Applications of the VR-ROI Project: ROI Estimates for Virginia and Maryland

Robert M. Schmidt
University of Richmond

Christopher M. Clapp
University of Chicago

John V. Pepper
University of Virginia

Steven N. Stern
Stony Brook University

September 2019

The following manuscript was published as:

Schmidt, R. M., Clapp, C. M., Pepper, J. V., & Stern, S. (2019). Applications of the VR-ROI project: ROI estimates for Virginia and Maryland. *The Journal of Rehabilitation Administration*, 40(1), 57-72.

Abstract

This paper briefly describes and then implements the VR-ROI (Vocational Rehabilitation Return on Investment) Project's model for applicants in state fiscal year 2007 to Virginia's Department for Aging and Rehabilitative Services (VA DARS) and to Maryland's Division of Rehabilitation Services (MD DORS). We present results that account for differences across disability, agency, VR service type and source, applicant characteristics, and county as well as national economic conditions. This approach provides a rich set of estimates that display considerable heterogeneity within each agency across three disability types (mental illness, physical impairment, and cognitive impairment) and seven service categories (*diagnosis & evaluation*, *training*, *education*, *restoration*, *maintenance*, *job placement*, *job supports*) within each disability type. Five online appendices providing additional detail on the model and results can be found at https://scholarship.richmond.edu/economics-faculty-publications/56/.

Keywords: Return on investment estimates, vocational rehabilitation (VR), mental illness, physical impairment, cognitive impairment, multiple service types, multiple service sources.

This article provides an overview of the VR-ROI (Vocational Rehabilitation Return on Investment) Project's statistical model and summarizes new VR-ROI estimates for two agencies: the Virginia Department for Aging and Rehabilitative Services (VA DARS) and the Maryland Division of Rehabilitation Services (MD DORS). Focusing on applicants during State Fiscal Year 2007 (SFY 2007), we provide six separate sets of estimates. We provide estimates by agency for each of three disability types – mental illness (anxiety disorders, depressive and other mood disorders, personality disorders, schizophrenia and other mood disorders), physical impairment (internal and musculoskeletal), and cognitive impairment (intellectual disability and learning disability).

Prior to presenting those results, we describe various aspects of the six cohorts and provide an overview of the model. The second section presents results for VA DARS and MD DORS. The third section provides some concluding remarks.

Cohorts, Data Description, and Model Overview

Cohorts

We begin by considering 10,849 VA DARS applicants and 9,018 MD DORS applicants during SFY 2007. We excluded some applicants from our analysis for two reasons. First, our exclusive focus for this article is on individuals with a mental illness (MI), physical impairment (PI), and/or cognitive impairment (CI). Correspondingly, we exclude applicants without any of these conditions. Second, as discussed in Rowe, Ashley, Pepper, Schmidt, & Stern (2019), we exclude applicants in SFY 2007 for whom this case was not their first.

Given these two restrictions, 59% of Virginia applicants and 75% of Maryland applicants are included in one or more of our analysis samples. (Among Virginia applicants, 26% were dropped because they had none of the three disability types and another 15% were dropped because this was not their first case. The corresponding percentages for Maryland are 15% and 10% respectively.) Sample sizes for Virginia are 2,884 applicants for the MI cohort, 2,559 for PI, and 2,420 for CI. Maryland sample sizes are 3,712 for MI, 3,600 for PI, and 2,057 for CI. Note that the samples are not mutually exclusive, i.e., an applicant may be included in more than one sample. For example, an individual with a mental illness as well as a physical impairment is included in both the MI and PI samples, and we control for that co-morbidity in the model.

Data Description

Service choice. We model both the choice among and the labor market impacts of seven broad categories of VR services for each individual in the sample:

- *Diagnosis & evaluation*: services for assessing eligibility and developing an Individualized Plan for Employment (IPE) as well as medical diagnostics;
- *Training*: career & technical training including vocational, job readiness, on-the-job, GED (high school equivalency certificate);
- *Education*: post-secondary education;
- *Restoration*: medical/healthcare service and assistive technology;
- *Maintenance*: extra living expenses (such as shelter, food, clothing, incidentals) needed for an individual to participate in VR, as well as VR-related transportation and vehicle/home modifications;
- *Job Placement*: job search and placement assistance including resume preparation, developing interview skills, identifying job opportunities, referral to a specific job; and
- *Job Supports*: job coaching, supported employment.

SFY 2007 applicants to either agency could receive services from any combination of three sources: (1) agency counselors, (2) external vendors (purchased services), and (3) a state-operated comprehensive rehabilitation center (SOCRC, i.e., the Wilson Workforce Rehabilitation Center [WWRC] in Virginia and the Workforce Training Center [WTC] in Maryland). Although we would prefer to include all services from all sources when estimating individual-level impacts, we cannot do so for counselor-provided services because until recently, neither agency tracked specific VR counselor services provided to specific VR participants in its electronic database. On the other hand, agency staff report that nearly all applicants receive counselor services for eligibility assessment, developing an IPE, and/or counseling; all of which fall into our classification of *diagnostic and evaluation* services. To the extent that this is true, the group of individuals not receiving counselor-provided services would likely be too small to serve as a valid comparison group in gauging the effectiveness of those services, and their omission would have little impact on our estimates.

In contrast to counselor-provided services, both agency databases include sufficiently detailed service information for the other two sources to allow us to identify which SFY 2007 applicants received which service(s) from either or both sources. The choice between purchasing a service or providing it through a SOCRC depends upon a number of factors. Irrespective of why, the labor market impact of a service could differ depending upon that choice. Accordingly, we estimate differential labor market impacts depending upon whether each type of service is purchased only, provided by the SOCRC only, or both.

Table 1 presents the distribution of service choices by agency, service type, and source. Although the table cannot distinguish between the many possible reasons for agency differences in service provision, it does reveal that meaningful differences do exist. Three merit comment. First, in 2007, MD DORS purchased *diagnostic and evaluation services* at a much higher rate than did VA DARS. In Maryland, the norm was to purchase services to assist in eligibility assessment whereas in Virginia the norm was to undertake the initial assessment in-house. Soon after 2007, MD DORS substantially decreased the practice of purchasing these services after deciding that, in most cases, counselors were quite capable of performing that determination using existing and readily available information regarding the applicant's current functioning. Second, both SOCRCs provided *maintenance support* through their residential facilities. However, Table 1 does not show any *maintenance support* for the WTC. The explanation is that while WWRC recorded residential per diems for individuals, the WTC did not. Third, VA DARS did not purchase *placement services* but WWRC did provide them for 2% of the applicants. By contrast, MD DORS did provide them through purchase and/or through the WTC for 14% of its applicants. Irrespective of the reasons for such heterogeneity across agencies, we do model these service choices as well as their labor market impacts by agency, disability type, service type, source, and period.

