Journal of Rehabilitation Administration

Volume 40(1), 2019

Special Issue: Return on Investment for State Vocational Rehabilitation Programs

3 **Return on Investment for State Vocational Rehabilitation Programs: Introduction to the Special Issue**
Robert M. Schmidt, Kevin Hollenbeck, and Kirsten L. Rowe

5 **What is ROI?**
Kevin Hollenbeck

11 **The Return on Investment of Vocational Rehabilitation Services: Some Ethical Considerations**
Betsy Hopkins

17 **Conceptual Issues in Developing Return on Investment Estimates of Vocational Rehabilitation Programs**
Robert Froehlich, Olivia Bentley, Diona Emmanuel, and Maureen McGuire-Kuletz

23 **Conceptual Issues in Developing Valid ROI Estimates**
Christopher M. Clapp, John V. Pepper, Robert M. Schmidt, and Steven N. Stern

35 **Data Issues in Developing Rigorous and Valid ROI Estimates**
Steven N. Stern, Christopher M. Clapp, John V. Pepper, and Robert M. Schmidt

47 **An Overview of the VR ROI Project and its Approach to Estimating ROI**
Kirsten L. Rowe, Joseph Ashley, John V. Pepper, Robert M. Schmidt, and Steven N. Stern

57 **Applications of the VR-ROI project: ROI estimates for Virginia and Maryland**
Robert M. Schmidt, Christopher M. Clapp, John V. Pepper, and Steven N. Stern

73 **Extending the VR-ROI Approach to Measure the Return on Virginia's Investment in the Public Workforce System**
Jeffrey Brown and Joseph Ashley

Journal of Rehabilitation Administration

Begun in January 1977, the *Journal of Rehabilitation Administration (JRA)* was initiated to establish a knowledge base and improve the professional practice of rehabilitation administration and supervision. The publication is widely distributed to human service administrators, managers, supervisors, researchers, evaluators, students, support personnel, libraries in colleges, hospitals and medical schools as well as instructors in non-profit rehabilitation, vocational rehabilitation and the private sector in rehabilitation.

JRA is provided to all individual NARL and RPEN members and is published bi-annually. Additional copies of issues may be purchased from ElliottFitzpatrick.com

Guidelines for Authors

Manuscript Preparation

Authors should prepare manuscripts according to the Publication Manual of the American Psychological Association (6th ed.). All manuscripts must include an abstract containing a maximum of 150 words, typed on a separate sheet of paper. Typing instructions (all copy must be double-spaced) and instructions on preparing tables, figures, references, metric, and abstracts appear in the APA Publication Manual. Manuscripts should be concisely written in unambiguous language. In addition, manuscripts should present material in logical order, starting with a statement of purpose and progressing through an analysis of evidence to implication and conclusions, with the conclusions clearly related to the evidence presented. All manuscripts must avoid terminology that might imply bias against persons on the basis of racial or ethnic group, gender, disability, age, religion, or sexual orientation. Except in rare instances, as approved by the editors, manuscripts should be no longer than 25 double-spaced typed pages including abstract, text, references, appendices, tables, and figures. Manuscripts failing to meet minimally accepted standards will be returned to the author without review.

The use of footnotes should be avoided wherever possible. Use tables and figures sparingly and include only essential data. All tables should be typed on separate pages. At time of acceptance for publication, graphic items (figures, charts, etc.) will need to be provided as separate files in the form of a TIFF, JPG or PNG (one item per file), prepared with professional-level graphic software. Graphic files should exclude white space around figures/graphs. Figure/Title information should also be excluded from the graphic itself. Refer to the APA Publication Manual for style of tabular presentations and graphic material.

Lengthy quotations (300 cumulative words or more from one source) require written permission from the copyright holder for reproduction. Adaptation of tables and figures also requires reproduction approval from the copyright source. It is the author's responsibility to secure such permission, and a copy of the publisher's written permission must be provided to the Editor immediately upon manuscript acceptance for publication.

A short biographical statement (1-2 sentences) should be included at the bottom of the title page. Statement should include author(s)' terminal degree, professional affiliation, and other relevant (contact or professional) information.

Submission

Manuscripts may be submitted by e-mail to Prof. Carl Flowers, JRA Editor, as an attachment to cflowers@siu.edu.

Manuscripts for the column on Leadership Initiatives should be emailed to Roy Chen, Associate Editor, University of Texas-Rio Grande Valley at roy.chen@utrgv.edu.

Authors of accepted manuscripts are required to supply an electronic version of their manuscripts to the journal via e-mail. Authors should be aware that this electronic file is considered final material for the manuscript.

Review Policy

JRA is a refereed journal, and uses a blind review process. All submissions should include a cover page with each copy of the manuscript, indicating the title of the manuscript, authors' names, and organizational affiliations. Author's notes containing information about the authors' identity or affiliations should be on the cover page. Every effort should be made to ensure that the manuscript itself does not contain clues to the authors' identity. Although feedback is ordinarily given, the editors reserve the right to reject a manuscript for publication without providing a rationale for their decision.

Publication Policy

APA prohibits an author from submitting the same manuscript for concurrent consideration by two or more publications. As *JRA* primarily publishes original material only, APA policy prohibits, as well, publication of any manuscript that has already been published in whole or substantial part elsewhere without permission and acknowledgement as a reprint. Authors have an obligation to consult the Editors concerning prior publication of any data upon which their manuscript depends. Authors bear full responsibility for the accuracy of references, tables, quotations, figures, and the overall content of the manuscript. With the exception of reprints, manuscripts accepted for publication become the sole property of *JRA* for the purpose of copyright.

Introduction to the Special Issue

Robert M. Schmidt
University of Richmond

Kevin Hollenbeck
Upjohn Institute

Kirsten L. Rowe
Virginia Department for Aging and Rehabilitative Services

Return on investment (ROI) and cost-benefit analyses of the public vocational rehabilitation(VR) program in the United States have a long history, from the first published studies by Bellante (1972), Conley (1969), and Worrall (1978) to a number of more recent publications (see, for example, Bua-Iam & Bias, 2011; Bua-Iam, Hampton, Sink, & Snuffer, 2013; Cimera, 2010; Dean, Pepper, Schmidt, & Stern, 2015, 2017; Schmidt et al., 2019 (this issue); Uvin, Karaaslani,& White, 2004; Wilhelm, 2013).While they have taken different approaches, existing VR-ROI studies have attempted to accurately estimate the full cost of VR as well as its economic benefits for program participants and, in some cases, for employers and taxpayers. VR-ROI estimation methods have become increasingly more rigorous in recent decades, and have increasingly focused on state-specific program impacts, as access to state-level administrative data on VR program participants, the VR services they receive, and their employment histories has increased.

The work currently being carried out by the Vocational Rehabilitation Return on Investment (VR-ROI) Project – a Disability and Rehabilitation Research Project funded by the National Institute on Disability, Independent Living and Rehabilitation Research – represents perhaps the most rigorous approach to estimating ROI for state VR programs. The purpose of this special issue is threefold:

1. To highlight conceptual, methodological and ethical considerations for VR program administrators and others interested in developing and using VR-ROI estimates;
2. To provide an overview of the VR-ROI Project and key features of its approach to estimating ROI for state VR programs; and
3. To describe the results of recent VR-ROI Project analyses of the VR programs in two states.

The first article, by Hollenbeck, offers an introduction to basic economic concepts and calculations used in estimating ROI and applies them to the VR context. It is followed by Hopkins' article on social ROI and its potential importance for the public VR program. Froehlich, Bentley, Emmanuel, and McGuire-Kuletz then discuss ethical considerations for policymakers and practitioners in using the results of ROI analyses.

These articles are followed by Clapp, Pepper, Schmidt, and Stern's discussion of conceptual issues in developing ROI estimates for state VR programs and Stern, Clapp, Pepper, and Schmidt's discussion of data issues in developing valid VR-ROI estimates. Rowe, Ashley, Pepper, Schmidt, and Stern then provide an overview of the VR-ROI project and its approach to estimating ROI, and Schmidt, Clapp, Pepper, and Stern describe the results of the project's recent analyses of program impacts and ROI estimates for the state VR programs in Virginia and Maryland. The final article in this special issue by Brown and Ashley considers ROI estimation in the broader context of evaluating the performance of state workforce development systems.

As discussed by McGuire-Kuletz and Tomlinson, government agencies of all types are being asked by elected officials, the public, and the media:

> What kind of positive impact is being delivered by your agency's service, and how much are those services costing taxpayers? Return on investment (ROI) studies are one of several powerful tools for vocational rehabilitation (VR) to demonstrate relevance (2015, p. xii).

Our hope is that the articles in this special issue will serve as a resource for VR program administrators and others interested in making effective use of this powerful tool.

References

Bellante, D. M. (1972). A multivariate analysis of a vocational rehabilitation program. *Journal of Human Resources, 7*(2), 226–241. doi:10.2307/ 145185

Bua-Iam, P., Hampton, J. E., Sink, T., & Snuffer, D. W. (2013). Critical issues unique to the selection of a realistic return on investment method for public vocational rehabilitation programs. *Journal of Rehabilitation Administration, 37*(1), 25–33.

Bua-Iam, P., & Bias, T. K. (2011). Estimating worklife return on investment of WVDRS youth and older consumers with significant disabilities. *Journal of RehabilitationAdministration, 35*(1), 17–26.

Cimera, R. E. (2010). The national cost-efficiency of supported employees with intellectual disabilities: the worker's perspective. *Journal of Vocational Rehabilitation, 33*(2), 123–131.

Conley, R. W. (1969). A benefit-cost analysis of the vocational rehabilitation program. *Journal of Human Resources, 4*(2), 226–252. doi:10.2307 /144722

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

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

McGuire-Kuletz, M., & Tomlinson, P. (2015). *Return on investment and economic impact: Determining and communicating the value of vocational rehabilitation* (Institute on Rehabilitation Issues Monograph No. 38). Washington, DC: The George Washington University Center for Rehabilitation Counseling Research and Education.

Uvin, J., Karaaslani, D., & White, G. (2004). *Evaluation of Massachusetts' public vocational rehabilitation program: Final report.* Boston, MA: Massachusetts Rehabilitation Commission.

Wilhelm, S. (2013). The economic impact of Utah's vocational rehabilitation program. *Journal of Disability Policy Studies, 24*(3), 148–157.

Worrall, J. D. (1978). A benefit-cost analysis of the vocational rehabilitation program. *Journal of Human Resources, 13*(2), 285–298. doi: 10.2307/ 145363

Acknowledgements

We would like to thank the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR Grant 90DP0070) for generous funding to support the development of this special issue, the authors of the articles in this issue for their contributions, and the anonymous reviewers who provided important feedback to the authors.

What is ROI?

Kevin Hollenbeck
Upjohn Institute

Abstract. *An investment is a transaction in which an investor exchanges resources that the investor owns or controls for an asset that is expected to have value in the future. One type of investment is a vocational rehabilitation (VR) agency providing an individual with services, with the expectation that the individual will experience positive labor market or other lifestyle outcomes. The return on investment, ROI, for such an investment is calculated from the net payoff (benefits minus costs) that occurs as a result of the investment. This article lays out the basic principles used in calculating ROIs, and discusses how the ROIs for VR differ depending on the perspective: the VR agency, the individuals served, or society.*

Keywords: Return on investment (ROI), cost-benefit analysis, discount rates, present value, net present value, vocational rehabilitation, multiple perspectives about ROI

What exactly is this entity that is referred to as return on investment (ROI)? Let's take the last word first. An *investment* is a transaction in which an investor exchanges resources that the investor owns or controls for an asset that is expected to have value in the future. The transaction may be financial, in which an investor exchanges money in return for the right of ownership to an asset that is expected to increase in value in the future. The transaction may involve time, such as an individual spending the time to listen to his or her friend's issues with the expectation that the time and interest will help the friend resolve the issue. The transaction may involve program services such as a rehabilitation agency providing an individual with services, with the expectation that the individual will experience positive labor market or other lifestyle outcomes. Financial investments usually involve the exchange of one asset for another. The investment occurs at a point in time (when the rights of ownership are transferred), and the commodities that are exchanged are *stock* variables. The other types of investments mentioned in this paragraph involve the exchange of *flow* variables (time spent or provision of services to customers). The costs of flow variables require an accounting period. Note that the investor can be an individual, a group of individuals, a corporation or firm, a government agency, or even society as a whole.

The first word in ROI, i.e., *return*, is calculated from the net payoff (benefit minus cost) that occurs sometime after the investment is made. The expectation of a positive payoff is usually the *raison d'être* for the investment. As with the investment itself, the payoff may be financial or nonfinancial. An example of the former would be proceeds from the sale of a financial asset that appreciated in value. An example of the latter would be the improved self-confidence and functioning of an individual who has participated in a rehabilitation program.

Calculation of ROI requires two elements, each of which involve time. These are (1) the costs of the investment and when they occur and (2) the benefits to be received from the investment and when they accrue to the investor.

In general, cost is defined as the value of the resources that are given up in the investment transaction. Calculation of an ROI requires monetization of these costs. For financial investments, monetization is often an easy matter of knowing the price of the asset that is purchased. Of course, an investor may need to make a financial arrangement in order to make the transaction. The "cost" of any such arrangement (present value of the payments made to obtain the finances necessary for the investment minus the value of the commodity that was received) is part of the cost of the investment. For investments of time or services offered, monetization of the cost requires assumptions about the value of those entities over the period of time when the time was invested or services were rendered; that is, the accounting period.

Receiving benefits is generally the purpose of the investment. As is true from the cost side, calculation of an ROI requires monetization of the benefits. For financial investments, monetization is accomplished by knowing (or estimating) the price of the purchased asset at the particular time when the ROI is to be calculated. When benefits include nonfinancial characteristics, monetization requires assumptions about the value of the benefits received at that point in time.

In general, benefits or costs that occur in the future are considered stochastic (have an element of statistical uncertainty), and assuming that a reasonable set of probabilities can be assigned, ROI calculations will use the expected values of the benefits and costs. An ROI that is calculated at the time when the investment is being contemplated or taking place is referred to as an *ex ante* ROI. If it were to be calculated after all of the benefits and costs have occurred, then it would be an *ex post* ROI. If an investor contemplates buying a lottery ticket, the *ex ante* ROI might be $.49 on the dollar. From the lottery winner's viewpoint, the *ex post* ROI might be $1 million on the dollar. The *ex ante* ROI is useful in the decision-making process for the investment. The *ex post* ROI may be of use in gauging the success of an investment or for accountability reasons.

Oftentimes it will be the case that an investor will want an estimate of the ROI sometime after the investment has been made, but before all of the benefits or costs have occurred. In that case, the investor needs to specify whether to use only realized values of benefits and costs or to extrapolate benefits and costs beyond the data.

We typically think of investments being made in a current period and the returns on those investments accruing to the investors at a later time period. But, in general, a dollar in the future is worth less than a dollar today. This is because a dollar today can be saved and earn interest, and because the purchasing power of a dollar today is greater than a dollar in the future, assuming that there is some inflation. For a financial investment, we use interest rates to adjust for the changing time value of money. To make a fair and even comparison of the benefits and costs of an investment, we adjust the future returns with an interest rate.

In the mathematical terms displayed in the next few paragraphs, a vector is italicized and a scalar is in Roman typeface. For example, let I be an investment made in 2018 and suppose R, the return to that investment, will be received in 2019. To calculate the ROI of this investment, we need to compute a ratio that involves both R and I. But even though they are both measured in dollars, we cannot directly compute the ratio because a 2018 dollar is worth more than a 2019 dollar. In particular, the 2018 dollar is worth $(1 + r)$ times the 2019 dollar, where $r > 0$ and represents the interest rate that would be earned if the dollar were saved. Consequently, if we were to calculate the ratio of benefits to costs, i.e., R/I, from the perspective of 2018, we need to put both values in 2018 dollars by deflating R by $1/(1 + r)$. This deflation of R yields the *present Replace with* "*the equation*" that return by $[1/(1 + r)^2]$ because the investment in 2018 would have 2 years of interest and purchasing power.

The general formula for the present value of flow of monetary returns that will be received in the future that are adjusted with an interest rate r is as follows:

$$(1)\ \mathrm{PV}\,(R) = \mathrm{R}_1/(1+\mathrm{r}) + \mathrm{R}_2/(1+\mathrm{r})^2 + \mathrm{R}_3/(1+\mathrm{r})^3 + \ldots + \mathrm{R}_t/(1+\mathrm{r})^t$$

where

R= vector of future returns to be received in future periods 1 through t

R_t = return that is received in period t

r = interest rate

A more general version of this formula would allow the interest rate to vary across time periods. Because in practice, this is usually not done, and for ease of exposition, we present the less general version. If the costs of an investment—not just the benefits—flow into the future, the returns in equation (1) should be net returns, i.e., benefits minus costs. Within this framework, r is often referred to as the "discount rate" because future returns are discounted back to lower present values. An important consideration in calculating present value is the precise definition of the time period. In practice, time periods are usually monthly, quarterly, or annual.

The net present value of an investment, I, that generates a stream of future net benefits, *R*, is simply the present value of *R* minus I. The usual decision rule is that an investment is rational if its net present value is greater than or equal to 0. It is irrational to invest if the net present value is negative, a sign that the investment does not even result in a payoff that is as large as the investment.

Related to the concept of an ROI is the internal rate of return (IRR) of an investment. The IRR is the rate of interest that equilibrates the present value of the returns from an investment to the cost of the investment. In Equation (1), it is the r that would make the present value equal to the investment cost. In other words, it is the discount rate that makes the net present value equal to 0.

In general terms, the ROI at time t for an investment I is equal to the ratio of the present value of the net returns received or expected to be received in every period up until t to the investment. Note that embedded in the present value of the net returns is an assumed interest or discount rate. Equation (2) shows this general ratio:

$$(2)\ ROI_t = PV_t\,(R) / I$$

where

R = flow of actual or expected net returns from the investment

$PV_t\,(R)$ = present value of R through period t,

ROI_t = return or expected return on investment per period as calculated at time t

The investment's ROI would be calculated from equation (2) using a t that is sufficiently large that the net returns to be accrued from the investment after t are negligible. The ROI that results from equation (2) will be a decimal. It provides the investor with a net return in dollars per dollar invested. It can be expressed as a percentage by multiplying by 100. If the period used is not a year, the ROI can be converted to an annual rate. For example, if the time period is a quarter, then the annualized ROI will equal $(1 + ROI)^4 - 1.00$. Finally, it can be expressed as a payback period, i.e., the length of time it will take for the numerator to equal the denominator.

The next section of this article enumerates the unique circumstances that arise when attempting to estimate an ROI for any workforce development program, but especially vocational rehabilitation (VR). Programs typically have multiple stakeholders, and the final section discusses how ROIs may be calculated for each stakeholder group.

ROI in the VR Context

VR may be thought of as a type of capital investment, in particular, *human capital investment.* Individuals receive services from VR agencies that are intended to increase the individuals' human capital, i.e., skills and knowledge that are productive in the workforce. This article assumes that any resources spent by VR agencies on job development or other services to employers as dual customers would be appropriately apportioned to individual customers as the goal of those services is assumed to be improved labor market outcomes for individuals. The financial payoff for the individual comes from higher levels of earnings (through employment, hours, or wage rates), and in addition to financial payoffs, there are generally substantial nonfinancial or intangible benefits as well.

Calculating the ROI for VR programs is more complicated than calculating an ROI for a financial asset. That is because the VR agency is the primary investor, but the agencies' clients are the recipients of most of the benefits. That is, the beneficiaries of the investment are for the most part different from the investor. This complication results in the possibility of three separate ROIs – one for the VR agency, one for clients, and one for society as a whole. The next section discusses the issue of multiple perspectives and how these three are interrelated.

Program administrators are interested in agency ROIs because of the widespread interest in reducing government spending at all levels—federal, state, and local. In theory, a prudent investor or a policymaker focusing on their fiduciary responsibility for taxpayer funds should use ROIs to guide investment/budgetary decisions. Their marginal dollars should be invested in assets or programs that have the greatest ROI. Of course, policymakers may need to balance their fiduciary responsibilities with the goals of serving the broad interests of society. Thus, program advocates want to be able to show high ROIs in order to maintain or grow their programs.

The U.S. Department of Education noted: "Projects, initiatives and efforts should be prioritized based on the lifecycle return on investment to the agency while taking into account economic, environmental, social, and mission related costs and benefits" (U.S. Department of Education, 2011, p. 9). From the agency perspective, the costs of serving individuals represent an investment of taxpayers' funds. The return to that investment for the agency (taxpayers) comes from the taxes and other revenue generated from the additional earnings of clients who have increased their human capital and from the additional economic activity generated by those clients plus reduced social insurance costs. Hollenbeck and Huang (2003, 2006, 2014, 2016) have shown that these ROIs are usually negative.

For the individual, the investment is typically the time that is spent in interacting with the VR agency and in training. VR statutory and regulatory provisions allow state agencies to establish criteria for financial participation by individuals with disabilities in the cost of some services. So in some cases, individuals are financing their investment out of their own pocket. In general, however, these costs are not captured in state agency data, and generally have the effect of understating the total cost of purchased services data for that individual. Economists typically value this time as forgone earnings. The return to individuals, as noted above, comes in the form of increased earnings. Hollenbeck and Huang (2003, 2006, 2014, 2016) have estimated the individuals' ROIs to be positive and quite large.

The remainder of this section will examine ROI from society's viewpoint. The investment is made by taxpayers and the returns are, for the most part, garnered by clients served. The benefits that accrue to clients who receive VR services are manifold. They potentially include improved self-esteem, improved labor market outcomes, improved health outcomes, and improved quality of life. For the sake of argument, suppose these benefits could be monetized and that we call them *B*. To calculate the present value of the benefits, we slightly change equation (1) as follows:

$$(3)\ PV(B) = B_1/(1+d) + B_2/(1+d)^2 + B_3/(1+d)^3 + \ldots + B_t/(1+d)^t$$

where

PV(*B*) = present value of benefits received by client

B_t = client's benefit from the VR services received in period *t*

d = discount rate

Note that equation (3) is intended to show that the present value calculation for the receipt of benefits from program services is analogous to the present value calculation for a financial investment. The B_t terms are the monetized value of benefits received in period t. The "art" of an ROI calculation is to estimate the value of benefits, especially when intangible benefits are included. However it should be noted that a conservative approach is to use increased earnings as the B_t terms. If the ROI is positive with earnings as the only benefit received, then it would be even larger if other benefits could be monetized.

In equation (3), future benefits from the VR services provided are *discounted* at rate *d*, rather than adjusted by an interest rate *r*. The principle is exactly the same, however. Benefits are not worth as much in the future as they would be worth today. However, determining what discount rate to use in calculating an ROI is not as easy as looking up an interest rate. The discount rate *d* reflects customers' time preferences (for money). If a customer has immediate, dramatic financial concerns, then his or her discount rate will be high (he or she prefers receiving the financial benefits sooner). On the other hand, if the customer has more of a long-term perspective, then his or her discount rate will be lower. Typically ROI studies will use discount rates in the 0.03 to 0.05 range.

Suppose that a VR program in a state spends $10,000 to provide services to a customer, and then the customer earns $3,000 more than if he or she had not received the services, every year for 5 years. Further suppose that this customer's discount rate is 0.05. The present value of these services using earnings as the only monetized benefit would be $12,988.43. The net present value of the services would be $2,988.43. The ROI of the services would be 29.88% for a 5-year period, or 5.37% annually. Calculations are as follows: PV = $\$3000 * (1/(1 + 0.05) + 1/(1 + 0.05)^2 + 1/(1 + 0.05)^3 + 1/(1 + 0.05)^4 + 1/(1 + 0.05)^5) =$ $12,988.43. NPV = $12,988.43 - $10,000 = $2,988.43. ROI for 5 years = $2,988.43 / $10,000 = 29.88%. Annual ROI = $1.2988^{.2} = 1.0537$. The IRR for this investment is 15.24% annually. This means that the investment is rational for taxpayers as long as the discount rate is less than or equal to 15.24%.

Multiple Perspectives

As noted, VR can be thought of as a co-investment between the customer and the agency. The costs of providing services are generally borne by the agency, although in some cases, the customer may bear part of the financial costs. Again, this article assumes that resources spent by VR agencies on job development or other services to employers as dual customers would be appropriately apportioned to individual customers as the goal of those services is assumed to be improved labor market outcomes for individuals. More generally, participants are investing their time and effort when they receive services, which is valued by using forgone earnings, using the individuals' actual wages if they are employed or imputed wages based on education and other characteristics. As noted above, the benefits to the customer may include increases in earnings as well as other intangible benefits. The benefits to the state agency are the potential of increased tax receipts or reduced social insurance costs. With two investors, it is possible to estimate two separate ROIs (as well as a single ROI that combines the two investors.)

Table 1 summarizes the costs and benefits of VR services for customers and the government. Note that the government column is labeled as being the agent for taxpayers. So government expenditures are assumed to be financed by taxpayers and thus are costs to them; government revenues are assumed to benefit taxpayers. In addition, the table has a column that adds the first two. The final row of the table shows the net benefits to each of the parties and is derived by summing the columns. The entries in the table represent the expected monetary costs (-) or benefits (+) to the group.

