Measurement of Learning Disabilities and Intellectual Disabilities: Racial Patterns and Labor Market Biases in RSA-911 Data

Christopher M. Clapp
University of Chicago

John Pepper
University of Virginia

Robert Schmidt
University of Richmond

Steven Stern*
Stony Brook University

June 2023

Abstract

This paper provides novel evidence on how racial differences in the classification of learning and intellectual disabilities bias inferences on labor market outcomes of vocational rehabilitation program clients. Estimates using Rehabilitation Services Administration data from Virginia imply that whites with learning disabilities have worse labor market outcomes than non-whites. We argue this unusual finding reflects racial differences in how disabilities are classified. In particular, the thresholds used to classify intellectual disabilities are more stringent for white than black students. We are the first to model and estimate the disability classification process to understand racial gaps in labor market outcomes of the disabled. Using an endogenous disability classification model, we find substantial biases in the estimated labor market coefficients. At minimum, the estimated white-black employment gap is biased down by 3.2% and the earnings gap by 10%.

1 Introduction

Research on students in special education programs has consistently documented the tendency for blacks to be disproportionately labeled as having an intellectual

*We would like to thank Joe Ashley, Kirsten Rowe, Vlad Mednikov, and Jim Rothrock from Virginia Department of Aging and Rehabilition Services (DRS) and Rachel Fowley, Frank Keat, and Dan Yu for excellent research assistance. The DRS, the Virginia Department of Medical Assistive Services, the National Institute for Disability and Rehabilitation Research, and the University of Virginia Bankard Fund for Political Economy provided generous financial support. All errors are ours.

disability.[1] For example, Finn (1982) finds that 3.85% of black children were labeled as having an intellectual disability compared to 1.25% of white children.[2] Although some of these patterns are thought to reflect true underlying differences in the incidence of disabilities, the diagnostic standards are known to differ by race.[3] That is, blacks and whites with identical underlying impairments are classified differently (Reschly, 1997; Oswald, Coutinho, and Best, 2002). NRC (1982), Harry and Anderson (1994) and Chin (2021) argue that placement of students in special education was a way to segregate children by race even after Brown vs. Board of Education (1954). A number of court cases argued exactly that point (e.g., Diana vs. State Board of Education, 1970; Johnson vs. San Francisco Unified School District, 1974; Larry P. et al. vs. Wilson Riles et al., 1979), and, in Virginia, the focus of our analysis, this issue reached the state legislature at the end of the last century (Ladner and Hammons, 2001).

In this paper, we examine how systematic racial differences in classifying learning disabilities (*LD*) and intellectual disabilities (*ID*) impact inferences on the white-black employment and earnings gaps.[4,5] While the most disabled students are classified as *ID* rather than *LD*, the threshold for classifying whites is known to be more stringent. Since whites classified as *ID* are more negatively selected than blacks, they are more disabled than their black peers, on average. Thus, the observed white-black labor market gaps for those classified as *ID* will be smaller than gaps conditional on the true latent disability. Similarly, all else equal, we expect the white-black gap conditional on *LD* classification to also be biased down because more severely-disabled whites are classified as *LD* instead of *ID*. In addition, it may be that the disability threshold for being labeled as *LD*

[1] See, for example, Finn, 1982; Chinn and Hughes, 1987; Harry and Anderson, 1994; Coulter, 1996; Oswald et al.,1999; Ladner and Hammons, 2001; Skiba et al., 2001; Oswald, Coutinho, and Best, 2002; Parrish, 2002; Zhang and Katsiyannis, 2002; and Elder et al. (2021) .

[2] The term mental retardation was used to describe this condition, but it is no longer used. Larson et al. (2001) note that the proportion of students receiving special education services who were labeled as having mental retardation declined from 24% in 1977 to 16% in 1986 to 11% in 1995. They attribute some of the decline to a tendency to replace mental retardation with learning disabilities because of the stigma associated with mental retardartation (US DoE, 1998) and cite other research suggesting that it is differentially changing with respect to race (Mercer 1973; McDermott and Altekruse, 1994; Andrews et al., 1995; Murphy et al., 1995).

[3] In addition, a number of papers provide evidence that disability classification standards are impacted by a range of non-diagnostic factors including state rules for special education (Cullen, 2003; Figlio and Getzler, 2006; Dhuey and Lipscomb, 2011), the Individuals with Disabilities Education Act (Lewit and Baker, 1996; and Hanushek, Kain and Rivkin, 2002), and the liberalization of SSI eligibility and generosity of benefits (Kubik, 1999).

[4] This is related to the literature on the consequences of group differences in disability classifications and medical treatments. For example, recent work by Low and Pistaferri (2020) demonstrates that erroneous rejections of disability insurance applications is more common among women than among men. A large literature in health economics explores racial differences in medical treatments for patients with equal underlying health status (see, e.g., Chandra and Staiger 2010 and the references therein).

[5] There is much evidence showing that learning disabilities affect labor market outcomes. See, for example, Peraino (1992), Blackorby and Wagner (1997), Reder and Vogel (1997), Wagner et al. (2005), Barkley (2006), Gregg (2007, 2013), and Dapeppo (2009). These studies do not explain the racial differences in labor market outcomes we see in our data.

(instead of not disabled) is higher for black students rather than white students, meaning that the least-disabled blacks labeled *LD* are less disabled than the least-disabled whites labeled *LD*. This would lead to further downward bias in labor market gaps conditional on *LD* classification. We are the first to model and estimate this endogenous disability classification process to understand racial gaps in labor market outcomes among workers classified as *ID*/*LD*.

We address this question through the lens of the vocational rehabilitation (VR) program. The VR program provides an ideal setting to study labor market outcomes of persons with disabilities. Administered by the Rehabilitation Services Administration (RSA), the federal-state VR program gives approximately \$3 billion annually to state agencies to provide a variety of vocational rehabilitation services to individuals with a wide range of disabilities. State VR agencies have closed an average of over $600,000$ cases annually, with slightly more than 10% of these being cases with a diagnosis of intellectual or learning disability (Butterworth et al., Table 8, 2011). The VR program is the dominant program aimed at helping people with disabilities prepare for and enter the labor market.

Our basic methodological approach extends the analysis Dean et al. (2015) of the impact of Virginia's VR program on clients with intellectual disabilities (also see Dean et al., 2017, 2018 and 2019, hereafter referred to as DPSS).[6] DPSS analyze the effect of VR services on labor market outcomes using a structural instrumental variable model and administrative data collected for the Rehabilitation Services Administration, RSA-911 data. To examine the impact of classification biases on the white-black labor market gaps, we disaggregate the data on Virginia's VR clients with cognitive impairments into the mutually exclusive and exhaustive sets of individuals with an intellectual disability (*ID*)[7] and individuals with a learning disability (*LD*).

These two disabilities are generally diagnosed in high school by a social worker in order to receive special education services. The diagnosis involves some subjective assessments, and many of the same objective factors cause both *LD* and *ID* (NCLD, 2014). MacMillan and Siperstein (2002) define a learning disability as "a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, which may manifest itself in an imperfect ability to listen, think, speak, read, write, spell, or to do mathematical calculations." Types of learning disability include dyslexia, dyscalculia, dysgraphia, dyspraxia, executive functioning, and (sometimes) ADHD.[8] Intellectual disabilities, which are considered to be more

[6] Dean et al. (2015) focuses on clients with intellectual disabilities. They find VR services provided to clients with intellectual disabilities have an annual median rate of return of around 20%. For these VR clients, they estimate that the employment propensity for whites is 0.025 greater than non-whites and, conditional on employment, quarterly earnings are 8.3% smaller for whites than non-whites. DPSS find that the problem causing whites to have lower employment and conditional earnings does not exist for people with mental illness or physical impairments.

[7] To avoid confusion, a different font is used when using a variable name.

[8] There has been a long debate on how to diagnose learning disabilities with one camp preferring a diagnosis based on poor educational performance with an IQ in a normal range and

limiting than *LD*, are defined as a condition that includes below-threshold intellectual function and a lack of skills necessary for daily living (NIH, 2013). Causes of intellectual disability include infections (present at birth or occurring after birth), chromosomal abnormalities (such as Down syndrome and Fragile X syndrome), environmental (such as lead poisoning), metabolic (such as hyperbilirubinemia), nutrition, toxic intrauterine exposure to alcohol, cocaine, amphetamines, and other drugs, and trauma (before and after birth) (Miranda, Maxson, and Edwards, 2009; ARC, 2011; Johnson, 2012; NIH, 2013). The largest cause of intellectual disability is called sociocultural intellectual disability, which usually results in borderline or mild *ID*.[9] Theoretically, the lack of a stimulating environment and opportunity, particularly in infancy and early childhood, results in diminished brain development and thus poorer cognitive function (Alexander, 1998; Miranda, Maxson, and Edwards, 2009).

To evaluate biases in the white-black labor market gap caused by this subjective classification process, we reestimate the Dean et al. (2015) model with two critical innovations: first, we introduce an endogenous classification process and, second, we allow for interactions between race and an *ID/LD* disability indicator. To do this, we estimate a selection model that allows for the unobserved factors impacting the *ID/LD* classification to be related to labor market outcomes. This is similar to the classic problem of interpreting racial or gender differences in wage rates among workers when there is selection into labor force participation that differs by race or gender. For example, Neal (2004) argues that the observed white-black wage gap among female workers is artificially small because the distribution of ability among black female workers stochastically dominates the distribution of ability among white female workers. Also, see Chandra (2000), Blundell et al. (2007), and Mulligan and Rubinstein (2008).

In Section 2, we summarize the data and present some basic descriptive statistics that illustrate the classification issues and labor market outcomes. Section 3 presents our endogenous classification model. To identify this selection model, we use instrumental variables – the classification rates in Virginia School districts by race - which are assumed to impact the classification probability but not the labor market outcomes.

Section 4 presents the results. When classification is exogenous, whites with *LD* are estimated to have worse labor market outcomes than blacks, on average, suggesting substantial biases in the estimated white-black labor market gap. For those with *ID*, whites are more likely to work and earn more, on average, than blacks. Allowing for endogenous classification, we find evidence of substantial negative labor market biases for clients with *LD* and *ID*. That is, the white-black labor market gap gets notably larger once we account for the selection problem. However, the model does not seem to fully account for classification biases as whites with *LD* still have worse labor market outcomes than blacks, on average.

In Section 5, we discuss and analyze the selection model results. First, we

---

the other preferring one based on positive responses to services aimed at learning disabilities.

[9] There are many causes of ID, but physicians find a specific cause in only 25% of cases (NIH, 2013).

present a simple illustrative model to demonstrate that classification biases can lead to the unusual labor market results documented in Section 4 where whites with *LD* have worse labor market outcomes than blacks. In this illustration, the threshold for *ID* classification is more stringent for whites than for blacks (i.e., whites are less likely to be labeled *ID* conditional on the true latent ability level) so that for a fixed ability distribution whites labeled as having an *ID* are more disabled than blacks with the same diagnosis. Similarly, the threshold of *LD* classification is more stringent for whites and, among those diagnosed as having an *LD*, blacks will be more able than whites because more severely disabled whites are inappropriately classified as having a *LD* rather than an *ID*. The RSA-911 data summarized in Section 2 and the parameter estimates reported in Section 4 are generally consistent with the illustrative model: white VR clients with a *LD* have worse labor market outcomes, lower levels of schooling and higher rates of the most significant disabilities.