For brevity, Table 1 does not differentiate by disability type because in general, the percent receiving a service are within a few percentage points across disability cohorts. The exceptions are job supports (7-9 points lower for PI than MI or CI in both agencies) and restoration (4-5 points higher in VA DARS and 7-9 points higher in MD DORS for PI than MI or CI). Each value in the table represents the percent of all individuals across the MI, PI, and CI disability cohorts. The percentages receiving services likely appear to be low to most readers because practitioners commonly focus on VR program participants who have developed an IPE. By contrast, our analysis samples include individuals whose cases were closed for various reasons prior to eligibility determination (3% of VA applicants, 15% for MD) as well as those who were eligible but did not complete an IPE (25% for VA, 28% for MD). For further discussion, see Rowe et al., 2019.

Employment and earnings. The unemployment insurance agency in each state provided us with quarterly earnings and employment status (defined as employed in the quarter if earnings are above zero) for up to three years prior to the application quarter and five years after. Thus, we have at least 30 separate quarterly observations for each individual regarding employment status and nominal earnings (if employed). We separate the quarterly observations into four distinct periods.

1. Two or more quarters prior to the application quarter. We use this period as a baseline against which to measure impacts for each service type.
2. The quarter immediately preceding the application quarter. Employment and earnings are known to drop in the periods just before individuals apply to many workforce development programs. To account for this decline (known as the Ashenfelter (1978) dip), we explicitly allow service effects to vary in this quarter.
3. The first eight quarters after application (the "short" run), during which time many individuals are receiving VR services
4. More than eight quarters post-application (the "long" run), by which time most participants' cases have closed.

For each agency and disability type, we estimate separate labor market impacts for each of the seven service types, each of the three sources (purchased services only, WWRC/WTC only, or both), and each of the four periods. Distinguishing by service source differs from our previous work (Dean, Pepper, Schmidt & Stern, 2015, 2017, 2018) and provides more nuance in service impacts. In total, we estimate 84 service impacts for employment propensity and another 84 for nominal earnings (if employed). The model controls for as many observable (individual and labor market) characteristics as possible in an attempt to ensure that our estimated changes result from provision of the service rather than from extraneous factors that are correlated with provision of the service. This subsection describes trends in the labor market variables over these periods. However, we caution against over-interpreting these trends because they do not control for anything other than disability type and agency. For example, these rates are influenced by the health of the United States economy. The economy strengthened in the years before and during the application year of SFY 2007 that preceded the 2008 financial crisis and ensuing Great Recession. The effects can be observed in the average U.S. unemployment rates of 5.3% for SFY 2004-2006 (roughly the pre-application period), 4.7% for SFY 2007-2008 (roughly the short run), and 8.9% for SFY 2009-2011 (roughly the long run) (Bureau of Labor Statistics, 2018). The decline of the U.S. economy likely played an important role in the fall of employment rates between the short and long runs for these groups of SFY 2007 applicants to VR.

Table 2 reports employment rates and mean nominal earnings (if employed) by agency, disability type, and period, and shows that both measures vary by agency, by disability type, and over time. With respect to employment rates, the PI cohort for both agencies exhibits a notable Ashenfelter dip in the quarter preceding application. By contrast, employment rates remain stable in both agencies for the MI cohort and actually rise in the quarter before application for the CI cohort. Employment rates rise across all cohorts and both agencies in the eight quarters following application but then fall in the quarters after that.

The second portion of Table 2 shows mean quarterly earnings for those applicants who were employed in the quarter. The Ashenfelter dip, ranging from about $400 to about $1,200, is evident for all disability cohorts across both states. As was the case with employment rates, the PI cohort exhibits the largest dip of about $1,200 in each state. However, the Great Recession does not appear to have affected mean earnings for those with a job as it did the employment rate. Mean earnings rose between the short run and long run in both states for all disability types, ranging from about $600 to about $800.

Figures 1-4 spotlight trends during the pre-application period by charting quarterly employment rates and mean nominal earnings (if employed) for the quarters leading up to application, separately by disability cohort and agency. Indeed, the patterns and levels of employment and earnings vary dramatically by disability and, for employment rates, by agency.

Although many factors affect trends in employment and earnings, we reiterate that the data in these simple tables and charts only account for an applicant’s disability and the state agency. For that reason, the main lessons at this stage are that there are significant differences across agencies and disability types in their pre-service employment and earnings. Any model that estimates labor market impacts of service provision must include controls for pre-service employment, pre-service earnings if employed, agency, and disability.

Overview of the Estimation Model

Our complete model for estimating impacts of VR services on labor market outcomes includes seven equations to model the probability of provision for each of the service categories, one equation to model the probability of employment in a quarter and one equation to model the *log* of nominal quarterly earnings (conditional on being employed). Although we refer simply to “earnings” elsewhere in this article, we use the *log* of earnings in the model for several reasons, including the interpretation of coefficients in percentage terms. For example, a coefficient of .02 for education would indicate that each additional year of education leads to 2% higher earnings if employed.

Each of these nine equations includes a set of over twenty explanatory variables that we believe to be correlated with, but not influenced by, either service provision or labor market performance. Most of these explanatory variables relate to an individual’s characteristics and disability. We provide additional explanation and descriptive statistics for these variables in online Appendix A.

A more thorough description of the model, including explicit mathematical representations of the model's relationships, can be found in online Appendix B. Although discussed more thoroughly in that appendix, two points merit discussion here. First, the model does use techniques to control for "selection bias." In this model, selection bias might occur when one or more variables that influence both the probability of service provision and labor market performance are excluded from the analysis, often because they are not available or not measurable. Examples might include motivation, family support, and access to transportation. For additional discussion, see Clapp, Pepper, Schmidt & Stern, 2019.

Second, the labor market equations use quarterly observations for the two outcome variables, employment in the quarter (1 for employed, 0 otherwise) and *log* of nominal quarterly earnings conditional on being employed. As described previously, in each labor market equation we estimate separate service coefficients by agency and disability type for (a) three service sources, (b) seven service types, and (c) four periods relative to the application quarter. How are these coefficients to be interpreted? As an example, consider coefficients in either labor market equation for *training* services purchased solely from an external vendor. A negative coefficient for the pre-application period (through the second quarter prior to application) would indicate that, after controlling for everything in the model, those applicants enter VR with lower levels of employment and earnings than those who do not receive *training* services at all. A positive coefficient would indicate higher levels of employment and earnings.