The top two rows of the table show the investment costs for the entities. The costs of direct program services, including purchased services, and stipends are in the first row of the table. Stipends are a benefit for participants, although they may be lessened by any payments as noted in footnote 8. The program services are beneficial for customers, but their monetary value comes from any labor market or intangible outcomes that result from the services. So the plus in that entry are relatively small. These services and stipends are unambiguously a large investment cost for the government and taxpayers. So the first row's entry for the second and third columns are negative. The second row of the table shows that participants are investing their time and effort in receiving education and training or other services. This opportunity cost does not affect the government (or taxpayers).

Rows 3-8 of the table show the potential benefits to customers and the government. Receiving program services increases the human capital of customers and leads to better labor market outcomes. When individuals become employed, they become productive members of the workforce. If program participants are incumbent workers, then program activities will improve their productivity. In rows 3 and 4, we show that rehabilitated/ rehabilitating workers receive higher earnings (through increased employment, wages, or hours) as well as higher fringe benefits.

In the fifth row, we show that increased levels of employment and turnover due to skills learned or accommodations received. We indicate that this is a cost to program participants because they are losing nonwork or leisure time, although it is likely to be a small cost that customers don't even recognize. The sixth row shows that customers may lose unemployment compensation benefits or lower social insurance or disability program benefits (DI, SSI,

Table 1

Components of an ROI Analysis

Benefit or cost	Customers	Government[a]	Society[b]
Investment			
1. Direct and purchased services/stipends	+	-	-
2. Time spent in education/training or receiving other services	-	0	-
3. Higher earnings	+	0	+
4. Fringe benefits	+	0	+
5. Less leisure time with increased employment	-	0	-
6. Lower unemployment/Social insurance/ disability payments	-	+	0
7. Income/payroll taxes	-	+	0
8. Value of intangible	+	0	+
Net returns			
9. Total	+	-	+/0/-

Note. [a] including VR, acting as agent of taxpayers; [b] Sum of columns 1 and 2

Medicaid, SNAP, TANF, and others). The reduction in these benefits is a positive for government agencies. In general, we assume that transfers such as these net out to be zero.

With higher levels of earnings and employment come higher tax liabilities. These are denoted in row 7. Workers will pay higher payroll taxes, which is a cost to them. The government receives the tax revenue, which is a plus for it. As with the transfers in the sixth row, these additional tax payments are assumed to offset.

The eighth row, which is often ignored in cost-benefit or ROI analyses shows that customers receive benefits that are not easily valued such as improved self-esteem, improved health, improved housing situations, or others. As noted, the conservative approach to calculating ROIs is to ignore this row (or equivalently assume it is zero). This will underestimate the true ROI.

The table shows that the expected net benefits, and thus ROIs, for VR participants are positive. The stipends that they receive and their increased earnings (net of taxes) will exceed their time costs and reduced social insurance benefits. The expected net benefits, and thus ROIs, for the government are negative, however. The direct cost of providing services is unlikely to be totally offset by eventual increased payroll taxes, or reductions in social insurance benefits.

When one looks at the ROI for the third column, i.e., the government is the investor and the customers are the beneficiaries of the investment, the result cannot be unambiguously predicted. The ROI will be positive if the sum of the after-tax labor market outcomes and value of the intangibles exceed the costs of providing rehabilitation services and stipends.

References

Hollenbeck, K., & Wei-Jang Huang, W. -J. (2003). *Net impact and benefit-cost estimates of the workforce development system in Washington state*. Kalamazoo, MI: W. E. Upjohn Institute for Employment Research.

Hollenbeck, K., & Huang, W. -J. (2006). *Net impact and benefit-cost estimates of the workforce development system in Washington state*. Kalamazoo, MI: W. E. Upjohn Institute for Employment Research.

Hollenbeck, K., & Huang, W. -J. (2014). *Net impact and benefit-cost estimates of the workforce development system in Washington state*. Kalamazoo, MI: W. E. Upjohn Institute for Employment Research.

Hollenbeck, K., & Huang, W. -J. (2016). *Net impact and benefit-cost estimates of the workforce development system in Washington state*. Kalamazoo, MI: W. E. Upjohn Institute for Employment Research.

U. S. Department of Education. (2011). *Strategic sustainability performance plan.* Washington, DC: Author. Retrieved from https://www2.ed.gov/about/reports/strat/sustainability/edstrat plan 2011.pdf

Author Notes

Kevin M. Hollenbeck has written extensively about the return on investment of workforce programs and is an Emeritus Senior Economist for the W. E. Upjohn Institute for Employment Research. The author gratefully acknowledges the helpful comments received from Bob Schmidt and two anonymous referees. The usual caveat applies.

Social Return on Investment: An Important Consideration for State Vocational Rehabilitation Programs

Betsy Hopkins
National Disability Institute

Abstract. *A more traditional method of estimating the ratio of net benefits and costs or a return on investment (ROI) is one way for agencies to indicate the financial benefits that these programs realize. In contrast, social return on investment (SROI) is a relatively new approach that can be used to help determine the value of social benefits gained by individuals with disabilities who obtain competitive integrated employment (CIE). This paper will discuss and give examples of the benefits of specific SROI methods in assessing employment outcomes for individuals with disabilities, which may be particularly important in estimating the overall value of the public vocational rehabilitation (VR) program.*

Keywords: Return on investment (ROI), social return on investment (SROI), cost benefit analysis, vocational rehabilitation, competitive integrated employment, social value

In a time of dwindling resources, federal and state agency administrators, along with the taxpayers who fund those agencies, want to ensure that funding is provided to entities that will minimally meet or preferably exceed program expectations and outcomes while staying within their budgets. One way for agencies to show how they are doing this is to estimate the ratio of the net benefits and costs involved with these programs, often called the return (benefits) on investment (costs) or ROI. Extant ROI analyses typically limit the estimation of benefits to items or activities that are measured financially. This paper will discuss the relationship between and importance of including social as well as financial benefits in estimating the value of the public vocational rehabilitation (VR) program.

As someone who has worked in the field of rehabilitation for over 25 years, and most recently as the director for the Maine general VR agency, I have firsthand experience with the need to be able to show taxpayers, legislators, and others that the program had a positive impact. While Maine VR never developed formal ROI estimates, we were able to show positive results. For example, the 758 individuals who exited our program in CIE during federal fiscal year 2017 collectively earned over $13,500,000 in annual wages, with their weekly wages from application to closure increasing an average of $215 (Maine Department of Labor, 2017).

This kind of information, while important in demonstrating the **benefits** of the public VR program to participants, does not include consideration of the **costs** associated with serving those individuals. When speaking to individuals, family members, or providers of services of individuals with disabilities, particularly those with the most significant disabilities, it would be helpful to be able to show the social as well as the financial benefits of working in CIE. Using both methods could help encourage individuals by raising expectations so they and their families believe that they will succeed in CIE the same as anyone else. Measuring and showing positive social gains for individuals with significant disabilities who work in CIE has mostly been done to date by sharing individual success stories of those who talk about the differences working in CIE has made in their lives. These stories are important; however, developing and using a more formal approach to estimating the social return on investment (SROI) could help clearly measure social gains to show another side of the effectiveness of the public VR program.

Historical Perspective

In the United States, the passage of the Rehabilitation Act of 1973 and subsequent amendments to the Act in 1978, 1986, and 2014 ensured that the funding streams and the majority of available resources and services were to be funneled toward people with more significant disabilities within the public VR programs (Rubin & Roessler, 1987). Two important pieces of civil rights legislation for people with disabilities were the Americans with Disabilities Act of 1990 (ADA) and the Rehabilitation Act Amendments of 1992. These critical pieces of legislation represented the start of a philosophical and practical view of people with disabilities as individuals valued by society who should be able to realize full civil rights under the law. Specifically, these Acts included language that encouraged full participation for individuals with significant disabilities in rehabilitation programs and access to employment in their communities (Rubin & Roessler, 1987). Even with the passage of these important pieces of legislation and the increased focus of supports for people with significant disabilities in employment, only 18.7 percent of working age people with disabilities were employed in 2017. In contrast, the employment-population ratio for those without a disability was 65.7 percent (Bureau of Labor Statistics (BLS), 2018).

Return on Investment (ROI)

When estimating an ROI for a state agency such as a public VR program, the "financial benefits or gains when receiving services from VR often take the form of increases in earnings that accrue because customers become employed, change jobs, get increases in hours of employment, increases in wage rates and/or increases in benefits" (McGuire-Kuletz & Tomlinsen, 2015, p. 11). When considering the costs associated with an individual going through a state VR program, there is the investment cost, which is the resource cost of initiating the investment, and then there are the ongoing costs that occur after the investment is made (McGuire-Kuletz & Tomlinsen, 2015). Again, using VR as an example, a participant may have received training and time with a VR counselor to oversee their program (investment cost), and then be referred to a vendor for assistance in finding a job and job coaching (ongoing costs).

One challenge with solely using financial cost benefit analyses or strict ROI formulas may be that they do not take into account the added social value that is crucial to consider when evaluating an organization, program, or strategy that supports the inclusion of individuals with the most significant disabilities in CIE (Sturmey & Didden, 2014). Cimera (2000) points out that on the benefit side of a cost benefit analysis (CBA), some of the benefits will have "obvious dollar values (e.g., wages earned); others will not (e.g. increased self-worth). In a CBA, program outcomes that do not translate into monetary values are not included and this is often cited as a weakness of this type of evaluation" (Cimera, 2000, p. 146).

Social Return on Investment (SROI)

The concept of SROI has developed over the last 20 years. It was first articulated in the United States in 1996 by Jed Emerson at the Roberts Enterprise Development Foundation (REDF) (Nicholls, 2016). Some reasons given by REDF for the development of SROI concepts and methods included the need to answer questions such as: "how do we as practitioners and philanthropists know whether we are accomplishing the goals we set for ourselves, how can we continue to make informed decisions about the on-going use of our resources, and how can we determine that funding is truly having an impact that results in quantifiable benefits to individuals and to society?" (Nicholls, 2016, p. 128). REDF felt the need to think about measurement of their programs beyond the use of a more traditional CBA to address these questions.

In 2004, with a grant from the Hewlett Foundation, a small working group of SROI practitioners in the United States and Europe broadened the scope of SROI and developed a number of standards for accountability, including the Principles Standard and Stakeholder Engagement Standard. These standards helped SROI methods move away from the more traditional CBA and ROI in that they incorporated the importance of stakeholder involvement in determining the pertinent issues related to specific social gains to be measured (Nicholls, 2016). this ultimately resulted in new guidance where SROI was defined as "a process of understanding, measuring and reporting on the social, environmental and economic value created by an organization" (Scholten, 2006, p. 12).

In determining a simple SROI today, one would attempt to account for the nonfinancial benefits and costs not often included in a traditional ROI estimate. "An SROI puts those affected at the heart of the process and uses a mix of qualitative and quantitative data. SROI is seen as an approach that monetizes outcomes, which can then be aggregated to create a measure of efficiency with which resources are turned in to value" (Nicholls, 2016, p. 127). Some examples of nonfinancial benefits might include the areas found in the eight life domains for considering quality of life developed by R. Schalock (Brown, Hatton, & Emerson, 2013): emotional well-being, interpersonal relations, material well-being, personal development, physical well-being, self-determination, social inclusion and rights.

There are no current SROI methods created specifically to measure the social returns for people with disabilities served by the public VR program. However, there is a SROI guide which came out of the work of the new economics foundation (NEF) and the development of the European SROI Network, formed in 2006 to help practitioners share their experiences (Nicholls, 2016). This guide, developed by the Scottish Government (Social Value UK, 2012), lays out a step-by-step method which could conceivably be applied to measuring a variety of social returns. This guide sets out six stages with a chapter devoted to each stage for conducting an SROI analysis:

- Establish scope of analysis and identify key stakeholders
- Map inputs, outputs and outcomes
- Identify outcome indicators and give them a value
- Establish impact of the program/activity
- Calculate the SROI
- Report and share findings with stakeholders and embed outcome processes in order to verify the report

In addition to the stages listed above, this guide uses a set of core principles when developing and carrying out an SROI: Involve stakeholders, understand what changes, do not over-claim, only include what is material or relevant, value the things that matter, be transparent and verify the result (Cabinet Office, 2009). Based on the work of Nicholls (2016), Table 1 compares these principles to those of CBA or ROI.

Valuing Social Benefits

Perhaps the most difficult part of conducting an SROI is doing the valuation of social benefits. A promising technique that has the added advantage of involving stakeholders is contingent valuation. Used widely in environmental economics and real estate appraisals (see, for example, Carson, 2012; Hoehn & Randall, 1987; Schkade & Payne, 1994), contingent valuation is a survey-based approach to determining willingness to pay or willingness to accept certain non-market situations. Essentially, individuals are surveyed and asked to place a value on a certain events or situations (e.g., avoidance of environmental degradation). What is suggested here is that stakeholders could be asked to value social benefits such as the feelings of self-worth that might be accomplished with competitive integrated employment. The use of contingent valuation relies on appropriate survey techniques: avoidance of strategic response behavior and constraining responses to realistic values. This approach, while promising, has not yet been used to estimate the value of potential social returns for the public VR program.

Conclusions

The overarching themes of a recent review of the literature on employment services for individuals with disabilities (Gidugu & Rogers, 2012), and the continued significant underemployment of people with disabilities across the country (BLS, 2018), point to the need to improve employment services, opportunities, and outcomes for individuals with disabilities. There are discrepancies in the field about how best to do that. Overall, it appears that the financial and social benefits (for individuals, their families and society) outweigh the alternatives, when individuals with significant disabilities are working in their communities.

Reliable ROI methods can show how quality employment for individuals with significant disabilities benefit consumers, families, taxpayers and society as a whole. In addition to ROI methods, SROI methods can also be used to measure and support social gains that are made by individuals with significant disabilities working in competitive integrated settings. There is some concern that public policy makers, state and federal agencies may use a financial ROI method to ascertain that the financial investment in helping individuals with the most significant disabilities gain independence does not produce a good enough financial return to warrant funding agencies to do this type of work.

Research has shown a clear connection between employment, economic self-sufficiency, and better health. In addition, there is a connection between em-

Table 1
Comparison of Principles for SROI to CBA/ROI

Social Value Principles	CBA/ROI	SROI
Involve stakeholders	No stakeholder involvement required in determining value, although stakeholders may be consulted	Stakeholder involvement is key to help in determining which activities will be measured and how much value to place on each
Understand what changes	Focus on impact of change on financial wellbeing.	Also focuses on changes in financial well-being, however it goes a step further by giving value to things that are routinely left out of more traditional economic appraisals
Do not over claim	Documents sources and assumptions made in valuing benefits and costs; often conservatively overestimates costs and underestimates benefit	Once the activities are chosen that will be measured, then a monetary value is attached. This monetary value can be somewhat subjective. As much as possible, it is important to choose credible financial proxies for different indicators and only count them once.
Only include what is material	Documentation of sources and assumptions rationalizes materiality and relevance.	Clearly outlining the initial scope of the project and then consulting with stakeholders help to ensure material/ relevant factors are considered
Value what matters	Money is used as a measure of the relative value of outcome	It is important to choose the activities that matter the most to the stakeholders impacted by the change rather than those easiest to measure
Be transparent	Requires transparency in order to be considered accurate - financial information used in CBA and ROI formulas must be backed up by agency audit and fiscal reporting requirements	All decisions (regarding identification and involvement of stakeholders, sources and methods of data collection, indicators and benchmarks, outcomes, and communication of results) must be explained and documented.
Verify the result	Estimate usually uses verifiable information (e.g., costs of services), although sometimes includes projections of economic or employment benefits into the future (which may not be readily verifiable)	There must be appropriate independent assurance of an SROI analysis by checking that the principles of good practice (including how data was collected and assessed) in SROI were followed

ployment and community integration for people with disabilities (Miller, Molina, Grossman, & Golonka, 2004). Furthermore, community integration is strongly tied to increased quality of life, (Brown et al., 2013) and employment services are more beneficial than any other service in supporting community integration (Brown & Brown, 2009). This is another reason to consider the broader social benefits by using SROI methodology in addition to ROI methodology to show how both financial and social gains positively impact individuals with significant disabilities, their families, communities and employers when work in competitive integrated work settings is encouraged and supported.

The Rehabilitation Act of 1973 states “disability is a natural part of the human experience and in no way diminishes the right of individuals to live independently, enjoy self-determination, make choices, contribute to society, pursue meaningful careers, and enjoy full inclusion and integration in the economic, political, social, cultural, and educational main-stream of American society" (Rehabilitation Act of 1973, Section 701 (a) (3)). Under this act, public VR programs are tasked with providing services that help individuals with disabilities enjoy and gain "full inclusion and integration" into society. In order to show that these programs are successful, further exploration and study into the use of SROI methods will help create sustainable evidence-based practices.

Regardless of whether people have disabilities, the social benefits of work are similar. Working in a job that has meaning can provide a sense of self-worth, of belonging by being included in a community of others who have the same or similar goals, and can provide an opportunity for growth, learning, and challenge that improves connections in society and community that one may not otherwise have (Canadian Association for Community Living, 2014). Even if one's current job is not the ultimate career goal, earning wages in the right job can give people a sense of pride, purpose, and goals and set them on a path to maintaining self-sufficiency (Brown et al., 2013). With the continued movement toward integration of people with disabilities in communities and employment settings, in addition to understanding the financial benefits that inclusion can bring, measuring social benefits can help policy makers, family members, agency providers and others more clearly understand why individuals with significant disabilities can benefit from being included in competitive integrated employment.

References

Brown, I., & Brown, R. I. (2009). Choice as an aspect of quality of life for people with intellectual disabilities. *Journal of Policy and Practice in Intellectual Disabilities, 6*(1), 11–18.

Brown, I., Hatton, C., & Emerson, E. (2013). Quality of life indicators for individuals with intellectual disabilities: Extending current practice. *American Association on Intellectual and Developmental Disabilities, 51*(5), 315–332.

Bureau of Labor Statistics, U. S. Department of Labor. (2018, June 21). *Persons with a disability: Labor force characteristics–2017* [Press release]. Retrieved from https://www.bls.gov/news.release/pdf/disabl.pdf

Cabinet Office, U. (2009, April). *Guide to social return on investment.* Retrieved from http://www.socialvalueuk.org/resource/a-guide-to-social-return-on-investment-2012/

Canadian Association for Community Living. (2014, October). *Ready, willing, and able, Volume 4, Issue 2.* Retrieved from Canadian Association for Community Living: http://www.cacl.ca/sites/default/files/Vol4Issue22014Final.pdf

Carson, R. T. (2012). Contingent valuation: A practical alternative when prices are not available. *Journal of Economic Perspectives*, *26*(4), 27–42.

Cimera, R. E. (2000). Improving the cost efficiency of supported employment programs. *Journal of Disability Policy Studies*, *11*(3), 145–151.

Gidugu, V., & Rogers, E. (2012). *Review of employment services for individuals with intellectual and developmental disabilites: A comprehensive review of the state-of-the-field from 1996-2011*. Retrieved from Boston University Center for Psychiatric Rehabilitation Disability Research Right to Know: http://www.bu.edu/drrk/research-syntheses/developmental-disabilities/employment-services

Hoehn, J. P., & Randall, A. (1987). A satisfactory benefit cost indicator from contingent valuation. *Journal of Environmental Economics and Management*, *14*(3), 226–247.

Maine Department of Labor (2017). *Bureau of Rehabilitation services highlights.* Retrieved from https://www.maine.gov/rehab/brs_highlights/index.shtml

McGuire-Kuletz, M., & Tomlinsen, P. (2015). *Return on investment and economic impact: determining and communicating the value of vocational rehabilitation.* Washington D.C.: The George Washington University Center for Rehabilitation Counseling Research and Education.

Miller, J., Molina, F., Grossman, L., & Golonka, S. (2004). *Building bridges to self-sufficiency: Improving services for low-income working families.* Maryland: NGA Center for Best Practices.

Nicholls, J. (2016). Social return on investment: Development and convergence. *Evaluation and Program Planning*, *64*, 127–135.

Rehabilitation Act of 1973, 29 U.S.C. §§ 701 et seq. (2016).

Rubin, S. E. , & Roessler, R. T. (1987). *Foundations of the vocational rehabilitation process* (3rd ed.). Austin, TX: Pro-Ed.

Schkade, D. A., & Payne, J. W. (1994). How people respond to contingent valuation questions: A verbal protocol analysis of willingness to pay for an environmental regulation. *Journal of Environmental Economics and Management*, *26*(1), 88–109.

Scholten, P. N. (2006). *Social return on investment: A guide to SROI analysis.* Amsterdam, Netherlands: Lenthe Publishers.

Social Value UK (2012). *A guide to social return on investment*. Retrieved from http://www.socialvalueuk.org/resource/a-guide-to-social-return-on-investment-2012/

Sturmey, P., & Didden, R. (2014). *Evidence-based practice and intellectual disabilites.* New York, NY: John Wiley & Sons.

Author Notes

Betsy Hopkins, PhD candidate, is a Training and Technical Assistance Manager for the Workforce Innovation and Technical Assistance Center with the National Disability Institute, Inc. in Washington, DC.

Journal of Rehabilitation Administration, 40(1), pp. 17–22, 2019

The Return on Investment of Vocational Rehabilitation Services: Some Ethical Considerations

Robert Froehlich
Olivia Bentley
Diona Emmanuel
Maureen McGuire-Kuletz
The George Washington University

Abstract. *Ethical codes and principles guide research and practice. From the feedback provided by its Advisory Council, the Vocational Rehabilitation Return on Investment (VR-ROI) Project recommends eight ethical considerations to guide the analysis, interpretation, and application of ROI data. These considerations are grounded in various principles and sections of the Commission on Rehabilitation Counselor Certification's (CRCC) Code of Professional Ethics for Rehabilitation Counselors. The CRCC Code and suggested ethical decision-making tools provide a framework for Vocational Rehabilitation administrators and decision-makers.*

Keywords: Professional codes of ethics, ethical tests, rehabilitation counselors, return on investment

Return on investment (ROI) is a term that has been used in both the private and public sector to describe the eventual payoff for spending money (McGuire-Kuletz & Tomlinson, 2015). The public vocational rehabilitation (VR) program is an integral part of the statewide workforce development system that provides a financial ROI and other benefits to the individuals it serves as well as to society at large (United States Department of Education, 2017). With the implementation of the 2014 Workforce Innovation and Opportunity Act (WIOA) and its expanded common performance measures, there is a greater focus on evaluating program efficacy, particularly regarding the cost of such programs relative to their benefits (United States Department of Education, 2014).

In response to WIOA and the growing emphasis on program evaluation as a way of showing a financial benefit, ROI methods are applied to illustrate the monetary value of programs like VR. As discussed elsewhere in this special issue, the Vocational Rehabilitation Return on Investment (VR-ROI) Project has developed rigorous methods for estimating ROI based on readily available administrative data from multiple sources. Discussions with participating state VR agencies and the VR-ROI Project Advisory Council regarding the project's approach to estimating ROI, and appropriate uses of the estimates, led the project team to recognize the importance of explicitly addressing various ethical considerations associated with this type of program evaluation. This article describes the VR-ROI Project's recommendations relative to the ethics of analysis, interpretation, and application of ROI data for VR administrators and decision-makers.

Ethical Codes

Successful VR brings individuals with disabilities into a collaborative relationship with an interdisciplinary team which can be comprised of counseling, medical, school, and allied health professionals. Business sector information and services (career development, employment issues, and labor market information) must be considered by these professionals when providing services. Most professions and organizations within academia, the private sector, government, and nonprofit sectors have individual codes of ethics, but the purpose of the codes remains the same. A code of ethics establishes a set of standards and principles for a profession or organization. Ethical codes are important in any profession because they inform decision-making processes and provide guidance and safeguard a profession by outlining responsibilities, values, and standards for individuals and organizations (Jamnik, 2011). Since the VR-ROI Project integrates various disciplines, multiple codes are relevant to the ethical considerations of the Project including those from the business, evaluation, and counseling disciplines.

Ethical codes vary by organization, and wider dissemination of the benefits of adopting these codes is necessary to increase acceptance of business ethics (Boda & Zsolnai, 2016). Financially based organizations have policies and procedures regulating ethical business. All employees of the United States Food and Drug Administration (FDA) are expected to follow the Code of Ethics for Government Service which was adopted by Congress in July 1958 (United States Department of Health and Human Services, 2017). Some relevant business and evaluation-based codes include the American Economic Association, the American Evaluation Association, the American Institute of Certified Public Accountants (AICPA), the Certified Financial Planner Board (CFPB), the International Economic Development Council (IEDC), and the International Federation of Accountants (IFAC) with its' International Ethics Standards Board for Accountants (IESBA). Among these business and evaluation-based codes, integrity is a common principle listed (American Economic Association, 2018; American Evaluation Association, 2018; AICPA, 2014; CFPB, 2018; IEDC, 2015; IESBA, 2016). Other guiding principles listed by most organizations include professionalism, objectivity, fairness, trust, and honesty (American Economic Association, 2018; CFPB, 2018; IESBA, 2016). Confidentiality and protecting the public interest are listed as important by several organizations (CFPB, 2018; IEDC, 2015; IESBA, 2016).

