Online Appendices for
The Effects of Vocational Rehabilitation for People with Mental Illness

David Dean
University of Richmond

John Pepper
University of Virginia

Robert Schmidt
University of Richmond

Steven Stern*
University of Virginia

2016

1 Covariance Structure

The covariance matrix of the errors $u_i' = (u_{i1}^y, u_{i2}^y, .., u_{iJ}^y, u_{i1}^z, u_{i1}^w, u_{i1}^r, .., u_{iT}^z, u_{iT}^w, u_{iT}^r)$ implied by the structure in equation (5) of the paper is

$$\underset{(J+3T)\times(J+3T)}{\Omega} = \begin{pmatrix} \underset{J\times J}{A} & \underset{J\times 3T}{B'} \\ \underset{3T\times J}{B} & \underset{3T\times 3T}{C} + \underset{3T\times 3T}{D} \end{pmatrix}$$

where

$$A = \begin{pmatrix} \sum_k (\lambda_{1k}^y)^2 & \sum_k \lambda_{1k}^y \lambda_{2k}^y & \cdots & \sum_k \lambda_{1k}^y \lambda_{Jk}^y \\ \sum_k \lambda_{1k}^y \lambda_{2k}^y & \sum_k (\lambda_{2k}^y)^2 & \cdots & \sum_k \lambda_{2k}^y \lambda_{Jk}^y \\ \vdots & \vdots & \ddots & \vdots \\ \sum_k \lambda_{1k}^y \lambda_{Jk}^y & \sum_k \lambda_{2k}^y \lambda_{Jk}^y & \cdots & \sum_k (\lambda_{Jk}^y)^2 \end{pmatrix},$$

$$\begin{aligned} C &= H \otimes Q_T \\ \underset{3\times 3}{H} &= \begin{pmatrix} \sum_k (\lambda_k^z)^2 & \sum_k \lambda_k^z \lambda_k^w & \sum_k \lambda_k^z \lambda_k^r \\ \sum_k \lambda_k^z \lambda_k^w & \sum_k (\lambda_k^w)^2 & \sum_k \lambda_k^w \lambda_k^r \\ \sum_k \lambda_k^z \lambda_k^r & \sum_k \lambda_k^w \lambda_k^r & \sum_k (\lambda_k^r)^2 \end{pmatrix}, \\ \underset{T\times T}{Q_T} &= \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix} \end{aligned}$$

$$D = \underset{3\times 3}{\Omega_\zeta} \otimes \frac{1}{1-\rho_\eta^2} \begin{pmatrix} 1 & \rho_\eta & \cdots & \rho_\eta^{T-1} \\ \rho_\eta & 1 & \cdots & \rho_\eta^{T-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_\eta^{T-1} & \rho_\eta^{T-2} & \cdots & 1 \end{pmatrix},$$

*We would like to thank Joe Ashley, John Phelps, Kirsten Rowe, Ann Stanfield, Vlad Mednikov, and Jim Rothrock from Virginia Department of Aging and Rehabilitation Services (DARS), Paul O'Leary of the Social Security Administration, David Stapleton, other members of the National Institute of Disability and Rehabilitation Research (NIDRR) Advisory Group, Jeff Smith, and workshop participants at the ASSA Meetings, the Summer Econometric Society Meetings, APPAM, Clemson, Western Ontario, Washington University, University of Southern California, University of Louisville, University of Zurich, Hong Kong University of Science and Technology, National University of Singapore, Singapore Management University, and the University of Maryland for excellent help and advice and Rachel Fowley for excellent research assistance. The DARS, the Virginia Department of Medical Assistive Services, the NIDRR, and the University of Virginia Bankard Fund for Political Economy provided generous financial support. All errors are ours.

Table A.1: Other Covariance Terms

Variable	Estimate		Std Err	Variable	Estimate		Std Err
Var(ζ_1)	0.041	**	0.000	Var(ζ_2)	0.008	**	0.004
Cov(ζ_1,ζ_2)	0.018	**	0.005	Cov(ζ_2,ζ_3)	-0.018	**	0.005
Cov(ζ_1,ζ_3)	-0.040	**	0.000	Var(ζ_3)	0.041	**	0.000
			0.000	σ_w	1.281	**	0.003

Note: Double-starred items are statistically significant at the 5% level.

and

$$
\begin{aligned}
B &= q_T \otimes F, \\
\underset{T\times 1}{q_T} &= \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \\
\underset{3\times J}{F} &= \begin{pmatrix} \sum_k \lambda^y_{1k}\lambda^z_{1k} & \sum_k \lambda^y_{2k}\lambda^z_{1k} & \cdots & \sum_k \lambda^y_{Jk}\lambda^z_{1k} \\ \sum_k \lambda^y_{1k}\lambda^w_{1k} & \sum_k \lambda^y_{2k}\lambda^w_{1k} & \cdots & \sum_k \lambda^y_{Jk}\lambda^w_{1k} \\ \sum_k \lambda^y_{1k}\lambda^r_{1k} & \sum_k \lambda^y_{2k}\lambda^r_{1k} & \cdots & \sum_k \lambda^y_{Jk}\lambda^r_{1k} \end{pmatrix}.
\end{aligned}
$$

The estimates of the primitives, other than the factor loading estimates in Table 8, associated with the covariance matrix are reported in Table A.1.

2 Measuring Non-Purchased Services

Services can be provided to an individual in any combination of three ways: a) as a "purchased service" through an outside vendor using DARS funds, b) as a "similar benefit" purchased or provided by another governmental agency or not-for-profit organization with no charge to DARS, and/or c) internally by DARS personnel. The DARS administrative data do not, however, provide the same detailed information for in-house services or similar benefits. Instead, we measure non-purchased service provision using two additional sources of service information. First, DARS reports on the provision of similar benefits (but not timing or cost) for the Rehabilitation Service Administration RSA-911 Case Service Report due at the end of the federal fiscal year for all cases closed during that year. Use of this information is complicated by several factors, the most important being that the two indicators included for each service category sometimes provide inconsistent information. We impose the condition that this source identifies the provision of similar benefits only if both indicators designate service provision. Second, we observe data on in-house benefits provisions from the Woodrow Wilson Rehabilitation Center (WWRC), a state agency that provides comprehensive, individualized services with an employment objective. The WWRC receives an annual block grant from DARS which it administers autonomously. When appropriate, DARS refers individuals to the WWRC for rehabilitative services. The WWRC provided us with service information for this type of in-house benefit. Because there may be some classification errors between in-house services and similar benefits, we identify them simply as "non-purchased services." These two sources of information cover all non-purchased service expenses except for in-house counselor services. Thus, our data do not fully measure non-purchased *diagnosis* & *evaluation* services provided by counselors.

Although purchased services and in-house services provided by WWRC map uniquely into the six service categories used in our analysis, 4 of the 22 categories used for the RSA-911 do not. For example, the RSA category *diagnostic* & *treatment* includes both the *diagnosis* & *evaluation* category as well as the *restoration* category. Using *diagnostic* & *treatment* as an example, 6 of the 75 DARS purchased service categories map into *diagnosis* & *evaluation*, and 14 map into *restoration*. For the individuals flagged by RSA codes as having received *diagnosis* & *evaluation*, we count the number of sample individuals who received a service in one or more of the 6 *diagnosis* & *evaluation* purchased service codes (D) and the number of sample individuals in one or more of the 14 *restoration* codes (R). We then assign a probability that an individual designated in the RSA-911 file as receiving *diagnosis* & *evaluation* receives *diagnosis* & *evaluation* as $0.56 = D/(D+R)$ and restoration as $0.44 = R/(D+R)$.

Table A.2: Moments of Inverse Normal Transformed Office Effects

Service	Mean	Std Dev	Minimum	Maximum
Diagnosis & Evaluation	-0.396	0.264	-1.512	0.785
Training	-0.149	0.255	-1.133	0.707
Education	-1.200	0.388	-2.699	-0.189
Restoration	-0.719	0.463	-2.469	0.511
Maintenance	-0.502	0.410	-1.754	0.419
Other Service	-0.828	0.595	-2.668	0.405

Note: # Obs = 1489.

