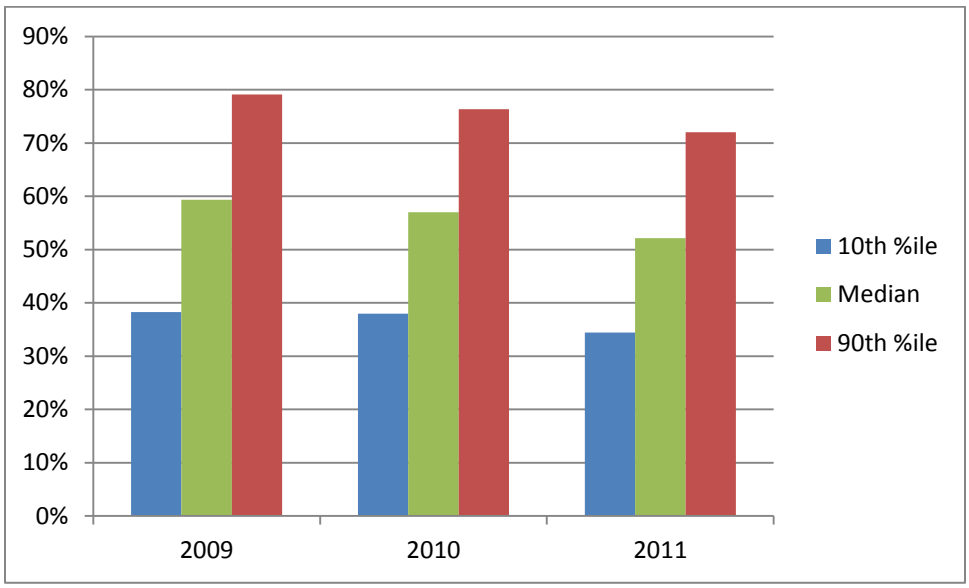


Table A- 5: Percent of ALJs with Extreme Allowance Rates (Fully Favorable + Partially Favorable)

FY	Percent of ALJs below 20% Favorable	Percent of ALJs below 30% Favorable	Percent of ALJs above 70% Favorable	Percent of ALJs above 80% Favorable
2009	1%	5%	25%	9%
2010	1%	4%	21%	7%
2011	2%	6%	13%	4%

Note: The values in this table were obtained by interpolating between values at the 1st, 5, 10th, 75th, 90th, 95th, and 99th percentiles.

C. Time to Disposition

Time to disposition was defined as the number of days between the hearing and issuing the disposition. On average it took ALJs 54.1 days after a hearing to issue a disposition (Median=47, SD=34.9). 95% of the monthly time-to-disposition values were in the range 15 to 116 days. Only 1% of ALJs took over 170 days and only 1% took under 7 days.

Approximately 10% of the time-to-disposition values were above 90 days and 10% were below 21 days. About half (53%) of the ALJs had at least one month with an average over 90 days and over half (61%) had at least one month with an average less than 21 days. Many ALJs had 1-6 months over 90 (35%) or under 21 days (45%). Only 2% had more than 24 months over 90 days, and only 1% had more than 24 months under 21 days.

About 8% of the ALJs had at least one month with average time-to-disposition over 180 days. The majority of these (6% of ALJs) had 1-3 months with high time to disposition. Six ALJs had over 12 months with an average time to disposition over 180 days, and one was over 180 days in each of 35 months. At the other end of the distribution, 17% of ALJs had at least one month under 7 days, the majority of these (16% of ALJs) having 1-3 months. Only three ALJs have more than 12 months with average time-to-disposition under 7 days.

D. Variance Decomposition Analysis

Because outcomes can be defined at multiple levels of aggregation, such as monthly, yearly, or ALJ total allowance rates, it is useful to examine the extent of variability at each level of analysis.[6] Such analyses can direct attention toward levels of analysis where there is sufficient variability to merit further investigation.

[6] J. HOX, MULTILEVEL ANALYSIS: TECHNIQUES AND APPLICATIONS (Routledge 2010) [hereinafter MULTILEVEL ANALYSIS].

Variance in adjudication outcomes will be influenced by factors that vary across cases, such as the merits of the claims and the quality of evidence provided. Thus, outcomes will be expected to vary across time within an ALJ due to the particular collection of cases reviewed in a particular month. These case-level characteristics will also create some variability across ALJs due to differences in the cases reviewed by each ALJ. Importantly, it is possible to predict the level of between-ALJ variance that is to be expected due to the amount of variance observed across months.[7] If the observed between-ALJ variance is greater than the amount predicted from the within-ALJ variance, then the data provide justification to investigate ALJs as a source of the variability. That is, some ALJs may evaluate claims in an idiosyncratic fashion, producing consistently higher or lower allowance rates relative to other ALJs.

Analysis of data with multiple levels of aggregation typically begins by calculating the proportion of the total variance that can be attributed to within- and between-group components. This is operationalized in terms of the intraclass correlation coefficient (ICC), which is defined as

$$ICC = \frac{\sigma^2_{Between}}{\sigma^2_{Between} + \sigma^2_{Within}},$$

where $\sigma^2_{Between}$ and σ^2_{Within} are estimates of the unique within-ALJ and between-ALJ variance, respectively. The ICC can be interpreted as the proportion of the observed variance that can be attributed to between-ALJ factors. Values close to zero indicate that most of the variance occurs within-ALJs, while values close to 1.0 indicate that most of the variance is at the ALJ level.

Estimates of the two variance components were obtained from a general linear mixed model analysis with monthly data nested within ALJ. An unconditional model was specified with no predictor variables, but a randomly varying intercept that captured the mean for each ALJ.

[7] *Id.*

The analysis was conducted using the MIXED procedure in the SAS software package. A separate variance analysis was conducted for the number of dispositions and the proportion of dispositions falling into each outcome category. The results are summarized in Table A-6.

Table A- 6: Intraclass Correlation Coefficients (ICC) for Within- and Between-ALJ Variance

Variable	ICC
Number of Dispositions	.45
Fully Favorable Rate	.60
Partially Favorable Rate	.38
Unfavorable Rate	.55
Dismissal Rate	.27
Time to Disposition	.55

The finding of moderate ICCs on most of the variables suggests that both within-ALJ and between-ALJ factors contribute to disparities in adjudication outcomes. All of the variables showed non-trivial between-ALJ variance, although ALJs had a bigger impact on some outcomes than others. Nearly half of the variance in the number of dispositions was found to occur between ALJs, and greater than 50% of the variance was at the ALJ level for the proportion of Fully Favorable and Unfavorable decisions. In contrast, the lower ICC for dismissals suggests that these are more strongly determined by situational factors than the other outcomes.

E. Identifying Outliers

From a purely statistical perspective, an outlier is defined as a score that is atypical given the overall distribution.[8] This approach to identifying outliers does not provide information on why a score is atypical. Instead, outliers reflect observations that warrant further attention to determine why the unusual event has occurred.

[8] J. Cohen, P. Cohen, L. S. Aiken & S.G. West, APPLIED MULTIPLE REGRESSION/CORRELATION FOR THE BEHAVIORAL SCIENCES (3rd ed. (2003) [hereinafter APPLIED MULTIPLE REGRESSION/CORRELATION].

Outliers are defined relative to an assumed distribution of scores. Some variability is expected due to normal fluctuations in the process. For example, differences in allowance rates are expected due to the specifics of individual cases, which will be distributed somewhat unevenly across ALJs just due to chance. The variability in a distribution is quantified by the SD.

Many variables are found to have a "normal" distribution, where scores are distributed symmetrically around the mean, and the density of scores decreases with distance from the mean, forming a bell-shaped curve. An example of a normal distribution is depicted by the smooth curve in Figure A-8, which shows that allowance rates had an approximately normal distribution.

If a distribution is normal, the majority of scores will fall within 2 SD of the mean. Only about 5% of scores will be more than 2 SD from the mean (with 2.5% more at the high end of the distribution, and 2.5% at the low end), and only 1% of scores will be more than 3 SD from the mean. Because only a small fraction of observations will naturally fall outside this range, scores outside these bounds are not likely to have occurred due to chance, and are therefore likely to have been produced by some process that differs from the rest of the distribution.

With any decision rule, there is a risk of false positive and false negative decisions. A false positive would occur if an observation is classified as an outlier when it occurred simply due to chance. Identifying outliers using a 2 SD rule will have a 5% false positive rate, because 5% of scores in a normal distribution are expected to fall outside 2 SD from the mean. Setting a more stringent criterion will tend to reduce the false positive error rate.

A false negative error would occur if an observation falls within the 2 SD interval, yet was actually generated by a different distribution. That is, a hypothetical ALJ who is applying overly lenient standards might nevertheless still produce an allowance rate within the 2 SD bounds, simply due to chance. The probability of a false negative error cannot be known

precisely in this context, but as the criterion for identifying an outlier becomes more stringent, the probability of a false negative error increases.

The decision of where to put the cutoff for defining outliers depends on the desired balance of false positive and false negative errors. Setting the cutoff too low will flag many observations as outliers when they are simply due to chance; while setting to cutoff too high will fail to identify many true outliers. It is common for statisticians to define outliers as observations fall 2 or 3 SD above the mean.[9]

Establishing a cutoff is complicated by changes in the distribution over time. Although the average number of dispositions was relatively stable from FY 2009 to FY 2011, SD has declined. Mean allowance rates declined over the three years of the study, suggesting that standards for defining an outlier should be tied to yearly statistics. Table A-7 provides the cutoffs corresponding to 2 and 3 SD above and below the mean for yearly disposition frequency and allowance rates.

Because the variables did not exactly reflect a normal distribution, the 2 or 3 SD rules need to be applied with caution. For the number of dispositions, the presence of large positive outliers will slightly inflate both the mean and the SD, so that the cutoff established using these rules may be too high. Further, because the allowance rate is a percentage, it is constrained by a maximum value of 100%, which may cause compression of values near the upper limit. For this reason, Table A-7 also includes percentile values, which do not rely on an assumed shape of the distribution. The percentile is the value of the variable such that the same percent of values fall below that value. For example, the 99th percentile of the number of dispositions was 959, indicating that 99% of ALJs had fewer than 959 dispositions, and only 1% had more than this value.

[9] *See* APPLIED MULTIPLE REGRESSION/CORRELATION, *supra* note 8.

