Outcome Variation in the Social Security Disability Insurance Program: The Role of Primary Diagnoses

by Javier Meseguer*

Based on the adjudicative process, the author classifies claimant-level data over an 8-year period (1997–2004) into four mutually exclusive categories: (1) initial allowances, (2) initial denials not appealed, (3) final allowances, and (4) final denials. The ability to predict those outcomes is explored within a multilevel modeling framework, with applicants clustered by state and primary diagnosis code. Variance decomposition suggests that medical diagnoses play a substantial role in explaining individual-level variation in initial allowances. Moreover, there is statistically significant high positive correlation between the predictions of an initial allowance and a final allowance across the diagnoses. This finding suggests that the ordinal ranking of impairments between these two adjudicative outcomes is widely preserved. In other words, impairments with a higher expectation of an initial allowance also tend to have a higher expectation of a final allowance.

Introduction

The purpose of the Disability Insurance (DI) program is to replace part of a worker's earnings in the eventuality of a physical or mental impairment preventing the individual from working. The disability portion of the Old-Age, Survivors, and Disability Insurance (OASDI) program, administered by the Social Security Administration (SSA), protects workers and their eligible dependents against such risk. SSA administers a second program, Supplemental Security Income (SSI), which has no employment or contribution requirements, but imposes strict income and asset limits. It is designed to be a program of last resort, assisting aged, blind, or disabled individuals who have very limited resources.

The goal of this study is to explore the extent to which medical diagnoses and state of origin may explain observed heterogeneity in disability decisions. One instance of heterogeneity is manifest at the state level. The DI program is federally administered and is operated in collaboration with the states. When a local Social Security field office establishes that an applicant meets all of his or her nonmedical requirements, the case is forwarded to the state Disability Determination Service (DDS) for a decision. The DDS follows a sequential process to evaluate the medical evidence and decide if the applicant meets the definition of disability. In doing so, a DDS examiner considers the severity of the impairment(s), along with vocational factors that take into account age, education, and work experience. SSA guidelines to determine disability are uniform across all 50 states. In practice, however, there can be wide variation in state allowance rates.

A second instance of variation in DI outcomes occurs through the adjudicative process. If a disability claim is denied, the applicant has a number of opportunities to appeal the decision. There are three stages of appeal within SSA: (1) a reconsideration by the

Selected Abbreviations

ALJ	administrative law judge
DDS	Disability Determination Service
DI	Disability Insurance
DIC	deviance information criterion
DRF	Disability Research File
RFC	residual functional capacity

* *Javier Meseguer is an economist with the Office of Economic Analysis and Comparative Studies, Office of Research, Evaluation, and Statistics, Office of Retirement and Disability Policy, Social Security Administration.*

Note: Contents of this publication are not copyrighted; any items may be reprinted, but citation of the Social Security Bulletin as the source is requested. To view the Bulletin online, visit our website at http://www.socialsecurity.gov/policy. The findings and conclusions presented in the Bulletin are those of the authors and do not necessarily represent the views of the Social Security Administration.

Selected Abbreviations—*Continued*

SGA	substantial gainful activity
SSA	Social Security Administration
SSI	Supplemental Security Income

state DDS, (2) a hearing before an administrative law judge (ALJ), and (3) a review by the Appeals Council. If those stages are exhausted, the claimant can always seek legal redress in a federal district court. While few initial denials are reversed at the reconsideration level, a substantial portion of claimants who appeal at the hearing level or above are eventually allowed.

The two referenced sources of variation in disability outcomes (by state and adjudicative level) have been a cause of concern to SSA and Congress regarding the practical implementation of the disability programs. My hunch is that the collection of impairments in particular might shed some light in explaining a portion of the observed variation. Thus, I investigate heterogeneity in disability outcomes along three dimensions: state of origin, medical diagnosis, and adjudicative stage. That objective is pursued by working with a random sample of the Disability Research File (DRF). The DRF follows a cohort of applicants through the various stages of the determination process, identifying decisions made at different adjudicative levels. The disability determinations in the file are separated into four mutually exclusive categories: (1) initial allowances, (2) initial denials not appealed, (3) final allowances, and (4) final denials. This classification of the data implicitly reduces the adjudicative process to two stages (*initial* and *final*).

The data is fitted to various Bayesian hierarchical multinomial logit specifications, with two different groups or clusters nesting the claimant-level observations. One group is the 50 states. The other group comprises 181 medical impairments, which represent the unique administrative four-digit primary diagnosis codes. This modeling approach offers several advantages. First, the framework is multivariate, meaning that instead of estimating a separate model for each stage, the adjudicative outcomes are estimated jointly. Second, the multilevel or hierarchical nature of the models enables the distinction to be made between claimant-level effects on one hand and state or diagnosis-level effects on the other hand. In other words, I can decompose heterogeneity in the adjudicative outcomes by source into "between-group" and "within-group" variance. For instance, at one end of the spectrum, it is possible that claimants within a state are rather uniform in their characteristics, so that most of the variance in initial allowances is due to unique differences between the states. Alternatively, a large portion of the total variance could be attributed to claimant-level heterogeneity within the states (that is, the states are not that different from one another, but the population within a given state varies greatly in its characteristics). Finally, a third advantage in this modeling approach is the ability to estimate correlation patterns that may exist between the disability adjudicative outcomes.

The next section in this article provides background information about the Social Security disability programs, including the disability determination and appeals processes. I then briefly review some of the literature regarding the modeling of allowance rates. The data and modeling approach are discussed next, emphasizing the observed variation in adjudicative outcomes by such factors as age, diagnosis group, state of origin, and mortality. The inferential results are presented in the following section, where the "goodness-of-fit" of the various models and the "average effect" of various explanatory variables are evaluated and discussed. Two other important issues addressed in this section involve variance decomposition and correlation, where I describe the interpretation and implications of my estimates. The last section concludes with a summary of the main findings.

Social Security Disability Programs

SSA operates two different programs that offer cash benefits to the disabled: the Disability Insurance program, which was enacted in 1956, and the Supplemental Security Income program, which began in 1972. The two programs share the same disability determination process, but have different objectives. DI is funded through payroll tax contributions and is designed to protect workers contributing to the program from earnings losses that are due to impairment. SSI, on the other hand, is not contributory. General revenues fund it, and the main goal of the program is to guarantee a minimal level of income to the poorest of the aged, blind, or disabled population.

The DI program provides benefits to disabled workers who are younger than their respective full retirement ages and to their spouses, surviving disabled spouses, and disabled children, although workers account for the largest share of beneficiaries (typically, over 80 percent of the DI rolls). At the end of 2010, about 8.8 million workers and their dependents

were receiving DI benefits and 4.7 million individuals were receiving SSI payments. Under both programs, the definition of disability is one of long-term work disability. It involves the inability to engage in substantial gainful activity (SGA) because of a medically determinable physical or mental impairment that is expected to last at least 12 months or result in death.

Eligibility for DI benefits requires a worker to be insured, younger than his or her full retirement age, and to meet the definition of disability. The applicant must have worked long enough in employment covered by Social Security (approximately 10 years) and recently enough (about 5 of the past 10 years). Those requirements are relaxed for younger applicants who have shorter employment histories. An applicant who is employed must also have monthly earnings below the SGA threshold ($1,640 for a blind person and $1,000 for a nonblind individual in 2010). However, there are no restrictions on nonwage income. Upon approval, benefits are received after a 5-month waiting period from the onset of disability. In addition, the beneficiary is entitled to Medicare coverage after receiving benefits for 2 years.

Disability benefits continue for as long as the beneficiary remains disabled or reaches full retirement age, in which case there is a conversion to retirement benefits. Upon death of a worker, some dependent benefits may convert into survivor benefits. SSA conducts periodic continuing disability reviews (CDRs) to determine if an individual remains disabled. Review frequency depends on the severity and likelihood of improvement of the disability and can range from 6 months from the initial finding to as long as 7 years. A finding that a beneficiary is engaging in SGA will result in termination.[1]

From 1970 through 2009, the number of beneficiaries in the DI program more than tripled, while DI expenditures increased by almost seven times in inflation-adjusted figures (Congressional Budget Office 2010). According to the Social Security Advisory Board (2012a), that expansion can be traced to several factors in addition to an increase in the general population. One factor has been an increase in the share of lower mortality impairments with earlier onset (such as musculoskeletal and mental disorders). Applicants with those types of impairments tend to enter the program at younger ages and remain as beneficiaries for longer periods of time. Another factor has been an increase in female labor force participation. The rapid pace at which women have joined the ranks among workers has considerably expanded the pool of applicants. Indeed, the gender composition of beneficiaries today is much closer to that of the population at large. A third factor has been an increase in earnings replacement rates. Rising income inequality coupled with the average wage indexing of benefits has increased the portion of potential earnings replaced by DI benefits. Younger low-skilled workers in particular have experienced the highest increase in the value of DI benefits at a time of reduced demand for their labor. Exacerbating the gap between potential earnings and disability benefits is a reduction in private health insurance coverage. Eventual access to Medicare after 2 years on the DI rolls may provide an additional enticement to apply.

The Sequential Disability Determination Process

A claimant typically files an application for DI or SSI in a Social Security field office. The field office gathers a variety of information from the applicant regarding entitlement status, impairment(s), and medical records. The disability determination follows a five-step sequential evaluation process that considers employment, medical, and vocational factors, in that order.

- Step 1: If the applicant is employed and earning more than the SGA amount, an SSA employee denies the claim. Otherwise, the field office sends the claim to the DDS.
- Step 2: If a medical impairment (or combination of impairments) is not severe enough to interfere with basic work-related activities for at least 1 year, a DDS examiner denies the claim. Otherwise, the evaluation proceeds to the next step.
- Step 3: Impairments that meet the criteria in SSA's medical listings or are found to be of equal severity result in an allowance determination. Otherwise, the claim is referred to the next step.
- Step 4: An applicant found with the capacity to engage in relevant employment performed in the past is denied. If not, the evaluation proceeds to the next step.
- Step 5: Based on the applicant's residual functional capacity (RFC), age, education, and work experience, the DDS determines if the applicant could engage in other types of employment. If so, the claim is denied. Otherwise, the claim results in a disability finding.

Motivating the sequential disability determination process is a screening strategy designed to deal first

with cases that can be easily decided on the basis of fairly objective medical tests. If the claimant does not meet or equal the severity requirements in the listings of impairments, the vocational grid is used to determine whether he or she is disabled. The grid incorporates a combination of the following factors: age, RFC, education, and the skill level involved in past work as well as the degree to which those skills can be transferred to another job. Age is divided along four thresholds (younger than age 50, aged 50–54, aged 55–59, and aged 60 or older). RFC is graded into five different categories that assess the exertional limitations of the filer for work-related activities (sedentary, light, medium, heavy, and very heavy work). For the purpose of the vocational grid, SSA divides educational level into four categories (illiterate or unable to communicate in English, limited education or less, high school graduate or more, and recent education that trained the applicant for a skilled job). Assessment of previous relevant work experience leads to the categories of unskilled, semiskilled, and skilled. Finally, the determination process takes into account whether the skills the applicant learned from a past job can be transferred to a new, similar position.

Lahiri, Vaughan, and Wixon (1995) and Hu and others (2001) used household survey data matched to Social Security's administrative records to model the sequential disability determination process. Their findings indicate that the predictive ability of particular variables is linked to their relevance within the stage of determination. For instance, information on activity limitations and medical variables are significant to steps 2 and 3, while the explanatory power of age, past work, and education are manifest in steps 4 and 5.

The Appeals Process

Within 60 days from the notice of denial, the applicant has a number of sequential chances to appeal the decision. There are four stages of appeal. The first stage is a reconsideration by the state DDS, where the case is reviewed by a different examiner and the applicant has the opportunity to submit additional evidence. The second stage involves the Office of Disability Adjudication and Review (ODAR), where the claimant can request a hearing before an ALJ.[2] The ALJ considers any documentary evidence introduced, evaluates the testimony of the applicant, and witnesses that testimony under oath. The third stage in the appeals process is to request a review by the Appeals Council, which is comprised of a panel of ALJs. The Council may choose to grant, deny, or dismiss the request. Upon review, the Council can uphold, reverse, or modify the decision. It can also send the case back to the ALJ for a new hearing. Finally, if the applicant is dissatisfied with the outcome, the fourth stage available is to appeal the case outside of SSA in a federal district court.

Table 1 presents allowance, denial, and appeal rates for disability determinations made at various adjudicative stages by year of application. The table reflects 100 percent of the determinations for workers applying to the DI program only, excluding concurrent applicants to DI and SSI. Results are shown for the combined 8-year period spanning the random data sample in my modeling effort (applications

Table 1.
Allowance, denial, and appeal counts and rates for disability determinations at various adjudicative levels, by selected years 1997, 2004, and the 1997–2004 period

Count and rate of disability determination	1997	2004	1997–2004
		Initial level	
Number			
Determinations	551,909	736,987	5,151,351
Allowances	228,793	329,523	2,319,171
Denials	323,116	407,464	2,832,180
Appeals	206,148	248,232	1,778,805
Percent			
Allowance rate	41.45	44.71	45.02
Denial rate	58.55	55.29	54.98
Appeal rate	63.80	60.92	62.81
		Reconsideration level	
Number			
Determinations	206,148	248,232	1,778,805
Allowances	33,373	28,707	255,201
Denials	172,775	219,525	1,523,604
Appeals	141,021	185,672	1,288,257
Percent			
Allowance rate	16.19	11.56	14.35
Denial rate	83.81	88.44	85.65
Appeal rate	81.62	84.58	84.55
		Hearing level or above	
Number			
Determinations	141,021	185,672	1,288,257
Allowances	107,539	151,122	1,009,799
Denials	33,482	34,550	278,458
Percent			
Allowance rate	76.26	81.39	78.38
Denial rate	23.74	18.61	21.62

SOURCE: Author's tabulations based on the *Annual Statistical Report on the Social Security Disability Insurance Program, 2008.*

from 1997 through 2004), as well as separately for 2 individual years (the first (1997) and last (2004)).[3] The initial disability allowance rate within the 8-year period considered stands at about 45 percent. Roughly, 63 percent of initial denials are appealed at the reconsideration stage, which results in a fairly small portion of reversals by the DDS (about 14 percent). However, 85 percent of the reconsideration denials are appealed. Once the third and fourth stages in the appeals process are reached (at the hearing level or in a federal court), denials are reversed at a rate of 78 percent. As a result, after the appeals process takes its course, the 45 percent initial disability allowance rate increases to an overall allowance rate of 70 percent.

Multiple factors can contribute to the high reversal rate of initial denials. The most obvious explanation is that many impairments can worsen over time, particularly disorders that are of a degenerative nature. One feature of the DI program is that at every stage of the appeals process the claimant has an opportunity to introduce additional medical evidence. Therefore, it is possible that ALJs are making decisions based on a more extensive information set that was simply not available to state DDS examiners. Moreover, unlike with the DDS appeals procedure, applicants at the hearing level or above are much more likely to retain legal counsel. Claimant representation benefits from detailed knowledge of the rules and process. This can be helpful in developing medical evidence that may include additional symptoms and impairments not claimed at the DDS level. In this context, the Social Security Advisory Board (2001) has made a number of recommendations addressing some of the procedural differences between the adjudicative levels (such as the fact that most claimants lack any face-to-face interaction with an adjudicator until they get to an ALJ hearing). Finally, by its very nature, the appeals process could be inducing a selection bias effect, where only the applicants with the strongest evidence appeal a denial. In fact, one possible route to selection bias is the use of legal counsel. After all, attorneys are likely to prescreen potential clients in order to represent those with the highest probability of an allowance.[4]

Previous Literature

SSA's statutory definition of disability in terms of "ability to work" is inevitably open to subjective judgment on the part of decision makers. In a minority of cases, proof of a specific impairment will qualify the filer for expedited case processing under the Compassionate Allowance (CAL) initiative, based on minimal, but sufficient objective medical information. Roughly, about a third of allowances are decided on the medical evidence alone (step 3), but even physicians may disagree over the interpretation of diagnostic tests. Most claimants are unlikely to neatly fit precisely defined eligibility criteria, and program guidelines can be subject to interpretation. In some instances, federal courts have issued decisions that at least for a while resulted in different disability policies for different parts of the country.[5] Moreover, individuals vary in their ability to withstand pain and in their response to treatment, so that one person facing a specific set of limitations may be able to work, while another may not. Once vocational considerations such as RFC, relevant past work experience, and transferable skills are criteria in the determination process, the decision becomes increasingly complex. For these reasons alone, one would expect some degree of heterogeneity in disability outcomes.