Second, we discuss a key limitation of our endogenous classification model that might help explain why whites with *LD* are estimated to have worse average labor market outcomes than blacks. Although the VR-data are well-suited for examining the labor market outcomes of people with *ID/LD*, the administrative RSA-911 data only include VR clients. Thus, as in many evaluations of job training programs, we cannot identify the parameters associated with the decision to take-up VR and the classification of any disability (see Imbens and Wooldridge 2009; Heckman et. al. 1999; and DPSS).

Finally, we conclude Section 5 by analyzing the bias in the observed associations between race and labor market outcomes using the results presented in Section 4. Overall, this analysis implies substantial biases in the estimated race parameters created by the different diagnostic standards used to classify disabilities across race. Section 6 concludes by summarizing the key arguments, results, and lessons of the paper.

2 Data

Our primary data source is the RSA-911 administrative records from the Virginia Department of Aging and Rehabilitative Services (DARS) for a cohort of applicants in 2000 merged with state unemployment insurance (UI) data on quarterly earnings. DPSS use these data to evaluate how the receipt of vocational rehabilitation (VR) services affects labor market outcomes. In this paper, we focus on how the measurement of disability affects estimates of labor market outcomes by race.

2.1 Our Sample and Key Variables

We begin with the administrative records of the Virginia DARS for the 10,176 individuals[10] who applied for VR services in SFY 2000 (July 1, 1999 - June

[10] This number is a little different than that reported in Dean et al. (2015) because of small differences in what was counted as part of the original sample and what was counted as being

Table 1: Sample Selection Analysis

Cause	# Obs	Remaining
Start		10176
State Missing	59	10117
Neither LD nor ID	6612	3505
Education Missing	0	3505
Primary Disability Code Missing	51	3454
Secondary Disability Code Missing	5	3449
Early DARS Episode	851	2598
Age < 19	340	2258
Age > 64	0	2258
Incomplete Observations	81	2177

Note: This table shows the number of VR clients dropped from our sample for different sequential sample selection criteria.

30, 2000). Table 1 provides information about sequential selection into our sample. The major causes of selection out of the sample are that the individual does not have an *ID/LD* diagnosis or that the individual had a prior DARS service spell.[11] There are a number of other selection criteria, listed in Table 1, that have minor impacts on the final sample. After selection, we have a sample size of 2,177 individuals.

For each individual in our sample, we observe a rich set of explanatory variables described in Table 2. A key variable for this study is an indicator for the respondent being white. All estimates from this paper treat race as binary: white or not white. While most of the literature on bias distinguishes between white and black, one should think of our results as treating not white as a close approximation of black in this population. In our sample, 63.6% of the sample is white, 34.7% is black, and 1.7% is neither white nor black.[12]

Besides the usual demographic variables such as gender, race, and age, we observe some disability measures not commonly included in other data sets and some other variables particularly relevant for this population.[13] In particular,

screened out of the original sample.

[11] We exclude people with prior DARS spells to avoid left-censoring issues discussed in Heckman and Singer (1984). In particular, an individual participating in a subsequent service episode may be doing so because she was (endogenously) unsuccessful in the labor market or because she found the first service episode (endogenously) unusually productive. In either case, inclusion of such individuals causes estimation bias. See Dean et al. (2015) for an estimate of the bias caused by left-censoring among clients with cognitive impairments.

[12] By comparison with the 2000 DARS data, Ipsen, Jain, and Stern (2022) find in national RSA-911 data for 2015 that the sum of black and white VR clients is 97.5%.

[13] A significant number of observations have missing information about education. Rather than delete them, we include a dummy variable for when education is missing.

While some of the explanatory variables such as marital status, access to transportation, and number of dependents may be endogenous, we follow DPSS as well as the related literature (e.g., Ettner, Frank, and Kessler, 1997) by including them as indicators of inclusion in society

Table 2: Explanatory Variable Sample Means and Standard Deviations

Variable	Mean	Std. Dev.	Variable	Mean	Std. Dev.
Male	0.570	0.495	Education	7.703	5.321
White	0.636	0.481	Special Education Certificate	0.226	0.418
Intellectual Disability	0.435	0.496	Education Missing	0.079	0.270
Learning Disability	0.565	0.496	Age (Qtrs/100)	0.900	0.343
Internal Disability	0.043	0.202	Married	0.047	0.211
Mental Illness	0.189	0.392	# Dependents	0.263	0.767
Other Disability	0.089	0.284	Government Assistance	0.114	0.215
Significant Disability	0.594	0.491	Transportation Available	0.613	0.487
Most Significant Disability	0.270	0.444	Has Driving License	0.410	0.492

Note: Means and standard deviations of the explanatory variables used in the model.

we have measures of the individuals' disabilities. We aggregate codes used by DARS into five broad disability categories as follows: a) intellectual disability is identified by the code for mental retardation (surprisingly, RSA continued to use the term at least through the 2016 RSA-911 handbook); b) learning disability is identified as those having a cognitive impairment but not an intellectual disability; c) internal disability includes respiratory impairments, general physical debilitation, and physical impairments not listed separately; d) mental illness includes anxiety disorders, depressive and other mood disorders, personality disorders, schizophrenia and other psychotic disorders, and mental illness not listed elsewhere; and e) other disability includes an assortment of impairments with too few observations to estimate impacts separately with any precision. The last group includes alcohol or drug abuse, attention deficit/ hyperactivity disorder, autism spectrum disorder, and musculo-skeletal impairments. In our sample of VR clients with cognitive impairments, 43.5% are diagnosed with *ID* and 56.5% with *LD*. These two are mutually exclusive and exhaustive categories. The other disability categories (e.g., physical impairment) are included as exogenous explanatory variables. We also have a measure of severity of the individual's disability, evaluated by the counselor, recorded as either *not significant disability* (the reference case), *significant disability*, or *most significant disability*. Over one-quarter of the VR clients with *ID/LD* are classified as having a most significant disability.

We also include six measures of VR service types (see Dean et al., 2015): diagnosis, training, education, restoration, maintenance, and other services. Summary statisitcs for these service variables can be found in an online appendix (Stern, 2023).[14]

and responsibility. Likewise, the variable measuring government financial assistance may be endogenous because the rules associated with receipt depend critically on involvement in the labor market. However, for our population, most individuals can participate in the labor market to some degree without losing their benefits. For this population, the income thresholds at which government benefits are reduced or eliminated is relatively high.

[14] Diagnosis & evaluation are provided at intake in assessing eligibility and developing an IPE; training includes vocationally-oriented expenditures for on-the-job training, job coach training, work adjustment, and supported employment; education includes tuition and fees

To estimate coefficients relating the explanatory variables in Table 2 to employment and log quarterly earnings, we merge the DARS data with state UI data. While it would be valuable to decompose quarterly earnings into wage level and hours, this is not possible in the UI data. The UI data provide information about individual quarterly earnings prior to, during, and after VR service receipt. In particular, we have a minimum of 16 quarters of pre-application employment history (1996-2000) and a minimum of 38 quarters post-application history (2000-2009). For people with either *LD* or *ID*, before service receipt, the mean employment rate is 27.5%, and the mean log quarterly earnings (conditional on working) is 6.613. After service receipt, they are respectively 47.9% and 7.541. See Dean et al. (2015) for further details on the UI data.

Data from the Virginia Department of Education (VDEC, 2022) and the Bureau of Labor Statistics (BLS, 2022) are used to create a set of instrumental variables for the disability classification model. Recall that ID/LD are generally diagnosed in high school by a social worker in order to receive special education services. VDEC data from 2010, the earliest available, are used to estimate the school system classification propensity with respect to ID/LD labeling. In particular, *pblackLD* is the proportion of blacks in the county special education program who have a *LD* diagnosis multiplied by $(1-white)$ and *pwhiteLD* is the proportion of whites in the county special education system who have a *LD* diagnosis multiplied by *white*. There are timing differences between the disability diagnosis (year varies based on the individual's age), the VR application (2000), and VDEC classification data (2010).[15] To control for these differences, the BLS data are used to measure the log ratio of per capita county income in the year an individual graduated from high school relative to the per capita income in 2000, the year an individual applied for VR services.[16] This variable is called *PCI adjustment*. One should think of it as a measure of how well-off people were in the county around the time the individual was being labeled with an ID/LD diagnosis. *PCI adjustment* is used to control for the fact that our special education classification variables are measured in 2010 while individuals graduated from high school (and were labeled) at different times.[17] While there

for a GED (graduate equivalency degree) program, a vocational or business school, a community college, or a university; restoration covers a wide variety of medical expenditures including dental services, hearing/speech services, eyeglasses and contact lenses, drug and alcohol treatments, psychological services, surgical procedures, hospitalization, prosthetic devices, and other assistive devices; maintenance includes cash payments to facilitate everyday living and covers such items as transportation, clothing, motor vehicle and/or home modifications, and services to family members; and other services consists of payments outside of the previous categories such as for tools and equipment.

[15] The VDEC classification rate data are not available prior to 2010, and we do not observe when an individual's disability diagnosis occurs. As reported in Table 2, the mean age of individuals in the sample when applying for VR benefits in 2000 was $(0.900 \times 100 \div 4 =)$ 22.5.

[16] We do not observe the age at graduation from high school in the data. Instead, we assume that all *ID* and *LD* students graduated high school when they were 21 (which is a reasonable assumption given the distribution of ages of graduation from high school).

[17] Intuitively, one can think of each school-district classification rate as an intercept and each *PCI adjustment* as the slope of a line that translates each of the other instruments into an appropriately timed variable.

Table 3: Classification Model Instrumental Variable Means and Standard Deviations

Variable	Mean	Std Dev
pblackLD	0.298	0.393
pwhiteLD	0.534	0.413
PCI Adjust	-0.194	0.150

Notes:
1) Means and standard deviations of the classification model (Equation 2) instrumental variables.
2) pblackLD is the proportion of blacks in special education programs who have a LD diagnosis by county multiplied by (1-white).
3) pwhiteLD is the proportion of whites in special education programs who have a LD diagnosis by county multiplied by (1-white).
4) PCI Adjustment measures the log ratio of per capita income in the year an individual graduated from HS relative to per capita income in 2010.

is a persistent component of district diagnosis procedures, classification biases are thought to be changing over time and possibly less pronounced for younger generations. Thus, to some degree, *PCI adjustment* is also a proxy for age. Table 3 shows the descriptive statistics for these explanatory variables.

Finally, in order to control for unobserved variation in local labor markets, we construct a variable, *local employment rate*, defined as the number of people working in a county divided by the number of people living in the county. The data come from the Bureau of Economic Analysis (2010).