3 Counselor and Field Office Effects

We use as an instrument in equation (1) of the paper, a transformation of the proportion of other clients of the same counselor provided service j, i.e., a counselor effect. We also use a transformation of the proportion of other clients from the same office provided service j, i.e., an office effect. We transform the counselor and office effects using an inverse normal distribution function to make it more likely that, as the counselor and office effects vary, their effect on service probabilities can vary by approximately the same amount. To consider why this is attractive, consider a counselor who almost always uses a particular service. We want to allow for the possibility that this will imply that all of the clients of the counselor are very likely to receive that service. Limiting the counselor effects to vary between $(0, 1)$ makes it harder for that to occur. On the other hand, using an inverse distribution function for a distribution with the real line as support makes the range $(-\infty, \infty)$.

While such a transformation makes sense analytically, in practice, it might cause problems for values of the untransformed effect at or near the boundaries. We propose a "fix" that both makes sense and solves the boundary problem. In particular, we propose replacing the untransformed effect c_{ij} with

$$c^*_{ij} = (1 - \omega_i)\, c_{ij} + \omega_i \bar{c}_j \tag{1}$$

where $\bar{c}_j$ is the mean value of c_{ij} across all counselors (offices), $\omega_i = \kappa_i^{-1}$, and κ_i is the number of clients seen by counselor i (office i). This specification allows the counselor effect and office effect to be more important for those counselors (offices) who have many observed clients. In fact, it has a certain Bayesian flavor to it.

There are some respondents who either have missing counselor or office information or who have a counselor (or office) with no other clients. Because of such cases, we include a set of dummies for missing counselor and/or missing office effects.[1] It turns out that these dummies are very highly correlated, and the missing office effects must be excluded from the model to avoid a singular Hessian.

Tables A.2 and A.3 provide information about the moments of the transformed counselor and office effects. One can see that there is significant variation in both. There is some evidence of left-tailed skewness but no unreasonable outliers. The lack of outliers occurs despite zeroes for some services for some counselors and field offices because of the weighted average inherent in equation (12).

4 VEC Data Match

DARS provided the VEC with identifiers from the universe of 10323 applicants for DARS services in SFY 2000. The VEC returned to DARS a longitudinal file containing employment data for 9041 individuals having at least one quarter of "covered" employment during the 47-quarter period spanning July 1995 through March 2009, a "hit rate" of 88%. The remaining 12% in this cohort were either a) unemployed or out of the labor force for this entire interval or b) employed in jobs that are not covered by the VEC (e.g., were self-employed or worked out of state, for federal employers, for very small-sized firms, or at contingent-type jobs that do not provide benefits).

We explored the coverage issue through an arrangement with the Social Security Administration (SSA) whereby they matched VEC earnings (aggregated to a calendar year) to calendar-year SSA earnings for all SFY 2000 applicants.[2] Table

[1] In fact, when a counselor (office) has only one other client, we treat it as missing also.
[2] This analysis was not limited to applicants with mental illness diagnoses.

Table A.3: Moments of Normal Logistic Transformed Counselor Effects

Service	Mean	Std Dev	Minimum	Maximum
Diagnosis & Evaluation	-0.412	0.424	-2.061	1.045
Training	-0.173	0.513	-1.795	1.472
Education	-1.351	0.625	-2.542	0.66
Restoration	-0.805	0.615	-2.298	0.735
Maintenance	-0.549	0.564	-2.105	0.802
Other Service	-0.883	0.697	-2.303	1.054

Note: # Obs = 1485.

Table A.4: Comparison between SSA and VEC Employment Records

	2001	2002
Neither SSA nor VEC show earnings	31%	35%
SSA shows earnings, VEC does not	12%	12%
VEC shows earnings, SSA does not	1%	1%
Both SSA & VEC show earnings	57%	52%
Mean SSA Earnings	$9,117	$9,859
Mean SSA - VEC Difference	$510	$616

A.4 summarizes these results for the 9913 individuals with an identification match. For the two calendar years following SFY 2000 (the fiscal year of application), the SSA and VEC agreed on employment status for 87% of individuals. VEC records missed employment covered by SSA for 12% of the individuals in both 2001 and 2002. For those individuals where both SSA and VEC report earnings, VEC earnings levels fall short of SSA levels by 5.6% in 2001 and 6.2% in 2002. Although formally accounting for these coverage errors is beyond the scope of this paper, the results in Table A.4 suggests that any resulting biases should be minimal for the earnings equations but may be more important for the employment regressions. Unfortunately, our agreement with the SSA did not allow us to investigate whether these errors varied by VR service receipt.

5 Local Labor Market Conditions

Virginia is unique among states in that it has both counties and independent cities. While BEA provides data for almost all counties and independent cities, there is a small number of mostly rural counties for which BEA provides data only after some aggregation. We create 11 aggregated regions to deal with this problem listed in Table A.5.

We construct the employment rate by dividing number of people employed by working age population. We do this both at the county/independent city level and at the MSA level.[3] Significant variation in these measures exists across time, across geography, and across the two separate measures. One should note that there are some counties with employment rates greater than one. This occurs because the population numbers are based on county of residence while the employment numbers are based on county where one works. Thus, these rates reflect variation in net commuting patterns across counties.

6 Bias Caused by Unobserved Heterogeneity in Measured Mental Illness

There are many possibilities to explain the results with respect to the *diagnosis & evaluation* effects, including, but not limited to, the possibility that a) the instrument is correlated with the errors; b) the estimated net effect of *diagnosis & evaluation* on long-run outcomes is negative is a statistical anomaly and one might reject the (one-sided) null hypothesis that all of the long-run outcomes are positive; and c) extra diagnosis and evaluation requires much time for people with mental

[3] In the paper, we use only the county/independent city level because the two measures are very highly correlated.

Table A.5: Aggregated Regions

Region	Component Counties	Component Independent Cities
Eastern Shore	Accomack, Northhampton	
Rural Shenandoah, South	Bath, Highland, Rockbridge	Alleghany, Covington, Buena Vista, Lexington
Dinwiddie	Brunswick, Lunenberg, Nottoway	Dinwiddie, Colonial Heights, Petersburg, Greensville, Emporia
Bluefield, WV-VA Micropolitan SA	Bland, Buchanan, Dickerson, Lee, Norton, Smyth, Tazewell, Wythe, Wise	
Lynchburg Rural	Buckingham, Prince Edward	
Danville Rural	Charlotte, Halifax, Mecklenburg	
Northern Neck	Essex, King George, Lancaster, Middlesex, Northumberland, Richmond, Westmoreland	
Martinsville Rural	Floyd, Grayson, Patrick	Carroll, Galax
Culpeper, VA Micropolitan SA	Culpeper, Madison, Orange, Rappahannock	
Franklin/ Southhampton	Franklin, Southhampton	
Harrisonburg Rural	Page, Shenandoah	

illness and thus slows down the rehabilitation process (note that the negative effect is due solely to employment effects).The problem with (a) is that the result is specific to *diagnosis & evaluation*, and it disappears for other disability groups (e.g., Dean, et al., 2015a). The problem with (b) is that the null hypothesis would be rejected. We have no information on (c).

The bias explanation we prefer, which is also confirmed by DARS counselors, is the explanation included in the text of the paper. The idea is that, for people with mental illness, receipt of purchased diagnosis & evaluation services is an indicator that the individual's mental health problem is particularly difficult to deal with in a way unobserved in the DARS administrative data. This unobserved heterogeneity in mental illness is an error in measurement of an explanatory variable, and it causes *diagnosis & evaluation* to be correlated with the errors in the labor market outcomes equations. More explicitly, but in a simpler linear context, consider the model,

$$y_i = X_i\beta + w_{1i}\alpha_1 + w_{2i}\alpha_2 + u_i$$

where y_i is an outcome variable of interest for observation i, X_i is a vector of exogenous explanatory variables, w_{1i} is a potentially endogenous explanatory variable such as receipt of *diagnosis & evaluation*, and w_{2i} is another explanatory variable measured with error such as degree of mental illness with

$$plim\left(n^{-1}\sum_i w_{1i}w_{2i}\right) > 0.$$

In particular, for simplicity, assume that

$$w_{2i} \in \{0, 1, 2\},$$

but w_{2i} is not observed, and, instead,

$$x_i = 1\left(w_{2i} > 0\right)$$

is observed. Then the model to be estimated

$$y_i = X_i b + w_{1i}a_1 + x_i a_2 + v_i.$$

Let Z be a matrix of instruments. Then the asymptotic properties of the IV estimator are

$$\begin{aligned}
plim\begin{pmatrix}\widehat{b}\\ \widehat{a}_1\\ \widehat{a}_2\end{pmatrix} &= plim\begin{pmatrix}Z'X/n\\ Z'w_1/n\\ Z'x/n\end{pmatrix}^{-1} plim\left(Z'y/n\right)\\
&= plim\begin{pmatrix}Z'X/n\\ Z'w_1/n\\ Z'x/n\end{pmatrix}^{-1}\left[plim\begin{pmatrix}Z'X/n\\ Z'w_1/n\\ Z'w_2/n\end{pmatrix}\begin{pmatrix}\beta\\ \alpha_1\\ \alpha_2\end{pmatrix}+Z'u/n\right]\\
&= plim\begin{pmatrix}Z'X/n\\ Z'w_1/n\\ Z'x/n\end{pmatrix}^{-1} plim\begin{pmatrix}Z'X/n\\ Z'w_1/n\\ Z'w_2/n\end{pmatrix}\begin{pmatrix}\beta\\ \alpha_1\\ \alpha_2\end{pmatrix}\neq\begin{pmatrix}\beta\\ \alpha_1\\ \alpha_2\end{pmatrix}.
\end{aligned}$$