The literature evaluating factors that affect allowance rates in Social Security's disability programs is extremely sparse. More effort has been devoted to investigating the determinants of application rates. Rupp and Stapleton (1995) summarized earlier contributions, while Rupp (2012) discussed more recent work. A growing body of evidence using different methodology and various sources of data suggests that application rates increase with labor market shocks. Higher unemployment reduces the opportunity cost of applying for marginally qualifying individuals, who must weigh their current earnings and future labor opportunities against the present value of benefits. Thus, application rates are expected to rise in response to a labor market shock. Additionally, the increase in marginally qualified applicants is anticipated to produce a decline in allowance rates, as those filers have a harder time qualifying through the determination process.

For over a decade, the Social Security Advisory Board (2001, 2006, 2012a) has been tracking the two main sources of variation in allowance rates referenced in this article (by state and adjudicative stage), calling for a major overhaul to the disability programs. Among its suggestions, the Board advocates strengthening the federal/state arrangement to decrease the large disparities that exist between different states regarding staff salaries, educational requirements, training, and attrition rates. The Board also recommends reforming the hearing process by establishing uniform procedures for claimant representatives;

having the government represented at the ALJ hearing level or above; and closing the record after the ALJ decision, so that cases do not change substantially at each level of appeal.

Using a combination of aggregate time-series and cross-sectional methodology, Rupp and Stapleton (1995) found a positive relationship between the state unemployment rate and both initial applications and awards. Their modeling of allowance rates suggested the presence of lagged effects. Specifically, the authors estimated that a 1 percentage point increase in the unemployment rate was associated with a 1 percent decline in the initial allowance rate in the first and second years following the year in which the unemployment rate changed.

State allowance rates depend on the economic, demographic, and health characteristics of the applicants, which vary among the states. For instance, states with older populations are anticipated to have higher disability allowance rates on average. Older applicants are more likely to qualify because of the higher prevalence of age-related disabilities and the fact that they face less stringent program standards than do younger individuals. Using state-level data over a 3-year period (1997–1999), Strand (2002) estimated that as much as half of the variation in initial allowance rates may have been attributable to state differences in economic and demographic factors. The author found a negative association between filing rates and allowance rates and a statistically significant negative impact of unemployment on allowance rates. Institutional considerations can also play a role in explaining observed heterogeneity in disability outcomes. For instance, Coe and others (2011) found that states with mandated health insurance and longer duration for Unemployment Insurance benefits were associated with lower application rates.

In a recent article, Rupp (2012) used individual-level data over the 1993–2008 period to investigate three factors affecting initial allowance rates: (1) the demographic characteristics of applicants, (2) the diagnostic mix of applicants, and (3) local labor market conditions. The modeling approach involved a binary logit process with fixed-effects for state of origin and year of determination. Explanatory variables included the state unemployment rate and indicators for sex, age group, impairment type,[6] and the presence of a secondary diagnosis code in the data. The author found these three sets of variables statistically significant. All else equal, male and older adult applicants had a higher likelihood of an initial allowance. Likewise, an increase in the state unemployment rate was associated with a decline in the probability of an initial allowance, with the size of the effect changing substantially by body system. The size of the state fixed-effects suggested that a substantial portion of the variation in state initial allowance rates could be attributed to permanent differences among the states.

Keiser (2010) explored the variation in self-reported (as opposed to actual) allowance rates among DDS examiners in three undisclosed states. The study approached the subject of outcome variation in disability decision making from the perspective of the theory of bounded rationality. The surveys mailed to DDS examiners considered a number of factors, including: (1) ideological identification; (2) adherence to conflicting goals (aiding disabled individuals, while protecting US tax payers from fraud); (3) perception about applicants' honesty in representing their limitations; and (4) the expectations of examiners' immediate supervisors (a focus on allowances, denials, or both equally). The model was able to account for only 12 percent of the variation in self-reported allowance rates. One aspect of the study relevant to the objectives here relates to the evidence of a possible policy feedback mechanism. In particular, knowledge of the extent to which ALJs reverse initial denials was found to be a factor in explaining higher reported allowance rates among examiners.

Data and Methodology

The Disability Research File (DRF) is a data file designed to longitudinally track a cohort of filers through 10 years of the disability decision and appeal process. Prompted by concern from Congress regarding the size of the disability rolls, the file—originally built in 1993—is updated once a year, with the 3 most recent years of claims data completely built from scratch. Because of differences in the structure of DI and SSI records (Title II and Title XVI, respectively, under the Social Security Act), two separate files are compiled that draw from multiple administrative data sources in a process that usually takes several months to complete. The file is unique in its ability to provide information about the status of a claim in its progression throughout the adjudicative stages, as well as activity about claimants who file multiple disability applications.

For this study, I work with a 10 percent random sample of an abbreviated version of the DRF, tracking 10 years of longitudinal disability claims (1997–2006). The analysis is restricted to medical determinations involving workers aged 18–65 who applied to the DI

program during the 8-year period from 1997 through 2004. The latter is the most recent year in the file for which the percentage of pending applications is negligible. Moreover, the focus is on DI medical claims only. In particular, technical denials are excluded because they generally lack the evaluation of any medical evidence.[7] Concurrent applicants to the DI and SSI programs are also excluded, as they represent a unique population that has enough work experience to qualify under DI, but that is poor enough to meet SSI's criteria. A look at the *Annual Statistical Report on the Social Security Disability Insurance Program* (SSA 2009, Tables 60 and 62) validates this decision. Nonconcurrent DI workers systematically experience higher allowance rates at the initial and hearing levels than concurrent workers. Furthermore, Rupp (2012, Table 1) illustrates how the age structure and diagnostic mix of both populations can differ substantially. Concurrent filers tend to be younger and have a much larger share of mental diagnoses. Thus, it seems appropriate to treat DI-only, concurrent, and SSI-only claimants as separate populations.

Formally, the adjudicative-level process can be thought of as a sequential interaction between two parties (Social Security and the applicant). Conditional on a claimant applying to the disability program, Social Security makes a decision to allow or deny. Likewise, conditional on a denial, the applicant decides whether or not to appeal. The sequence continues, with the process ending upon an allowance, a decision not to appeal, or exhaustion of all appeals opportunities. While the appeals decision is always made by the same individual (the applicant), the decision to allow or deny can be made by a field office representative, an examiner at the DDS, an ALJ, or even a federal judge. Complicating matters further is the Prototype program, which breaks the order of the sequence by allowing several states to skip the reconsideration adjudicative level.

This article focuses on the prediction of outcomes as a purely statistical classification problem. I do not model the sequential structure of the decision-making process. For purposes of this study, the disability determinations in the file are separated into four mutually exclusive categories: (1) initial allowances, (2) initial denials not appealed, (3) final allowances, and (4) final denials. This classification of the data implicitly reduces the adjudicative process to two stages. Specifically, the first two categories (initial allowances and initial denials not appealed) represent outcomes at the initial DDS level. The last two categories (final allowances and final denials) result once the applicant decides to stop appealing or exhausts the appeals process. This can occur at the reconsideration DDS level, at the hearing level, or in a federal court. In other words, what triggers the difference between the two adjudicative stages is a decision to appeal an initial denial. However, because of the low allowance rate and high appeal rate at the reconsideration stage (see Table 1), the large majority of decisions falling into the final allowance and final denial categories occur at the hearing level or above.

Table 2 breaks down the count and proportion of sample observations by adjudicative disability outcome. In the top panel of the table, out of a random sample comprising 462,578 observations, 46.2 percent of applicants receive an initial allowance, while 19.4 percent decide not to appeal an initial denial. The percentages of claimants that end up in the final allowance and final denial categories are 24.9 percent and 9.5 percent, respectively. For comparison, the bottom panel of the table displays equivalent quantities

Table 2.
Number and percent of sample observations, by adjudicative disability category, 1997–2004

Count and proportion	Initial		Final		Total
	Allowances	Denials not appealed	Allowances	Denials	
			10 percent random sample		
Number	213,851	89,796	115,112	43,819	462,578
Percent	46.23	19.41	24.88	9.47	99.99
			100 percent data file		
Number	2,319,171	1,053,375	1,265,000	513,805	5,151,351
Percent	45.02	20.45	24.56	9.97	100.00

SOURCE: Author's calculations based on a 10 percent sample of the DRF and Table 1.

NOTE: Values may not sum to 100 because of rounding.

corresponding to the full data set. The outcome proportions in the 10 percent random sample suggest an adequate approximation to the population of DI claimants over the 8-year period.[8]

Summary statistics of the explanatory variables used in my modeling effort appear in Table 3. Age at filing is the only continuous predictor. As illustrated in a later section of this article, the age profiles associated with the disability outcomes are highly nonlinear. In the models, I include both age and its square as a means to capture the nonlinearity. The mean age of all filers in the sample is about 50, but on average, claimants receiving an initial allowance tend to be 2 years older, while those in the final denials category have a mean age of less than 47. All else equal, it is expected that an increase in age would be positively associated with the likelihood of an initial allowance.

The models include binary indicators for sex (1 if male), for reapplication (1 if the claimant has applied to the DI program before), and for having zero earnings in the year before application (1 for zero earnings). Males comprise 52 percent of all filers in the sample, but make up 56 percent of claimants receiving an initial allowance. All else equal, it is expected that males would have a higher probability of an initial allowance. The two remaining indicator variables (reapplicants and claimants with zero earnings in the year before filing) are included because of their potential to serve as proxies for marginally qualified applicants, however imperfectly.

Following the DRF documentation, I use a 10-year window to classify an individual as having previously applied. That is, a new claimant is a person who is actually a first-time applicant or whose previous DI application dates back at least 10 years. About 17 percent of filers in the sample are reapplicants, compared with only 12 percent of those receiving an initial allowance. Notice how outcomes in the final adjudicative stages tend to have a higher share of claimants with a prior application history. Thus, it is expected that new applicants would have a higher likelihood of an initial allowance. Finally, the focus turns to a claimant's lack of earnings in the year before filing to identify those with the highest immediate financial incentive to apply. Throughout this study, such applicants are referred to as unemployed (Table 3). About 19.5 percent of claimants in the sample had zero earnings in the year before applying, compared with 24 percent and 28 percent of those in the initial denials not appealed and final denials categories, respectively. All else equal, it is anticipated that applicants with nonzero earnings in the year before filing would have a higher probability of an initial allowance.

The last explanatory variable used here is a derived field in the DRF, representing a discrete earnings index. The earnings index is constructed using the Department of Labor's official minimum wage and Social Security's Office of the Chief Actuary's national income averages. An applicant's individual earnings are compared with the minimum wage and the national income average in order to assign a numerical value (from 1–5) that indicates whether the claimant's earnings are below or above the national average. Among allowed claims, the index encompasses the 2nd through 6th years of earnings prior to

Table 3.
Summary statistics of explanatory variables (in percent)

Variable	Initial		Final		Total (variable category)
	Allowances	Denials not appealed	Allowances	Denials	
Male	56.29	49.21	49.20	48.15	52.38
Reapplicant	11.82	18.62	22.09	23.41	16.79
Unemployed	15.13	24.21	20.72	27.63	19.47
Earnings					
Marginal	22.47	36.91	23.63	37.45	26.98
Low	25.50	29.46	29.96	28.88	27.70
Average	26.34	19.63	25.34	19.59	24.15
High	18.44	10.77	15.88	10.86	15.60
Very high	7.25	3.23	5.19	3.22	5.58
Age					
Mean	52.15	47.39	49.58	46.76	50.08
Standard deviation	10.10	10.80	8.52	9.31	10.03

SOURCE: Author's calculations based on a 10 percent random sample of the DRF.

the established date of disability onset. Among the denied claims, the earnings index comprises the 2nd through 6th years of earnings before the filing date. The rationale in choosing this time frame is based on a desire to avoid potential bias that is due to a sharp decline in earnings in the most recent years because of the gradual onset of disability. The earnings index categories are as follows:

1. Marginal earnings.
2. Low earnings—mean earnings exceed marginal earnings, up to 75 percent of the national average.
3. Average earnings—mean earnings fall between 75 percent and 125 percent of the national average.
4. High earnings—mean earnings fall between 125 percent and 200 percent of the national average.
5. Very high earnings—mean earnings above 200 percent of the national average.

While zero earnings in the year before filing (defined here as unemployed) reflects a claimant's immediate incentive to apply, the earnings index encompasses the future earnings potential that the applicant must renounce in order to receive DI benefits. Roughly, 27 percent of filers have marginal earnings, which tend to distribute more heavily among the denial categories (36.9 percent of initial denials not appealed and 37.5 percent of final denials). That trend reverses for average, high, and very high earners. For instance, 15.6 percent of claimants in the sample are high earners. However, among applicants receiving an initial or a final allowance, their shares are 18.4 percent and 15.9 percent, respectively. Meanwhile, the proportion of high-income filers in each of the initial denials not appealed and final denials categories is less than 11 percent. All else equal, it is anticipated that higher earnings would be associated with a higher probability of an initial allowance.

The Models

The Bayesian approach to inference embodies the idea of learning from experience, through which new evidence is integrated with existing knowledge. Given observed data, a researcher (classical or Bayesian) makes probabilistic assumptions about how that data were generated (the data distribution or data model). The model contains a number of unknown parameters and the goal is typically to reach statistical conclusions about their values. Bayesian statisticians include a second element to the model (the prior distribution), which reflects prior uncertainty about the parameter values. Those two elements are combined through a mechanism known as Bayes's theorem to derive the so-called posterior distribution. The *posterior probability distribution* results from conditioning on the observed sample and reflects how the information in the data modifies prior knowledge. Once available, it can be used to report point estimates of the parameters, construct credible intervals and regions of the parameter space associated with some posterior probability, and estimate the posterior predictive density associated with future observations.

The *prior probability distribution* (often called the prior) provides a formal mechanism to explicitly incorporate available nonsample information. The prior might be specified to accommodate the empirical evidence of previous studies or for purely economic or statistical theory considerations. It may also aim at simply reflecting the views of the researcher. These are examples of informative priors. On the other hand, diffuse or noninformative priors aim at representing a lack of prior knowledge, by minimizing the influence of the prior on the resulting posterior distribution. At any rate, when a large sample of observations is involved, the data density usually dominates the prior, so that the choice of prior is inconsequential in terms of the derived posterior inference.[9]

The Bayesian models estimated in this analysis closely follow the description and algorithmic implementation in Rossi, Allenby and McCulloch (2005). I estimate separate hierarchical multinomial logit models that cluster the claimant-level data into states and into diagnoses. Appendix Tables A-1 and A-2 present sample counts by disability outcome for the 181 primary impairments and 50 states, respectively. Following Congdon (2005), a hierarchical multinomial logit model is often defined by the nature of the individual-level explanatory variables entertained. In this application, all of the available predictors are invariant with respect to the adjudicative disability outcome. As a result, the specification becomes a pure multinomial logit model with category-specific parameters. The parameters for a baseline outcome are typically set to zero to avoid model indeterminacy. In all cases, final denials represent the baseline. Thus, for a particular cluster (a specific state or diagnosis) and a particular outcome (an initial allowance, an initial denial not appealed, or a final allowance), there is a distinct set of parameters associated with the following explanatory variables:

- An intercept.
- A binary variable taking the value of 1 if the individual has applied to the DI program before.

- A binary variable taking the value of 1 if male.
- A discrete earnings index taking values of 1 through 5.
- A binary variable taking the value of 1 if the individual had zero earnings in the year before applying.
- The applicant's age at filing.
- The square of the applicant's age at filing.

One way to think of a hierarchical model is as a compromise between two extreme solutions. On the one hand, I could disregard the state of origin and the primary diagnosis codes and estimate a multinomial logit model that pools all the claimants together. For comparison, estimates from such a model are provided. Alternatively, I could estimate a separate model for every state and every impairment. That approach would be problematic for those groups with few observations, which is the case for many of the individual impairments. Instead, the hierarchical version of the model can be seen as a set of multinomial logit processes that are linked together through a common distributional assumption. That is, the individual parameters are assumed to derive from a multivariate normal distribution (often referred to as the heterogeneity distribution), with unknown mean and covariance matrix. Estimates of the covariance matrix can be used to decompose outcome variation into its within-group and between-group components (see for instance, Raudenbush and Bryk (2002)). Moreover, unlike the nonhierarchical version of the multinomial logit model, my approach can accommodate the possibility of correlation between the groups, although not within the groups. Finally, one virtue of hierarchical models lies in their ability to diminish the influence of outlying observations. That property (often referred to as shrinkage) is desirable in circumstances where many of the clusters contain few observations. The result is usually more reasonable parameter estimates that are not skewed by the scarcity of data or the influence of outliers in specific groups.