2.2 Descriptive Analysis

Before formally estimating a model of the disability classification process and labor market outcomes, it is useful to examine the descriptive evidence comparing how mean labor market outcomes vary by *ID/LD* and race. Figure 1 displays the average employment rate and the mean log quarterly earnings (conditional on working) by race and disability (*ID/LD*). In general, whites have slightly worse mean labor market outcomes than blacks. For example, comparing the first pair of bars to the third, the average employment rate for whites with an *ID* is 0.288, nearly 0.05 less than the analogous rate of 0.339 for non-whites. The one exception to this finding is the quarterly earnings for clients with *LD* where the mean for whites of 7.612 is nearly 0.3 greater than for blacks. Overall, the fact these differences are either negative or close to zero for VR clients with *LD* and *ID* suggests that the disability classification process may lead to biased estimates of the white-black labor market gap. After all, based on established findings in the literature, we would expect the true differences in employment

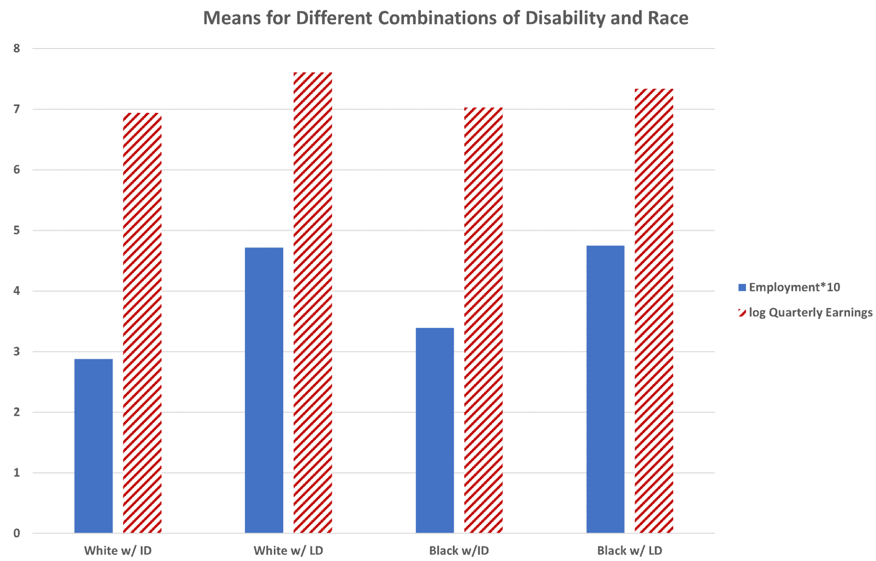

Figure 1: Means for Different Combinations of Disability and Race

and earnings to be positive.[18]

Direct evidence of systematic differences in "ability" across race and ID/LD classification can be found by examining the empirical relationship of the ID/LD choice with other observed covariates. Table 4 displays the mean and standard deviations of the labor market outcomes and the covariates by race and disability classification of the VR clients with ID or LD. Some of the differences across race are consistent with the notion that whites labeled as having a learning disability were less able than non-whites with the same diagnosis. In particular, whites have slightly fewer years of education (8.4 vs 8.8), are more likely to receive special education services (0.141 vs 0.122), and are more likely to be classified as having a most significant disability (0.172 vs 0.148).[19] Other differences are harder to reconcile with our hypothesis. In particular, whites are more likely to have access to transportation and a driver's license. In general, these observed differences are consistent with the notion that whites labeled with LD are less able. Of course, this comparison is based on observed covariates that

[18] That whites have better average labor market outcomes (employment and earnings) than blacks has been documented in the general population of working age adults (see Section 4.1), as well as across nearly every subpopulation of interest including those with disablities. For example, DPSS observe this basic outcome for adults with different types of disabilities. For VR clients with physical disabilities, Dean et al. (2018) estimate the employment propensity for whites is 0.036 greater than non-whites and, conditional on employment, quarterly earnings is 16.5% higher for whites than non-whites. The same qualitative patterns are found for VR clients with mental illness.

[19] Measures of significance of disability, which are based on a counselor's evaluation, may reflect racial biases. We are not aware of any evidence on whether this is an important issue in classifying the severity of VR client disabilities.

Table 4: Variable Means and Standard Deviations by Race and Disability Type

Variable	White with Intellectual Disability			White with Learning Disability			Black with Intellectual Disability			Black with Learning Disability		
	# Obs	Mean	Std Dev	# Obs	Mean	Std Dev	# Obs	Mean	Std Dev	# Obs	Mean	Std Dev
Dependent Variables												
Employment	30914	0.288	0.453	49358	0.472	0.499	24070	0.339	0.473	21924	0.475	0.499
log Quarterly Earnings	3838	6.937	1.414	10800	7.612	1.334	3368	7.031	1.410	3995	7.334	1.402
Explanatory Variables												
Male	533	0.510	0.500	851	0.616	0.486	415	0.477	0.499	378	0.651	0.477
Visual Disability	533	0.015	0.122	851	0.001	0.034	415	0.010	0.098	378	0.000	0.000
Hearing/Speech Disability	533	0.049	0.215	851	0.019	0.136	415	0.031	0.174	378	0.021	0.144
Musculo/Skeletal Disability	533	0.069	0.254	851	0.043	0.204	415	0.063	0.242	378	0.029	0.168
Internal Disability	533	0.062	0.241	851	0.027	0.162	415	0.060	0.238	378	0.032	0.175
Mental Illness	533	0.197	0.398	851	0.213	0.409	415	0.171	0.377	378	0.146	0.353
Substance Abuse Disability	533	0.006	0.075	851	0.008	0.090	415	0.017	0.129	378	0.008	0.089
Other Disability	533	0.141	0.348	851	0.063	0.244	415	0.108	0.311	378	0.050	0.218
Disability Significant	533	0.563	0.496	851	0.596	0.491	415	0.576	0.494	378	0.653	0.476
Disability Most Significant	533	0.415	0.493	851	0.172	0.377	415	0.395	0.489	378	0.148	0.355
Education	533	6.737	5.412	851	8.435	5.181	415	6.448	5.465	378	8.796	4.806
Special Education Certificate	533	0.332	0.471	851	0.141	0.348	415	0.357	0.479	378	0.122	0.327
Education Missing	533	0.030	0.171	851	0.119	0.323	415	0.041	0.198	378	0.101	0.301
Age (Quarters/100)	533	1.042	0.445	851	0.819	0.250	415	0.967	0.369	378	0.809	0.223
Married	533	0.056	0.230	851	0.063	0.244	415	0.024	0.153	378	0.021	0.144
# Dependents	533	0.261	0.781	851	0.177	0.602	415	0.508	1.066	378	0.190	0.614
Government Assistance	533	0.211	0.261	851	0.026	0.118	415	0.229	0.244	378	0.047	0.151
Transportation Available	533	0.480	0.500	851	0.805	0.396	415	0.424	0.494	378	0.577	0.494
Has Driver's License	533	0.205	0.403	851	0.678	0.467	415	0.128	0.334	378	0.407	0.491

Note: Means and standard deviations of the labor market outcomes and the covariates by race and disability classification of the VR clients with ID or LD.

are included in the regression analysis. The possible bias caused by labeling differences has to do with unobserved factors.

Figure 2 displays the proportion of VR applicants who have either LD or ID disaggregated by race and age at time of VR application. First, it shows the rapid increase over age in the proportion of clients with ID. This is almost surely caused by the increased priority of helping people with intellectual disabilities. For young applicants ($age < 26$), i.e. those who finished high school recently, proportions are equal between blacks and whites. For older applicants, i.e. those who finished high school under earlier regimes, the proportions are much smaller, and they vary in size between the two races. Second, the proportion with LD has also increased over age and thus time and more so for whites than for blacks. These are important facts that are relevant for this analysis in that they illustrate the nature of the composition effects driving the result. It also suggests that we need to control for age (or really cohort) when analyzing the data. We include client's age and the $PCI\ adjustment$ variable to do so.

3 Model

For this analysis, we extend the model in Dean et al. (2015) to account for an endogenous disability classification process applied to a sample of VR clients with either LD or ID. Here, we present the details of the disability classification, labor market and service provision models.

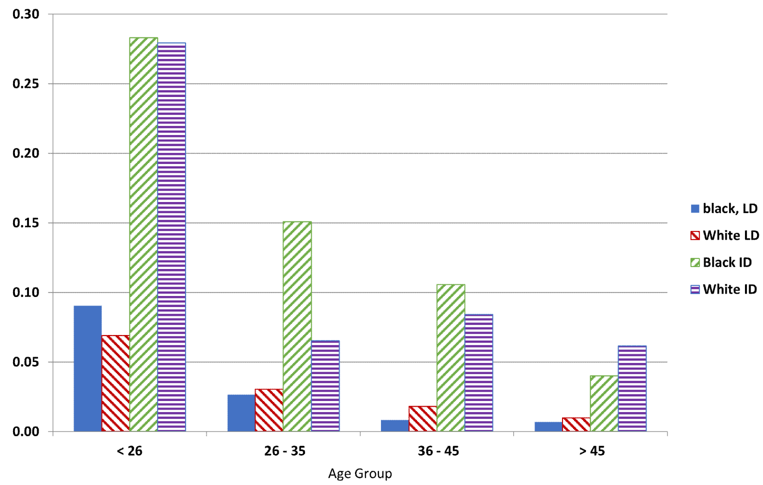

Figure 2: Learning Disability and Intellectual Disability Density by Age and Race

We begin with our model of the disability classification process. Let d_i be a dummy equal to one if i reports having an ID and zero if i reports having a LD. Assume that

$$\begin{aligned} d_i &= 1\left(\widetilde{d}_i \leq 0\right), \\ \widetilde{d}_i &= d_i^* + X_i^{\widetilde{d}}\widetilde{\psi} + u_i^{\widetilde{d}} + \widetilde{\xi}_i, \\ d_i^* &= X_i^{d*}\psi^* + u_i^{d^*} + \xi_i^* \end{aligned} \tag{1}$$

where $\widetilde{d}_i$ is a latent variable associated with disability reports, d_i^* is a latent, continuous measure of true disability (conditional on having either a LD or an ID), $X_i^{\widetilde{d}}$ and X_i^{d*} are vectors of exogenous explanatory variables of the type discussed in Sections 2 that affect both reporting decisions and true disability, and $\left(u_i^{\widetilde{d}}, u_i^{d^*}, \widetilde{\xi}_i, \xi_i^*\right)$ are unobserved factors whose structure is specified below. We can plug the definition of d_i^* into the equation for $\widetilde{d}_i$ in equation (1) to get

$$\begin{aligned} \widetilde{d}_i &= X_i^{d*}\psi^* + X_i^{\widetilde{d}}\widetilde{\psi} + u_i^{\widetilde{d}} + u_i^{d^*} + \widetilde{\xi}_i + \xi_i^* \\ &= X_i^d\psi + u_i^d + \xi_i \end{aligned} \tag{2}$$

(where $X_i^d = X_i^{d*} \cup X_i^{\widetilde{d}}$, $\psi = \psi^* + \widetilde{\psi}$, $u_i^d = u_i^{\widetilde{d}} + u_i^{d^*}$, and $\xi_i = \widetilde{\xi}_i + \xi_i^*$). It is clear from this equation that, for any variable x_i in both $X_i^{\widetilde{d}}$ and X_i^{d*}, we can estimate only the total association, $x_i\psi_{k^*}^* + x_i\widetilde{\psi}_{\widetilde{k}}$, where k^* indicates the element of ψ^* associated with x_i and $\widetilde{k}$ indicates the element of $\widetilde{\psi}$ associated with x_i. Since race is in both X_i^{d*} and $X_i^{\widetilde{d}}$, we cannot distinguish between the effect of race on d_i^* and its effect on $\widetilde{d}_i \mid d_i^*$ (which is the effect we are most interested in). Thus, instead of estimating the $\widetilde{d}_i$ and d_i^* in equation (1), we estimate the second line of equation (2).[20]

In practice, we use a very parsimonious specification of X_i^d in equation (2). In particular, we include four variables in X_i^d: *white*, *pblackLD*, *pwhiteLD,* and *PCI adjustment*. Moreover, *white*, is included in the two labor market equations (see below), but the other three variables are excluded. Thus, to identify the labor market parameters in our endogenous classification model, *pblackLD*, *pwhiteLD* and *PCI adjustment* are assumed to be instrumental variables that are associated with disability classification, do not directly affect employment or earnings, and are unrelated to the unobserved labor market and disability classification factors. Similar instrumental variables are applied in Doyle (2007), Maestas, Mullen, and Strand (2013), and DPSS.