Now, in the interest of making more progress, consider a special case where

$$w_{1i}=\gamma_0+\gamma_1 w_{2i}+e_i.$$

Then

$$\begin{aligned}
plim\begin{pmatrix}\widehat{b}\\ \widehat{a}_1\\ \widehat{a}_2\end{pmatrix} &= = plim\begin{pmatrix}Z'X/n\\ Z'\left[\gamma_0 1+\gamma_1 w_2+e\right]/n\\ Z'\left[1\left(w_2>0\right)\right]/n\end{pmatrix}^{-1}\cdot\\
&\qquad plim\begin{pmatrix}Z'X/n\\ Z'\left[\gamma_0 1+\gamma_1 w_2+e\right]/n\\ Z'w_2/n\end{pmatrix}\begin{pmatrix}\beta\\ \alpha_1\\ \alpha_2\end{pmatrix}.
\end{aligned}$$

Next, in the same spirit, assume that $\beta=0$; i.e., there are no X's (without this assumption, as is the case in any measurement error problem, the sample correlation of X with w_1 and w_2 contaminates the analysis relative to the simpler case). Then

$$\begin{aligned}
plim\begin{pmatrix}\widehat{a}_1\\ \widehat{a}_2\end{pmatrix} &= plim\begin{pmatrix}z_1'\left[\gamma_0 1+\gamma_1 w_2+e\right]/n & z_2'\left[\gamma_0 1+\gamma_1 w_2+e\right]/n\\ z_1'x/n & z_2'x/n\end{pmatrix}^{-1}\cdot\\
&\qquad plim\begin{pmatrix}z_1'\left[\gamma_0 1+\gamma_1 w_2+e\right]/n & z_2'\left[\gamma_0 1+\gamma_1 w_2+e\right]/n\\ z_1'w_2/n & z_2'w_2/n\end{pmatrix}\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix}\\
&= \frac{plim\begin{pmatrix}z_2'x/n & -z_2'\left[\gamma_0 1+\gamma_1 w_2+e\right]/n\\ -z_1'x/n & z_1'\left[\gamma_0 1+\gamma_1 w_2+e\right]/n\end{pmatrix}}{plim\left[\left(\frac{z_1'\left[\gamma_0 1+\gamma_1 w_2+e\right]}{n}\right)\left(\frac{z_2'x}{n}\right)-\left(\frac{z_1'x}{n}\right)\left(\frac{z_2'\left[\gamma_0 1+\gamma_1 w_2+e\right]}{n}\right)\right]}\cdot\\
&\qquad plim\begin{pmatrix}z_1'\left[\gamma_0 1+\gamma_1 w_2+e\right]/n & z_2'\left[\gamma_0 1+\gamma_1 w_2+e\right]/n\\ z_1'w_2/n & z_2'w_2/n\end{pmatrix}\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix}\\
&= plim D^{-1} plim\begin{pmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{pmatrix}\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix}
\end{aligned}$$

where

$$\begin{aligned}
A_{11} &= \left(\frac{z_2'x}{n}\right)\left(\frac{z_1'\left[\gamma_0 1+\gamma_1 w_2+e\right]}{n}\right)-\left(\frac{z_1'w_2}{n}\right)\left(\frac{z_2'\left[\gamma_0 1+\gamma_1 w_2+e\right]}{n}\right),\\
A_{12} &= \left[\left(\frac{z_2'x}{n}\right)-\left(\frac{z_2'w_2}{n}\right)\right]\left(\frac{z_2'\left[\gamma_0 1+\gamma_1 w_2+e\right]}{n}\right),\\
A_{21} &= \left[\left(\frac{z_1'w_2}{n}\right)-\left(\frac{z_1'x}{n}\right)\right]\left(\frac{z_1'\left[\gamma_0 1+\gamma_1 w_2+e\right]}{n}\right),\\
A_{22} &= \left(\frac{z_2'w_2}{n}\right)\left(\frac{z_1'\left[\gamma_0 1+\gamma_1 w_2+e\right]}{n}\right)-\left(\frac{z_1'x}{n}\right)\left(\frac{z_2'\left[\gamma_0 1+\gamma_1 w_2+e\right]}{n}\right),\\
D &= \left[\left(\frac{z_2'x}{n}\right)\left(\frac{z_1'\left[\gamma_0 1+\gamma_1 w_2+e\right]}{n}\right)-\left(\frac{z_1'x}{n}\right)\left(\frac{z_2'\left[\gamma_0 1+\gamma_1 w_2+e\right]}{n}\right)\right].
\end{aligned}$$

Note that, in the case where $x = w_2$ (i.e., there is no measurement error),

$$plim\begin{pmatrix}\widehat{a}_1\\ \widehat{a}_2\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix} = \begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix}.$$

In general,

$$plim\begin{pmatrix}\widehat{a}_1\\ \widehat{a}_2\end{pmatrix} = \begin{pmatrix}1 & \eta_{12}\\ \eta_{21} & 1\end{pmatrix}\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix}$$

where

$$\begin{aligned}\eta_{12} &= \frac{plim\left[\left(\frac{z_2'x}{n}\right)-\left(\frac{z_2'w_2}{n}\right)\right]\left(\frac{z_2'[\gamma_0 1+\gamma_1 w_2+e]}{n}\right)}{plim\left[\left(\frac{z_2'x}{n}\right)\left(\frac{z_1'[\gamma_0 1+\gamma_1 w_2+e]}{n}\right)-\left(\frac{z_1'x}{n}\right)\left(\frac{z_2'[\gamma_0 1+\gamma_1 w_2+e]}{n}\right)\right]},\\ \eta_{21} &= \frac{plim\left[\left(\frac{z_1'w_2}{n}\right)-\left(\frac{z_1'x}{n}\right)\right]\left(\frac{z_1'[\gamma_0 1+\gamma_1 w_2+e]}{n}\right)}{plim\left[\left(\frac{z_2'x}{n}\right)\left(\frac{z_1'[\gamma_0 1+\gamma_1 w_2+e]}{n}\right)-\left(\frac{z_1'x}{n}\right)\left(\frac{z_2'[\gamma_0 1+\gamma_1 w_2+e]}{n}\right)\right]}.\end{aligned}$$

Without loss of generality, we can assume that

$$\begin{aligned}plim\left(\frac{z_1'1}{n}\right) &= plim\left(\frac{z_2'1}{n}\right) = plim\left(\frac{z_1'e}{n}\right) = plim\left(\frac{z_2'e}{n}\right) = 0;\\ plim\left(\frac{z_2'w_2}{n}\right) &> 0\end{aligned}$$

which implies that

$$\begin{aligned}\eta_{12} &= \frac{plim\left(\frac{z_2'(x-w_2)}{n}\right)\left(\frac{z_2'w_2}{n}\right)}{plim\left[\left(\frac{z_2'x}{n}\right)\left(\frac{z_1'w_2}{n}\right)-\left(\frac{z_1'x}{n}\right)\left(\frac{z_2'w_2}{n}\right)\right]};\\ \eta_{21} &= \frac{plim\left(\frac{z_1'(w_2-x)}{n}\right)\left(\frac{z_1'w_2}{n}\right)}{plim\left[\left(\frac{z_2'x}{n}\right)\left(\frac{z_1'w_2}{n}\right)-\left(\frac{z_1'x}{n}\right)\left(\frac{z_2'w_2}{n}\right)\right]}.\end{aligned}$$

If the denominator is positive (z_2 close to w_2 and z_1 close to w_1) and $plim\left(\frac{z_2'(x-w_2)}{n}\right)$ is a better instrument for w_2 than for x (note that these assumptions would all be true if ($z_2 = w_2$ or $z_2 = x$) and $z_1{=}w_1$), then

$$\eta_{12} < 0, \eta_{21} > 0,$$

which implies that

$$\begin{aligned}plim\widehat{a}_1 &< \alpha_1,\\ plim\widehat{a}_2 &> \alpha_2.\end{aligned}$$

In words, the estimate on diagnosis & evaluation would be negatively biased, and the estimate on *mental illness* or *SMI* would be biased upwards.