Applying a 2 SD rule to allowance rates would classify outliers as ALJs with allowance rates below 23% (approximately 37, or 3% ALJs in 2011) and ALJs with allowance rates above 82% (approximately 30, or 2% ALJs). Similarly, 3% of ALJs (approximately 38) had disposition frequencies over 878.

Table A- 7: Extreme Values for Number of Dispositions and Allowance Rates (Fully + Partially Favorable), FY 2011

	Mean	SD	3 SD Below Mean	2 SD Below Mean	2 SD Above Mean	3 SD Above Mean	1st %tile	5th %tile	95th %tile	99th %tile
Number of Dispositions	537.5	170.1	-	-	878	1048	-	-	780	959
Allowance Rate	53%	15%	9%	23%	82%	96%	18%	29%	77%	88%

Note: Because data from ALJs with fewer than 200 dispositions were excluded from the analysis, it was not possible to use the distribution to identify outliers corresponding to low number of dispositions.

Another factor that can be used to identify outliers is the consistency in outcomes over time. An ALJ who is in the top or bottom 1% one year might be the result of chance variation. However, if the same individual appears in the top or bottom 1% in multiple years, this is more likely to be due to something unique to that individual. Therefore, special attention should be paid to ALJs who have consistent extreme scores in multiple years. Table A-8 reports the number of ALJs appearing for multiple years in the top and bottom 1% of the disposition and allowance rates.

Table A- 8: Number of ALJs with Multiple Years in Top and Bottom 1% of the Disposition Frequency and Allowance Rate Distributions (FY 2009 - FY 2011)

Years in Top or Bottom 1%	Number of ALJs With Disposition Frequencies in Top 1%	Number of ALJs With Favorable Rates (Full + Partial) in Top/Bottom 1%	
	Top 1%	Bottom 1%	Top 1%
At least 1 year	22	24	25
At least 2 years	10	10	10
All 3 years	6	5	3

Based on these considerations, we recommend that ALJs who are more than 2 SD above or below the mean in two consecutive years should be flagged for further examination. The mean and SD used to define these cutoffs should be adjusted each year based on the most recent data available.

Ideally, outliers would be defined relative to a predictive model that accounts for important characteristics of the portfolio of cases reviewed by an ALJ, and other relevant variables. This appendix describes an initial attempt to identify correlates of adjudication outcomes, but additional work is needed to build a model that incorporates all relevant factors.

III. ANALYSIS OF OUTLIER ALJS

A. Outliers Based on Disposition Frequency

We examined whether individuals with atypically high disposition frequencies in a FY tended to issue different dispositions than less extreme ALJs. The top 1% of yearly disposition frequencies (1,079 or more dispositions) consisted of 38 data points. These comprised 22 ALJs, six of whom were outliers in all 3 years, four were outliers in 2 years, and 12 were outliers in a single year.

Comparison of the top 1% to other ALJs indicates that outliers issued fewer Partially Favorable (2% vs. 5%), fewer Unfavorable (16% vs. 29%), and more dismissed dispositions (23% vs. 15%). The difference in Fully Favorable rates (58% vs. 50%) was not statistically significant.

ALJs in the top 1% of the disposition frequency did not differ from other ALJs in the average number of days to reach a disposition. ALJ with high disposition frequency were less likely to use a decisionwriter (53%) than other ALJs (63%). They tended hear cases that were less complex, that is, they heard cases where the average number of claims was slightly lower (1.54 claims on average vs. 1.58) and they were less likely to have cases involving mental impairments (36% vs. 42%).

Some differences were found between ALJs who were in the top 1% for multiple years. There was a trend of higher Fully Favorable rates among those with 2 or 3 years in the top 1% (65%) compared to those in the top 1% for 0 or 1 year (50% Fully Favorable). Those with 2 or 3 years in the top 1% had lower Partially Favorable rates. Those with 1 or 2 years in top 1% had lower Unfavorable rates. Those with 1 year in the top 1% had the highest dismissal rate, and those with either 0 or 2 years had the lowest dismissal rate. No significant differences were found on other variables.

B. Outliers Based on Allowance Rates

Outliers can also be identified in terms of the decision outcomes. For this analysis, we combined Fully Favorable and Partially Favorable decisions to identify those ALJs who tended to have atypically high or low grant rates.

The top 1% of allowance rates (those issuing Fully or Partially Favorable decisions in over 89% of cases) consisted of 25 ALJs, 15 with 1 year in the top 1%, seven with 2 years in the

top 1%, and three with 3 years in the top 1%. Groups defined by number of years in the top 1% of allowance rate tended to differ in the number of dispositions issued, $F(3, 1505)=9.29$, $p<.01$. ALJs with 3 years of high allowance rates tended to issue a high number of dispositions per year (M=866, SD=360), as did ALJs with 2 years of high allowance rates (M=724, SD=338). Those with a single year in the top 1% issued a similar number of dispositions (M=557, SD=220) to those who were never in the top 1% (M=537, SD=177).

The bottom 1% of allowance rates (those under 19%) consisted of 24 ALJs, 14 with one year in the bottom 1%, five with 2 years in the bottom 1%, and five with 3 years in the bottom 1%. There was no consistent pattern in the number of dispositions issued. Those with 1 or 3 years on the bottom 1% issued fewer dispositions on average than those with 0 or 2 years. The number of dispositions tended to be below average for ALJs with one year in the bottom 1% (M=405, SD=158) and those with 3 years in the bottom 1% (M=417,SD=87).141 However, those with 2 years in the bottom 1% (M=530, SD=81) were similar to those who were never in the bottom 1% (M=541, SD=181).

C. Outliers based on Time-to-Disposition

There were 26 ALJs with more than 3 months of average time-to-disposition over 180 days. These individuals tended to have a substantially lower number of dispositions (28.8 dispositions per month, SD=10.4) compared to other ALJs, (44.5 dispositions per month, SD=13.5), $t(1507)=5.88$, $p<.0001$. They tended to have lower Fully Favorable rates (40% vs. 50%), $t(1507)=3.30$, $p<.01$, and higher Unfavorable rates (37% vs. 29%), $t(1507)=-2.99$, $p<.01$.

IV. Correlates of Disposition Outcomes

A. Analysis

For the following analyses, disposition outcomes were classified as Fully Favorable, Partially Favorable, Unfavorable, or Dismissed. For those variables where monthly ALJ data

were available, we examined both within- and between-ALJ relationships. Within-ALJ relationships represent differences in how the same ALJ responds to different types of cases. Between-ALJ relationships represent the tendency of ALJs assigned to a higher or lower proportion of certain types of cases to reach particular decisions.

Two sets of analyses were conducted due to the time frame for reporting different sets of variables. Data on disposition outcomes and some predictor variables was reported monthly, while data on other variables was only available on a yearly basis. The monthly data provides a more fine-tuned examination of the stability of outcomes over time, but this analysis was only available for disposition outcomes and a limited number of predictor variables. A second analysis was able to include a larger set of variables using yearly data. These analyses are reported separately below.

B. Correlates of Monthly Disposition Outcomes

Monthly data were available for the number of dispositions, dispositions outcomes, time from hearing to disposition, use of a decisionwriter, and whether the case was from a prototype state. In addition, monthly data were provided on two indicators of case complexity: percent of cases with mental impairments, and average number of claims per claimant.

Correlations among disposition outcomes and related variables were computed both within- and between-ALJs. Between-ALJs were computed using the mean for each ALJ as the unit of analysis. Within-ALJ correlations were computed using within-ALJ residual scores (*i.e.*, the mean for a month minus the ALJ average). The correlation matrix is presented in Table A-9 on the following page.

Table A- 9: Within- and Between-ALJ Correlations

Variable	1	2	3	4	5	6	7	8	9	10
1. Dispositions per month	--	0.15	-0.14	-0.15	0.02	-0.16	0.04	-0.07	0.06	0.04
2. Percent Fully Favorable	-0.05	--	-0.31	-0.88	-0.43	-0.18	-0.11	0.29	-0.29	-0.10
3. Percent Partially Favorable	-0.03	-0.21	--	0.10	-0.03	0.13	0.04	0.01	0.05	-0.01
4. Percent Unfavorable	0.07	-0.67	-0.07	--	0.02	0.14	-0.09	-0.13	0.31	-0.04
5. Percent Dismissals	-0.01	-0.42	-0.14	-0.30	--	0.07	0.41	-0.42	0.03	0.33
6. Hearing Days Average Mean	0.04	-0.19	0.05	0.24	-0.06	--	0.02	-0.16	0.13	-0.00
7. Percent Drafted byALJ	-0.07	0.20	-0.01	-0.20	-0.01	-0.09	--	-0.25	-0.10	0.14
8. Percent Mental	0.00	0.13	0.02	0.21	-0.44	0.08	-0.02	--	-0.05	-0.22
9. Average Claims Mean	-0.03	-0.10	0.04	0.08	0.02	0.04	-0.03	0.03	--	-0.04
10. Prototype State	0.01	-0.06	-0.02	0.02	0.07	0.03	-0.01	-0.06	0.01	--

Note: Within-ALJ correlations are displayed below the diagonal, between-ALJ correlations are displayed above the diagonal. Within ALJ correlations greater than .02 and between-ALJ correlations above .07 are statistically significant (α = .01).

The number of dispositions issued in a month was not strongly correlated with any of the disposition outcome rates. Within-ALJ correlations were very small (-.07 to .07). Weak but statistically significant correlations in the between-ALJ data indicate a very slight positive relationship between number of dispositions and Fully Favorable rate (.15, p<.001), and very slight negative correlations of number of dispositions with Partially Favorable (-.14, p<.001) and Unfavorable (-.15, p<.001) rates.

Average time to disposition had a weak but statistically significant correlations with the percent Fully Favorable, both within-ALJs (-.19, p<.001) and between-ALJs (-.18, p<.001). Average time to disposition also correlated positively with percent Unfavorable within-ALJ (.24, p<.001), but the correlation was much weaker between-ALJs (.14, p<.001). There was also a weak between-ALJ correlation with percent Partially Favorable (.13, p<.001).

Percent of dispositions drafted by the ALJ had weak but statistically significant within-ALJ correlations with percent Fully Favorable (.20, p<.001) and percent Unfavorable (-.20, p<.001). A substantial between-ALJ correlation was found between disposition writing and percent dismissed (-.41, p<.001).