Three key features of our model and data support these standard instrumental variable assumptions of relevance, an exclusion restriction, and independence. First, as shown subsequently in Table 6, these instruments are associated with the *ID/LD* classification probability. Second, the school district *ID/LD* classification propensities in 2010 (*pblackLD*, *pwhiteLD*), along with the *PCI adjustment* anchored to a client's age at graduation, should be unrelated to the quarterly labor market outcomes of VR clients from 1996 to 2009. This means that school district-level factors only affect individual labor-market outcomes through the individual's actual classification (but not directly).[21]

Finally, the school district disability classification propensities along with the *PCI adjustment* are plausibly independent of the unobserved factors associated with an individual's disability classification in high school and the labor market outcomes. Our model of employment and earnings (see equations (3) and (4)) include a rich set of exogenous control variables (see Table 2) including the VR clients' contemporaneous limitations and county-level employment rates. Furthermore, our factor model (see equation (6)) accounts for persistent, unobserved effects. Together, these rich controls mitigate concerns about possible correlations between the instruments and the unobserved factors.

Although we cannot test whether the instruments are unrelated to these unobserved factors, we can compare the instruments to the observed covariates. Using a type of balance table, we find the instruments are only weakly correlated with the many observed explanatory variables (see the on-line appendix, Stern, 2023): most of these correlations are less than 0.1 in absolute value, while a few are between 0.1 and 0.2. The one exception is the nearly perfect correlation between the (*pblackLD*, *pwhiteLD*) instruments and the *white* indicator variable.

Next, we consider the modelling of labor market outcomes. Let z_{it}^* be the

[20] The same issue occurs in Stern (1989), and the solution is the same.

[21] This identification assumption would be threatened if, for instance, employers infer something about true ability not only from an individual's classification, but also knowledge of the nature of the classification process in the individual's high school. We develop a factor model that allows us to condition on fixed unobserved factors (e.g., ability, motivation) across our disability classification, labor market and service provision models that mean there is unlikely additional information to be gained from the district level.

value to i of working at quarter t, and define $z_{it} = 1\left(z_{it}^{*} > 0\right)$. Let w_{it} be the log quarterly earnings of i at t. Assume that

$$z_{it}^{*} = X_{it}^{z}\gamma^{x} + d_i\gamma^{d} + u_{it}^{z} + \eta_{it}^{z} \tag{3}$$

and that

$$w_{it} = X_{it}^{w}\delta^{x} + d_i\delta^{d} + u_{it}^{w} + \eta_{it}^{w} \tag{4}$$

where X_{it}^{z} and X_{it}^{w} are vectors of (possibly) time-varying, exogenous explanatory variables listed in Table 2, and u_{it}^{z}, u_{it}^{w}, η_{it}^{z}, and η_{it}^{w} are errors whose structure is specified below.

Except for the race indicator *white*, the explanatory variables included in the labor market equations (3) and (4) are excluded from the disability classification equation. These exclusions reflect timing differences between VR receipt in 2000 and ID/LD classification in high school at various times before 2000 depending on the client's age. In particular, covariates such as marital status, other disabilities, number of dependents, and whether the client has access to transportation or a driver's license in 2000 should not directly affect prior ID/LD classification in high school. In addition to excluding these variables, we also assume they are independent of the unobserved determinants of disability classification, ξ_i. In part, this assumption reflects the timing differences between disability assignment and VR receipt; observed covariates in 2000 will not affect random unobserved factors influencing previous disability classification in high school. Moreover, our factor model in equation (6) accounts for more persistent or fixed unobserved factors such as ability or motivation that might be correlated with the 2000 covariates such as access to transportation or marital status (see the factor model estimates in Section 4.2.3). Given that our model accounts for these critical factors, we believe the exogeneity assumption is credible. Finally, note that these exclusion restrictions are not required to identify the parameters of the disability classification equation (2) both because of the nonlinearity of the model and also there are no endogenous variables in the disability equation.

Following Dean et al. (2015), we also include a VR service provision equation. As described Section (2), we use six different measures of VR service types; $J = 6$. Let y_{ij}^{*} be the value for individual i of participating in VR service j, $j = 1, 2, .., J$, and define $y_{ij} = 1\left(y_{ij}^{*} > 0\right)$ be an indicator for whether i receives service j. Assume that

$$\begin{aligned} y_{ij}^{*} &= X_i^{y}\beta_j + u_{ij}^{y} + \varepsilon_{ij}, \\ \varepsilon_{ij} &\sim iidLogistic \end{aligned} \tag{5}$$

where X_i^{y} is a vector of exogenous explanatory variables and u_{ij}^{y} is an error whose structure is specified in equation (6). All of the covariates in Table 2 are included in the service model in (5). In addition, following DPSS, we address the possible endogeneity of service receipt using two instrumental variables for each of the six binary service provision variables. The first is the proportion of other clients of the individual's counselor who were provided with the service and the second is the proportion of other clients of the individual's VR field office who were provided with the service. These instruments are associated with service receipt and only weakly correlated with the observed covariates (See the online appendix, Stern 2023).

Finally, assume that

$$\begin{aligned} u_i^d &= \lambda_1^d e_{i1} + \lambda_2^d e_{i2}, \quad u_{ij}^y = \lambda_{j1}^y e_{i1} + \lambda_{j2}^y e_{i2} \\ u_{it}^z &= \lambda_1^z e_{i1} + \lambda_2^z e_{i2} + \eta_{it}^z, \quad u_{it}^w = \lambda_1^w e_{i1} + \lambda_2^w e_{i2} + \eta_{it}^w, \\ \eta_{it}^z &= \rho_\eta^z \eta_{it-1}^z + \zeta_{it}^z, \ \eta_{it}^w = \rho_\eta \eta_{it-1}^w + \zeta_{it}^w, \\ \xi_i &\sim iidN\left[0,1\right], \\ \begin{pmatrix} \zeta_{it}^z \\ \zeta_{it}^w \end{pmatrix} &\sim iidN\left[0, \sigma_\zeta^2 \begin{pmatrix} 1 & \rho_\zeta \\ \rho_\zeta & 1 \end{pmatrix}\right], \\ \begin{pmatrix} e_{i1} \\ e_{i2} \end{pmatrix} &\sim iidN\left[0, I\right], \\ \upsilon_{it}^z &\sim iidN\left[0,1\right], \text{ and } \upsilon_{it}^w \sim iidN\left[0, \sigma_w^2\right]. \end{aligned} \tag{6}$$

We include (e_{i1}, e_{i2}) to allow for two common factors affecting all dependent variables with factor loadings $\left(\lambda_k^d, \lambda_{jk,}^y \lambda_k^z, \lambda_k^w\right)_{k=1}^2$. We also allow for serial correlation and contemporaneous correlation in the labor market errors $(\eta_{it}^z, \eta_{it}^w)$. The model presented in equations (3) through (6) is the same as that in Dean et al. (2015). Equation (2) is the addition to the model to allow for endogenous disability classification.

Appendix A details how this model is estimated.

4 Results

In this section, we present the estimated coefficients of the explanatory variables on employment and conditional earnings. We first consider the case where disability classification is exogenous. That is, we do not include equation (2) in the model. We then allow for endogenous classification. In each case, we interact *white* with indicators for *ID* and *LD* classification. Our focus is on these interaction variables.

4.1 Assuming Exogenous Disability Classification

Table 5 displays the estimated coefficients on the explanatory variables for employment and conditional earnings when we do not include equation (2) in the model.[22] Our focus is on the estimated white-black labor market gaps. However, first we discuss the other explanatory variable estimates. The results show

[22] We also estimate the parameters associated with VR service receipt (see equations (5)). See the on-line appendix (Stern, 2023) for estimates of the associated coefficients. The estimates associated with the receipt of vocational rehabilitation services are consistent with those reported in Dean et al. (2015).

Table 5: Labor Market Equation Estimates Assuming Exogenous Classification

Variable	Employment			log Quarterly Earnings		
	Estimate		Std Err	Estimate		Std Err
Male	0.088	**	0.008	0.172	**	0.012
White*Intellectual Disability	0.240	**	0.011	0.170	**	0.016
White*Learning Disability	-0.219	**	0.011	-0.219	**	0.017
Intellectual Disability	-2.491	**	0.027	4.311	**	0.052
Learning Disability	-2.203	**	0.027	4.656	**	0.051
Internal Disability	-0.261	**	0.019	-0.247	**	0.027
Mental Illness	0.052	**	0.011	-0.087	**	0.017
Other Disability	-0.076	**	0.014	-0.169	**	0.020
Significant Disability	0.184	**	0.013	0.244	**	0.018
Most Significant Disability	0.405	**	0.015	0.408	**	0.022
Local Employment Rate	1.739	**	0.152	-1.254	**	0.152
Education	0.064	**	0.001	0.077	**	0.003
Special Education Certificate	0.675	**	0.017	0.799	**	0.038
Education Missing	0.232	**	0.020	0.565	**	0.043
Age (# Quarters/100)	1.996	**	0.012	1.645	**	0.023
Married	-0.587	**	0.016	-0.301	**	0.031
# Dependents	-0.001		0.006	-0.080	**	0.008
Government Assistance	-0.861	**	0.021	-0.878	**	0.029
Transportation Available	0.002		0.010	0.109	**	0.014
Has Driver's License	0.151	**	0.010	0.258	**	0.016

Notes:

1) Estimated coefficients and standard errors on the explanatory variables for the labor market outcome equations when disability classification is exogenous (i.e., Equation 2 is not included).

2) Double-starred items are statistically significant at the 5% level, and single-starred items are statistically significant at the 10% level.

3) Labor market variables are employment propensity and conditional log quarterly earnings.

that the effects of education on employment propensity and conditional quarterly earnings are respectively 0.064 and 0.077. Getting a special education certificate also has a large effect on both. The effect of age on employment propensity is 8% for every extra year ($1.196 \times 4/100$), and it's effect on earnings is 6.6%. The penalty for having an ID or LD is large relative to that for other disabilities. However, surprisingly, increasing the severity of one's disability has beneficial effects on both labor market outcomes. We include the local employment rate as an explanatory variable to control for unobserved variation in the labor markets across Virginia. The variable we use has a large, positive effect on employment and a large, negative effect on conditional earnings.