7 Nonstructural Model Estimates

Tables A.11 and A.12 display estimated parameters for the impact of services on employment (see Equation 2) using a probit model and log-earnings (see Equation 3) using a linear mean regression model, respectively. Table A.13 compares the estimated effects from these models to those found using the structural equations.

Table A.11: Probit Model Estimates of the DARS Purchased Survice Effects on Employment Propensity

Variable	Two or More Quarters Prior to Service	Quarters Prior Prior to Service Participation	First 2 Years After Service Participation	More than 2 Years After Service
Diagnosis & Evaluation	0.212 **	0.053	0.023	-0.120 **
	(0.015)	(0.061)	(0.022)	(0.013)
Training	-0.149 **	-0.070	0.190 **	0.070 **
	(0.019)	(0.079)	(0.027)	(0.015)
Education	0.172 **	0.104	-0.006	0.142 **
	(0.024)	(0.103)	(0.036)	(0.019)
Restoration	0.180 **	0.360 **	0.215 **	0.108 **
	(0.018)	(0.073)	(0.026)	(0.014)
Maintenance	-0.141 **	-0.148 *	-0.120 **	-0.191 **
	(0.020)	(0.084)	(0.029)	(0.016)
Other Services	0.040 *	-0.091	0.194 **	0.086 **
	(0.021)	(0.088)	(0.030)	(0.016)

Notes:
1. Standard errors are in parentheses.
2. Single-starred items are statistically significant at the 10% level, and double-starred items are statistically significant at the 5% level.

Table A.12: Linear (OLS) Model Estimates of the DARS Purchased Survice Effects on Log Quarterly Earnings

Variable	Two or More Quarters Prior to Service Participation	Quarters Prior Prior to Service Participation	First 2 Years After Service Participation	More than 2 Years After Service Participation
Diagnosis & Evaluation	-0.102 **	-0.480 **	-0.128 **	0.070 **
	(0.028)	(0.116)	(0.041)	(0.026)
Training	-0.014	-0.154	-0.212 **	-0.046
	(0.035)	(0.150)	(0.048)	(0.030)
Education	0.030	0.017	-0.082	0.196 **
	(0.042)	(0.183)	(0.062)	(0.035)
Restoration	-0.046	-0.098	0.111 **	0.236 **
	(0.032)	(0.129)	(0.045)	(0.029)
Maintenance	-0.304 **	0.058	-0.146 **	-0.089 **
	(0.037)	(0.156)	(0.052)	(0.032)
Other Services	-0.017	-0.244	0.107 **	0.199 **
	(0.037)	(0.159)	(0.051)	(0.031)

Notes:
1. Estimates are effects on log quarterly earnings conditional on employment.
2. Standard errors are in parentheses.
3. Single-starred items are statistically significant at the 10% level, and double-starred items are statistically significant at the 5% level.

Table A.13: Comparision of Linear/Probit Model Estimates to the Structural Model Estimates

	Employment			
	Probit Model		Structural Model	
	Short Run	Long Run	Short Run	Long Run
Diagnosis & Evaluation	-0.190	-0.332	-0.228	-0.462
Training	0.339	0.218	0.631	0.541
Education	-0.177	-0.030	-0.299	-0.113
Restoration	0.034	-0.072	-0.017	-0.127
Maintenance	0.021	-0.050	0.054	-0.074
Other Services	0.154	0.046	0.128	0.049
	Conditional Log Earnings			
	Linear Model		Structural Model	
	Short Run	Long Run	Short Run	Long Run
Diagnosis & Evaluation	-0.027	0.171	-0.085	0.032
Training	-0.198	-0.032	-0.055	0.136
Education	-0.112	0.166	-0.085	0.146
Restoration	0.158	0.282	0.092	0.206
Maintenance	0.158	0.215	0.106	0.217
Other Services	0.124	0.215	0.084	0.146

8 Estimates of the Impact of Covariates

Table A.6 displays the estimates of the effects of demographic characteristics on the propensity to use different services (y^*_{ijt} in equation 1). For the most part, the observed characteristics do not have statistically significant effects on service receipt, but there are some interesting exceptions. We find that clients with learning disabilities (0.680) and those receiving government assistance (0.491) are more likely to receive diagnosis & evaluation services. The probability of receiving training is higher for persons with government assistance (0.777) but lower for men (-0.315) and for those with musculoskeletal disabilities (-0.489) and/or substance abuse problems (-0.373). The receipt of education increases for those with more education (0.082) and for those with access to transportation and a driver's license (0.679). Interestingly, however, there is no statistically significant effect associated with having a serious mental illness or a significant disability.[4]

Table A.7 reports the effects of the demographic, socioeconomic, and disability-related characteristics on the three labor market outcomes of interest (z^*_{it} in equation 2), w_{it} in equation (3), and r^*_{it} in equation (4). For labor market outcomes, almost all of the estimates are statistically significant. Many of the estimates are as expected including positive effects of being white on *employment* propensity (0.157) and *log quarterly earnings* (0.362) as well as positive effects of *education* on *employment* propensity (0.024) and *log quarterly earnings* (0.053). The two transportation variables also have positive impacts on both labor market outcomes. The local labor market employment rate increases employment probabilities but decreases conditional earnings, suggesting that it might have been useful to include a measure of local wage rates. Some of the demographic and socioeconomic parameter estimates are counterintuitive. In particular, having a serious mental illness (SMI) increases *employment* propensity (0.194), receipt of special education services increases *log quarterly earnings* (0.259),[5] while being married decreases both *employment* propensity (-0.322) and *log quarterly earnings* (-0.138). The marriage effects can occur through income effects associated with having a spouse.

The diagnosis of a mental illness in the "base case" versus being initially diagnosed with mental illness in a subsequent application for VR services has a negative effect on *employment* propensity (-0.210) while increasing *log quarterly earnings* (0.399). Meanwhile, the disability severity-related variables have the expected signs, with negative effects of *significant* and *most significant* disabilities (relative to mild) on both labor market outcomes. Unlike its impact on service provision, the *SMI* estimates are explaining a significant amount of variation in labor market outcomes. *SMI*, by itself, increases employment (0.194) and increases *log quarterly earnings* (0.964). For males and whites, there are added interaction effects, all adversely affecting labor market outcomes. However, overall, the estimates with respect to *SMI* effects are hard to explain.[6] *Education* interacted with *SMI* has negative effects, and *age* interacted with *SMI* has mixed but statistically

[4] The *education missing* variable is statistically significantly negative across almost all services. It turns out that almost all of the individuals with *education missing* were closed during the application process. Thus, in an important sense, causation for this variable runs the other way.

[5] Special education programs have been found to improve schooling outcomes (Hanushek et al., 2002) and are associated with the use of supported employment services linked to higher earnings (Drake et al., 2009).