Multiple professional codes of ethics are also available for counseling and related professionals. These include: American Counseling Association, American Psychological Association, Commission on Rehabilitation Counselor Certification (CRCC), and National Association of Social Workers (Ivey, Ivey, & Zalaquett, 2018). The CRCC Code of Professional Ethics for Rehabilitation Counselors (CRCC Code) is most directly related to the operation of the VR program. The CRCC Code addresses issues likely to arise for VR professionals when applying employment-related techniques in their work with individuals with disabilities. These techniques include career counseling, adjustment to the medical and psychosocial impact of disability, program evaluation and research, and employment-related tasks such as job analysis, development, and placement services, job accommodations, and rehabilitation technology (CRCC, 2017).

Application of CRCC *Code* and Ethical Principles to ROI

When an ROI approach is used, rehabilitation professionals should consider some essential concepts including the ethical principles that form the foundation of the CRCC *Code*. The principles include autonomy, beneficence, fidelity, justice, nonmaleficence, and veracity. The *Code*'s core principles echo and elaborate on principles found in codes used in the business and evaluation sectors. Autonomy refers to respecting the client's choice. Beneficence means prioritizing client interests above all. Fidelity refers to trustworthiness and being faithful. Justice refers to the equitable use of available resources. Nonmaleficence ensures that results are not used in a manner that is harmful to clients, and veracity refers to being honest (CRCC, 2017). The codes and underlying principles are helpful tools when making decisions relative to scarcity of resources, confidentiality, evaluation and model selection and development, communication with other professionals, and documentation. When conducting ROI evaluations, it is important to utilize ethical decision-making models, and to consider various sections of the CRCC *Code*.

The analysis, interpretation, and application of VR-ROI estimates require ongoing and discerning feedback from multiple stakeholders. The VR-ROI Project approach relies heavily on feedback from multiple project constituencies in order to gain access to data, analyze that data in the correct context, and present the estimates using dissemination practices grounded in the *Code's* six core ethical principles. One important feedback mechanism is the VR-ROI Advisory Council, a group consisting of constituents with disabilities, administrators of VR Programs, State Rehabilitation Council members, economists, and policy experts. Throughout the life of the project, the Advisory Council has provided feedback on the ethical use and application of ROI data and estimates. Educating others about the VR-ROI approach can help guide future projects. Some ethical considerations can be found below. It is important to be aware that these ethical considerations not only apply to ROI estimates; they should be explored with any research/evaluation findings.

Do not use ROI to Screen Out Individuals

ROI estimates should not be used as a screening tool for services. While estimates can be analyzed to identify a potential profile of services based on a certain set of variables, this profile should not be considered a recipe for all service recipients. Consistent with the ethical principle of justice, ROI estimates are intended to best direct resources toward effective rehabilitation options. As outlined in Section A.1 of the CRCC *Code*, results should be applied in a manner that is respectful of clients and promotes their best interests. Diversity and the avoidance of bias should be prioritized in rehabilitation counseling research (CRCC, 2017). ROI estimates can be a powerful tool for justifying services so it is important they are applied in an ethical manner.

Use Must be Compatible With the Individualized Nature of VR Services

Rehabilitation outcomes are achieved and considered successful when an individualized approach is used, incorporating strengths and accommodations, and addressing barriers and functional limitations, as outlined in Section A.1.b of the CRCC *Code*. This approach keeps to the principle of autonomy by respecting a consumer's right to influence the development of their counseling plan (CRCC, 2017). Using a lock-step type of recipe or formula for identifying a combination of services is not an effective use of ROI estimates and tools. Rather than using a standardized approach, it is important to consider the characteristics of the consumer. Theories and assessments based upon observations are most effective relative to generating discussion about options. ROI estimates are no different. Any profile of services should meet the needs of the consumers.

Report Findings Accurately to Ensure Credibility

ROI estimates are calculated to tell the entire story behind observed occurrences. They must be reported even in instances where negative rates of return are observed. Identifying the reasons behind positive and negative returns meets the ethical principles of beneficence and nonmaleficence (CRCC, 2017). The CRCC *Code* recommends that results not be misused, and are reported accurately even if they are not favorable (2017; Sections G.2.a, I.3.a, and I.3.b). If it is determined that mistakes have been made, steps must be taken to correct them (Section I.3.c).

Obtain Feedback From Key Stakeholders

Key stakeholders play a significant role in providing feedback for a project. The VR-ROI Project seeks feedback and consultation with multiple entities, including economists, evaluation experts, legislators, administrators, field staff supervisors, direct service providers, consumers, and family members. The CRCC *Code* recommends that researchers who do not have oversight from an Institutional Review Board (IRB) should seek consultation (CRCC, 2017; Section I.1.e). Because IRB membership typically does not represent the full range of stakeholders for any given research project, we believe that such consultation should be sought even when one or more IRBs are involved, especially since the focus of IRBs is on protecting the rights and welfare of human participants in research. Involving stakeholders is consistent with the commitment to upholding values outlined in the CRCC *Code*, including "enhancing the quality of professional knowledge and its application to increase professional and personal effectiveness" (CRCC, 2017, p. 5). Section E.2 of the CRCC *Code* provides guidelines for organization and team relationships including the use of interdisciplinary teams, decision-making, documentation, and the inclusion of consumers as team members. Involving stakeholders is also consistent with multiple ethical decision-making models that indicate the importance of gathering external feedback or consultation (Corey, Corey, Corey, & Callahan, 2014; Cottone, 2001; Herlihy & Watson, 2007).

Select ROI Method Carefully Based Upon Needs

There is a need for continuous dialogue and consultation when reviewing and presenting estimates. Selection of ROI methods differs depending on the topic and variables. Careful consideration is needed when choosing a specific ROI method. A quicker calculation for some trends might be appropriate (for example an initial exploration or evaluation of a service or topic), or a more rigorous approach might be necessary (to make a more definitive statement about a trend). It is important to acknowledge limitations associated with the chosen method. Researchers or evaluators must be competent in understanding and applying the chosen method as supported by the ethical principles of justice, fidelity, and veracity outlined in Section D.1 of the CRCC *Code* (2017).

Evaluate Results to Avoid Broad Generalizations

The generalizability of estimates needs to be explored. For example, if the data suggest that there is a much lower rate of return for males receiving education services than for females, it would be inappropriate to decide to no longer provide such services to men. Gender may be related, however other plausible factors may include differences in regional public school, pre-application earnings, or other regional labor market issues. Before implementing policy changes, thoughtful consideration of all factors and findings need to be examined. There are a variety of methods for avoiding broad generalizations including considering the impact of applying any ROI finding to a population, and carefully communicating results. All six ethical principles apply to this consideration because they promote "the fundamental spirit of caring and respect" of all individuals (CRCC, 2017, p.5).

Develop Safeguards to Investigate Unanticipated Results

When results from analyses seem to not make sense, it is important to investigate the underlying meaning. ROI estimates may highlight a trend that a program is meeting a need unrelated to the original charge. For instance, a state VR agency may be providing disproportionate funding for a medical need (for instance hearing aids or eyeglasses) that cannot be covered by other social service and allied health agencies. In order to avoid inaccurate assumptions, it is important to identify the context of such trends and engage in dialogue with decision-makers. This is consistent with the ethical principles of nonmaleficence and veracity, and Sections C.1, D.1.d, and I.3.a of the CRCC *Code* addressing advocacy, avoiding harm, and reporting accurate results.

Do not Confuse Settings

Failure to acknowledge a false equivalence between ROI for business investments versus workforce training programs can lead to a misuse of ROI estimates. For example, VR provides important services that can greatly increase a person's quality of life. In such cases, the financial cost of providing the services may outweigh the financial benefits to individuals or to society at large. Failing to consider the value of non-monetary positive outcomes associated with VR programs, such as quality of life and increased independence, ignores the fact that such programs are not meant to serve a sole business investment function. To ignore positive outcomes that are difficult to measure is incongruent with CRCC *Code* Section C: Advocacy and Accessibility. Applying such a false equivalence contradicts Section C imperatives such as "promote opportunity and access, improve the quality of life of individuals with disabilities, and remove potential barriers to the provision of or access to services" (CRCC, 2017, p. 12).

Use of Ethical Tests

There are several resources to aid in the ethical decision-making process. Professionals should consult their relevant codes of ethics and seek consultation when encountering ethical dilemmas. Ethical decision-making models and ethical tests can be utilized to provide guidance. Questions include "Is the choice I propose legal?" and "Does this choice comply with our agency's values?" There are various ethical tests (Stadler, 1986); these include universality (Would you recommend the chosen course of action to someone else?); publicity (How would you feel if the chosen course of action was reported in your local newspaper?); and justice (Would you want to see everyone treated in the same manner as you selected?). When making decisions, it is important to be guided by an ethical decision-making tool. More information can be found on the CRCC website at https://www. crccertification.com/decision-making-models.

Conclusion

Ethical codes are important, and can provide a framework for decision-making. The VR-ROI Project identified ethical considerations after engaging in conversations with Advisory Council members on the application of findings. The CRCC *Code* provides guidance for understanding and applying the ROI findings in accurate and responsible manner. The proposed ethical considerations focus on the goals of promoting and protecting consumers' best interests. The VR-ROI Project puts forth these ethical considerations to help VR professionals understand the power of ROI findings and the need for ethically applying information.

References

American Economic Association. (2018). *AEA code of professional conduct*. Retrieved from https://www.aeaweb.org/about-aea/code-of-conduct

American Evaluation Association (2018). Preface to evaluators' ethical guiding principles. Retrieved from https://www.eval.org/p/cm/ld/fid=51

American Institute of Certified Public Accountants. (2014). Code of professional conduct. Retrieved from https://www.aicpa.org/research/standards/code of conduct.html

Boda, Z., & Zsolnai, L. (2016). The failure of business ethics. *Society and Business Review, 11*(1), 93–104.

Certified Financial Planner Board. (2018). *Code of ethics & professional responsibility.* Retrieved from https://www.cfp.net/for-cfp-professionals/professional-standards-enforcement/current-standards-of-professional-conduct/standards-of-professional-conduct/code-of-ethics-professional-responsibility

Commission on Rehabilitation Counselor Certification. (2017). *Code of professional ethics for rehabilitation counselors.* Retrieved from https://www.crccertification.com/filebin/pdf/ethics/CodeOfEthics_01-01-2017.pdf

Corey, G., Corey, M. S., Corey, C., & Callahan, P. (2014). *Issues and ethics in the helping professions* (9th ed.). Belmont, CA: Brooks/Cole.

Cottone, R. R. (2001). A social constructivism model of ethical decision-making in counseling. *Journal of Counseling Development, 79,* 39–45.

Herlihy, B., & Watson, Z. E. (2007). Social justice and counseling ethics. In C. C. Lee (Ed.), *Counseling for social justice* (2nd ed., pp. 181–199). Alexandria, VA: American Counseling Association.

International Economic Development Council. (2015). *International economic development council code of ethics*. Retrieved from https://www.iedconline.org/web-pages/inside-iedc/iedc-code-of-ethics

International Ethics Standards Board for Accountants. (2016). *Handbook of the code of ethics for professional accountants.* Retrieved from http://www.ire.lu/fileadmin/media/Env_normatif_D%C3%A9ontologie/IESBA-2016-Handbook.pdf

Ivey, A. E., Ivey, M. B., & Zalaquett, C. P. (2018). *Intentional interviewing and counseling: Facilitating client development in a multicultural society* (9th ed.). Boston, MA: Cengage Learning.

Jamnik, A. (2011). Ethical code in the public accounting profession. *Journal of International Business Ethics, 4*(2), 17–29.

McGuire-Kuletz, M., & Tomlinson, P. (2015). *Return on investment and economic impact: Determining and communicating the value of vocational rehabilitation* (Institute on Rehabilitation Issues Monograph No. 38). Washington, DC: The George Washington University Center for Rehabilitation Counseling Research and Education (GW-CRCRE).

Stadler, H. A. (1986). Making hard choices: Clarifying controversial ethical issues. *Counseling and Human Development*, *19*, 1–10.

United States Department of Education. (2014*). The workforce innovation and opportunity act overview of title II: Adult education and literacy*. Retrieved from http://www2.ed.gov/about/ offices/list/ovae/pi/AdultEd/wioaoverview.pdf

United States Department of Education. (2017*). Programs: Vocational rehabilitation state grants*. Retrieved from https://www2.ed.gov/programs/rsabvrs/ index.html

United States Department of Health and Human Services. (2017). *Code of federal regulations*. Retrieved from https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?fr=19.6

Author Notes

Robert J. Froehlich, EdD, LPC, CRC is the Vocational Rehabilitation Return On Investment (VR ROI) Project Manager and a Project Director/Adjunct Professor at The George Washington University. Olivia Bentley, MA, is a Licensed Professional Clinical Counselor in Kentucky, doctoral candidate at The George Washington University, and Senior Research Assistant at the GWU Center for Rehabilitation Counseling Research and Education. Diona Emmanuel, PhD, MPH, is an Adjunct Professor at

The George Washington University and Clinician in D.C. Maureen McGuire-Kuletz, EdD, CRC is Co-Director of the GWU Center for Rehabilitation Counseling, Research and Education and Associate Professor with The George Washington University.

Journal of Rehabilitation Administration, 40(1), pp. 23–24, 2019

Conceptual Issues in Developing Return on Investment Estimates of Vocational Rehabilitation Programs

Christopher M. Clapp
University of Chicago

John V. Pepper
University of Virginia

Robert M. Schmidt
University of Richmond

Steven N. Stern
Stony Brook University

Abstract. *We provide an overview of the basic conceptual issues involved in estimating the return on investment (ROI) of state vocational rehabilitation (VR) programs. Our aim is to highlight some of the key issues in ROI evaluations, especially those associated with estimating the benefits and costs of VR. Finally, we discuss different ways of implementing ROI calculations and suggest that rate-of-return type analysis is appealing for VR evaluations where there is no widely accepted discount rate.*

Keywords: Return on investment (ROI), rate of return (ROR), vocational rehabilitation, benefit-cost analysis, discount rates, alternative evaluation designs, long-run benefits

Each year, over 1.3 million disabled adults receive services from public-sector vocational rehabilitation (VR) programs at a cost of around $3 billion per year. These federally mandated state programs are administered by 80 different state-level VR agencies that work to ensure that clients "achieve high-quality employment outcomes" (U.S. Department of Education, 2018). In the last decade, there has been heightened interest in producing credible evaluations of whether these VR programs have been effective in meeting that goal. Several recent reports from the U.S. Government Accountability Office (2005, 2007, 2012) highlight the need for improved data and evaluation methodologies. Additionally, the 2014 Workforce Innovation and Opportunity Act (WIOA), which requires formal reports on VR clients' post-program employment and earnings, has further increased the need for updated and credible data and evaluations. To do this, researchers often turn to a return on investment (ROI) analysis.

Return on investment analyses of state VR programs provide a succinct and useful measure of program efficacy. ROI is a measure of investment performance that compares the amount of financial return or benefit relative to program cost (Hollenbeck, 2019; McGuire-Kuletz & Tomlinson, 2015). For example, one commonly used ROI measure is the benefit cost ratio (BCR) which is the ratio of the present value of selected monetizable program benefits to the present value (PV) of costs:

$$\frac{PV\ of\ Benefits}{PV\ of\ Costs}$$

For every dollar spent on a VR client, the BCR shows how many extra dollars (in present value terms) the client earns as a result. Therefore, if the BCR exceeds one, the ROI is positive This BCR may vary across individual VR clients. One may compute the BCR for each individual or provide a summary measure such as the mean, median, or some other quantile. Another related ROI measure called the rate of return is discussed in Section IV.

While the idea is straightforward, developing a credible ROI estimate is a difficult undertaking. Given the available data, one must first estimate and monetize the present value of the benefits and costs of VR services and then, using basic mathematical formulas such as the BCR, determine the ROI.

In this paper, we provide an overview of the basic conceptual issues involved in estimating the ROI of VR programs. McGuire-Kuletz and Tomlinson (2015), and Hollenbeck (2019) provide a more detailed and technical guide to VR-ROI analyses. Our aim is to highlight some of the key issues in ROI evaluations, not to provide an exhaustive how-to guide. As such, this paper should help VR administrators, state program evaluators, policymakers, and others appreciate the complexities involved in developing a credible analysis and interpreting ROI results. Much of this paper focuses on the central issues involved in estimating the benefits of VR. This is the most critical and demanding part of the ROI analyses. After reviewing the issues involved in estimating the benefits of VR in the first section, we then turn to the more mundane albeit important issues involved in determining the costs of VR and the ROI estimates in the third and fourth sections, respectively. The last section draws conclusions.

Estimating the Benefits of VR on Labor Market Outcomes

As discussed in Stern, Clapp, Pepper, and Schmidt (2019), impact evaluations of VR typically focus on labor market outcomes (i.e., employment, wages and earnings). Employment outcomes are of interest to policymakers, and the primary objective of VR programs is to improve labor market outcomes (U.S. Department of Education, 2018). Moreover, labor market outcomes are easily quantified because they are monetized. VR may also have important effects on other outcomes such as self-esteem and independent living skills but these outcomes are difficult to measure and quantify.

The basic idea of an impact evaluation is simple and appealing. Program outcomes – for example, employment and earnings/wages – are measured and compared to the outcomes that would have resulted in the absence of the program. In practice, however, it is difficult to design a credible evaluation where this comparison can be made. The fundamental difficulty is that client outcomes in the absence of the program are counterfactual and not observable. What would have happened to VR recipients had they not received services?

The data alone cannot answer this question. This fundamental methodological problem, labeled the counterfactual outcomes or selection problem, requires that the evaluation design provide some basis for constructing a credible estimate of the counterfactual outcome. This is difficult in practice because VR clients (or their counselors) choose services based on unobservable confounding characteristics that may bias the counterfactual estimate. For instance, highly motivated individuals may seek out and take full advantage of multiple VR services, then find success in the labor market because of both that training and their motivated attitudes. Had the individuals not received VR services, they still might have enjoyed a good deal of job-market success because of their hard-working ways. In this scenario, the researcher has no way of knowing whether positive labor market outcomes are due to VR services or unobserved client motivation because the counterfactual scenario without VR assistance is unobservable. A positive association between VR services and labor market outcomes may reflect unobserved client attitudes and motivation. This would result in an overstatement of VR benefits (positive selection bias). Alternatively, clients with significant impairments that limit their potential returns in the labor market may attempt to overcome the significance of their disabilities by making use of multiple VR services. If the clients' disabilities would have resulted in poorer than average labor market outcomes in the absence of VR services, the effects of those services will be understated (negative selection bias).

More generally, unobserved characteristics such as ability, attitude towards work (e.g., motivation), health status, family support, local labor market conditions, access to transportation, and support from other related programs may affect both the decision to receive substantial VR services and labor market outcomes. Thus, any observed relationships between VR service receipt and labor market outcomes could be spurious. A *selection problem* results from the facts that a) individuals may *select* themselves into a treated group that receives substantial VR services or an untreated group that does not receive substantial services based on their expectation of the resulting labor market outcome and b) the data alone cannot reveal what the counterfactual labor market outcomes would have been.

In a randomized controlled trial (RCT) research design, concerns about selection are negated by randomly assigning subjects into either a treatment group that receives substantial VR services or a control group that does not receive substantial services. In this setting, the decision to assign services is exogenous or unrelated to the labor market outcomes. In practice, selection bias may also impact an RCT if some individuals assigned to the treatment group do not follow through on treatment (dropout bias) and/or individuals assigned to the control group obtain similar treatment outside of the program (contamination bias). Although a useful benchmark to keep in mind, the RCT design is infeasible in most VR settings where counselors and agencies are reluctant to, or possibly even prohibited from, randomly assigning different VR services to clients.

With administrative data on VR clients (see Stern et al., 2019), conclusions about the counterfactual outcomes depend critically on what data are available and what assumptions the evaluator brings to bear. Although this problem can be resolved if the employment data are combined with sufficiently strong assumptions, there is no established solution to the counterfactual outcomes problem that is valid in all settings. Labor economists have long recognized this as the central problem in addressing the impact of job training programs (Friedlander, Greenberg, & Robins, 1997; LaLonde, 1995). Hotz (1992) provided a framework for the Governmental Accountability Office that laid out several options for evaluation of the public-sector VR program in a non-experimental setting that presents a variety of techniques to control for the problem of selection bias inherent in such voluntary programs. Imbens and Wooldridge (2009) provide a summary of some of the recent developments in program evaluation methodologies. As such, establishing credible estimates of what the outcomes would have been without the program is the most critical and demanding part of an impact evaluation. When those estimates are convincing, the effects found in the evaluation can be attributed to the program rather than to any of the many other possible influences on the outcomes (e.g., unobserved motivation, health issues or functional limitations, imperfect controls for local economic conditions, or unobserved support from other programs). Otherwise, the evaluation may be misleading. For example, a simple comparison of the employment outcomes of treated and untreated clients may not estimate the true impact of VR services. Any differences in labor market outcomes could be due easily to one or more of the aforementioned preexisting differences between the groups. The job of a good impact evaluation design is to neutralize or rule out such problems.

Three Simple Evaluation Designs

To illustrate the counterfactual outcomes problem in a relatively simple setting, we reexamine the data from Dean, Pepper, Schmidt, and Stern's (2018) analysis of the Virginia General VR program on clients diagnosed with physical impairments. Since clients receive services for an average of about two years, we focus on employment outcomes three years after the application quarter. This analysis is based on a pre-WIOA period and uses pre-WIOA data.

Table 1 displays the quarterly employment rates one year before and three years after the application for VR services in state fiscal year (SFY) 2000 for clients who received substantial VR services and those who did not receive substantial services. Manski and Pepper (2018) provide a similar illustration in their analysis of right-to-carry gun laws. Following the literature, we refer to these two groups as the treated and untreated, respectively.

These data may be used to compute three simple estimates of the effect of VR services on employment rates. A "before-and-after" analysis compares employment rates for treated clients, yielding the estimate -0.11 (0.41 – 0.52). This estimate suggests VR reduces the employment probability by 11 percentage points. Contemporaneous comparison of the treated and untreated yields the estimate 0.13 (0.41 – 0.28), suggesting VR increases the employment probability by 13 percentage points. The difference-in-difference (DID) estimate compares the time-series changes in employment rates for the treated and untreated, yielding the estimate 0.14 [(0.41-0.52) - (0.28 – 0.53)]. This estimate suggests VR increases the employment probability by 14 percentage points.

These three estimates yield different empirical findings. Given the validity of certain assumptions, each

Table 1

Quarterly Employment Rates by Application Quarter and Treatment Status, SFY 2000 Virginia General VR Agency Clients with Physical Impairments

	Group[2]	
Period[1]	Untreated	Treated
Pre-Application	0.53	0.52
Post-Application	0.28	0.41

Note:

1. The period is four quarters before (pre) or twelve quarters after (post) the date when the VR clients applied for services in SFY 2000.
2. The treated group received substantial VR purchased services. The untreated group did not.

might appropriately measure the effect of VR on the employment rate of Virginia's clients with physical impairments in SFY 2000. However, the assumptions that justify the interpretations differ across estimates, and there is no guarantee that any of the requisite assumptions are valid. Moreover, even if the underlying assumptions are valid, there are several reasons this type of analysis may not reflect the true social benefits of VR services (Dean et al., 2017). First, these estimates do not account for the potential displacement of non-VR participants, particularly if VR services do not improve the VR participant skills or the job matching process. Second, VR services may lead to improved self-esteem and other social benefits associated with increased attachment to the labor market as well as a resulting reduction in use of the social welfare system. While society does not benefit from reduced transfer payments or increased tax revenues – taxpayer gains exactly offset VR participant losses (except for changes in deadweight loss) – social benefits may result from reduced administrative cost associated with welfare programs and increased VR participant utility due to reduced welfare dependence, improved health status, and access to health care insurance (LaLonde, 1995). At the same time, the deadweight costs of taxation may change if welfare receipt and tax payments change.

The "before-after" analysis is correct if one can credibly assume that no determinant of employment, including health status or the local labor market, changed over the four-years between the pre- and post-application periods except for receipt of substantial VR services. In this illustration, the assumption does not appear to hold, at least for the untreated. The employment rate for the untreated fell from 0.53 one year prior to the application quarter to 0.28 three years after the application quarter. Since the untreated group did not receive substantive VR services, something else must have changed, possibly their health and/or local labor market conditions. This casts doubt on the validity of the "before-after" assumption and analysis.

The contemporaneous comparison of employment rates is correct under the assumption that the treated and untreated had the same employment propensities and faced the same labor market environments except for the fact that the treated received substantial VR services. This is commonly referred to as the exogenous or random selection assumption that is credible in RCTs, but it is not generally credible in observational studies where treatments (i.e., VR service receipt) are self-selected. A particular concern is that the collaboration between counselors and clients in determining a plan for services (i.e., the Individualized Plan for Employment) may be influenced by a client's propensity to find employment. In this case, the observed association would be spurious: treated clients would have higher or lower employment rates regardless, depending on whether the selection bias is positive or negative.

Finally, the DID finding is correct if one can plausibly make the assumption that, in the absence of VR services, the treated and untreated would have experienced the same change in employment rates. As with the before-after analysis, the DID model alone only estimates the effect of VR for treated clients. To use this model to estimate the effect for the full population of clients, one needs to combine the DID assumptions with a homogeneity assumption that the effect of VR on employment is the same for the treated and untreated clients. This often is formalized using a linear mean regression model that assumes the effect is the same for all clients.