Now consider how we estimate service impacts on labor market performance. We estimate them as the change between the pre-application coefficient and the post-application coefficient.

Specifically, we calculate the short-run (first eight quarters after application) impact as the difference between the short-run and pre-service coefficients. We calculate the long-run (more than eight quarters after application) impact in an analogous manner. Thus, a positive long-run coefficient could result in either a positive change when it is larger than the pre-application coefficient or a negative change when it is smaller. Conversely, a negative long-run coefficient could result in either a positive change when it is less negative than the pre-application coefficient or a negative change when it is more negative or when the pre-application coefficient is positive.

Estimation Results

This section presents results by disability cohort for VA DARS and MD DORS. For each, we present impacts of the service categories on employment as well as earnings (if employed). We then present ROI results for each cohort and agency. Before doing that, we consider the context in which to interpret these results.

ROI analysis restricts itself to readily quantifiable outcomes. In this and many other studies, those outcomes are employment and earnings. Given that restriction, these analyses exclude qualitative, non-market impacts of VR such as levels of independent living, community integration, and self-efficacy. To the extent that VR exerts a positive influence on these, ROI estimates will underestimate VR's total impact. We recognize that the magnitude of the underestimate varies by client and may be substantial for some individuals and disabilities. However, we do not have access to data on these outcomes. Researchers have not yet developed a widely accepted methodology to collect data on, and assign dollar values to, quality of life improvements. (For additional insight, see Hopkins, 2019.) In the meantime, it is important to recognize that ROI estimates for VR and other workforce programs likely underestimate the program's overall impact.

More specific to the SFY 2007 cohort of VR applicants, individuals applied for VR services during a period of economic prosperity. At 4.5%, SFY 2007's unemployment rate was the lowest it had been in six years (Bureau of Labor Statistics, 2018). That growing economy informed the cohort's decisions to apply for VR services as well as their work with counselors to develop IPEs. That trend ended with the 2008 financial crisis that culminated in the deepest recession since the Great Depression. Upon exiting VR, individuals faced a depressed labor market as evidenced by national unemployment rates that climbed to 9.8% by 2010, and slowly declined after that (Bureau of Labor Statistics, 2018). Although we include controls for the state of the economy, those might not be adequate to capture the full impacts on this cohort's ability to find employment and/or increase earnings.

Plausibly, the declining state of the economy affected transition-age youth (traditionally defined as 16-24 years of age) even more negatively than it affected VR program participants of prime working age. Transition-age youth are concentrated most heavily in the CI cohorts where individuals under the age of 24 comprise 61% of the VA DARS applicants and 77% of the MD DORS applicants (compared with percentages ranging from 16-31% for the MI and PI cohorts). Although we include a dummy variable for applicants who are "very young," that variable is unlikely to fully control for the effects of the Great Recession on transition-age youth. We anticipate the differential effect of the economy on transition-age youth to be most substantial for the CI cohort.

Preface: Interpreting the Results for Virginia and Maryland

Clapp et al. (2019) describe a number of approaches to estimating VR service impacts, including difference-in-differences (diff-in-diff). To provide a contrast to estimates through the VR-ROI model presented in the next subsection for Virginia DARS and in the subsection after that for Maryland DORS, we calculated treated vs. untreated, diff-in-diff service impacts for each agency. Specifically, we calculated mean labor market performance over the pre-application period as well as the long-run period (more than eight quarters after application) for both the "treated" (applicants who received any VR services) and the "untreated" (applicants who did not). Diff-in-diff VR service impacts are then calculated as [(LR mean - Pre-app mean)$_{\text{Treated}}$ - (LR mean - Pre-app mean)$_{\text{Untreated}}$]. Virginia's diff-in-diff results are negative for both employment and earnings while Maryland's are both positive. Specifically, relative to the "untreated," the "treated" change in employment rates is 0.6 percentage points lower in Virginia and 3.3 percentage points higher in Maryland while the change in mean earnings (if employed) is 16.1% lower in Virginia and 6.2% higher in Maryland. Although simple to calculate, these numbers can be misleading and provide no insight into the forces that lie behind them. They do not control for any covariates (individual characteristics, severity of disability, selection issues) and provide estimates for a generic VR participant receiving a generic VR service. By controlling for covariates and providing separate results for three disability types and seven VR service types, the figures presented and discussed in the next two subsections indicate that diff-in-diff estimates mask the considerable heterogeneity in effects.

The next two sections present and discuss long-run (more eight quarters after application) service impacts by disability type when the agency purchases services from an external vendor but does not provide them by the SOCRC (Figures 5-7 for Virginia and 8-10 for Maryland). Each figure reports these estimated impacts separately by service category for employment propensity as well as conditional earnings. We also estimate impacts when providing the service only by the SOCRC or by both an external vendor and the SOCRC. We include their impacts and costs in our ROI estimates but do not show them in this paper. Rather, online Appendix C provides estimates and statistical significance for all service impacts, the influence of explanatory variables in the service provision and labor market equations, and many other model parameters.

Several considerations are important to understanding the estimates provided in Figures 5-10.

- As noted earlier, we focus on long-run impacts (more than eight quarters after application) in this article. We do this to allow program participants to exit from VR before measuring service impacts. The period includes twelve or more quarters for each individual.
- The service impacts depicted in this chart are calculated as changes from the twelve-quarter period prior to application (but excluding the quarter immediately preceding application), relative to individuals in the cohort not receiving the service either by purchase or from the SCORC. Thus, a positive value indicates that the change in labor market performance was stronger for those receiving the service than for those who did not, not necessarily that the change itself was positive for service recipients. Conversely, a negative value indicates that the change was not as strong, not necessarily that the change itself was negative for service recipients.
- We consider an individual to be in the "treatment" group for those service categories in which the individual received a service and in the "comparison" group for the others. Thus, not only do we allow for differential service regimens, we also take a more nuanced view than simply "received substantial VR services" versus "did not

receive substantial VR services." We do not estimate interaction effects across service types because Dean et al. (2015, 2017, 2018) found interactions between pairs of services to be statistically insignificant.