[6] The estimates imply that the average person in the sample has negative impacts of SMI. But the results are still problematic, for example,

Table A.6: Effects of Client Characteristics on Service Receipt by Type

Variable	Diagnosis & Evaluation	Training	Education
Constant	0.631	1.199 *	-1.183
Male	-0.102	-0.315 *	-0.118
White	-0.056	-0.077	-0.251
Education	-0.011	-0.005	0.082 **
Special Education	0.458	-0.529	0.125
Education Missing	-1.008 **	-4.000 +	-2.583 *
Age/100	0.143	-0.247	-0.110
Married	-0.145	-0.299	-0.019
# Dependents	-0.044	-0.198 **	-0.095
Transportation Available	0.182	0.152	0.493 *
Has Driving License	0.231	-0.252	0.679 **
Receives Govt Assistance	0.491 **	0.777 **	0.313
Musculoskeletal Disability	0.241	-0.489 **	0.009
Learning Disability	0.680 **	0.074	-0.628
Mental Illness	-0.640 **	-0.976 **	-0.464
Substance Abuse	0.027	-0.373 *	0.265
Disability Significant	0.350 *	0.136	-0.355
Disability Most Significant	0.466 *	0.502	-0.406
SMI	0.028	0.714	0.350
Male * SMI	-0.146	0.041	-0.664
White * SMI	-0.081	0.153	0.448
Education * SMI	0.005	-0.063	0.008
Age/100 * SMI	-0.062	0.316	-0.443

Table A.6 (continued)

Variable	Restoration	Maintenance	Other Services
Constant	0.860 *	1.672 **	0.113
Male	-0.301 *	0.038	-0.095
White	-0.155	-0.393 *	-0.283
Education	-0.008	-0.032	0.011
Special Education	-0.020	-0.921 *	-0.124
Education Missing	-1.386 **	-4.000 +	-2.429 **
Age/100	-0.039	-0.386 *	0.010
Married	-0.254	-0.405 *	-0.235
# Dependents	0.095	-0.015	-0.134 *
Transportation Available	-0.066	-0.149	0.382 *
Has Driving License	0.142	-0.337 *	0.058
Receives Govt Assistance	-0.119	0.378	0.101
Musculo/Skeletal Disability	-0.134	0.128	-0.004
Learning Disability	2.071 **	0.036	0.556 *
Mental Illness	-0.432	-0.534	-0.592 *
Substance Abuse	-0.103	0.329 *	-0.192
Disability Significant	0.095	0.022	0.004
Disability Most Significant	0.335	0.396	0.172
SMI	-1.441	-0.558	0.469
Male * SMI	0.228	-0.277	0.042
White * SMI	0.030	0.231	0.088
Education * SMI	0.028	-0.047	-0.031
Age/100 * SMI	-0.123	0.818 *	-0.053

Notes:

1.Standard errors not presented to save space but are available from the corresponding author.

2.Single-starred items are statistically significant at the 10% level, and double-starred items are statistically significant at the 5% level, and items with + were restricted.

significant effects on outcomes. Baldwin (2005) estimates the effect of mood disorder, anxiety disorder, and adjustment disorder on employment probabilities and finds an average reduction in employment probability on the order of 0.3. Our estimates imply smaller effects, at least for significant mental health problems similar to those considered by Baldwin. A big part of the reason for this is probably that our sample consists only of people who have been identified as having a mental health problem while Baldwin (2005) uses the SIPP sample.

For DI/SSI receipt, almost all of the effects are statistically significant and with expected signs, also. For example, the probability a client takes-up DI/SSI is estimated to decrease for *white* (-0.114), *education* (-0.044), and *transportation available* (-0.068). Surprises are *married* (0.484), *learning disability* (-0.166), and *local employment rate* (0.050).

So far all of the discussion has concerned the effect of purchased services on labor market outcomes. In fact, DARS also provides some services in-house, and other services sometimes are paid for by other organizations, and, as discussed in Section III.A.2, we have some information about those other services. Using this data, we allow the effects of covariates on the receipt of such services to be proportionate to their effect for service choice in equation (1) and their effect for employment propensity in equation (2) as reported in Table 4, for conditional log quarterly earnings in equation (3) as reported in Table 5, and for DI/SSI receipt in equation (4) as reported in Table 6. The estimated proportion for service choice propensity is 0.824^{**} (0.280) which is not significantly different from 1.0. Thus, decisions about using non-purchased services are similar to those for purchased services. By contrast, the estimated proportion for employment propensity, conditional log quarterly earnings, and DI/SSI receipt propensity is 0.432^{**} (0.033) which is significantly different from 1.0. Thus, the effect of non-purchased services on labor market outcomes and DI/SSI receipt is 43.2% of that for purchased services.

9 Smoothed Sample DI/SSI Probabilities

Smoothed sample DI/SSI probabilities conditional on predicted probabilities are displayed in Figure 6. As was similar for employment, the predicted probabilities fit pretty well except for values above 0.7.

for black women with *SMI*.

Table A.7: Labor Market and DI/SSI Effects

Variable	Employment Estimate	Employment Std Err	Log Quarterly Earnings Estimate	Log Quarterly Earnings Std Err	DI/SSI Receipt Estimate	DI/SSI Receipt Std Err
Constant	0.052 **	0.021	5.311 **	0.032	-3.338 **	0.026
Male	0.024 **	0.006	0.360 **	0.008	-0.036 **	0.007
White	0.157 **	0.007	0.362 **	0.009	-0.114 **	0.009
Education	0.024 **	0.001	0.053 **	0.001	-0.044 **	0.001
Special Education	-0.014	0.027	0.259 **	0.032	-0.019	0.024
Education Missing	0.423 **	0.013	0.626 **	0.020	-0.438 **	0.015
Age/100	-0.701 **	0.007	0.043 **	0.009	1.358 **	0.010
Married	-0.322 **	0.007	-0.138 **	0.009	0.484 **	0.009
# Dependents	0.061 **	0.002	0.064 **	0.003	-0.063 **	0.003
Transportation Available	0.122 **	0.007	0.089 **	0.009	-0.068 **	0.007
Has Driving License	0.213 **	0.007	0.333 **	0.010	-0.175 **	0.007
Receives Govt Assistance	-0.332 **	0.008	-0.237 **	0.012	1.753 **	0.009
Musculoskeletal Disability	0.097 **	0.007	0.130 **	0.009	0.036 **	0.009
Learning Disability	0.406 **	0.012	0.225 **	0.015	-0.166 **	0.015
Mental Illness	-0.210 **	0.012	0.399 **	0.018	-0.016	0.013
Substance Abuse	0.213 **	0.007	0.086 **	0.010	-0.366 **	0.009
Disability Significant	-0.027 **	0.009	-0.170 **	0.012	0.664 **	0.011
Disability Most Significant	-0.112 **	0.010	-0.280 **	0.013	1.147 **	0.012
SMI	0.194 **	0.028	0.964 **	0.038	1.083 **	0.032
Male * SMI	-0.094 **	0.012	-0.521 **	0.017	0.513 **	0.014
White * SMI	-0.535 **	0.012	-0.482 **	0.018	0.815 **	0.014
Education * SMI	-0.019 **	0.001	-0.047 **	0.002	0.044 **	0.001
Age/100 * SMI	0.236 **	0.015	-0.083 **	0.021	-1.249 **	0.016
Local Employment Rate	0.185 **	0.067	-0.185 **	0.068	0.050 **	0.011

Note: Single-starred items are statistically significant at the 10% level, and double-starred items are statistically significant at the 5% level.

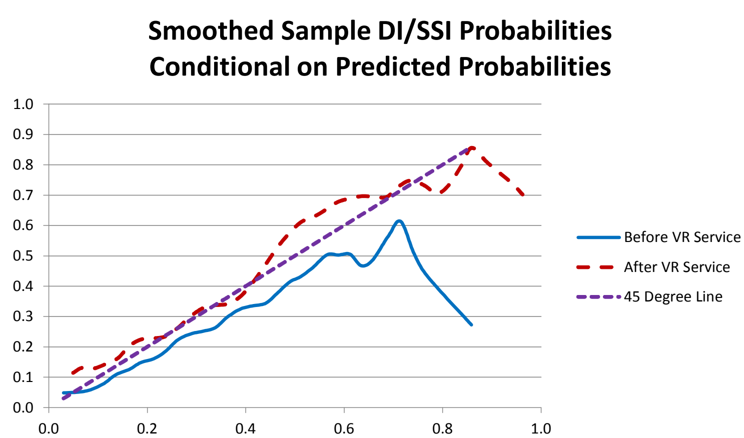

Figure 1: Smoothed Sample DI/SSI Probabilities Conditional on Predicted Probabilities

10 Application Test

In this Appendix, we present a test for whether the decision to apply for VR services is exogenous. We begin by formalizing the test, then describe the data, and conclude by summarizing the test results.