Percent of mental claims showed a strong correlation with percent dismissal, both within-ALJ (-.44, p<.001) and between-ALJs (.-.42, p<.001). Percent of mental claims also showed a weak correlation with percent Unfavorable within-ALJ (.21, p<.001), but not between-ALJ (-.13, p<.001). Conversely, percent of mental claims correlated somewhat with percent Fully Favorable between-ALJ (.29, p<.001), but not within-ALJ (.13, p<.001).

Average number of claims did not have any meaningful correlations with disposition outcomes within-ALJs. Between-ALJ correlations were found with percent Fully Favorable (-.29, p<.001), percent Unfavorable (.31, p<.001).

Whether case was from a prototype state could only be examined as a between-ALJ variable because it did not vary within-ALJs. Prototype state claims were found to correlate with percent dismissed (.33, $p<.001$), indicating that ALJs from prototype states were more likely to dismiss claims.

ALJs from prototype states issued a similar number of dispositions as those from non-prototype states. Prototype state ALJs issued Fully Favorable dispositions at a slightly higher rate (51% vs. 48%), and had a higher dismissal rate (18% vs. 13%).

C. Correlates of Yearly Disposition Outcomes

Yearly data were available on a greater number of case characteristics, including: proportion of claims with mental impairment, average number of claims per claimant, prototype state, proportion of bench decisions, proportion of on-the-record (OTR) decisions, type of claim (title II, title XVI or combination), percent of claims with a medical expert present, proportion of claims with a vocational expert present, proportion of claims a with claimant representative present, proportion of hearings using video, proportion of decisions drafted by the ALJ (as opposed to a decisionwriter), and proportion of cases where new evidence was introduced. Descriptive statistics are provided in Table A-10.

Table A- 10: Descriptive Statistics on Yearly Disposition Outcomes and Case Characteristics

Variable	Mean	Median	SD	Minimum	Maximum
Dispositions Per Year	539	530	181	200	3620
Percent Fully Favorable	51%	50%	16%	4%	97%
Percent Partially Favorable	5%	4%	4%	0%	51%
Percent Unfavorable	29%	29%	13%	0%	80%
Percent Dismissal	15%	14%	7%	2%	72%
Time to Disposition	53.96	49.08	28.58	4.33	397.08
Percent Mental	42%	42%	9%	11%	79%
Average Claims Mean	1.58	1.58	.11	1.03	2.05
Percent Prototype State	30%	0%	46%	0%	100%
Percent Bench	5%	0%	11%	0%	81%
Percent OTR	9%	5%	10%	0%	84%
Percent Title2	31%	30%	9%	4%	99%
Percent Title16	28%	27%	8%	0%	68%
Percent Combo	41%	42%	6%	1%	57%
Percent Medical Expert	15%	2%	26%	0%	100%
Percent Vocational Expert	77%	90%	30%	0%	100%
Percent Representative	77%	79%	10%	9%	99%
Percent Video Hearing	19%	9%	24%	0%	100%
Percent ALJ Draft	10%	0%	19%	0%	100%
Percent New Evidence	23%	20%	18%	0%	79%

Correlations among these variables and disposition outcomes are summarized in Table A-11. The table reports both within- and between-ALJ correlations. Between-ALJs were computed using the mean for each ALJ as the unit of analysis. Within-ALJ correlations were computed using within-ALJ residual scores (*i.e.*, the mean for a month minus the ALJ average).

The yearly number of dispositions was not strongly correlated with any of the case characteristics. None of the within- or between-ALJ correlations exceeded .20. The largest correlations were with proportion of OTR decisions (-.19) and proportion of cases with a vocational expert present (.17).

Within-ALJs, years with a higher proportion of bench decisions tended to correlate with slightly higher Fully Favorable rates (.22) and slightly lower Unfavorable rates (-.19).

The between-ALJ correlations indicated that ALJs who issued more OTR decisions tended to have higher Fully Favorable rates (.41) and lower Unfavorable rates (-.40). A similar pattern was observed within ALJs, with higher Fully Favorable rates (.34) and lower Unfavorable rates (-.34) in years with a higher proportion of OTR decisions.

ALJs who heard a higher proportion of cases with the claimant representative present tended to have higher Fully Favorable rates (.30). A similar correlation (.28) was found within ALJs.

ALJs who heard a higher proportion of cases with a vocational expert present tended to have a higher Unfavorable rate (.22). All within-ALJ correlations between presence of a vocational expert and disposition outcomes were close to zero.

Within ALJs, type of claim was correlated with allowance rates. In years with more title II claims, ALJs tended to have higher Fully Favorable rates (.30) and lower Unfavorable rates (-.40). In contrast, in years with more title XVI claims, ALJs tended to have lower Fully Favorable rates (-.31) and higher Unfavorable rates (.34). This trend was similar, but much weaker between ALJs. ALJs who heard a higher proportion of title II claims showed a slight tendency toward higher Fully Favorable (.19) and lower Unfavorable rates (-.19). Table A-11 on the following page presents yearly within-ALJ and between-ALJ correlations among yearly disposition outcomes and case characteristics.

Table A- 11: Within- and Between-ALJ Correlations Among Yearly Disposition Outcomes and Case Characteristics

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. # Dispositions	--	0.15	-0.06	-0.17	0.03	0.03	-0.12	0.13	0.03	0.07	0.17	0.05	-0.03	-0.04	-0.06	-0.19	-0.03	0.03	0.02	-0.06
2. % Fully Favorable	-0.23	--	-0.31	-0.88	-0.43	-0.18	0.29	-0.29	-0.10	0.17	0.41	0.19	-0.21	0.02	0.10	-0.14	0.30	0.04	0.05	0.13
3. % Partially Favorable	-0.08	-0.15	--	0.10	-0.03	0.12	0.01	0.05	-0.01	-0.06	-0.16	0.00	-0.03	0.04	0.01	0.00	-0.07	-0.05	0.01	-0.06
4. % Unfavorable	0.04	-0.68	-0.09	--	0.02	0.14	-0.13	0.31	-0.04	-0.15	-0.40	-0.19	0.12	0.11	-0.14	0.22	-0.11	0.05	-0.10	-0.08
5. % Dismissed	0.31	-0.49	-0.16	-0.22	--	0.07	-0.42	0.03	0.33	-0.07	-0.09	-0.08	0.28	-0.28	0.04	-0.12	-0.44	-0.17	0.08	-0.11
6. Time to Disposition	0.03	-0.07	0.05	0.08	-0.03	--	-0.16	0.13	0.00	-0.10	-0.15	-0.11	0.09	0.03	0.08	-0.03	-0.03	-0.10	-0.10	0.00
7. % Mental Impairment	-0.12	0.22	0.01	0.14	-0.47	-0.03	--	-0.05	-0.22	-0.02	0.06	0.04	-0.17	0.18	0.05	0.04	0.28	0.05	-0.08	0.20
8. Ave. # Claims	-0.06	-0.11	0.04	0.12	-0.01	0.09	-0.02	--	-0.04	-0.12	-0.09	-0.35	-0.11	0.66	-0.15	0.08	0.16	0.00	-0.13	0.06
9. Prototype State	-0.01	-0.12	-0.03	0.06	0.11	0.07	-0.13	0.02	--	-0.03	0.02	-0.16	0.36	-0.26	-0.14	-0.15	-0.46	-0.13	0.00	-0.08
10. % Bench Dec.	-0.05	0.22	0.02	-0.19	-0.07	-0.02	0.01	-0.03	-0.09	--	-0.01	0.06	-0.02	-0.06	0.18	0.07	0.06	0.05	0.45	0.08
11. % OTR	0.04	0.34	0.04	-0.34	-0.07	0.00	-0.04	-0.07	-0.02	-0.05	--	0.22	-0.19	-0.06	-0.07	-0.12	0.09	0.08	0.03	0.03
12. % Title II	0.05	0.30	0.05	-0.40	0.03	0.03	-0.07	-0.27	-0.17	0.12	0.26	--	-0.75	-0.42	0.09	-0.03	0.20	0.07	-0.03	-0.17
13. % Title XVI	0.01	-0.31	-0.11	0.34	0.07	-0.01	0.09	-0.08	0.21	-0.13	-0.20	-0.73	--	-0.29	0.02	-0.15	-0.52	-0.16	0.11	0.03
14. % Title II and Title XVI	-0.08	-0.06	0.05	0.16	-0.14	-0.02	-0.01	0.48	0.00	-0.03	-0.14	-0.58	-0.14	--	-0.15	0.25	0.43	0.11	-0.11	0.21
15. % Medical Expert Present	-0.08	0.13	-0.01	-0.11	-0.04	0.06	0.07	-0.06	-0.07	0.15	-0.07	0.14	-0.05	-0.15	--	0.10	0.01	-0.05	0.08	0.02
16. % Voc. Expert Present	-0.01	0.05	-0.01	-0.02	-0.03	-0.05	0.02	0.06	-0.13	-0.01	-0.06	0.05	-0.16	0.13	0.04	--	0.25	0.09	0.07	0.10
17. % Claimant Rep. Present	0.02	0.28	0.06	-0.11	-0.28	-0.03	0.15	0.00	-0.29	0.08	0.09	0.29	-0.42	0.08	0.06	0.15	--	0.13	-0.08	0.20
18. % Video Hearing	0.02	-0.08	-0.03	0.09	0.02	0.00	-0.03	0.09	0.10	-0.10	0.00	-0.15	0.04	0.18	-0.17	0.00	-0.06	--	-0.02	0.01
19. %Drafted by ALJ	-0.08	0.18	0.02	-0.22	0.01	-0.03	-0.02	-0.05	-0.02	0.40	0.03	0.15	-0.10	-0.09	0.05	0.01	0.02	-0.08	--	0.07
20. % with New Evidence	-0.04	0.09	-0.02	0.03	-0.14	-0.05	0.17	0.05	-0.03	0.00	-0.02	-0.18	0.11	0.13	0.05	0.04	0.06	0.04	0.00	--

Note: Within-ALJ correlations are displayed below the diagonal, between-ALJ correlations are displayed above the diagonal. Within-ALJ correlations above .06 and between-ALJ correlations above .07 are statistically significant (α = .01).

D. Multiple Regression Analysis

The preceding analyses explored case characteristics individually and found that many case characteristics are related to disposition outcomes. Another important question is to what extent these variables, when taken together, account for the variability in ALJ adjudication outcomes.