Clearly, the credibility of this approach depends on whether the "untreated" are a reasonable comparison group – that is, do the untreated clients provide information on the counterfactual trends in the employment rates for the treated clients? To proxy for those trends, researchers have used a number of different internal comparison groups in practice. Those groups include individuals who apply but drop out of the program after being determined eligible and applicants who are ''screened-out'' (e.g., persons whose disabilities are too significant for them to benefit from VR services or those whose disabilities do not constitute or result in a substantial barrier to employment). As with the contemporaneous comparison analysis, a common concern with this approach is that the treatment decision – whether it is made by the client deciding to drop out or the VR counselor who screens out – may, in part, be based upon beliefs about either a client's propensity to find employment or the efficacy of services for that client.

All three of these research designs are commonly used in the literature on the impact of VR programs, frequently in the same evaluations. To determine the impact of workforce development programs in Texas, King, Tang, Smith, Schroeder, and Barnow (2008), and Smith, Christensen, and Cumpton (2015) use a before-after design to evaluate the effects of low-intensity services and contemporaneous comparison to evaluate the effects of high intensity services (relative to low intensity services). Hollenbeck and Huang (2006), and Maryns and Robertson (2015) use both contemporaneous comparison and DID methods to evaluate Washington state's and Minnesota's workforce programs, respectively, while Uvin, Karaaslani, and White (2004), and Wilhelm and Robinson (2013) use all three methods to evaluate the VR programs in Massachusetts and Utah.

While these three approaches are widely used, it may be difficult to credibly address the selection problem using the internal comparison groups they are all based on. VR services are not likely to be randomly assigned, and any imaginable control group is likely to differ in ways that may lead to spurious correlations in the observed data and biased employment impacts.

One common but potentially problematic approach for addressing this concern is to statistically account for observed factors such as age, gender, disability status and severity, and so forth. In this case, researchers assume that VR service receipt is exogenously or randomly assigned conditional on the set of observed covariates even if it may not be exogenous when excluding such control variables from the analysis. A related approach statistically matches clients to untreated individuals based on observable characteristics to construct the most similar counterfactual group (Hollenbeck & Huang, 2006). Yet, the fact that clients with the same covariates receive different services suggests that confounding unobserved factors may play a role in the selection process.

Other Evaluation Designs

Given concerns that VR services are generally not randomly assigned, other model-based evaluation designs have been applied in the literature assessing VR programs. Dean and Schmidt (2005a), for example, address the selection problem by modeling the joint relationship between earnings and VR service receipt using the Heckman (1979) two-stage selection model. Aakvik, Heckman, and Vytlacil (2005) use similar statistical modelling approaches to evaluate VR programs in Norway. More recently, Dean et al. (2015, 2017, 2018, 2019) combine the basic structure of the DID model of labor market outcomes with a model of VR service receipt decisions. By formalizing and estimating a model jointly describing how treatments are selected and outcomes determined, these studies can evaluate the impact of VR services in the presence of the selection problem. While these nonlinear simultaneous equations models allow researchers to formally model the selection problem, they are theoretically, statistically, and computationally complex. This makes them difficult to estimate and evaluate. In contrast, the before-after, contemporaneous comparison, and DID models in the previous section that take realized treatments as given and only model outcomes are relatively more straightforward. For example, Dean et al. (2015, 2017, 2018, 2019) include three jointly determined equations to reflect the mix of services provided, clients' choices to work, and their earnings conditional on working. Since the selection problem occurs because unobserved characteristics may affect both service and labor market outcomes, the researchers model all three relationships as a function of random, unobserved components or error terms. Using this model, they allow services to be assigned based in part on expected labor market outcomes through those unobserved components.

Finally, a well-established approach to address the selection problem exploits some observed covariate, termed an instrumental variable (IV), which has no direct effect on employment outcomes but does influence VR service receipt. In statistical terminology, the IV is said to be independent of employment outcomes but not service receipt. This type of exogenous variation has been shown to help estimate the impact of the treatment. A number of possible observed variables might serve as credible instruments for evaluating the impact of VR services. For example, a client's distance to a VR field office and service provider capacity in a specific geographic area might be related to whether a VR applicant receives services but unrelated to labor market outcomes. Likewise, an order of selection regime may serve as an instrument that is correlated with service receipt but not labor market outcomes.

Dean et al. (2015, 2017, 2018, 2019) use the propensity of a client's VR counselor to assign specific services as an instrument, arguing that counselor tendencies impact VR service receipt but are not directly related to labor market outcomes. As a simplified but intuitive example to illustrate how this IV addresses the selection problem, one can think of there being two types of counselors with respect to a particular service type: high and low propensity. High-propensity counselors decide that every client requires substantial VR services of that type, and low-propensity counselors decide that no client should receive substantial services of that type. If counselors are randomly assigned to clients, or at least if the assignment is unrelated to future labor market outcomes as the researchers argue, then the unobserved factors associated with the assignment to VR services are effectively exogenous, just as in a RCT (Dean et al., 2015, 2017, 2018, 2019).

Two Other Issues

We highlight two other issues related to impact evaluations that are particularly salient for ROI analyses. First, it is important to recognize there is variation in the types of VR services and the types of impairments of VR clients. Second, there may be differences between the short- and long-run impact of VR.

Accounting for heterogeneity in VR services and in the client population. VR agencies provide a wide range of different services to clients with a wide range of disabilities and other characteristics. The decision of how to account for this variation, or heterogeneity, in services and client circumstances is a key issue in designing an impact evaluation. If the estimated impacts differ by type(s) of service received and the type of limitation, the ROI is likely to vary across services and individuals.

Most evaluations classify clients as either receiving or not receiving substantial VR services. Dean et al. (2002) and Dean et al. (2015, 2017, 2018, 2019) aggregate VR services into six types: (1) diagnosis and evaluation, (2) training, (3) education, (4) restoration, (5) maintenance, and (6) other; and allow these six services to have different labor market effects. Moreover, the authors evaluate the impact of VR services on clients with specific types of impairments (e.g., mental illness, cognitive impairments, and physical impairments) rather than the entire caseload, and account for a number of different observed factors, including age, race, gender, years of schooling, and the severity of the disability. Except for Dean and Dolan (1991), and Dean et al. (2015, 2017, 2018, 2019), the existing state-level evaluations of VR services either ignore differences in limitations entirely (Bua-Iam & Bias, 2011; King et al., 2008; Maryns & Robertson, 2015; Wilhelm & Robinson, 2010) or distinguish among clients with different disabilities only by including dummy variables for type of impairment in regression models (Hollenbeck & Huang, 2006; Uvin et al., 2004).

Measuring long run benefits. VR services are thought to have long-run labor market benefits that may be important to account for in an ROI calculation. Dean and Schmidt (2005b), for example, argue that the 10-year ROI is too conservative since earnings gains may be incurred many years after the program. The problem with conducting a lifetime ROI estimate is that the data used to evaluate VR programs do not include lifetime labor market profiles. The longest panel used in the literature evaluating VR programs is the Dean et al. (2015, 2017, 2018, 2019) analysis of applicants to the Virginia general VR agency in SFY 2000 which uses the quarterly labor market outcomes of clients for ten years post-application. Mann, Honeycutt, Bailey, and O'Neill (2017) track VR client outcomes for up to seven years after service receipt. Without the full lifetime labor market profile, which may be too time consuming and costly to assemble, analysts face the problem of trying to use near-term, observed labor market data to draw conclusions about lifetime, unobserved labor market outcomes. To resolve this problem, researchers impute the longer-run benefits from the shorter-run outcome data. Imputing long-run benefits requires assumptions mapping observed data and benefit estimates to future benefit forecasts. The problem is that there is not a single set of assumptions for the extrapolation problem that credibly applies in all settings (Manski, Newman, & Pepper, 2002). This problem may be mitigated in cases where short- and intermediate-run outcomes imply a high rate of return. In such cases, the discounted longer run outcomes may not matter enough to change the basic qualitative conclusion.

Estimating the Costs of VR

Relative to estimating the impact that VR services have on client outcomes, determining the cost of providing VR is straightforward. ROI studies of VR generally ignore the counterfactual outcomes problem when assessing costs. In this case, one merely assembles the realized costs data on VR services and administration. (Yet, there could be a selection problem if there is heterogeneity in costs related to unobserved client characteristics.) Data from the state agency's client data system and from the Rehabilitation Service Administration's Annual Vocational Rehabilitation Program/Cost Report (also known as the RSA-2) provide the necessary information on the costs of services and administrative costs.

Services are provided to clients in any combination of three ways: (1) as a "purchased service" through an outside vendor using agency funds, (2) as a "similar benefit" purchased or provided by another governmental agency or not-for-profit organization with no charge to the VR agency, and/or (3) internally by agency personnel ("in-house benefits"). The section entitled, "Data on Purchased Services and In-House Services," in Stern et al., 2019, provides more detail. VR administrative data provide actual purchased service costs but may not contain the same detailed information for in-house services or similar benefits. Instead, Dean et al. (2015, 2017, 2018, 2019) measure non-purchased service provision costs and administrative costs using data from the RSA-2.

To be clear, there is some uncertainty about the cost estimates derived using the RSA reports, especially for the costs of in-house and similar benefit services. A more detailed analysis of these costs would be useful. In the absence of these details, Dean, Pepper, Schmidt, and Stern (2015, 2017, 2018, 2019) report a range of ROI estimates under different costs estimates.

Computing ROI

Given estimated benefits and costs of VR services, one can then compute a ROI. The basic computations are well-known and largely standardized. Still, there are number of steps in the process that are worth reviewing.

The first step is to discount the dollar values of future benefits and costs to a present value. Benefits and possibly the costs of VR services may be accrued over many years, and a dollar today is worth more than a dollar tomorrow. Discounting is a way to standardize the units of future dollars so they are comparable with current dollars. This allows for an apples-to-apples comparison that reflects the different periods when benefits and costs may be realized. Importantly, this is not an adjustment for inflation but rather a way to account for the real gains that could be realized by investing a dollar today.

Formally, the present value of money recieved in periods in the future equals

$$PV_0 = F\,V_n/(1 + r)^n$$

where PV_0 is the present value in year 0 (i.e., the base year), FV_n is the value n periods into the future (i.e., the future value of benefits), and r is the discount rate. When future streams of money accrue over multiple periods, one adds the discounted stream of money from each period. See Hollenbeck (2019) for more details.

To illustrate, suppose that, five years from today, one will receive \$15,000. How much is that \$15,000 worth today? If the discount rate equals 0.05, then the present value equals \$11,753 $(15{,}000/(1+.05)^5)$. That is, with a five-percent discount rate, \$15,000 in five years (future value of benefits) is worth \$11,753 today (present value of benefits). In other words, investing \$11,753 compounded annually at five percent would yield \$15,000 in five years.

Figure 1 displays the present value of \$15,000 five years from today for a range of discount rates from 0.00 to 0.25. For instance, the figure shows that, with a 0.02 discount rate, the present value of \$15,000 in five years is \$13,586, and, for a discount rate of 0.10, the present value is \$9,314.

Clearly, the present value is sensitive to the choice of the discount rate, r. The discount rate represents the foregone value of money spent today. Stated another way, it is the opportunity cost of not saving or investing capital in the current period. It is chosen by the researcher and is often set to or at least centered on

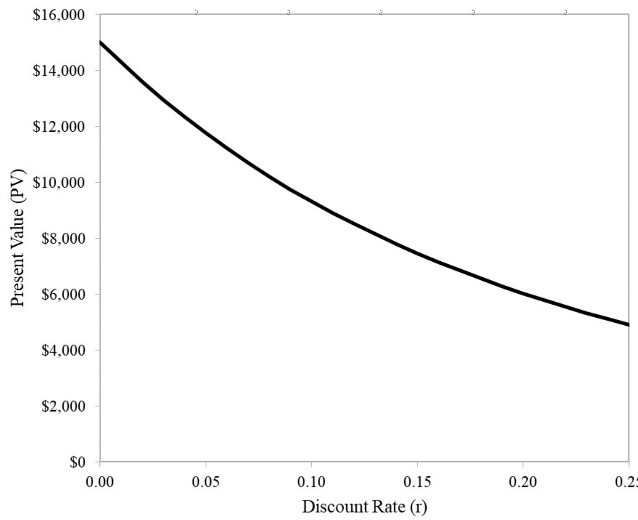

Figure 1. Illustration of How Discount Rate Affects Present Value. For this example, we assume a future value of $15,000 and a time horizon (n) of five years.

some basic interest rate (e.g., savings account interest rate).

After discounting the stream of benefits and costs to the present, a straightforward way to assess the ROI is to compare the present value of benefits to costs. In particular, as noted previously, the BCR equals

$$\frac{\textit{PV of Benefits}}{\textit{PV of Costs}}$$

If the present value of benefits exceeds the present value of costs, the return to VR services is positive and the BCR > 1. Otherwise, the return to VR services is negative and the BCR < 1. The BCR can be interpreted as the "bang per buck." In the VR context, this means that, for every dollar of VR service provision, the customer earns BCR extra dollars (in present value terms). For example, suppose the present value of the costs of VR services is $10,000 and the present value of the benefits is $11,753. Then, the BCR is 1.18, implying that a dollar of VR services results in $1.18 in additional earnings.

Although the BCR is easy to interpret, it is sensitive to the choice of the discount rate. Lower values of the discount rate make the investment look better, and higher values make it look worse. To illustrate, note that the previous hypothetical example calculates the present value of benefits by assuming that VR results in $15,000 in benefits in five years and the discount rate is 0.05. Yet, if the discount rate is 0.084, then the present value of benefits equals $10,000 and the BCR = 1. If the discount rate is 0.10, then the present value of benefits is $9,314, and the BCR is less than one.

The sensitivity of the BCR to the discount rate may be problematic for evaluating workforce training programs. Businesses typically use some measure of their financing costs (i.e., "cost of capital") as a discount rate when evaluating an investment. By contrast, there is no widely accepted "cost of capital" or discount rate for evaluating workforce training programs. Moore, Boardman, Vining, Weimer, and Greenberg (2004) present a discussion of the issues surrounding the use of discount rates in program evaluation and guidance on how to choose an appropriate rate. The choice is largely arbitrary, and, given the sensitivity of BCR to the discount rate, the rate used can make a program look good or bad.

In this setting, the rate of return (ROR) provides an alternative approach that may be preferred. The ROR is the discount rate that equilibrates the returns from an investment to the cost of the investment. That is, the ROR is the interest rate where the BCR = 1 or the present value of benefits equals the present value of costs. This calculation does not require the choice of an arbitrary discount rate. To illustrate, Figure 2 shows the ROR that results from a range of benefits realized five years after $10,000 of costs were incurred. The figure shows that, if the $10,000 investment returns $15,000 in 5 years (see Figure 1), the ROR is 0.084 $(15{,}000/10{,}000)^{1/5}$-1). That is, for a

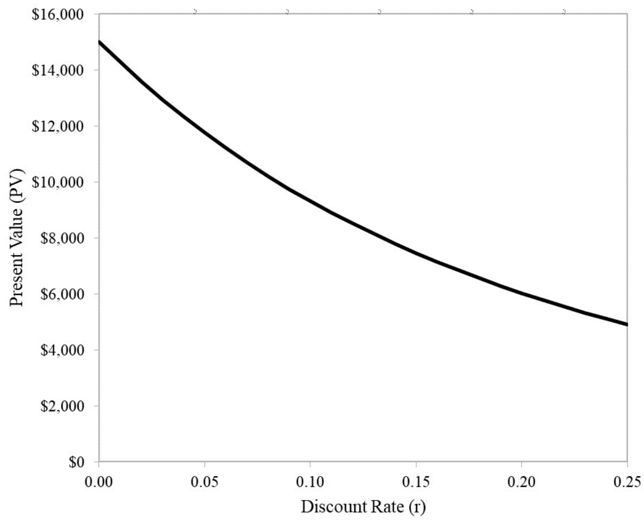

Figure 1. Illustration of How Discount Rate Affects Present Value. For this example, we assume a future value of $15,000 and a time horizon (n) of five years.

discount rate of 0.084, the present value of benefits equals the present value of costs.

The ROR can be compared to that of other government programs or well-known returns in the private sector. For example, current annual returns on money market accounts are 2% or less and the long-run annual rate of return to the U.S. stock market is about 10%. Thinking of a discount rate as the "opportunity cost of capital" and using the ROR of 8.4% from our hypothetical example, purely profit maximizing individuals would choose to "invest" their money in VR instead of a money market account, but would prefer the long-run returns from the stock market to either of the other two investments. Alternatively, the U.S. Office of Management and Budget (OMB) sets guidelines for evaluating public sector programs (OMB, 1992). Those guidelines include discount rates by time horizon that are updated each year. According to OMB (2018), current discount rates vary from 1% for 3-year horizons to 2.6% for 30-year horizons.

Dean et al.'s (2015, 2017, 2018, 2019) recent analyses of applicants to the Virginia general VR agency in SFY 2000 estimate the long run ROR of VR services for each client. They report a median annualized rate of return of 20% for clients with mental illness, 19% for clients with cognitive impairments, and 169% for clients with physical impairments. Thus, by any conventional standard, the ROI of VR services for this cohort is positive and substantial. Schmidt, Clapp, Pepper, and Stern (2019) summarize more recent ROI evaluations.

Conclusion

In this paper, we highlight key conceptual issues involved in ROI evaluations of VR programs. Most notably, we focus on estimating the benefits and costs of VR in light of the counterfactual outcomes problems. This is the most critical and demanding part of ROI analyses. We then discuss different ways of implementing ROI calculations and suggest that the ROR analysis is appealing for VR evaluations where there is no widely accepted discount rate.

There are many other critical steps involved in undertaking such an evaluation. For example, analysts must decide whether to report returns at a client or program level; what outcomes to monetize (e.g., labor market, disability insurance, others); whether the return should be measured for society, the taxpayer, the client, or some other group; what the relevant time period should be; and how to account for statistical uncertainty. These and other issues shape the details of an ROI analysis. Readers interested in a more complete and in-depth analysis of VR-ROI might turn to McGuire-Kuletz and Tomlinson (2015), and Hollenbeck (2019).

References

Aakvik, A., Heckman, J. J., & Vytlacil, E. J. (2005). Estimating treatment effects for discrete outcomes when responses to treatment vary: an application to Norwegian vocational rehabilitation programs. *Journal of Econometrics*, *125*(1-2), 15–51.

Bua-Iam, P., & Bias, T. K. (2011). Economic impacts of West Virginia Division of Rehabilitation Services on consumers with significant disabilities: Realistic return-on-investment models for state-federal VR programs. *Journal of Rehabilitation*, *77*(3), 25–30.

Dean, D. H., & Dolan, R. C. (1991). Fixed-effects estimates of earnings impacts for the vocational rehabilitation program. *Journal of Human Resources*, *26*(2), 380–391.

Dean, D. H., Dolan, R. C., Schmidt, R., Wehman, P., Kregel, J., & Revell, G. (2002). A paradigm for evaluation of the Federal-State vocational rehabilitation program. In J. Kregel, D. H. Dean, & P. Wehman (Eds.), *Achievements and challenges in employment services for people with disabilities: The longitudinal impact of workplace supports (*pp. 1–40). Richmond, VA: Virginia Commonwealth University Rehabilitation Research and Training Center on Workplace Supports.

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.

Dean, D. H., Pepper, J. V., Schmidt, R. M., & Stern, S. N. (2019). The effects of youth transition programs on labor market outcomes. *Economics of Education Review*, *68*, 68–88.

Dean, D. H., & Schmidt, R. (2005a). *Determining the 'return on investment' for VR participants. Data on employment of people with disabilities: Deliverable 6.* Richmond, VA: Virginia Department of Rehabilitative Services.

Dean, D. H., & Schmidt, R. (2005b). *An outcome-based assessment of the Chapter 31 Program.* Final Report, U.S. Department of Veterans Affairs, Contract No. 101-Y27247.

Friedlander, D., Greenberg, D. H., & Robins, P. K. (1997). Evaluating government training programs for the economically disadvantaged. *Journal of Economic Literature*, *35*(4), 1809–1855.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, *47*(1), 153–161.

Hollenbeck, K. M. (2019). What is ROI? *The Journal of Rehabilitation Administration*, *40*(1), 5–10.

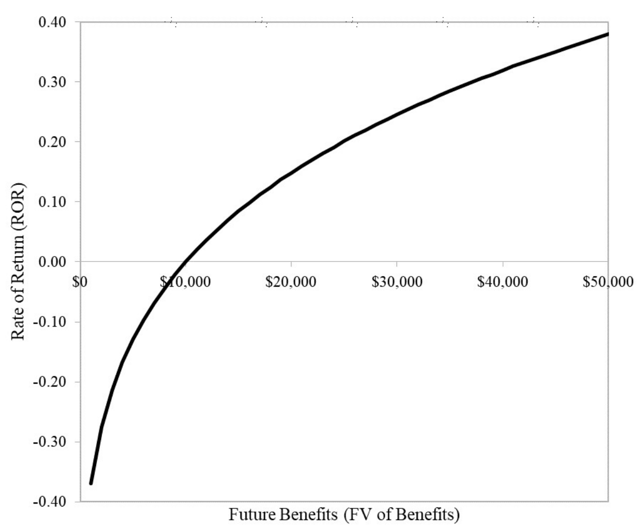

Figure 2. Illustration of How the Future Value of Benefits Affect Rate of Return. For this example, we assume the present value of costs is $10,000 and a time horizon (n) of five years.

Hollenbeck, K. M., & Huang, W. J. (2006). *Net impact and benefit-cost estimates of the workforce development system in Washington state*. Kalamazoo, MI: W. E. Upjohn Institute for Employment Research.

Hotz, V. J. (1992). *Design analysis plan for the GAO study of the vocational rehabilitation program*. Correspondence with the GAO Program Evaluation and Methodology Division, 1–10.

Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, *47*(1), 5–86.

King, C. T., Tang, Y., Smith, T., Schroeder, D., & Barnow, B. S. (2008). *Returns from Investments in Workforce Services: Texas Statewide Estimates for Participants, Taxpayers, and Society*. Austin, TX: Ray Marshall Center for the Study of Human Resources.

LaLonde, R. J. (1995). The promise of public sector-sponsored training programs. *Journal of Economic Perspectives*, *9*(2), 149–168.

Mann, D. R., Honeycutt, T., Bailey, M. S., & O'Neill, J. (2017). Using administrative data to explore the employment and benefit receipt outcomes of vocational rehabilitation applicants years after program exit. *Journal of Vocational Rehabilitation*, *46*(2), 159–176.

Manski, C. F. & Pepper, J. V. (2018). How do right-to-carry laws affect crime rates? Coping with ambiguity using bounded-variation assumptions. *Review of Economics and Statistics, 100*(2), 232–244.

Manski, C. F., Newman, J., & Pepper, J. V. (2002). Using performance standards to evaluate social programs with incomplete outcome data: general issues and application to a higher education block grant program. *Evaluation Review*, *26*(4), 355–381.

Maryns, N., & Robertson, R. (2015). *Smart investments, real results: A net impact evaluation framework for Minnesota's workforce Development System and Initial Findings*. St. Paul, MN: Minnesota Department of Employment and Economic Development (DEED).

McGuire-Kuletz, M., & Tomlinson, P. (2015). *Return on investment and economic impact: Determining and communicating the value of vocational rehabilitation* (Institute on Rehabilitation Issues Monograph No. 38). Washington, DC: The George Washington University Center for Rehabilitation Counseling Research and Education (GW-CRCRE).

Moore, M. A., Boardman, A. E., Vining, A. R., Weimer, D. L., & Greenberg, D. H. (2004). "Just give me a number!" Practical values for the social discount rate. *Journal of Policy Analysis and Management*, *23*(4), 789–812.

Office of Management and Budget. (1992, October 29). *Memorandum for heads of executive departments and establishments: Guidelines and discount rates for benefit-cost analysis of federal programs* [Circular No. A-94 Revised]. Retrieved from: https://obamawhitehouse.archives.gov/omb/circulars_a094/

Office of Management and Budget. (2018, February 8). *Discount rates for cost-effectiveness Analysis of Federal Programs* [Revisions to Appendix C of OMB Circular A-94]. Retrieved from https:// www.federalregister.gov/documents/2018/02/08/2018-02520/discount-rates-for-cost-effectiveness-analysis-of-federal-programs/

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

Smith, T., Christensen, K., & Cumpton, G. (2015). *An evaluation of local investments in workforce development: 2014 Update*. Austin, TX: Ray Marshall Center for Human Resources.

Stern, S. N., Clapp, C. M., Pepper, J. V., & Schmidt, R. (2019). Data issues in developing valid ROI estimates. *The Journal of Rehabilitation Administration*, *40*(1), 35–46.

U.S. Department of Education, Office of Special Education and Rehabilitative Services. (2018) *Rehabilitation Services Administration Report for Fiscal Years 2014–15, Report on Federal Activities under the Rehabilitation Act of 1973, as Amended*, Washington, DC.