The empirical labor economics literature provides a very robust result that being white increases one's employment probability and wages. Oaxaca and Ransom (1994) find that black men over 25 earn 12.5% less than white men over 25 after controlling for a full set of other observables. Johnson and Neal (1996) report a reduction in wages of 19.6% for black men relative to white men and 15.5% for black women relative to white women (Table 1). Carneiro, Heckman, and Masterov (2005) find a reduction of 19.4% for men and 8.6% for women after controlling for a small set of explanatory variables similar to those in Johnson and Neal (1996) (Table 5). Chetty et al. (2019) report employment rates for black men between 11.4 and 18.9 percentage points less than white men depending upon the income of their parents.

In our case, when we do not include the disability classification equation (2) in the model, the estimates for $white \times LD$ are negative. For those classified as having a *LD*, being white decreases one's employment propensity by 0.219 and decreases one's conditional quarterly earnings by 21.9%.[23] These negative estimates suggest some classification bias. In contrast, the results for $white \times ID$ are positive. For those classified as having an *ID*, being white increases one's employment propensity by 0.240 and increases one's conditional quarterly earnings by 17.0%.

4.2 Assuming Endogenous Disability Classification

In this section, we estimate the full model, including equation (2) which allows for the endogenous classification of *ID*. (relative to those with *LD*).

4.2.1 Classification Model Estimates and Discussion

Table 6 reports the estimates for equation (2). The estimate for *white* is negative (-2.029) and statistically significant, implying that whites are less likely to be classified as having an *ID* than non-whites. This result is consistent with previous analyses of the special education population (see, for example, Finn, 1982; and Elder et al., 2021) that find the threshold used for *ID* classification is more stringent for whites than blacks so that whites labeled as *ID* are more disabled, on average, than their black *ID*-labeled peers. In other words, conditional on the true latent disablity level, whites are less likely to be labeled as *ID*.

The estimate for *pblackLD* is negative (-1.794), statistically significant and has the expected sign; as the proportion of blacks with a *LD* in the black individual's county increases, the probability of having an *ID* diagnosis decreases. The estimate for *pwhiteLD* is small and statistically insignificant. The estimate for *PCI adjustment*, which is statistically significant at the 10% level, implies that individuals living in counties that were significantly poorer when the individual was in high school are less likely to have a *LD* diagnosis. As seen in Figure 2, *PCI adjustment* is highly correlated with the individual's age implying that, over time, *LD* diagnoses are becoming more prevalent than *ID* diagnoses (which is consistent with other papers in the literature such as Mercer (1973), McDermott and Altekruse (1994), Andrews et al. (1995), Murphy et al. (1995), and Larson et al. (2001).

Although our specification of the selection equation is parsimonious, we use a Lagrange Multiplier (LM) test to assess whether *education*, *special ed-*

[23] To translate employment propensity effects into approximate marginal employment probability, multiply by $\phi(0) \approx 0.4$.

Table 6: Intellectual Disability Classification Equation Estimates

	Estimate		Std Err
Constant	1.451	**	0.437
White	-2.029	**	0.615
pblackLD	-1.794	**	0.538
pwhiteLD	0.252		0.509
PCI Adjustment	-0.324	*	0.194

Notes:
1) Estimated coefficients and standard errors for the disability classification model (equation 2) parameters.
2) Double-starred items are statistically significant at the 5% level, and single-starred items are statistically significant at the 10% level.
3) The dependent variable is a dummy equal to one iff the individual has an ID diagnosis.

ucation, and *most significant disability* along with each of the three variables interacted with *white* has a statistically significant association with the *ID* classification probability. The LM statistic for the joint null is 402.5, and individual t-statistics are -11.6 for *education*, 7.2 for *special education*, 7.6 for *most significant disability*, -11.8 for *education*$\times$*white*, 3.2 for *special education*$\times$*white*, and 3.9 for *most significant disability*$\times$*white*. The mean derivatives of the log likelihood are -1.83 for *education*, 0.06 for *special education*, 0.07 for *most significant disability*, -1.50 for *education*$\times$*white*, 0.02 for *special education*$\times$*white*, and 0.03 for *most significant disability*$\times$*white*. So, for *special education*, *most significant disability*, and interactions of the two with *white*, the improvement in the log likelihood function is small but statistically significant, while, for *education* and *education*$\times$*white*, the improvement in the log likelihood function is large and statistically significant. This implies that there are racial differences in the relationship between human capital characteristics and disability classification.

We also perform an LM test to determine if *age*, *white*, and *age*$\times$*white* should be part of the model. Figure 2 shows an interaction of *age* (or cohort) with *white*; as *age* increases, whites are less likely to be classified as having an *ID*. The LM statistic for the joint null is 274.3, and individual t-statistics are -1.3 for *age*, -8.8 for *white*, and -4.2 for *age*$\times$*white*. The mean derivatives of the log likelihood are -0.02 for *age*, -0.12 for *white*, and -0.06 for *age*$\times$*white*. Thus, white VR clients have a lower probability of being classified as having an *ID* (relative to *LD*), and this difference decreases with age, as illustrated by Figure 2.

Table 7: Labor Market Equation Estimates Assuming Endogenous Classification

Variable	Employment			log Quarterly Earnings		
	Estimate		Std Err	Estimate		Std Err
Male	0.036	**	0.008	0.209	**	0.012
White*Intellectual Disability	0.238	**	0.011	0.412	**	0.018
White*Learning Disability	-0.124	**	0.012	-0.118	**	0.017
Intellectual Disability	-2.460	**	0.031	3.812	**	0.060
Learning Disability	-2.212	**	0.030	4.163	**	0.058
Internal Disability	-0.485	**	0.024	-0.336	**	0.031
Mental Illness	0.120	**	0.011	0.191	**	0.020
Other Disability	0.100	**	0.014	-0.090	**	0.021
Significant Disability	0.044	**	0.013	0.183	**	0.021
Most Significant Disability	-0.090	**	0.016	0.208	**	0.024
Local Employment Rate	2.280	**	0.161	-1.795	**	0.161
Education	0.054	**	0.001	0.080	**	0.003
Special Education Certificate	0.325	**	0.018	0.549	**	0.041
Education Missing	-0.239	**	0.023	0.468	**	0.046
Age (# Quarters/100)	1.994	**	0.013	1.804	**	0.028
Married	-0.741	**	0.019	-0.295	**	0.033
# Dependents	-0.103	**	0.006	-0.155	**	0.009
Government Assistance	-0.575	**	0.021	-0.822	**	0.033
Transportation Available	0.284	**	0.010	0.278	**	0.014
Has Driver's License	0.280	**	0.010	0.297	**	0.016

Notes:

1) Estimated coefficients and standard errors on the explanatory variables for the labor market outcome equations when disability classification is endogenous (i.e., Equation 2 is included).

2) Double-starred items are statistically significant at the 5% level, and single-starred items are statistically significant at the 10% level.

3) Labor market variables are employment propensity and conditional log quarterly earnings.

4.2.2 Employment and Earnings Model Estimates

Table 7 displays the estimated coefficients on the explanatory variables when we include equation (2) in the model. The key variable we focus on is *white*. In this model, the estimates for $white \times LD$ are still negative but substantially larger (closer to zero) than those reported when classification is assumed to be exogenous. In particular, the estimated employment propensity coefficient increases from -0.219 to -0.124, and the estimated conditional log-quarterly earnings coefficient increases from -0.219 to -0.118. Likewise, the estimates for $white \times ID$ for the conditional log-quarterly earnings equation increase substantially, from 0.170 to 0.412. The estimate for the employment equation, at 0.238, is nearly identical to the one found when classification is assumed to be exogenous. Overall, these results imply that allowing for endogenous classification substantially increases the estimated coefficients on *white* for clients with *LD* and *ID*.

Allowing for endogenous classification, we find evidence of substantial negative labor market biases for clients with *LD* and *ID*. In particular, we find that whites classified as *ID* are more negatively selected than blacks, causing the observed white-black gap in labor market outcomes to be too small (e.g, the estimates for conditional log-quarterly earnings increase substantially, from 0.170 to 0.412). Similarly, all else equal, the white-black gap conditional on *LD* classification also is biased downward because more severely disabled whites are classified as *LD* instead of *ID*.[24] In addition, it may be that a larger fraction of black students are labeled as *LD*, meaning that the least-disabled blacks labeled *LD* are less disabled than the least-disabled whites labeled *LD*. This leads to further downward bias in labor market gaps conditional on *LD* classification. Once we account for the endogenous classification problem, the estimated white-black labor market gap gets notably larger. However, the model does not seem to fully account for classification biases as whites with *LD* still have worse labor market outcomes than blacks, on average.

4.2.3 Covariance Parameters and Factor Model Estimates

Finally, Table 8 presents the covariance parameter estimates associated with equation (6). We focus on the factor loadings for each of the two included factors. The first factor has a positive estimated factor loading for *employment* and a negative estimated factor loading for *log quarterly earnings*. The difference in sign suggests that this factor is a measure of eagerness to work. The factor loading for *ID* diagnosis is negative (-0.097) which suggests that people with an *ID* diagnosis have unobserved weaker preferences to work than people with an *LD* diagnosis (because it has a different sign than for *employment*). The second factor has positive estimated factor loadings for both *employment* and *log quarterly earnings*. Having the same sign suggests that this factor is a measure of ability. The estimated factor loading for *ID* diagnosis is statistically significant and positive (0.087) implying that ability and the unobserved characteristics that affect the probability of having *ID* relative to *LD* diagnosis are positively correlated. Given that *ID*s are generally considered to be more limiting that *LD*s, this result is somewhat surprising.

5 Discussion: Estimated Bias of the Race Coefficients

In this section, we further discuss the possible biases in our estimated race coefficients in Table 5. We begin by providing a simple model to illustrate how racial classification differences may bias inferences on the association between race and

[24] Interpreting how the changes in the coefficients between Tables 5 and 7 map into changes in the white black gap is complicated by the opposing signs of the positive $white \times ID$ and negative $white \times LD$ estimates. It may help the reader to visualize changes in the gap in both cases as shifts on a number line. For example, the *ID* conditional log-earnings gap shifts to the right on a number line from 0.170 to 0.412 when the bias is mitigated. For *LD*, the log-earnings gap again shifts to the right from -0.219 to -0.118.