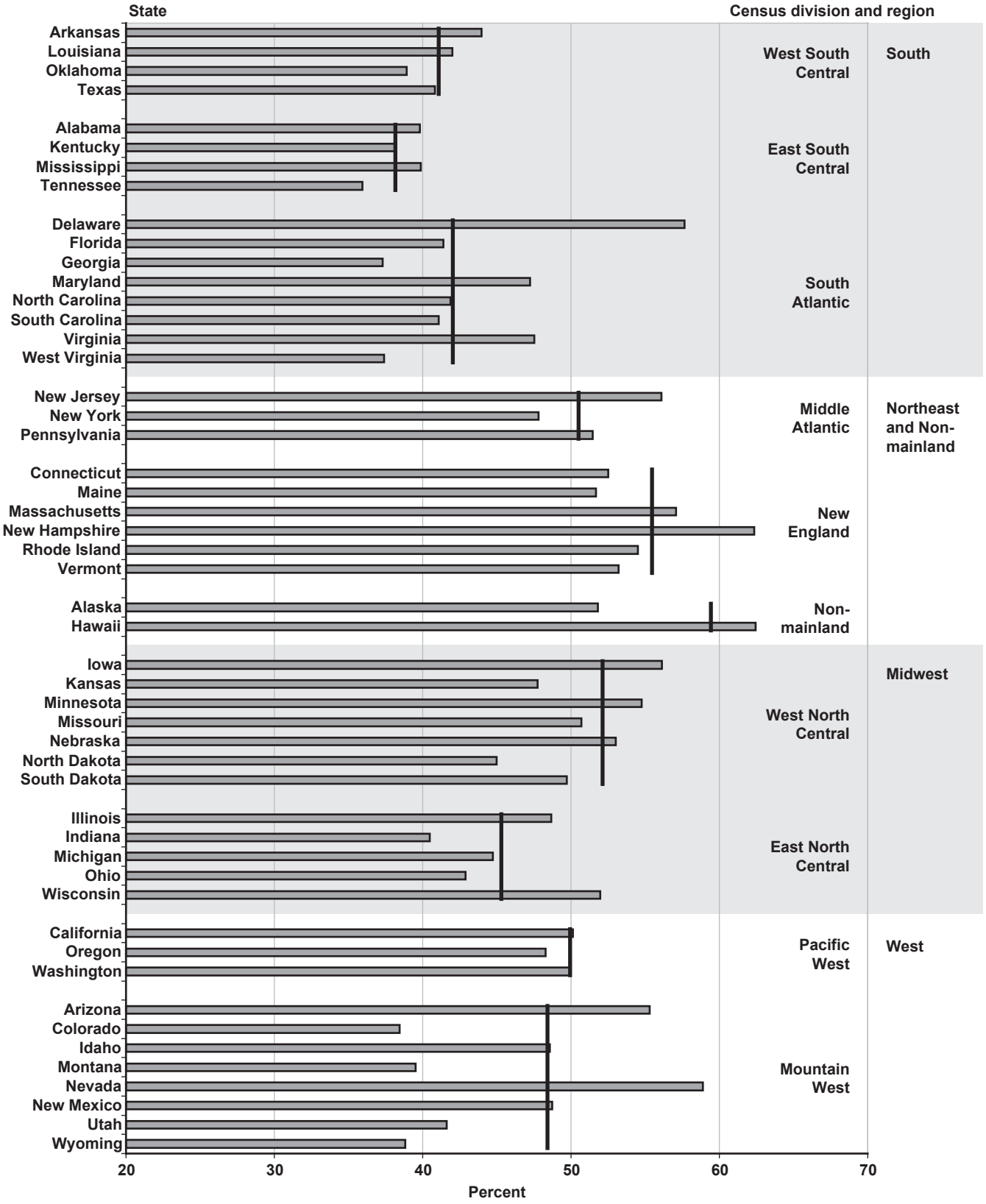
Once posterior estimates of the parameters are available, the models can be used to generate probability predictions.[10] Given specific values of the explanatory variables, three separate equations generate linear predictions for an initial allowance, an initial denial not appealed, and a final allowance (by default, the linear prediction for a final denial takes a 0 value). These linear predictions can be transformed into probabilities using standard formulae associated with the logit model. It is important to keep in mind the distinction between a linear prediction and a probability. For a given outcome (say an initial allowance), the linear predictions allow comparison of how all the clusters (the states or diagnoses) rank within that outcome. On the other hand, the probability that the *i*-th applicant in the *j*-th group falls into say the initial allowance category is computed using the linear predictions for all four adjudicative disability outcomes combined. Thus, within a given cluster, the estimated probabilities of an initial allowance, an initial denial not appealed, a final allowance, and a final denial add to 100 percent, as they track the observed proportions in the data sample.

State Variation

The disability outcomes in the sample for all 50 states are listed in Appendix Table A-2. In terms of sample size, California contributes 10.1 percent of total applicants, followed by New York, Florida, and Texas. These four states combined account for more than a quarter of all claimants. At the other end of the spectrum, Alaska comprises a mere 0.12 percent of the total observations (552), followed by Wyoming, North Dakota, and South Dakota. The graphs in Chart 1 display initial allowance rates by state, grouped according to the Census Bureau regions and divisions. The black vertical lines denote the overall initial allowance rate for a particular division, with the horizontal bars corresponding to each individual state. For geographical reasons, I place Alaska and Hawaii in the Nonmainland category, although technically, those two states are counted as part of the Pacific-West division.

In terms of initial allowance rates, the four states with the lowest values are southern states: Tennessee (35.9 percent), Georgia (37.3 percent), West Virginia (37.4 percent), and Kentucky (38.1 percent). On the other hand, Hawaii leads with the highest initial allowance rate at 62.5 percent, followed by New Hampshire (62.3 percent), Nevada (58.9 percent), and Delaware (57.7 percent). Thus, the range of state variation in initial allowances (the difference between Hawaii with the highest initial allowance rate and Tennessee with the lowest rate) is roughly 25 percentage points. Chart 1 does not appear to reveal any clear-cut geographical patterns other than perhaps the contrast between the South and New England. Specifically, the three divisions with the lowest initial allowance rates are the southern ones (West South Central, East South Central, and South Atlantic). Clearly, Delaware and to a lesser extent Maryland and Virginia appear to be outliers in the South Atlantic division and more at

Chart 1.
Percentage of initial allowances, by state and Census division and region



SOURCE: Author's calculations based on a 10 percent random sample of the DRF.

NOTE: The black vertical lines indicate the percentage for each Census division.

home in the Middle Atlantic division. Overall, however, it is fair to say that southern states tend to have low initial allowance rates. New England, on the other hand, is the Census division with the highest allowance rate.

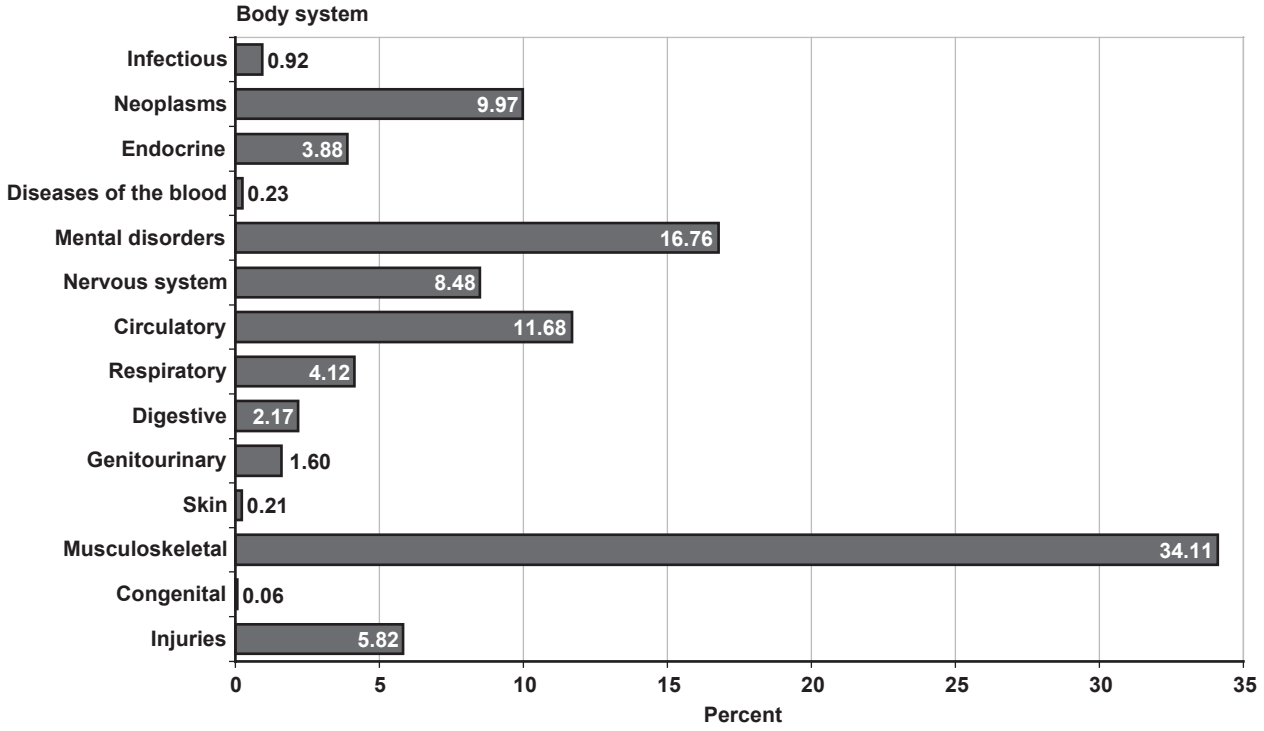
Diagnosis Variation

SSA maintains a classification of impairments that identify the medical conditions on which disability-related claims are based. Since 1985, the coding of primary and secondary diagnoses has approximately followed the *International Classification of Diseases: 9th Revision* (ICD-9) taxonomy. Appendix Table A-1 summarizes the disability outcomes for 181 medical impairments, which are grouped into 14 body systems.[11] Notice that I employ the body system for descriptive purposes only, as a means of grouping individual diagnoses. To this end, each impairment is uniquely matched to a single body group, following the description in the *SSA Program Data User's Manual* (Panis and others 2000).

The primary diagnosis field in the data is generally based on the latest Form SSA-831 at the DDS level, but will be assigned based on an alternative source if that field is incomplete. There is evidence that on appeal, some claimants will be evaluated on the basis of a different primary diagnosis. That may occur for a number of reasons. Typically an adjudicator designates the primary impairment at the time of the decision, based on the medical evidence. However, many disability claims allege multiple impairments. Moreover, impairments may worsen and new diagnoses develop over time. As a result, additional medical evidence introduced on appeal can lead an adjudicator to change the primary impairment. Unfortunately, the DRF does not identify changes in the primary diagnosis throughout the adjudicative process. Such events are not accommodated in this analysis. An audit report from Social Security's Office of the Inspector General (SSA 2010) found that a switch in the primary diagnosis was common for three of the four impairments most likely to be denied at the initial level and allowed at the hearing level in the 2004–2006 period. These three impairments (diabetes mellitus; osteoarthrosis and allied disorders; and muscle, ligament, and fascia disorders) are prone to worsen over time and affect other body systems.[12]

Chart 2 displays the percentage of claimants in each body system for the entire sample. Musculoskeletal impairments account for 34 percent of the diagnoses,

Chart 2.
Percentage of claimants, by body system



SOURCE: Author's calculations based on a 10 percent random sample of the DRF.

followed by mental disorders with 17 percent. Those two body systems combined make up slightly over half of all observations. Circulatory diseases and neoplasms represent 12 percent and 10 percent of all outcomes, respectively. The nervous system and sense organs category comprises 8 percent of the impairments, while injuries make up 6 percent. Both the respiratory and the endocrine, nutritional, and metabolic body systems account for about 4 percent of claimants each. Likewise, each of the digestive and genitourinary body systems represents 2 percent of all diagnoses. Infectious and parasitic diseases contribute almost 1 percent of the observations. Finally, the remaining body groups (congenital anomalies and both diseases of the skin and subcutaneous tissue and blood and blood forming organs) represent well below 1 percent of cases combined.

A cursory look at Appendix Table A-1 reveals that one or a few primary diagnoses codes may sometimes account for the bulk of diagnoses within a body system. The tabulation below highlights selected cases. For example, disorders of the back and osteoarthrosis represent 56 percent and 21 percent of all musculoskeletal impairments, respectively, while affective and mood disorders make up more than half of the mental diagnoses. Diabetes and obesity respectively contribute 63 percent and 31 percent of claimants to the endocrine, nutritional, and metabolic body system. Four types of cancers (lung, breast, colon, and genital organs) comprise over 50 percent of the neoplasms.[13] Similarly, symptomatic HIV infections are more than half of all infectious and parasitic disorders. Chronic liver disease and cirrhosis accounts for 56 percent of digestive impairments, while about 67 percent of respiratory ailments involve chronic pulmonary insufficiency. Finally, 85 percent of the genitourinary impairments are chronic renal failure, which explains the high initial allowance rate of this body system.

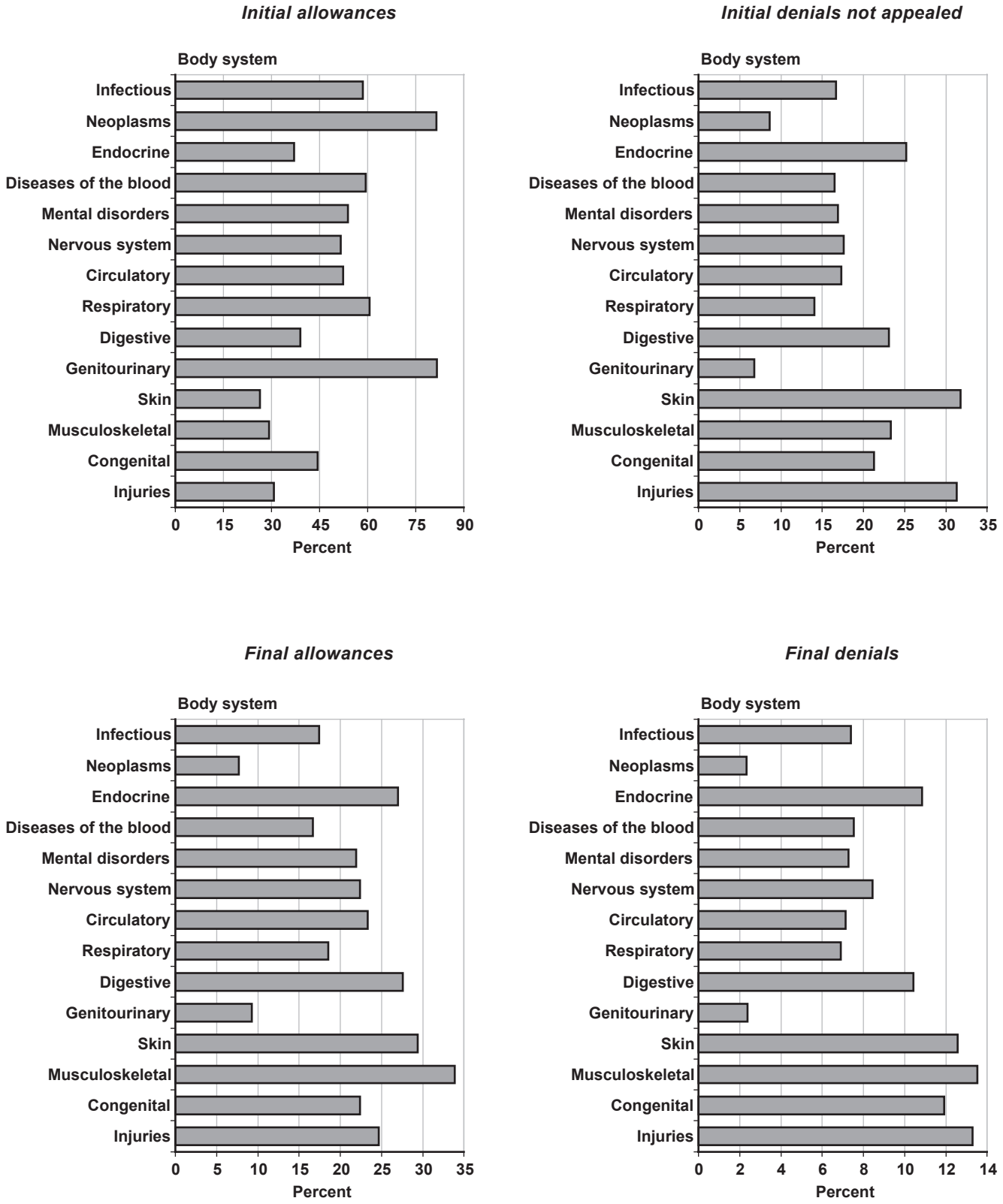
Impairment	Percent
Musculoskeletal	
Disorders of the back—discogenic and degenerative	55.7
Osteoarthrosis and allied disorders	20.8
Mental	
Affective/mood disorders	55.7
Neoplastic	
Malignant cancers of the—	
Trachea, bronchus, or lung	19.0
Breast	15.5
Colon, rectum, or anus	10.0
Genital organs	9.2
Respiratory	
Chronic pulmonary insufficiency	66.7
Endocrine, nutritional, and metabolic	
Diabetes	62.6
Obesity and other hyperalimentation disorders	30.7
Digestive	
Chronic liver disease and cirrhosis	55.7
Genitourinary	
Chronic renal failure	84.9
Infectious and parasitic	
Symptomatic HIV infections	52.8

There is huge variation in disability outcomes by primary diagnosis. Chart 3 illustrates the proportion of decisions that correspond to each body system. The overall proportion of initial allowances in the sample is 46.2 percent (Table 2). However, over 80 percent of genitourinary and neoplastic impairments receive an initial allowance, while the share drops to 26.3 percent for skin disorders and to about 30 percent for musculoskeletal diagnoses. Thus, the range of variation in initial allowances among the body systems is roughly 55 percentage points. In general, the genitourinary and neoplastic body systems have the highest initial rates of allowance, exceeding any other group by at least 20 percentage points. As a result, those two groups also have the lowest proportions of initial denials not appealed, final allowances, and final denials. Applicants with injuries and skin impairments appear most likely not to appeal an initial denial, with about 31 percent of the outcomes. Musculoskeletal diagnoses have the highest proportion of final allowances, with about 34 percent of the outcomes, followed by skin disorders. In addition to injuries, however, musculoskeletal and skin impairments also exhibit the highest rates of final denials.