U.S. Government Accountability Office. (2005). *Vocational rehabilitation: Better measures and monitoring could improve the performance of the VR program*. (GAO-05-865). Washington, DC.

U.S. Government Accountability Office. (2007). *Vocational rehabilitation: improved information and practices may enhance state agency earnings outcomes for SSA beneficiaries*. (GAO-07-521). Washington, DC.

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, DC.

Uvin, J., Karaaslani, D., & White, G. (2004). *Evaluation of Massachusetts' public vocational rehabilitation program: Final report*. Boston, MA: Massachusetts Rehabilitation Commission.

Wilhelm, S. & Robinson, J. L. (2013). The economic impact of Utah's vocational rehabilitation program. *Journal of Disability Policy Studies*, *24*(3), 148–17.

Author Notes

Christopher Clapp, PhD, is a Senior Research Associate in the Kenneth C. Griffin Department of Economics at the University of Chicago. John Pepper, PhD, is the Merrill H. Bankard Professor of Economics at the University of Virginia. Robert Schmidt, PhD, is an Emeritus Professor of Economics at the University of Richmond. Steven Stern, PhD, is a full Professor of Economics at Stony Brook University. Drs. Pepper, Clapp, 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, Kirsten Rowe and two anonymous reviewers for helpful comments and the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR Grant 90DP0070) for generous funding. All errors are ours.

Journal of Rehabilitation Administration, 40(1), pp. 35–46, 2019

Data Issues in Developing Valid ROI/ROR Estimates

Steven N. Stern
Stony Brook University

Christopher M. Clapp
University of Chicago

John V. Pepper
University of Virginia

Robert M. Schmidt
University of Richmond

Abstract. *This paper discusses issues associated with using readily available administrative data in estimating ROI for vocational rehabilitation services. It starts with a discussion of longitudinal outcomes data. The discussion is divided up into labor market outcomes data, other types of outcomes data, necessary sample sizes (power analysis), and ways to deal with people systematically excluded from the outcomes data. Next, the paper focuses on services data. The topics covered include the need for control groups, using service cohort data, different sources of service, and merging service data with outcomes data. Finally, the paper moves to the need for other controlling explanatory variables including discussions of inclusion of demographic explanatory variables and data from local labor markets. Two online appendices to this paper provide additional details through (a) an example of a power analysis to illustrate sample size issues and (b) a discussion of Institutional Review Board issues associated with conducting empirical investigations using administrative data.*

Keywords: Return on investment (ROI), vocational rehabilitation, longitudinal data, unemployment insurance data, control group, VR service data, personally identifiable data, explanatory variables

An important step for return on investment/rate of return (ROI/ROR) analysis of state vocational rehabilitation (VR) programs is constructing and cleaning data. Significant time, care, and expense must go into planning for and execution of the process for collecting, merging, and cleaning of the data. Unfortunately, data construction is expensive, sometimes in acquisition costs and always in time spent merging and cleaning. There are no easy and appropriate ways to avoid the investment necessary to have good data. Furthermore, without good data, estimation and analysis are meaningless. Therefore, it is important for any administrator who is considering an ROI/ROR study to make sure that there are adequate resources to develop appropriate data.

ROI/ROR essentially involves a comparison of benefits and costs of a proposed investment. For state VR programs, the benefits are derived mainly from participants' labor market outcomes, and the costs are borne by the state agency providing services to them. A key feature of VR services is that there are many different types of investments. The first section discusses one of the two key data pieces necessary for estimation: longitudinal outcomes data. The section is divided up into discussions of labor market outcomes data, other types of outcomes data, necessary sample sizes (power analysis), and ways to deal with people systematically excluded from the outcomes data. The second section discusses the other key data piece, services data. The section is divided up into discussions of the need for control groups, using service cohort data, different sources of service, and merging service data with outcomes data. The third section discusses the need for other controlling explanatory variables. This includes discussions of inclusion of demographic explanatory variables and data from local labor markets. Measuring cost is outside the scope of this article; however, one can see Clapp, Pepper, Schmidt, and Stern (2019) for a discussion of necessary cost measurement issues.

Longitudinal Outcomes Data

One of the two key data components for an empirical analysis of the effectiveness of a VR program is data on outcomes that are potentially influenced by that program. These data must be linked or merged with data on VR services received by each person. Once merged, the combined data can be used to estimate how receipt of VR services affects the outcomes of interest over a long enough period to produce solid ROI/ROR estimates. Without longitudinal data, one must assume that the labor market effects are constant over time, and this assumption is rejected by the data. In the next section, entitled "Service Receipt Data," we discuss the necessary characteristics of service data and issues concerning merging. In this section, we limit our discussion to issues associated with the outcomes data.

Labor Market Outcomes

A key set of outcomes for policy makers is labor market outcomes. This arises partially because the stated goal of VR services is to help people with disabilities prepare for, develop skills for, and gain access to the labor market. Four important labor market outcomes are employment, hours worked, wage, and earnings. Employment is a binary measure of involvement in the labor market; one is either employed or not. Hours worked per week, per quarter, or per year provide information on intensity of involvement in the labor market. Wage is a measure of how much one is paid for each fixed unit of time worked. Labor economists argue that wage is a measure of productivity in the labor market (e.g., Ehrenberg & Smith, 2012, chapter 3). One of the goals of a VR program might be to help participants become more productive and therefore increase their wage. Earnings, defined as the wage multiplied by the units of time working, are the outcome that policymakers tend to focus on the most.

Maybe the best available source of data on labor market outcomes that can be merged with information on service receipt is unemployment insurance earnings (UI) data. Each state's UI program collects data from most business establishments on the quarterly earnings of each of their employees. The main purpose of the collection is to determine eligibility and benefit amounts for unemployment insurance benefits. However, as has been recognized (Dean & Dolan, 1991; Dean, Pepper, Schmidt, & Stern, 2015, 2017, 2018, 2019; Hollenbeck & Huang 2006; Wilhelm & Robinson, 2010), the data also can be used as a longitudinal panel for all workers in the state who are eligible for unemployment insurance benefits.

There are three prominent problems with UI data: a) most state UI programs do not collect information on wages or hours worked; b) a significant number of people in each state are not covered by their state's UI program; and c) the researcher cannot distinguish among exit from the labor market, exit from a covered job to an uncovered job, exit from the state, and death. There are no good solutions to (a) and (c). Using only UI data for outcomes prevents the researcher from saying anything about wages (and, therefore, productivity) and hours worked. Both of these pieces of information would be valuable to observe but are not available in most states, nor can they be imputed from the available data. In addition, when a person disappears from the UI data, the researcher cannot determine the reason. This is unfortunate as it would be useful to know the circumstances leading to attrition; different causes of attrition have very different policy implications (e.g., Chang, Yang, Tang, & Gianguli, 2009; de Graaf, Bijl, Smit, Ravelli, & Vollebergh, 2000). For example, death would imply a loss of the accumulated skills, while migration would not. With respect to (b), there are two large groups not covered in a single state's UI data: (1) federal government workers; and (2) people who cross state lines to work, who would be covered in the state where they work but not in the state whose program is being analyzed. The planned State Wage Interchange System (U.S. Department of Labor [USDOL], 2016), which will enable state VR agencies to access other states' UI data on their VR program participants and largely alleviate the issue of people crossing state lines for work. The Federal Employment Data Exchange System (FEDES; Jacob France Institute, 2018) previously provided data on federal employment to participating states to help them meet their reporting requirements and could potentially address the UI data gap for federal workers. However, the FEDES was suspended in January 2018 (USDOL, 2017) while the U.S. Department of Labor assesses its feasibility. In the section entitled, "Dealing with People not Covered by Unemployment Insurance," we discuss some alternative methods to adjust for these current data gaps.

The big advantages of the UI data are that i) the researcher can observe multiple quarters of data both before and after service receipt; and ii) the data are high quality because it has an important administrative purpose. With respect to (i), Dean et al. (2015, 2017, 2018) used three years of UI data prior to VR service receipt and ten years after service receipt. As seen in Table 1, the results in Dean et al. (2015, 2017, 2018) show that labor market outcomes right after service receipt are not predictive about longer-term effects of VR. The short-run estimates include the first two years after service receipt, and the long-run estimates are for quarters after two years. In almost all cases, the short-run and long-run estimates, both for employment and earnings, are quite different. For example, for people with mental illness, quarterly earnings (if employed) are 5.5% lower in the short run than before service but 13.6% higher in the long run.

Other potentially useful data sets that do not rely on UI coverage have their own problems. One possibility is to use specially provided Social Security Administration (SSA) earnings data (see, for example, Honeycutt, Thompkins, Bardos, & Stern, 2015, 2017; Schley, Weathers, Hemmeter, Hennessey, & Burkhauser, 2011). However, gaining access to such data is difficult and must be used under restrictive rules. National surveys typically collect self-reported earnings data that is known to be inaccurate. For example, Bricker and Engelhardt (2007) report standard deviations of measurement error in reported earnings data of 5.9% (for men) and 6.7% (for women) in the Health and Retirement Survey. Kim and Tamborini (2014) report systematic measurement error by gender and race in the Survey of Income and Program Participation. Both of these used administrative data to assess the measurement error in earnings. Kapteyn and Ypma (2006) provide a survey of the research on this topic and suggest that administrative data might have errors as well. However, overall,

Table 1

Short-Run and Long-Run Changes in Employment and Earnings Due to Receipt of Training Services

	Cognitive Impairment	Mental Illness	Physical Impairment
Change in Employment Rate			
Short Run	10.3%	9.4%	5.0%
Long Run	6.8%	8.0%	6.4%
Change in Quarterly Earnings (if employed)			
Short Run	20.9%	-5.5%	0.9%
Long Run	28.5%	13.6%	17.2%

Notes:

1) Short-run numbers are for the first two years after service, and long run numbers are after two years

2) All numbers come from Dean et al. (2015, 2017, 2018). Employment Rate numbers translate the reported results into rates.

3) Both short-run and long-run numbers above are the difference in estimates after service receipt minus estimates prior to service receipt. The change in the outcome is the effect of the service.

administrative earnings data are usually taken as "the truth."

Other Outcomes

Besides labor market data, there might be other VR outcomes of interest to policy makers, including job-connected benefits, especially health benefits, and receipt of SSI and SSDI (and other government transfers). Job-connected health benefits are not observable in UI data. They are frequently observed in other national data sets, but these data lack information on receipt of VR services. Information on SSI and SSDI receipt can be acquired from SSA, but at very large cost with significant binding restrictions. For example, Dean et al. (2017) used individual SSI/SSDI data attained from SSA on Virginia VR program participants. However, it took about eight years to get the data after initial application, and we were no longer allowed to use them about five years later. Thus, while including such outcomes would provide valuable information, getting meaningful estimates of the outcomes is difficult and expensive.

Other potential VR outcomes include independent living skills, community integration, and emotional health. These are all outcomes that both can be affected by VR services and are of significant value to individuals with disabilities. They should also be important to policy makers even if they are concerned only with government fiscal health. For example, a person with good independent living skills requires less help from expensive government services. However, no administrative body collects data on any of these non-labor market outcomes, and the cost and effort to acquire them through expansion of administrative data systems or survey research would be prohibitive. In addition, by their very nature, measures of independent living skills, community integration, and emotional health are not as objective as earnings and employment and are difficult to monetize.

Necessary Sample Sizes

The first step in ROI/ROR analysis is to estimate statistically significant effects of services on outcomes of interest. The bigger the sample, the more likely the estimates will be statistically significant. A reasonable rule of thumb is to require at least 1000 observations and have a strong preference for approximately 2000 observations. Online Appendix A (available at: https://scholarship.richmond.edu/economics-faculty-publications/55/) provides some detail in a simplified example of a power analysis. In some analyses for people with small prevalence rates (e.g., people who are blind or people with autism), the researcher can increase sample size by using multiple cohorts of service recipients (see the section entitled, "Using a Service Receipt Cohort").

Dealing with People not Covered by Unemployment Insurance

As previously mentioned, there are two large groups of employed people who are not covered by UI and therefore not included in a given state's UI data: a) people who work for the federal government; and b) people who commute across state lines to work. The numbers of VR participants who are not included in their state's UI records can be substantial. Online Appendix 7 for Dean et al. (2017) report on an analysis of over 9,000 VR participants in Virginia with records in both the Virginia UI system and SSA earnings files. In 12% of those cases, the SSA showed earnings in 2001 when UI did not. Plausibly, the majority of these instances are due to federal employment and cross-state commuting. This proportion may be larger than in most states because Virginia has many military installations (leading to many jobs with the federal government), and shares borders with Washington, DC (leading to many jobs with the federal government) and Maryland (leading to many jobs in that state).

Both of these problems have two possible solutions. The first (and best) solution, if possible, is to collect labor market data for the missing people. Hollenbeck and Huang (2006) used this approach, and Wilson (2005) proposes using this approach. The second solution is to use statistical techniques to control for the missing data. One can use publicly available data from the federal government on the number of people in each county who work for the federal government and on the number of people who commute across state lines. These can be turned into per capita numbers and then used as regressors to control for variation in prevalence of federal government jobs and across-state-lines commuting (Schmidt, Clapp, Pepper, & Stern, 2019).

Service Receipt Data

While having outcomes data is essential to estimating the effects of VR service receipt, so is having service receipt data that vary across program participants. The service data should reflect the choices made by VR participants and their counselors in a parsimonious way without losing the sense of flexibility in state VR programs. In this section, we begin by discussing why variation in service receipt is essential. Then we discuss issues associated with choosing the source of service data and merging the service data with the outcomes data.

Need for a Control Group

In order to say anything about the effects of VR services on labor market outcomes, one needs a control group; a subset of people who did not receive the services for which the effects are being estimated. It is only through comparison of outcomes for people who did and did not receive the service that the researcher can make any statements about causation (or even correlation). Without a control group, the researcher has no information about the difference in outcomes between those who received service and those who did not; although Bua-Iam, Hampton, Sink, and Snuffer (2013) argued in this journal that features unique to state VR programs make it unreasonable to use control groups. Bua-Iam and Bias (2011) estimated effects of VR programs essentially assuming that VR participants would have earned nothing without the VR services. This is an extreme assumption and certainly not as good as using a flawed but reasonable control group such as those constructed by Dean et al. (2015, 2017, 2018, 2019).

One of the interesting features of state VR programs is that there are multiple types of services. The existence of people choosing from and sometimes using multiple services allows the researcher to identify the effect of each service even if there are no people using no services (see Clapp et al., 2019). For example, Dean et al. (2015, 2017, 2018) aggregated the set of services offered by the Virginia Department of Aging and Rehabilitative Services (DARS) into six separate service groups: diagnosis & evaluation, training, education, restoration, maintenance, and other services. They modeled service choice as well as service impact, allowing for any combination of service categories including none at all. Other authors acknowledge the existence of multiple services but seldom address the issue directly. For example, Aakvik, Heckman, and Vytlacil (2005) note that there are multiple service types but do not model them. Rather, they allow for heterogeneous treatment effects of an observed binary treatment variable. Frolich, Heshmati, and Lechner (2004) observe the existence of multiple programs but only focus on comparing them. By contrast, Hollenbeck and Huang (2006) evaluate multiple services separately.

In fact, many VR applicants receive no VR services. These people are either declared ineligible for services or choose to withdraw. One might argue that this group is not a good control group as their non-receipt of services is probably correlated with their labor market outcomes (Dean & Dolan, 1991). The potential correlation causes a bias in the estimate although the direction of the bias is unclear. For example, one might argue that the people who choose to utilize a particular service are those who would benefit most from it. If this were the case, then the observed labor market outcomes for those who used the service would provide an upwardly biased estimate of the service's effect for those who chose not to use it (or for a randomly chosen person). Alternatively, one might argue that the people who choose to utilize a particular service are at greatest need for any service because they have poor market skills. If this were the case, then the observed labor market outcomes for those who used the service would provide a downwardly biased estimate of the service's effect for those who chose not to use it (or for a randomly chosen person). These endogeneity concerns (or "selection" issues) are almost surely valid and are the concern behind much of the economic literature on estimating treatment effects (Aakvik et al., 2005; Dean & Dolan, 1991; Doyle, 2007; Heckman, Ichimura, & Todd, 1998; Heckman, LaLonde, & Smith, 1999). However, the same concern applies to all recipients of any service as well, and the solution is to find or construct instrumental variables (Aakvik et al., 2005; Dean & Dolan, 1991; Doyle, 2007; Heckman et al., 1999; Clapp et al., 2019). An instrumental variable is correlated with the likelihood of VR service receipt but not with employment and earnings, except through the effect of that service on employment and earnings. For example, closing certain order-of-selection categories influences the likelihood of those individuals receiving services, but its only impact on their employment is (1) if it precludes them from getting VR service or (2) if order of selection is correlated with labor market conditions not already controlled for.

Using a Service Receipt Cohort

There are various ways to organize VR administrative data. Dean et al. (2015, 2017, 2018, 2019) used observations from the cohort of VR participants who applied for VR services during the fiscal year (FY) 2000 disaggregated into three large disability groups: people with (1) cognitive impairments, (2) mental illness, and (3) physical impairments. They further limited their analysis to people whose first VR case was in FY 2000, because Dean et al (2015) showed that there was bias associated with using people who had had prior cases (as one would expect).

There are two reasons a researcher might want to deviate from such a strategy. First, given changes in VR priorities, the proportion of individuals who are youth in transition has increased significantly. In fact, the Workforce Innovation and Opportunity Act (WIOA) of 2014 requires VR agencies (see section 419 of WIOA) to reserve at least 15% of their allotted funds for pre-employment transition services to students with disabilities. With such participants, using only the first case results either in no labor market outcomes (because they are too young) or deleting a high proportion of them from the estimation sample (because their first service case occurred while they were in school). Thus, other statistical adjustments must be used as an alternative to dealing with the bias suggested in Dean et al. (2015).

Second, there are some disability groups with small enough prevalence so that there would not be enough observations if the researcher limited analysis to one cohort year. One way to address this problem is to use multiple year cohorts from the same state. Another is to use single year cohorts across multiple states. For example, in on-going work, Clapp, Pepper, Schmidt, & Stern (2018a) use applicant cohorts from FY 2007 from Maryland, Oklahoma, and Virginia to estimate the effect of VR service receipt on labor market outcomes for people who are blind, and Clapp, Pepper, Schmidt, & Stern (2018b) use applicant cohorts from multiple fiscal years in Virginia (FY 2000 - FY 2007) to estimate the effect of VR service receipt on labor market outcomes for adults with autism.

Another interesting possibility is to use available data on multiple cases for each individual rather than exclude individuals with prior cases as was done in Dean et al. (2015, 2017, 2018). While including individuals with prior cases increases sample sizes, there are serious data and statistical issues that need to be addressed. First, and most importantly, the researcher must think less in terms of discrete VR cases that begin with application and end with closure, and more in terms of "service spells," each of which potentially could encompass multiple VR cases. Thus, one must take a stand on when a service spell begins and ends. One approach is to use service receipt dates to define a spell. However, our understanding is that dates recorded for services suffer from serious measurement error. Frequently, an authorization record for a specific service includes the earliest date a service can begin as well as the date when the service is expected to be completed (which can be well after the case closure date). Neither identifies the precise delivery date. A second approach is to use case application and closure dates to define service stints. A service stint could be defined as encompassing all cases where a case's application date is within a researcher-defined period (e.g., within the same quarter) of the prior case's closure date. Also, there are interesting econometric issues associated with using this approach. Dean et al. (2015) found significantly different service impacts for individuals with prior case(s) versus those for whom the case was their first. Thus, this heterogeneity must be accounted for when including both in a single analysis. Additionally, one might check, for example, Lancaster (1992) for a discussion of issues when using data with varying service spell lengths.

Data on Purchased Services and In-House Services

VR provides services to participants through any combination of four sources — internally from their VR counselors or other VR staff, through comparable benefits, externally through purchased services, and/or through a state-operated comprehensive rehabilitation facility. This section begins by considering each in turn.

In the past, VR staff typically maintained written case notes about the services they provided in lieu of entering such information into VR administrative databases. Because ROI/ROR analyses tend to use retrospective cohorts and readily available administrative data, the extent and nature of in-house services have rarely been included in such analyses. Similarly, the provision of and precise nature of comparable benefits have seldom been available to the researcher for previous ROI analyses. However, as VR data systems grow in sophistication, better information on in-house services and comparable benefits may become available.

By contrast, state VR agencies have long tracked purchased services for both accounts payable and required federal reporting purposes. Agencies typically record a rich array of information for every purchase, including authorization date, date paid, vendor, service type (generally, with hundreds of categories), dollar amount, and so forth. From an ROI/ROR perspective, the key missing item in these records is the date of service delivery. Although agencies typically record the beginning and ending dates for service delivery, these are set to provide flexibility in their provision and do not provide a clear indication of the actual service date. As a result, researchers must make their own decision/assumption with respect to the period during which services were provided. In the case of Dean et al. (2015, 2017, 2018, 2019), services were presumed to begin (potentially) at application. When they ended was not important to determine and was left as an open question.

Eight states also provide services through a state-operated comprehensive rehabilitation facility (e.g., the Maryland's Workforce Technology Center and Virginia's Wilson Workforce and Rehabilitation Center). Generally, these facilities provide services as part of an Individualized Plan for Employment and in coordination with the participant's counselor. In our experience, all of the items recorded for purchased services are also recorded for these facilities with one exception that relates to the cost of the service rather than its provision. For purchased services, the recorded dollar amount reflects the actual price paid by the agency for the specific service. For services provided by these facilities, either no dollar amount is recorded for individual services, or a "charge" for the service is recorded based on a schedule that identifies either a blanket charge or a per-unit charge that is coupled with the number of units provided to the participant.

Service Aggregation

Dean et al. (2015, 2017, 2018, 2019) show that different types of VR services are likely to have very different labor market effects. They classify services into six types: diagnostic, training, education, restoration, maintenance, and other. Schmidt et al. (2019) reference recent work that replaces the "other" type with two new types, placement and job supports. Irrespective of the service types chosen, mapping the agency's detailed service categories is one of the most important tasks the researcher performs. The rationale by which ROI/ROR analyses group services differs from that by which VR staff think about service types. For example, consider the diagnostic category. From a VR perspective, receipt of diagnostic services suggests eligibility determination; from an ROI/ROR perspective, receipt of diagnostic services identifies other services that would improve functioning in labor markets. Thus, diagnostic services would include eligibility determination as well as medical diagnostics and vocational evaluation. Performing the crucial task of aggregating service categories well requires the combined efforts of VR staff with agency knowledge and a researcher with an ROI/ROR mindset.

An alternative source for identifying service provision is the standardized Case Service Report (also known as the RSA-911) submitted by state VR agencies to the federal Rehabilitation Services Administration (RSA). Over time, RSA has required the reporting of increasing amounts of service-related data. Conceptually, the use of the RSA-911 has an important advantage over the above-described data sources: it includes a code for provision of each service type by agency personnel. As a practical matter, however, if information on in-house services resides strictly in case notes, then the validity of these data relies upon VR counselors to enter this information accurately. To use this source, the researcher must also address how well the RSA-911 categories work for an ROI/ROR analysis. For example, the RSA-911 combines both medical diagnostics and treatment into a single category. However, as discussed above, diagnostic services may have different labor market effects than treatments.

Merging Process

UI data and VR administrative data are not useful for estimating labor market outcomes of VR service receipt unless the researcher can merge the two data sets together. In other words, for each person in the VR administrative data, the researcher must find the labor market earnings history of that person in the UI data. The key to doing this involves merging by Social Security number (SSN) since it is included in both data sets. The level of complexity is increased by the need for maximum possible anonymity required by VR agencies as well as Institutional Review Board (IRB) rules (see online Appendix B, available at: https://scholarship.richmond.edu/economics-faculty-publications/55/) for a discussion of IRB issues associated with conducting empirical investigations using state VR program data).

A good merging methodology used for SSA data merger in Dean et al. (2017) is:

1. VR agency staff member draws one large-digit random number (PIN) for each person in the VR administrative data set, then creates and stores a crosswalk table connecting each SSN in the VR administrative data to the assigned PIN.

2. VR agency staff member strips each VR administrative data observation of the SSN, replaces it with the PIN for that observation (from step 1), removes other personally identifiable information (PII), and provides the amended data to the researchers.

3. VR agency staff member provides SSA staff member with the crosswalk connecting each SSN in the VR administrative data to the assigned PIN for that observation (from step 1).

4. SSA staff member gathers the SSA administrative data observations with the identified SSNs, strips each SSA observation of its SSN and other PII, replaces it with the PIN for that observation (from step 1), and provides the constructed SSA data to the researchers.

Note that no individual in this process sees everything. The VR agency staff member sees none of the SSA data, the SSA staff person sees none of the VR data, and the researchers see no Social Security numbers. Other algorithms could be used, but an alternative algorithm should have the same features: maximum anonymity and minimum complexity with merged, anonymous data going to the researcher at the end of the process. For data needs in Dean et al. (2015, 2017, 2018, 2019), a similar but less rigorous method was used for merging DARS and UI data. DARS sent SSNs to a UI staff member and got back the quarterly UI earnings records. DARS then replaced the SSNs with the PINs provided to the researcher in step 1. It is useful to note that, because of the program performance reporting requirements of WIOA, state VR agencies are increasingly coordinating with their UI agencies to get earnings data.