Results for Virginia DARS

With these considerations in mind, what does these charts reveal? We begin by examining the effects of VR services on MI individuals reported in Figure 5. First, consider *Training* as an example for interpretation. Changes in employment rates were 27 percentage points higher when *Training* services were purchased and changes in earnings (conditional on being employed) were 22% higher. Both employment and earnings changes were also higher when services were purchased in the *Restoration*, *Maintenance*, and *Job Supports* categories. The magnitudes were small for *Restoration* and *Maintenance*; however, they were large for *Job Supports* (50 percentage points higher for employment and 19% for conditional earnings).

By contrast, the negative values for *Diagnosis and Evaluation* (*D&E*) indicate that changes were lower for individuals for whom *D&E* services were purchased than for those for who did not receive *D&E* services. While we cannot rule out that D&E services cause clients to have worse labor market outcomes, we find this interpretation of our estimates to be implausible. An alternative explanation for these negative values is that the model does not fully control for "selection bias," i.e., a variable that influences both the probability of service provision and labor market performance. Conferring with agency staff in such an instance has sometimes provided additional insights into the nuances of the agency and often gives us ideas to improve the model further. In the case of *D&E*, we have observed negative values across agencies and disability cohorts. Agency staff have suggested that combining eligibility determination and medical diagnostics into *D&E* services might be the source of the problem. Accordingly, we are exploring their separation as well as adding a dummy variable to identify individuals who received no services of any kind after eligibility determination. The *Education* results also appear to be anomalous. Individuals who received support for education beyond high school appear to enjoy higher employment rates but lower earnings once employed. However, these results are unlikely to have a big impact on overall ROI estimates due to the small numbers of individuals involved – 66 or 2% of the 2,884 individuals in this cohort. Indeed, any service from any source that is provided to a small proportion of applicants will not have as large of an impact on ROI as the more prevalent service types. As shown in Table 1, this applies most often to services provided by the WWRC (or, for MD DORS, from the WTC).

Figures 6 and 7 display long-run purchased service impacts by service category for the PI and CI cohorts, respectively. The service impacts for the PI cohort in Figure 6 are particularly strong. With the exception of *D&E*, those receiving purchased services exhibit improved labor market changes for all service categories compared with those who did not receive the service. With respect to employment impacts, *Job Supports* are the strongest (43 percentage points higher) with *Training* second (29 percentage points higher). For conditional earnings impacts, *Training* is strongest (19% higher) with *Education* and *Job Supports* tied for second (16% higher).

For the CI cohort, Figure 7 shows less consistency across service categories than do the results for the MI and PI cohorts. CI service recipients for whom services are purchased for *Training*, *Education*, and *Job Supports* enjoy stronger labor market changes than for those for whom those services are not provided. These results are encouraging in light of the fact that 61% of these individual were under twenty-four at application and many were attempting to enter the job market in the face of the Great Recession. Less encouraging are the negative values for those receiving *D&E*, *Restoration*, and *Maintenance*.

Turning to our estimates of ROI, the labor market impacts of VR service provision from any source represent the “benefits” side of ROI while VR agency costs represent the “investment” side. We estimate total VR costs including “fixed cost” (administrative, counseling, and placement) which are not tracked at the individual level as well as “variable costs” (services both purchased and provided by the WWRC) which are tracked for individuals. Thus, our estimates allow us to measure the benefits of VR relative to its costs. We make ROI estimates on a individual basis to allow us to provide a full set of descriptive statistics as well as aggregate results for any desired group of individuals (e.g., a disability cohort, using a particular mix of services, or even the full agency).

Table 3 presents 20-year ROI results in the form of in the form of internal rate of return (ROR) for the three VA DARS disability cohorts. (See Hollenbeck, 2019 and Clapp et al., 2019 for a discussion of alternative ROI measures as well as their relative strengths and weaknesses.) We base the 20-year annualized ROR estimates on 5+ years of post-application employment and earnings data. We extrapolate the labor market impacts to 20 years using standard techniques. that are discussed in online Appendix D. The appendix provides a detailed discussion of the methodology and an illustration of the relative impact on ROR for 5-year, 10-year, and 20-year RORs. When interpreting these ROI results note that a positive ROR indicates that the labor market gains more than offset the agency’s costs and the higher the ROR the better. Additionally, a ROR that is not positive does <u>not</u> necessarily mean that there were no labor market gains from VR services. It might simply indicate that those gains did not exceed the agency’s costs.

Table 3 shows strong ROR results for both the MI and PI cohorts. While we would not expect all program participants to enjoy labor market gains from VR services, much less gains that exceed VR costs, 67% of the MI cohort and 58% of the PI cohort did just that. The median values indicate that half of them enjoyed annual RORs exceeding 17.5% and 15.5%, respectively. By comparison, the long-run annual return in the United States is about 1% in money market accounts and 10% in the stock market. Individuals in the CI cohort did not fare nearly as well – 25% enjoyed labor market gains that exceeded VR costs. Note that these results are based on employment and earnings data that run through 2012. We plan to extend that to 2017 which will provide ten years of post-application data, a period well into the recovery and when most in the CI cohort will have entered their prime working ages.

Results for Maryland DORS

As in the previous subsection, we first present and discuss service impacts on labor market outcomes before presenting ROI estimates. Figures 8-10 display long-run service impacts (more than eight quarters after application) for the MI, PI, and CI cohorts when DORS purchased services from an external vendor but does not provide them through the state-operated comprehensive rehabilitation center (WTC). We show these impacts separately by service category and for employment propensity and conditional earnings. Impacts when the service was provided only by the WTC or by both an external vendor and WTC have also been calculated and are shown in online Appendix C.

The variation across these three figures provides a much more nuanced perspective on the impacts of the VR program than would an approach that compares a generic VR program participant who receives "substantive" services to one who does not without regard to disabling condition. These figures exhibit considerable heterogeneity across service categories and disabling conditions. With respect to service types, participants for whom the agency purchased *Training* and *Education* services show employment and earnings changes that were higher for all three cohorts than for those not receiving the service. The magnitude of the *Training* impacts are consistent across disabilities while those for *Education* are notably larger for the PI cohort. Although the magnitudes are not large, participants receiving purchased *Restoration* services had lower employment and earnings changes for each cohort. Purchased *Maintenance* services exhibited higher employment changes but lower earnings changes across all cohorts. Labor market changes for *Placement* services are mixed – higher for the PI and CI cohorts but lower for the MI cohort. Finally, purchased *Job Supports* resulted in notably higher employment impacts, but minimal or lower earnings changes.