1. Analysis

Multiple regression analysis estimates how well a set of variables are able to predict a single outcome variable.[10] It allows examination of the overall ability of the set of predictors as a whole to explain variance in the outcome, which is indexed by the R^2 statistic. In addition, the regression coefficients resulting from this analysis provide information on the unique contribution of each predictor to the model. The analysis produces a regression equation that predicts the outcome as a function of predictor variables. The unique contribution of each predictor is reflected in the regression coefficient, which determines the weight given to a predictor in the equation.

Raw regression coefficients indicate the predicted change in the outcome when the predictor in increased by one point, while holding all other predictors constant. Due to differences in the scaling of the variables, raw regression coefficients can be difficult to interpret. To facilitate interpretation, standardized regression coefficients are also reported. Standardized coefficients present the strength of the relationship on a standardized scale, similar to a correlation coefficient, and thus allow interpretation of the strength of the relationship that can be more readily compared across variables. For both raw and standardized coefficients, positive coefficients indicate that increases in the predictor are associated with higher values of the

[10] *See* APPLIED MULTIPLE REGRESSION/CORRELATION, *supra* note 8.

outcome, while negative coefficients indicate that increases in the predictor are associated with lower levels of the outcome.

Each of the analyses were repeated to estimate both within-ALJ and between-ALJ relationships. Within-ALJ analyses were conducted using a linear mixed model analysis with random ALJ effects. The used of mixed model was to control for repeated measurement of each ALJ, which if not properly modeled can produce biased results.[11] The intercept of the model was allowed to vary randomly across ALJs, which accounts for ALJ differences on the disposition and allowance rates, while all predictors were estimated with fixed coefficients. The use of fixed coefficients assumes that the relationships are the same for all ALJs, and was needed due to the small number of data points within ALJ (*i.e.*, only three years of data). The analysis was conducted using the MIXED routine in the SAS software package.[12]

The between-ALJ analyses were conducted using the mean values for each ALJ on both the predictor and outcome variables. These analyses were conducted using standard ordinary least squares regression techniques.[13] The analyses were conducted using the general linear model (GLM) routine in the SAS software package.[14]

Predictor variables included proportion of claims with mental impairment, average number of claims per claimant, prototype state, proportion of bench decisions, proportion of OTR decisions, type of claim (title II, title XVI or combination), proportion of claims with a medical expert present, proportion of claims with a vocational expert present, proportion of claims with a claimant representative present, proportion of hearings using video, proportion of decisions drafted by the ALJ (as opposed to a decisionwriter), and proportion of cases where new

[11] *See* MULTILEVEL ANALYSIS, *supra* note 6.
[12] *See supra* note 3.
[13] *See* APPLIED MULTIPLE REGRESSION/CORRELATION, *supra* note 8.
[14] *See supra* note 3.

evidence was introduced. Type of claim consisted of three mutually exclusive categories (title II, title XVI or combination), which were represented in the analysis with two predictor variables (proportion title II and proportion title XVI). Because the three categories are mutually exclusive, the proportion of combined claims is a function of the other two, and therefore, its effect is implied by the other two.[15] Year of data collection was also included to control for changes in outcomes over time. Each predictor was centered around its mean, except for year, which was coded as number of years since FY 2009 (*i.e.*, FY 2009=0, FY 2010=1, FY 2011=2).

2. Predictors of Number of Dispositions

Within ALJs, the number of dispositions was not strongly related to the predictor variables. Although several of the effects are statistically significant, they are quite small in magnitude (*see* Table A-12). "Year" accounted for less than1% of the variance in the number of dispositions, and the full set of predictors explained only 3% of the within-ALJ variance.

The between-ALJ regression resulted in only weak prediction of the number of dispositions (*see* Table A-13). The predictor variables accounted for 11% of the between-ALJ variances. ALJs with higher disposition frequencies tended to have more claims per claimant, more bench decisions, more OTR decisions, fewer cases with medical or vocational experts, and more video hearings. The positive coefficient for both title II and title XVI claims indicates that either type of claim produced a higher rate of dispositions than combined cases.

[15] *See* APPLIED MULTIPLE REGRESSION/CORRELATION, *supra* note 8.

Table A- 12: Within-ALJ Predictors of the Number of Dispositions

Effect	Std. Coeff.	Raw Coeff.	SE	df	t Value	p-value
Intercept	--	519.13	4.73	1508	109.86	<.0001
Year After 2009	0.03	7.05	2.13	2221	3.31	0.0009
% Mental Claims	-0.10	-201.88	33.00	2221	-6.12	<.0001
Average Number of Claims	0.02	40.10	28.84	2221	1.39	0.1645
Prototype State	-0.02	-8.22	7.70	2221	-1.07	0.2857
% Bench Decision	0.04	63.02	31.87	2221	1.98	0.0481
% OTR Decision	0.11	198.65	32.79	2221	6.06	<.0001
% Title II	0.11	234.21	55.61	2221	4.21	<.0001
% Title XVI	0.09	205.07	64.27	2221	3.19	0.0014
% Medical Expert Present	-0.12	-83.24	14.50	2221	-5.74	<.0001
% Vocational Expert Present	-0.09	-57.07	13.22	2221	-4.32	<.0001
% Representative Present	0.03	44.54	32.59	2221	1.37	0.1719
% Video Hearing	0.04	29.11	12.03	2221	2.42	0.0156
% Decisions Drafted by ALJ	-0.05	-47.02	17.11	2221	-2.75	0.006
% Cases with New Evidence	-0.01	-5.14	17.14	2221	-0.3	0.7643

Note: Percentage indicates that the predictor was defined as the proportion of cases.

Table A- 13: Between-ALJ Predictors of the Number of Dispositions

Effect	Std. Coeff.	Raw Coeff.	SE	t Value	p-value
Intercept	--	525.16	4.06	129.39	<.0001
% Mental Claims	-0.05	-100.36	54.94	-1.83	0.0679
Average Number of Claims	0.17	290.08	60.65	4.78	<.0001
Prototype State	-0.02	-6.85	10.39	-0.66	0.5097
% Bench Decision	0.08	131.46	45.21	2.91	0.0037
% OTR Decision	0.19	353.80	47.51	7.45	<.0001
% Title II	0.19	380.93	115.78	3.29	0.001
% Title XVI	0.14	317.22	128.62	2.47	0.0138
% Medical Expert Present	-0.11	-77.20	17.17	-4.5	<.0001
% Vocational Expert Present	-0.12	-69.30	15.22	-4.55	<.0001
% Representative Present	-0.02	-26.52	57.90	-0.46	0.647
% Video Hearing	0.07	51.08	18.00	2.84	0.0046
% Decisions Drafted by ALJ	-0.01	-7.23	26.65	-0.27	0.7861
% Cases with New Evidence	-0.01	-10.24	24.53	-0.42	0.6765

Note: Percentage indicates that the predictor was defined as the proportion of cases.

3. Predictors of Allowance Rates

A regression analysis was conducted on within- and between-ALJ predictors of the total allowance rate (either Fully or Partially Favorable). Separate analyses of the Fully Favorable rates produced similar results, and therefore are not reported here.

The within-ALJ analysis revealed that, after controlling for the trend over time, the case characteristics accounted for 20% of the within-ALJ variance in allowance rates. The results are summarized in Table A-14. Higher allowance rates tended to occur in years with a higher proportion of mental claims, fewer claims per claimant, more bench and OTR decisions, more cases with a medical expert present, fewer cases with a vocational expert present, more cases with the claimant representative present, and more cases where new evidence was introduced. The negative coefficients for both title II and title XVI claims indicate that both types of claims had lower allowance rates than combined claims.

In the between-ALJ analysis, a substantial percentage (44%) of the between-ALJ variance in allowance rates was attributable to case characteristics. Results of the between-ALJ analysis are summarized in Table A-15. ALJs who had higher allowance rates tended to have more cases with mental claims, fewer claims per claimant, more bench and OTR decisions, more cases with a medical expert present, fewer cases with a vocational expert present, and more cases with a claimant representative present. The negative coefficients for both title II and title XVI claims indicate that both types of claims had lower allowance rates than combined claims. Allowance rates were also higher in prototype states.

Table A- 14: Within-ALJ Predictors of Total Allowance Rate (Fully + Partially Favorable)

Effect	Std. Coeff.	Raw Coeff.	SE	df	t Value	p-value
Intercept	--	0.57	0.00	1508	178.77	<.0001
Year since 2009	-0.09	-0.02	0.00	2221	-14.06	<.0001
% Mental Claims	0.19	0.33	0.02	2221	16.88	<.0001
Average number of Claims	-0.13	-0.19	0.02	2221	-11.66	<.0001
Prototype State	0.02	0.01	0.00	2221	1.26	0.2087
% Bench Decision	0.13	0.17	0.02	2221	8.87	<.0001
% OTR Decision	0.25	0.39	0.02	2221	19.43	<.0001
% Title II	-0.12	-0.21	0.03	2221	-6.45	<.0001
% Title XVI	-0.21	-0.40	0.04	2221	-10.63	<.0001
% Medical Expert Present	0.09	0.05	0.01	2221	5.64	<.0001
% Vocational Expert Present	-0.12	-0.06	0.01	2221	-6.73	<.0001
% Representative Present	0.13	0.20	0.02	2221	10.32	<.0001
% Video Hearing	-0.02	-0.02	0.01	2221	-2.14	0.0323
% Decisions drafted by ALJ	0.02	0.02	0.01	2221	1.91	0.0567
% Cases with New Evidence	0.04	0.04	0.01	2221	3.5	0.0005

Note: % indicates that the predictor was defined as the proportion of cases.

Table A- 15: Between-ALJ Predictors of Total Allowance Rate (Fully + Partially Favorable)

Effect	Std. Coeff.	Raw Coeff.	SE	t Value	p-value
Intercept	--	0.55	0.00	195.22	<.0001
% Mental Claims	0.2	0.29	0.04	7.56	<.0001
Average Number of Claims	-0.5	-0.66	0.04	-15.69	<.0001
Prototype State	0.1	0.02	0.01	3.25	0.0012
% Bench Decision	0.1	0.17	0.03	5.45	<.0001
% OTR Decision	0.3	0.51	0.03	15.37	<.0001
% Title II	-0.4	-0.75	0.08	-9.33	<.0001
% Title XVI	-0.5	-0.86	0.09	-9.59	<.0001
% Medical Expert Present	0.1	0.06	0.01	5.12	<.0001
% Vocational Expert Present	-0.2	-0.11	0.01	-10.45	<.0001
% Claimant Representative Present	0.2	0.34	0.04	8.3	<.0001
% Video Hearing	0.0	-0.03	0.01	-2.23	0.026
% Decisions Drafted by ALJ	0.0	0.00	0.02	0.11	0.9164
% Cases with New Evidence	0.0	0.01	0.02	0.68	0.4969

Note: % indicates that the predictor was defined as the proportion of cases.