A question of increasing interest motivated by WIOA is the possibility of merging VR data and UI data with administrative data from other agencies providing complementary services, such as the other core WIOA programs. There has been a push to accommodate such data merging partially because of the large number of workforce-related programs existing in many states (e.g., Ganzglass, Reamer, Roberts, Smith, & Unruh, 2010; Jenkins & Harmon, 2010). In principle, this can work as long as there is a common merging identification code across such agencies, the most obvious of which would be the SSN. However, many agencies are cautious about providing Social Security numbers in data available to outside entities, and, in fact, many are moving away from using SSNs even internally. Without such a common merging code, it is impossible to combine agency databases. With such a code, there remain confidentiality issues that can be addressed using carefully constructed merging algorithms as suggested above.

Other Explanatory Variables

Individual Characteristics

It is important to control for other explanatory variables that might influence service choice or labor market outcomes. The online Appendix 6 from Dean et al. (2017) as well as Schmidt et al. (2019) show that controlling for other explanatory variables is critical for estimation and in particular, has a large effect on estimates of VR service effects. Fortunately, VR administrative data sets provide a wealth of information about each individual. Using such data, Dean et al. (2015, 2017, 2018) controlled for gender, race, education, age, marital status, number of dependents, measures of transportation availability, a rich description of the person's disabilities, and labor market history prior to service receipt. Frolich, Heshmati, and Lechner (2004) controlled for age, gender, citizenship, occupation, some disability variables, and some local community characteristics. Hollenbeck and Huang (2006) controlled for gender, race, age, disability, education, veteran status, English proficiency, and labor market history prior to service receipt.

Controlling for such explanatory variables reduces omitted variables bias (e.g., Stock & Watson, 2007, Chapter 6.1; Wooldridge, 2010, Chapter 4.3). It also may help reduce bias caused by service receipt choices being related to unobserved characteristics directly affecting labor market outcomes. In much of the literature, this is called using propensity scores (Frolich et al., 2004; Heckman et al., 1998; Hollenbeck & Huang, 2006), which help only if they include an instrumental variable. Finally, it improves the precision of parameter estimates associated with the effect of service provision on labor market outcomes as discussed in the section above entitled, "Dealing with People not Covered by Unemployment Insurance."

Local Labor Market Conditions

Another set of useful variables to collect, merge, and use are measures of local economic activity. VR administrative data provide information about the location of residence of each program participant. This information can be combined with county-level measures of economic activity (primarily available through federal government data sources) such as employment rates, unemployment rates, levels of income, and demographic composition. For example, Dean et al. (2015, 2017, 2018) used information on ratios of number of jobs to population in each county from the Bureau of Economic Analysis, and Dean et al. (2019) used Virginia data on the number of youth with Individualized Education Programs (IEPs) by county. Another example of such data is described in the section above entitled, "Dealing with People not Covered by Unemployment Insurance."

Summary

Constructing quality data is a necessary step in any good ROI/ROR analysis. Doing it well requires skill, knowledge, and care. However, available data provide one with the opportunity to learn much about the operation and effectiveness of the VR agency. Such an analysis is useful both for providing evidence to policy makers interested in the effectiveness of the programs being funded and for use in informing decision-making and continuous improvement methods by agency administration. Our experience is that administrative VR agency data and UI earnings data are very high quality. It is still the user's responsibility to ensure that data are collected in an ethical manner, data elements are considered appropriately, and data are protected.

References

Aakvik, A., Heckman, J., & Vytlacil, E. (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.

Bricker, J., & Engelhardt, G. (2007). *Measurement error in earnings data in the health and retirement study*. Unpublished manuscript.

Bua-Iam, P., & Bias, T. (2011). Economic impacts of West Virginia Division of Rehabilitation Services on consumers with significant disabilities: Realistic return-on-investment models for state-federal VR programs. *Journal of Rehabilitation*, *77*(3), 25–30.

Bua-Iam, P., Hampton, J., Sink, T., & Snuffer, D. (2013). Critical issues unique to the selection of a realistic Return on investment method for public vocational rehabilitation programs. *Journal of Rehabilitation Administration*, *37*(1), 25–33.

Chang, C., Yang, H., Tang, G., & Ganguli, M. (2009). Minimizing attrition bias: A longitudinal study of depressive symptoms in an elderly cohort. *International Psychogeriatrics Journal*, *21*(5), 869–878.

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. *Journal of Rehabilitation Administration*, *40*(1), 23–34.

Clapp, C. M., Pepper, J. V., Schmidt, R. M., & Stern, S. N. (2018a). *Ongoing work on effects of vocational rehabilitation for people who are blind*. Unpublished manuscript.

Clapp, C. M., Pepper, J. V., Schmidt, R. M., & Stern, S. N. (2018b). *Ongoing work on effects of vocational rehabilitation for adults with autism*. Unpublished manuscript.

Dean, D. H., & Dolan, R. C. (1991). Fixed-Effects estimates of earnings impacts for the vocational rehabilitation program. *Journal of Human Resources*, *26*, 380–391.

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.

Dean, D. H., Pepper, J. V., Schmidt, R. M., & Stern, S. N. (2019). The effects of youth transition programs on labor market outcomes. *Economics of Education Review*, *68* (February), 68–88.

de Graaf, R., Bijl, R., Smit, F., Ravelli, A., & Vollebergh, W. (2000). Psychiatric and sociodemographic predictors of attrition in a longitudinal study: The Netherlands mental health survey and incidence study (NEMESIS). *American Journal of Epidemiology*, *152*(11), 1039–1047.

Doyle, J. (2007). Child protection and child outcomes: Measuring the effects of foster care. *American Economic Review*, *97*, 1583–1610.

Ehrenberg, R., & Smith, R. (2012). *Modern Labor Economics: Theory and Public Policy* (11th ed.). Boston, MA: Pearson Education.

Frolich, M., Heshmati, A., & Lechner, M. (2004). A microeconometric evaluation of rehabilitation of long-term sickness in Sweden. *Journal of Applied Econometrics*, *19*, 375–396.

Ganzglass, E., Reamer, A., Roberts, B., Smith, W., & Unruh, R. (2010). *Recommendations for incorporating postsecondary and workforce data systems in statewide longitudinal data systems*. Retrieved from http://www.clasp.org/resources_and_publications/publication?id=0734&list=publications

Heckman, J., Ichimura, H., & Todd, P. (1998). Matching as an econometric evaluation estimator. *Review of Economic Studies*, *65*, 261–294.

Heckman, J., LaLonde, R., & Smith, J. (1999). The economics and econometrics of active labor market programs. In O. Ashenfelter & D. Card (Eds.), *Handbook of labor economics* (Vol. 3, pp. 1865–2097). Amsterdam, Netherlands: North-Holland.

Hollenbeck, K., & Huang, W. (2006). *Net impact and benefit-cost estimates of the workforce development system in Washington state*. Kalamazoo, MI: W. E. Upjohn Institute for Employment Research.

Honeycutt, T., Thompkins, A., Bardos, M., & Stern, S. (2015). State differences in the vocational rehabilitation experiences of transition-age youth with disabilities. *Journal of Vocational Rehabilitation*, *42*, 17–30.

Honeycutt, T., Thompkins, A., Bardos, M., & Stern, S. (2017). Youth with disabilities at the crossroads: The intersection of vocational rehabilitation and disability benefits for youth with disabilities. *Rehabilitation Counseling Bulletin*, *60*(3), 131–144. doi: 10.1177/0034355215621897

Jacob France Institute. (2018). *Federal employment data exchange system*. Retrieved from http://www.ubalt.edu/jfi/fedes/

Jenkins, D., & Harmon, T. (2010). *Using state data to promote continuous improvement of workforce programs: Guidance for states preparing applications to the U.S. DOL Workforce Data Quality Initiative*. Retrieved from https://ccrc.tc.columbia.edu/media/k2/attachments/using-state-data-workforce.pdf

Kapteyn, A., & Ypma, J. (2006). *Measurement error and misclassification: A comparison of survey and register data*. Unpublished manuscript.

Kim, C., & Tamborini, C. (2014). Response error in earnings: An analysis of the survey of income and program participation matched with administrative data. *Sociological Methods and Research*, *43*(1), 39–72.

Lancaster, T. (1992). *The Econometric Analysis of Transition Data*. Econometric Society Monograph. Cambridge, UK.: Cambridge University Press.

Schley, S. G., Weathers, W. R., Hemmeter, J., Hennessey, J., & Burkhauser, R. (2011). Effect of postsecondary education on the economic status of persons who are deaf or hard of hearing. *Journal of Deaf Studies and Deaf Education*, *16*(4), 524–536.

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

Stock, J., & Watson, M. (2007). *Introduction to Econometrics*. Boston, MA: Addison-Wesley.

United States Department of Labor. (2016). Retrieved from https://wdr.doleta.gov/directivesattach/TEGL/TEGL_26-16_Acc.pdf

United States Department of Labor. (2017). Retrieved from https://wdr.doleta.gov/directives/attach/TEN/TEN_3-17_Acc.pdf

Wilhelm, S., & Robinson, J. (2010). Utah State Office of Rehabilitation economic impact study. *Journal of Disability Policy Studies*, *24*(3), 148–157.

Wilson, B. (2005). *Integrated performance information for workforce development*. Retrieved from https://www.nga.org/files/live/sites/NGA/files/pdf/0603IPIBLUEPRINT.pdf

Wooldridge, J. (2010). *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.

Workforce Innovation and Opportunity Act (2014). Retrieved from https://www.congress.gov/113/bills/hr803/BILLS-113hr803enr.pdf

Author Notes

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

We would like to thank Joe Ashley, Robert Froehlich, Maureen McGuire-Kuletz, Kevin Hollenbeck, and especially Kirsten Rowe for thoughtful assistance and comments and the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR Grant 90DP0070) for generous funding. All errors are ours.

An Overview of the VR-ROI Project and its Approach to Estimating ROI

Kirsten L. Rowe
Virginia Department for Aging and Rehabilitative Services

Joseph M. Ashley
Ashley Consulting, LLC

John V. Pepper
University of Virginia

Robert M. Schmidt
University of Richmond

Steven N. Stern
Stony Brook University

Abstract. *The public vocational rehabilitation program (VR) has faced increasing demands to demonstrate its effectiveness in recent decades. Consistent with these rising expectations, the 2014 Workforce Innovation and Opportunity Act (WIOA) introduced new performance accountability requirements, one of which extends agency tracking of employment from the time of case closure to a full year after closure. The VR-ROI Project supplements these requirements by developing and testing a model to address questions of long-term VR impacts and cost-effectiveness. This paper provides an overview of the VR-ROI Project before describing seven key features of the model. In the process, the paper illustrates why rigorous ROI estimates require a partnership between researchers/evaluators well versed in appropriate statistical procedures and VR agency personnel. Through that partnership, the proposed approach offers methodologically sound and feasible strategies for state VR programs to demonstrate their employment impacts and effectiveness for people with disabilities.*

Keywords: Return on investment, public vocational rehabilitation, program evaluation

The public vocational rehabilitation (VR) program, as well as other programs that support employment for people with disabilities, have faced increasing scrutiny and heightened expectations of credible, data-based evidence of their effectiveness (Bua-Iam & Bias, 2011; Bua-Iam, Hampton, Sink, & Snuffer, 2013; Chan, Rosenthal, & Pruett, 2008; Lewis, 2005; McGuire-Kuletz & Tomlinson, 2015; Migliore & Butterworth, 2008; Millington, 2009; Stensrud, Bruinekool, & Vandergoot, 2009; U.S. Government Accountability Office [GAO] 2005, 2007, 2012). Consistent with these rising expectations, the 2014 Workforce Innovation and Opportunity Act (WIOA) introduced new performance accountability requirements for state VR programs that include post-exit employment and earnings (U.S. Department of Education, 2014.).

The Vocational Rehabilitation Return on Investment (VR-ROI) Project, which predates WIOA, was established to develop, refine, and implement a rigorous model for estimating VR long-term employment and earnings impacts as well as ROI. Supported by grant funding from the National Institute on Disability, Independent Living and Rehabilitation Research (NIDILRR), the VR-ROI Project began in 2010 to develop and test an ROI model for state VR programs in cooperation with four partner agencies – the general and blind agencies in Virginia and the combined agencies in Maryland and Oklahoma. The current project objectives are to (a) refine and test the VR-ROI model with a more heterogeneous set of state agencies and (b) explore the feasibility of a VR-ROI Calculator for use by state VR agencies. The project now includes partner agencies in six states – Virginia (both general and blind), Maryland (combined), Delaware (general), Kentucky (both general and blind), North Carolina (both general and blind), and Texas (combined in recent years under the Texas Workforce Commission) – and uses data from three cohorts: applicants for VR services during state fiscal years 2000, 2007, and 2012.

Consistent with NIDILRR's mission, a guiding principle for the project is to produce research results that are both rigorous and relevant. To that end, we have collaborated with staff from all partner agencies in developing a framework that is general enough to be applied to diverse state VR agencies yet adaptable enough to allow for institutional differences across them. We have based our approach on seven core features that we believe to be important to obtaining rigorous and credible ROI estimates, the first two of which are congruent with the requirements of WIOA.

1. Use readily-available longitudinal VR data to focus on individuals, not cases.
2. Link VR data with readily-available administrative data from other systems to examine employment outcomes over time.

The remaining key features incorporate well-established research methods to ensure that studies of VR's impact produce credible, data-based results:

3. Examine VR's impact on groups of individuals from the time they applied for VR services rather than the time after their cases were closed from VR.
4. Include everyone who applies to VR in studying the program's impact, not just those determined to be eligible, those completing an individualized plan for employment, those receiving substantial services, or those closing with a successful outcome.
5. Examine VR's impact starting with each person's first application.
6. Estimate impacts of specific types of VR services separately for people with specific types of disabling conditions.
7. Control for "selection bias" to enhance the credibility of program evaluations and ensure that observed outcomes are the result of what VR provides rather than other factors.

This article describes and provides a brief rationale for each of these key features. Other articles in this volume provide a more detailed justification for our approach and recent estimation results (Clapp, Pepper, Schmidt, & Stern, 2019; Schmidt, Clapp, Pepper, & Stern, 2019; Stern, Clapp, Pepper, & Schmidt, 2019). Other project-related publications provide a more technical discussion, mathematical exposition of the full model, and earlier estimation results (Dean, Pepper, Schmidt, & Stern, 2015, 2017, 2018).

Use Readily-Available Longitudinal VR Data to Focus on Individuals, Not Cases

Substantial numbers of VR program participants have more than one VR case. Table 1 illustrates this point for three project cohorts – individuals who applied for VR services to the Virginia Department for Aging and Rehabilitative Services (DARS) in State Fiscal Years (SFY) 2000 and 2007, and individuals who applied to the Maryland Division of Rehabilitative Services (DORS) in SFY 2007. The table shows the percent of individuals with at least one additional case in prior or subsequent years as well as the number of years examined for each cohort. As can be seen, a substantial percentage of individuals have at least one prior or subsequent case and the percentage rises the longer the period considered.

Given that substantial proportions of VR program applicants in any given year may have previously received VR services, it is important in estimating program impacts to consider their entire VR experience rather than their experience with a single case. Fortunately, the increasing use of relational database software enables program analysts and researchers to link data across cases and across time and even

Table 1
Percentage of Individuals with Prior or Subsequent Cases

Applicant Cohort	Prior Case		Subsequent Case	
	# Years Examined	% with Case	# Years Examined	% with Case
VA DARS SFY 2000	10	23%	7	20%
VA DARS SFY 2007	5	17%	6	19%
MD DORS SFY 2007	3	13%	5	17%

Source: Unless noted otherwise, the values in this table and elsewhere in this paper are based on calculations by the authors using data collected for the VR ROI Project.

across agencies. Of course such an approach requires collaboration between agency personnel and the evaluator/researcher as well procedures to ensure the confidentiality of personally-identifiable data. (See Stern et al., 2019, for suggestions on maintaining individual privacy.) Additional implications of individuals who receive VR services across multiple cases are discussed below in the subsection entitled, "Start with the First VR Application."

Link VR Data with Readily-Available Administrative Data from Other Systems to Examine Outcomes over Time

Until recently, employment data collected by state VR agencies and reported to the Rehabilitation Services Administration (RSA) included self-reported work status and earnings at only two points in time: when an individual applied for VR services and when their VR case was closed. Furthermore, data on employment at closure was reported routinely only for individuals who the VR agency knew were employed at closure. Early studies relying on these data (e.g., Bellante, 1972; Conley, 1969; Worrall, 1978) implicitly assumed that those known to have a job at closure remained employed thereafter, while those whose job status was unknown at closure were unemployed and remained unemployed thereafter. However, Hayward and Schmidt-Davis (2005) and Ashley et al. (2011) found that the proportion of all participants who had jobs declined steadily over the three years following case closure. Indeed, using a cohort of individuals closing from Virginia DARS in SFY 2006, Ashley et al. (2011) found that the decline occurred across all closure conditions (as used in that era's RSA-911 Case Service Report). Their results are reproduced in Table 2. Interestingly, the closure conditions with the highest Year 1 employment rates ("exited with an employment outcome" at 81% and "no impediment to employment" at 65%) exhibited the largest declines (16 and 13 percentage points, respectively) over those years.

VR agencies now track program participants' employment and earnings for a full year after closure for all individuals who have signed an Individualized Plan for Employment (IPE), typically by obtaining quarterly earnings data maintained by their state Unemployment Insurance (UI) program. The links developed between state UI programs and state VR agencies for reporting purposes simplifies the tracking of employment and earnings for ROI purposes as well. As discussed in subsequent sections, the use of a reasonable comparison group that includes individuals who have not signed an IPE as well as the use of employment information from before and after VR are important to obtain rigorous ROI estimates, and UI data requests could be extended to include both. For the VR-ROI Project, partner VR agencies provide UI data on all VR applicants covering a period of at least two years prior to application through at least five years (ten if possible) following the application quarter.

Examine Impacts for Groups of Applicants, Not Closures

The VR-ROI Project models both the service choices that clients make in consultation with their counselors as well as the subsequent effects of those service choices on employment and earnings (if employed). Although many factors enter into the service choice process, they are made within the broader context of institutional and economic environments. From an institutional standpoint, budgetary challenges may lead at times to the imposition of an order of selection

Table 2
Post-VR Employment Rates by Closure Condition for Virginia DARS SFY 2006 Closures

Closure Status	# Cases	Year 1	Year 2	Year 3
No impediment to employment	167	65.3%	62.3%	52.7%
Disability too Significant	142	27.5%	26.1%	19.0%
Exited as applicant (other reason)	1,116	46.4%	42.4%	38.7%
After eligibility but before services	2,067	46.1%	43.5%	40.0%
Exited without employment outcome	3,146	39.5%	37.6%	34.6%
Exited with employment outcome	3,785	81.2%	72.0%	65.3%
All Individuals	*10,423*	*56.9%*	*52.0%*	*47.3%*

Source: Ashley et al. (2011)

and may require eligible individuals to be placed on a waiting list for VR services; during periods without such fiscal issues, all eligible individuals may be served right away. Additionally, agencies may change policies and emphases over time regarding various types of services (e.g., higher education, restoration, placement, and job support services) as well as clientele (e.g., an increasing focus on transition-age youth arising out of WIOA). Institutional variations such as these, as well as variations in economic conditions, should be accounted for in estimating the effects of VR services.

From an economic standpoint, VR clients and their counselors plausibly take into account the clients' recent employment experiences as well as their job prospects with and without various types of services. For example, in a strong economy with low unemployment rates, more employable individuals (i.e., those with more human capital) might conclude that they can obtain suitable employment without following through on an IPE. To the extent this is true, a group of individuals who did not receive VR services but obtained suitable employment likely would compare favorably with the individuals with less human capital who did receive services. We do not intend to argue that this accurately represents the way in which the economy could influence service decisions, only that it is likely that economic conditions do influence service choices.

Many studies of the labor market impacts of VR services and ROI have used closure cohorts, individuals with cases closing in a given fiscal year (see, for example, Bellante, 1972; Hemenway & Rohani, 1999; Kisker, Strech, Vetter, & Foote, 2008; Uvin, Karaaslani, & White, 2004.) However, we have found that a single closure cohort may include cases with applications dating back seven years or more as well as cases that closed within the year. Therefore, we prefer to use applicant cohorts, i.e., individuals applying to a VR agency in a given fiscal year. An applicant cohort has the virtue of having all individuals within the cohort making service choice decisions within the same institutional and economic environments. Although the VR-ROI model does include control variables relating to an individual's characteristics as well as type and severity of disabling condition(s), identifying and quantifying relevant institutional and economic factors presents a challenge. The homogeneity of those factors in an applicant cohort eliminates that concern.

There is a second, pragmatic reason for using an applicant cohort. An important aspect in our estimation of VR service impacts is comparison to a baseline comprised of two to three years of individual employment and earnings (if employed) prior to application. (For additional discussion, see Schmidt et al., 2019). In working with VR agencies from six different states, we have found variability in how long state UI programs retain records. For the cases in a closure cohort that have applications dating back seven or more years, the relevant pre-application employment and earnings period may be ten years or more before the closure date. In some states, UI data that old may not be available.

Include Everyone Who Applies for VR Services

We include all VR applicants in our analysis sample for three reasons. First, individuals who have not received substantive services provide an admittedly imperfect but useful comparison group to address the question of how those served would have fared in the labor market had they not received their particular mix of VR services. The use of a comparison group is not a new concept in evaluations of VR and other workforce programs (see, for example, Kisker et al., 2008; Wilhelm & Robinson, 2013). As discussed by the GAO (2012), the challenge is to observe people who are as similar as possible to VR participants but who do not receive VR services to ensure that the observed effects are due to the program itself rather than other factors. It is important to note that "similar" does not mean that the treatment and control groups be identical, because any observable differences between the two groups (such as demographic and disability-related characteristics) can be controlled for statistically. Rather, the two groups should be as similar as possible in "unobservable" characteristics (i.e., traits such as motivation, family support, and self-efficacy) that are important in jobseekers' choices to seek VR services and their ability to become and remain employed. Even after controlling for observable differences, substantial unobserved differences between individuals who receive VR services and those who do not (known as a selection effect) likely will remain. We discuss this issue further in the subsection below entitled, "Control for the Problem of Selection Bias."

The second reason for including all applicants is that they significantly increase the overall sample size. Larger samples can produce estimates that are more precise. In other words, they provide greater statistical confidence in the estimated impacts and ROI. Third, it is important to recognize that the VR program's impact may not be limited to individuals with a signed IPE, even though existing program performance measures emphasize those types of cases. The national longitudinal study commissioned by RSA revealed that over one-third of individuals who had an IPE, but who did not receive substantial services, were employed 1-3 years after exiting VR (Hayward & Schmidt-Davis, 2003). Services received during the eligibility determination process (e.g., medical evaluation, psychological evaluation, discussions with VR counselors about barriers to employment) or during the development of an IPE (e.g., vocational evaluation, discussions with VR counselors about career goals, information about the skills required for specific jobs) may influence an individual's employability even if the person's case ends before planned services are provided.

Start With the First VR Application

It is standard practice in workforce program evaluations to begin examining potential impacts as soon as the intervention or treatment – VR services, in this case – is started (see, for example, Aakvik, Heckman, & Vytlacil, 2005; Dean et al., 2015, 2017, 2018, 2019; Heinrich, Mueser, Troske, Jeon, & Kahvecioglu, 2009). Following this practice in VR research is complicated because of many clients in a cohort for whom the case is not their first (between 13 and 23% for the three cohorts summarized in Table 1). Ignoring the fact that a case is not a client's first would make little difference to ROI estimates if VR service impacts were the same in initial and subsequent VR cases. However, Dean et al. (2015) found significantly different employment impacts for individuals who had a prior VR case in a cohort of applicants to Virginia DARS in SFY 2000 with cognitive impairments. Although no generalizations can be made, because these differences varied by employment and earnings as well as by service type, anecdotal information provided by VR staff engaged in the VR-ROI Project suggest that program participants with prior cases tend to have different employment outcomes than participants exiting after their first case.

A researcher can address this issue in two ways. The first (and our preferred) approach is to include only first cases in the analysis sample. Of course, this requires a large enough cohort to afford the loss of individuals for whom this is not the first case. In the VR-ROI Project's work to date, we have had the luxury of large enough cohorts to allow us to use this approach.

The second approach is to include clients with prior cases but control statistically for the apparent differences observed in Dean et al. (2015). We are exploring modeling changes to enable this approach because we be believe that the percent of individuals with multiple VR cases could well rise in the future. WIOA includes an increased focus on serving students and other youth with disabilities. Honeycutt, Thompkins, Bardos, and Stern (2013) argued that since they enter VR at an early age, it is reasonable to assume that many such youth will return for additional services during the decades following their initial case. If true, the proportion of individuals with multiple VR cases is likely to rise to the point where it will be necessary to include clients who have returned for additional VR services in future ROI analyses.

Estimate Program Impacts Separately by Service Type and Disability Type

The vocational rehabilitation program provides a customized set of services to participants. Counselors work collaboratively with participants to develop an IPE that specifies the services to which both parties agree. In other words, VR provides its customers with an individualized service plan with consideration given to the type and severity of the individual's disability or disabilities, needs and desires, family support, job availability, and many other factors. Given the customized nature of the VR program, we would prefer not to model a "generic" VR participant who receives a "generic" type of service. That is, we do not simply split VR participants into "treated" (those who receive substantive services of any type) and "untreated" (those who do not) groups.