Table 4 presents annualized internal rates of return (ROR) estimates for the three MD DORS disability cohorts. Results are particularly strong for clients in the MI cohort. Seventy-eight percent exhibited labor market impacts that exceeded VR costs (i.e., positive RORs). RORs exceeded 14% for half of the cohort, exceeded 26% annually for 25%, and exceeded 42% for 10%. ROR results were not as strong for the PI and CI cohorts with less than half of those cohorts enjoying positive RORs.

For reasons discussed earlier, weaker RORs are not particularly surprising for the CI cohort. What is surprising, however, is that ROR results for the PI cohort are the weakest of the three even though their labor market impacts appear to be the strongest (see Figures 6-8). The explanation lies with *D&E* and *Restoration*. For both of these, PI clients receiving the service had lower employment and conditional earnings changes than those who did not. That was true irrespective of whether the service was purchased, provided by the WTC, or both. In particular, *D&E* has an outsized effect on these ROR estimates because MD DORS routinely purchased services in that era to assist with eligibility determination and IPE development. As a result, *D&*E services were purchased for over 56% of all applicants. Shortly thereafter, the agency emphasized that their counselors could use existing and readily available information to perform eligibility determinations and provide vocational guidance and counseling competently and at a lower cost. Purchases of D&E services were much less common from then on.

As an experiment, we asked what would be the results if *D&E* were excluded from the ROR calculations. In other words, what is the ROR for core VR services? In that case, 66% of the PI cohort would have positive RORs with a median of 8.7% annually, a 75th percentile of 26.6%, and a 90th percentile of 49.4%. Online Appendix E shows side-by-side comparisons of ROR that include or exclude *D&E* services for each disability type and both agencies.

Concluding Remarks

The VR-ROI model provides a framework for estimating the labor market impacts of VR service provision differentially across type of service, type of disability, time period, and agency. When applying the VR-ROI model to applicants during SFY 2007 to VA DARS and MD DORS, we find considerable heterogeneity within each agency of estimated service impacts on employment versus earnings (if employed) as well as across disability type (mental illness, physical impairment, cognitive impairment) and service type (*diagnosis & evaluation*, *training*, *education*, *restoration*, *maintenance*, *job placement*, *job supports*) within each disability type. We also find noteworthy differences in ROI across disability types.

The approach provides a much more nuanced perspective on VR than do simpler approaches such as a difference-in-differences approach applied to a generic VR participant receiving a generic VR service. For example, that approach indicates a negative impact of VR services on both employment and earnings (if employed) for VA DARS. However, the VR-ROI model provides generally positive results for participants with mental illness as well as those with a physical impairment. For reasons discussed elsewhere in this paper, results for participants with a cognitive impairment appear to be negative. We observe notable differences across service types within each disability as well.

The VR-ROI model employs state-of-the-science statistical techniques in an attempt to ensure that labor market impacts result from the provision of VR services rather than are simply correlated with them. Nevertheless, we do emphasize two important qualifications. First, we reiterate some of the lessons upon the ethical use of these results as discussed in Froehlich, Bentley, Emmanuel, & McGuire-Kuletz (2019). Paramount among these is that such results not be used to deny services to any group of applicants. Rather those insights might be used in conjunction with other program information for broader management decisions. Additionally, given the major differences across agencies and the nuances of these results, they should never be used to make comparisons across agencies.

Second, we emphasize that these results are preliminary. We continue to work on our understanding of the Great Recession as well as controlling for applicants who receive no services beyond *diagnosis & evaluation* services for eligibility determination. Additionally, we recently received additional data that provides ten full years of earnings beyond application. This extends well into the recovery from the Great Recession and further into the working life of the younger participants with a cognitive impairment. Our planned analyses using these newer data are likely to produce different results.

Table 1

Service Choice by Service Type, Agency, and Provider (VR Applicants during SFY 2007)

	VA DARS (6,380 applicants)				MD DORS (6,749 applicants)			
Service Type	PS Only	WWRC Only	Both	Neither	PS Only	WTC Only	Both	Neither
Diagnostic	28%	8%	5%	58%	43%	8%	13%	35%
Training	11%	6%	1%	82%	13%	3%	1%	83%
Education	2%			98%	6%			94%
Restoration	18%	6%	2%	74%	8%	10%	3%	79%
Maintenance	24%	4%	3%	69%	24%			76%
Placement		2%		98%	10%	3%	1%	86%
Job Supports	20%			80%	13%			87%

Table 2

Employment Rates and Mean Nominal Quarterly Earnings (if employed) by Agency, Disability Type, and Period (VR Applicants during SFY 2007)

Descriptive Statistic	VA DARS			MD DORS		
	MI	PI	CI	MI	PI	CI
# of Applicants in Cohort	2,884	2,420	2,559	3,712	3,600	2,057
% Employed						
2 or More Qtrs Before Application	31%	37%	22%	30%	31%	28%
1 Quarter Before Application	32%	33%	32%	28%	25%	34%
First 8 Qtrs After Application (short run)	39%	36%	42%	35%	30%	42%
More than 8 Qtrs After App. (long run)	29%	29%	39%	25%	22%	37%
Mean Nominal Earnings (if employed)						
2 or More Qtrs Before Application	$3,219	$4,440	$1,988	$2,958	$4,283	$2,437
1 Quarter Before Application	$2,420	$3,250	$1,580	$2,086	$3,092	$2,087
First 8 Qtrs After Application (short run)	$2,589	$3,329	$2,154	$2,712	$3,521	$2,401
More than 8 Qtrs After App. (long run)	$3,335	$3,933	$2,954	$3,394	$4,234	$3,143

Note: Each observation in the table represents one quarter. Thus, for example, the number of observations for the row labeled, “First 8 Qtrs After Application,” is eight times the number of applicants.