E. Effects of Individual Case Characteristics

To further investigate the above findings, additional analyses were conducted on each case characteristic. These analyses provide an estimate of the difference in allowance rates for the same ALJ when evaluating cases of different types.

1. Analysis

For this analysis, disposition outcomes were dichotomized into Favorable (*i.e.*, Fully or Partially Favorable) and Unfavorable decisions. Dismissed cases were excluded from the analysis. In order to account for multiple data points from each ALJ, analyses were conducted using a generalized linear mixed model with a logit link function.[16] Fiscal year of the disposition was included to control for changes in allowance rates over time. In order to provide a consistent frame of reference, allowance rates are discussed below in terms of 2011 rates, although data from all three years was included in the analysis.

The logit link function is a transformation of the outcome variable that improves the statistical performance of models with dichotomous outcomes (*i.e.*, Favorable or Unfavorable). Due to this transformation, the coefficients of the model are difficult to interpret directly. Therefore, results were transformed into predicted allowance rates to facilitate interpretation. Predicted allowance rates are discussed below in terms of FY 2011 rates, although data from all three years was included in the analysis.

For each analysis, the predictor was entered both as a within-ALJ and between-ALJ effect. The within-ALJ effect estimates whether the probability of an allowance decision is different across levels of the predictor. The between-ALJ effect estimates whether ALJs assigned a higher percentage of a case type tend to issue more or less allowance decisions. All predictors were centered around their means.

[16] Analyses were conducted using the MIXED routine in the SAS data analysis package. *See supra* note 3.

The within-ALJ effects were estimated with randomly varying regression coefficients. In a random coefficient model, the regression coefficient is allowed to take on a unique value for each ALJ. The analysis provided an estimate of the average difference between video and non-video hearings, as well as the variance of these differences across ALJs. This allowed us to test whether the effect of each predictor was consistent across ALJs.

2. Video Hearings

Video hearings were used in 20% of dispositions. The analysis found a small but statistically significant negative effect, indicating that video hearings tended to have slightly lower allowance rates than non-video hearings (*see* Table A-16). Video hearings on average produced allowance rates that were about 3% lower than non-video hearings (*see* Figure A-10).

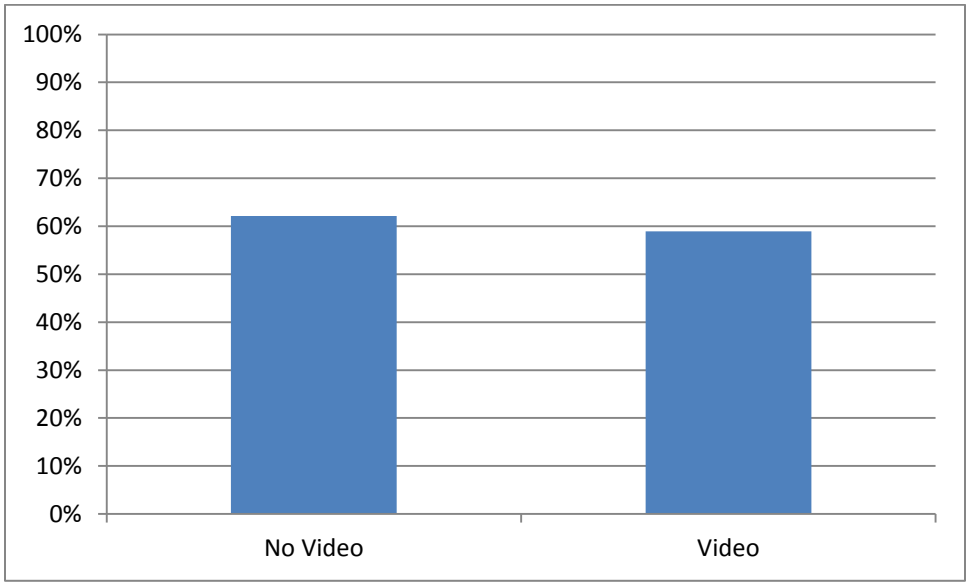
The variance component for the video hearing slope was significant (slope SD=.28, $p<.01$), indicating that the video effect differed across ALJs. Although the average effect of video hearing was small, some ALJs showed a larger effect, while others showed an effect in the opposite direction. While the majority of ALJs (80%) showed lower allowance rates, 20% showed higher allowance rates on video hearings.

A 95% prediction interval for the video hearing slope (*i.e.*, slope +/- 1.96 SD) ranged from -0.68 to 0.42. ALJs at the low end of the distribution had a 55% allowance rate for video hearings and 71% for non-video hearings. ALJs at the upper end of the distribution had an allowance rate of 75% for video hearings and 66% for non-video hearings.

Table A- 16: Effect of Video Hearing on Allowance Decision Rate

Effect	Estimate	SE	DF	t Value	p-value
Intercept	0.76	0.020	1508	37.04	<.0001
Year after 2009	-0.15	0.002	1540000	-60.42	<.0001
Video Hearing	-0.13	0.011	1360	-12.53	<.0001
ALJ % Video	-0.05	0.089	1508	-0.56	n.s.
Random Effects					
Variance of Intercept	0.61	0.023			
Variance of Video Hearing Coefficient	0.08	0.005			

Figure A- 10: Predicted Allowance Rate for Video Hearings (FY 2011)



An additional analysis was conducted to examine whether video hearings produced more consistency (less variance) in allowance rates. Data provided by SSA on yearly disposition outcomes indicates the opposite pattern, with larger variance across ALJs in allowance rates for video hearings (SD=.25) than for non-video hearings (SD=.17). However, these numbers are

not directly comparable, due to differences in the number dispositions used in the calculation of rates for these two conditions. The variance due to sampling error for rate statistics is influenced by the number of data points. Specifically, the SD of a rate *p* computed from *n* data points is $\sqrt{p(1-p)/n}$. As the number of data points increases, the SD of the rate will decrease. From our data, the average yearly allowance rate for video hearings was 0.50 based on an average of 101 dispositions, while the average yearly allowance rate for non-video hearings was .55 based on an average of 350 dispositions. Given these values, the expected SD due to sampling error would be .05 for video hearings, and .03 for non-video hearings.

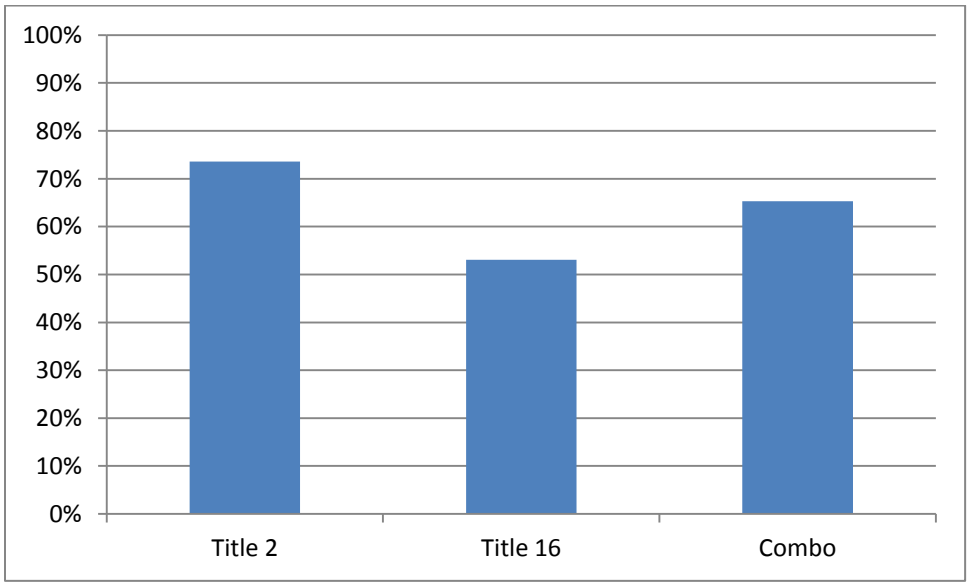
3. Claim Type

Thirty-six percent of dispositions were filed under title II, 28% under title XVI, and 41% under both. Results for type of claim are summarized in Table A-17. On average, allowance rates were higher for title II claims and lower for title XVI claims (*see* Figure A-11). Significance variance was found across ALJs in both effects (title II slope SD=.26, $p<.05$; title XVI slope SD=.29, $p<.01$). For both slope coefficients, a 95% prediction interval indicated a stronger effect for some ALJs while a near-zero effect for others (ranging from -0.12 to 0.90 for the title II slope and -1.08 to 0.06 for the title XVI slope). Thus, while the strength of the relationships varied across ALJS, there was a consistent tendency to have higher allowance rates in title II cases and lower allowance rates in title XVI cases.

Table A- 17: Effect for Title II and Title XVI Claims on Allowance Decision Rate

Effect	Estimate	SE	DF	t Value	p-value
Intercept	0.92	0.021	1507	43.29	<.0001
Year after 2009	-0.15	0.002	1710000	-65.31	<.0001
Title II Claims	0.39	0.008	1508	47.45	<.0001
Title XVI Claims	-0.51	0.009	1509	-57.26	<.0001
ALJ % Title II	1.07	0.380	1503	2.81	0.005
ALJ % Title XVI	-0.02	0.425	1503	-0.04	0.9652
Random Effects					
Intercept Var.	0.66	0.024			
Variance of Title II Coeff.	0.07	0.004			
Variance of Title XVI Coeff.	0.09	0.004			

Figure A- 11: Predicted Allowance Rate by Claim Type (FY 2011)



4. New Evidence

New evidence was introduced after the hearing in 23% of the dispositions. The actual frequency of new evidence is probably even higher, because this factor is recorded in CPMS

only if the new evidence changed the disposition outcome. An unknown number of claimants submitted new information that did not change the final disposition.