Table 8: Covariance Estimates

	Factor 1			Factor 2		
Variable	Estimate		Std Err	Estimate		Std Err
Diagnosis & Evaluation	0.428	**	0.077	-0.024		0.071
Training	0.599	**	0.086	-0.073		0.084
Education	0.465	**	0.159	0.131		0.127
Restoration	1.481	**	0.081	-0.095		0.088
Maintenance	0.892	**	0.102	0.166	*	0.098
Other	0.592	**	0.092			
Employment	1.493	**	0.008	1.182	**	0.013
log Quarterly Earnings	-0.170	**	0.004	0.519	**	0.007
ID/LD Diagnosis	-0.097	**	0.038	0.087	**	0.039
log Earnings Std Dev	1.100	**	0.003			

Notes:
1) Covariance parameter estimates and standard errors (see Equation 6).
2) Double-starred items are statistically significant at the 5% level, and single-starred items are statistically significant at the 10% level.

labor market outcomes. The VR data are generally consistent with this model. We then discuss two key limitations of our endogenous classification model that might explain why whites with *LD* are estimated to have worse average labor market outcomes than blacks. Finally, we provide a more speculative analysis of the magnitude of the labor market biases in the estimated race coefficients reported in Table 5.

5.1 Simple Illustration

Our results in Table 5 on the association between race and labor market outcomes are likely to be impacted by classification or composition bias (see, for example, Solon, Barsky, and Parker, 1994; Lemieux, 2006). To see this, it is helpful to consider a simplified classification process where the density of abilities is the same for blacks and whites, as drawn in Figure 3. Without classification biases, whites and blacks have the same rates of *ID*s and *LD*s. Moreover, assuming preferences for leisure/work do not vary by race and no racial discrimination, the labor market outcomes would not differ across the races. Yet, in the VR data we find that whites are less likely to be classified as having an *ID*, and whites diagnosed with a *LD* work and earn less, on average, than *LD*-labeled blacks. Also, the estimates of the white coefficient in the endogenous classification model are substantially larger than those when classification is exogenous, for both the *LD* and *ID* interactions variables.

A systematic disability classification process can lead to these results. To generate these qualitative results in this simple setting, further assume that a white person is labeled as having an *ID* if his ability is to the left of the first (solid red) vertical line, he is labeled as having a *LD* if his ability is between

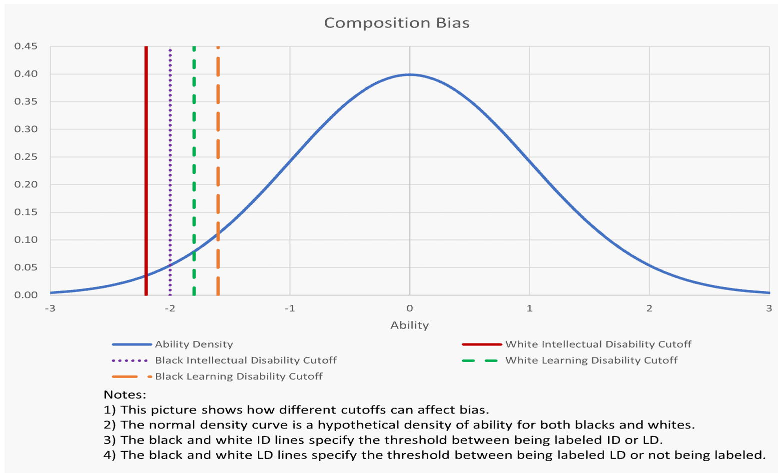

Figure 3: Composition Bias

the first (solid red) vertical line and the third (dashed green) vertical line, and he is labeled as having no cognitive disability if he is to the right of the third (dashed green) vertical line. Analogously, assume that a black person is labeled as having an *ID* if his ability is to the left of the second (dotted purple) vertical line, he is labeled as having a *LD* if his ability is between the second (dotted purple) vertical line and the fourth (long-dashed orange) vertical line, and he is labeled as having no cognitive disability if he is to the right of the fourth (long-dashed orange) vertical line.

These assumptions (and the uniform spacing of the vertical lines) imply that blacks are more likely to be labeled with an *ID* than whites and that blacks and whites have the same probability of having a *LD* label. On average, blacks with an *ID* have higher ability than whites with an *ID* because the cutoff point for blacks is to the right of the cutoff point for whites. In particular, the blacks between the first two vertical lines are labeled as having an *ID* and the whites as having a *LD*. This group of people raises average ability of blacks with an *ID* above that for whites. Also, on average, blacks with a *LD* have higher ability than whites with a *LD*. The blacks and whites between the second and third line are labeled as having a *LD*. However, only blacks between the third and the fourth vertical lines (with higher average ability) are labeled with a *LD*, and only whites between the first and section vertical lines (with lower average ability) are labeled with a *LD*. Thus, the ability composition of both the *ID* group and the *LD* group are biased in favor of blacks. This is consistent with what we see in our data.

Importantly, this illustration demonstrates how biases in classifying disabilities by race lead to biased inferences about the association between race and labor market outcomes. Without racial differences in standards used to classify disabilities, the labor market outcomes of blacks and whites would be identical in this model. With the classification differences, whites with a *LD* have lower ability and, therefore, will have worse labor market outcomes than blacks.

Of course, this illustration does not account for many important factors. Most notably, labor market discrimination and environmental and socioeconomic factors that contribute to the incidence of disabilities are known to lead to racial differences in labor market outcomes and disability classification. Yet, these differences are inconsistent with the finding that, among VR clients classified as having a *LD*, whites have lower employment rates and earnings than blacks. In Section 5.3, we estimate the magnitude of the bias in the labor market coefficient associated with white on labor market outcomes.

5.2 Limitations of the Classification Model and Data

The RSA-911 data summarized in Section 3 and the model estimates reported in Sections 4.1 and 4.2 are generally consistent with this simple illustration: white VR clients with a *LD* have worse labor market outcomes, lower levels of schooling and higher rates of the most significant disabilities.[25] Likewise, the LM tests discussed in Section 4.2 imply the racial differences in *ID/LD* disability classification vary with *education*, *special education* and *most significant disability*. Thus, blacks and whites with the same observed human capital characteristics are classified differently.

Classification biases result from unobserved factors jointly associated with race, disability classification, and labor market outcomes. While our classification model in Section 3 allows for unobserved confounders, this model does not seem to fully account for classification biases as whites with an *LD* are estimated to have worse labor market outcomes than blacks, on average. Our model of the classification of *ID* or *LD* may miss other important selection mechanisms. In particular, although the VR-data are well-suited for examining the labor market outcomes of people with *ID/LD*, the administrative RSA-911 data only include VR clients. Thus, we cannot identify parameters associated with the decision to take-up VR and with the classification of any disability.

As in DPSS, our analysis follows the conventional approach in evaluations of job training programs of ignoring selection problems associated with the application process (see Imbens and Wooldridge, 2009; Heckman et. al., 1999; and DPSS). Yet, focusing on applicants may create biases and limit the external validity of the estimates if the decision to apply for services is related to unobserved factors associated with labor market outcomes. There are, in fact, observed differences in the racial and gender composition of the selected sample of applicants (see Table 4) and the full population of special education students in VA (see VA-DOE, 2022).[26] In particular, VR applicants are more likely to be white than the population of special education students with *ID/LD* (e.g., for the *ID*, 56% versus 50%, and for the *LD*, 69% versus 60%). For the gender

[25] A similar argument could be made about differences across race in the labeling of people as disabled (or not). See, for example Cullen (2003) or Figlio and Getzler (2006).

[26] Beyond race and gender, the VA-DOE data do not include other covariates that are in the RSA-911 data.

composition, VR applicants with *ID* are less likely to be male (for whites, 51% versus 56% and for blacks, 48% versus 59%), while the composition is similar for those with *LD*. In addition, a much smaller fraction of whites and women are classified as having any disability (US DOE, 1998). These differences between VR applicants and the general population suggest that the application decision is correlated with important factors that are associated with race, disability classification, and labor market outcomes.

While this may, in part, explain the *LD* results discussed in Section 4.2, the extent to which this selection problem biases the estimates is uncertain. There may be differences in observed covariates of applicants versus the full population, but the central issue is whether there are important unobserved confounders. The model controls for race, gender, and many other factors. In addition, the application probabilities appear to be similar for whites and blacks, suggesting that this extensive margin may not lead to big differences in the selection probabilities.[27] Finally, for clients with mental illnesses, Dean et al. (2017) conclude that, after controlling for the rich set of covariates observed in the RSA-911 data, the application decision to apply for VR services is exogenous.

5.3 Estimated Bias of Race Coefficients

In this section, we evaluate the bias in the race coefficient estimator caused by the differences in classification standards used for whites and blacks. We begin by comparing the estimates on the white coefficients in Tables 5 and 7 under the assumption that the classification model properly accounts for the selection bias problem. Focusing on clients diagnosed with a *LD*, the estimates on *white* from Tables 5 and 7 imply composition biases of 3.2% for the employment probability[28] and $(-0.118 + 0.219 =)$ 10.1% for quarterly earnings. For clients diagnosed with *ID*, the estimates imply a composition bias for quarterly earnings of $(0.412 - 0.170 =)$ 24.2%.

Arguably, the estimated biases for $white \times LD$ coefficients are a lower bound as negative coefficient estimates on $white \times LD$ imply that our model may not fully account for the selection problem. To examine the sensitivity of the bias to conjectured values of the true parameter value, we estimate the bias allowing the true parameters on $white \times LD$ to range from 0 to 0.25. A parameter of 0 means that race has no association with the employment propensity or log quarterly earnings, while a parameter of 0.25 implies that whites have 0.25 higher employment propensity and expected log-quarterly earnings than blacks. Figure 4 shows the estimated bias for the employment probability and the percent change mean quarterly earnings as a function of conjectured values of the true

[27] Using the VA-DOE (2022) data, we estimate the application probabilities by *race* and *ID*/*LD*. For *ID*, 12% of whites and 10% of blacks apply for VR services. For *LD*, the analogous probabilities are 3% and 2%.

[28] The average employment rate is 33.5%. Thus, if the true coefficient is -0.124 (see Table 7) the employment rate for whites is estimated to equal $\Phi\left[\Phi^{-1}(0.335) - 0.124\right] = 0.291$. The estimates without accounting for classification imply $\Phi\left[\Phi^{-1}(0.335) - 0.219\right] = 0.259$. Thus, the bias on the employment coefficient is $0.032 = 0.291 - 0.259$.

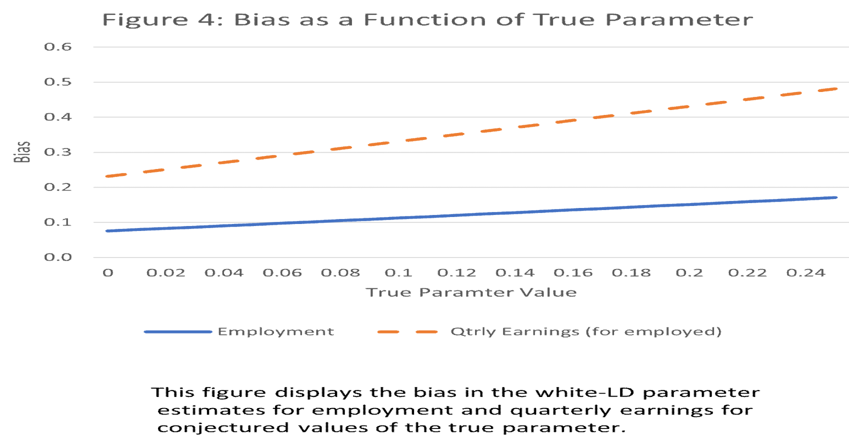

Figure 4: Bias as a Function of the True Parameter Value

parameter value. Most notably, even if race has no association with labor market outcomes, the biases of the estimated coefficients on employment (0.08) and earnings (0.219) are substantial. Moreover, the biases monotonically increase with the conjectured value of the true parameter. At 0.25, the bias is 19% for the employment probability parameter and 46.9% (=21.9%+25%) for quarterly earnings.