Table 3

20-year Annualized ROR for 4,121 VA DARS 2007 Participants

	MI	**PI**	**CI**
% with Positive ROR	67%	58%	25%
ROR at Median	17.5%	15.5%	0.0%
75th Percentile	42.8%	42.1%	0.0%
90th Percentile	77.0%	77.0%	16.1%

Table 4

20-year Annualized ROR for 5,197 MD DORS 2007 Participants

	MI	**PI**	**CI**
% with Positive ROR	78%	26%	33%
ROR at Median	14.0%	0.0%	0.0%
75^{th} Percentile	26.1%	1.1%	8.7%
90^{th} Percentile	42.5%	23.0%	30.2%

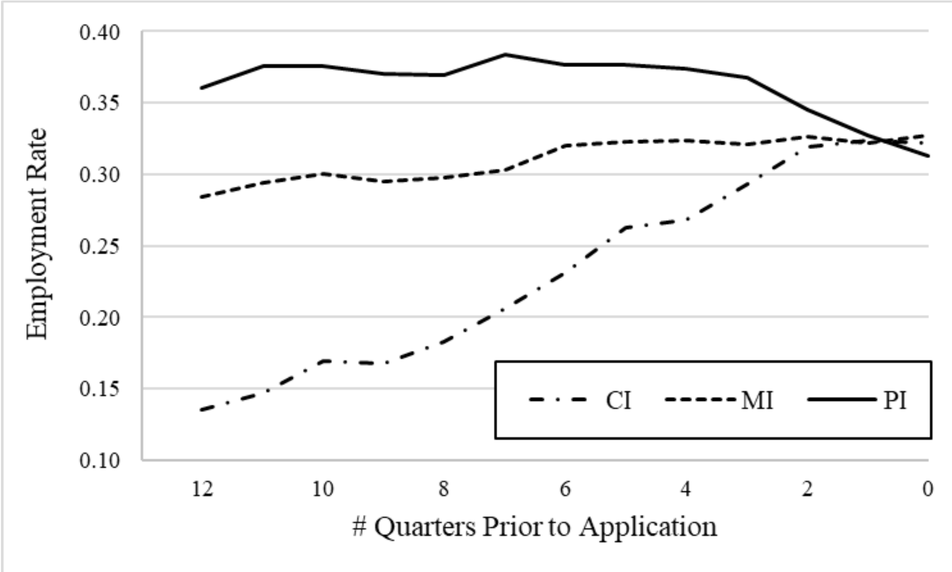

Figure 1. VA DARS – Quarterly Employment Rates Prior to Application by Agency and Cohort

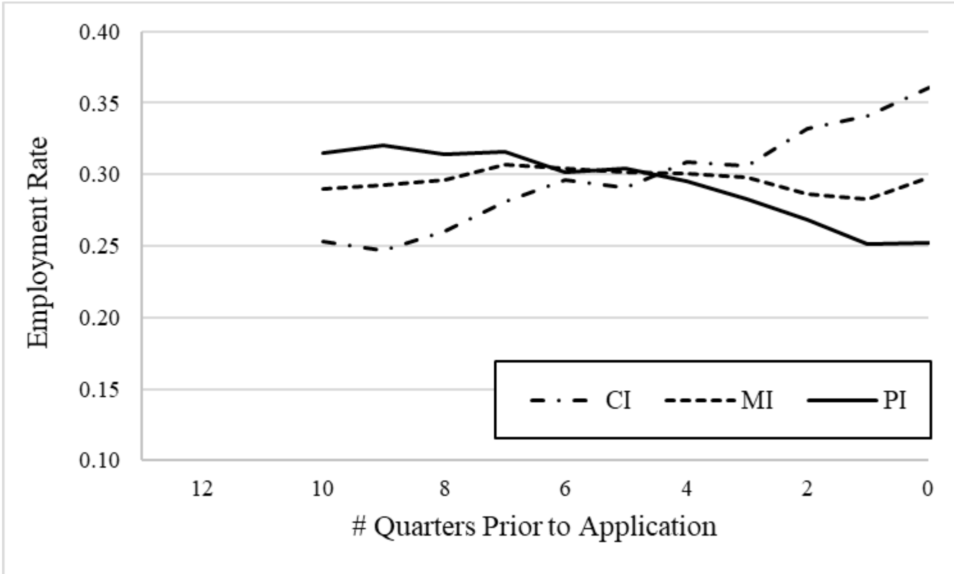

Figure 2. MD DORS – Quarterly Employment Rates Prior to Application by Agency and Cohort

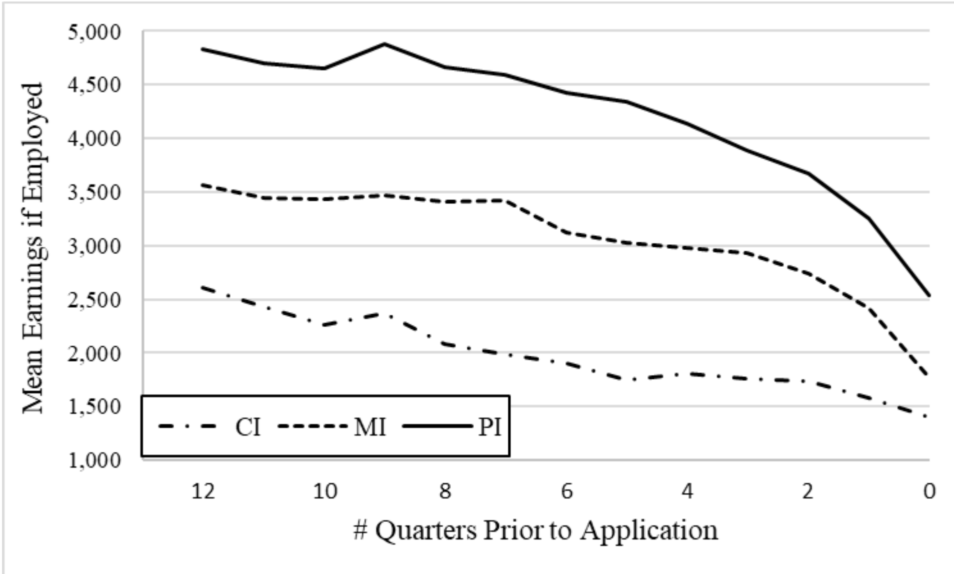

Figure 3. VA DARS – Average Quarterly Earnings (if employed) Prior to Application by Agency and Cohort