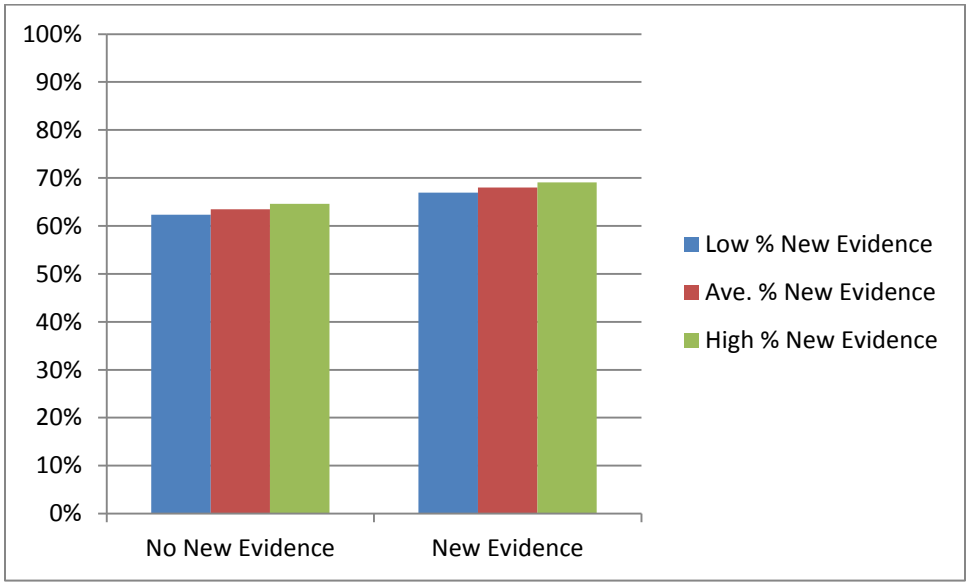
The introduction of new evidence had a positive effect on allowance rates both within- and between-ALJs. ALSs issued allowance decisions in 68% of cases where new evidence was introduced, compared with a 63% allowance rate for cases without new evidence (*see* Figure A-12). In addition, even after controlling for this effect, ALJs who had a higher percentage of cases where new evidence was introduced tended to issue more allowance dispositions than ALJs with a low percent of cases with new evidence. The combined within- and between-ALJ effects are depicted in Figure A-12.

The effect of new evidence varied considerably across ALJs (slope SD=.76, p<.01). The 95% prediction interval indicated that the slope ranges from strongly negative -1.29 to strongly positive (1.70). Thus, while the effect of new evidence was small on average, it had a big impact (either positive or negative) for some individual ALJs (*see* Table A-18).

Table A- 18: Effect of New Evidence on Allowance Decision Rate

Effect	Estimate	SE	DF	t Value	p-value
Intercept	0.93	0.021	1508	44.14	<.0001
Year after 2009	-0.16	0.002	1710000	-70.26	<.0001
New Evidence	0.20	0.021	1495	9.67	<.0001
ALJ % New Evidence	0.26	0.104	1507	2.52	0.012
Random Effects					
Variance of Intercept	.64	0.024			
Variance of New Evidence Slope	.58	0.025			

Figure A- 12: Predicted Allowance Rates by Introduction of New Evidence (FY 2011)



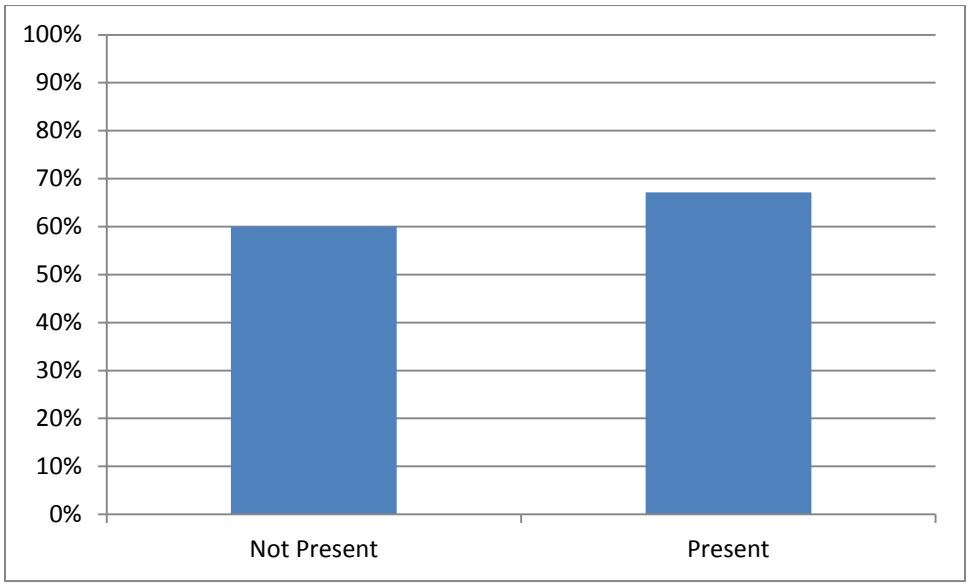
5. Medical Expert

A medical expert was recorded as present in 14% of hearings. Results from the analysis of a medical expert presence are summarized in Table A-19. Figure A-13 shows allowance rates tended to be higher for cases where a medical expert was present (67% allowance) than when a medical expert was not present (60% allowance). There was considerable variability in this effect across ALJs (slope SD=67. 95% prediction interval ranged from -1.0 to 1.6). For ALJs at the low end of the distribution, hearings with a medical expert present had a 24% lower allowance rate that hearings without a medical expert, while for those at the upper end of the distribution, hearings with a medical expert had a 27% higher allowance rate than hearings without a medical expert.

Table A- 19: Effect of Medical Expert Presence on Allowance Decision Rate

Effect	Estimate	SE	DF	t Value	p-value
Intercept	0.74	0.020	1507	36.63	<.0001
Year after 2009	-0.14	0.002	1540000	-60.23	<.0001
Medical Expert	0.31	0.023	1275	13.55	<.0001
ALJ % Medical Expert	0.14	0.073	1506	1.86	0.0625
Random Effects					
Variance of Intercept	.74	0.020			
Variance of Medical Expert Coefficient	.45	0.025			

Figure A- 13: Predicted Allowance Rate by Presence of Medical Expert



6. Vocational Expert

A vocational expert was recorded as present in 76% of hearings. The distribution of vocational expert use was negatively skewed. For half of the ALJs, vocational experts were present in over 90% of the hearings, and for 75% of the ALJs, vocational experts were used in over 80% of the cases. In contrast, for about five percent of the ALJs, vocational experts were used in less than 2% of the cases.

The analysis revealed a complex pattern of results (*see* Table A-20). When looking within an ALJ, cases with a vocational expert present had higher allowance rates than cases with no vocational expert (*see* Figure A-14). However, ALJs who had a vocational expert in a high percent of cases (95% with vocational expert) tended to have lower allowance rates than ALJs who had a low use of vocational expert use (50% with vocational expert).

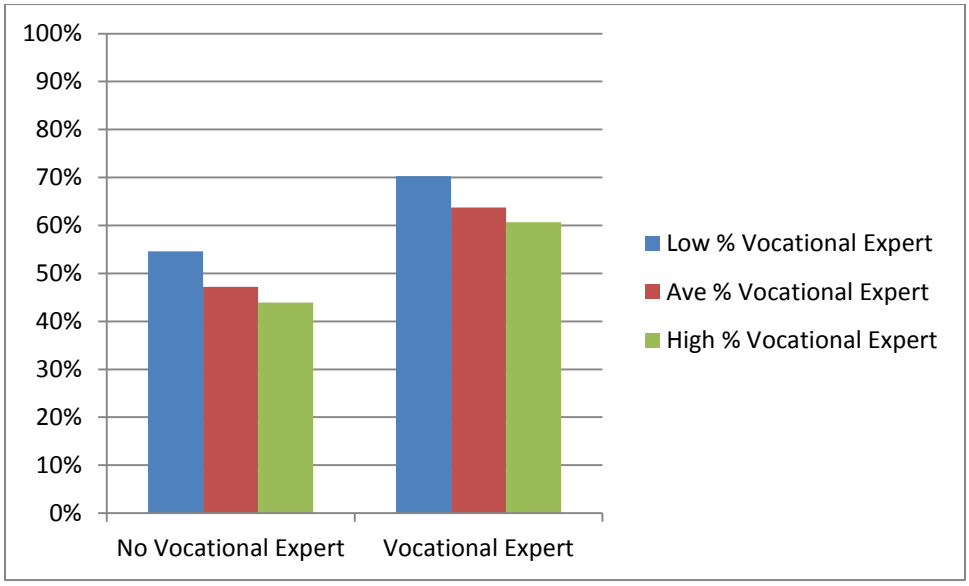
This within-ALJ result differs from the multiple regression analysis reported above, which showed a negative impact of vocational expert. The positive correlation may be due in part to the confounding of vocational expert use with other case characteristics. Vocational experts were more likely to be present in cases that involved combined title II and title XVI claims (r=.25) and where a claimant representative was present (r=.25). Because these factors tended to be associated with higher allowance rates, the positive correlation of vocational expert in the separate analysis may be influenced by its co-occurrence with these other factors. In contrast, the multiple regression analysis controlled for levels of these variables, and suggests a negative relationship.

The within-ALJ coefficient for presence of a vocational expert showed considerable variance across ALJs (slope SD=0.66). A 95% prediction interval indicated that the slope ranged from -0.61 to 1.97. Thus, the for some ALJs, the effect was substantially stronger than the average, while for others, presence of a vocational expert was associated with lower allowance rates.