Mortality Variation

One source of concern regarding the categorization of outcomes in this analysis is a potential biasing effect that is due to death. Specifically, claimants with an initial denial could die before having a chance to appeal. Our DRF sample identifies an applicant's date of death over the 11-year period from 1997 through 2007. It is of course impossible to determine from the data which deaths occurred as a direct result of the underlying disability impairment. Nevertheless, this information is used to compute raw death rates (adjusted neither by age or sex) over the period in question. For the different body systems, Table 4 shows the proportion of applicants in every adjudicative outcome that passed away. About 17 percent of all claimants died during this period. However, while 28.4 percent of

Chart 3.
Percentage of adjudicative disability categories, by body system



SOURCE: Author's calculations based on a 10 percent random sample of the DRF.

the applicants in the initial allowance category died, only about 7 percent of claimants who did not appeal an initial denial did not survive to 2007. Among those, two-thirds passed away at least 3 years after their application. Consequently, the potential fraction of applicants who died before having the chance to appeal would be too marginal to affect this analysis in any material way.

Deaths occurred more frequently among the most medically serious diagnoses. In terms of all outcomes, the body system with the lowest rate of mortality during the 11-year period is musculoskeletal, which is followed by injuries, mental disorders, and skin impairments. The diagnostic groups with the highest proportion of deceased claimants are neoplasms, followed by genitourinary impairments, diseases of the blood and blood forming organs, respiratory diagnoses, and digestive disorders. Given the DI program's goal to serve claimants in greater need more expeditiously, it is reassuring to see that the proportion of deceased claimants in every single body system is highest among those initially allowed and second highest for filers in the final allowance category.

It is also worth recalling that disability in the DI program is defined on the basis of long-term inability to work. As a result, death proportions and initial allowance rates are not expected to always go hand in hand. For instance, 82 percent of claimants with a neoplasm disorder who receive an initial allowance die within the 11-year period under consideration. For corresponding applicants with a genitourinary disorder (85 percent of whom have a diagnosis of chronic renal failure), mortality is lower (39 percent). Nevertheless, both body systems have similar initial allowance rates of roughly 81 percent. Standard treatments for those two impairments (such as chemotherapy and dialysis) likely pose equally severe barriers to work, even if one kind of diagnosis is much more deadly in the short run.

Age Variation

Another relevant factor of variation in disability adjudicative outcomes is age. Three important characteristics are identified in the data:

1. The proportion of outcomes by single year of age is both highly nonlinear and pretty regular from one year to the next.
2. There are distinct patterns at ages 50 and 55, which represent threshold points in the vocational grid.
3. There is an age-62 effect that results from an influx of early retirement applicants. As pointed out by Leonesio, Vaughan, and Wixon (2003), it is a common procedure at SSA field offices to compare the potential benefits to which an applicant is entitled under more than one program. What this means in practice is that early retirees with health problems often apply concurrently for retirement and disability benefits.

Table 4.
Percentage of applicant deaths, by adjudicative disability category and body system, 1997–2007

Body system	Initial		Final		Total (claimant deaths in the period)
	Allowances	Denials not appealed	Allowances	Denials	
All	28.37	6.94	8.75	5.62	17.17
Infectious	30.55	10.06	14.88	7.03	22.66
Neoplasms	82.27	21.88	38.78	16.54	72.21
Endocrine	23.84	11.82	14.54	10.48	16.86
Diseases of the blood	43.50	9.83	18.86	8.86	31.24
Mental disorders	8.49	5.41	6.74	5.25	7.35
Nervous system	16.03	5.40	7.59	5.58	11.39
Circulatory	25.30	12.49	14.31	10.05	19.44
Respiratory	37.83	11.42	15.08	9.29	27.95
Digestive	47.50	11.67	18.36	10.77	27.37
Genitourinary	39.20	9.80	20.26	10.80	34.79
Skin	15.69	5.19	8.42	4.10	8.76
Musculoskeletal	7.72	4.01	4.97	3.68	5.37
Congenital	18.70	8.47	9.68	3.03	12.64
Injuries	12.33	4.39	5.61	3.91	7.07

SOURCE: Author's calculations based on a 10 percent random sample of the DRF.

Chart 4 displays the number of claimants for each adjudicative disability outcome by single year of age (18–65). Because the focus here is on workers covered by the DI program, the total number of applicants at the youngest ages represents a tiny fraction of the sample (239 claimants at age 18 out of more than 462,000 observations). At ages 30–49, the rate at which applicants join the initial allowance category is fairly constant, but increases sharply by age 50 (top graph on the left). There are also noticeable spikes at ages 55 and 62, the latter representing a peak with over 14,000 observations. On the other hand, the number of claimants initially denied who decide not to appeal rises at a fairly constant rate up until about age 42, but levels off subsequently. The most remarkable feature in the top right graph of Chart 4 is the huge spike at age 62. The number of applicants at age 62 in this category totals more than twice that of filers at ages 61 or 63. This suggests that a substantial portion of concurrent early retirement and DI applicants receive an initial denial and decide against filing an appeal. The graph on final allowances (bottom left) shows visible spikes at ages 50 and 55, while final denials experience a jump at age 62 (bottom right).

The proportion of outcomes (rates) by single year of age is shown in Chart 5. The thin discontinued lines in the chart denote the age profiles for each individual year from 1997 through 2004, while the continuous thick line corresponds to the full 8 years of data combined. The proportion of initial allowances by age displays a distinct convex "u-shape," while initial denials not appealed, final allowances, and final denials roughly follow a concave profile in the form of an "inverted-u." These patterns exhibit a great deal of regularity from one year to the next.

For the youngest claimants, the initial allowance rate is very high, ranging from 60 to 70 percent at ages 18–23 (top graph on the left). Then, the rate declines rapidly, reaching 34 percent by age 30, where it remains stable in the low-to-mid 30 percent range until age 49. The subsequent increase resembles a piece-wise linear function with discontinuities at ages 50 and 55 and a dip at the early retirement age. The rate of initial denials not appealed (top graph on the right) rises from about 20 percent at age 20 to its peak of 35.5 percent by age 27. It steadily declines from this point forward, reaching its lowest value of 11 percent at age 59. As retirement nears, the rate increases again, with the early retiree effect inducing a sizeable jump at age 62. The final allowance rate (bottom graph on the left) rises steadily to its peak of 34 percent at age 50, declining rapidly afterwards. Finally, the rate of final denials (bottom graph on the right) hovers below 15 percent at ages 32–48, declining to about 5 percent by age 55.

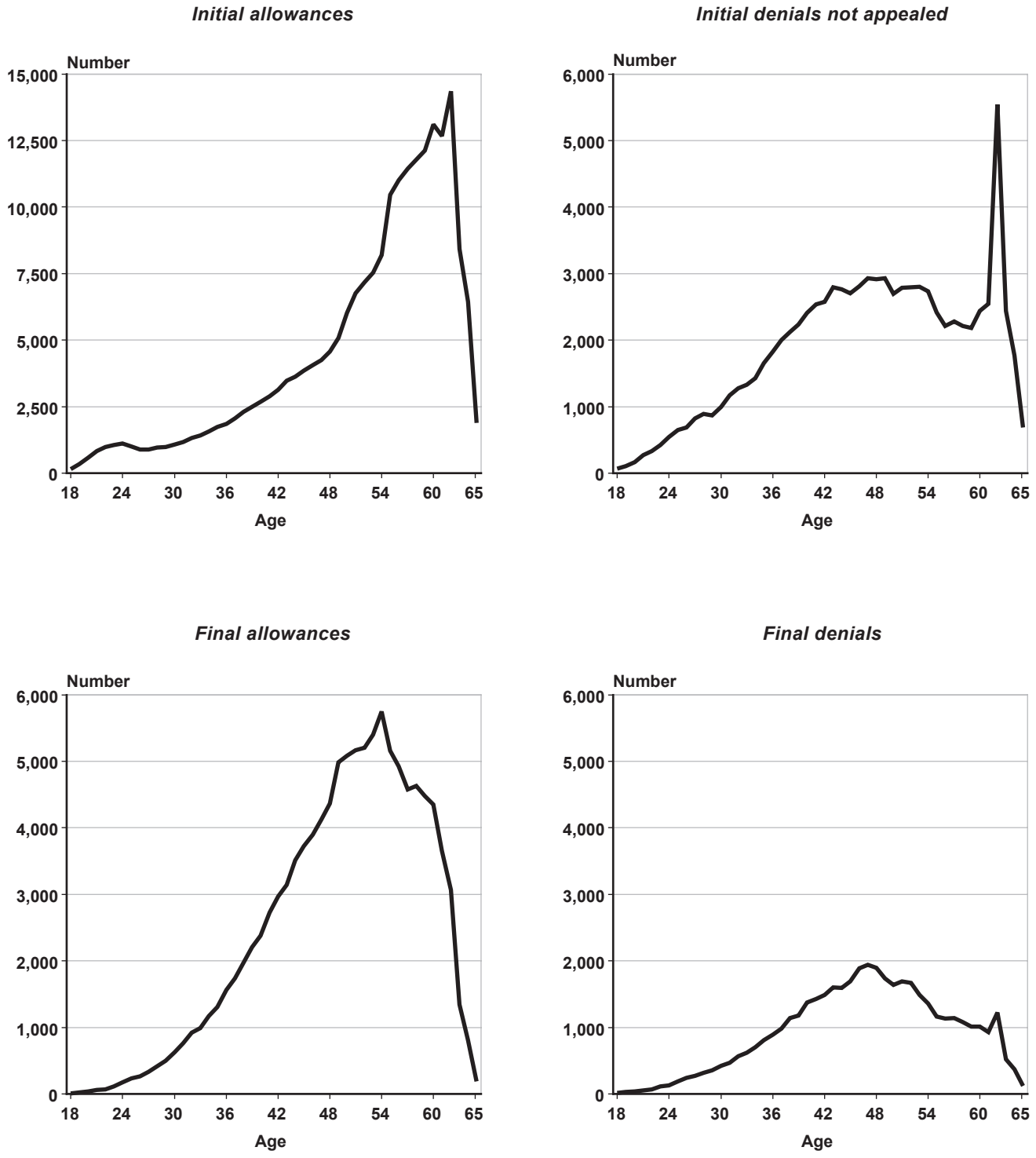
One interesting aspect of the age profiles is their nonlinearity. Specifically, the convex shape in the proportion of initial allowances might appear at odds with the notion that age is a reasonable proxy for health. Beyond some threshold age range, it is reasonable to expect the initial allowance rate to rise. After all, the increasing prevalence of serious age-related disabilities and less stringent vocational standards of the program are bound to push allowance rates upward. But what explains the high initial allowance rates for claimants at a very young age? One plausible answer is that the high allowance rates are driven by the impairment severity of a tiny number of applicants from an otherwise very healthy pool of workers. In addition, the contributory requirements of the DI program could be creating a bottleneck effect, with young disabled workers waiting to reach insured status. A look at the diagnostic makeup of claimants by age reveals some insights.

Chart 6 displays the distribution of claimants for the most common body systems by single year of age. About 60 percent of the small fraction of applicants aged 18–23 receive a mental diagnosis. Because mental impairments tend to have a very early onset, they indeed dominate the composition of claimants until about age 30. From age 31 forward, musculoskeletal impairments become the most common diagnosis. On the other hand, the share of mental impairments declines steadily with age. By ages 55 and 57, circulatory disorders and neoplasms surpass mental impairments to respectively become the second and third leading groups of diagnoses.

Inferential Results

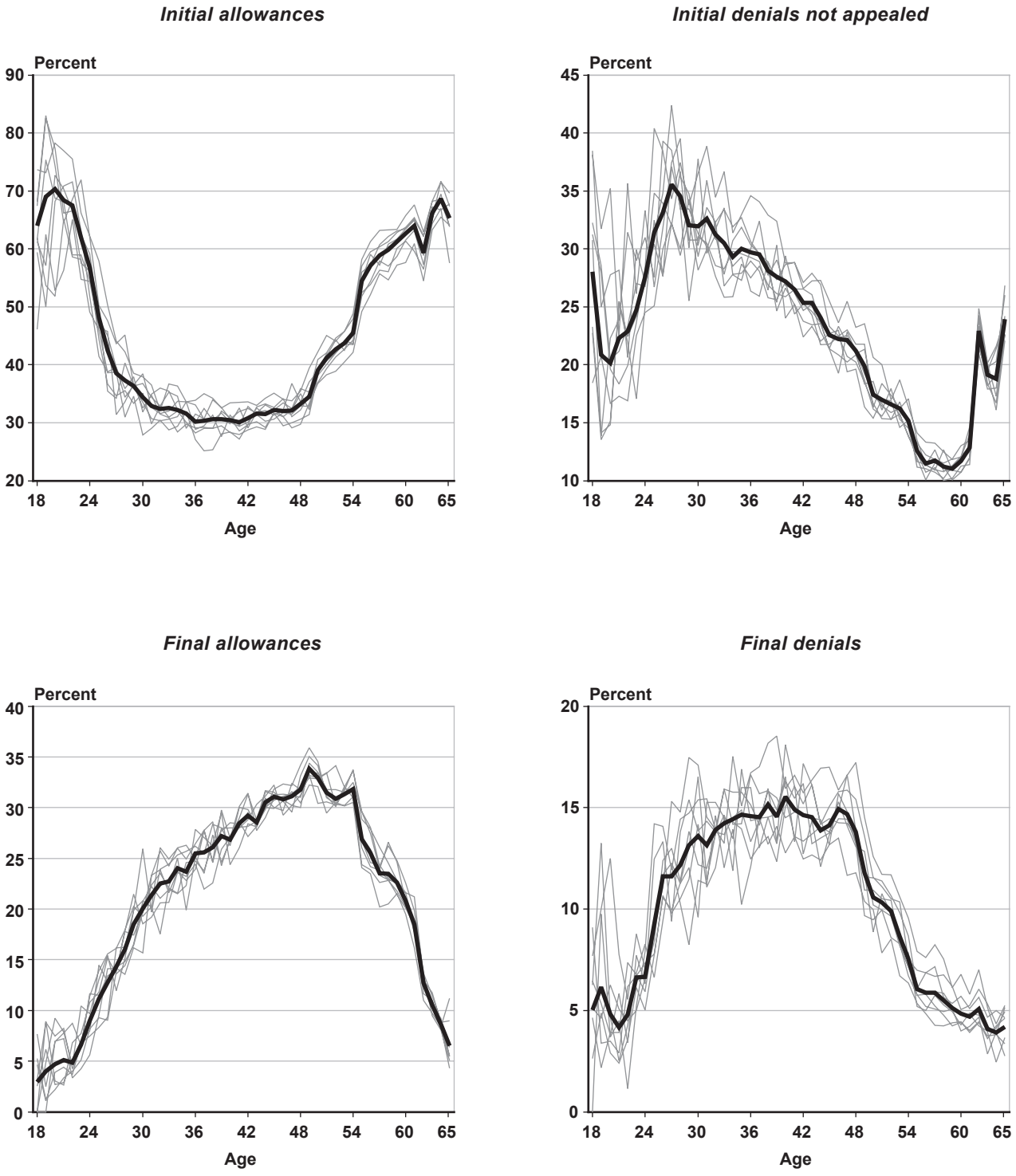
For each hierarchical structure (claimants nested by state or diagnosis), two model specifications are contemplated. Each model is estimated initially with no explanatory variables other than intercepts. The intercepts-only specification is useful to apportion unconditional data variance between hierarchical levels. It also provides a benchmark lower bound to goodness-of-fit criteria, which can be used for comparison purposes. The second specification entertains the previously described individual-level predictors. In addition, estimates are provided for a pooled or nonhierarchical model that does not entertain any grouping of the data.