Rather, our approach estimates the effects of services on employment and earnings (if employed) separately by disability type (e.g., intellectual disability, learning disability, mental illness, physical disability, and blindness or low vision). Furthermore, we "crack the black box" of the overall impact of VR services by estimating employment and earnings impacts of six to nine broad service types separately for individuals with several different types of disabilities. For all disability types, these include diagnostic, training, higher education, restoration, maintenance, placement, and supported employment services. We add two categories of particular importance to the service regimen for individuals who are blind or have low vision – assistive technology and orientation and mobility. Our categorization of services is not as detailed as the service categories reported on the RSA-911, but it is a manageable number for analysis and is quite revealing as demonstrated by Dean et al. (2015, 2017, 2018, 2019), and Schmidt et al. (2019). We find that the mix of service choices varies substantially across disability types and that the employment impact of the various service types varies dramatically for individuals with specific types of disabilities. For example, among individuals with mental illness who applied to Virginia DARS in SFY 2000, the most effective category of purchased services was vocational training/supported employment (even though supported employment for individuals with mental illness was not yet widely available at that time). The results reveal that the VR program indeed provides a rich and varied set of services in a manner that is customized for the individual.

Control for the Problem of Selection Bias

In estimating the impact of VR on employment outcomes, it is important to recognize that there may be a number of factors influencing both program participation and employment outcomes that cannot readily be accounted for by the information that VR agencies collect on the individuals they serve. For example, it might be the case that individuals who are highly motivated, have strong family support, and/or have ready access to transportation are more likely to choose to participate in VR than those with lower skills, and these individuals would be more likely to have positive employment outcomes even without VR services. Another possibility is that individuals with lower employment likelihood and lower earnings are more likely to be determined eligible for VR. In both of these cases, employment outcomes are correlated with the likelihood of VR participation through influences that are not included in agency databases and therefore not included as control variables in the statistical model. This causes the estimates of VR's impact on participants' employment and earnings to be biased (i.e., to systematically under- or over-estimate their true impact).

A number of ways to address the selection bias problem have been suggested in the workforce program evaluation literature (see Heckman, Ichimura, & Todd, 1998, for a discussion on the relative merits of some alternatives). Indeed, a number of state-level evaluations of VR in the past 15 years (e.g., Bua-Iam & Bias, 2011; Hemenway & Rohani, 1999; Hollenbeck & Huang, 2006; Kisker et al., 2008; Uvin et al., 2004; Wilhelm & Robinson, 2013) have acknowledged and attempted to statistically control, to varying degrees, for this "selection bias" issue. The approach utilized by the VR-ROI Project has been reviewed through a double-blind refereeing process within the academic economics literature (see Dean et al., 2015, 2017, 2018). We provide an overview of that approach here and refer the interested reader to these articles for a more detailed justification and exposition.

The VR-ROI Project model includes a service choice (received or did not) equation for each of 6-9 service types (depending upon the cohort), an employment (employed or not) equation, and an earnings-if-employed (using a log functional form) equation. Even after including around twenty explanatory variables, the selection issue affects each of these equations and we employ statistical techniques to address the issue. Most importantly, we include a separate and distinct "instrumental variable" (IV) in each equation. For a service choice equation, an IV is an independent factor that affects an individual's likelihood of VR participation and yet has no direct effect on the individual's labor market outcome (and vice versa for the labor market equations).

The alternative we use for the service choice equations is an instrument based on the variations among counselors and field offices in the types of purchased VR services their participants receive. This approach acknowledges participant-counselor collaboration and takes advantage of the individualized nature of VR which highly values clinical judgment, individual choice, and local variation in availability of services. Although this instrument is a new concept for VR program evaluation, it makes effective use of the program's broad array of service options to provide convincing evidence of the impact of VR services on employment outcomes. A more complete discussion of the instrumental variable and the VR-ROI model is presented in online Appendix B to Schmidt et al. (2019) at https://scholarship.richmond.edu/economics-faculty-publications/56/

Summary

Congress, federal agencies, and state legislatures have increasingly pushed for a stronger evidence base of the ongoing effectiveness and value of VR services. In the current era of intense scrutiny of publicly-funded human service programs, the methods that VR used in the past to demonstrate its impact are insufficient to address questions of the program's long-term impact. The passage of WIOA and its amendments to the Rehabilitation Act have resulted in major changes in the ways in which the public VR program is being evaluated. The features of our approach to evaluating state VR programs are largely congruent with the new VR program performance expectations and offer opportunities for VR agencies to be well-positioned to respond to the ongoing scrutiny of VR's effectiveness in serving individuals with disabilities. Through a partnership between researchers well versed in appropriate statistical procedures and VR agency personnel, implementation of our paradigm is also feasible at the state level. It can strengthen the reliability and validity of VR outcome evaluations, while providing meaningful information for those operating state VR programs as well as decision makers who legislate funding and determine program guidelines. As such, they may be considered "promising practices" for VR program evaluation, and we encourage state VR agencies and VR researchers to consider incorporating these features when modeling the effects of VR services on participants' employment and earnings (if employed).

References

Aakvik, A., Heckman, J., & Vytlacil, E. (2005). Estimating treatment effect for discrete outcomes when responses to treatment vary: An application to Norwegian vocational rehabilitation programs. *Journal of Econometrics, 125*, 15–21.

Ashley, J. M., Rowe, K. L., Dean, D. H., Schmidt, R. M., Pepper J. V., & Stern, S. N. (2011, November). *VR Economics: Developing a new model for vocational rehabilitation return on investment*. Paper presented at the meeting of the National Council on Rehabilitation Education, Arlington, VA.

Bua-Iam, P., & Bias, T. (2011). Impacts of West Virginia Division of Rehabilitation Services on consumers with significant disabilities: Realistic return-on investment models for state-federal VR programs. *Journal of Rehabilitation, 77,* 25–30.

Bua-Iam, P., Hampton, J. E., Sink, T., & Snuffer, D. W. (2013). Critical issues unique to the selection of a realistic return on investment method for public vocational rehabilitation programs. *Journal of Rehabilitation Administration*, *37*, 25–34.

Bellante, D. (1972). A multivariate analysis of a vocational rehabilitation program. *Journal of Human Resources*, *7*(2), 227–241.

Chan, F., Rosenthal, D. A., & Pruett, S. R. (2008). Evidence-based practice in the provision of rehabilitation services. *Journal of Rehabilitation, 74*, 3–5.

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. *Journal of Rehabilitation Administration*, *40*(1), 23–34.

Conley, R. (1969). A benefit-cost analysis of the vocational rehabilitation program. *Journal of Human Resources*, *4*(2), 226–252.

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.

Dean, D. H., Pepper, J. V., Schmidt, R. M., & Stern, S. N. (2019). The effects of youth transition programs on labor market outcomes. *Economics of Education Review*, 68(February), 68–88.

Hayward, B. J., & Schmidt-Davis, H. (2003). *Longitudinal Study of the Vocational Rehabilitation (VR) Services Program, Second Final Report: VR Services and Outcomes.* Washington, DC: U.S. Department of Education, Office of Special Education and Rehabilitative Services, Rehabilitation Services Administration.

Hayward, B. J. & Schmidt-Davis, H. (2005). *Longitudinal Study of the Vocational Rehabilitation (VR) Services Program, Third Final Report: The Context of VR Services.* Washington, DC: U.S. Department of Education, Office of Special Education and Rehabilitative Services, Rehabilitation Services Administration.

Heckman, J., Ichimura, H., & Todd, P. (1998). Matching as an econometric evaluation estimator. *Review of Economic Studies, 65*, 261–294.

Heinrich, C., Mueser, P., Troske, K., Jeon, K., & Kahvecioglu, D. (2009). New estimates of public employment and training program net impact: A non-experimental evaluation of the Workforce Investment Act program. *Working papers series at the Department of Economics at the University of Missouri*. Retrieved from https://mospace.umsystem.edu/xmlui/handle/10355/2389

Hemenway, D., & Rohani, F. (1999). *A cost-benefit analysis of the employment of people with disabilities in Florida: Final Report*. Tallahassee, FL: Florida State University Educational Services Program.

Hollenbeck, K., & Huang, W. (2006). *Net impact and benefit-cost estimates of the workforce development system in Washington state.* Kalamazoo, MI: W. E. Upjohn Institute for Employment Research.

Honeycutt, T., Thompkins, A., Bardos, M., & Stern, S. (2013). *State differences in the vocational rehabilitation experiences of transition-age youth with disabilities*. Washington, DC: Mathematica Policy Research Working Paper (WP 2013-02).

Kisker, E., Strech, G., Vetter, J., & Foote, C., (2008). *Evaluating rehabilitation services in Oklahoma: An analysis of program impacts and of benefit/costs*. Oklahoma Department of Rehabilitative Services. Retrieved from http://eteam. ou.edu/www-eteam/file.axd?file=/media/DRS% 20Exec%20Summary.pdf

Lewis, A. (2005). Using data to improve outcomes in rehabilitation practice. *Journal of Rehabilitation Administration, 29*(1), 43–56.

McGuire-Kuletz, M., & Tomlinson, P. (2015). *Return on investment and economic impact: Determining and communicating the value of vocational rehabilitation*. (Institute on Rehabilitation Issues Monograph No. 38). Washington, DC: The George Washington University Center for Rehabilitation Counseling Research and Education.

Migliore, A., & Butterworth, J. (2008). Trends in outcomes of the vocational rehabilitation program for adults with developmental disabilities: 1995-2005. *Rehabilitation Counseling Bulletin, 52*, 35–44.

Millington, M. (2009). Introduction to the special issue: Drawing content from an emerging QA community. *Journal of Rehabilitation Administration*, *33*(2), 65–68.

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

Stensrud, R., Bruinekool, M., & Vandergoot, David. (2009). Program management, quality assurance, and research needs for public rehabilitation. *Journal of Rehabilitation Administration*, *33*(2), 83–90.

Stern, S. N., Clapp, C. M., Pepper, J. V., & Schmidt, R. M. (2019). Data issues in developing valid ROI estimates. *Journal of Rehabilitation Administration*, *40*(1), 35-46.

United States Department of Education (July 2014*). The Workforce Innovation and Opportunity Act Overview of Title IV: Amendments to the Rehabilitation Act of 1973*. Retrieved from http://www2.ed.gov/about/offices/list/osers/rsa/publications/wioa-changes-to-rehab-act.pdf

United States Government Accountability Office. (2005). *Vocational rehabilitation: Better measures and monitoring could improve the performance of the VR program*. (GAO-05-865). Washington, DC: Author.

United States Government Accountability Office. (2007). *Vocational rehabilitation: improved information and practices may enhance state agency earnings outcomes for SSA beneficiaries*. (GAO-07-521). Washington, DC: Author.

United States 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, DC: Author.

Uvin, J., Karaaslani, D., & White, G. (2004). *Evaluation of Massachusetts' public vocational rehabilitation program: Final report.* Boston, MA: Massachusetts Rehabilitation Commission.

Wilhelm, S., & Robinson, J. (2013). The economic impact of Utah's vocational rehabilitation program. *Journal of Disability Policy Studies, 24*(3), 148–157.

Worrall, J. D. (1978). A benefit-cost analysis of the vocational rehabilitation program. *Journal of Human Resources*, *13*(2), 285–298.

Author Notes

Kirsten L. Rowe, PhD, is the Vocational Rehabilitation Return on Investment (VR ROI) Project Coordinator for the Virginia Department for Aging and Rehabilitative Services. Joseph M. Ashley, PhD, CRC, is a co-Principal Investigator for the VR ROI Project and recently retired Assistant Commissioner for the Virginia Department for Aging and Rehabilitative Services. He is currently Principal at Ashley Consulting, LLC. John Pepper, PhD, is a researcher for the VR ROI Project and the Merrill H. Bankard Professor of Economics at the University of Virginia. Robert Schmidt, PhD, is a co-Principal Investigator for the VR ROI Project and Emeritus Professor of Economics at the University of Richmond. Steven Stern, PhD, is a researcher for the VR ROI Project and a full Professor of Economics at Stony Brook University.

We are particularly indebted to the late Dr. David Dean whose foundational work made the VR-ROI Project possible. Using his knowledge of the VR program and labor economics, he first brought together the members of the VR ROI Project team. We are also grateful to Dr. Dean and Dr. Michael O'Brien for their efforts on earlier versions of this paper as well as Dr. Kevin Hollenbeck and an anonymous reviewer for insightful comments. Finally, we are grateful to the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR Grant 90DP0070) for generous funding of the VR-ROI Project.

Journal of Rehabilitation Administration, 40(1), pp. 57–72, 2019

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

*A**bstract**. 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 exclude 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, and 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% are dropped because they have none of the three disability types and another 15% are dropped because this is 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 among 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

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%

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.

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.

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 do 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., caused by the existence of 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 an 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 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, an ROR that is not positive does not 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 10% in the stock market and about 1% currently in money market accounts. 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 did 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

Table 3

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

	MI	PI	CI
% with Positive ROR	67.0%	58.0%	25.0%
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%

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 is 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 decided 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.

Table 4

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

	MI	PI	CI
% with Positive ROR	78.0%	26.0%	33.0%
ROR at Median	14.0%	0.0%	0.0%
5th Percentile	26.1%	1.1%	8.7%
90th Percentile	42.5%	23.0%	30.2%

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 and cognitive impairment*) and service type (*diagnosis & evaluation, training, education, restoration, maintenance, job placement and 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, and 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.

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. *Journal of Rehabilitation Administration, 40*(1), 35–46.

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. *Journal of Rehabilitation Administration, 40*(1), 17–22.

Hollenbeck, K. (2019). What is ROI? *Journal of Rehabilitation Administration*, *40*(1), 5-10.

Hopkins, K. (2019). Social return on investment: An important consideration for state vocational rehabilitation programs. *Journal of Rehabilitation Administration*, *40*(1), 11–16.

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. *Journal of Rehabilitation Administration*, *40*(1), 47–56.

Author Notes

Robert Schmidt, PhD, is an Emeritus Professor of Economics at the University of Richmond. Christopher Clapp, PhD, is a Senior Research Associate in the Kenneth C. Griffin Department of Economics at the University of Chicago. John Pepper, PhD, is the Merrill H. Bankard Professor of Economics at the University of Virginia. Steven Stern, PhD, 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.

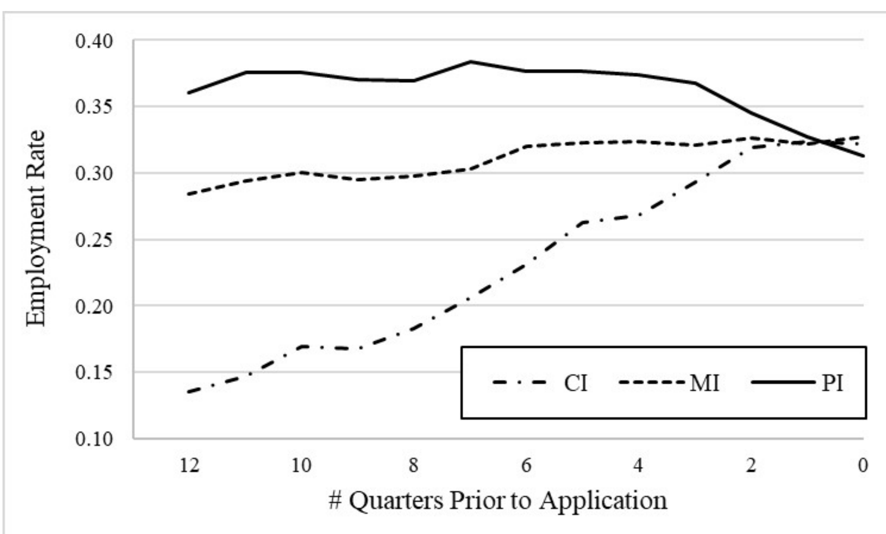

Figure 1. VA DARS – Quarterly Employment Rates Prior to Application by Agency and Cohort

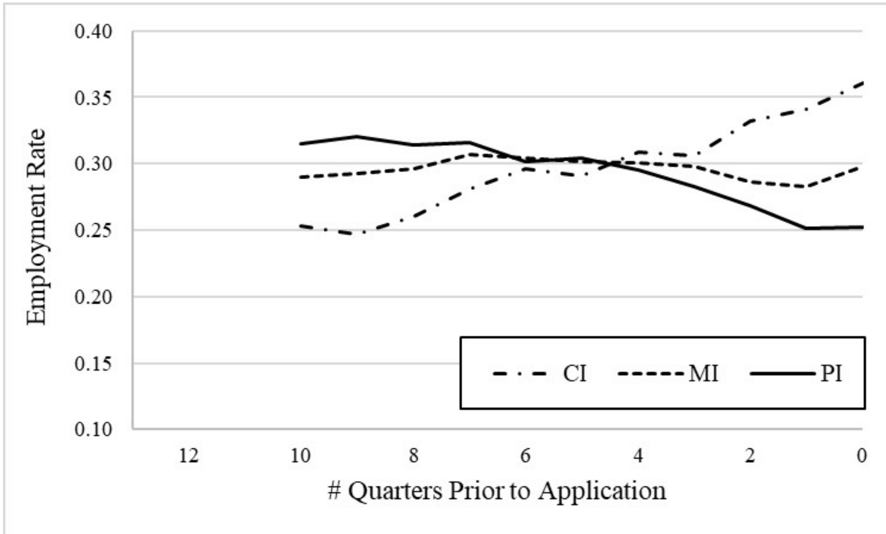

Figure 2. MD DORS – Quarterly Employment Rates Prior to Application by Agency and Cohort

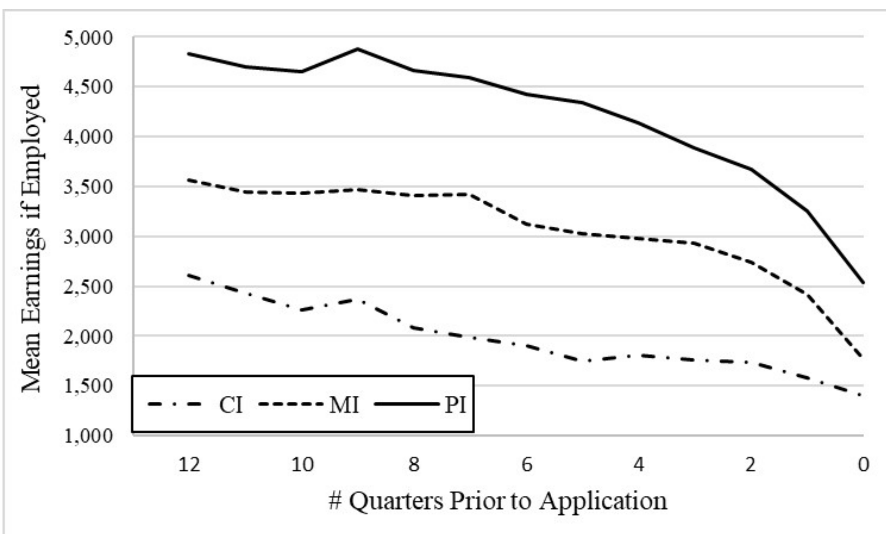

Figure 3. VA DARS – Average Quarterly Earnings (if employed) Prior to Application by Agency and Cohort

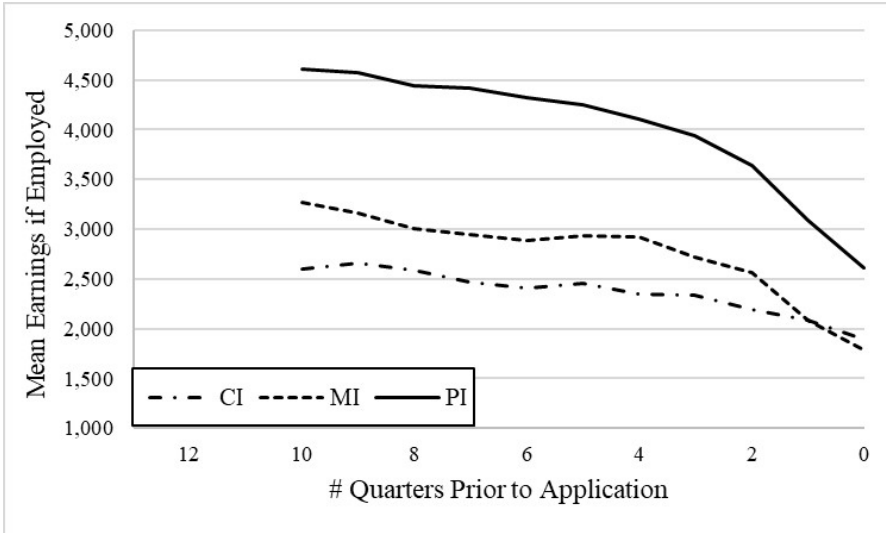

Figure 4. MD DORS – Average Quarterly Earnings (if employed) Prior to Application by Agency and Cohort

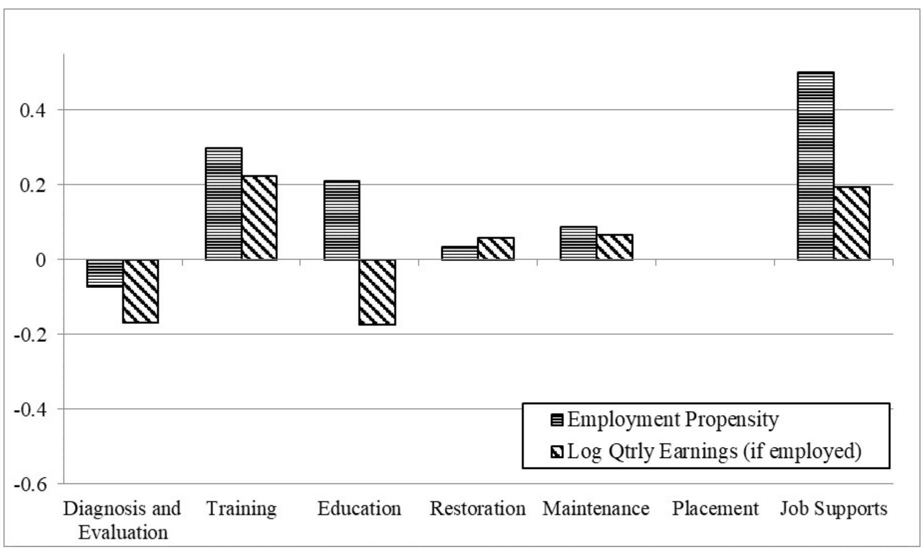

Figure 5. VA DARS MI Cohort – Long-run Impact of Purchased Services on Labor Market Outcomes by Service Category. (Note: VA DARS does not purchase *Placement* services although they are provided by WWRC.)

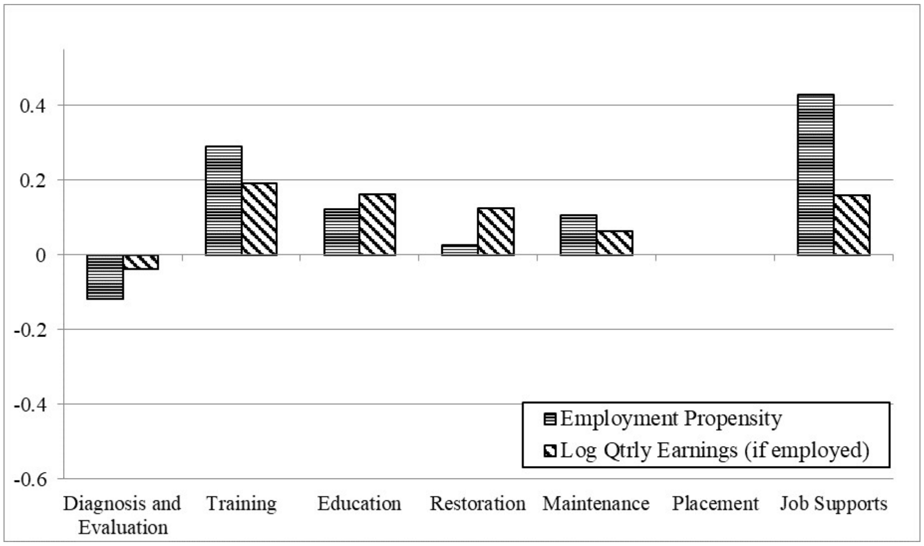

Figure 6. VA DARS PI Cohort – Long-run Impact of Purchased Services on Labor Market Outcomes by Service Category (Note: VA DARS does not purchase *Placement* services although they are provided by WWRC.)