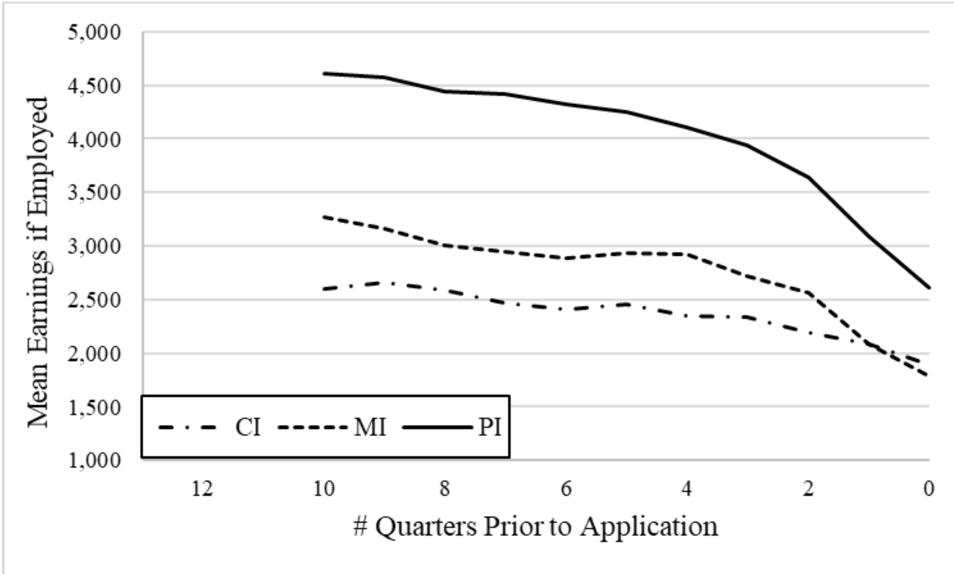

Figure 4. MD DORS – Average Quarterly Earnings (if employed) Prior to Application by Agency and Cohort

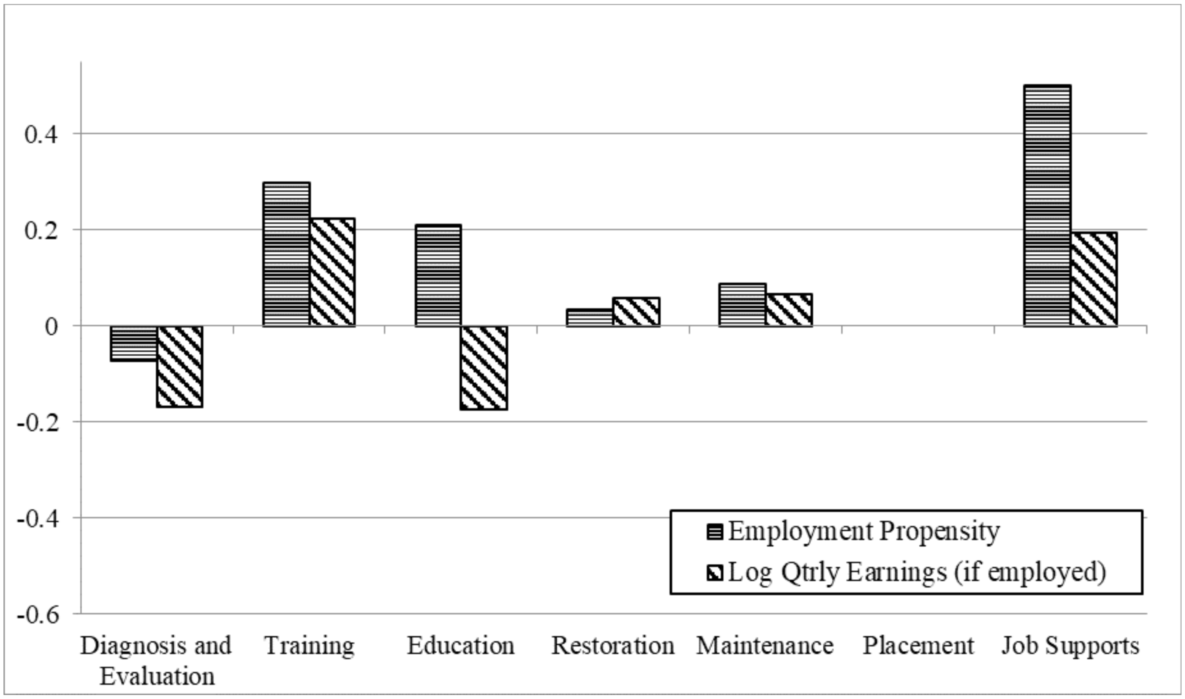

Figure 5. VA DARS MI Cohort – Long-run Impact of Purchased Services on Labor Market Outcomes by Service Category. (Note: VA DARS does not purchase *Placement* services although they are provided by WWRC.)

Figure 6. VA DARS PI Cohort – Long-run Impact of Purchased Services on Labor Market Outcomes by Service Category (Note: VA DARS does not purchase *Placement* services although they are provided by WWRC.)

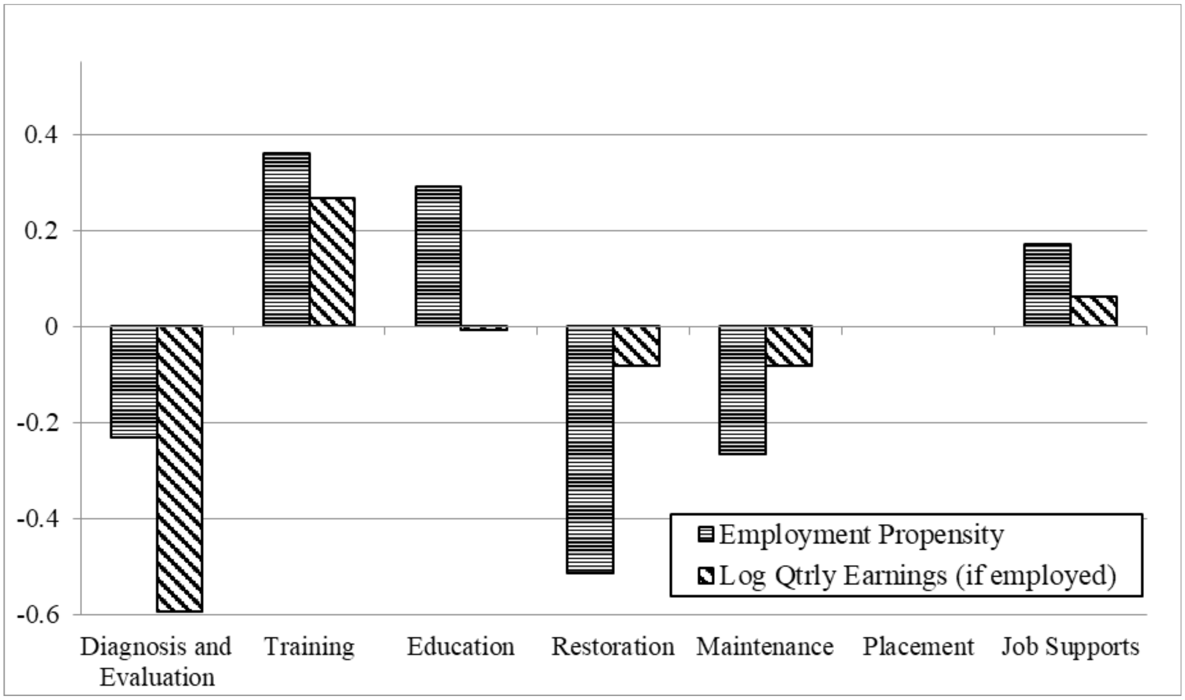

Figure 7. VA DARS CI Cohort – Long-run Impact of Purchased Services on Labor Market Outcomes by Service Category (Note: VA DARS does not purchase *Placement* services although they are provided by WWRC.)