10.1 Methodology

Let

$$p(x_i, \gamma) = \Phi(x_i \gamma) \tag{2}$$

be the probability that a disabled person i applies for DARS services where $x_i \in X$ is a set of exogenous variables. Divide X into subsets, indexed by k. Define $P_k = P(X_k)$ to be the proportion of the population belonging to subset k, and let $\widehat{P}_k$ be a consistent estimate of P_k from some sample. Johnson et al. (2015) provides such an estimate for "composite" CSBs based on work in Stern (2014). Then, the expected number of applicants from subset k is

$$N_k = M \int_{x \in X_k} p(x, \gamma) \, dF_x(x)$$

where M is the population of disabled people and $F_x(x)$ is the distribution of x.

Assume we have a consistent estimate of $M(x) = M f_x(x)$ denoted as $\widehat{M}(x) \; \forall x \in X$. Then we can write the expected number of DARS applicants belonging to subset k as

$$N_k(\gamma) = \int_{x \in X_k} p(x, \gamma) \, M_x(x) \, dx$$

and approximate it as

$$\widehat{N}_k(\gamma) = \int_{x \in X_k} p(x, \gamma) \, \widehat{M}_x(x) \, dx.$$

In the data, we observe the actual number of people applying to DARS from subset k which we call n_k. Then, a MOM estimate of γ is

$$\widehat{\gamma} = \min_{\gamma} \sum_k \omega_k \left[n_k - \widehat{N}_k(\gamma) \right]^2 \tag{3}$$

where ω_k is an appropriately chosen weight.

Let e_i be the error in the latent variable equation corresponding to equation (13). Note that the generalized residual for $e_i \mid apply$ is

$$\widehat{e}_i = E(e_i \mid apply) = \frac{\int_{e > -x_i \gamma} e dF_e(e)}{\int_{e > -x_i \gamma} dF_e(e)} = \begin{cases} \frac{\phi(x_i \gamma)}{\Phi(x_i \gamma)} & \text{if apply} \\ \frac{-\phi(x_i \gamma)}{1 - \Phi(x_i \gamma)} & \text{if not apply} \end{cases}.$$

Next, let $L_i(u_i)$ be the likelihood function associated with person i, conditional on error u_i and application, and let $\widehat{u}_i$ be the generalized residual for person i. We can test for selection effects into the sample by constructing[7]

$$\widehat{\rho}_{eu} = \frac{n^{-1} \sum_i \widehat{e}_i \widehat{u}_i}{\sqrt{(n^{-1} \sum_i \widehat{e}_i^2)(n^{-1} \sum_i \widehat{u}_i^2)}}.$$

Under H_0, $plim \widehat{\rho}_{eu} = 0$ and

$$\sqrt{n} \left(\widehat{\rho}_{eu} - 0 \right) \sim N(0, 1).$$

10.1.1 ACS Data

Data on Virginia residents from the 2012 American Community Survey (ACS) are used to estimate the model described in equation (13) (N = 64009; see Johnson et. al. (2015) for details).[8] Moments of the data are presented in Table A.8. The surprising number is the large mean family income of \$84.95K. This is caused by a fat right tail; the median family income is only \$62.5K.

[7] $\widehat{u}_i$ is a vector, so $\widehat{\rho}_{eu}$ is as well. Thus, we need to adjust the equation to include the inverse covariance matrix of $\widehat{u}_i$. However, here we consider the case where $\widehat{u}_i$ is a scalar to simplify discussion.

[8] We use the same 2012 data as Johnson et al. (2015) despite our DARS data coming from 2000 because the earliest ACS data with the necessary variables to predict mental health with any precision is from 2008, and the difference between 2008 and 2012 in the distribution of demographic characteristics was not large.

Table A.8: Weighted Moments of Explanatory Variables from ACS Used in Estimation

Variable	Mean	Std Dev	Variable	Mean	Std Dev
Female	0.516	0.500	High School Diploma	0.489	0.500
Age/100	0.084	1.056	College Degree	0.394	0.489
Black	0.185	0.389	Family Income	84.950	85.216
Other Race	0.077	0.267	Dummy: Fam Inc > $50K	0.585	0.493
Hispanic	0.073	0.260	In MSA	0.711	0.453
Married	0.550	0.497	Health Condition	0.086	0.280
Divorced	0.108	0.310	# ADLs	0.030	0.170
Widowed	0.058	0.234	# IADLs	0.055	0.228
Veteran	0.116	0.320	Functional Limitation	0.077	0.266

Notes

1) Sample size is 59207.

2) Sample weights are used to replicate the population joint density of variables.

3) Median family income = $62.5K.

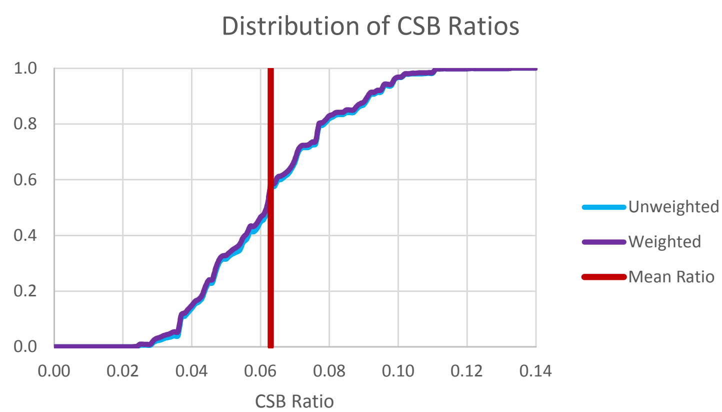

Figure 2: Distribution of CSB Ratios

The state is divided up into 40 public mental health care regions called community service boards (CSBs) (Johnson, et al., 2015). Our estimation strategy relies on the existence of significant variation in the ratios of DARS applicants to people with mental health problems across the CSBs. Figure 2 presents the distribution of ratios. One can see immediately that there is significant variation and that the distribution is continuous and well behaved.

10.2 Estimation Results

Table A.9 shows the results of the estimation process associated with equation (14) Each of the estimated coefficients should be interpreted as $\partial p^* (x_i, \gamma) / \partial x_i$ for each of the variables in the table where $p^* (x_i, \gamma)$ is the latent variable associated with $p (x_i, \gamma)$; i.e., $p (x_i, \gamma) = \Pr [p^* (x_i, \gamma) > 0]$. The estimates in Table A.9 are interesting in their own right in that they provide information about how the propensity to apply for DARS services varies with individual characteristics. Women are more likely to apply than men (0.285), and whites are more likely to apply than blacks (-0.252) or Hispanics (-0.237). Veterans (0.107) have higher application rates, and application rates decline with education (-0.074, -0.141). Health conditions have large effects on application rates as people in poor health (1.445), with health conditions (0.525), or with functional limitations (0.488) are significantly more likely to apply for DARS services.

The other feature of this estimation process is the extremely high proportion of the data variation explained by the variables in Table A.9. The weighted sum of squared residuals when $\gamma = 0$ is 3641.2, and, at $\widehat{\gamma}$ (reported in Table A.9), it is 0.07. This implies that there is very little room for unobserved characteristics, possibly correlated with errors in the model described in the text, to exist. More precisely, there may be other variation v_{ki} across observations i with each homogenous

Table A.9: Estimates of Determinants of Walking in the Front Door

Variable	Estimate	Std Error	Variable	Estimate	Std Error
Constant	-1.152 **	0.437	High School Diploma	-0.074 **	0.000
Female	0.285 **	0.000	College Degree	-0.141 **	0.001
Age/100	-0.148 **	0.000	Family Income/1K	0.000 **	0.000
(Age/100)**2	-0.089 **	0.000	Dummy: Big Fam Inc	-0.292 **	0.001
(Age/100)**3	0.069 **	0.000	Health Fair	-0.026 **	0.003
Black	-0.252 **	0.001	Health Poor	1.445 **	0.005
Other Race	-0.023 **	0.001	Weight/Height	-0.215	0.178
Hispanic	-0.237 **	0.002	In MSA	0.090 **	0.000
Married	-0.213 **	0.001	Health Condition	0.525 **	0.000
Divorced	0.126 **	0.001	# ADLs	0.148 **	0.000
Widowed	-0.022 **	0.001	# IADLs	0.320 **	0.000
Veteran	0.107 **	0.001	Functional Limitation	0.488 **	0.001