Chart 4.
Number of claimants, by adjudicative disability category and single year of age



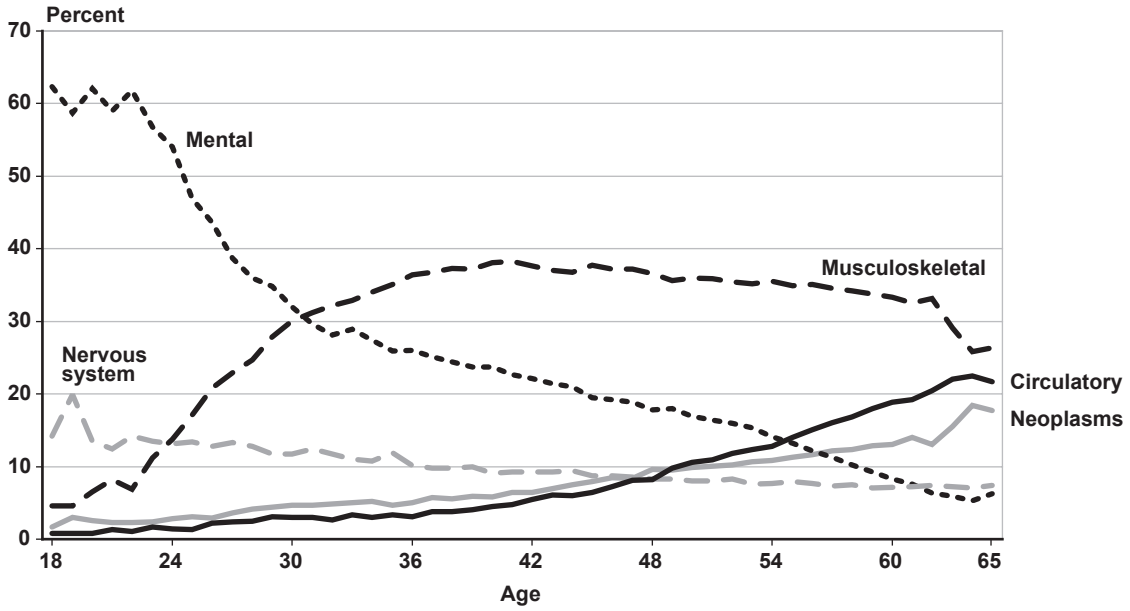
SOURCE: Author's calculations based on a 10 percent random sample of the DRF.

Chart 5.
Percentage of adjudicative disability categories, by single year of age



SOURCE: Author's calculations based on a 10 percent random sample of the DRF.

Chart 6.
Percentage of claimants, by selected body systems and single year of age



SOURCE: Author's calculations based on a 10 percent random sample of the DRF.

Next, I consider two different metrics for goodness-of-fit assessment. One measure that is particularly convenient in the context of Bayesian hierarchical models is the deviance information criterion (DIC), proposed by Spiegelhalter and others (2002). The DIC can be seen as the Bayesian analogous to the classical Akaike information criterion. It incorporates cross-validation and penalizes excess complexity. When comparing multiple specifications, the smaller the DIC value, the better the model's fit. DIC estimates are presented in the following tabulation. Additionally, I compute the percentage of observations correctly predicted by each model, shown in Table 5. In this case, an observed outcome is treated as a correct prediction if its estimated posterior mean probability is higher than the mean classification probabilities of the three other remaining outcomes.

Model specification	DIC estimate
Intercepts only	
Pooled	1,151,155.30
State	1,140,108.40
Diagnosis	1,038,875.60
Individual-level inputs	
Pooled	1,093,989.10
State	1,080,995.40
Diagnosis	980,212.70

Both measures of model fit provide a consistent picture. First, for a given set of variables, there is an unequivocal advantage in grouping claimants by state rather than pooling them together and in grouping them by impairment rather than clustering them by state. Consider for instance the top entry in Table 5, which corresponds to the intercepts-only pooled multinomial logit specification. As there are no explanatory variables, the estimated probability of any observation within a category is simply the sample proportion. All claimants are predicted to receive an initial allowance because this is the outcome that occurs most often. As a result, all of the initial allowances, but none of the other outcomes, are correctly categorized. This provides a lower predictive bound of 46.23 percent of the decisions correctly classified.

One way to think of a model with only intercepts is as a naive classification rule. In a hierarchical context, all individual outcomes within say a state or a diagnosis are predicted to be equal to the disability category with the highest sample proportion for that state or diagnosis. In grouping claimants by state, the intercepts-only model variant achieves some very modest gains relative to the pooled specification (46.26 percent). On the other hand, prediction improves more significantly if claimants are clustered by diagnosis (51.45 percent). When claimant-level explanatory

Table 5.
Percentage of observations correctly predicted, by model and adjudicative disability category

Model	Initial		Final		Total (correctly categorized)
	Allowances	Denials not appealed	Allowances	Denials	
			Intercepts only		
Pooled	100.00	0	0	0	46.23
State	97.18	0	5.38	0	46.26
Diagnosis	83.68	18.06	37.18	0	51.45
			Individual-level inputs		
Pooled	90.80	6.35	17.07	0	47.46
State	85.89	9.43	27.95	0.03	48.50
Diagnosis	87.84	24.80	39.50	0.20	55.27

SOURCE: Author's calculations based on a 10 percent random sample of the DRF.

variables are accommodated, the hierarchical diagnosis model can accurately classify 55.27 percent of the observations. The DIC estimates result in a similar ranking of the models.

A second conclusion can be drawn from Table 5. Notice how the diagnosis model with only intercepts correctly predicts a larger share of observations (51.45 percent) than the state model with claimant-level explanatory variables (48.50 percent). The same conclusion is reached when comparing the DIC estimates in the tabulation on the previous page (1,038,875 versus 1,080,995). This suggests that the primary diagnosis codes carry greater predictive ability than all other explanatory variables that are entertained combined. To put it differently, grouping a sample of claimants by diagnosis alone (the naive classification rule implied by an intercept-only model) will predict the adjudicative disability decision outcomes more accurately than knowing everything else, including age, sex, state of origin, application history, earnings history, and employment status in the year before filing. This finding is hardly unexpected, considering the role medical impairments play in the disability determination process. However, the result suggests that the full range of primary diagnosis codes (which are often overlooked for the purpose of research) is a crucial piece of information among the limited set of useful variables typically available from administrative data extracts.

Average Effects

The top portion of Table 6 presents posterior means and standard deviations of the regression coefficients in the pooled multinomial logit model.[14] The bottom part of the table displays estimates corresponding to the so-called average effects of the hierarchical diagnosis model. These parameters represent the mean of the distribution of the diagnosis-specific coefficients (that is, the estimated means of the multivariate normal heterogeneity distribution). For both models (pooled and hierarchical), the estimates tend to have similar signs and magnitudes, although as expected, the standard deviations are much higher in the hierarchical version of the process.

Given a particular observation and model, three equations yield continuous linear predictions of an initial allowance, an initial denial not appealed, and a final allowance. Those linear predictions are defined in reference to the benchmark category of final denials, which has a zero linear prediction by design. All else equal and relative to an initial denial, the sign of the estimated coefficients implies the following effects at the claimant level:

- The linear prediction of an initial allowance: (1) increases for males and higher earners, and (2) decreases for unemployed applicants and claimants who have applied before.
- The linear prediction of an initial denial not appealed: (1) increases for males; and (2) decreases for higher earners, unemployed applicants, and claimants who have applied before.
- The linear prediction of a final allowance: (1) increases for higher earners and claimants who have applied before, and (2) decreases for males and unemployed applicants.

At the individual level, the estimated effects for the explanatory variables match my a priori expectations. The results also appear consistent with research

Table 6.
Posterior parameter means and standard deviations, by adjudicative disability category

Variable	Initial allowances		Initial denials not appealed		Final allowances	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
			Pooled multinomial logit			
Intercept	1.252581	0.007014	0.568303	0.007492	1.060417	0.007195
Reapplicant	-0.427993	0.013526	-0.207347	0.014696	0.143670	0.013884
Male	0.121928	0.011101	0.026245	0.012352	-0.086577	0.011496
Earnings	0.236173	0.005119	-0.013485	0.005773	0.215522	0.005406
Unemployed	-0.614464	0.013845	-0.159045	0.013977	-0.292570	0.013669
Age	0.085465	0.000775	0.031423	0.000837	0.031753	0.000817
Age^2	0.004253	0.000051	0.002366	0.000055	-0.000036	0.000058
			Hierarchical diagnosis multinomial logit (average effects)			
Intercept	1.682253	0.131225	0.698521	0.046508	1.213439	0.057412
Reapplicant	-0.363732	0.060715	-0.198089	0.061695	0.202028	0.061481
Male	0.200885	0.054438	0.107004	0.052963	-0.069526	0.054504
Earnings	0.242488	0.040234	-0.029794	0.039516	0.220765	0.041297
Unemployed	-0.655286	0.064591	-0.213411	0.062811	-0.332018	0.064286
Age	0.081470	0.030319	0.027600	0.030594	0.016404	0.029204
Age^2	0.011028	0.027359	0.001629	0.027164	0.002343	0.027872

SOURCE: Author's calculations based on a 10 percent random sample of the DRF.

by Rupp (2012), who also used claimant-level data. Specifically, Rupp's "fixed-effects" binary logit model for initial determinations yielded qualitatively similar conclusions about the impact of sex and unemployment on the initial allowance rate. Of course, there are substantial differences in the two modeling approaches. Rupp (2012) used the time-varying state unemployment rates, while I do not control for year-effects and instead define unemployment at the individual level (as having zero earnings in the year prior to application). All else equal, the higher the earnings category, the higher the opportunity cost of filing for DI benefits, which may explain the positive association I find between earnings and the predictions of both an initial and a final allowance. Meanwhile, a history of previous applications shows a negative impact on the likelihood of an initial allowance, but a positive impact on the likelihood of a final allowance. In addition, I find that reapplicants are more likely to appeal an initial denial.

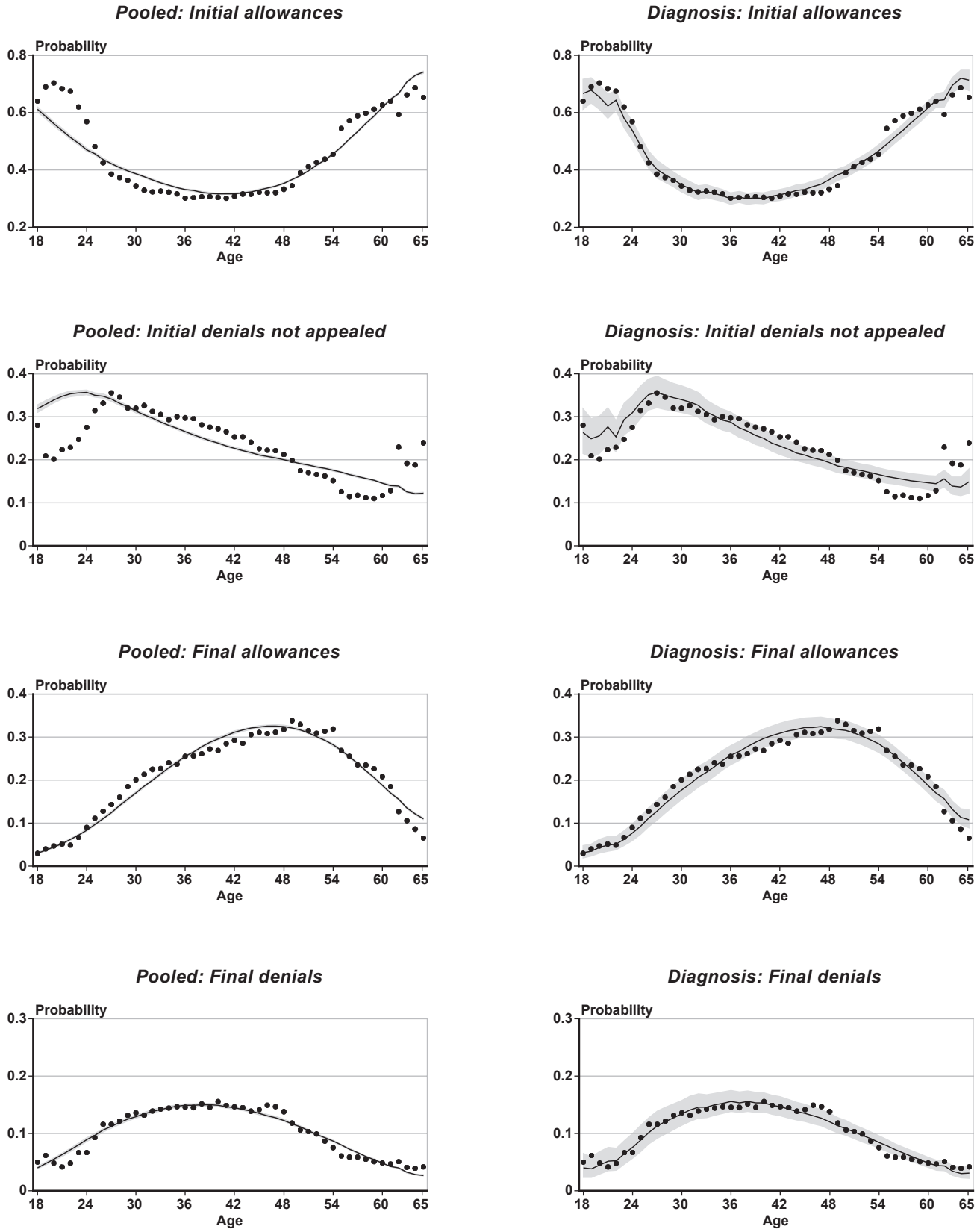
The interpretation of the parameters associated with age is less tractable because of the fact that those parameters represent the coefficients of a quadratic polynomial. Aggregate point and interval probability predictions for each outcome by single year of age are presented in Chart 7. Those predictions are obtained by averaging over the estimated probabilities of all the claimants in the sample who are the same age. The shaded areas in the graphs represent 90 percent posterior credible intervals (in other words, intervals containing 90 percent posterior probability). The thin dark lines along the intervals correspond to the posterior mean of each prediction. In addition, the solid dots show the actual proportions observed in the sample.

In general, it appears that the square term for age does a reasonably good job at capturing the nonlinear shape of the age profiles. The left and right columns of graphs in Chart 7 correspond to the pooled and hierarchical diagnosis models, respectively. The interval estimates for the pooled specification seem inadequately narrow, seriously underrepresenting uncertainty, as they miss most of the actual proportions. The point and interval predictions for the hierarchical diagnosis process clearly provide an improvement in fit. This is particularly evident in both the greater width of the intervals and at the youngest ages, where the shape of the age profiles is defined by relatively small numbers of claimants with a predominance of mental impairments.

Variance Decomposition

One issue of particular interest in this analysis is variance decomposition; that is, the portion of total variation in outcomes that the models attribute to

Chart 7.
Aggregate point and interval probability predictions for each adjudicative disability category, by single year of age



SOURCE: Author's calculations based on a 10 percent random sample of the DRF and model estimates.

the groups rather than the claimants. The top panel of Table 7 presents posterior means and standard deviations of between-group variances for the specifications with intercepts only. Consider for instance the first entry in the table, which corresponds to an initial allowance in the state hierarchical specification. The model has 50 intercept parameters per equation, each representing a state's mean linear prediction of an initial allowance. The posterior mean of the variance among those predictions is 0.22. Likewise, the between-state variance estimate for the linear prediction of an initial denial not appealed is 0.16.

In a similar fashion, the middle panel of Table 7 shows between-group variances corresponding to the models with claimant-level explanatory variables. Now the intercepts represent mean linear predictions of the outcomes when the explanatory variables take their average values in the sample.[15] Thus, the adjusted mean linear prediction of an initial allowance has a between-state variance of 0.59. Likewise, the variance of the mean-adjusted predictions for an initial denial not appealed between the states is 0.51.

One pattern emerges from the estimates in Table 7. For a given specification, the between-state variances corresponding to the prediction of all three outcomes are small and close in magnitude to one another. On the other hand, things are quite different when claimants are grouped by their impairments. In particular, variation in the prediction of an initial allowance between the diagnoses is very large (2.6 for the model with only intercepts and 2.8 for the variant with individual explanatory variables). Those magnitudes dwarf the variances associated with the other adjudicative categories (initial denials not appealed and final allowances). The implication is one of considerable heterogeneity in the prediction of an initial allowance among the impairments. This is of course consistent with the description of the data, where some primary diagnosis codes have initial allowance rates of over 95 percent, while others are close to zero.

In hierarchical models, total data variance is the sum of the within-group and the between-group variances. A useful statistic of variance decomposition is the intraclass correlation coefficient (ICC), which measures the proportion of variance in the outcomes between the groups. A value close to zero indicates a good deal of homogeneity between the clusters, so that most of the data variance can be attributed to individual-level variation within the groups. Conversely, an ICC close to 100 percent suggests a high degree of between-group heterogeneity, which implicitly favors a hierarchical modeling structure.