Figure 8. MD DORS DARS MI Cohort – Long-run Impact of Purchased Services on Labor Market Outcomes by Service Category

Figure 9. MD DORS DARS PI Cohort – Long-run Impact of Purchased Services on Labor Market Outcomes by Service Category

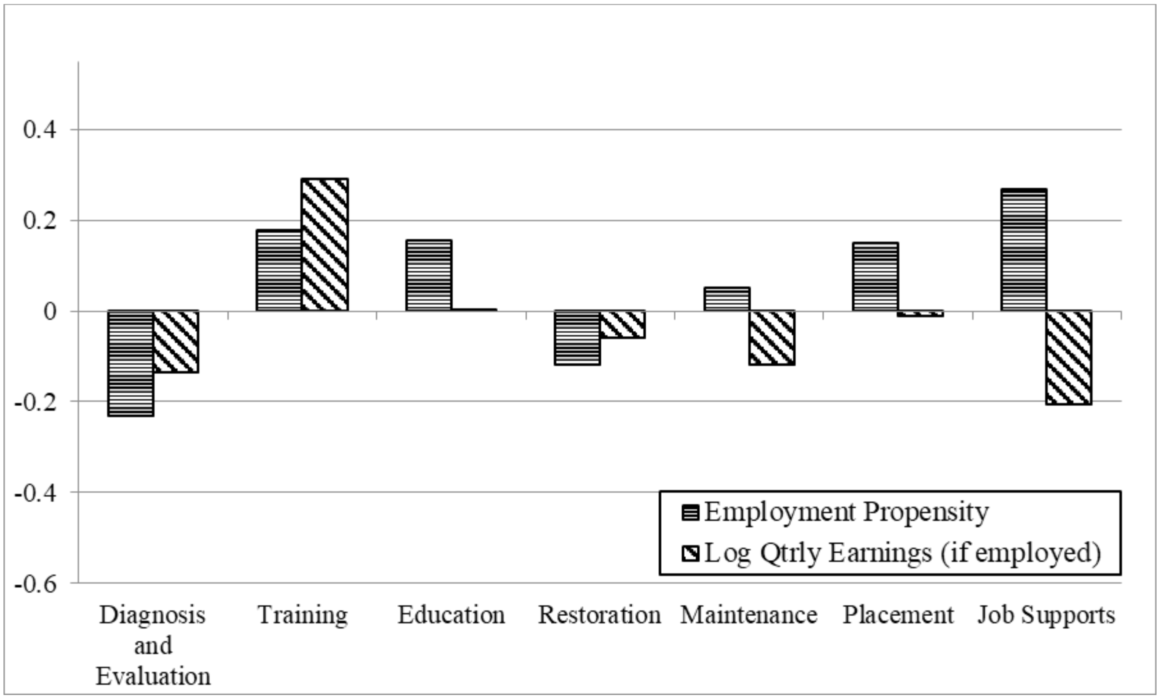

Figure 10. MD DORS DARS CI Cohort – Long-run Impact of Purchased Services on Labor Market Outcomes by Service Category

References

Ashenfelter, O. (1978). Estimating the effect of training programs on earnings. *Review of Economics and Statistics*, *60*, 47–57.

Bureau of Labor Statistics (2018). Unemployment rate - civilian labor force - LNS14000000. U.S. Department of Labor. Retrieved from http://data.bls.gov/cgi-bin/surveymost?ln.

Clapp, C. M., Pepper, J. V., Schmidt, R. M., & Stern, S. N. (2019). Conceptual issues in developing return on investment estimates of vocational rehabilitation programs. *The Journal of Rehabilitation Administration*, forthcoming.

Dean, D. H., Pepper, J. V., Schmidt, R. M., & Stern, S. N. (2015). The effects of vocational rehabilitation for people with cognitive impairments. *International Economic Review*, *56*(2), 399-426.

Dean, D. H., Pepper, J. V., Schmidt, R. M., & Stern, S. N. (2017). The effects of vocational rehabilitation for people with mental illness. *Journal of Human Resources*, *52*(3), 826-858.

Dean, D. H., Pepper, J. V., Schmidt, R. M., & Stern, S. N. (2018). The effects of vocational rehabilitation for people with physical impairments. *Journal of Human Capital*, *12*(1), 1-37.

Froehlich, R., Bentley, O., Emmanuel, E., & McGuire-Kuletz, M. (2019). The return on investment of vocational rehabilitation services: some ethical considerations. *The Journal of Rehabilitation Administration*, forthcoming.

Hollenbeck, K. (2019). What is ROI? *The Journal of Rehabilitation Administration*, forthcoming.

Hopkins, K. (2019). Social return on investment: An important consideration for state vocational rehabilitation programs. *Journal of Rehabilitation Administration*, forthcoming.

Rowe, K. L., Ashley, J. M., Pepper, J. V., Schmidt, R. M., & Stern, S. N. (2019). An overview of the VR-ROI Project and its approach to estimating ROI. *The Journal of Rehabilitation Administration*, forthcoming

Authors' Notes

Robert Schmidt, Ph.D., is an Emeritus Professor of Economics at the University of Richmond. Christopher Clapp, Ph.D., is a Senior Research Associate in the Kenneth C. Griffin Department of Economics at the University of Chicago. John Pepper, Ph.D., is the Merrill H. Bankard Professor of Economics at the University of Virginia. Steven Stern, Ph.D., is a full Professor of Economics at Stony Brook University. Drs. Clapp, Pepper, and Stern are consultants and Dr. Schmidt is co-Principal Investigator for the VR Return on Investment Project.

We would like to thank Kevin Hollenbeck and Kirsten Rowe for insightful comments and the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR Grant 90DP0070) for generous funding. All remaining errors are ours.