To focus on a particular value of the true parameter, we rely on results from DPSS on VR clients with other disabilities. Consider, for example, clients with physical disabilities. Reschly (1997) shows a very small racial bias in physical disability labeling. Dean et al. (2018) estimate a coefficient on *white* in a probit-like equation for employment as 0.036 and estimate a 16.5% increase in quarterly earnings associated with being white.[29]. These estimates imply composition biases of 8% for employment probability and a (21.9% + 16.5% =) 38% for quarterly earnings.

6 Conclusion

Diagnostic standards used to classify disabilities for special education services are known to differ across race. White and black students with the same underlying impairments are classified differently; whites are less likley to be classified as *ID*. In this paper, we provide the first evidence on how racial differences in classification of disabilities bias the associations between race and labor market outcomes for persons with learning and intellectual disabilities.

Using RSA-911 data on vocational rehabilitation services provided to clients with *ID* or *LD*, we find evidence of systematic racial classification of disabilities

[29] This implies an increase in employment rates of 1.3%. The average employment rate was 33.5%. Thus, $\Phi\left[\Phi^{-1}(0.335) + 0.036\right] - 0.335 = 0.013$.

that lead to downward-biased estimates of the white-black gap in average labor market outcomes. Following the model developed in Dean et al. (2015), we estimate that white clients with LD are less likely to be employed and earn less, on average, than non-white clients. After accounting for endogenous selection, the estimates increase substantially but are still negative. For clients with ID, the estimates for white are positive but, for quarterly earnings, substantially increase after accounting for disability classification.

The key question is how the classification standards might lead to the biased labor market findings. A simple illustrative model shows how different classification standards lead to a selected sample where whites with intellectual disabilities have lower ability than blacks with LD and thus have worse average labor market outcomes than blacks. Likewise, whites with learning disablities have lower unobserved ability than blacks.

Finally, we evaluate the biases in the labor market coefficients by comparing the estimates from the RSA-911 data with and without the endogenous classification model. These imply that, for those with LD, the employment propensity is biased by at least 3.2%. For quarterly earnings, the bias is at least 10.1% for those with LD and 24.2% for those with ID. To be clear, these estimated biases are lower bounds for the LD sample as our endogenous disability equation does not appear to fully capture the selection process. If race were to have no real impact on labor market outcomes, the employment propensity coefficient would be biased by 0.219 and the log earnings coefficient by 0.219. Thus, the differences in the classification of disabilities substantially biases inferences on the white-black gap of labor market outcomes of persons with learning and intellectual disabilities.[30]

7 Appendix A: Estimation

The model we use is similar to the model in Dean et al. (2015) and is estimated the same way. However, the model described in Section 3 can be estimated to gain information about the parameters in equation (2). Define

$$\theta = \left(\psi, \gamma, \delta, \lambda, \rho_d, \rho_\xi, \sigma_\zeta^2, \rho_\zeta, \sigma_w^2\right)$$

as the vector of parameters to estimate where $\psi = (\psi_1, \psi_1)$ are the effects of X_i^d on reported disability in equation (2), $\gamma = \left(\gamma^x, \gamma^d\right)$ are the effects of explanatory variables and disability on labor market participation in equation (3), $\delta = \left(\delta^x, \delta^d\right)$ are the effects of explanatory variables and disability on conditional log quarterly earnings in equation (4), λ are the factor loadings in equation (6),

[30] Whether these results can be generalized to other state VR agencies and other time periods is an open question and an important issue for future research to consider. Previous research on VR programs shows substantial cross-state and temporal heterogeneity, suggesting the labor market results reported in this paper and, more generally, in DPSS and Schmidt et al. (2019) might not be generalizable across states or over time and should be interpreted with caution. However, this concern is relevant for all papers on the effects of VR (e.g., see Aakvig et al., 2005).

and the other parameters are elements of the error structure. We can think of the data available to us as $\left\{d_i, [z_{it}, w_{it}, X_{it}]_{t=1}^{T_i}\right\}_{i=1}^{n}$.[31]

The log likelihood contribution for an observation is

$$\log L_i = \log \int L_i\left(e, d\right) dM\left[\zeta \mid 0, \Omega_\zeta\right] dB\left[e \mid 0, I\right]$$

where $e = (e_1, e_2)$ is the vector of factors with bivariate normal distribution $B\left[\cdot \mid 0, I\right]$, ζ_i is the vector of innovations defined in equation (6) with multivariate normal distribution $M\left[\cdot \mid 0, \Omega_\zeta\right]$,

$$\Omega_\zeta = \sigma_\zeta^2 \begin{pmatrix} 1 & \rho_\zeta \\ \rho_\zeta & 1 \end{pmatrix} \otimes I_{T_i},$$

and[32]

$$L_i\left(e, d\right) = L_i^d\left(e, d\right) \prod_{t=1}^{T_i} L_{it}^{lm}\left(e, d\right),$$

$$L_i^d\left(e, d\right) = \Phi\left(X_i^d \psi + u_i^d\right)^{d_i} \left[1 - \Phi\left(X_i^d \psi + u_i^d\right)\right]^{1-d_i},$$

$$\begin{aligned} L_{it}^{lm}\left(e, \zeta, d\right) &= \left[1 - \Phi\left(X_{it}^z \gamma^x + d_i \gamma^d + u_{it}^z\right)\right]^{1-z_{it}} \cdot \\ &\left[\frac{1}{\sigma_w} \phi\left(\frac{w_{it} - X_{it}^w \delta^x - d_i \delta^d - u_{it}^w}{\sigma_w}\right) \Phi\left(X_{it}^z \gamma^x + d_i \gamma^d + u_{it}^z\right)\right]^{z_{it}} \end{aligned}$$

where (u_{it}^z, u_{it}^w) are defined in equation (6). The log likelihood contribution can be simulated easily as in DPSS.

References

[1] Aakvik, Arild., James Heckman, and Edward Vytlacil (2005). "Estimating Treatment Effects for Discrete Outcomes When Responses to Treatment Vary: An Application to Norwegian Vocational Rehabilitation Programs." *Journal of Econometrics*. 125: 15-51.

[31] The actual data structure deviates from this in that we observe log quarterly earnings only when employed. The likelihood function that follows reflects the deviation.

[32] Note that

$$\begin{aligned} \Pr\left[d_i = 1 \mid X_i^d, u_i^d\right] &= \Pr\left[X_i^d \psi + u_i^d + \xi_i > 0\right] \\ &= \Pr\left[-\xi_i < X_i^d \psi + u_i^d\right]; \\ \Pr\left[d_i = 0 \mid X_i^d, u_i^d\right] &= \Pr\left[X_i^d \psi + u_i^d + \xi_i < 0\right] \\ &= \Pr\left[\xi_i < -\left(X_i^d \psi + u_i^d\right)\right]. \end{aligned}$$

[2] Alexander, Duane (1998). "Prevention of Mental Retardation: Four Decades of Research." *Mental Retardation and Developmental Disabilities Research Reviews*. 4: 50-58.

[3] Andrews, Howard, John Pittman, Doris Goldberg, Elmer Struening, and Norman Wellen (1995). "Prediction of Special Education Placement from Birth Certificate Data." *American Journal of Preventative Medicine*. 11(3): 55–61.

[4] ARC (2011). "Causes and Prevention of Intellectual Disabilities." http://www.thearc.org/page.aspx?pid=2453.

[5] Barkley, Russell (2006). *Attention-Deficit Hyperactivity Disorder: A Handbook for Diagnosis and Treatment, Third Edition*. New York: Guilford Press.

[6] Blackorby, Jose and Mary Wagner (1997). "The Employment Outcomes of Youth with Learning Disabilities: A Review of Findings from the National Longitudinal Transition Study of Special Education Students." *Learning Disabilities and Employment*. (eds.) P. Gerber and D. Brown. Austin, TX: PRO-ED.

[7] Blundell, Richard, Amanda Gosling, Hidehiko Ichimura, and Costas Meghir (2007). "Changes in the Distribution of Male and Female Wages Accounting for Employment Composition Using Bounds." *Econometrica*. 75(2): 323-363.

[8] Brown vs. Board of Education (1954). 347 U.S. 483, 74 S.Ct. 686, 98 L.Ed. 873.

[9] Bureau of Economic Analysis (2010). http://www.bea.gov/regional/docs/footnotes.cfm?tablename=CA04.

[10] Butterworth, John, Frank Smith, Allison Hall, Alberto Migliore, and Jean Winsor (2011). *State Data: The National Report on Employment Services and Outcomes: 2008*. Institute for Community Inclusion, University of Massachusetts, Boston, MA.

[11] Carneiro, Pedro, James Heckman, and Dimitriy Masterov (2005). "Labor Market Discrimination and Racial Differences in Premarket Factors." *Journal of Law and Economics*. 48(1): 1-39.

[12] Chandra, Amitabh (2000). "Labor-Market Dropouts and the Racial Wage Gap: 1940-1990." *American Economic Review*. 90(2): 333-338.

[13] Chandra, Amitabh and Douglas Staiger (2010). "Identifying Provider Prejudice in Healthcare" (No. w16382). National Bureau of Economic Research.

[14] Chetty, Raj, Nathaniel Hendren, Maggie Jones, and Sonya Porter (2020). "Race and Economic Opportunity in the United States: An Intergenerational Perspective." *Quarterly Journal of Economics*. 135(2): 711–783.

[15] Chin, Mark, (forthcoming). "JUE Insight: Desegregated but Still Separated? The Impact of School Integration on Student Suspensions and Special Education Classification," *Journal of Urban Economics*, 103389.

[16] Chinn, Phillip and Selma Hughes (1987). "Representation of Minority Students in Special Education Classes." *Remedial and Special Education.* 8(4): 41–46.

[17] Coulter, W. (1996). *Alarming or Disarming: The Status of Ethnic Differences within Exceptionalities.* 1-27.

[18] Cullen, Julie Berry (2003). "The Impact of Fiscal Incentives on Student Disability Rates." *Journal of Public Economics.* 87(7-8): 1557-89.

[19] Dapeppo, Lisa (2009). "Integration Factors Related to the Academic Success and Intent to Persist of College Students with Learning Disabilities." *Learning Disabilities Research and Practice.* 24(3): 122-131.

[20] Dean, David, John Pepper, Robert Schmidt, and Steven Stern (2015). "The Effects of Vocational Rehabilitation for People with Cognitive Impairments." *International Economic Review.* 56(2): 399-426.

[21] Dean, David, John Pepper, Robert Schmidt, and Steven Stern (2017). "The Effects of Vocational Rehabilitation for People with Mental Illness." *Journal of Human Resources.* 52(3): 826-858.

[22] Dean, David, John Pepper, Robert Schmidt, and Steven Stern (2018). "The Effects of Vocational Rehabilitation for People with Physical Disabilities." *Journal of Human Capital.* 12(1): 1-37.