The bottom panel of Table 7 displays estimated ICC values.[16] On average, only about 6.2 percent of total variance in initial allowances can be attributed to differences between the states. Most of the observed heterogeneity in initial allowances (over 90 percent) seems to be due to disparities among claimants within the states. The decomposition suggests that applicants within any given state can be very heterogeneous in

Table 7.
Posterior estimates of group-level variances and ICCs, by adjudicative disability category

Disability outcome	State		Diagnosis	
	Mean	Standard deviation	Mean	Standard deviation
	Between-group variances: Intercepts only			
Initial allowances	0.219	0.044	2.587	0.286
Initial denials not appealed	0.160	0.032	0.122	0.016
Final allowances	0.180	0.036	0.269	0.035
	Between-group variances: Individual-level inputs			
Initial allowances	0.594	0.120	2.824	0.338
Initial denials not appealed	0.514	0.104	0.274	0.034
Final allowances	0.543	0.108	0.409	0.052
	ICCs (percent)			
Initial allowances	6.22	1.17	43.89	2.70
Initial denials not appealed	4.65	0.90	3.59	0.46
Final allowances	5.17	0.98	7.56	0.91

SOURCE: Author's calculations based on a 10 percent random sample of the DRF.

NOTE: ICC = intraclass correlation coefficient.

their disability characteristics. In fact, once claimants are grouped by primary diagnosis, a large portion of variation previously attributed to the individuals can now be explained by the differences between the impairments. About 44 percent of total variation in initial allowances is attributed to the different diagnosis groups. These results do not extend to the other outcomes (initial denials not appealed and final allowances), where group-level heterogeneity does not exceed 10 percent of total variance.

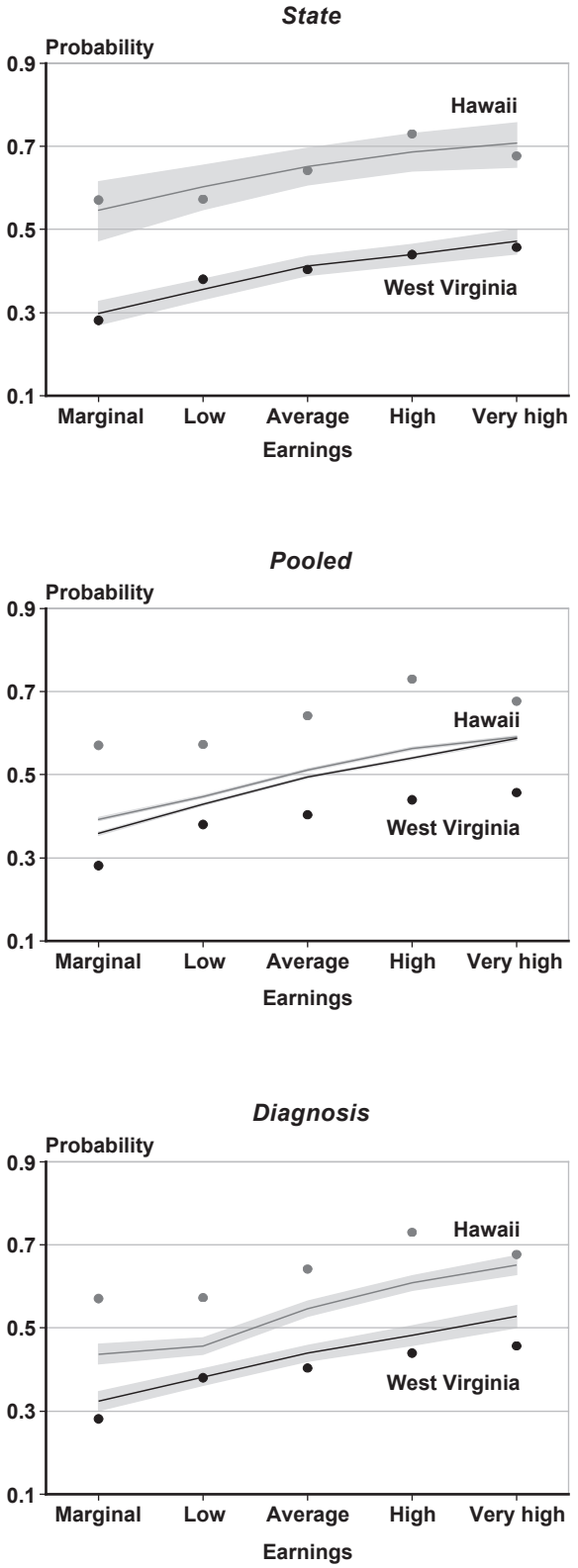
One of the implications of the ICC estimates is that the primary diagnoses can account for a great deal of the observed variation in initial allowances among claimants. To the extent that it is possible, parallels are drawn between the findings in this article and those in Rupp (2012). Fixed-effects models are not designed to apportion variance into between-group and within-group sources. Rupp (2012, Table 9) looked at the decomposition of overall variation in initial allowance rates across states by three sources. For adult DI-only claimants, the state fixed-effects accounted for 52 percent of the variation, while the year fixed-effects and the demographic and diagnostic characteristics of claimants contributed 14 percent and 10 percent of variation, respectively. The large size of the state fixed-effects in Rupp's article suggested that long-term unique differences among the states were substantial. That might seem at odds with this article's finding of small between-state, but large within-state variation in the outcomes. Notice, however, that the hierarchical state model here tracks with a great deal of accuracy the four adjudicative outcomes for each one of the states. This is by design because the model accommodates state-specific parameters. In other words, the hierarchical state model does a much better job at predicting the observed allowance and denial rates by state than does the hierarchical diagnosis model. Nevertheless, as the DIC tabulation and Table 5 confirm, the hierarchical diagnosis model unquestionably fits the overall data much better. First, it yields a significantly smaller DIC estimate. Second, for all claimants, it correctly predicts a higher share of each of the four adjudicative outcomes than does the state model.

The results in Rupp (2012) hinted at the diagnostic mix playing a role (although a small one), in explaining state heterogeneity in initial allowance rates.[17] The findings here (values not shown) are consistent with that view, in that the diagnostic mix is not a major factor at accurately predicting initial allowance rates in most states, except in some cases, despite the fact that state variation in the composition of impairments is substantial in the sample under study. For instance, the proportion of musculoskeletal diagnoses ranges from 27 percent in Hawaii to 42.9 percent in Montana. Mental disorders comprise 26.9 percent of the diagnoses in New Hampshire, but only 12.1 percent of those in Arkansas. Neoplasms vary from 13.6 percent in Iowa to 6.3 percent in West Virginia. Mississippi has the highest composition of circulatory diagnoses at 15.8 percent, while Idaho has the lowest at 7.1 percent. For the nervous system and sense organs group, Colorado has a proportion of diagnoses (12.3 percent) that is three times the size of that corresponding to Vermont. Injuries also vary from 2.5 percent in South Dakota to 10.6 percent in West Virginia. Coe and others (2011) cited substantial variation in age-adjusted mortality rates by state and even greater variation in self-reported disability.

In the context of my modeling effort, one way to further illustrate state heterogeneity in disability outcomes is through a specific example. Chart 8 provides a comparison between the states of Hawaii and West Virginia. The graphs display point and interval probability predictions (90 percent posterior probability) of an initial allowance as a function of earnings for both states. Hawaii exhibits the highest initial allowance rate in the sample at 62.5 percent. In addition, it also happens to have the lowest proportion of musculoskeletal impairments of any state. By contrast, West Virginia has the third-lowest initial allowance rate (37.4 percent) and incidentally, the lowest proportion of neoplasms and the highest share of injuries among the 50 states.

The top graph in Chart 8 corresponds to the hierarchical state model, which by design, accurately reproduces the observed state proportions. Notice that Hawaii has a smaller number of observations than West Virginia (Appendix Table A-2), resulting in state-specific parameter estimates with greater variance (and as a result, a wider probability interval). The middle graph in Chart 8 presents the predictions associated with the pooled model. In this case, there is a wide gap between observed and predicted outcomes. Over all claimants, Hawaii and West Virginia differ in their proportion of initial allowances by about 25 percentage points (see Chart 1). Instead, the pooled model predicts a mean gap of about 3 percentage points, despite the fact that the predictions take into account the different mix of characteristics between the applicant populations in the two states (age, sex, employment status, application history, and earnings history).

Chart 8.
Aggregate point and interval probability predictions for an initial allowance, by earnings: Hawaii compared with West Virginia



SOURCE: Author's calculations based on a 10 percent random sample of the DRF and model estimates.

The graph at the bottom of Chart 8 shows the probability predictions resulting from the hierarchical diagnosis model. This specification incorporates the same individual-level predictors as the pooled multinomial logit model. The only difference, of course, is that claimants are grouped according to their impairments. Relative to the observed proportions, the diagnosis model slightly overpredicts the probabilities corresponding to West Virginia, but significantly underpredicts the probabilities associated with Hawaii. On average, the predicted gap in the probability of an initial allowance between the two states is 11 percentage points. In other words, discrepancies in claimant-level characteristics (differences in the impairment mix specifically) seem to account for a little less than half of the observed difference in the initial allowance rate between these two states. This result, however, does not generalize to comparisons among other states.

Correlation Across Outcomes

Table 8 presents posterior estimates of the correlation between the disability adjudicative outcomes. The top panel of the table corresponds to the intercepts-only specification, while the bottom panel comprises the estimates for the models with claimant-level predictors. For example, the mean correlation between the average linear predictions of an initial allowance and an initial denial not appealed among the 50 states is 0.25. Likewise, the mean correlation between those two outcomes among the 181 primary diagnosis codes is 0.31. When the individual explanatory variables are included in the models, the corresponding correlation for the adjusted linear prediction of an initial allowance and an initial denial not appealed is 0.1 among the states and 0.13 among the impairments.

A look at Table 8 reveals that after controlling for individual-level predictors, the correlations in the state hierarchical models are small in magnitude and statistically insignificant. However, when claimants are grouped by diagnosis, there is very high statistically significant positive correlation between the linear predictions of an initial and a final allowance. For instance, with only intercepts, the posterior mean correlation among the impairments is 0.74. After controlling for claimant-level explanatory variables, a mean estimate of 0.56 is obtained. To the best of my knowledge, the finding of high significant positive correlation when impairments are used as a criterion for grouping claimants has never been reported in the literature. The finding is important for several reasons. First, it indicates that the zero correlation property implicit in

a pure multinomial logit model (the so-called independence from irrelevant alternatives property) is an unrealistic restriction. More generally, any effort to model the adjudicative process using the impairments should accommodate this pattern in the data.

My classification of claimants roughly corresponds to a two-stage adjudication (decisions at the DDS level versus decisions made mostly at the hearing level or above). In this context, the estimation results suggest a substantial degree of dependence between the two adjudicative outcomes. Across the impairments, the high positive correlation between the predictions of an initial and a final allowance is important for a second reason. Normatively speaking, the more disabling a diagnosis, the greater the linear predictions of both an initial and a final allowance should be, relative to less disabling impairments. In this very narrow sense, the correlation result here appears to suggest a degree of consistency within the adjudicative process.

Consider the top graph on the left in Chart 9, which plots posterior means of the intercepts for the 181 primary diagnosis codes corresponding to the model with claimant-level predictors. Those coefficients represent adjusted mean linear predictions of an initial denial not appealed and a final allowance. There is no apparent relationship between the two outcomes, as a statistically insignificant mean correlation estimate of 0.12 bears out in Table 8. Transforming the linear predictions into actual probabilities results in the top graph on the right. Unlike the linear predictions, the probabilities show an upward trend. Impairments that have a higher classification probability of an initial denial not appealed also tend to have a higher probability of a final allowance.

The bottom-left graph in Chart 9 plots the relationship between the linear predictions of an initial and a final allowance for each of the impairments. In this case, the mean correlation is 0.56 (shown in Table 8). However, the corresponding probabilities in the bottom-right graph indicate the opposite effect (negative correlation). In other words, diagnoses that have a higher classification probability of an initial allowance tend to have a lower classification probability of a final allowance. The reason for the correlation inversion has to do with the fact that the probability of an outcome is a nonlinear function of the linear prediction of all the possible outcomes. As the linear prediction of an initial allowance dominates the magnitude of the other predictions, the classification probabilities of an initial denial not appealed, a final allowance, and a final denial can only decline.

The implications of high positive correlation between the linear predictions of an initial and a final allowance (bottom-left graph in Chart 9) can be further clarified with a somewhat extreme example involving the two impairments that are presented in Chart 10. The most common diagnosis in the musculoskeletal body system is a disorder of the back (discogenic and degenerative). The proportions in the entire sample of initial and final allowances for that impairment are about 23 percent and 38 percent, respectively. On the other hand, based on its effect on mortality alone, a highly disabling diagnosis is lung cancer (malignant neoplasm of the trachea, bronchus, or lung). In this case, 94 percent of the decisions result in an initial allowance, while only 3 percent of the outcomes represent a final allowance.

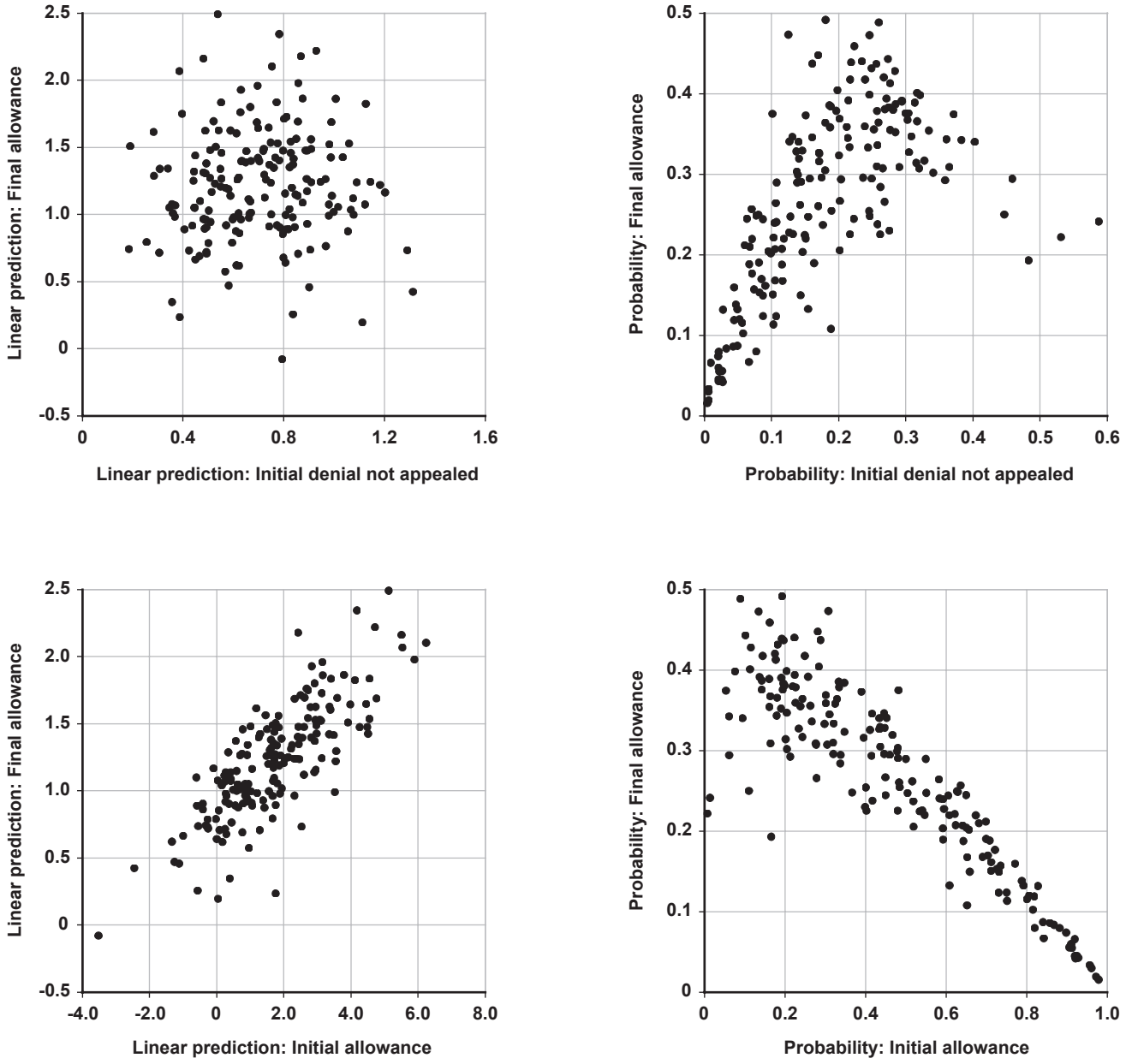
Table 8.
Posterior correlations, by model specification

Correlation sequence of disability outcome	State		Diagnosis	
	Mean	Standard deviation	Mean	Standard deviation
		Intercepts only		
Initial allowance—initial denial not appealed	0.249	0.129	0.307	0.087
Initial allowance—final allowance	0.063	0.136	0.737	0.041
Initial denial not appealed—final allowance	0.015	0.135	0.177	0.092
		Individual-level inputs		
Initial allowance—initial denial not appealed	0.100	0.133	0.125	0.096
Initial allowance—final allowance	0.048	0.135	0.561	0.064
Initial denial not appealed—final allowance	0.006	0.134	0.119	0.087

SOURCE: Author's calculations based on a 10 percent random sample of the DRF.