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

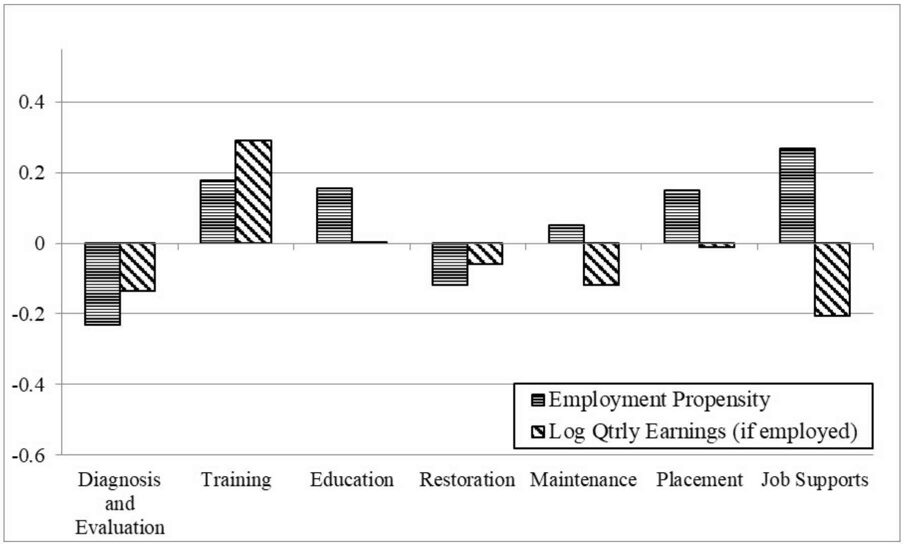

Figure 10. MD DORS DARS CI Cohort – Long-run Impact of Purchased Services on Labor Market Outcomes by Service Category

Extending the VR-ROI Approach to Measure the Return on Virginia's Investment in the Public Workforce System

Jeffrey Brown
Manhattan Strategy Group

Joseph Ashley
Ashley Consulting, LLC

Abstract. *This article discusses recent efforts to evaluate the return on investment (ROI) of Virginia's public workforce development system and provides examples of similar efforts in other states. It considers various challenges in attempting to meet the common expectation of a measurable return on public investments in the workforce system for both participants and taxpayers, and argues for extending to other workforce programs the approach being used in Virginia to develop ROI estimates for the state vocational rehabilitation program.*

Keywords: Return on investment (ROI), cost-benefit analysis, net impact evaluation, vocational rehabilitation, Virginia, workforce development

In June 2016, former Virginia Governor Terry McAuliffe met with several of his cabinet secretaries and state agency heads to discuss Virginia's workforce. The economy in the state was in the midst of historic change: Virginia's long reliance on defense spending was diminishing in Northern Virginia and Hampton Roads, as was her reliance on coal mining from the Appalachian Mountains in the southwest. Promising new industry sectors were sprouting up near the capital city of Richmond and along the bedroom communities outside Washington, DC. The workforce itself was undergoing a dramatic shift, as older workers were leaving the workforce, taking their knowledge and experience with them. Reminding everyone present that "Virginia began 400 years ago as a business venture," the governor settled in with his staff to focus on addressing the needs of what he called "The New Virginia Economy" and the imperative before all of the Commonwealth's workforce leaders to find solutions for the modern economic and workforce development challenges (Commonwealth of Virginia, 2016).

Governor McAuliffe asked his team to help him address five workforce challenges he believed held the key for success in meeting Virginia's workforce needs: (1) attaining workforce credentials to make job seekers more competitive in the job market, (2) emphasizing career pathways toward "middle skills jobs," (3) identifying avenues for family-sustaining wages for Virginia's workers, (4) accelerating business engagement in the workforce system, and (5) realizing a high return on the investment taxpayers make in the system.

It is this last challenge – the expectation of a measurable return on public investment to workforce system participants and to taxpayers – that we shall address in this article. The task of defining and standardizing an objective measure capturing return on investment (ROI) has been particularly elusive for the workforce development system, and the quest to find consensus around ROI has been remarkably instructive to the workforce partners representing each of the agencies whose leaders sat down with the governor.

The Need to Measure ROI

It seemed like a simple proposition when the governor said that he wanted to measure (and publicize) the ROI of Virginia's public workforce system. After all, he knew that approximately $365 million flowed into the workforce system in June 2016 (Dunnigan, 2017). The question then became a simple one any good investor in the public good would ask: did more dollars flow into the Commonwealth's economy as a result of our investment than were spent to make the investment in the first place? If a program cost $1 to administer, could we objectively show that the Commonwealth realized, say, $1.25 in tax revenue, lower support and social services costs, and long-term community wealth building?

Moreover, the Governor was particularly interested in customer outcomes. Like governors and administrators in many states, he was eager to know what happened to people after they encountered our public workforce system. Did they get timely and relevant job training? Did that training lead to a job in the field for which they had trained? Were their wages higher or lower than when they entered the system? Were businesses actually getting the workers they wanted and needed? These, too, were rudimentary questions, but questions on which the viability – and credibility – of Virginia's public workforce development system depended. And at the heart of every question was the stubborn issue of ROI.

A pesky conundrum arises when Virginia's workforce stakeholders bring up the topic of ROI: should the Commonwealth support the expectation that taxpayers receive a positive return on their investment in the public workforce system, or is it more important that system participants (job seekers, students, and clients) see a return on the time and effort they dedicate to workforce programs? Is "investment" in the public sphere defined as turning a pot of money into a larger pot of money (as it is in the private sector), or is "investment" the cost of providing opportunity to those who might not otherwise get it? Reasonable people disagree about whether the effectiveness of public programs can be evaluated fairly by their ability to achieve revenue neutrality (or better). Perhaps the answer to Virginia's conundrum is that ROI should strive to measure both the state's fiscal growth and the clients' personal growth. (See Hollenbeck, 2019, for a discussion of cost-benefit analyses from different perspectives.)

After many months of debate, Virginia still does not have a universally accepted method of measuring ROI across our entire workforce system. We have found the greatest success (and the highest potential for universality) in our Vocational Rehabilitation Return on Investment (VR-ROI) project, which was rolled out at Virginia's Department for Aging and Rehabilitative Services (DARS) in 2015. The award to the University of Richmond by the National Institute for Disability, Independent Living, and Rehabilitation Research (NIDILRR) funded the VR-ROI project, which provides a systematic measurement of returns that we discuss in greater detail later in this article. (See Rowe, Ashley, Pepper, Schmidt, & Stern, 2019, for an overview of the VR-ROI project.)

Accountability or Program Improvement?

The context of Governor McAuliffe's challenge was to calculate ROIs in order to be accountable to taxpayers. However, it should be recognized that many scholars urge caution. In discussing basic ROI concepts in workforce development programs, Hollenbeck (2012) provides an important perspective:

> calculating ROIs for workforce development programs requires considerable data and careful analyses of benefits and costs. We believe that it is folly to think that a simple, one-size-fits-all tool can be developed that can estimate a program's ROI with minimal data inputs and with quick turnaround. Even though the data and analytical burdens are great, we believe that analyses of ROIs (or, equivalently, benefits and costs) are a tool that administrators should use ***to monitor their program's performance*** [emphasis added] (p. 16).

Wilson (2005) discusses the difference between the program evaluation effort and a net impact assessment for the purpose of program improvement and outcome determination. The return on investment focus on net impacts takes longer to develop and is not a usable strategy for ongoing continuous quality improvement, but it is appropriate for developing policy and demonstrating value.

Similarly, Harper-Anderson and Jin (2014) underscore the importance of measuring ROI when they argue:

> Public investment in Virginia's workforce development programs is necessary to create a strong workforce, connect job seekers to work opportunities, and provide businesses with the talent necessary to keep them competitive. Under cer-

> tain circumstances these programs have the capacity to yield future returns on the taxpayers' investment. While each program has a different cost structure and serves a particular demographic, which both play a role in the ROI outcome, under the right circumstances, the likelihood of positive returns can be increased. . . . ROI analysis can help direct resources to capitalize on proven practices and address inherent challenges in the system (p. 48).

Wilson (2005) notes that there is an inherent propensity to compare one workforce program to another, but as noted by Harper-Anderson and Jin (2014) there is variability in the cost of programs and Wilson notes that differences in program participants affects outcomes. Workforce participants with barriers to employment typically require additional services and programs focused on training may have different cost structures so that comparison across programs becomes a challenge if not appropriate. Comparing within a program over time is a strategy that can assist in program improvement and policy review.

These thought leaders and others have shaped the national discussion about the importance of considering the ROI of workforce programs. They have also provided Virginia policy makers with a jumping-off point from which to contemplate our own state's approach to ROI.

Prior State Efforts

In fact, states have long grappled with the issue of return on investment for workforce development programs. It is a very good question to ask whether public investment in economic development actually "works," as development of the workforce is a key pillar in the development of economies. It is also productive to ensure that questions about the viability and success of workforce programs are asked within the proper context. In fact, defining this context appears to have been a primary occupation of studies of ROI in other states, and several approaches to the estimation of ROI have emerged thanks to this work.

In 1995, the Washington legislature enacted legislation that required outcome evaluations of the state's training system. Specifically, this law requires non-experimental net-impact and cost-benefit evaluations of secondary vocational-technical education, work-related adult basic skills education, postsecondary workforce training, Job Training Partnership Act programs, as well as of the system as a whole.

The Washington State Training and Education Coordinating Board has published several studies conducted by the Upjohn Institute for Employment Research that estimate the return on investment of the programs comprising the state's workforce training system (Hollenbeck & Huang, 2003; 2006; 2014; 2016). With the exception of their analyses of vocational rehabilitation (VR), these studies used a non-experimental approach in which individuals who exited from one of the state's training programs were matched statistically to a comparison set of individuals who encountered the state's Wagner-Peyser employment service. For the state's VR program, these researchers used a multivariate regression approach in which individuals who applied for VR services, but did not receive them, were used as the comparison group.

Hollenbeck and Huang (2007) provided the Virginia workforce system with an assessment of the state workforce programs, analyzing outcomes for all four workforce titles and for other related workforce programs such as social services and correctional education programs. One issue that became clear during the process was the difficulty in developing an appropriate comparison group for net impact comparisons. The authors were familiar with Dean and Dolan (1991). In addition, their prior experience in evaluating the Washington State workforce programs led them to use VR program applicants and other groups of VR program participants who did not complete the program for the comparison group, rather than Wagner-Peyser program participants (who were used as the comparison group for other Virginia workforce programs). By using what was perceived to be a more appropriate comparison group, the Virginia VR agencies were more comfortable with the resulting outcome study and net impacts analysis. The selection of a suitable testing cohort highlights the importance of addressing the program participant comparison groups required for net impact studies.

In 2008, Texas measured four contextual factors in its study of workforce programs' ROI. These factors were Participant ROI, Taxpayer ROI, Societal ROI, and "Low- Versus High-Intensity" ROI (King, Tang, Smith, & Schroeder, 2008). Texas's argument was that ROI means different things to different people, and chose to measure several different returns rather than attempting to settle on a single, all-encompassing return. It is, however, a difficult task to objectively measure concepts like "societal return," and safeguard these measures against political influences over the long run.

In 2013, Oregon conducted a study to estimate the ROI of its vocational rehabilitation system (Renfro, Kessi, Khalid, & Munoz, 2013). An interesting aspect of this study is that it used a macroeconomic (input-output) model to estimate the impact of clients' purchased services on the state's economy.

Minnesota uses established diagnostic tools to measure workforce development ROI, and in 2015 the state codified these tools into law. Similar to Washington, state law requires analysis of the net impact of workforce services on individual employment, earnings, and public benefit use, as well as a cost-benefit analysis of the impacts of workforce services from the participant and taxpayer points of view (Maryns & Robertson, 2015).

While these approaches all have merit, none of them is quite as sophisticated as Virginia's Vocational Rehabilitation ROI (VR-ROI) model (Rowe et al., 2019; Schmidt, Rowe, Stem, & Sizemore, 2018). The pragmatic approach of using only readily available administrative data from services records and UI earnings records is an important aspect of model feasibility. The richness of the VR data, combined with long-term unemployment insurance (UI) wage data, provides more in-depth information to determine service employment impact. The use of applicant cohorts offers similar advantages to those in VR, because services are provided under the same policies for a more homogeneous service environment and more consistent economic conditions. In addition, the VR-ROI model includes state-of-the-science statistical techniques to ensure that results can be attributed to participation in VR. Begun in 2015, this approach to measuring ROI is also a net impact approach. However, VR-ROI takes a very long view of each participant, conducting longitudinal analyses using up to three years of pre-VR employment data and at least five years of post-application data (Ashley, Rowe, & Stern, 2015; Rowe et al., 2019). For us, this distinction is an important one: modern workforce development is not a transactional undertaking. Little evidence exists to suggest that meaningful, sustainable employment can be successfully obtained through fleeting "access to services," or isolated interactions with the public workforce system.

Additionally, the VR-ROI model estimates labor force impacts and ROI at the individual level. That is, the model is not concerned with overall "employment rates." This is a critical feature of the approach to ROI, because it compels the public workforce system to address employment successes and failures by evaluating each individual employment outcome. In our opinion, it is better public policy to strive for fruitful outcomes for all system participants than to focus only on those participants who are most likely to succeed: we should adjust our programs to meet customers' needs, rather than adjust our customers to meet program needs. This individual participant approach to ROI is aligned with the strength of the vocational rehabilitation system: individual availability of services based on participant informed choice. Also focusing on participant services received, the VR-ROI Model is able to identify the impacts of group services on labor force attachment as well as earnings levels for those who gain employment. The richness of the data sets enhances the usefulness in program improvement but does require a long view.

Many key features of the VR-ROI approach are appropriate for use with other workforce programs. For example, the model uses longitudinal wage data from quarterly employment records from several years of pre-VR earnings, as well as five or more years of data following start of services. The model uses an applicant cohort that begins to estimate VR's impact when services begin, not when they end. Another key feature the model provides is an estimation of employment and earnings impacts, as well as service costs at the individual level. It recognizes that services might have differential impacts depending upon the type of service and disability. Of particular importance to workforce programs are several state-of-the-science statistical controls to account for the variability of services and individual nature of services by participant. These VR ROI features are a part of ensuring that program impacts are attributable to Vocational Rehabilitation program services rather than other factors.

Extending VR-ROI to the Workforce System

Our challenge now is to scale the VR-ROI model across the Commonwealth's workforce development system. There is a clear need in Virginia to have the capacity to measure and evaluate the effectiveness of our public workforce system as a whole. Yet our system – like the systems in every other state – is comprised of several distinct programs and funding streams, none of which was originally designed to work in concert with any of the others. Although the 2014 Workforce Innovation and Opportunity Act (WIOA) provided states with a collaborative framework with which to build workforce policy, the law did not go so far as to combine funding streams (or make them more flexible), nor did it compel states to consolidate administrative control of existing programs. In many ways, the prospect of defining ROI becomes an exercise akin to building Babel's infamous tower, with the primary impediment to success being the multiplicity of languages spoken by workforce professionals administering what in Virginia are eight different programs under four different secretariats.

As researchers Harper-Anderson and Jin (2014) noted just before the passage of WIOA,

> . . . costs are very different across the programs. While the overall budget for WIA [Workforce Investment Act Title I programs] is more than twice that of TAA [Trade Adjustment Assistance program], the WIA program serves about five times more clients per year. . . .WP [Wagner-Peyser employment service], on the other hand, has the highest budget of the three programs but because it has no training the costs per person are extremely low. The average cost for WIA training is about $435, for TAA the average cost is almost triple at $1,725. To put this into perspective, the average WP cost per participant is less than 1% of the cost for a trained TAA participant and about 10% the cost of one who does not receive training (p. 37).

VR-ROI Value

The VR-ROI Project has developed a set of tools that describes the economic value of VR services while providing information for improving service delivery and policy development. Project results are not intended to be 'the one answer' or to be used without context but to provide additional information to existing data for policy analysis and program accountability (Schmidt et al., 2018).

Both Virginia and Maryland VR agencies participate in the VR-ROI Project and have experience reviewing ROI data for program improvement, policy analysis, and accountability purposes. VR-ROI researchers strive to provide as accurate as possible ROI estimates and at the same time provide as meaningful as possible information to VR agencies (Rowe et al., 2019; Schmidt et al., 2018). Critical to achieving this objective is the engagement of key VR staff including those at Maryland DORS and Virginia DARS to refine the ROI model. The partnership between VR-ROI researchers and VR agencies has been essential to the project development and continual model refinements. This commitment of the economic researchers to work with program experts to ensure an accurate ROI model is a strength of the project and would work well with other workforce programs.

A recent review of VR-ROI data from Maryland and Virginia (Schmidt et al., 2018) provided a platform to reflect on agency involvement and model impact. VR agency staff noted value from project participation. The in-depth service review paired with labor force attachment information provided a new perspective for them. The data preparation process also had benefits. Maryland staff noted improvement in Individualized Plans for Employment and increased consistency in service coding across the agency as a result of VR-ROI project participation. Maryland staff further reported getting a head start on internal controls documentation for RSA 911 reporting and for WIOA data collection (Schmidt et al., 2018). In addition to benefits from data preparation, Maryland staff noted ROI data provided information for policy analysis and service improvement. For example, the 2007 VR-ROI data analysis identified diagnostic services as having a negative ROI. This prompted a review of related services and recent policy decisions. Contributing factors were discussed as well as recent policy changes that seemed to address the potential factors (Schmidt et al., 2018). Using the VR-ROI data analysis, Maryland staff were able to compare programs outcomes of the Workforce Technology Center (a Maryland operated state rehabilitation center) This was of great interest to the Maryland DORS. For example, Maryland 2007 ROI results indicate employment propensity and earnings were substantially higher when service was provided exclusively by the Center (Schmidt et al., 2018).

The Virginia VR agency was also interested in outcome information on a school-to-work transition program at the Wilson Rehabilitation and Workforce Center (WWRC), a state-operated rehabilitation center. The Postsecondary Education Rehabilitation Transition (PERT) program is a collaborative school to work transition service that uses intensive career assessment and independent living skills from a twelve-day residential assessment at WWRC as a foundation for community-based education and VR teams service planning. Participants may receive additional VR services following PERT participation (Ashley et al., 2015). The PERT study focused on a cohort of transition-age youth who applied to Virginia DARS during state fiscal year 2000 and included agency data as well as 12 years of employment and wage data. Compared with youth who did not participate in

PERT, the study found that PERT increased job finding and keeping by 12%. If the student participated in at least an additional year in high school, job finding and keeping were increased by another 38%. If employed, a participant can expect to earn twice as much as if they did not participate in PERT. The information had the unexpected impact in that it was a morale boost for the PERT staff, confirming with data their perception that the service had tremendous impact. PERT staff were interested in using the ROI results as information for parents, students and education partners. VR-ROI research economists worked with PERT staff to create informational statements that were consistent with the study data. The result of the collaboration was an informational brochure that presented the PERT outcome data in an understandable format (for example job finding vs. labor force attachment) consistent with study data. Targeting the information to parents, students and education partners proved helpful for explaining the value of the service to perspective participants.

Extending the VR-ROI Project Approach

The VR-ROI project provides a starting point for developing a model for ROI estimates for all workforce programs. For example, the use of applicant cohorts seems appropriate for all workforce programs as it creates a process that has all members of an analysis cohort served using the same policies within each program. It also ensures other aspects, such as economic conditions and other external factors affecting individuals' decisions to seek employment or apply for services, are consistent. Developing comparison groups specific to each program can appropriately address the "as compared to whom" issue and not require comparison to Wagner-Peyser program participants. Other aspects to the VR-ROI model that would be beneficial to apply across programs are the use of readily available administrative data and the longitudinal approach to go beyond the first year following program exit. This approach is particularly appropriate for programs that promote training and other longer-term services. Finally, the individual service focus of the VR ROI model and the use of state-of-the-science statistical controls and other factors to ensure that service impacts are attributable to the workforce program seem extremely relevant for a workforce ROI model.

Governor McAuliffe left office in January 2018, along with all of the cabinet secretaries and many of the agency heads involved in transforming Virginia's public workforce system during his administration. The new governor, Governor Ralph Northam, has picked up where his predecessor left off, empowering his team to grapple with the public workforce system's perennial challenges. The question of how authorities or agencies should be held accountable for ROI has been a lingering topic, and the team ultimately decided to avoid assigning accountability to workforce partners for the moment because of inconsistencies in approaches and data across programs. The Workforce Innovation Opportunity Act has new data reporting requirements for workforce programs that are consistent across programs. As workforce programs develop data infrastructure and processes to respond to these new requirements, there may be renewed interest in developing a cross-program ROI model.

Because public workforce systems are not inherently self-sustaining – they rely on tax dollars to operate – there will always be some degree of scrutiny about their effectiveness. This scrutiny is not unreasonable, and states should embrace it as a catalyst for continuous improvement and collective impact. We are convinced that one of the best methods for addressing scrutiny is a reliable and accessible method of measuring workforce system ROI. We are also convinced that the best path forward is to adopt the methodology of the VR-ROI initiative statewide through a pilot designed to enable scaling the model appropriately over time. For us, making the effort to establish credible measurements of return on investment is one of the strongest investments we can make.

The nature of the change of administration demonstrates one of the issues in developing a workforce system return on investment model. Although the various agency heads will change, the interest from policy makers on the relative value of investments in the public workforce system remains constant. Developing a consensus across the workforce programs on embracing the accountability provided by a return on investment process is a challenge that must be accepted. The VR-ROI model offers an opportunity to develop a long-term view of workforce outcomes that also provides a richness of information on which to make program adjustments to enhance services.

References

Ashley, J. M., Rowe, K. L., & Stern, S. N. (2015, June). *VR-ROI: The return on investment of Virginia's vocational rehabilitation program*. Paper presented at the VLDS Insights Conference, Fredericksburg, VA.

Commonwealth of Virginia. (2016). *Governor Terry McAuliffe. Two years. Building a new Virginia economy*. Retrieved from https://wayback. archive-it.org/4186/20161220151525/http://governor.virginia.gov/media/5252/gtm_two years.pdf

Dean, D. H., & Dolan, R. C. (1991). Fixed-Effect estimates of earnings impacts for the vocational rehabilitation program. *Journal of Human Resources*, 26(2), 380–391

Dunnigan, S. (2017). *Virginia Workforce System 2016 Budget Review: Preliminary Report*. Retrieved from http://www.elevatevirginia.org/wp-content/uploads/2014/04/2016-Workforce-System-Budget-Review.pdf

Harper-Anderson, E., & Jin, M. (2014). *Return on investment for Virginia's workforce programs*. Retrieved from https://vlds.virginia.gov/media/1014/vcureturnoninvesment04032014final.pdf

Hollenbeck, K. (2012) *Return on investment in workforce development programs* (Working paper 12-188). Kalamazoo, MI: W. E. Upjohn Institute for Employment Research.

Hollenbeck, K. (2019). What is ROI? *Journal of Rehabilitation Administration*, *40*(1), 5–10.

Hollenbeck, K., & Huang, W. (2003). *Net impact and benefit-cost estimates of the workforce development system in Washington State* (Technical report 03-018). Kalamazoo, MI: W. E. Upjohn Institute for Employment Research.

Hollenbeck, K., & Huang, W. (2006). *Net impact and benefit-cost estimates of the workforce development system in Washington State* (Technical report 06-020). Kalamazoo, MI: W. E. Upjohn Institute for Employment Research.

Hollenbeck, K., & Huang, W. (2007). *Workforce program performance indicators for the Commonwealth of Virginia* (Technical report 08-024). Kalamazoo, MI: W. E. Upjohn Institute for Employment Research.

Hollenbeck, K., & Huang, W. (2014). *Net impact and benefit-cost estimates of the workforce development system in Washington State* (Technical report 13-029). Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Hollenbeck, K., & Huang, W. (2016). *Net impact and benefit-cost estimates of the workforce development system in Washington State* (Technical report 16-033). Kalamazoo, MI: W. E. Upjohn Institute for Employment Research.

King, C. T., Tang, Y., Smith, T. C., & Schroeder, D. G. (2008) *Returns from Investments in Workforce Service: Texas Statewide Estimates for Participants, Taxpayers, and Society*. Austin TX: Ray Marshall Center for the Study of Human Resources.

Maryns, N., & Robertson, R. (2015). *Smart investments, real results: A net impact evaluation framework for Minnesota's workforce development system and initial findings*. St. Paul, MN: Department of Employment and Economic Development.

Renfro, J., Kessi, J., Khalid, A., & Munoz, H. (2013). *Vocational rehabilitation: Return on investment in Oregon* (Report 19). Portland, OR: Northwest Economic Research Center.

Rowe, K. L., Ashley, J. M., Pepper, J. V., Schmidt, R. M., & Stern, S. (2019). An overview of the VR-ROI Project and its approach to estimating ROI. *Journal of Rehabilitation Administration*, *40*(1), 47–56.

Schmidt, R. M., Rowe, K. L., Stem, J., & Sizemore, R. (2018, September). *The VR ROI project: Recent results for Virginia and Maryland*. Paper presented at the 11th Annual Summit on Performance Management Excellence, Oklahoma City, OK.

Wilson, B. (2005). *Integrated performance information for workforce development: A blueprint for states*. Olympia, WA: Workforce Training and Education Coordinating Board.

Author Notes

Jeffrey Brown is the former director of the Workforce Services Division of the Virginia Employment Commission. He is currently a senior consultant for the Manhattan Strategy Group. Joseph M. Ashley, RhD, CRC, is co-Principal Investigator for the VR-ROI Project and recently retired Assistant Commissioner for Virginia Department for Aging and Rehabilitative Services. He is currently Principal of Ashley Consulting, LLC.