[23] Dhuey, Elizabeth and Stephen Lipscomb (2011). "Funding Special Education by Capitation Evidence from State Finance Reforms." *Education Finance and Policy.* 6(2): 168-201.

[24] Diana vs. State Board of Education (1970). CA 70 RFT (N.D. Cal. 1970).

[25] Doyle, Joseph (2007). "Child Protection and Child Outcomes: Measuring the Effects of Foster Care." *American Economic Review.* 97(5): 1583-1610.

[26] Elder, Todd E., David N. Figlio, Scott A. Imberman, and Claudia L. Persico, (2021). "School Segregation and Racial Gaps in Special Education Identification," *Journal of Labor Economics*, 39 (S1), S151–S197.

[27] Ettner, Susan, Richard Frank, and Ronald Kessler (1997). "The Impact of Psychiatric Disorders on Labor Market Outcomes." *Industrial Labor Relations Review.* 51(1): 64-81.

[28] Figlio, David and Lawrence Getzler (2006). "Accountability, Ability, and Disability: Gaming the System?" *Improving School Accountability (Advances in Applied Microeconomics, Volume 14).* (eds.) Timothy Gronberg and Dennis Jansen. Emerald Group Publishing Limited: 35-49.

[29] Finn, Jeremy (1982). "Patterns in Special Education Placement as Revealed by the OCR Survey." *Placing Children in Special Education: A Strategy for Equity.* (eds.) K. A. Heller, W. Holtzman, and S. Messick. Washington, DC: National Academy Press.

[30] Gidugu, Vasudha, E. Sally Rogers, Mihoko Maru, Lauren Mizock, Philippe Bloch, Marci McCoy-Roth, and Bridgette Gavin (2011). *Review of Employment Services for Individuals with Intellectual and Developmental Disabilities: A Comprehensive Review of the State-of-the-Field from 1996-2011.* Center for Psychiatric Rehabilitation, Boston.

[31] Gregg, Noell (2007). "Underserved and Underprepared: Postsecondary Learning Disabilities." *Learning Disabilities Research and Practice.* 22(4): 219-228.

[32] Gregg, Noell (2013). "Adults with Learning Disabilities: Factors Contributing to Persistence." *Handbook of Learning Disabilities, Second Edition.* (eds.) H. Swanson, K. Harris, and S. Graham. New York: Guilford Press.

[33] Hanushek, Eric, John Kain, and Steven Rivkin (2002). "Inferring Program Effects for Special Populations: Does Special Education Raise Achievement for Students with Disabilities?" *Review of Economics and Statistics.* 84(4): 584-599.

[34] Harry, Beth and Mary Anderson (1994). "The Disproportionate Placement of African American Males in Special Education Programs: A Critique of the Process." *Journal of Negro Education.* 63: 602–619.

[35] Heckman, James and Burton Singer (1984). "Econometric Duration Analysis." *Journal of Econometrics.* 24: 63-132.

[36] Heckman, James, Robert LaLonde, and Jeffrey Smith (1999). "The Economics and Econometrics of Active Labor Market Programs." *Handbook of Labor Economics Volume 3A.* (eds.) Orley Ashenfelter and David Card. Amsterdam: North-Holland. 1865-2097.

[37] Imbens, Guido and Jeffrey Wooldridge (2009). "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economic Literature.* 47(1): 5–86.

[38] Ipsen, Catherine, Kamini Jain, and Steven Stern (2022). "Vocational Rehabilitation Service Receipt, Service Expenditures, and Ruralness." *Journal of Vocational Research*, forthcoming.

[39] Johnson vs. San Francisco Unified School District (1974). 500 F. 2d 349 (9th Cir. 1974).

[40] Johnson, Shannon (2012). "Mental Retardation." *HealthLine.* http://www.healthline.com/symptom/mental-retardation.

[41] Johnson, William and Derek Neal (1996). "The Role of Premarket Factors in Black-White Wage Differences." *Journal of Political Economy.* 104(5): 869-895.

[42] Kubik, Jeffrey (1999). "Incentives for the Identification and Treatment of Children with Disabilities: The Supplemental Security Income Program." *Journal of Public Economics.* 73(2): 187-215.

[43] Ladner, Matthew and Christopher Hammons (2001). "Special but Unequal: Race and Special Education." *Rethinking Special Education for a New Century.* (eds.) C. Finn, A. Rotherham, and C. Hokanson. Washington, DC: The Progressive Policy Institute and The Thomas B. Fordham Foundation.

[44] Larry P. et al. vs. Wilson Riles et al. (1979). 495 F. Supp. 926 (N.D. Cal. 1979).

[45] Larson, Sheryl, K. Lakin, Lynda Anderson, Nohoon Kwak Lee, Jeoung Hak Lee, and Deborah Anderson (2001). "Prevalence of Mental Retardation and Developmental Disabilities: Estimates From the 1994/1995 National Health Interview Survey Disability Supplements." *American Journal on Mental Retardation.* 106(3): 231-252.

[46] Lemieux, Thomas (2006). "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?" *American Economic Review.* 96(3): 461-498.

[47] Lewit, Eugene and Linda Baker (1996). "Child Indicators: Children in Special Education." *The Future of Children.* 6(1): 139–151.

[48] Low, Hamish and Luigi Pistaferri (2020). "Disability Insurance: Error Rates and Gender Differences." Unpublished manuscript.

[49] MacMillan, Donald and Gary Siperstein (2002). "Learning Disabilities as Operationally Defined by Schools." *Identification of Learning Disabilities: Research to Practice.* (eds.) R. Bradley, L. Danielson, and D. Hallahan. Mahwah, NJ: Lawrence: 287-333.

[50] Maestas, Nicole, Kathleen Mullen, and Alexander Strand (2013). "Does Disability Insurance Receipt Discourage Work? Using Examiner Assignment to Estimate Causal Effects of SSDI Receipt." *American Economic Review.* 103(5): 1797-1829.

[51] McDermott, Suzanne and Joan Altekruse (1994). "Dynamic Model for Preventing Mental Retardation in the Population: The Importance of Poverty and Deprivation." *Research in Developmental Disabilities.* 15: 49–65.

[52] Mercer, Jane (1973). *Labeling the Mentally Retarded: Clinical and Social Systems Perspectives in Mental Retardation.* Berkeley: University of California Press.

[53] Miranda, Marie, Pamela Maxson, and Sharon Edwards (2009). "Environmental Contributions to Disparities in Pregnancy Outcomes." *Epidemiologic Reviews*. 31(1): 67-83.

[54] Mulligan, Casey and Yona Rubinstein (2008). "Selection, Investment, and Women's Relative Wages over Time." *Quarterly Journal of Economics*. 123(3): 1061-1110.

[55] Murphy, Catherine, Marshalyn Yeargin-Allsopp, Pierre Decoufle, and Carolyn Drews (1995). "The Administrative Prevalence of Mental Retardation in 10-Year Old Children in Metropolitan Atlanta, 1985–1987." *American Journal of Public Health*. 85: 319–323.

[56] National Center for Learning Disabilities (2014). "What Causes Learning Disabilities?" http://www.ncld.org/types-learning-disabilities/what-is-ld/what-causes-learning-disabilities.

[57] National Institute of Health (2013). "Intellectual Disability." Medline Plus. http://www.nlm.nih.gov/medlineplus/ency/article/001523.htm.

[58] National Research Council (1982). *Placing Children in Special Education: A Strategy for Equity*. Washington, DC: The National Academies Press. https://doi.org/10.17226/9440.

[59] Neal, Derek. (2004). "The Measured Black-White Wage Gap among Women is too Small." *Journal of Political Economy*. 112(S1): S1-S28.

[60] Oaxaca, Ronald and Michael Ransom (1994). "On Discrimination and the Decomposition of Wage Differentials." *Journal of Econometrics*. 61(1): 5-21.

[61] Oswald, Donald, Martha Coutinho, and Al Best (2002). "Community and School Predictors of Overrepresentation of Minority Children in Special Education." *Racial Inequity in Special Education*. (eds.) D. Losen and G. Orfield. Cambridge, MA: Harvard Civil Rights Project.

[62] Oswald, Donald, Martha Coutinho, Al Best, and Nirbhay Singh (1999). "Ethnic Representation in Special Education: The Influence of School-Related Economic and Demographic Variables." *Journal of Special Education*. 32: 194–206.

[63] Parrish, Thomas (2002). "Racial Disparities in the Identification, Funding, and Provision of Special Education." *Racial Inequity in Special Education*. (eds.) D. Losen and G. Orfield. Cambridge, MA: Harvard Civil Rights Project.

[64] Peraino, Joseph (1992). "Post-21 Follow-Up Studies: How do Special Education Graduates Fare?" *Life Beyond the Classroom: Transition Strategies for Young People with Disabilities*. (ed.) P. Wehman. Baltimore: Brookes: 21-70.

[65] Reder, Stephen and Susan Vogel (1997). "Life-Span Employment and Economic Outcomes for Adults with Self-Reported Learning Disabilities." *Learning Disabilities and Employment.* (eds.) P. Gerber and D. Brown. Austin, TX: PRO-ED.

[66] Reschly, Daniel (1997). *Disproportionate Minority Representation in General and Special Education: Patterns, Issues, and Alternatives.* Des Moines: Iowa Department of Education.

[67] Skiba, Russell, Tony Wu, Kristin Kohler, Choong-Geun Chung, and Ada Simmons (2001). *Disproportionality and Discipline among Indiana's Students with Disabilities.* Bloomington: Indiana Education Policy Center.

[68] Solon, Gary, Robert Barsky, and Jonathan Parker (1994). "Measuring the Cyclicality of Real Wages: How Important Is Composition Bias?" *Quarterly Journal of Economics.* 109(1): 1-25.

[69] Stern, Steven (1989). "Measuring the Effect of Disability on Labor Force Participation." *Journal of Human Resources.* 24(3): 361-395.

[70] Stern, Steven (2023). "Complete Estimation Results for Effects of Vocational Rehabilitation Services on People with Learning and Intellectual Disabilities." https://sites.google.com/site/stevensterneconomics/research-interests/vocational-rehabilitation/full-set-of-on-line-estimates.

[71] U.S. Bureau of Labor Statistics (2022). *Historical CPI-U.* https://data.bls.gov/timeseries/CUUR0000SA0.

[72] U.S. Department of Education. (1998). *Twentieth Annual Report to Congress on the Implementation of the Individuals with Disabilities Act.* Washington, DC.

[73] Virginia Department of Education (2022). December 1 Special Education Child Count, 2010-11 Child Count Data. https://p1pe.doe.virginia.gov/apex_captcha/home.do?apexTypeId=307#tagline (accessed 8/2022).

[74] Wagner, Mary, Krista Kutash, Albert Duchnowski, and Michael Epstein (2005). "The Special Education Elementary Longitudinal Study and the National Longitudinal Transition Study." *Journal of Emotional and Behavioral Disorders.* 13(1): 25-41.

[75] Zhang, Dalun and Antonis Katsiyannis (2002). "Minority Representation in Special Education: A Persistent Challenge." *Remedial and Special Education.* 23: 180–181.