Notes
1) Age variables are transformed into 1st-, 2nd-, and 3rd-order orthogonal polynomials.
2) Double-starred items are statistically significant at the 5% level.

cell k used in equation (14) that average out over the cell and therefore have no effect on estimation. However, if one thinks that the $N_k(\gamma)$ people from cell k applying to DARS are those with the greatest (latent) value of participating, then the standard deviation of the aggregated error (across the cell) is the standard deviation of sample average of the first $N_k(\gamma)$ sample order statistics of the cell;

$$StDev = \sqrt{\frac{1}{N_k(\gamma)} \sum_{(i)\in k} \left(\upsilon_{k(i)} - \overline{\upsilon}_k\right)^2} \tag{4}$$

where (i) is the υ_{ki} term for the ith largest latent value and $\overline{\upsilon}_k$ is the sample average of the υ_{ki}'s associated with the $N_k(\gamma)$ largest latent values. As an illustration, consider the case where $\upsilon_{ki} \sim iidN(0,1)$ and $\gamma = 0$. Then, the simulated[9] standard deviation in equation (15) is presented in Figure 8. The bars for "proportion applying is 1.00" is the standard deviation of the mean: $1/M_k$ where

$$M_k = \int_{x\in X_k} M(x)\, dx$$

is the number of disabled people in cell k. The bars for other proportions are for different values of $N_k(\gamma)/M_k$. As seen in Figure 7, the median proportion is approximately 0.06. Thus we simulate results for 0.04, 0.06, and 0.08. Heterogeneity $(\gamma \neq 0)$[10] changes the sample of applicants from those with the highest value of υ_{ki} to those with the highest value of the latent variable. The results in Figure 8 show that the standard deviation of the mean value of υ_{ki} among those who apply for DARS services decreases with sample size M_k, decreases with the proportion applying, and increases with heterogeneity. Compared to a random sample of 10000 disabled people (proportion applying = 1.00) with or without heterogeneity, a sample of 10000 with 6% applying and with heterogeneity has a standard deviation of υ_{ki} among those applying 2.5 times larger $(0.01 \Rightarrow 0.025)$. If the cell sample size is 5000, then the ratio increases to 2.8 $(0.014 \Rightarrow .040)$. Thus, the standard deviation of υ_{ki} consistent with our results based on cell-mean application rates would have to be on the order of 2.8 times smaller than if all disabled people in the cell applied.

10.3 Test Results

Given the estimates in Table A.9, we can construct generalized residuals and use them in the proposed test statistic described in Section 10.1. The test statistics and critical values are reported in Table A.10. Note the range of test statistics associated with service choices is $(0.006, 0.076)$, and the range for labor market and SSI/DI outcomes is $(0.018, 0.051)$, all which are very small. All of the test statistics fall between the 2.5% and 97.5% critical values. Thus, there is no evidence of selection on unobservables caused by the DARS application decision.

[9] As can be seen in, e.g., Headrick and Pant (2012), it is quite difficult to analytically evaluate the covariance matrix of multiple order statistics, especially for large sample sizes. Our need for the standard deviation of the mean error conditional on application can be computed using formulae for moments of truncated normal variables available, for example, in Heckman (1979). However, it is completely straightforward and inexpensive (CPU time) to simulate them.

[10] Heterogeneity is simulated for this example as $x = 3U(0,1)$.

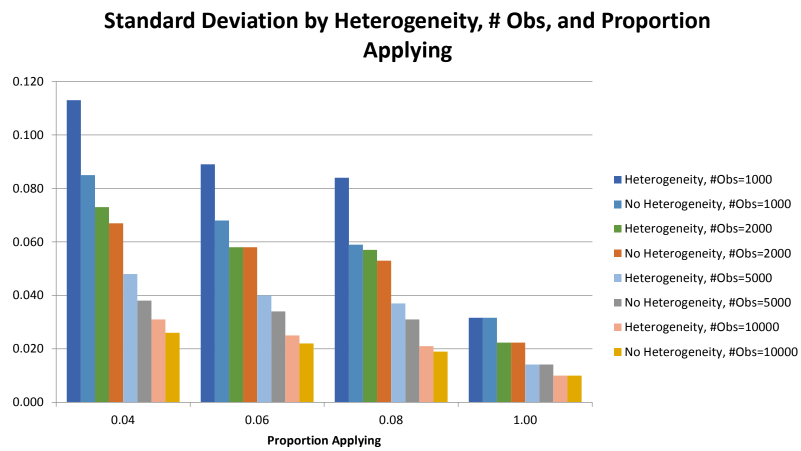

Figure 3: Standard Deviation by Heterogeneity, # Obs, and Proportion Applying

Table A.10: Test Statistics and Statistical Significance

	Test Statistic	p-Value	.025 Critical Value	.975 Critical Value
Diagnosis & Evaluation	0.033	0.315	-0.137	0.115
Training	0.006	0.063	-0.034	0.215
Education	0.024	0.350	-0.140	0.115
Restoration	0.076	0.102	-0.135	0.115
Maintenance	0.039	0.262	-0.140	0.115
Other Service	0.060	0.157	-0.140	0.115
Employment Prior to Application	0.033	0.307	-0.140	0.115
Short-Run Employment After Application	0.018	0.389	-0.140	0.115
Long-Run Employment After Application	0.043	0.258	-0.140	0.110
SSI/DI Receipt Prior to Application	0.051	0.210	-0.130	0.120
Short-Run SSI/DI Receipt After Application	0.046	0.239	-0.135	0.120
Long-Run SSI/DI Receipt After Application	0.047	0.332	-0.140	0.120
log Conditional Earnings Prior to Application	0.046	0.258	-0.145	0.130
Short-Run log Conditional Earnings After Application	0.030	0.318	-0.140	0.130
Long-Run log Conditional Earnings After Application	0.037	0.292	-0.140	0.125

References

[1] Aakvik, A., J. Heckman, and E. 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.

[2] Ashenfelter, Orley (1978). “Estimating the Effect of Training Programs on Earnings.” *Review of Economics and Statistics*. 60: 47-57.

[3] Autor, David and Mark Duggan (2010). *Supporting Work: A Proposal for Modernizing the U.S. Disability Insurance System.* Hamilton Project.

[4] Baldwin, Marjorie (1999). “The Effects of Impairments on Employment and Wages: Estimates from the 1984 and 1990 SIPP.” *Behavioral Sciences and the Law*. 17: 7-27.

[5] Baldwin, Marjorie (2005). “Explaining the Differences in Employment Outcomes Between Persons With and Without Mental Disorders.” Unpublished manuscript.

[6] Bellante, D. (1972). “A Multivariate Analysis of a Vocational Rehabilitation Program.” *Journal of Human Resources*. 7: 226-241.

[7] Benitez-Silva, Hugo, Moshe Buchinsky, Hiu Man Chan, John Rust, and Sofia Sheidvasser. (1999). “An Empirical Analysis of the Social Security Application, Appeal and Award Process.” *Labour Economics*. 6: 147-178.

[8] Börsch-Supan, Axel and Vassilis Hajivassiliou (1992). “Health, Children, and Elderly Living Arrangements: A Multiperiod-Multinomial Probit Model with Unobserved Heterogeneity and Autocorrelated Errors.” *Topics in the Economics of Aging*. (ed.) David Wise. Chicago and London: U. of Chicago Press, 79-104.

[9] Bratberg, E., A. Grasdal, and A. Risa. (2002). “Evaluating Social Policy by Experimental and Nonexperimental Methods.” *Scandinavian Journal of Economics*. 104(1): 147-171.

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

[11] Bureau of Economic Analysis (2010b). http://www.bea.gov/regional/ docs/footnotes.cfm?tablename=CA25.

[12] Cimera, R. (2010). “The National Cost-Efficiency of Supported Employees with Intellectual Disabilities: 2002 to 2007.” *American Journal on Intellectual and Developmental Disabilities*. 115: 19-29.

[13] Conley, R. (1969). “A Benefit Cost Analysis of the Vocational Rehabilitation Program." *Journal of Human Resources*. 4: 226-252.

[14] Conti, G., J. Heckman, and S. Urzúa (2010). “The Education-Health Gradient.” *American Economic Review: Papers and Proceedings*. 100: 1–5.

[15] Dean, David and Robert Dolan (1991). “Fixed-Effects Estimates of Earnings Impacts for the Vocational Rehabilitation Program,” *Journal of Human Resources*. 26(2): 380-391.