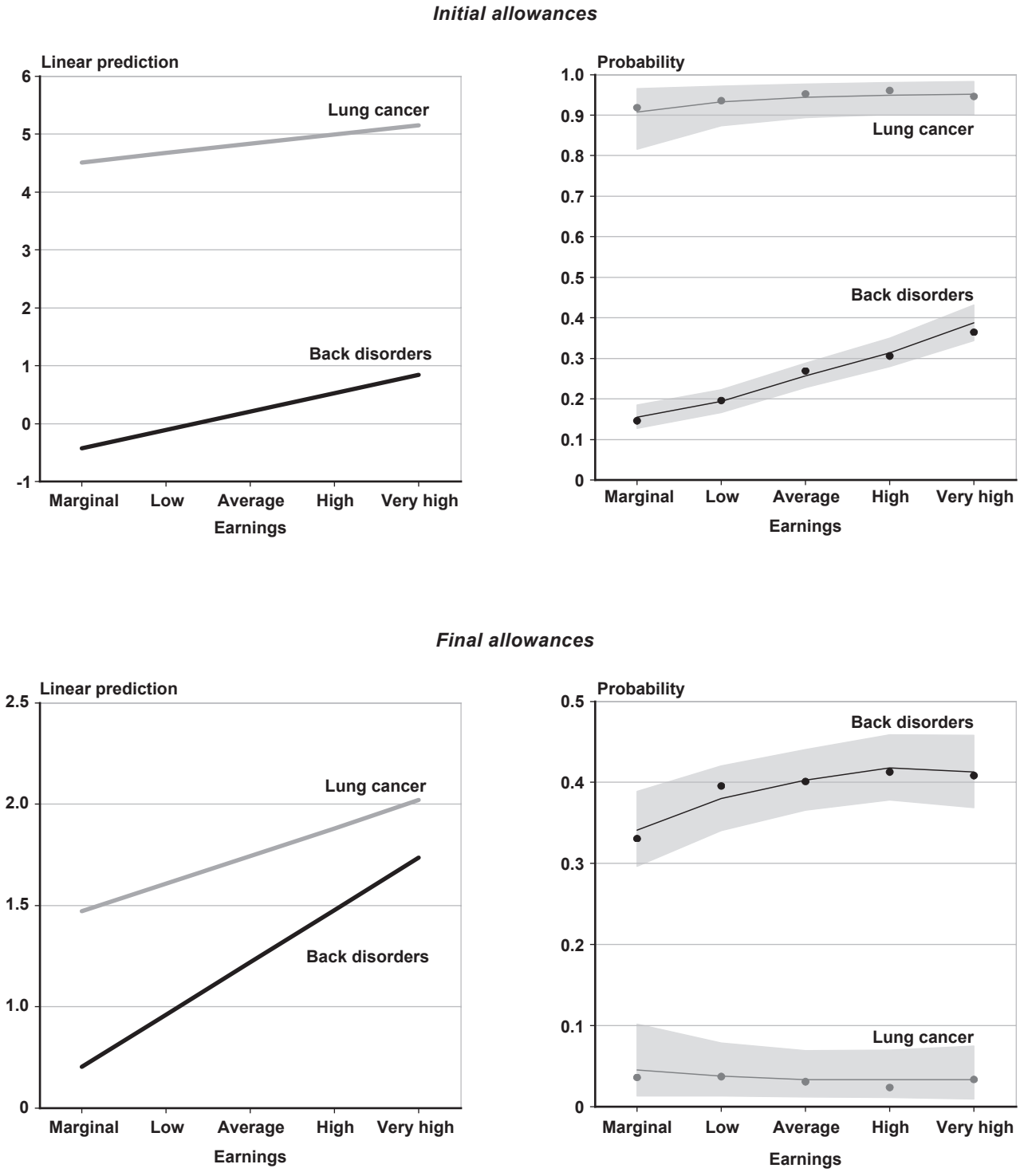
Suppose two claimants were identical in all measured characteristics (having the sample mean features), except one was diagnosed with lung cancer and the other had a back disorder. Linear predictions for those two claimants as a function of earnings appear on the left (top and bottom) graphs of Chart 10. Notice in particular how the predictions of an initial and a final allowance for the claimant with lung cancer exceed the predictions corresponding to the applicant with a back disorder. By contrast, the two graphs on the right side of the chart display point and interval probability predictions (90 percent posterior probability), which closely follow the observed sample proportions. For any outcome different from an initial allowance, the classification probabilities of lung cancer lie well below the probabilities of a disorder of the back. This, of course, is due to the extremely high probability of an initial allowance

Chart 9.
Linear predictions compared with probabilities in the diagnosis model



SOURCE: Author's calculations based on a 10 percent random sample of the DRF and model estimates.

Chart 10.
Lung cancer versus disorders of the back, by earnings: Linear predictions compared with probabilities



SOURCE: Author's calculations based on a 10 percent random sample of the DRF and model estimates.

associated with a diagnosis of lung cancer in the first place.

In the two-impairment (lung cancer/back disorder) example, a significant fraction of claimants with back disorders are initially denied, but eventually allowed. Yet, claimants with lung cancer have a higher prediction of both an initial and a final allowance. Put differently, it is simply not the case that ALJs are favoring applicants with back disorders over those with lung cancer. Whether it is at the DDS or at the hearing level or above, lung cancer is determined to be a more disabling diagnosis than a back disorder. In general, the high positive correlation implies that in going from an initial to a final allowance, decision makers are largely preserving the ordinal ranking of impairments (a finding that is only evident when looking at the linear predictions and not the probabilities).

One might be tempted to conclude that this correlation finding provides evidence that decision makers are uniformly adhering to SSA's disability guidelines at the various adjudicative levels. However, other possible explanations cannot be ruled out. For example, Keiser (2010) hinted at evidence of a policy feedback mechanism, where knowledge of ALJ reversal rates affected the self-reported initial allowance rate of DDS examiners.[18] If there was a feedback effect, it could also flow in either direction (from the DDS to the Office of Disability Adjudication and Review (ODAR) and vice versa), or from both directions simultaneously. The bottom line is that it is important not to overreach when it comes to interpreting my results. The positive correlation between the predictions of an initial and a final allowance could be potentially explained by a feedback effect, where decision makers at the two stages are influenced by each other's ranking of impairments. Nevertheless, whether a feedback mechanism or adherence to the guidelines explains the positive correlation, the result implies some degree of consistency.

Conclusion

This article explores the roles that primary diagnoses and state of origin play in explaining observed heterogeneity in disability outcomes by adjudicative stage. Disability determinations are separated into four mutually exclusive categories: (1) initial allowances, (2) initial denials not appealed, (3) final allowances, and (4) final denials. The main findings are as follows:

- The primary diagnosis codes carry greater predictive ability for placing claimants into adjudicative categories than all other explanatory variables that are entertained combined. Knowing the impairments of a sample of applicants yields more accurate classification probabilities than knowing their age, sex, state of origin, earnings, employment status in the year before filing, and application history combined.
- The prediction of an initial allowance (1) increases for males and higher earners, and (2) decreases for unemployed applicants and claimants who have applied before.
- The prediction of an initial denial not appealed (1) increases for males; and (2) decreases for higher earners, unemployed applicants, and claimants who have applied before.
- The prediction of a final allowance (1) increases for higher earners and claimants who have applied before, and (2) decreases for males and unemployed applicants.
- As a function of single year of age, the initial allowance rate has a *u-shape* defined at very young ages by small numbers of claimants with a predominance of mental impairments. A quadratic polynomial seems to reproduce the age profiles accurately.
- When claimants are grouped by state, variance decomposition suggests that most of the variation in outcomes is driven by individual-level heterogeneity within the states. On the other hand, almost half of the variation in initial allowances can be attributed to the various primary diagnoses. In some cases, the different mix of impairments in the population of claimants may explain a significant portion of the difference in initial allowances between two states. Still, a great deal of variation in outcomes remains unaccounted for by the models, particularly when it comes to identifying final denials.
- When applicants are grouped by diagnosis, there is high positive correlation between the predictions of an initial and a final allowance. To the best of my knowledge, that finding has never been documented in the literature. Impairments that are considered to be more disabling at the DDS level tend to also be considered more disabling at the hearing level or above. In other words, when moving from an initial to a final allowance, the severity ranking of the diagnoses is preserved to a good extent.

Table A-1.
Sample distribution, by adjudicative disability category, body system, and primary diagnosis

Body system and primary diagnosis	Initial		Final		Total
	Allowances	Denials not appealed	Allowances	Denials	
Infectious/parasitic diseases	2,478	706	739	313	4,236
Pulmonary tuberculosis	(X)	13	(X)	(X)	27
Symptomatic HIV	1,559	298	283	98	2,238
Asymptomatic HIV	30	186	130	80	426
Neurosyphilis	19	(X)	(X)	(X)	36
Mycobacterial, other chronic infections	32	18	(X)	(X)	71
Other infectious and parasitic disorders	83	45	34	14	176
Late effects of acute poliomyelitis	568	50	115	31	764
Neoplasms	37,526	3,968	3,533	1,070	46,097
Malignant neoplasm of tongue	254	21	24	9	308
Malignant neoplasm of salivary glands	(X)	(X)	(X)	(X)	21
Malignant neoplasm of esophagus	1,123	(X)	36	(X)	1,179
Malignant neoplasm of stomach	641	(X)	24	(X)	687
Malignant neoplasm of small intestine	144	(X)	13	(X)	176
Malignant neoplasm of colon or rectum	3,528	514	435	126	4,603
Malignant neoplasm of liver	1,667	10	37	(X)	1,718
Malignant neoplasm of gallbladder	139	(X)	(X)	(X)	148
Malignant neoplasm of pancreas	1,357	(X)	24	(X)	1,394
Malignant neoplasm of digestive system	176	(X)	(X)	(X)	196
Malignant neoplasm of trachea or lung	8,249	161	281	50	8,741
Malignant neoplasm of pleura	332	(X)	(X)	(X)	347
Malignant neoplasm of heart	(X)	(X)	(X)	(X)	30
Malignant neoplasm of bone and cartilage	445	(X)	41	(X)	525
Malignant neoplasm of connective tissue	198	30	(X)	(X)	256
Malignant melanoma of skin	801	(X)	26	(X)	857
Other malignant neoplasm of skin	50	15	(X)	(X)	79
Malignant neoplasm of breast	4,717	1,370	731	345	7,163
Kaposi's sarcoma	(X)	(X)	(X)	(X)	(X)
Malignant neoplasm of bladder	451	65	57	12	585
Malignant neoplasm of kidney	977	64	63	23	1,127
Malignant neoplasm of eye	(X)	(X)	(X)	(X)	11
Malignant neoplasm of brain	2,507	55	111	21	2,694
Malignant neoplasm of nervous system	(X)	(X)	(X)	(X)	13
Malignant neoplasm of thyroid gland	87	38	33	13	171
Malignant neoplasm of endocrine glands	34	(X)	(X)	(X)	44
Malignant neoplasm of other sites (head, neck)	1,383	165	222	49	1,819
Secondary malignant neoplasms	232	(X)	(X)	(X)	244
Malignant neoplasm of unspecified site	47	(X)	(X)	(X)	58
Lymphoma	1,769	494	431	137	2,831
Multiple myeloma	900	45	123	12	1,080
Leukemias	1,626	89	128	25	1,868
Benign neoplasm of brain	430	159	208	73	870
Neoplasm of uncertain behavior	(X)	(X)	(X)	(X)	15
Neoplasm of unspecified/unknown nature	(X)	(X)	(X)	(X)	(X)
Malignant neoplasm of genital organs	3,191	520	395	123	4,229

Continued

Table A-1.
Sample distribution, by adjudicative disability category, body system, and primary diagnosis—*Continued*

Body system and primary diagnosis	Initial		Final		Total
	Allowances	Denials not appealed	Allowances	Denials	
Endocrine, nutritional, and metabolic	6,635	4,517	4,842	1,947	17,941
All disorders of thyroid	42	146	129	67	384
Diabetes mellitus	3,014	3,490	3,345	1,387	11,236
All disorders of parathyroid gland	(X)	(X)	(X)	(X)	(X)
All disorders of pituitary gland	(X)	(X)	15	(X)	28
All disorders of adrenal glands	(X)	(X)	(X)	(X)	22
Malnutrition (weight loss)	113	(X)	32	(X)	164
Disorders of plasma protein metabolism	(X)	(X)	(X)	(X)	(X)
Gout	65	75	75	37	252
Disorders of metabolism (cystic fibrosis)	85	(X)	11	(X)	112
Obesity and other hyperalimentation	3,229	725	1,140	416	5,510
Disorders of the immune mechanism	77	38	89	18	222
Diseases of the blood	623	173	175	79	1,050
Deficiency anemias	48	23	26	14	111
Hereditary hemolytic anemias	143	35	27	12	217
Aplastic anemia	152	(X)	21	(X)	184
Other anemias	148	53	39	15	255
Coagulation defects	(X)	(X)	(X)	(X)	28
Purpura and other hemorrhagic conditions	14	14	(X)	(X)	47
Other diseases of blood-forming organs	109	36	40	23	208
Mental disorders	41,770	13,117	17,007	5,641	77,535
Organic mental disorders	8,024	740	1,878	308	10,950
Schizophrenic, paranoid, psychotic disorders	3,963	650	665	186	5,464
Affective/mood disorders	19,678	8,466	11,290	3,768	43,202
Autistic disorders	75	(X)	(X)	(X)	89
Anxiety disorders	3,817	1,477	2,096	736	8,126
Personality disorders	457	280	177	119	1,033
Substance addiction (alcohol)	(X)	439	(X)	186	777
Substance addiction (drugs)	(X)	218	(X)	64	342
Somatoform disorders	216	61	162	32	471
Eating and tic disorders	(X)	(X)	(X)	(X)	(X)
Attention deficit disorder	44	32	11	19	106
Learning disorder	54	103	17	21	195
Mental retardation	5,079	335	311	81	5,806
Borderline intellectual functioning	354	310	184	119	967
Nervous system and sense organs	20,239	6,891	8,773	3,313	39,216
Cerebral degenerations	36	(X)	(X)	(X)	48
Brain atrophy	713	97	163	37	1,010
Parkinson's disease	1,315	137	300	50	1,802
Anterior horn cell disease	690	(X)	(X)	(X)	740
Other diseases of spinal cord	763	43	110	18	934
Disorders of autonomous nervous system	155	67	107	33	362
Multiple sclerosis	3,543	588	1,612	311	6,054
Cerebral palsy	549	68	70	29	716
Epilepsy	727	1,290	901	592	3,510
Migraine	349	446	488	217	1,500
Other neurological conditions	1,365	888	1,135	485	3,873
Carpal tunnel syndrome	117	174	213	103	607
Diabetic and other peripheral neuropathy	2,478	554	1,186	292	4,510

Continued

Table A-1.
Sample distribution, by adjudicative disability category, body system, and primary diagnosis—*Continued*