Table A- 20: Effect of Presence of Vocational Expert on Allowance Decision Rate

Effect	Estimate	SE	DF	t Value	p-value
Intercept	0.70	0.020	1507	34.48	<.0001
Year after 2009	-0.15	0.002	1540000	-61.81	<.0001
Vocational Expert	0.68	0.019	1497	36.07	<.0001
ALJ %Vocational Expert	-0.99	0.070	1506	-14.14	<.0001
Random Effects					
Variance of Intercept	0.59	0.023			
Variance of Vocational Expert Coefficient	0.43	0.019			

Figure A- 14: Predicted Allowance Rates by Presence of Vocational Expert (FY 2011)



7. Claimant Representative

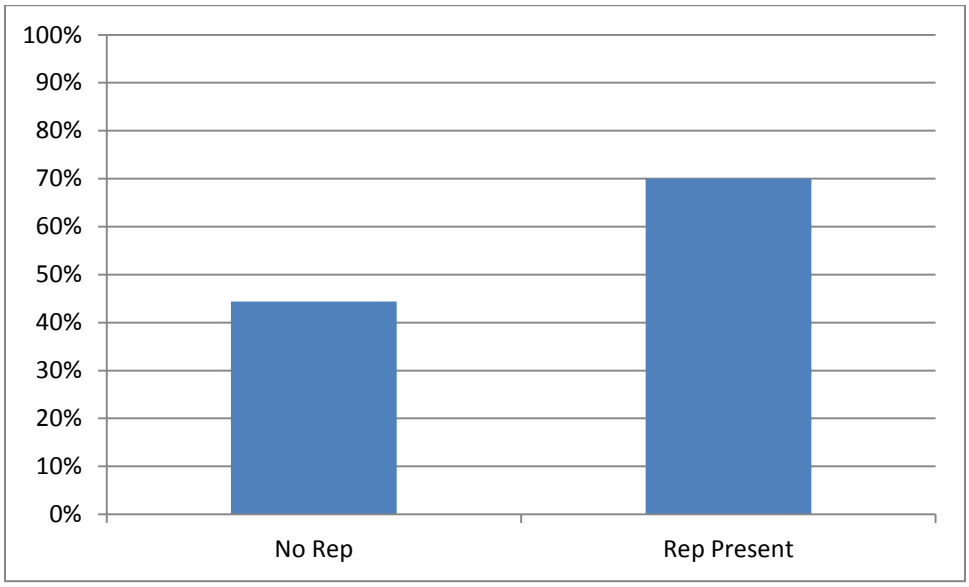
Seventy-seven percent of hearings involved a claimant representative. Results from the analysis of claimant representative are summarized in Table A-21. Figure A-15 shows that in general, allowance rates were higher when a claimant representative was present (64%) than when a representative was not present (47% allowance). The analysis also revealed significant variance in slopes across ALJs (slope SD=.46, p<.01). A 95% prediction interval on the slope

ranged from -.17 to 1.97. Thus, while the presence of a representative was associated with higher allowance rates for most ALJs, for some this trend was attenuated.

Table A- 21: Effect of Presence of Claimant Representative on Allowance Decision Rate

Effect	Estimate	SE	DF	t Value	p-value
Intercept	0.94	0.022	1507	43.16	<.0001
Year after 2009	-0.17	0.002	1710000	-71.83	<.0001
Claimant Representative	1.07	0.013	1509	81.41	<.0001
ALJ % Claimant Representative	0.27	0.245	1506	1.09	0.2779
Random Effects					
Variance of Intercept	0.69	0.026			
Variance of Claimant Representative Coefficient	0.21	0.009			

Figure A- 15: Predicted Allowance Rate by Presence of Claimant Representative (FY 2011)



8. Bench Decisions, OTR Decisions, and Use of Decisionwriters

These three variables were all strongly correlated with allowance rates, due to the nature of these situations. Bench and OTR decisions are only permitted for allowance decisions. Similarly, the data indicate that ALJs write decisions almost exclusively for allowance decisions (80% of ALJ-authored dispositions) and dismissals (13% of ALJ-authored dispositions). Given the consistency in these findings, additional within-ALJ analyses were not conducted for these predictors.

9. Prototype States

For the most part, the claims reviewed by an ALJ were from either all prototype or all non-prototype states. Therefore, analysis of this variable was conducted at the ALJ level. In the few cases where an ALJ had data from both types of states, the majority of claims were generally from one type or the other. Therefore, ALJs were classified into prototype and non-prototype states based on majority of claims.

Thirty percent of the claims reviewed by ALJs were from prototype states (*i.e.*, states that lack the reconsideration stage between the initial (Disability Determination Service) stage and the ALJ hearing stage. Means and SD of disposition outcomes for prototype and non-prototype states are summarized in Table A-21. An independent-groups t-test indicated significant differences on Fully Favorable and dismissal rates. Figure A-16 and Table A-22 show that Fully Favorable rates were slightly lower in prototype states (48%) than non-prototype states (51%). The interpretation of this difference is complicated by the finding of higher dismissal rates for prototype states (18%) than non-prototype states (14%). The inclusion of more dismissed claims in prototype states will tend to increase the total number of claims used in the calculations and drive down the rates for all of the other outcomes. If one considers only cases were a decision was reached (excluding dismissed claims), the Fully Favorable rate was quite similar for

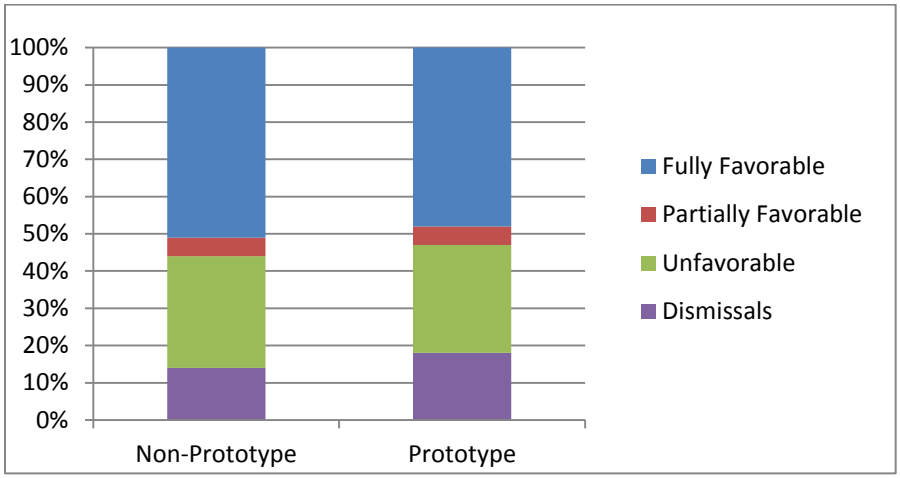
prototype and non-prototype states (58% and 59%, respectively), and no differences were observed for Partially Favorable (6% in both) or Unfavorable rates (35% in both).

The interpretation of the prototype effect is complicated by moderate correlations with other case characteristics. Specifically, ALJs from prototype states tend to have a higher proportion of title XVI claims (r=.36) and a lower percentage of hearings with a claimant representative present (r=-.46). Both of these factors (title XVI cases and absence of representative) are associated with lower allowance rates, these may partially explain the lower allowance rate in prototype states. In fact, the multiple regression analysis presented above (Table A-15) showed that, after controlling for these effects, ALJs from prototype states had 2% higher allowance rates than those from non-prototype states.

Table A- 22: Disposition Outcomes for Prototype and Non-Prototype States

	Non-Prototype (N=1057)		**Prototype (N=452)**				
	M (SE)	**SD**	**M (SE)**	**SD**	**t Value**	**df**	**p-value**
Dispositions Per Year	526 (5.03)	163	542 (7.61)	162	-1.72	1507	0.09
Fully Favorable Rate	51% (0.48)	16%	48% (0.66)	14%	4.02	1507	<.0001
Partially Favorable Rate	5% (0.12)	4%	5% (0.15)	3%	0.54	1507	0.59
Unfavorable Rate	30% (0.41)	13%	29% (0.56)	12%	1.52	1507	0.13
Dismissals Rate	14% (0.17)	6%	18% (0.33)	7%	-13.72	1507	<.0001

Figure A- 16: Comparison of Rulings for Non-Prototype and Prototype States



10. Type of Impairment

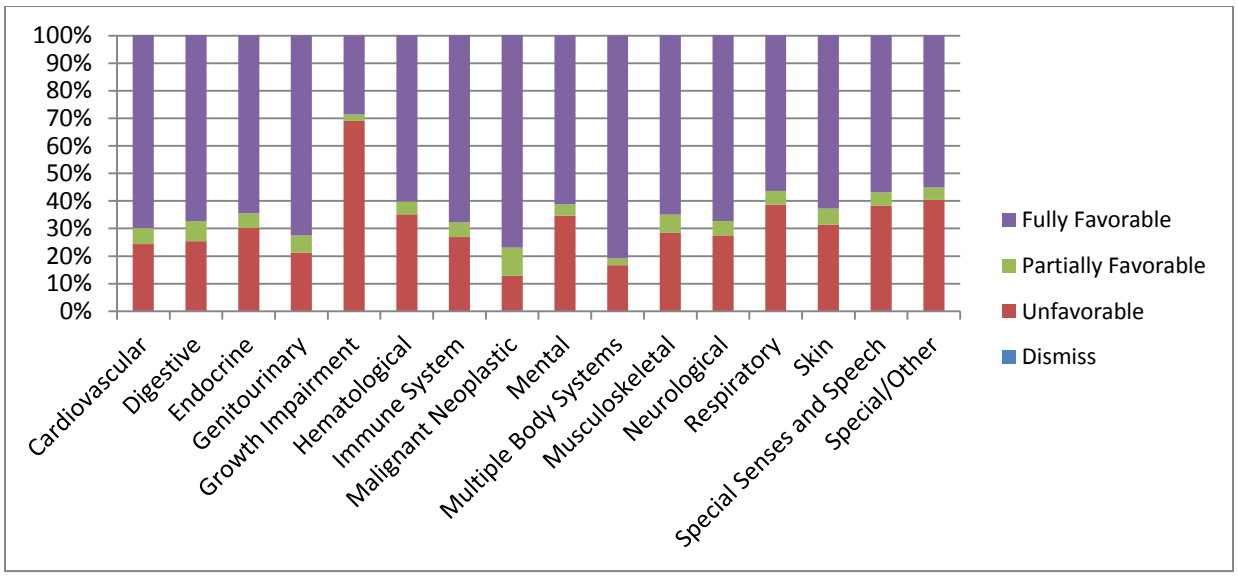
Data on type of impairment was not linked to individual ALJs, and therefore could not be included in the correlation and regression analyses reported above.

Disposition outcome rates by impairment type are summarized in Table A-23 and Figure A-17 on the following page. Allowance rates differed across type of impairment. Higher Unfavorable rates were found for growth impairments. Higher allowance rates were found for genitourinary, malignant neoplastic, and impairments involving multiple body systems. There was a significant effect for impairment type on Fully Favorable decisions [F(15, 560)=163.03, p=<.0001], Partially Favorable decisions [F(15, 560)=26.41, p=<.0001], and Unfavorable decisions [F(15, 560)=183.00, p=<.0001].