[16] Dean, David, Robert Dolan, and Robert Schmidt (1999). “Evaluating the Vocational Rehabilitation Program Using Longitudinal Data: Evidence for a Quasi-experimental Research Design.” *Evaluation Review*. 23(2): 162-189.

[17] Dean, David, Robert Dolan, Robert Schmidt, Paul Wehman, John Kregel, and Grant Revell (2002). “A Paradigm for Evaluation of the Federal-State Vocational Rehabilitation Program.” *Achievements and Challenges in Employment Services for People with Disabilities: The Longitudinal Impact of Workplace Supports*. (eds.) John Kregel, David Dean, and Paul Wehman. Richmond: Virginia Commonwealth University Rehabilitation Research and Training Center on Workplace Supports.

[18] Dean, David, John Pepper, Robert Schmidt, and Steven Stern (2014). “State Vocational Rehabilitation Programs and Federal Disability Insurance: An Analysis of Virginia's Vocational Rehabilitation Program.” *IZA Journal of Labor Policy*. 3(7): 1-19.

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

[20] Dean, David, John Pepper, Robert Schmidt, and Steven Stern (2015b). "The Effects of Youth Transition Programs on Labor Market Outcomes." Unpublished manuscript.

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

[22] Drake, Robert, Jonathan Skinner, Gary Bond, and Howard Goldman (2009). "Social Security and Mental Illness: Reducing Disability With Supported Employment." *Health Affairs*. 28(3): 761-770.

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

[24] French, Eric and Jae Song. (2014). "The Effect of Disability Insurance Receipt on Labor Supply: A Dynamic Analysis." *American Economic Journal: Economic Policy*. 6(2): 291-337.

[25] Friedlander, Daniel, David Greenberg, and Philip Robins (1997). "Evaluating Government Training Programs for the Economically Disadvantaged." *Journal of Economic Literature*. 35(4): 1809-1855.

[26] Frolich, M., A. Heshmati, and M. Lechner (2004). "A Microeconometric Evaluation of Rehabilitation of Long-Term Sickness in Sweden." *Journal of Econometrics*. 19: 375–396.

[27] Geweke, John (1988). "Antithetic Acceleration of Monte Carlo Integration in Bayesian Inference." *Journal of Econometrics*. 38(1/2): 73-89.

[28] Gilleskie, D. and D. Hoffman.(2014). "The Impact of Human and Health Capital on Wages: An Examination of Occupation, Employer, and Disability Transitions." *Journal of Human Capital.* 8(3): 235-379.

[29] 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.

[30] Headrick, Todd and Mohan Pant (2012). "On the Order Statistics of Standard Normal-Based Power Method Distributions." *ISRN Applied Mathematics*. 2012: Article ID 945627.

[31] Heckman, James (1979). "Sample Selection Bias as a Specification Error." *Econometrica*. 47(1): 153-162.

[32] Heckman, J. and V. Hotz (1989). "Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training." *Journal of the American Statistical Association*. 84: 862-874.

[33] 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.

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

[35] Heckman, J., J. Stixrud, and S. Urzúa (2006). "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics*. 24: 411–482.

[36] Heminway, Derek and Faranak Rohani (1999). *A Cost-Benefit Analysis of the Employment of People with Disabilities in Florida: Final Report*. Educational Services Program.

[37] Hollenbeck, Kevin and Wei-Jang Huang (2006). *Net Impact and Benefit-Cost Estimates of the Workforce Development System in Washington State*. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

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

[39] Johnson, W., M. LaForest, B. Lissenden, and S. Stern (2015). "Variation in Mental Health Illness and Provision of Public Services in Virginia." Unpublished manuscript.

[40] Kessler, R., P. Berglund, M. Bruce, J. Koch, E. Laska, P. Leaf, R. Manderscheid, R. Rosenheck, E. Walters, and P. Wang (2001). "The Prevalence and Correlates of Untreated Serious Mental Illness." *Health Services Research*. 36(6 Pt 1): 987-1007.

[41] Kisker, E., G. Strech, J. Vetter, and C. Foote (2008). *Evaluating Rehabilitation Services in Oklahoma: An Analysis of Program Impacts and of Benefit/Costs*. Oklahoma Department of Rehabilitative Services.

[42] Kreider, Brent (1999). "Social Security Disability Insurance: Applications, Awards, and Lifetime Income Flows." *Journal of Labor Economics*. 17(4): 784-827.

[43] Kreider, B. and R. Riphahn (2000). "Explaining Applications to the U.S. Disability System: A Semiparametric Approach." *Journal of Human Resources*. 35(1): 82-115.

[44] LaLonde, Robert (1995). "The Promise of Public Sector-Sponsored Training Programs." *Journal of Economic Perspectives*. 9(2): 149-168.

[45] Loprest, P. (2007). *Strategic Assessment of the State of the Science in Research on Employment for Individuals with Disabilities*. National Institute for Disability and Rehabilitation Research. Contract GS23F8198H.

[46] 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.

[47] Marcotte D., V. Wilxox-Gok, and D. Redmond (2000). "The Labor Market Effects of Mental Illness: The Case of Affective Disorders." *The Economics of Disability*. (eds.) D. Salkever and A. Sorkin. Greenwich, CT: JAI Press.

[48] McFadden, Daniel and Kenneth Train (2000). "Mixed MNL Models for Discrete Response." *Journal of Applied Econometrics*. 15(5): 447-470.

[49] Meyer, Bruce (1995). "Natural and Quasi-Experiments in Economics." *Journal of Business and Economic Statistics*. 13(2): 151–162.

[50] Mueser, K., M. Salyers, and P. Mueser (2001). "A Prospective Analysis of Work in Schizophrenia." *Schizophrenia Bulletin*. 27: 281-296.

[51] Mueser, Peter, Kenneth Troske, and Alexey Gorislavky (2007). "Using State Administrative Data to Measure Program Performance." *Review of Economics and Statistics*. 89(4): 761-783.

[52] Raaum, O. and H. Torp (2001). "Labor Market Training in Norway – Effects on Earnings." *Labour Economics*. 8(5): 207-247.

[53] Rogers, J., M. Bishop, and M. Crystal (2005). "Predicting Rehabilitation Outcome for Supplemental Security Income and Social Security Disability Insurance Recipients: Implications for Consideration with the Ticket to Work Program." *Journal of Rehabilitation*. 71(3): 5-10.

[54] Stapleton, David and Frank Martin (2012). "Vocational Rehabilitation on the Road to Social Security Disability: Longitudinal Statistics from Matched Administrative Data." Unpublished manuscript.

[55] Stern, Steven (1992). "A Method for Smoothing Simulated Moments of Discrete Probabilities in Multinomial Probit Models." *Econometrica*. 60(4): 943-952.

[56] Stern, Steven (1997). "Simulation-Based Estimation." *Journal of Economic Literature*. 35(4): 2006-2039.

[57] Stern, Steven (2014). "Estimating Local Prevalence of Mental Health Problems." *Health Services and Outcomes Research Methodology*. 14: 109-155.

[58] U.S. Dept. of Health and Human Services, Substance Abuse and Mental Health Services Administration (1999). "Estimation Methodology for Adults With Serious Mental Illness (SMI)." *Federal Register*. 64(121): 33890-33897.

[59] U.S. Dept. of Health and Human Services, National Center for Health Statistics (2004). *National Health Interview Survey*.

[60] U.S. Government Accountability Office (2005) “Vocational Rehabilitation: Better Measures and Monitoring Could Improve the Performance of the VR Program.” Washington, D.C.

[61] U.S. Government Accountability Office. (June 2012). Employment for people with disabilities: Little is known about the effectiveness of fragmented and overlapping programs. (GAO-12-677). Washington, D.C.

[62] Uvin, J., D. Karaaslani, and G. White (2004). *Evaluation of Massachusetts' Public Vocational Rehabilitation Program: Final Report*. Massachusetts Rehabilitation Commission.

[63] Wilhelm, S. and J. Robinson (2010). *Utah State Office of Rehabilitation Economic Impact Study*. The University of Utah Center for Public Policy and Administration.

[64] Worrall, John (1978). “A Benefit-Cost Analysis of the Vocational Rehabilitation Program.” *Journal of Human Resources*. 13: 285-298.