Body system and primary diagnosis	Initial		Final		Total
	Allowances	Denials not appealed	Allowances	Denials	
Nervous system and sense organs (cont.)					
Myoneural disorders	430	260	376	159	1,225
Muscular dystrophies	532	65	162	44	803
Retinal detachments and defects	207	103	78	42	430
Other retina disorders	644	151	212	51	1,058
Glaucoma	200	126	94	65	485
Cataract	61	99	44	26	230
Visual disturbances	437	400	326	148	1,311
Blindness and low vision	2,838	554	552	226	4,170
Cardiac transplantation	62	(X)	(X)	(X)	75
Disorders of eye movements	(X)	(X)	(X)	(X)	13
Disorders of vestibular system	284	194	299	122	899
Other disorders of ear	38	127	58	54	277
Deafness	1,704	439	229	202	2,574
Circulatory	28,256	9,336	12,593	3,852	54,037
Rheumatic fever with heart involvement	(X)	(X)	(X)	(X)	(X)
Diseases of aortic valve	297	192	221	84	794
Other rheumatic heart disease	70	(X)	25	(X)	124
Essential hypertension	412	1,706	1,192	696	4,006
Hypertensive vascular disease	538	467	453	172	1,630
Hypertensive vascular and renal disease	14	(X)	(X)	(X)	30
Acute myocardial infarction	566	435	385	119	1,505
Angina without ischemic heart disease	126	106	113	37	382
Chronic ischemic heart disease	7,977	3,132	5,107	1,395	17,611
Chronic pulmonary heart disease	378	42	68	14	502
Valvular heart disease/other defects	229	169	217	81	696
Cardiomyopathy	2,514	554	1,016	264	4,348
Cardiac dysrhythmias	412	282	363	141	1,198
Heart failure	2,972	480	737	153	4,342
Late effects of cerebrovascular disease	7,786	984	1,478	371	10,619
Aortic aneurysm	201	72	100	26	399
Peripheral vascular (arterial) disease	2,373	254	550	96	3,273
Periarteritis nodosa, allied conditions	50	(X)	(X)	(X)	71
Disease of capillaries	(X)	(X)	(X)	(X)	(X)
Phlebitis and thrombophlebitis	106	72	79	38	295
Varicose veins of lower extremities	292	70	80	29	471
Other diseases of circulatory system	942	280	381	123	1,726
Respiratory	11,539	2,671	3,528	1,313	19,051
Chronic bronchitis	41	53	60	29	183
Emphysema	890	150	195	52	1,287
Asthma	786	1,148	996	564	3,494
Bronchiectasis	33	15	(X)	(X)	66
Chronic pulmonary insufficiency	9,271	1,014	1,898	509	12,692
Asbestosis	43	(X)	39	(X)	106
Pneumoconiosis	(X)	(X)	10	(X)	20
Other diseases of the respiratory system	471	270	318	144	1,203

Continued

Table A-1.
Sample distribution, by adjudicative disability category, body system, and primary diagnosis—*Continued*

Body system and primary diagnosis	Initial		Final		Total
	Allowances	Denials not appealed	Allowances	Denials	
Digestive	3,918	2,322	2,772	1,049	10,061
Diseases of esophagus	17	22	20	13	72
Peptic ulcer (gastric or duodenal)	28	41	21	15	105
Gastritis and duodenitis	(X)	48	44	(X)	128
Hernias	72	160	176	73	481
Crohn's disease	297	266	423	137	1,123
Idiopathic proctocolitis	89	94	114	51	348
Other diseases of gastrointestinal system	397	694	729	291	2,111
Chronic liver disease, cirrhosis	2,968	970	1,224	439	5,601
Gastrointestinal hemorrhage	41	27	(X)	(X)	92
Genitourinary	6,043	500	686	176	7,405
Nephrotic syndrome	219	56	79	23	377
Chronic renal failure	5,731	144	376	36	6,287
Other diseases of the urinary tract	81	175	183	81	520
Disorders of the genital organs	12	125	48	36	221
Skin	255	308	285	122	970
Bullous disease	(X)	(X)	(X)	(X)	13
Ichthyosis	32	56	73	22	183
Dermatitis/psoriasis	80	99	77	26	282
Other disorders of the skin	138	149	133	72	492
Musculoskeletal	46,164	36,793	53,485	21,329	157,771
Diffuse diseases of connective tissue	1,075	483	919	295	2,772
Rheumatoid arthritis	4,138	1,093	1,904	504	7,639
Osteoarthrosis and allied disorders	14,398	6,341	8,852	3,208	32,799
Other and unspecified arthropathies	810	683	705	304	2,502
Ankylosing spondylitis	308	134	222	65	729
Disorders of back (discogenic and degenerative)	19,797	21,150	33,682	13,237	87,866
Disorders of muscle, ligament, and fascia	3,484	5,518	5,696	3,072	17,770
Osteomyelitis and other bone infection	258	86	99	30	473
Other disorders of bone and cartilage	1,761	1,165	1,265	518	4,709
Curvature of spine	135	140	141	96	512
Congenital	123	59	62	33	277
Spina bifida	44	(X)	(X)	(X)	60
Congenital anomalies of heart	60	40	31	19	150
Other congenital anomalies	19	(X)	25	(X)	67

Continued

Table A-1.
Sample distribution, by adjudicative disability category, body system, and primary diagnosis—*Continued*

Body system and primary diagnosis	Initial		Final		Total
	Allowances	Denials not appealed	Allowances	Denials	
Injuries	8,282	8,435	6,632	3,582	26,931
Multiple body dysfunctions	(X)	(X)	(X)	(X)	14
Sleep-related breathing disorders	85	107	150	74	416
Loss of voice	109	21	28	18	176
Fracture of vertebral column	912	108	117	29	1,166
Fracture of upper limb	597	1,069	692	398	2,756
Fracture of lower limb	2,178	2,309	1,836	835	7,158
Other fractures of bones	340	546	425	218	1,529
Dislocations (all types)	104	206	135	63	508
Sprains and strains (all types)	436	2,222	1,407	1,125	5,190
Intracranial injury	593	213	195	72	1,073
Internal injury	10	(X)	17	(X)	41
Open wound, except limbs	(X)	(X)	(X)	(X)	(X)
Open wound upper limb (soft tissue)	216	303	206	117	842
Open wound lower limb (soft tissue)	211	177	146	72	606
Amputations	1,292	770	718	350	3,130
Late effects of injuries to nervous system	1,039	225	301	119	1,684
Chronic fatigue syndrome	102	92	211	67	472
Burns (code 9480)	32	33	26	11	102
Burns (code 9490)	19	22	(X)	(X)	62

SOURCE: Author's calculations based on a 10 percent random sample of the DRF.

NOTE: (X) = suppressed to avoid disclosing information about particular individuals.

Table A-2.
Sample distribution, by adjudicative disability category and state

State	Initial		Final		Total
	Allowances	Denials not appealed	Allowances	Denials	
Alabama	3,858	1,583	3,620	625	9,686
Alaska	286	145	85	36	552
Arizona	4,707	1,492	1,725	588	8,512
Arkansas	2,589	913	1,853	533	5,888
California	23,358	10,492	8,279	4,456	46,585
Colorado	2,106	1,368	1,495	507	5,476
Connecticut	2,820	934	1,137	479	5,370
Delaware	828	239	254	115	1,436
Florida	11,372	5,180	8,082	2,839	27,473
Georgia	5,084	2,808	4,310	1,428	13,630
Hawaii	872	286	142	96	1,396
Idaho	918	377	421	174	1,890
Illinois	8,179	3,411	3,794	1,416	16,800
Indiana	4,822	2,555	3,112	1,420	11,909
Iowa	2,339	752	690	386	4,167
Kansas	1,817	814	804	371	3,806
Kentucky	3,552	1,355	3,291	1,119	9,317
Louisiana	2,934	1,388	1,956	709	6,987
Maine	1,313	366	678	183	2,540
Maryland	2,908	1,281	1,500	467	6,156
Massachusetts	5,163	1,280	1,955	646	9,044
Michigan	9,584	4,858	5,087	1,888	21,417
Minnesota	4,209	1,311	1,539	625	7,684
Mississippi	2,343	1,112	1,683	738	5,876
Missouri	5,336	1,846	2,499	846	10,527
Montana	537	291	353	177	1,358
Nebraska	1,261	514	382	221	2,378
Nevada	1,688	529	455	195	2,867
New Hampshire	1,377	320	406	106	2,209
New Jersey	6,863	1,964	2,549	856	12,232
New Mexico	1,285	530	584	238	2,637
New York	15,947	6,143	8,596	2,667	33,353
North Carolina	7,277	3,367	5,064	1,665	17,373
North Dakota	328	150	170	81	729
Ohio	8,028	3,871	4,658	2,150	18,707
Oklahoma	2,834	1,518	2,068	859	7,279
Oregon	2,939	1,278	1,226	644	6,087
Pennsylvania	11,635	4,056	4,866	2,050	22,607
Rhode Island	1,190	282	502	209	2,183
South Carolina	3,769	1,588	2,925	896	9,178
South Dakota	468	199	158	116	941
Tennessee	4,030	1,851	4,182	1,154	11,217
Texas	10,728	5,751	6,669	3,135	26,283
Utah	1,000	486	628	289	2,403
Vermont	490	171	188	72	921
Virginia	5,478	2,096	2,832	1,117	11,523
Washington	4,638	2,048	1,816	786	9,288
West Virginia	2,004	778	2,009	567	5,358
Wisconsin	4,478	1,700	1,664	775	8,617
Wyoming	282	169	171	104	726

SOURCE: Author's calculations based on a 10 percent random sample of the DRF.

Notes

[1] According to the Social Security Advisory Board (2012a), CDRs over the 1996–2008 period resulted on average in more than $10 of savings per $1 spent. Yet, because of budgetary constraints, the number of processed CDRs declined from its peak of more than 1.8 million in 2000 to about 1.1 million by 2009.

[2] In 10 states, a Prototype process initiated in 1999 allows claimants receiving an initial denial to appeal directly to the hearing level without having to go through the reconsideration stage.

[3] The figures in Table 1 are derived from SSA (2009, Tables 60, 61, and 62). Additional years of data appear in those tables. The reason why concurrent applicants are excluded is discussed in the data and methodology section of this article.

[4] The ability to test the impact of any of these factors on the reversal rate of initial denials falls outside the scope of this investigation because of the lack of readily available data. The focus here is on the capacity of primary diagnosis codes to successfully predict disability outcomes through the adjudicative process. A recent preliminary publication by the Social Security Advisory Board (2012b) suggested that third-party representation at the initial determination level increases the likelihood of an allowance substantially for SSI claimants, but only marginally for DI applicants.

[5] For a summary on litigation affecting the disability determination process, see the Social Security Advisory Board (2012a).

[6] Rupp's model did not use the individual primary diagnosis codes, but instead used 16 body systems, which group the specific impairments (15 dummy variables in addition to the musculoskeletal body group serving as the reference category).

[7] Technical denials can occur for a variety of nonmedical reasons, such as engaging in SGA or lacking the required amount of work credits.

[8] For estimation purposes, a 10 percent random sample is used instead of the full DRF because of the computational demands of the estimated models. The 100 percent figures reported in Table 2 are directly derived from the values in Table 1. There are small discrepancies between the two sets of figures. For instance, the 10 percent random sample culls any observations without a known primary diagnosis code or outside the 50 states (Puerto Rico, the District of Columbia, and other territories).

[9] Notice that when estimated from a classical perspective, random coefficient models like the ones in this article make distributional assumptions about subsets of parameters that are in effect no different from those of a prior density. In other words, classical statisticians may also use prior distributions, even if they do not refer to them as such.

[10] All of the models are estimated using Markov Chain Monte Carlo (MCMC) methods. The algorithm is an example of what is known as a Metropolis-within-Gibbs random sampler. A "noninformative" proper prior specification is adopted, with hyperparameter values as suggested by Rossi, Allenby, and McCulloch (2005).

[11] In this article, I focus exclusively on the primary diagnosis codes. A cross-classification of unique primary and secondary diagnosis code combinations would yield many thousands of clusters nesting the individual-level data. Forthcoming research by the author investigates the correlation patterns between primary and secondary diagnosis codes among initial determinations.

[12] To the best of my knowledge, the full extent to which the primary diagnosis change may occur on appeal across the full listing of impairments has never been documented.

[13] Because sex is an individual-level predictor in my models, I merge a few primary impairments that are gender specific. The single category "malignant neoplasm of the genital organs" combines four female diagnosis codes (malignant neoplasms of the uterus, cervix, ovaries, and other female genital organs) with three male diagnosis codes (malignant neoplasms of the prostate, testes, and penis and other male genital organs).

[14] In a Bayesian context, the mean and standard deviation of the posterior density can be used to compute approximate bounds on the posterior probability that a parameter changes sign (much like the t-statistics typically reported in the classical approach).

[15] If a model includes claimant-level predictors, there is a group variance parameter estimate associated with every explanatory variable and not just with the intercepts. However, because the claimant-level predictors have been centered around their grand mean, the intercepts carry the interpretation of adjusted mean linear predictions (see Raudenbush and Bryk (2002)).

[16] In discrete categorical models, a common identification restriction imposes a constant variance. For the multinomial logit case, the within-group variance has a logistic distribution with variance $\pi^2/3$. I follow the approach in Grilli and Rampichini (2007) to recover the ICC estimates.

[17] Notice that a fixed-effects model with the primary impairments rather than body systems would have required 180 indicator variables in the regression, potentially posing serious computational difficulties. In addition, it is unlikely that using the impairments would have substantially increased the share of explained state-level variation.

[18] Surprisingly, as many as 77 percent of the survey respondents were unaware of any activities at the hearing level or above, which appears to undercut the relevance of the result.

References

Coe, Norma B., Kelly Haverstick, Alicia H. Munnell, and Anthony Webb. 2011. "What Explains State Variation in SSDI Application Rates?" CRR Working Paper No. 2011-23. Chestnut Hill, MA: Center for Retirement Research at Boston College.

Congdon, Peter. 2005. *Bayesian Models for Categorical Data*. New York, NY: John Wiley & Sons.

Congressional Budget Office. 2010. "Social Security Disability Insurance: Participation Trends and Their Fiscal Implications." Economic and Budget Issue Brief. Washington, DC: Congressional Budget Office, Health and Human Resources Division (July 22).

Grilli, Leonardo, and Carla Rampichini. 2007. "A Multilevel Multinomial Logit Model for the Analysis of Graduates' Skills." *Statistical Methods & Applications* 16: 381–393.

Hu, Jianting, Kajal Lahiri, Denton R. Vaughan, and Bernard Wixon. 2001. "A Structural Model of Social Security's Disability Determination Process." *Review of Economics and Statistics* 83(2): 348–361.

Keiser, Lael R. 2010. "Understanding Street-Level Bureaucrats' Decision Making: Determining Eligibility in the Social Security Disability Program." *Public Administration Review* 72(2): 247–257.

Lahiri, Kajal, Denton R. Vaughan, and Bernard Wixon. 1995. "Modeling SSA's Sequential Disability Determination Process Using Matched SIPP Data." *Social Security Bulletin* 58(4): 3–42.

Leonesio, Michael V., Denton R. Vaughan, and Bernard Wixon. 2003. "Increasing the Early Retirement Age Under Social Security: Health, Work and Financial Resources." Health and Income Security for an Aging Workforce, Brief No. 7. Washington, DC: National Academy of Social Insurance.

Panis, Constantijn, Ronald Euller, Cynthia Grant, Melissa Bradley, Christin E. Peterson, Randall Hirscher, and Paul Steinberg. 2000. *SSA Program Data User's Manual*. Prepared by the RAND Corporation (contract no. PM-973-SSA) for the Social Security Administration.

Raudenbush, Stephen W., and Anthony S. Bryk. 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods,* 2nd edition. Thousand Oaks, CA: Sage Publications, Inc.

Rossi, Peter E., Greg M. Allenby, and Rob McCulloch. 2005. *Bayesian Statistics and Marketing*. New York, NY: John Wiley & Sons.

Rupp, Kalman. 2012. "Factors Affecting Initial Disability Allowance Rates for the Disability Insurance and Supplemental Security Income Programs: The Role of the Demographic and Diagnostic Composition of Applicants and Local Labor Market Conditions." *Social Security Bulletin* 72(4): 11–35.

Rupp, Kalman, and David Stapleton. 1995. "Determinants of the Growth in the Social Security Administration's Disability Programs: An Overview." *Social Security Bulletin* 58(4): 43–70.

Social Security Administration. 2009. *Annual Statistical Report on the Social Security Disability Insurance Program, 2008*. Washington, DC: Office of Retirement and Disability Policy, Office of Research, Evaluation, and Statistics.

———, Office of the Inspector General. 2010. *Disability Impairments on Cases Most Frequently Denied by Disability Determination Services and Subsequently Allowed by Administrative Law Judges*. Audit Report No. A-07-09-19083. Baltimore, MD: Office of the Inspector General.

Social Security Advisory Board. 2001. *Charting the Future of Social Security's Disability Programs: The Need for Fundamental Change*. Washington, DC: Social Security Advisory Board (January).

———. 2006. *Disability Decision Making: Data and Materials*. Washington, DC: Social Security Advisory Board (May).

———.2012a. *Aspects of Disability Decision Making: Data and Materials.* Washington, DC: Social Security Advisory Board (February).

———. 2012b. *Filing for Social Security Disability Benefits: What Impact Does Professional Representation Have on the Process at the Initial Application Level?* Washington, DC: Social Security Advisory Board (September).

Spiegelhalter, David J., Nicola G. Best, Bradley P. Carlin, and Angelika van der Linde. 2002. "Bayesian Measures of Model Complexity and Fit (with discussion)." *Journal of the Royal Statistical Society (Series B)* 64(4): 583–639.

SSA. *See* Social Security Administration.

Strand, Alexander. 2002. "Social Security Disability Programs: Assessing the Variation in Allowance Rates." ORES Working Paper Series No. 98. Washington, DC: Social Security Administration, Office of Research, Evaluation, and Statistics (August).