Table A- 23: Trends in Number of Dispositions by Impairment (FY 2009 - FY 2011)

	FY 2009		FY 2010		FY 2011	
Impairment	**Dispositions**	**%**	**Dispositions**	**%**	**Dispositions**	**%**
Cardiovascular	36076	7%	38401	6%	38941	6%
Digestive	12375	2%	14123	2%	14803	2%
Endocrine	23112	4%	25425	4%	26679	4%
Genitourinary	3569	1%	4082	1%	4215	1%
Growth Impairment	258	0%	379	0%	521	0%
Hematological	1806	0%	2143	0%	2186	0%
Immune System	15289	3%	17256	3%	17885	3%
Malignant Neoplastic	9273	2%	9950	2%	9715	1%
Mental	138817	25%	167217	27%	182673	27%
Multiple Body Systems	78	0%	74	0%	82	0%
Musculoskeletal	229921	42%	257461	41%	269864	40%
Neurological	34389	6%	39347	6%	41702	6%
Respiratory	20474	4%	23223	4%	24871	4%
Skin	1470	0%	1732	0%	1838	0%
Special Senses and Speech	8011	1%	9171	1%	10520	2%
Special/Other	14428	3%	20522	3%	23702	4%
Total	549346		630506		670197	

Figure A- 17: Comparison of Rulings Based on Impairment Type



V. ANALYSIS OF REMANDS

A. Analysis of Federal Court Remands

1. Data

SSA provided two datasets with information about court remands. Data from the Appeals Council Review Processing System (ARPS) database provided information summary information on the number of appeals and remands for both the Appeals Council and the courts. Data were available on Appeals Council appeals for FY 2009 through FY 2011. Summary data on the courts were available for the first half of FY 2009, and all of FY 2010 and FY 2011.

2. Analysis of Court Remands

It should be noted that the summary frequencies (*see* Table A-24) reflect the number of cases filed and court remands issued during a FY, and thus do not necessarily reflect the same cases. Cases filed in a particular year are often decided in subsequent years, and therefore the number of remands is not based only on the cases filed in that year. Assuming that the rate of court filings and court decisions is stable over time, the ratio of remands to cases filed should approximate the percentage of cases filed that are remanded.

Table A- 24: Frequency of Federal Court Filings and Remands

Fiscal Year	Court Filings During Fiscal Year	Remands Issued During Fiscal Year	Percent Remanded
FY 2009*	6441	3085	48%
FY 2010	13106	6182	47%
FY 2011	14648	6171	42%
Total	34195	15438	45%

*Note: Percentages are only approximate, because remand decisions may be issued in a different year than court filing. *Only 6 months of data were available for FY 2009.*

3. Remands by Region

Significant differences in the number of remands were found across regions, chi-square (9, N=14567)=3540.0, p<.001. As shown in Table A-25, lower remand frequencies were found in Region 8 (340), Region 1 (890) and Region 7 (915), while higher remand rates were found in Region 4 (3098), Region 6 (1969) and Region 9 (1,853). However, without information about the relative volume of dispositions across regions, these frequencies are difficult to interpret.

Table A- 25: Number of Remands by Region

Region	Number of Remands	%
1	890	6%
2	1562	11%
3	1456	10%
4	3098	21%
5	1440	10%
6	1969	14%
7	915	6%
8	340	2%
9	1853	13%
10	1044	7%

4. Other Case Characteristics

The data on court remands included limited information on the characteristics of the cases. Of the 11,240 remands where Continuing Disability Review (CDR) was coded, only 130 (1%) involved CDR. 5,460 remands (37%) involved title II claims, 3,098 (21%) involved title XVI claims, and 6,013 (41%) involved concurrent claims. Almost all (99%) of the title II claims were Disability Insurance – Worker/Child. Over 99% of the title XVI claims were disability related; very few involved age or blindness. Over 99% of the concurrent claims involved disability.

5. Reason for Court Remand

Detailed reason for remand was available for 14,571 cases remanded by the courts between FY 2009 and FY 2011. Some remands were excluded from this data by SSA due to inability to match the remands to other system data. Thus, the number of remands included in this analysis is less than total reported above.

Remand reasons were classified by a specific code, which fell into one of 10 categories: Substantial Gainful Activity (SGA), Severe/Non-severe, Adult Listings, Child Listings, Credibility Evaluation, Opinion Evidence Evaluation & Residual Functional Capacity (OEE & RFC), Past Relevant Work, Grid/Vocational Expert, Dismissal/Procedural, and Miscellaneous.

Each case listed up to three remand reasons, and a particular case was included in the frequency counts for each category listed. As such, each case could be included in multiple categories, and the listed frequencies are not mutually exclusive.

The frequency of remand reason categories are summarized in Table A-26 and displayed in Figure A-18. The most common reason was OEE & RFC, which comprised 54% of the remands. This category involves remands related to application of the treating physician rule, as well as other medical evidence and other issues related to residual functional capacity.[17]

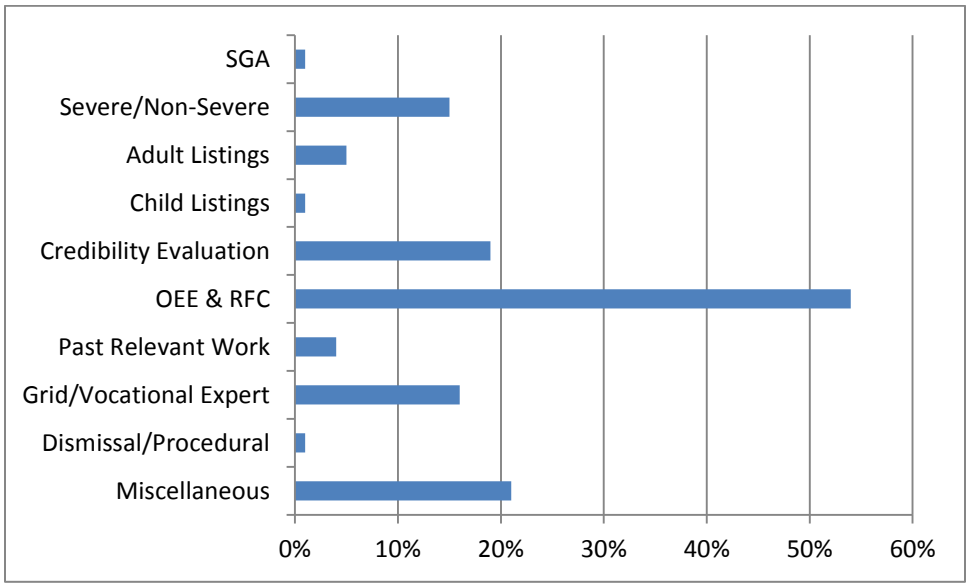
[17] For additional analysis of remand reasons related to OEE & RFC, *see* Appendix B to OFFICE OF THE CHAIRMAN, ADMIN. CONFERENCE OF THE U.S., SSA DISABILITY BENEFITS PROGRAMS: ASSESSING THE EFFICACY OF THE TREATING PHYSICIAN RULE (2013), *available at* http://www.acus.gov/sites/default/files/documents/Treating_Physician_Rule_Draft_Report_revised_3-3-2013.pdf.

Table A- 26: Relative Frequency of Reasons for Remand Categories

Category Number	Category Label	Frequency	Percent of Remands
1	SGA	161	1%
2	Severe/Non-Severe	2160	15%
3	Adult Listings	665	5%
4	Child Listings	158	1%
5	Credibility Evaluation	2727	19%
6	OEE & RFC	7864	54%
7	Past Relevant Work	639	4%
8	Grid/Vocational Expert	2306	16%
9	Dismissal/Procedural	124	1%
0	Miscellaneous	3086	21%
	Total Remands	14571	

Note: Remand categories are not mutually exclusive.

Figure A- 18: Frequencies of Reason for Remand Categories



6. Remands Related to New Evidence

Two reasons for remand cited introduction of new evidence. New evidence presented upon administrative appeal/review was cited in 355 remanded cases (2%). New evidence presented to federal court was cited in 139 remands (1%).

The rate of remands related to the introduction of new evidence differed slightly across regions, chi-square (9, N=14567) = 21.33, p=.01. Remands due to new evidence occurred at a lower rate for Regions 1, 5, 7 and 10 (3%). The highest remand rates were in Regions 2, 4, 6 and 8 (4%). Remands related to new evidence made up 2.6% of the remands from Region 1.

Differences across regions were more pronounced on remands for new evidence presented to the federal court, chi-square (9, N=14567) = 33.9, p<.0001. This reason made up less than 0.5% of the remands in Regions 1, 8, 9 and 10, and made up over 1.5% of remands in Regions 2 and 4. Remands for new evidence presented to federal court made up 0.45% of the remands in Region 1.

B. Appeals Council Remands

1. Frequency of Appeals Council Remands

Data on requests for review by the Appeals Council for FY 2009 - FY 2011 were obtained from ARPS. The data indicate that 24% of reviewed cases are remanded each year (*see* Table A-27). This trend has been stable over the three years investigated.

Table A- 27: Appeals Council Remand Rates

Fiscal Year	Dispositions Reviewed	Remands	Percent Remand
2009	89074	21797	24%
2010	102076	24810	24%
2011	127029	30044	24%

2. Reasons for Appeals Council Remand

Data on reason for remands was obtained from heat maps prepared by Ben Gurga, ODAS/OESS/DMIA, 10/19/2011. Note that these percentages are based on the total number of remand reasons given. Because each case could yield multiple remand reasons, these frequencies are greater than the total number of remands.

Table A- 28: Frequency of Appeals Council Remand Reason Categories (FY 2009 - FY 2011)

Remand Reason	Frequency	Percent of Total Reasons
SGA	1424	1%
Severe/Non-Severe	13365	11%
Adult Listings	1728	1%
Child Listings	314	0%
Credibility Evaluation	7087	6%
OEE & RFC	39525	34%
Past Relevant Work	5820	5%
Grid/Vocational Expert	8029	7%
Dismissal/Procedural	13214	11%
Misc.	20642	18%
Not Listed in Heat Map	5161	4%
Total	116,309	

3. Remands Related to New Evidence

New evidence was cited in 5,245 Appeals Council remands, which constitute 5% of the cited remand reasons. This was fairly stable over time: 4.5% in FY 2009 4.7% in FY 2010, and 4.4% in FY 2011.